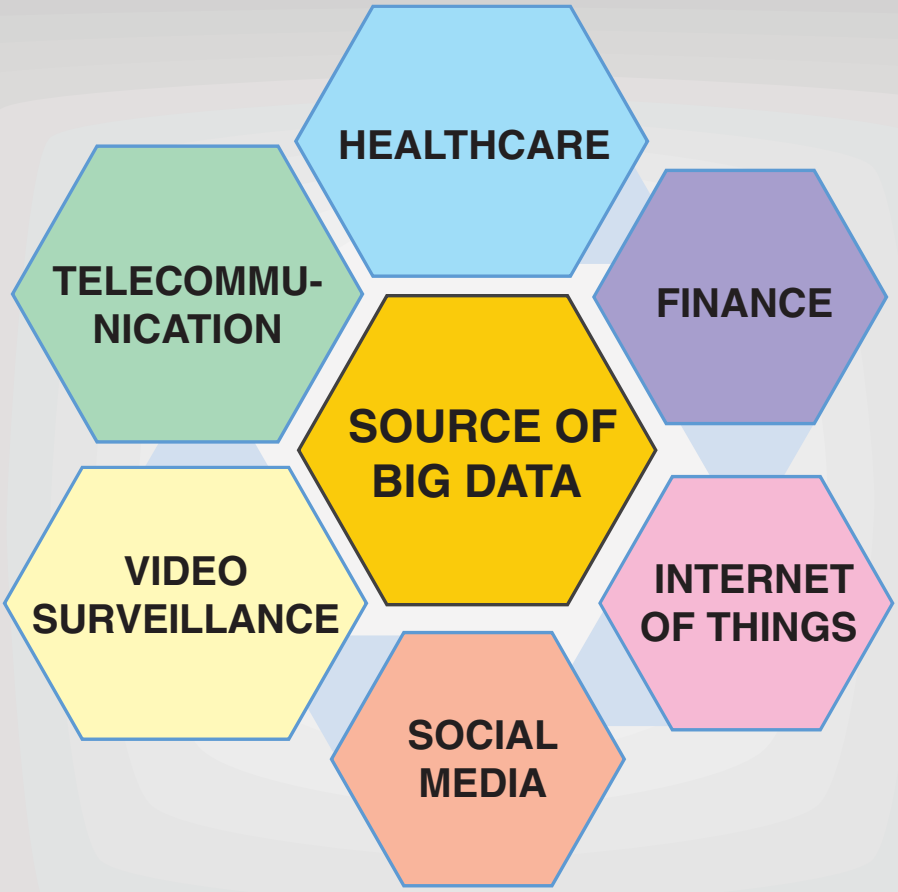
Volume 15 • Number 5 May 2018

Journal of Computational and Theoretical NANOSCIENCE

For all Computational and Theoretical Aspects of Nanoscience and General Mathematical Procedures

www.aspbs.com/ctn



ASP
AMERICAN
SCIENTIFIC
PUBLISHERS

Journal of
Computational and Theoretical Nanoscience
www.aspbs.com/ctn

VOLUME 15, NUMBER 5
MAY 2018

1429–1778 (2018)
ISSN: 1546-1955, EISSN: 1546-1963

Manuscript Submission

Submit your manuscript electronically as a PDF or MS Word file to the Editor-in-Chief, Prof. Rui-Hua Xie or one of the Associate Editors listed above.

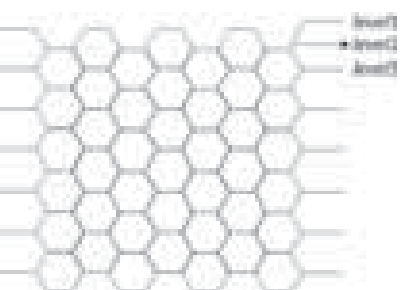
Journal of
Computational and Theoretical Nanoscience
www.aspbs.com/ctn

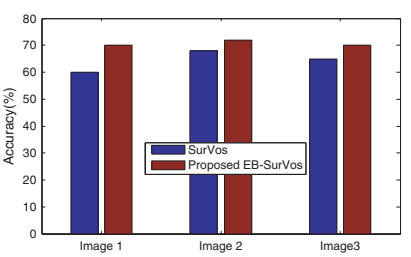
VOLUME 15, NUMBER 5
MAY 2018

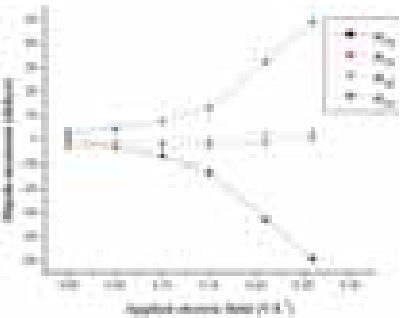
1429–1778 (2018)
ISSN: 1546-1955, EISSN: 1546-1963

ON THE COVER:

RESEARCH ARTICLES

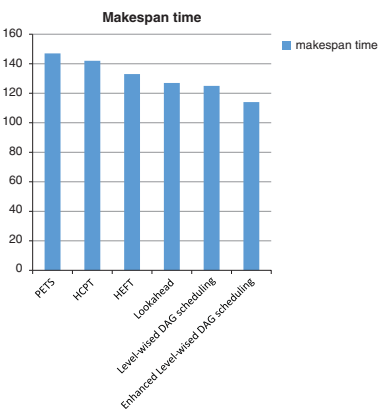


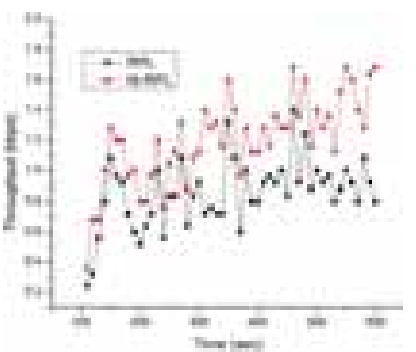




A SPECIAL SECTION

RESEARCH ARTICLES





Journal of
Computational and Theoretical Nanoscience
www.aspbs.com/ctn

Aims and Scope:

Journal of Computational and Theoretical Nanoscience is an international peer-reviewed journal with a wide-ranging coverage, consolidates research activities in all aspects of computational and theoretical nanoscience into a single reference source. This journal offers scientists and engineers peer-reviewed research papers in all aspects of computational and theoretical nanoscience and nanotechnology in chemistry, physics, materials science, engineering and biology to publish original full papers and timely state-of-the-art reviews and short communications encompassing the fundamental and applied research.

Research topics covered (alphabetical listing):

Assemblers, basic physics, biological systems, biochemical systems, bionics, biophysics, CAD, carbon systems, cellular mechanisms, chaotic systems, circuits, clusters, cluster systems, complex aggregates, computer codes, crystal growth, data analysis, defined chain length molecules, devices, diffusion processes, DNA, drug design, dynamics, electronics, electronic properties, enzyme reactivity and reactions, equation of state, friction, computational genomics, gene technology, genetics, holistic views, information theory, interactions, ion channelling, kinetics, macromolecules, molecular interactions, large scale simulations, liquids, liquid crystals, luminescence, magnetic structures, manufacturing, many-particle systems, metallurgy, materials, material properties, mechanical models, metals, mathematical methods, molecule design, molecular dynamics, molecular mechanics, Monte Carlo simulations, multiscale methods, nanomachines, nano-optics, nanorobotics, nanotechnology and ethics, noble gases, non-linear optics, numerical algorithm, numerical procedures, oligomers, optoelectronics, phase transitions, phenomenological theory, philosophical implications and positions, photonic crystals, polymers, potential development, protein folding, quantum chemistry, quantum computers, quantum dots, quantum electronics and optics, quantum technology, replicators, RNA, semiconductors, superconductors, solid state physics, statistical physics, structural chemistry, structures, structures on surfaces, surfaces, technological applications, theoretical biosciences, theoretical physics, thermodynamics thought experiments, wear, and much more.

Instructions for Authors:

Please visit the journal Web site (http://www.aspbs.com/ctn) or see the instructions in each issue of the journal. Authors should follow the instructions when submitting their manuscripts.

Journal Policy:

It is journal policy to publish only original and unpublished research work therefore Journal of Computational and Theoretical Nanoscience does not wish to receive any papers on research work that has already been reported in parts or contains already published text, data, figures, tables or other illustrations or any copyright materials whatsoever that has been submitted or accepted for publication either in a journal or conference proceedings elsewhere in any form, in print or in electronic media.

Manuscript Submission:

Authors should submit manuscript electronically either as a PDF or Microsoft Word file to the Editor-in-Chief.

Abstracting and Indexing

Chemical Abstracts, Chemistry Citation Index (CCI), Research Alert, Materials Science Citation Index (MSCI), Current Contents/Physical, Chemical & Earth Sciences (CC/PC&ES), Science Direct, Scopus, Compendex, Engineering Index (EI), Cambridge Scientific Abstracts/METADEX-Engineered Materials Abstracts, Biological Sciences Abstracts, Biotechnology and Bio-Engineering Abstracts, Biotechnology Research Abstracts, Bacteriology Abstracts (Microbiology B), Neurosciences Abstracts, Aerospace Database, Advanced Technologies Database with Aerospace, Aluminium Industry Abstracts, ANTE: Abstracts in New Technology & Engineering, Ceramic Abstracts, Civil Engineering Abstracts, Computer & Information Systems Abstracts, Copper Technical Reference Library, Corrosion Abstracts, Earthquake Engineering Abstracts, Electronics and Communications Abstracts, Engineered Materials, Materials Research Database, Materials Business File, Meteorological & Geoastrophysical Abstracts, Mechanical & Transportation Engineering Abstracts, Solid State and Superconductivity Abstracts, Engineering Research Database, Technology Research Database, Environmental Science and Pollution Management.

Books for Review:

Publications should be sent to the: *Journal of Computational and Theoretical Nanoscience*, American Scientific Publishers, 26650 The Old Road, Suite 208, Valencia, California 91381, USA; Tel.: (661) 799-7200; Fax: (661) 799-7230; E-mail: jctn@aspbs.com. Readers are encouraged to send books that may be of interest to other readers of *Journal of Computational and Theoretical Nanoscience*.

Referee's Report:

Please submit Reviewer's Reports on the *Journal of Computational and Theoretical Nanoscience* home page.

Annual Subscription Rates (Print Edition):

Volume 15, 2018, 12 issues
Institutional: US$ 3440 (USA), US$ 3440 (Foreign)
Postage and handling charges: US$ 300 (USA), US$ 600 (Foreign).
The publisher reserves the right to refuse nonqualified subscriptions.

Web Edition:

Journal of Computational and Theoretical Nanoscience is available to subscribers via the internet. For institutional subscriptions to the Web Edition or site licensing, please contact publisher.

Subscription Orders, Advertising, Renewals and Reprints:

For quotes, and other details, contact: American Scientific Publishers
Tel.: (661) 799-7200 Fax: (661) 799-7230
E-mail: order@aspbs.com

Honorable Claim Period:

Claims for issues not received will be honored only if submitted within 60 days of the issue date for subscribers in North America or 120 days for all other subscribers. Send your written request to American Scientific Publishers.

Digital Object Identifier (DOI):

The DOI identification system for digital media has been designed to provide persistent and reliable identification of digital objects. Information on the DOI and its governing body, the International DOI Foundation, can be found at http://www.doi.org. The DOI appears in the lower right of the first page of each article.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1429–1433, 2018

The Second *ABC* Index and Second *GA* Index of $TUAC_6[p, q]$

Hualong Wu[1], Mohammad Reza Farahani[2], Bo Zhao[1,*], and Wei Gao[1,*]

[1] School of Information Science and Technology, Yunnan Normal University, Kunming, 650500, China
[2] Department of Applied Mathematics, Iran University of Science and Technology, Narmak, Tehran, 16844, Iran

The atom-bond connectivity index and geometric-arithmetic index are vital distance-based topological indices which reveal certain structural features and chemical characteristics of organic molecules. In this report, in terms of the molecular structural analysis and mathematical derivation, we consider the second version of atom-bond connectivity index and geometric-arithmetic index for special classes of armchair polyhex nanotube. The computational formulas for calculating these indices of $TUAC_6[p, q]$ are presented.

Keywords: Theoretical Chemistry, Second Atom-Bond Connectivity Index, Second Geometric-Arithmetic Index, Armchair Polyhex Nanotube.

RESEARCH ARTICLE

1. INTRODUCTION

The studies of topological indices for molecular structures have been conducted for over 35 years. Degree-based topological indices are numerical parameters of molecular structure, and play important roles in physics, chemistry and pharmacology science.

Specifically, let G be a molecular graph, then a topological index can be regarded as a positive real function $f: G \to \mathbb{R}^+$. As numerical descriptors of the molecular structure deduced from the corresponding molecular graph, topological indices have found several applications in theoretical chemistry, like QSPR/QSAR study. For instance, harmonic index, Wiener index, PI index, Randic index and sum connectivity index are introduced to reflect certain structural features and chemical characteristics of organic molecules. Recently, several articles contributed to reporting certain distance-based or degree-based indices of special molecular graph (See Yan et al.,[1,2] Gao et al.,[3,4] Gao and Shi,[5] Gao and Wang,[6] Xi and Gao,[7,8] Xi et al.,[9] Gao et al.,[10] for more detail). The notation and terminology used but undefined in this paper can be found in Ref. [11].

The atom-bond connectivity index (shortly, *ABC* index) is defined by Estrada et. al.,[12] as

$$ABC(G) = \sum_{uv \in E(G)} \sqrt{\frac{d(u) + d(v) - 2}{d(u)d(v)}}$$

where $d(u)$ denotes the degree of vertex $u \in V(G)$.

Let $e = uv$ be an edge of the molecular graph G. The number of vertices of G whose distance to the vertex u is smaller than the distance to the vertex v is denoted by $n_u(e)$. Analogously, $n_v(e)$ is the number of vertices of G whose distance to the vertex v is smaller than the distance to the vertex u. Graovac and Ghorbani[13] defined a new version of the atom-bond connectivity index, i.e., the second atom bond connectivity index:

$$ABC_2(G) = \sum_{uv \in E(G)} \sqrt{\frac{n_u(e) + n_v(e) - 2}{n_u(e)n_v(e)}}$$

Rostami et al.,[14] calculated some upper bounds for the second atom-bond connectivity index. Gao and Wang[6] determined the second atom-bond connectivity index of unilateral polyomino chain and unilateral hexagonal chain. Also, they deduced the second *ABC* indices of V-phenylenic nanotubes and nanotori.

By considering the degrees of vertices in G, Vukicevic and Furtula[15] developed the Geometric-arithmetic index, shortly *GA* index, which is defined by

$$GA(G) = \sum_{uv \in E(G)} \frac{2\sqrt{d(u)d(v)}}{d(u) + d(v)}$$

Recently, Fath-Tabar et al.,[16] defined a new version of the geometric-arithmetic index, i.e., the second geometric-arithmetic index:

$$GA_2(G) = \sum_{uv \in E(G)} \frac{2\sqrt{n_u(e)n_v(e)}}{n_u(e) + n_v(e)}$$

In Zhan and Qiao.,[17] the maximum and the minimum second geometric-arithmetic index of the star-like tree are

*Authors to whom correspondence should be addressed.

learned in view of an increasing or decreasing transformation of the second geometric arithmetic index of trees. Furthermore, they determine the corresponding extremal trees.

Gao and Wang[18] yielded the computational formulas for calculating the second geometric-arithmetic index and introduced the general second geometric-arithmetic index as follows:

$$GA_2^{\gamma}(G) = \sum_{uv \in E(G)} \left(\frac{2\sqrt{n_u(e) n_v(e)}}{n_u(e) + n_v(e)} \right)^{\gamma}$$

where γ is a real number.

In the past 10 years, the computation of degree-based and distance-based indices for special structure of chemical molecular had raised many attention among researchers. Although there have been several advances in degree based indices of molecular graphs, the study of distance-based indices for special chemical structures are still largely limited. In addition, as widespread and critical chemical structures, armchair polyhex nanotubes are widely used in medical science and pharmaceutical field. For these reasons, many kinds of indices for special structures of armchair polyhex nanotubes have been studied in recent years.

Farahani[19] presented the fourth atom-bond connectivity index of armchair polyhex nanotube $TUAC_6$. Farahani[20] determined the fifth *GA* index of armchair polyhex nanotube. Eliasi and Taeri,[21] reported the Szeged index of armchair polyhex nanotubes. Eliasi and Salehi[22] computed the Schultz index of armchair polyhex nanotubes $TUVC_6[2p, q]$. Iranmanesh and Ashrafi[23] calculated the Balaban index of an armchair polyhex, $TUC_4C_8(R)$ and $TUC_4C_8(S)$ nanotorus. Yousefi and Ashrafi,[24] and Eliasi and Taeri[25] studied the distance of armchair polyhex nanotubes. Deng[26] raised the PI index of $TUVC_6[2p, q]$. Saheli and Ashrafi[27] discussed the eccentric connectivity index of armchair polyhex nanotubes $TUVC_6[p, q]$. Badakhshian and Loghman[28] presented the PI polynomial of armchair polyhex nanotubes and nanotorus. Diudea et al.,[29] considered the Wiener index of armchair polyhex nanotubes.

The contribution of our paper is to determine the computational formulas for calculating the second atom-bond connectivity index and second geometric-arithmetic index of nanotubes $TUAC_6[p, q]$.

2. MAIN RESULTS AND PROOFS

We consider the armchair polyhex nanotube $TUAC_6[p, q]$, where p is the number of horizontal lines and q is the number of levels. It has $2pq$ vertices, where the number of vertices with degree 2 and the number of vertices with degree 3 are $4p$ and $2p(q-2)$, respectively.

Let e be an arbitrary edge of armchair polyhex nanotube $TUAC_6[p, q]$. For convenience, we assume that $n_1(e)$ counts the vertices of G lying closer to one vertex than to other vertex and the notation $n_2(e)$ is analogues. The structure of $TUAC_6[p, q]$ can refer to Figure 1 for more details.

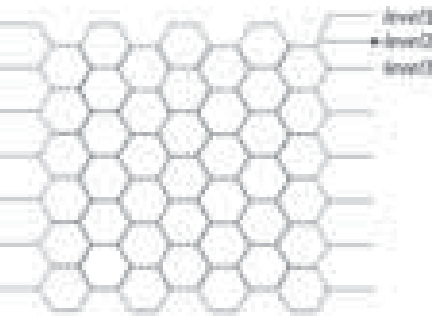


Fig. 1. The structure of $TUAC_6[p, q]$.

CLAIM 1. Let e be a horizontal edge of $TUAC_6[p, q]$. Then, $n_1(e) = n_2(e) = pq$.

PROOF. W. L. O. G., set $e = uv$ and R_1, R_2 are the regions have the vertices that belong to $n_1(e)$ and $n_2(e)$, respectively (see Fig. 2). Hence, we get $n_1(e) = n_2(e) = pq$. The result is obtained by the symmetry of $TUAC_6[p, q]$ for each horizontal edges. □

Set $a = \lfloor (q-m-1)/2 \rfloor$ and $b = \lfloor (m-1)/2 \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function and m is a label for level.

CLAIM 2. Assume that $p \equiv 0 (\mathrm{mod}\ 2)$ and e is an oblique edge in level m $(1 \le m \le q)$. We infer

(1) If $m \le p$ and $q - m \le p$, then

$$n_1(e) = p(q+m-1) + 2b(5-2m+3b-p) + 2a(q-m-a-2) + 2q - 6m + 2 = S_1 \quad (1)$$

(2) If $m \le p$ and $q - m > p$, then

$$n_1(e) = p\left(2q - \frac{p}{2} - 1\right) + 2b(5-p+3b-2m) - 4m + 4 = S_2 \quad (2)$$

(3) If $m > p$ and $q - m \le p$, then

$$n_1(e) = p\left(q - m + \frac{p}{2}\right) + 2a(q-a-m-2) + 2q - 2m - 2 = S_3 \quad (3)$$

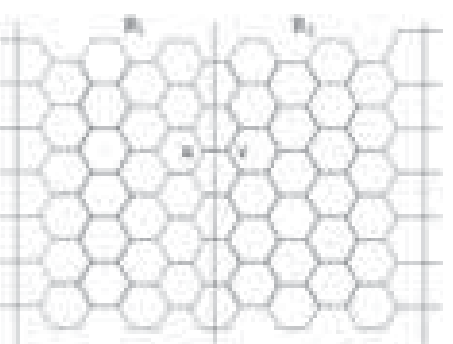


Fig. 2. The horizontal edge $e = uv$.

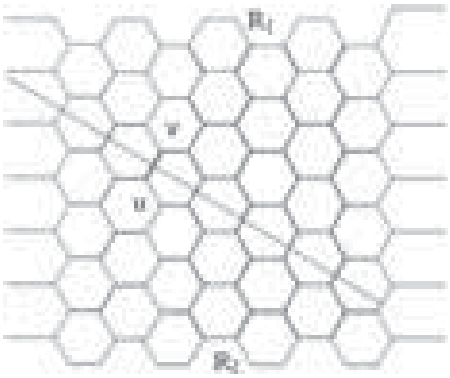


Fig. 3. The oblique edge $e = uv$.

(4) If $m > p$ and $q - m > p$, then

$$n_1(e) = 2p(q-m) + 2 = S_4 \tag{4}$$

Proof. Suppose $e = uv$ is an oblique edge of $TUAC_6[p, q]$ and R_1, R_2 are the regions have the vertices that belong to $n_1(e)$ and $n_2(e)$, respectively (see Fig. 3).

- If $m \le p$ and $q - m \le p$, then

$$\begin{aligned} n_1(e) &= p(q-m+1) + \sum_{i=1}^{a} 4i + \sum_{i=1}^{b}(2p-4i) \\ &\quad + (q-m-2a-1)(2a+2) + (m-2b-1) \\ &\quad \times (2p-4b-4) \\ &= p(q+m-1) + 2b(5-2m+3b-p) \\ &\quad + 2a(q-m-a-2) + 2q - 6m + 2 \end{aligned}$$

- If $m \le p$ and $q - m > p$, then

$$\begin{aligned} n_1(e) &= p(2q-2m-p+2) + \sum_{i=1}^{(p-2)/2} 4i + \sum_{i=1}^{b}(2p-4i) \\ &\quad + (m-2b-1)(2p-4b-4) \\ &= p\left(2q - \frac{p}{2} - 1\right) + 2b(5-p+3b-2m) - 4m + 4 \end{aligned}$$

- If $m > p$ and $q - m \le p$, then

$$\begin{aligned} n_1(e) &= p(q-m+1) + \sum_{i=1}^{a} 4i + \sum_{i=1}^{(p-2)/2}(2p-4i) \\ &\quad + (q-m-2a-1)(2a+2) + 2 \\ &= p\left(q - m + \frac{p}{2}\right) + 2a(q-a-m-2) \\ &\quad + 2q - 2m - 2 \end{aligned}$$

- If $m > p$ and $q - m > p$, then

$$\begin{aligned} n_1(e) &= p(2q-2m-p+2) + \sum_{i=1}^{(p-2)/2} 4i \\ &\quad + \sum_{i=1}^{(p-2)/2}(2p-4i) + 2 = 2p(q-m) + 2 \end{aligned}$$

Therefore, the results are obtained by the symmetry of $TUAC_6[p, q]$ nanotube for each oblique edge. □

Claim 3. Assume that $p \equiv 1 (\mathrm{mod}\ 2)$ and e is an oblique edge in level m $(1 \le m \le q)$. We infer

(1) If $m \le p$ and $q - m \le p$, then

$$\begin{aligned} n_1(e) &= p(q+m-1) + 2b(5-2m+3b-p) \\ &\quad + 2a(q-m-a-2) + 2q - 6m + 2 = S_5 \end{aligned} \tag{5}$$

(2) If $m \le p$ and $q - m > p$, then

$$\begin{aligned} n_1(e) &= p\left(2q - \frac{p}{2} - 1\right) + 2b(5-p+3b-2m) - 4m + \frac{7}{2} \\ &= S_6 \end{aligned} \tag{6}$$

(3) If $m > p$ and $q - m \le p$, then

$$\begin{aligned} n_1(e) &= p\left(q - m + \frac{p}{2}\right) + 2a(q-a-m-2) \\ &\quad + 2q - 2m - \frac{3}{2} = S_7 \end{aligned} \tag{7}$$

(4) If $m > p$ and $q - m > p$, then

$$n_1(e) = 2p(q-m) = S_8 \tag{8}$$

Proof. The proof is similar to the proof of Claim 2. □

Remark 1. By Figure 3, let e be an oblique edge in level m $(1 \le m \le q)$, then we have $n_1(e) + n_2(e) = 2pq$.

Theorem 1. *If p is even, we have*

(1) *If $q \le p$, then*

$$ABC_2(TUAC_6[p,q]) = \sqrt{2pq-2} + 2p\sum_{m=1}^{q-1}\sqrt{\frac{2pq-2}{S_1(2pq-S_1)}},$$

$$GA_2(TUAC_6[p,q]) = pq + 2\sum_{m=1}^{q-1}\frac{\sqrt{S_1(2pq-S_1)}}{q},$$

$$GA_2^{\gamma}(TUAC_6[p,q]) = pq + 2p\sum_{m=1}^{q-1}\left(\frac{\sqrt{S_1(2pq-S_1)}}{pq}\right)^{\gamma}$$

(2) *If $p < q \le 2p$, then*

$$\begin{aligned} &ABC_2(TUAC_6[p,q]) \\ &= \sqrt{2pq-2} + 2p\left\{\sum_{m=1}^{q-p-1}\sqrt{\frac{2pq-2}{S_2(2pq-S_2)}},\right. \\ &\quad \left. + \sum_{m=q-p}^{p}\sqrt{\frac{2pq-2}{S_1(2pq-S_1)}} + \sum_{m=p+1}^{q-1}\sqrt{\frac{2pq-2}{S_3(2pq-S_3)}}\right\} \end{aligned}$$

$$\begin{aligned} &GA_2(TUAC_6[p,q]) \\ &= pq + \frac{2}{q}\left\{\sum_{m=1}^{q-p-1}\sqrt{S_2(2pq-S_2)} + \sum_{m=q-p}^{p}\sqrt{S_1(2pq-S_1)}\right. \\ &\quad \left. + \sum_{m=p+1}^{q-1}\sqrt{S_3(2pq-S_3)}\right\}, \end{aligned}$$

$$GA_2^{\gamma}(TUAC_6[p,q]) = pq+2p\left\{\sum_{m=1}^{q-p-1}\left(\frac{\sqrt{S_2(2pq-S_2)}}{pq}\right)^{\gamma} + \sum_{m=q-p}^{p}\left(\frac{\sqrt{S_1(2pq-S_1)}}{pq}\right)^{\gamma} + \sum_{m=p+1}^{q-1}\left(\frac{\sqrt{S_3(2pq-S_3)}}{pq}\right)^{\gamma}\right\}$$

(3) *If* $q > 2p$, *then*

$$ABC_2(TUAC_6[p,q]) = \sqrt{2pq-2}+2p\left\{\sum_{m=1}^{q-p-1}\sqrt{\frac{2pq-2}{S_2(2pq-S_2)}} + \sum_{m=q-p}^{p}\sqrt{\frac{2pq-2}{S_4(2pq-S_4)}}+\sum_{m=p+1}^{q-1}\sqrt{\frac{2pq-2}{S_3(2pq-S_3)}}\right\},$$

$$GA_2(TUAC_6[p,q]) = pq+\frac{2}{q}\left\{\sum_{m=1}^{q-p-1}\sqrt{S_2(2pq-S_2)} + \sum_{m=q-p}^{p}\sqrt{S_4(2pq-S_4)}+\sum_{m=p+1}^{q-1}\sqrt{S_3(2pq-S_3)}\right\},$$

$$GA_2^{\gamma}(TUAC_6[p,q]) = pq+2p\left\{\sum_{m=1}^{q-p-1}\left(\frac{\sqrt{S_2(2pq-S_2)}}{pq}\right)^{\gamma} + \sum_{m=q-p}^{p}\left(\frac{\sqrt{S_4(2pq-S_4)}}{pq}\right)^{\gamma} + \sum_{m=p+1}^{q-1}\left(\frac{\sqrt{S_3(2pq-S_3)}}{pq}\right)^{\gamma}\right\}$$

Proof. Let A and B be the sets of all horizontal and oblique edges of $\mathrm{TUAC}_6[p,q]$, respectively. Then, in view of the definition of second atom-bond connectivity index, we have

$$ABC_2(\mathrm{TUAC}_6[p,q]) = \sum_{e\in A}\sqrt{\frac{n_1(e)+n_2(e)-2}{n_1(e)n_2(e)}} + \sum_{e\in B}\sqrt{\frac{n_1(e)+n_2(e)-2}{n_1(e)n_2(e)}}$$

The number of horizontal edges are pq. Thus, we have

$$\sum_{e\in A}\sqrt{\frac{n_1(e)+n_2(e)-2}{n_1(e)n_2(e)}} = \sqrt{2pq-2}$$

The number of oblique edges are $2p(q-1)$.

- If $q \le p$, then

$$\sum_{e\in B}\sqrt{\frac{n_1(e)+n_2(e)-2}{n_1(e)n_2(e)}} = 2p\sum_{m=1}^{q-1}\sqrt{\frac{2pq-2}{S_1(2pq-S_1)}}$$

- If $p < q \le 2p$, then

$$\sum_{e\in B}\sqrt{\frac{n_1(e)+n_2(e)-2}{n_1(e)n_2(e)}} = 2p\left\{\sum_{m=1}^{q-p-1}\sqrt{\frac{2pq-2}{S_2(2pq-S_2)}}+\sum_{m=q-p}^{p}\sqrt{\frac{2pq-2}{S_1(2pq-S_1)}} + \sum_{m=p+1}^{q-1}\sqrt{\frac{2pq-2}{S_3(2pq-S_3)}}\right\}$$

- If $q > 2p$, then

$$\sum_{e\in B}\sqrt{\frac{n_1(e)+n_2(e)-2}{n_1(e)n_2(e)}} = 2p\left\{\sum_{m=1}^{q-p-1}\sqrt{\frac{2pq-2}{S_2(2pq-S_2)}}+\sum_{m=q-p}^{p}\sqrt{\frac{2pq-2}{S_4(2pq-S_4)}} + \sum_{m=p+1}^{q-1}\sqrt{\frac{2pq-2}{S_3(2pq-S_3)}}\right\}$$

Similarly, the remaining results are obtained by the definition of second geometric-arithmetic index and its general version. □

Theorem 2. *If* p *is odd, we have*

(1) *If* $q \le p$, *then*

$$ABC_2(TUAC_6[p,q]) = \sqrt{2pq-2}+2p\sum_{m=1}^{q-1}\sqrt{\frac{2pq-2}{S_5(2pq-S_5)}},$$

$$GA_2(TUAC_6[p,q]) = pq+2\sum_{m=1}^{q-1}\frac{\sqrt{S_5(2pq-S_5)}}{q},$$

$$GA_2^{\gamma}(TUAC_6[p,q]) = pq+2p\sum_{m=1}^{q-1}\left(\frac{\sqrt{S_5(2pq-S_5)}}{pq}\right)^{\gamma}$$

(2) *If* $p < q \le 2p$, *then*

$$ABC_2(TUAC_6[p,q]) = \sqrt{2pq-2}+2p\left\{\sum_{m=1}^{q-p-1}\sqrt{\frac{2pq-2}{S_6(2pq-S_6)}} + \sum_{m=q-p}^{p}\sqrt{\frac{2pq-2}{S_5(2pq-S_5)}}+\sum_{m=p+1}^{q-1}\sqrt{\frac{2pq-2}{S_7(2pq-S_7)}}\right\},$$

$$GA_2(TUAC_6[p,q]) = pq+\frac{2}{q}\left\{\sum_{m=1}^{q-p-1}\sqrt{S_6(2pq-S_6)}+\sum_{m=q-p}^{p}\sqrt{S_5(2pq-S_5)} + \sum_{m=p+1}^{q-1}\sqrt{S_7(2pq-S_7)}\right\},$$

$$GA_2^{\gamma}(TUAC_6[p,q]) = pq+2p\left\{\sum_{m=1}^{q-p-1}\left(\frac{\sqrt{S_6(2pq-S_6)}}{pq}\right)^{\gamma} + \sum_{m=q-p}^{p}\left(\frac{\sqrt{S_5(2pq-S_5)}}{pq}\right)^{\gamma} + \sum_{m=p+1}^{q-1}\left(\frac{\sqrt{S_7(2pq-S_7)}}{pq}\right)^{\gamma}\right\}$$

(3) *If* $q > 2p$, *then*

$$ABC_2(TUAC_6[p,q]) = \sqrt{2pq-2}+2p\left\{\sum_{m=1}^{q-p-1}\sqrt{\frac{2pq-2}{S_6(2pq-S_6)}} + \sum_{m=q-p}^{p}\sqrt{\frac{2pq-2}{S_8(2pq-S_8)}}+\sum_{m=p+1}^{q-1}\sqrt{\frac{2pq-2}{S_7(2pq-S_7)}}\right\},$$

$$GA_2(TUAC_6[p,q]) = pq+\frac{2}{q}\left\{\sum_{m=1}^{q-p-1}\sqrt{S_6(2pq-S_6)} + \sum_{m=q-p}^{p}\sqrt{S_8(2pq-S_8)}+\sum_{m=p+1}^{q-1}\sqrt{S_7(2pq-S_7)}\right\},$$

$$GA_2^{\gamma}(TUAC_6[p,q]) = pq+2p\left\{\sum_{m=1}^{q-p-1}\left(\frac{\sqrt{S_6(2pq-S_6)}}{pq}\right)^{\gamma} + \sum_{m=q-p}^{p}\left(\frac{\sqrt{S_8(2pq-S_8)}}{pq}\right)^{\gamma} + \sum_{m=p+1}^{q-1}\left(\frac{\sqrt{S_7(2pq-S_7)}}{pq}\right)^{\gamma}\right\}$$

PROOF. The proof is similar to the proof of Theorem 1. □

3. CONCLUSION

In our article, by virtue of the molecular graph structural analysis and mathematical derivation, we mainly report the second version of atom-bond connectivity index and geometric-arithmetic index of armchair polyhex nanotube $TUAC_6[p, q]$. The theoretical results achieved in this article illustrate the promising prospects of the application for the chemical and pharmacy engineering.

Acknowledgments: We thank the reviewers for their constructive comments in improving the quality of this paper. This work was supported in part by the National Natural Science Foundation of China (11401519).

References

1. L. Yan, Y. Li, W. Gao, and J. S. Li, *Journal of Chemical and Pharmaceutical Research* 6, 477 **(2014)**.
2. L. Yan, W. Gao, and J. S. Li, *J. Comput. Theor. Nanosci.* 12, 3940 **(2015)**.
3. W. Gao, L. Liang, and Y. Gao, *Energy Education Science and Technology: Part A* 32, 8961 **(2014)**.
4. W. Gao, L. Liang, and Y. Gao, *The BioTechnology: An Indian Journal* 10, 3837 **(2014)**.
5. W. Gao and L. Shi, *Asian Journal of Chemistry* 26, 3397 **(2014)**.
6. W. Gao and W. F. Wang, *Journal of Chemistry* 2014, Article ID 906254.
7. W. F. Xi and W. Gao, *Journal of Advances in Chemistry* 10, 2254 **(2014)**.
8. W. F. Xi and W. Gao, *Journal of Applied Computer Science and Mathematics* 18, 43 **(2014)**.
9. W. F. Xi, W. Gao, and Y. Li, *Journal of Advances in Mathematics* 9, 2696 **(2014)**.
10. Y. Gao, W. Gao, and L. Liang, *International Journal of Applied Physics and Mathematics* 4, 417 **(2014)**.
11. J. A. Bondy and U. S. R. Mutry, Graph Theory, Spring, Berlin **(2008)**.
12. E. Estrada, L. Torres, L. Rodriguez, and I. Gutman, *Indian Journal of Chemistry* 37A, 849 **(1998)**.
13. A. Graovac and M. Ghorbani, *Acta Chimica Slovenica* 57, 609 **(2010)**.
14. M. Rostami, M. S. Haghighat, and M. Ghorbani, *Iranian Journal of Mathematical Chemistry* 4, 265 **(2013)**.
15. D. Vukicevic and B. Furtula, *Journal of Mathematical Chemistry* 4, 1369 **(2009)**.
16. G. F. Tabar, B. Purtula, and I. Gutman, *Journal of Mathematical Chemistry* 47, 477 **(2010)**.
17. F. Q. Zhan and Y. F. Qiao, *Mathematics in Practice and Theory* 44, 226 **(2014)**.
18. W. Gao and W. F. Wang, *Annals of Pure and Applied Mathematics* 13, 99 **(2017)**.
19. M. R. Farahani, *The Publishing House of the Romanian Academy* 15, 3 **(2013)**.
20. M. R. Farahani, *Le Mathematiche* LXIX, 69 **(2014)**.
21. M. Eliasi and B. Taeri, *MATCH Commun. Math. Comput. Chem.* 59, 437 **(2008)**.
22. M. Eliasi and N. Salehi, *International Journal of Molecular Sciences* 9, 2016 **(2008)**.
23. A. Iranmanesh and A. R. Ashrafi, *J. Comput. Theor. Nanosci.* 4, 514 **(2007)**.
24. S. Yousefi and A. R. Ashrafi, *Studia Univ. Babes-Bolyai, Chemia* 53, 111 **(2008)**.
25. M. Eliasi and B. Taeri, *MATCH Commun. Math. Comput. Chem.* 62, 295 **(2009)**.
26. H. Deng, *MATCH Commun. Math. Comput. Chem.* 55, 461 **(2006)**.
27. M. Saheli and A. R. Ashrafi, *Macedonian Journal of Chemistry and Chemical Engineering* 29, 71 **(2010)**.
28. L. Badakhshian and A. Loghman, *Digest Journal of Nanomaterials and Biostructures* 4, 183 **(2009)**.
29. M. V. Diudea, M. Stefu, B. Pârv, and P. E. John, *Croatica Chemica Acta* 77, 111 **(2004)**.

Received: 16 March 2016. Accepted: 29 March 2016.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1434–1440, 2018

RESEARCH ARTICLE

Enhanced Bat Algorithm with Super-Region Volume Segmentation Workbench for Markov Random Field Images

B. Basavaprasad[1,*] and Ravindra S. Hegadi[2]

[1] *Research Scholar, Research and Development Centre, Bharathiar University, Coimbatore 641046, Tamilnadu, India*
[2] *Department of Computer Science, Solapur University, Solapur 413255, Maharashtra, India*

Biological volume's separation is a critical step required completely in order to examine their scientific content. With the intention of consolidating the automatic approaches of power of the segmentation with the experienced and capable of annotating biological samples manually, provides a new board which is termed as Super-Region Volume Segmentation (SuRVoS). SuRVoS purifies the estimation by utilizing Markov Random Field (MRF) formulation, which in turn considers neighboring labels in order to estimate precisely consistent. Nevertheless, in the MRF formulation model, various separations play a vital part in order to raise the exactness of output of the segmentation. Hence, identifying that partition number is an important task and also it is a difficult problem. In proposed work, the edge weight is segregated into two disjoints sets, which means cut and off into all feasible disjoint segments. The proposed energy function is reduced by utilizing Enhanced Bat (EB) algorithm over the MRF and the standard deviation calculation among two edges. In our work, the EB-SuRVoS gives good segmentation output by generating good objective function values in the form of best supervoxels. So, it raises the accuracy of image segmentation greater instead of the earlier methods. The performance metrics were computed namely accuracy, precision, recall, Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE) and correlation values by utilizing the proposed method. The experimental output gives the greater segmentation accuracy which is accomplished by the proposed method.

Keywords: Segmentation, EB-SuRVoS, MRF, Super-Regions.

1. INTRODUCTION

Biological tissue was frequently collected by cells with same morphologies which are redundant all over the huge volumes and various biological applications depends on the exact detection these cells and their position from image data. Biological imaging techniques have navigated from two-dimensional to three-dimensional via tomography, or serial section/imaging to give spatial context, specifically for entire cells. In biological volume analysis, data is frequently denoted as a 3D grid of voxels with every voxel denoted as a consumption of few probe (X-rays or electrons) at that point. Proceeding these volumes were hard because of their low signal-to-noise ratio and low contrast, in the case of cryo-immobilized samples, and highly difficult datasets.[1]

A novel region-based active contour model (ACM) is executed in Ref. [2] for image segmentation, with a new level set. In this turn minimizes the lavish re-initialization of the conventional level set method to create as highly effective. This model is enforced with the method which links the advantages of the conventional geodesic active contours (GAC) and Chan–Vese (C–V) models, which constitutes the belongings of local or global segmentation. Extensive experiments on synthetic and real images explains the merits of the proposed method over the classical ACMs with the conventional level set methods like the GAC and C–V models. This method is potent which can be enforced for achieving the algorithms of few classical ACMs, namely GAC model, C–V model,[3] PS model and LBF model.

Random forest utilizes local features to classifiers by utilizing the adjacent neighbor who matches with the class models. This classifier can enhance the classification achievement of pre-learned class models, because the local features can combine spatial context, and the forest permits multi-class discriminative learning.[4] It gives the

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1434/007 doi:10.1166/jctn.2018.6972

computational effectiveness in both the training and classification and also it manages the huge collection of the visual features (e.g., color, texture, shape, depth etc.).

Gradient boosting is a highly matured algorithm to update the boosting classifiers online and frequently enhances the exactness of the segmentation.[5] In order to manage the labeling uncertainty for semi-supervised learning and it explains the online gradient boosting technique which depends on the various instance learning. Weak classifiers in boosting were usually easy linear functions but decision trees were at time utilized for good generalization;[6] the output of the classifier is termed as gradient boosting decision tree.

SVMs have been tested and measured as pixel-based image classifiers.[7] Navigating from pixel-based techniques to object-based representation are the dimensions of remote sensing imagery feature space, which gets raised considerably. The output of this, raises the difficulty of the classification process, and creates issues to the conventional sample-based classification schemes. The aim is to measure SVMs for effectiveness and anticipate for the object-based image classification like a modern computational intelligence method. The segmentation algorithm creates primitive objects of variable sizes and shapes. Then, choosing the feature is a step which takes place to give the features for classification, which, in turn complicates spectral, texture and shape information. Contextual information is not utilized. Next to the feature selection process, a module incorporated to an SVM classifier and the segmentation algorithm. In the proposed system, EB-SuRVoS method is brought-in to execute the good segmentation process for the provided images very effectively. The modules in this research work are data preparation, preprocessing, data representation and model training.

2. RELATED WORK

Kato et al.[8] examines a Markov random field (MRF) image segmentation model in Ref. [1], which targets at joining the color and texture features. The theoretical framework depends on the Bayesian estimation trough via combinatorial optimization (simulated annealing). The segmentation is acquired by segregating the pixels into various pixel classes. These classes are denoted by multi-variant Gaussian distributions. So, the only axiom regarding the nature of the features is that an added ingredients for the Gaussian noise model which is appropriate to explain the feature distribution which is own by the provided class. Here, it utilizes the perceptually uniform color values as color features and a set of Gabor filters as texture features. Gaussian parameters may be estimated by utilizing training data set or from the input image. It gives a parameter estimation method by utilizing the Expectation Maximization (EM) algorithm. Experimental results are given to explain the execution of the method on both synthetic and natural color images.

In Ref. [9] brings a novel interactive framework for segmenting images by utilizing the probabilistic hyper graphs which model the spatial and appearance relations between the image pixels. The probabilistic hyper graph gives a means to position of the image segmentation as a machine learning issues. Specifically, this considers a small set of pixels, which is named as an object and background. The seed pixels were utilized to measure the labels of the unnamed pixels through learning on a hyper graph by reducing the quadratic smoothness term which is created by a hyper graph Laplacian matrix which is liable to the known label restrictions. This gives a natural probabilistic interpretation of this smoothness term.

In Ref. [10] utilizes the hierarchical method in order to drew-out more object from the image. In every stage it selects one sub region that adds more than one object, and it separates it into two sub regions. For the provided image, it enforces the segmentation algorithm. Finally in this stage, it has a external layer which explains the edges of the segmented object. If in case further segmentation is required then it manually select the area created from the earlier step and it enforce the segmentation algorithm again specifically to the selected area. The same way, it separates the appealing part of the image.

In Ref. [11] proposed Bat Algorithm (BA) is a biological algorithm, advanced and BA is defined as highly effective. BA depends on echo locations nature for multilevel thresholds for choosing, which utilizes the maximum entropy criterion. The experimental results explain the BA algorithm which search for multiple thresholds and it is very close to the optimal ones that are defined by the exhaustive search method. When distinguished with the current algorithms, the BA algorithm's Excellency is exceeded. This algorithm is to explain the possibility of BA method for multilevel threshold. Along with this, it gives a new option to the traditional methods because of its modesty and effectiveness.

3. PROPOSED METHODOLOGY

In the proposed methodology, Enhanced Bat Algorithm based Super-Region Volume Segmentation (EB-SuRVoS) is proposed. The general block diagram is explained in Figure 1.

Fig. 1. Overall block diagram of the proposed system.

3.1. Data Collection

Cryo light microscopy is utilized to screen grids for spreading the cell, manually manage the loss and suitable ice thickness prior to soft Xray tomography (SXT) imaging. SXT data were gathered on an UltraXRM-S220c microscope which utilizes the B24 beamline at 500 eV (2.4 nm wavelength) at Diamond Light Source (B24, n.d.). Images were gathered by utilizing the 40 nm zone plate, at 812× magnifications (16 nm/pixel), on a 1 K by 1 K back-thinned direct detection CCD. Tilt series were gathered.

Two samples were utilized as test cases for segmentation with SuRVoS. The first sample is a neuronal-like mammalian cell line. The cells are implemented onto gold finder Transmission Electron Microscopy (TEM) grids with holey carbon support (Ted Pella), grown and differentiated, before cryo-immobilization utilized an EM-GP (Leica) plunge freezer. The second sample is Trypanosoma brucei procyclic cells. The cells were experienced and gently fixed priory in glutaraldehyde to include 200 nm gold fiducials, pipetting onto copper TEM grids with lacey carbon support, and cryo-immobilization using a Mark IV (FEI) plunge freezer.

3.2. SuRVoS Framework

SuRVoS favors the MRC file format (.mrc) and HDF5 file format (.h5, .hdf5) at present. After loading the data into the tool, the variations can be selected mechanically and adapted manually. The user interface of the workbench adds three various regions, the Plugins, the Visualization Pane and the Tool Column. Parameters can be selected and enforced by utilizing the Plugins and evaluated in the Visualisation Pane, while the Tool Column houses shortcuts were utilized in tools often. Multiple RoIs can be generated in the design of bounding cuboids. Remaining actions of the tool can be restricted to the entitled RoI.

This permits the user to choose the RoI to validate the de-noising and textural features in a smaller area before extending then to the entire volume or preferably it permits the user to restrict the automatic segmentation to a particular Region of Interest (RoI).

3.3. Data Pre-Processing

The data pre-processing step links both de-noising and textural feature extraction methods to improve the attitude of the data and create future classification much easier. De-noising adds Gaussian and Total Variation (TV) filters. While Gaussian filters generally creates over-smoothed results, TV methods maintain strong volume edges and it leads to split the volume in piece-wise smooth regions. Hence, a Gaussian-denoised volume is better adapted for super-region extraction, while TV de-noising offers good execution in extracting the features for classification.

Along with these primary filters, Gaussian derivative filters, variations of Gaussians and Laplacian of Gaussians, rotation and scale invariant filters and the eigenvalue of the Hessian matrix and Structure tensor are also accessible because they are highly powerful feature extraction methods that assist detecting concealed characteristics of the data. For instance, a TV filter will be suitable for determining the hue areas like the nucleus from the cytoplasm, while the Gradient Magnitude or Laplacian of Gaussian filters will be suitable for segmenting smaller, more nuanced regions like organelles.

3.4. Data Representation

The volumetric data is denoted as super regions inside a 3-layer hierarchical structure. This structure is a collection of voxels, supervoxels and megavoxels. Every layer is produced by gathering similar and neighboring elements of the earlier layer. So, voxels denotes the standard volume voxels, supervoxels are groups of adjacent voxels gathered into an essential region that reserve strong volume limits (i.e., limits are among the various biological features in the image). Likewise, megavoxels are a set of neighboring supervoxels which has same look. With this hierarchical partitioning, huge regions of the volume own the similar object which was denoted by:a set of thousands of voxels, tens of supervoxels or some megavoxels. This hierarchical structure has various merits when distinguished to the standard voxel grid:

(1) Every layer of the hierarchy denotes the same volume with many fewer elements than the earlier layer, so, minimizes the difficulty of annotating or segmenting the volume by many orders of magnitude.

(2) Supervoxels and megavoxels are outlined to have a strong limits adherence and are therefore able to represent the same biological feature without significant loss of information.

(3) The shape and size parameters of the supervoxels are user definable in order to correctly model the volumes or areas with various physical characteristics.

SuRVoS executes a custom GPU version of SLIC supervoxels which by default drew-out the supervoxels of size $10 \times 10 \times 10$ voxels from the volume. This means that the same volume is denoted by small groups of around 1000 voxels without considerable loss of information as homogeneous regions are denoted by single supervoxels. Further huge regions can be drawn-out as megavoxels. These megavoxels are also drawn-out in 3D and merge large, similar areas while reserving strong edges and small cellular structures as unique entities. Commonly speaking, supervoxels were suitable for segmenting smaller, more differed features, while megavoxels may be more helpful for huge similar regions.

3.5. Model Training

In this section, EB algorithm is brought-in for effective model training. The current system has many noise and fake labels which doesn't have spatial coherence.

The proposed research considers the neighboring labels to create the predictions more spatially consistent.

The MRF formulation

$$E(c) = \sum_{p \in V} \psi_p(C_p) + \lambda \sum_{p,q \epsilon E} w_{pq} \cdot \psi_{pq}(C_p, C_q) \quad (1)$$

From the above functions, similarity weight among voxels or supervoxels is evaluated as variations among the headings of nodes p and q weighted by the amount of boundary voxels they share. This inspires the neighboring nodes to have the same.

In the proposed work, new MRF formulation is performed with the assistance of intelligent behavior. In this work it proceeds the procedure of Bat Algorithm based MRF formulation. Nevertheless, the proposed BAMRF algorithm resolves the energy function into an image-labeling issue to control the label consistency conditions

$$E(c) = \sum_{p,q \in E} \delta_{pq}(c_p, c_q) \quad (2)$$

$$\delta_{pq}(c_p, c_q) = (W_{pq} - \sigma_{C_{\text{not}}})^2 \quad \text{if } c_p \neq c_q \quad (3)$$

$$\delta_{pq}(c_p, c_q) = (W_{pq} - \sigma_{C_{in}})^2 \quad \text{if } c_p = c_q \quad (4)$$

The energy function can be resolved by bat algorithm. This BA enhances the label MRFs by continuously extend the individual label by utilizing the graph cut. Graph cut can identify the optimal expansion if the expansion is sub modular.

Behaviour of Microbats: Most of microbats have progressive ability of echolocation. These bats can diffuse a highly loud and short sound pulse; the echo that gives back from the surrounding objects is obtained by their amazing big auricle. After this, the feedback details of the echo are examined in their subtle brain. They not only segregate the directions for their belonging flight pathway based on the echo, but also it compares the various different insects and barriers in order to search for food and eliminate the collision efficiently in the day or night.

Every bat utilizes echolocation in order to feel the distance, and they also "know" the difference among the food/prey and background obstacles in few magical way.

Bats fly without any order but with velocity Vi at position xi with a fixed frequency $f\min$, varying wavelength λ, and loudness $A0$ to search for prey. They can mechanically adapt the wavelength (or frequency) of their emitted pulses and adapt the rate of pulse emission $r \in [0, 1]$, based on the closeness of their aim.

Though the loudness differs in various ways, we consider that the loudness differs from a large (positive) $A0$ to a minimum constant value $A\min$.

Along with this, for simplicity upcoming approximations were also utilized: Commonly, the frequency f in a range $[f\min, f\max]$ agrees to a range of wavelengths $[\lambda\min, \lambda\max]$. In fact, they just differ in the frequency while fixed in the wavelength λ and consider $f \in [0, f\max]$ in their implementation. This is due to λ and f are related because of the fact that $\lambda f = V$ is constant.

In simulations, they utilize the virtual bats commonly to determine the updated rules of their positions xi and velocities Vi in a D-dimensional search space. The new solutions xti and velocities Vti at time step t are provided by

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta \quad (5)$$

$$v_i^t = v_i^{t-1} + (x_i^t - x_*) f_i \quad (6)$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (7)$$

Where $\beta \in [0, 1]$ is a random vector extracted from a uniform distribution. Here, $x*$ is the current global best location (solution) which is pin-pointed after distinguishing all the solutions between all the n bats.

For the local search part, after choosing the solution between the existing best solutions, a new solution for every bat is created locally by utilizing random walk:

$$x_{\text{new}} = x_{\text{old}} + \varepsilon A_t \quad (8)$$

Where $\varepsilon \in [-1, 1]$ is a random number, while $At = \langle Ati \rangle$ is the average loudness of whole bat simultaneously.

The loudness Ai and the rate ri of pulse emission have to be updated based on the repetition. These formulas are

$$A_i^{t+1} = \alpha A_i^t$$

$$r_i^{t+1} = r_1^0[1 - \exp(-\gamma t) \quad (9)$$

Where α and γ are constants.

Levy Flights: Levy flights are Markov processes, which varies from the regular Brownian motion, whose individual jumps have lengths that are spread with the probability density function (PDF) $\lambda(x)$ decaying at large x as $\lambda(x) = |x| - 1 - \alpha$ with $0 < \alpha$.

(1) Stability: distribution of the sum of independent identically distributed stable random variables equal to distribution of each variable.

(2) Power law asymptotics ("heavy tails").

(3) Generalized Central Limit Theorem: The central limit theorem states that the sum of a number of independent and identically distributed (i.i.d.) random variables with finite variances will tend to a normal distribution as the number of variables grows.

(4) Which has an infinite variance with an infinite mean values

Objective function $f(x) = [x_1, x_2, \ldots x_d]^T$

Initialize the bat population $x_i = (i = 1, 2, \ldots n)$ and v_i

Define pulse frequency f_i and x_i initialize pulse rates r_i and the loudness A_i

While ($t <$ Max number of iterations)

Generate new supervoxels (features) by adjusting frequency, and updating velocities (descriptors) and locations/solutions (7)

RESEARCH ARTICLE

if (rand > r_i)
Select a supervoxels among the best feature solutions
Generate a local solution around the selected best solution
end if
Generate a new solution by flying randomly
If (rand < A_i & $f(x_i) < f(x_*)$)
Accept the new supervoxels remove the noise features
Increase r_i and reduce A_i
End if
Rank the bats and find the current best x_i
End while
Postprocess results and visualization

The EB algorithm is utilized to choose the best supervoxels features from the provided image dataset. The optimal supervoxels are created by providing the best fitness values.

4. EXPERIMENTAL RESULT

For the evaluation of the proposed tumor classification, the DICOM brain images are utilized. A total of 760 images are utilized of which 500 are for training and 260 are for testing. In the training images, 272 are tumor and 228 are non-tumor images while in testing images 150 are tumor and 110 are non-tumor. The size of the image is 256×256 pixels have a total of 65536 pixels size with resolution of 96 ppi. The 3D data are exotic into SuRVoS, where the contrast is adapted and the volume is pruned to the appropriate area. In the pre-processing tab, the data is scaled and suitable supervoxels and megavoxels are detected. The experiments were proceeded through MATLAB. Matlab is a programming language which can be utilized for a broad array of numerical and computing applications. It measures the efficiency of the image compression performance. The computation of the proposed EB-SuRVoS algorithm is distinguished with the current SuRVoS method. To the measure the implementation of the current and the proposed approaches, various parameters were utilized like MSE, PSNR, correlation, accuracy, precision and, recall.

4.1. Accuracy

Accuracy is determined as the exactness of the model and it is evaluated as the sum of actual classification parameters (T_p+T_n) segregated by the total number of classification parameters (T_p+T_n+F_p+F_n)

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{10}$$

From the above Figure 2 the graph explains the accuracy comparison for the provided image dataset. In x-axis the number of images is taken and in y-axis the accuracy is plotted. The experimental output shows that the existing SuRVoS approach has less accuracy and the proposed EB-SuRVoS has greater accuracy. The proposed EB-SuRVoS algorithm is used to extract the optimal features and cluster the similar pixels efficiently. Thus it increases the accuracy higher for given DICOM brain images rather than the existing algorithm.

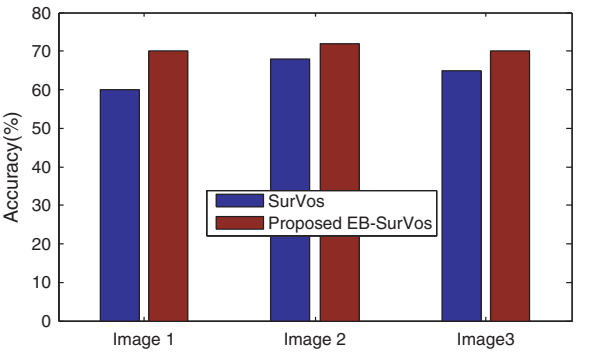


Fig. 2. Accuracy comparison.

4.2. Precision

Precision is determined as the proportion of the true positives against both true positives and false positives output for interruption and the real features. It is determined as follows

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{11}$$

From the above Figure 3 the graph explains the precision metric comparison for the provided image dataset. In x-axis the number of images is taken and in y-axis the precision value is plotted. The experimental output shows that the existing SuRVoS approach has lower precision value and the proposed EB-SuRVoS has higher precision values. The proposed EB-SuRVoS algorithm is used to optimize the significant pixels and cluster the similar pixels optimally. Therefore it provides the higher precision value for given DICOM brain images rather than the existing algorithm.

4.3. Recall

Recall value: Recall value is measured according to the details gathered at the true positive prediction, false

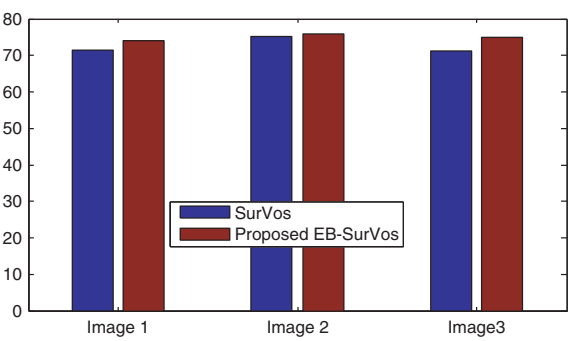


Fig. 3. Precision comparison.

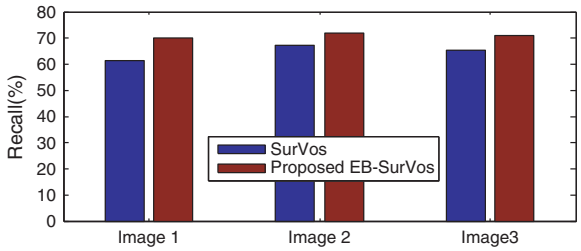


Fig. 4. Recall comparison.

negative. Commonly it can be determined as true positive prediction, false negative. Generally it can be defined as

$$\text{RECALL} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \tag{12}$$

Recall is termed as sensitivity. Recall in gathering the details represents the fraction of the pages which have pertinence to the query which is obtained successfully.

$$\text{RECALL} = \frac{|\{\text{relevant pages}\} \cap \{\text{retrieved pages}\}|}{\{|\text{relevant pages}|\}} \tag{13}$$

From the above Figure 4 the graph explains the recall metric comparison for the provided image dataset. In x-axis the number of images is taken and in y-axis the recall value is plotted. The experimental output shows that the existing SuRVoS approach has lower recall value and the proposed EB-SuRVoS has higher recall values. The proposed EB-SuRVoS algorithm is used to optimize the significant pixels and cluster the similar pixels optimally. Therefore it provides the higher recall value for given DICOM brain images rather than the existing algorithm.

4.4. Correlation

The proximity of two images can be evaluated with the correlation coefficient. The correlation coefficient is assumed by utilizing the following equation

$$\text{corr}\left(\frac{A}{B}\right) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(A_{i,j}-\bar{A})(B_{i,j}-\bar{B})}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(A_{i,j}-\bar{A})^2 \sum_{i=1}^{M}\sum_{j=1}^{N}(B_{i,j}-\bar{B})^2}} \tag{14}$$

From the above Figure 5 the graph explains that the correlation comparison for the provided image dataset. In x-axis the number of images is taken and in y-axis the recall value is plotted. The experimental output shows that the existing SuRVoS approach has lower correlation value and the proposed EB-SuRVoS has higher correlation values. The proposed EB-SuRVoS algorithm is used to optimize the significant pixels and cluster the similar pixels optimally. Therefore it provides the higher correlation value for given DICOM brain images rather than the existing algorithm.

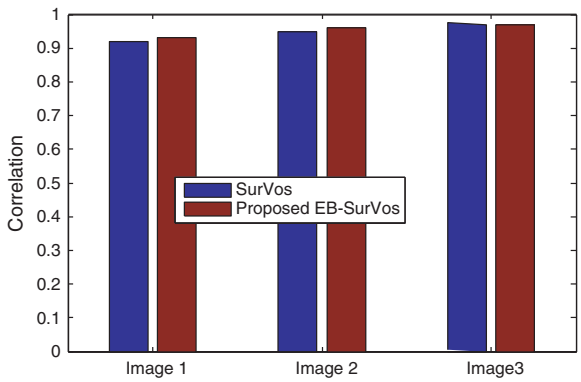


Fig. 5. Correlation comparison.

4.5. PSNR

If the Peak Signal to Noise Ratio (PSNR) is small then the quality of the image will be poor. Commonly, a high quality rebuilt image is of high PSNR and low MSE. PSNR is provided by:

$$\text{PSNR} = \frac{10\log 255^2}{\text{MSE}} \tag{15}$$

Figure 6 explains the PSNR comparison output among the proposed EB-SuRVoS, and the current SuRVoS approach. The proposed method has great value of Peak Signal to Noise Ratio (PSNR). The output shows that the proposed EB-SuRVoS gives high PSNR showing the good reconstructed image. The proposed EB-SuRVoS algorithm segments the pixels effectively and it is used to optimize the significant pixels using bat optimization. Therefore it provides the higher PSNR value for given DICOM brain images rather than the existing algorithm.

4.6. MSE

Mean Square Error (MSE) is an evaluation of the simple image quality. The greater value of MSE denotes that the quality of the image is very poor. MSE is provided by:

$$\text{MSE} = \frac{1}{\text{MSE}} \sum_{m=1}^{M}\sum_{n=1}^{N}(x(m,n)-\hat{x}(m,n))^2 \tag{16}$$

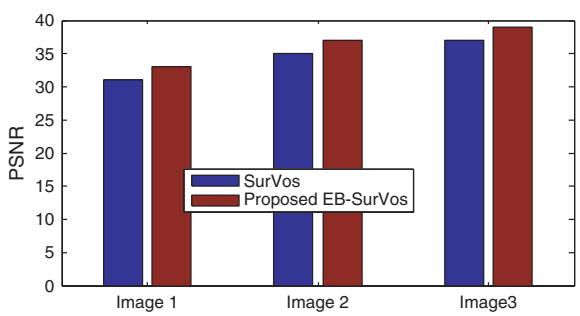


Fig. 6. PSNR comparison.

RESEARCH ARTICLE

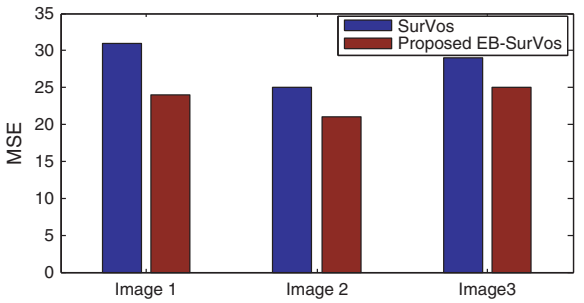


Fig. 7. MSE comparison.

Figure 7 explains that the MSE comparison output among the proposed EB-SuRVoS, and the current RWS, ACM and SuRVoS approaches. The proposed method has lower value of MSE. The output denotes that proposed EB-SuRVoS holds less MSE presents the good reconstructed image. The proposed EB-SuRVoS is used to optimize the important pixels and segment the most similar pixels for the given DICOM brain images.

5. CONCLUSION

The proposed EB-SuRVoS method models are a combination of the initialization techniques to segment structures in volume data. The user interface is to mutually segment huge 3D volumes. Super-Regions (supervoxels and megavoxels) minimize the manual segmentation and eliminate manual delineation of boundaries, conceivably reduce the subjectivity. Furthermore, EB-SuRVoS gives a broad variety of de-noising and textural filters which is tuned to every sample and utilized to guide the machine learning model to expand the user's manual segmentations to the remaining set of volume. Supervoxels and megavoxels were utilized in combination with the model training to segment the nucleus, cytoplasm and extracellular area. Furthermore to the above advantages by utilizing the features which is brought-in in EB-SuRVoS, user segmentation time is minimized by 5 times when compared with the manual segmentation. Along with this, these features provide a way for good and very quantitative utilization of the 3D volume datasets in biological sciences. The improved bat algorithm is utilized to give an enhanced model by producing the best fitness function values. The result confirms that the proposed system gives excellent performance output in terms of higher accuracy, precision, recall, correlation, PSNR and lower MSE values.

References

1. V. Lucic, A. Rigort, and W. Baumeister, *J. Cell Biol.* 202 **(2013)**.
2. K. Zhang, et al., *Image and Vision Computing* 28.4, 668 **(2010)**.
3. T. Chan and L. Vese, *IEEE Transaction on Image Processing* 10, 266 **(2001)**.
4. A. Bosch, A. Zisserman, and X. Munoz, Image classification using random forests and ferns, *Proc. ICCV* **(2007)**.
5. J. Son, et al., Tracking-by-segmentation with online gradient boosting decision tree, *Proceedings of the IEEE International Conference on Computer Vision* **(2015)**.
6. J. H. Friedman, *Annals of Statistics* 29, 1189 **(2000)**.
7. A. Tzotsos and D. Argialas, Support vector machine classification for object-based image analysis, Object-Based Image Analysis, Springer, Heidelberg, Berlin **(2008)**, pp. 663–677.
8. Z. Kato and T.-C. Pong, *Image and Vision Computing* 24.10, 1103 **(2006)**.
9. L. Ding and A. Yilmaz, *Pattern Recognition* 43.5, 1863 **(2010)**.
10. M. Holtzman-Gazit, D. Goldsher, and R. Kimmel, Hierarchical segmentation of thin structures in volumetric medical images, *International Conference on Medical Image Computing and Computer-Assisted Intervention* Springer, Heidelberg, Berlin **(2003)**.
11. A. Alihodzic and M. Tuba, *Recent Researches in Telecommunications, Informatics, Electronics and Signal Processing* 17 **(2013)**.

Received: 6 April 2017. Accepted: 29 April 2017.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1441–1444, 2018

On Upper and Lower Supra-Continuous Multifunction

M. Kamel El-Sayed

Department of Basic Science, Higher Institute of Engineering and Technology, Kafr El-Sheikh, 33514, Egypt

In this paper, a multifunction $F: (X, \tau) \to (Y, \sigma)$ is said to be upper (resp. lower) supra-continuous if $F^+(V)$ (resp. $F^-(V)$) is supra-open in (X, τ) for each $V \in \sigma$. These multifunctions are considered as a generalization of some types of near-continuous multifunction, which are known before. Several characterizations and fundamental properties of these new classes of multifunctions are obtained. The relationships between supra-continuous multifunctions and supra-closed graphs are also discussed.

Keywords: Upper Supra-Continuous, Lower Supra-Continuous, Supra-Regular Set, Supra-Compact Space, Punctually Supra-Regular Multifunction, Supra-Closed Graph.

RESEARCH ARTICLE

1. INTRODUCTION

In 1977, Husain[3] was initiated the concept of supra-open set which is considered as a wider class of some known types of near-open sets. In 1983, Mashhour et al.[8] defined the concept of S-continuity but for a single-valued function $f: (X, \tau) \to (Y, \sigma)$. Many topological properties of the above mentioned concepts and others have been established in Refs. [6–8]. The purpose of this paper is to present the upper (lower) supra-continuous multifunction as a generalization of each of upper (lower) semi-continuous multifunction in the sense of Berge[2] the upper (lower) quasi-continuous and the upper (lower) precontinuous multifunction due to Popa[11, 12] and also, upper (lower) α-continuous and upper (lower) β-continuous multifunctions given by Popa and Noiri.[13, 14] Moreover, these new multifunctions are characterized with many of their properties have also been established.

2. PRELIMINARIES

The topological space or simply space which are used here will be given by (X, τ) and (Y, σ), while $\tau - cl(W)$ and $\tau - \text{int}(W)$ are denoted to the closure and the interior of any subset W of X with respect to a topology τ. In (X, τ), the class $\tau^* \subseteq P(X)$, is called supra-topology on X if $X \in \tau^*$ is closed under arbitrary union,[3] (X, τ^*) is a supra-topological space or simply supra-space, each member of τ is supra-open and its complement is supra-closed,[8] in (X, τ^*), the supra-closer, the supra-interior and supra-frontier of any $A \subseteq X$ will be denoted by supra-$cl(A)$, supra-int(A) and supra-fr(A), which are defined in Ref. [8] likewise the corresponding ordinary ones. While, for any $x \in X$; $\tau^*(x) = \{W \subseteq X: W \in \tau^*, x \in W\}$. In (X, τ), $A \subseteq X$ is called semi-open[4] if there exists $U \in \tau$ such that $U \in \tau$ such that $U \subseteq A \subseteq \tau - cl(U)$, while A is preopen[1, 5] if $A \subseteq \tau - \text{int}(\tau - cl(A))$. The family of all semi-open and pre-open sets in (X, τ) are denoted by $SO(x, \tau)$ and $PO(x, \tau)$, respectively. Moreover, $\tau^\alpha = SO(x, \tau) \cap PO(x, \tau)$ and $\beta O(x, \tau) \Rightarrow SO(x, \tau) \cup PO(x, \tau)$. While $A \in \tau^\alpha$ and $A \in \beta O(x, \tau)$ are called α-set[9] and β-open set,[1] respectively.

A single-valued function $f: (X, \tau) \to (Y, \sigma)$ is called S-continuous,[8] if the inverse image of each open set (Y, σ) is τ^*-super open in (X, τ). For a multifunction $f: (X, \tau) \to (Y, \sigma)$, the upper and the lower inverses of any $B \subseteq Y$ given by $F^+(B) = \{x \in X: F(x) \subseteq B\}$ and $F^-(B) = \{x \in X: F(x) \cap B \neq \phi\}$, respectively. Moreover, $f: (X, \tau) \to (Y, \sigma)$ is called upper (resp. lower) semi-continuous,[2] if for each $V \in \sigma$, $F^+(V) \in \tau$, (resp. $F^-(V) \in \tau$). If τ in semi-continuity is replaced by $SO(X, \tau)$, τ^α, $PO(X, \tau)$ and $\beta O(X, \tau)$, then F is upper (lower) quasi-continuous,[11] upper (lower) α continuous,[12] upper (lower) precontinuous[13] and upper (lower) β-continuous,[14] respectively. A space (X, τ) is called supra-compact,[6] if every supraopen cover of X admits a finite subcover.

3. SUPRA-CONTINUOUS MULTIFUNCTIONS

DEFINITION 3.1. A multifunction $f: (X, \tau) \to (Y, \sigma)$ is said to be:

(a) upper supra-continuous at a point $x \in X$ if for each open set V containing $F(x)$, these exists $W \in \tau^*(x)$ such that $F(W) \subseteq V$,

(b) lower supra-continuous at a point $x \in X$ if for each open set V containing $F(x)$, there exists $W \in \tau^*(x)$ such that $F(W) \cap V \neq \phi$,

(c) upper (lower) supra-continuous if F has this property at every point of X.

Any single-valued function $f:(X,\tau)\to(Y,\sigma)$ can be considered as a multi-valued one which assigns to any $x\in X$, the singleton $\{f(x)\}$. Applying the above definition of both of upper and lower supra-continuous multifunction to a single-valued. It's clearly that they coincide with the notion of S-continuous due to Mashhour et al.[8] One characterization of the above multifunction is established throughout the following result, which its proof is straightforward, so it is omitted.

Remark 3.1. For a multifunction $f:(X,\tau)\to(Y,\sigma)$, many properties of upper (lower) semi-continuity[2] (resp. upper (lower) α-continuity,[13] upper (lower) quasi-continuity,[11] upper (lower) precontinuity[12] and upper (lower) (β-continuity[14]) can be deduced from the upper (lower) supra-continuity by considering $\tau^*=\tau$ (resp. $\tau^*=\tau^\alpha$, $\tau^*=SO(X,\tau)$, $\tau^*=PO(X,\tau)$ and $\tau^*=\beta O(X,\tau)$).

Proposition 3.1. *A multifunction $f:(X,\tau)\to(Y,\sigma)$ is upper (resp. lower) supra-continuous at a point $x\in X$ if and only if for $V\in\sigma$ with $F(x)\subseteq V$ (resp. $F(x)\cap V\neq\phi$). Then $x\in$ supra-int$(F^+(V))$ (resp. $x\in$ supra-int$(F^-(V))$.*

Lemma 3.1. *For any $A\in(X,\tau)$, $\tau-int(A)\subseteq$ supra-int$(A)\subseteq A\subseteq$ supra-cl$(A)\subseteq\tau-cl(A)$.*

Theorem 3.1. *The followings are equivalent for a multifunction $f:(X,\tau)\to(Y,\sigma)$:*
(i) *F is upper supra-continuous,*
(ii) *for each $x\in X$ and each $V\in\alpha(F(x))$, $F^+(V)\in\tau^*(x)$,*
(iii) *for each $x\in X$ and each $V\in\alpha(F(x))$, there exist $W\in\tau^*$ such that $F(W)\subseteq V$,*
(iv) *$F^+(V)\in\tau^*$, for every $V\in\sigma$,*
(v) *$F^-(K)$ is supra-closed for every closet set $K\subseteq Y$,*
(vi) *supra-cl$(F^+(B))\subseteq F^+(\tau-cl(B))$, for every $B\subseteq Y$,*
(vii) *$F^+(\tau-int(B))\subseteq$ supra-int$(F^+(B))$, for every $B\subseteq Y$,*
(viii) *supra-fr$(F^-(B))\subseteq F(fr(B))$, for every $B\subseteq Y$,*
(ix) *$F:(X,\tau^*)\to(Y,\sigma)$ is upper semi-continuous.*

Proof. (i) $\Leftrightarrow$ (ii) and (i) $\Rightarrow$ (iv): Follow from Proposition 3.1.
(ii) $\Leftrightarrow$ (iii): This is obvious, since the arbitrary union of supra-open set is supra-open.
(iv) $\Rightarrow$ (v): Let K be closed in Y, the result satisfies since $F^+(Y\backslash K)=X\backslash F^-(K)$.
(v) $\Rightarrow$ (vi): By putting $K=\sigma-cl(B)$ and applying Lemma 3.1.
(vi) $\Rightarrow$ (vii): Let $B\subseteq Y$, then $\sigma-\text{int}(B)\in\sigma$ and so $Y\backslash\sigma-\text{int}(B)$ is closed in (Y,σ).

Therefore by (vi) we get $X\backslash$supra-int$(F^+(B))=$ supra-cl$(X\backslash F^+(B))\subseteq$ supra-cl$(X\backslash F^+(\sigma-\text{int}(B))=$ supra-cl$(F^-(Y\backslash\sigma-\text{int}(B))\subseteq F^-(Y\backslash\sigma-\text{int}(B))\subseteq X\backslash F^+(\sigma-\text{int}(B))$. This implies that $F^+(\sigma-\text{int}(B))\subseteq$ supra-int$(F^+(B))$.

(vii) $\Rightarrow$ (ii): Let $x\in X$ be arbitrary and each $V\in\sigma(f(x))$ then $F^+(V)\subseteq$ supra-int$(F^+(V))$.
(viii) $\Leftrightarrow$ (v): Clearly, for suprafrontier and frontier of any set is supra-closed and closed, respectively.
(ix) $\Leftrightarrow$ (iv): Follows immediately.

Theorem 3.2. *For a multifunction $f:(X,\tau)\to(Y,\sigma)$, the following statements are equivalents*:
(i) *F is lower supra-continuous,*
(ii) *for each $x\in X$ and each $V\in\sigma$ such that $F(x)\cap V\neq\phi$, $F^-(V)\in\tau^*(x)$,*
(iii) *for each $x\in X$ and each $V\in\sigma$ with $F(x)\cap V\neq\phi$, there exists $W\in\tau^*$ such that $F(W)\cap V\neq\phi$,*
(iv) *$F^-(V)\in\tau^*$ for every $V\in\sigma$,*
(v) *$F^+(K)$ is supra-closed for every closed set $K\subseteq Y$,*
(vi) *supra-cl$(F^+(B))\subseteq F^+(\sigma cl-(B))$ for any $B\subseteq Y$,*
(vii) *$F^-(\sigma-int(B))\subseteq F^+(fr(B))$, for any $B\subseteq Y$,*
(viii) *supra-fr$(F^+(B))\subseteq F^+(fr(B))$, for every $B\subseteq Y$,*
(ix) *$F(X,\tau^*)\to(Y,\sigma)$ is lower semi-continuous.*

Proof. The proof is a quite similar to that of Theorem 3.1.

Recalling that, the net $(x_i)_{i\in I}$ supra-convergent to x_0, if for each $W\in\tau^*$ there exists a $i_0\in I$ such that for each $i\geq i_0$ implies $x_i\in W$.

Theorem 3.3. *A multifunction $f:(X,\tau)\to(Y,\sigma)$ is upper supra-continuous if and only if for each net $(x_i)_{i\in I}$ supra-convergent to x_0 and for each $V\in\sigma$ with $F(x_0)\subseteq V$ there is $i_0\in I$ such that $F(X_i)\subseteq V$ for all $i\geq i_0$.*

Proof. Necessity, let $V\in\sigma$ with $F(x_0)\subseteq V$. By upper supra-continuous of F, there is $W\in\tau^*(x_0)$ such that $F(W)\subseteq V$. Since from hypothesis, a net $(x_i)_{i\in I}$ is supra-convergent to x_0 and $W\in\tau^*(x_0)$ there one $i_0\in I$ such that $x_i\in W$ for all $i\geq i_0$ and then $F(x_i)\subseteq V$ for all $i\geq i_0$.

Sufficiency, assume the converse, i.e., there is an open set V in Y with $F(x_0)\subseteq V$ such that for each $W\in\tau^*$ under inclusion relation $F(W)\not\subseteq V$, i.e., there is $x_w\in W$ such that $F(x_w)\not\subseteq V$. Then all of x_w will form a net in X with directed set W of $\tau^*(x_0)$, clearly this net is supra-convergent to x_0. But $F(x_w)\not\subseteq V$ for all $W\in\tau^*(x_0)$. This leads to a contradiction which completes the proof.

Theorem 3.4. *A multifunction $f:(X,\tau)\to(Y,\sigma)$ is lower supra-continuous if and only if for each $y_0\in F(x_0)$ and for every net $(x_i)_{i\in I}$ supra-convergent to x_0, there exists a subset $(z_j)_{j\in J}$ of the net $(x_i)_{i\in I}$ and a net $(y_i)_{(j,v)\in J}$ in Y so that $(y_i)_{(j,v)\in J}$ supra-convergent to y and $y_i\in F(z_j)$.*

Proof. For the necessity, suppose F is lower supra-continuous, $(x_i)_{i\in I}$ is a net supra-convergent to x_0, $y\in F(x_0)$ and $V\in\sigma(y)$. So we have $F(x_0)\cap V\neq\phi$, by lower supra-continuity of F at x_0, there is a supra-open set $W\subseteq X$ containing x_0 such that $W\subseteq(V)$. Supra-convergent of a net $(x_i)_{i\in I}$ to x_0 and for this W, there is a $i_0\in I$ such that for each $i\geq i_0$, then $x_i\in W$ and therefore $x_i\in F^-(V)$. Hence, for each $V\in\sigma(y)$ define the sets

$I_v = \{i_0 \in I : i \geq i_0 \Rightarrow x_i \in F(V)\}$ and $J = \{(i, V) : V \in D(y), i \in I_v\}$ and order $\geq$ on J given as $(i', V') \geq (i, V)$ if and only if $i' \geq i$ and $V' \subseteq V$. Also, define $\zeta : J \to I$ by $\zeta(j, V) = j$. Then ζ, increasing and cofinal in I, so ζ defines a subset of $(x_i)_{i\in I}$ denoted by $(z_j)_{(j,v)\in J}$. On the other hand for any $(j, V) \in J$, since $j \geq j_0$ implies $x_j \in F(V)$ we have $F(Z_j) \cap V = F(X_j) \cap V \neq \phi$. Then the net $(y_i)_{(j,v)\in J}$ is supra-convergent to y. To see this, let $V_0 \in \sigma(y)$, then there is $j_0 \in I$ with $j_0 = \zeta(j_0, V_0)$; $(j_0, V_0) \in J$ and $y_{j_0} \in V$. If $(j, V) \geq (j_0, V_0)$ this means that $j \geq j_0$ and $V \subseteq V_0$. Therefore, $y_j \in F(x_j) \cap F(x_j) \cap V \subseteq F(x_j) \cap V_0$ so $y_i \in V_0$. Thus $(y_i)_{(j,v)\in J}$ is supra-convergent to y, which shows the result.

To show the sufficiency, assume the converse, i.e., F is not lower supra-continuous at x_0. Then there exists $V \in \sigma$ that $f(x_0) \cap V \neq \phi$ and for any supra-neighborhood $W \subseteq X$ of x_0, there is $x_w \in W$ for which $F(x_0) \cap V = \phi$. Let us consider the net $(x_w)_{W\in\tau^*(x_0)}$ which obviously supra-convergent to x_0. Suppose $y_0 \in F(x_0) \cap V$, by hypothesis, there is a subnet $(x_k)_{k\in K}$ of $(x_w)_{W\in\tau^*(x_0)}$ and $y_k \in F(z_k)$ like $(y_k)_{k\in K}$ supra-convergent to y_0. As $y_0 \in V \in \sigma$ there is $k'_0 \in K$ so that $k > k'_0$ implies $y_k \in V$. On the other hand $(z_k)_{k\in K}$ is a subnet of the net $(x_w)_{W\in\tau^*(x_0)}$ and so there is a function $\Omega : K \to \tau^*(x_0)$ such that $z_k = x_{\Omega(k)}$ and for each $W \in \tau^*(x_0)$ there exists $k^*_0 \in K$ such that $\Omega(k^*_0) \geq W$. If $k \geq k^*_0$ then $\Omega(k) \geq \Omega(k^*_0) \geq W$. Considering $k_0 \in K$ so that $k_0 \geq k'_0$ and $k_0 \geq k^*_0$. Therefore $y_k \in V$ and the meaning of the net $(x_w)_{W\in\tau^*(x_0)}$, we have $F(z_k) \cap V = F(x_{\Omega(k)}) \cap V = \phi$. This gives $y_k \notin V$ which contradict the hypothesis and so the requirement holds.

Definition 3.2. A subset W of a space (X, τ) is called supra-regular, if for any $x \in W$and any $H \in \tau^*(x)$ there exists $U \in \tau$ such that $x \in U \subseteq \tau - cl(U) \subseteq H$. Therefore, recalling that $F : (X, \tau) \to (Y, \sigma)$ is punctually supra-regular, if for each $x \in X$, $F(x)$ is supra-regular.

Lemma 3.2. *In space (X, τ), if $W \subseteq X$ is supra-regular ad contained in a supra-open set H. Then there exists $U \in \tau$ such that $W \subseteq U \subseteq \tau - cl(U) \subseteq H$.*

For a multifunction $F : (X, \tau) \to (Y, \sigma)$, a multifunction supra-cl$(F) : (X, \tau) \to (Y, \sigma)$ is defined as: (supra-cl F) $(x) =$ supra-cl$(F(x))$ for each $x \in X$.

Proposition 3.2. *For a punctually α-paracompact and punctually supra-regular multifunction $F : (X, \tau) \to (Y, \sigma)$. Then (supra-cl$(F)^+(W)) = F^+(W)$, for each $W \in \sigma^*$.*

Proof. Let $x \in$ supra-cl$(F))^+(W)$ for any $W \in \sigma^*$, this means $F(x) \subseteq$ supra-cl$(F(x)) \subseteq W$ which leads to $x \in F^+(W)$. Hence one inclusion holds. To show the other, let $x \in F^+(W)$ where $W \in \sigma^*(x)$, by hypothesis of F and the fact that $\sigma \subseteq \sigma^*$, applying Lemma 3.2, there exists $G \in \sigma$ such that $F(x) \subseteq G \in \sigma - cl(G) \subseteq W$. Therefore, supra-cl$(F(x)) \subseteq W$ this means that $x \in$ (supra-cl $F)^+$. Hence the equality 6 verify.

Theorem 3.5. *Let, $F : (X, \tau) \to (Y, \sigma)$ be a punctually a paracompact and punctually supra-regular multifunction. Then F is upper supra-continuous if and only if (supra-cl F) : $(X, \tau) \to (Y, \sigma)$ is upper supra-continuous.*

Proof. Necessity, suppose $V \in \sigma$ and $x \in$ (supra-cl $F)^+(V) = F^+(V)$ (see Proposition 3.2). By upper supra-continuity of F, there exists $H \in \tau^*(x)$ such that $F(H) \subseteq V$. Since $\sigma \subseteq \sigma^*$, then by Lemma 3.2 and assumption of F, there exists $G \in \sigma$ such that $F(h) \subseteq G \subseteq \sigma - cl(G) \subseteq W$, for each $h \in H$. Hence, supra-cl$(F(h)) =$ (supra-cl $F)(h) \subseteq$ supra-cl$(G) \subseteq \sigma - cl(G) \subseteq V$, for each $h \in H$ which gives (supra-cl $F)(H) \subseteq V$. Thus (supra-cl F) is upper supra-continuous. Sufficiency, assume $V \in \sigma$ and $X \in F^+(V) =$ (supra-cl $F)^+(V)$. By hypothesis of F.

In this case, there is $H \in \tau^*(x)$ such that (supra-cl F) $(H) \subseteq V$ which obviously gives $F(H) \subseteq V$. This comp-letes the proof.

Lemma 3.3. *In a space (X, τ), any $x \in X$ and $A \subseteq X$, $X \in$ supra-cl(A) if and only if $A \cap W \neq \phi$, for each $W \in \tau^*(x)$.*

Proposition 3.3. *For a multifunction $F : (X, \tau) \to (Y, \sigma)$, (supra-cl $F)^-(W) = F^-(W)$, for each $W \in \sigma^*$.*

Proof. Let $x \in$ (supra-cl $F)^-(W)$, then $W \cap$ supra-cl$(F(x)) \neq \phi$. Since $W \in \sigma^*$ so, Lemma 3.3 gives $W \cap F(x) \neq \phi$ and hence $x \in F^-(W)$. Conversely, let $x \in F^-(W)$, then $\phi \neq F(x) \cap W \subseteq$ supra-cl $F)^-(x) \cap W$ and so $x \in$ (supra-cl $F)^-(W)$. Hence $x \in$ (supra-cl $F)^-(W)$ and this completes the equality.

Theorem 3.6. *A multifunction $F : (X, \tau) \to (Y, \sigma)$ is lower supra-continuous if and only if (supra-cl F) : $(X, \tau) \to (Y, \sigma)$ is lower supra-continuous.*

Proof. This is an immediate consequence of Proposition 3.2 taking in considering that $\tau \subseteq \tau^*$ and (iv) of Theorem 3.2.

Theorem 3.7. *If $F : (X, \tau) \to (Y, \sigma)$ is an upper supra-continuous surjection and for each $x \in X$, $F(x)$ is compact relative to Y. If (X, τ) is supra-compact, then (Y, σ) is compact.*

Proof. Let $\{V_i : i \in I, V_i \in \sigma\}$ be a cover of Y, and since $F(x)$ is compact relative to Y, for each $x \in X$. Then there exists a finite $I_0(x)$ of I such that $F(x) \subseteq U \{V_i : i \in I_0(x)\}$. Upper supra-continuity of F gives that there exists $W(x) \in \tau^*(X, x)$ such that $F(W(x)) \subseteq \cup \{V_i : i \in I_0(x)\}$. Since (X, τ) is supra-compact, then there exists $\{x_1, x_2, \ldots, x_n\}$ such that $X = U(W(x_i)) : I \leq j \leq n\}$. Therefore, $Y = F(X) = U(F(W(x_i))) : I \leq j \leq n\} \subseteq \boldsymbol{U} \{V_i : i \in I_0(x_i), 1 \leq j \leq n\}$. Hence (Y, σ) is compact.

4. SUPRA-CONTINUOUS MULTIFUNCTIONS AND SUPRA-CLOSED GRAPHS

Definition 4.1. A multifunction $F : (X, \tau) \to (Y, \sigma)$ is said to be have a supra-closed graph if for each pair

RESEARCH ARTICLE

$(x, y) \notin G(F)$ there exists $W \in \tau^*(x)$ and $H \notin \sigma^*(y)$ such that $(W \times H) \cap G(F) = \phi$.

A multifunction $F: (X, \tau) \to (Y, \sigma)$ is point-closed (supra-closed), if for each $x \in X$, $F(x)$ is closed (supra-closed) in Y.

Proposition 4.1. *A multifunction $F: (X, \tau) \to (Y, \sigma)$ has a supra-closed graph if and only if for each $x \in X$ and $y \in Y$ such that $y \notin F(x)$, there exists two supra-open sets H, W containing x and y, respectively, such that $F(H) \cap W = \phi$.*

Proof. Necessity, let $x \in X$ and $y \in Y$ with $y \notin F(x)$. Then by supra-closed graph of F, there are $H \in \tau^*(x)$ and $W \in \sigma^*$ containing $F(x)$ such that $(H \times W) \cap G(F) = \phi$. This implies that for every $x \in H$ and $y \in W$ where $y \notin F(x)$ and so $F(H) \cap W = \phi$.

Sufficiency, let $(x, y) \notin G(F)$, this means $y \notin F(x)$ then there are two disjoint supra-open sets H, W containing x and y, respectively, such that $F(H) \cap W = \phi$. This implies that $(H \times W) \cap G(F) = \phi$, which completes the proof.

Theorem 4.1. *If $F: (X, \tau) \to (Y, \sigma)$ is upper supra-continuous and point-closed multifunction. Then $G(F)$ is supra-closed if (Y, σ) is regular.*

Proof. Suppose $(x, y) \notin G(F)$, then $y \notin F(x)$. Since Y is regular, there exists disjoint $V_i \in \sigma, i = 1, 2$ such that $y \in V_1$ and $F(x) \subseteq V_2$. Since F is upper supra-continuous at x, there exists $W \in \tau^*(x)$ such that $F(W) \subseteq V_2$. As $V_1 \cap V_2 = \phi$, then $\bigcap_{i=1}^{2}$ supra-int$(V_i) = \phi$ and therefore, $x \in$ supra-int$(W) = W$, $y \in$ supra-int(V_1) and $(x, y) \in W \times$ (supra-int$(V_1) \subseteq (X \times Y) \backslash G(F)$. Thus $(X \times Y) \backslash G(F) \in \tau^*(X \times Y)$ which gives the result.

Definition 4.2. A subset W of a space (X, τ) is called α-paracompact[15] if for every open cover v of W in (X, τ) there exists a locally finite open cover ξ of W which refines v.

Theorem 4.2. *Let $F: (X, \tau) \to (Y, \sigma)$ be an upper supra-continuous multifunction from (X, τ) into a Hausdorff space (Y, σ). If $F(x)$ is α-paracompact for each $x \in X$, then $G(F)$ is supra-closed.*

Proof. Let that $(x_0, y_0) \notin G(F)$, then $y_0 \notin F(x_0)$. Since (Y, σ) is Hausdorff, then for each $y \in F(x_0)$ there exists $V_y \in \sigma(y)$ and $V_y^* \in \sigma(y_0)$ such that $V_y \cap V_y^* = \phi$. So, the family $\{V_y : y \in F(x_0)\}$ is an open cover of $F(x_0)$. Thus, by α-paracompactness of $F(x_0)$, there is a locally finite open cover $v = \{U_i : i \in I\}$ which refines $\{V_y : y \in F(x_0)\}$. Therefore, there exists $H_0 \in \sigma(y_0)$ such that H_0 intersects only finitely many members $U_{i_1}, U_{i_2}, \ldots, U_{i_n}$ of v. Choose $y_1, y_2, \ldots, y_n$ in $F(x_0)$ such that $U_{i_j} \subseteq V_{y_j}$ for each $1 < j < n$, and the set $H = H_0 \cap (\bigcup_{i \in I} V_{y_i})$. Then $H \in \sigma(y_0)$ such that $H \cap (\bigcup_{i \in I} V_i) = \phi$. The upper supra-continuityof F, means that there exists $W \in \tau^*(x_0)$ such that $x_0 \in W \subseteq F^+(\bigcup_{i \in I} V_i)$. It follows that $(W \times U) \cap G(F) = \phi$ and hence $G(F)$is supra-closed.

Lemma 4.1 (Noiri and Popa, 2000). *The following hold for $F: (X, \tau) \to (Y, \sigma)$, $A \subseteq X$ and $B \subseteq Y$;*
(i) $G_F^+(A \times B) = A \cap F^+(B)$;
(ii) $G_F^-(A \times B) = A \cap F^-(B)$.

Theorem 4.3. *For a multifunction $F: (X, \tau) \to (Y, \sigma)$, if G_F is upper supra-continuous, then F is upper supra-continuous.*

Proof. Let $x \in X$ and $V \in \sigma(F(x))$. Since $X \times V \in \tau \times \sigma$ and $G_F(x) \subseteq X \times V$, by Theorem 3.1, there exists $W \in \tau^*(x)$ such that $G_F \subseteq X \times V$. Therefore, by Lemma 4.1, we get $W \subseteq G_F^-(X \times V) = X \cap G_F^+(V) = F^+(V)$ and so $F(W) \subseteq V$. Hence Theorem 3.1 gives also that F upper supra-continuous.

Theorem 4.4. *If the graph G_F of a multifunction $F: (X, \tau) \to (Y, \sigma)$ is lower supra-continuous, then F is also so.*

Proof. Let $x \in X$ and $V \in \sigma(F(x))$ with $F(X) \cap V \neq \phi$, also since $X \times V \in \tau \times \sigma$ then $G_F(x) \cap (X \times V) = \{\{x\} \times F(x)\} \cap (X \times V) = \{x\} \times (F(x) \cap V) \neq \phi$. Theorem 3.2 shows that there exists $W \in \tau^*(x)$ such that $G_F(w) \subseteq (X \times V) \neq \phi$, for each $w \in W$. Hence Lemma 4.1 obtains that $W \subseteq G^-(X \times V) = X \cap G^-(V) = F^-(V)$. Therefore, $F(w) \cap V \neq \phi$, for each $w \in W$ and Theorem 3.2 completes the proof.

References

1. M. E. Abd El-Monsef, S. N. El-Deeb, and R. A. Mahmoud, *Bull. Fac. Sci. Assiut Univ.* 12, 77 **(1983)**.
2. C. Berge, Espaces Topologiques, Fonctions Multivoques, Dunod, Paris **(1966)**.
3. T. Husain, Topology and Maps, Plenum Press, New York **(1977)**.
4. N. Levine, *Amer. Math. Monthly* 70, 36 **(1963)**.
5. A. S. Mashhour, M. E. Abd El-Monsef, and S. N. El-Deeb, On precontinuous and weakprecontinuous mappings, *Proc. Math. Phys. Soc. Egypt* **(1983)**, Vol. 53, pp. 47–53.
6. A. S. Mashhour, A. A. Allam, and I. A. Hasanein, Some Supratopological Results, J. Hohannes Kepler Univ. Linz, Austria **(1983)**, p. 25.
7. A. S. Mashhour, A. A. Allam, I. A. Hasanein, and S. N. El-Deeb, Operations on supratopological spaces, *Proc. Math. Phys. Soc. Egypt* **(1987)**.
8. A. S. Mashhour, A. A. Allam, F. S. Mahmoud, and F. H. Khedr, *Indian J. Pure and Appl. Math.* 14, 502 **(1983)**.
9. O. Njastad, *Pacific J. Math.* 15, 961 **(1965)**.
10. T. Noiri and V. Popa, *Demonstratio Math.* 26, 363 **(1993)**.
11. V. Popa, *St. Cer. Mat.* 27, 323 **(1975)**.
12. V. Popa, *Problemy Mat.* 10, 9 **(1988)**.
13. V. Popa and T. Noiri, *Math. Slovaca* 43, 477 **(1993)**.
14. V. Popa and T. Noiri, *Real Analysis Exchange* 22, 362 **(1996/97)**.
15. J. D. Wine, *Glasnik Mat.* 10, 351 **(1975)**.

Received: 15 November 2017. Accepted: 16 November 2017.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1445–1451, 2018

RESEARCH ARTICLE

Activities of Daily Living Rehab Game Play System with Augmented Reality Based Gamification Therapy for Automation of Post Stroke Upper Limb Rehabilitation

S. J. Syed Ali Fathima[1,*], S. Shankar[2], and A. Ahamed Thajudeen[3]

[1] *Kumaraguru College of Technology, Coimbatore 641049, Tamil Nadu, India*
[2] *Hindusthan College of Engineering and Technology, Coimbatore 641032, Tamil Nadu, India*
[3] *The Oxford College of Physiotheraphy, Bangalore 560068, Karnataka, India*

The effectiveness of Post Stroke upper limb rehabilitation can be maximized through the patient involvement in the exercise training program regularly with proper motivation and engagement to speed up recovery. The stroke patient suffers from functional disability of independently executing their activities of daily living (ADL) like reaching, handling, grooming, eating, and dressing etc. In-spite of various external supports, an individual gets recovered only at his/her own pace of confidence. ADL Rehab Game Play System proposes an augmented reality based gamification therapy using Microsoft©Kinect and Unity 3D Gaming Engine Software to create three dimensional (3D) training simulation scenarios to carry out most basic activities of daily living (ADL). Gamification using natural user interface (NUI) enhances the data-driven techniques used by the designers to motivate the patient or the players, so as to engage in the action actively and adds significance to the faster recovery.

Keywords: Microsoft©Kinect, Unity 3D Gaming Engine, Rehabilitation, Gamification, Activities of Daily Living (ADL), Natural User Interface (NUI), Range of Motion (ROM), Human Computer Interaction (HCI), Augmented Reality (AR), Three Dimensional (3D).

1. INTRODUCTION

Upper Limb refer to hands, are the most important part in most of the human physical actions. According to the reports from the various agencies, the occupationally impaired upper limbs affect one-third functions of the human body. As per the research of American Medical Association, 60% of the mechanisms of the body are lost when one loses his/her arm. Drop of 90% of the arm function equates a hand loss or the entire body mechanism by 54%. Therefore, one cannot ignore having healthy upper limbs.[1] Several studies have been conducted to examine stroke patients on the recovery of the hemiplegic arm. Patients up to 85% show an initial deficit in the arm. Three quarters of strokes happen in the region supplied by the middle cerebral artery. As a consequence, the upper limb will be affected in a large number of patients.[2]

In Physical rehabilitation, Activities of Daily Living (ADL) is an axiom concerning human self care routine activities or tasks that is undertaken in their everyday life with no assistance. The six basic ADL's are bathing, eating, toileting, transferring, dressing and continence. Research shows that there is a high correlation between ADL of the patients with stroke and their Quality of Life (QOL).[3] The Range of Motion (ROM) activities of upper limb joints are restricted by stroke condition and this limitation can have a severe impact on performing ADL by the patient.[4] The rehabilitation of the upper limb is a challenge. Although, many therapeutic approaches are currently available, considerable controversy exists about their effectiveness.

There should be appropriate rehabilitation equipment for the patients with dysfunction of upper limb using which it may be trained again and again to recover. Exercise slate of the arm, incline board, care motion exercises, vertical tower, stacking cones and exercise skate of the hand are the most frequently used equipment that is used clinically in the therapy. There are also different therapeutic methods like augmented reality (AR) that improves and integrates the texts, video, sound and other graphics, mechanical arms used to make a positive training of the patient through a mechanical structure and video games are helpful to give the instructions on the screen so that

*Author to whom correspondence should be addressed.

the mechanical arms can move to help the neural rehabilitation. Using these methods, the patients always feel that they are experiencing the training sessions as a real part of life.[5] The shoulder joint is used in daily routine activities, however, the bruising or sprain causes shoulder abduction or the rotation disorders where the patients have the problem in moving the shoulder because of the fear of the joint pain. Elbow, forearms and shoulders are the main parts that require a serious training and exercises. The adoption and complex integration of muscle activity of arm from shoulder to fingers is required for functional recovery of the upperlimb to perform activities like grasping, holding, and manipulating objects.[4]

There is a relationship between the computer and a human being where a suitable current technology is to provide human oriented and a natural mean of interaction and is called Human-computer interaction (HCI) where human interface using natural gesture.[7] Gestures play a pivotal role for the upcoming new technologies for conveying the message or any other information. When the user is based on a real-time action which is responsive, they are augmented reality systems. These systems offer a modern strategy for a patient to perform a specific type of therapy activities through games or simulations which are the integrated tasks.[8]

Augmented Reality (AR) using NUI enables virtual objects to be apparently merged into the captured and rendered real world environment in which user interacts with the objects by the use of motion-tracking devices like Microsoft©Kinect. This facilitates the developer to create more interesting and realistic engaging training exercises as games that makes the patients to feel, touch and interact with tangible objects in both real and virtual environment with minimum assistance.[9] Gamification is nothing but application of gaming concept in any non-game context in order to motivate and engage people to accomplish their goal. Gamification is of more benefit with NUI when applied to the field of rehabilitation. The proposed AR based gamification therapy approach motivates the patients and the main aim is a creation of rehab system which sustains the exercise of the patient's performance through an interesting game play and to increase the speed of rehabilitation. Post Stroke upper limb functional recovery using gamification approach has the facility to perform rehabilitation at their own home environment which could provide better psychological feel to the patients.[10]

The further sections of this paper explains about the technologies used to develop proposed ADL Rehab Game play system with the significant features of Microsoft©Kinect for Xbox One and Unity 3D gaming engine software, methodology of the system and environment setup, implementation of ADL Rehab Game Play Algorithm, Game scenarios for ADL training and range of Motion (ROM) of upper limb training supported by the system.

2. TECHNOLOGIES

ADL Rehab Game Play system is implemented using Unity 3D Gaming Engine software for game development and Microsoft©Kinect for XBox One sensor that provides a means to interact with the training games where the body of the players becomes a controller. The Microsoft©Kinect is a device that recognizes the body movements and voice recognition. It consists of 3D camera technology and is enhanced with the development of the computer vision[a]. The Microsoft©Kinect and its Software Development Kit (SDK) were released together to recognize the player's joints and can be used to monitor and record the training results of the patient. Microsoft©Kinect supports the level of integration between unity and the device that tracks the body. It has built a new technology game that focuses on the rehabilitation paradigm.[11]

Unity 3D Gaming Engine software is used to develop interactive 2D and 3D games. The games can run in multiple platforms such as web, Android, IOS, consoles and PC desktop. The software is excellent to use and performance is quite good from the simple to complex games where workflow supports 2D and 3D games development[b]. C# programming language is used for coding part that is easy to learn and so the mechanism animation system is very powerful to implement. Unity 3D supports augmented reality (AR) with a free licensing model and gives the useful services like Unity Analytics, Unity Ads and other new feature are regularly released[c].

The hardware configurations recommended for implementation are:

- 64-bit ($\times$64) processor PC with Microsoft©Windows V8 or later with the Microsoft©Kinect V2 drivers installed
- Minimum 4 GB RAM Memory
- USB 3.0 controller for Kinect V2 sensor
- Physical dual-core 3.1 GHz (2 logical cores per physical) or faster processor.
- Microsoft©Kinect sensor V2 for Xbox One
- Windows PC adaptor for Microsoft©Kinect sensor for Xbox One which includes a USB cabling and power hub
- DX11 capable graphics adapter.

Software Packages recommended for implementation are:

- Unity Editor 64-bit (Win) 5.4 and Higher
- Microsoft©Face Tracking Software Development Kit(SDK) for Kinect for Windows – Unity Package
- C# Programming Language.

After the download and the installation of the Unity 3D and Kinect v2 SDK and to complete the setup, Kinect

[a]Microsoft©Kinect and windows: https://developer.microsoft.com/en-us/ windows/kinect, last accessed 10 May 2017.

[b]Unity 3D gaming engine: http://zigfu.com/en/zdk/unity3d/, last accessed 10 May 2017.

[c]Kinect and Unity 3D integration: http://www.imaginativeuniversal.com/blog/post/2015/03/27/unity-5-and-kinect-2-integration.aspx, last accessed 12 May 2017.

v2 Unity plug-in is to be extracted to get the location. There is 3plug-ins in a folder with 2 sample scenes. They are—the first one is basic one, the second one is gesture builder plug-in and the third is face recognition, each one is wrapped with functionalities from Kinect v2 SDK. The development of the new game scenarios can be done in Unity editor using the existing models of SDK with relevant customization using C# coding implementation in Mono Develop environment of Unity 3D.

3. ADL REHAB GAME PLAY SYSTEM

ADL Rehab game play system mainly concentrates on the developing training and monitoring sessions to enhance the daily activities of disabled people with impaired upper limb. It is comprised of three different augmented reality (AR) based game scenarios specifically designed using which the user will be able to rehabilitate or train on the activities like reaching and touching an object, holding and releasing an object and eating behavior that are most essential for day to day activity. It also concentrates on various aspects like building an aerobic, strength and cognitive capabilities.

Developing such a game environment for rehabilitation is challenging as it involves a complex task of handling both the functionalities of an interactive videogame as well as an exercise therapy with the performance monitoring and evaluation setup at its background. Also, the patient should be motivated to engage in to the game environment without any aggravation and an audio visual feedback should be instantly given to the patients while playing the game, as this feedback information will be useful for the patient to perform the exercise correctly.[12] Moreover, the game designed should be transparent enough in such a way that the user should not feel the treatment session. It is good that the therapist is always there to guide, but on the other hand, session training could be carried on alone and the therapist could configure the parameters of different types of game scenario exercises. There is a facility to read the report session online or offline, therefore, the therapist is always aware and very well knows that how much the patient has performed the session and acts accordingly to modify whatever once deems to what is necessary for the patient's therapy.

Patient Rehabilitation Database (PRD) is the used to maintain the information regarding the patients' personal data, medical data, training plans and hand gesture analysis records using which the effectiveness and the performance of the patient can be evaluated by the rehabilitation experts. Every patient is associated with the login credentials and a personalized training plan based upon his/her fitness condition. The training plan has the list of game scenarios that the patient should undertake, order of game scenario, time period for each play within the scenario, count that specifies the number of repetitions of the play for each game scenario. The PRD also provides therapist

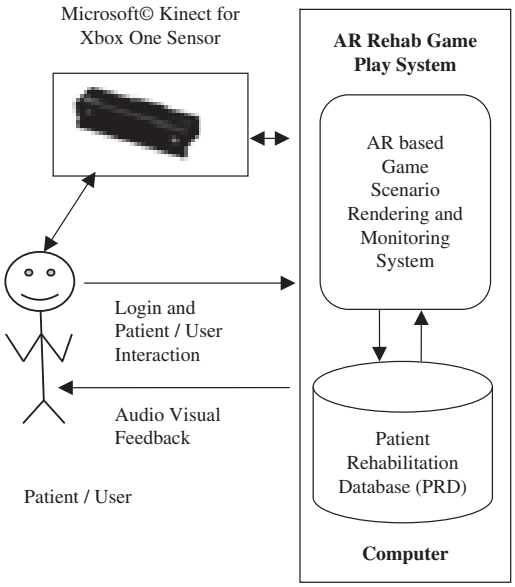


Fig. 1. ADL rehab game play environment.

with the information to access the recovery process that enables him/her to decide that whether other sequences of the training plan should be used or not, which the patient has to bring into the practice again. This ensures a proper guidance of the patient.

The game play environment setup consists of a computer installed with ADL Rehab Game Play System and a Microsoft©Kinect for Xbox one sensor device as shown in Figure 1. This setup can be done very easily either in clinical or home environment. The games are designed in such a way that it can be handled or played with very less assistance. The motion capturing feature of Microsoft©Kinect for Xbox One sensor is considerably accurate that the game can be played smoothly as per the expectation and measurement of the different Range of Motion (ROM) angles of joints with respect to shoulder, elbow and wrist of upper limb movements can be captured and recorded for further detailed analysis.

4. ADL REHAB GAME PLAY ALGORITHM

Initially to start the training session, the patient login to the ADL Rehab Game Play System using his/her login credential. As soon as the user login successfully, the new training session is started with the first game scenario associated for the patient and it is automatically rendered according to the order of training plan. The count and training period variables of the game scenarios are initialized from the training plan. While the patient plays the game scenario, for each play, the timer runs and the patient's hand gestures in terms of movement angles of shoulder, elbow, and wrist and game scores are captured and recorded into the hand gesture analysis PRD.

```
If (Login_check(patient))
{
  do
    {
      GameScenarios[] = get_GameScenarios(patient);
      foreach (GameScenario GS in GameScenarios[])
        {
//load and initialize values from Patient's Training plan DB
Load_and_Render_GS();
Initialize timer;
Initialize count;
          While (count! = 0)
{
  score = 0;
            ReInitialize timer;
            While (timer! = 0)
          {
            Run_timer();
          Game_Play();
          //for hand gesture analysis
          Capture_and_Record_ROM angles();
          While (game_hit)
              {
                Update_score();
              }
              //store game play score and ROM angles
              Update_gesture_analysis_DB();
          }
          }
          count–;
          }
          }While (New_Session);
}
```

Fig. 2. ADL Rehab game play algorithm.

The game score is calculated on the basis of hit or miss in each play and it varies with respect to each game scenario designed. Hit in the game refers to every correct action and miss refers to invalid or incorrect action. When the timer reaches 0, the count is decremented by 1 and next play for the same game scenario repeats. When the count equals 0, the next game scenario in the training plan gets loaded and rendered for the patient in the same manner. When all game scenarios are completed, the patient is given an option to start a new session of the training plan. ADL Rehab Game Play algorithm is illustrated in Figure 2.

5. ADL REHAB GAME PLAY SYSTEM AS AN INTELLIGENT AGENT

In the field of artificial intelligence, an agent is defined as anything that perceives its environment using sensors and acts upon the environment using effectors. An agent can also be defined as the combination of architecture and program. The system developed can be viewed as an utility based intelligent agent for rehabilitation application that uses augmented reality gaming environment in which the patients/users act upon the virtual objects in the real environment using a game play system through Microsoft Kinect sensing device. While designing any intelligent agent, the four factors are mainly considered are given as PAGE (Percepts, Actions, Goals, Environment) description.[13] Table I describes the PAGE description of ADL Rehab Game Play Intelligent Agent.

Table I. ADL Rehab game play intelligent agent – PAGE description.

Agent type	ADL Rehab game play intelligent agent
Percepts/sensors	ROM angles of hand with respect to shoulder, elbow and wrist captured using Microsoft Kinect Motion sensing device, game score
Actions/actuators	Patient/user play games by hand movements, reach, touch, hold and release objects, eating behavior
Goals/performance measures	Motivation, self/less support, improved training time, faster recovery, remote assistance
Environment	Home/health care centre—augmented reality gaming environment such as reach and touch object, hold and release object, eating environment

Table II. ADL rehab game play intelligent agent—environment properties.

Properties	Description
Accessible	Microsoft Kinect sensing device gives access to the complete state of the game play environment.
Deterministic	The next state of the game scenario is completely determined by the current state and the actions performed by the patient/user.
Episodic	Patient's experience is divided into "episodes" known as treatment sessions.
Static	Game play environment remains static as it is easy for the patients/users to deal but the performance score differs based on the action performed.
Discrete	Games can be played using fixed number of possible moves in each iteration.

The properties of ADL Rehab Game Play intelligent agent's environment[13] is given in Table II.

6. GAME SCENARIOS FOR ADL TRAINING

The main goal of the system is to maximize the speed of the patient's recovery and ability to live independently or with minimum aid. The three basic ADL training game scenarios of ADL game play system are (1) Reach and Touch objects, (2) Hold and Release objects and (3) Eating activity. The entertaining game scenarios are designed with audio visual feedback effects that allow the patient to practice the movements with more engagement and self motivation leading to a faster recovery to perform their self care ADL with less assistance.

6.1. Reach and Touch Objects

The scenario consists of the five jumper objects virtually scattered around the augmented reality (AR) environment.

Fig. 3. Reach and touch objects game scenario.

The AR environment in this case refers to blending of virtual objects into real interactive scene captured using Microsoft©Kinect's background camera. The game play is that the user should reach, touch and bump the object until it disappears, and this is to be repeated for all the 5 objects for a single play and timer is associated with each play of this scenario. For each correct touch and bump action the score is updated and the hand gestures are captured. When the timer reaches 0, the count is decremented and the game play repeats until count reaches 0. When Count is 0, the next game scenario is rendered according to the training plan of the patient. User Interface (UI) design of Reach and Touch Objects game scenario is shown in Figure 3.

A game problem is formulated using set of states and actions. The game formulation of Reach and Touch Object scenario is defined as follows.

States: a state description specifies the location of each of the five jumper object placed in the virtual space of user's AR environment

Operators: Move hands up, down, stretch, touch object.

Goal test: goal configuration for each repetition is to clear all five jumper objects in the space within the time limit using touch and bump action. A typical instance of Reach and Touch game scenario with initial and goal state is shown in Figure 4.

Fig. 5. Hold and release objects game scenario.

Path cost: for each object disappearance within the time limit, the score is set to be 10.

6.2. Hold and Release Objects

This scenario is completely a virtual interactive scene where five different objects like candy cane, Christmas ball, gift box, star and teddy bear are placed around a big box. The game play is that the user should move his hand towards the object, hold and drag the object and release to drop the object into the box. For each action of hold and release object correctly into the box, the score is updated and the hand gestures are captured. UI design of Hold and Release Objects game scenario is shown in Figure 5.

The game formulation of Hold and Release Object scenario is defined as follows.

States: A state description specifies the location of each of the five objects like candy cane, christmas ball, gift box, star and teddy bear and box object placed in the virtual space of user's AR environment.

Operators: Move hands up, down, stretch, bend, touch, grasp, hold, release object.

Goal test: Goal configuration for each repetition is to pick the object from space one by one and drop it into the box within the time limit using hold and release action. A typical instance of Hold and Release game scenario with initial and goal state is shown in Figure 6.

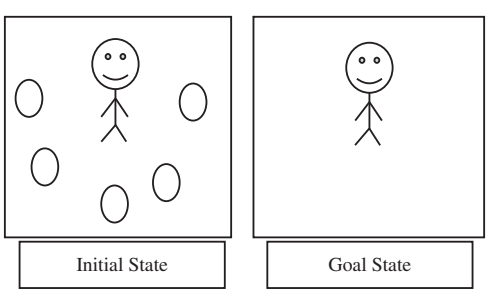


Fig. 4. Typical instance of reach and touch objects game scenario.

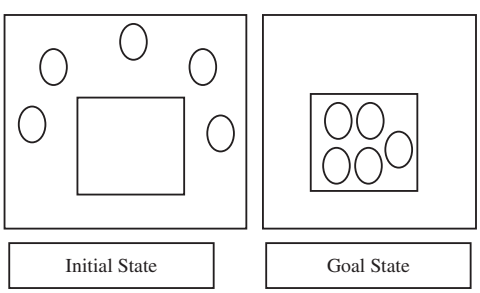


Fig. 6. Typical instance of hold and release objects game scenario.

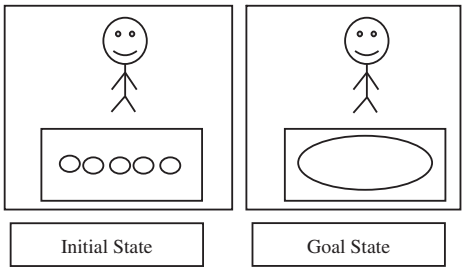


Fig. 7. Typical instance of eating activity game scenario.

Path cost: For each object correctly dropped into the box within the time limit, the score is set to be 10.

6.3. Eating Activity

In this scenario, AR environment is created in which a basket with set of fruits resembling bananas and oranges are placed on the table as the virtual objects and the user is made to interact with the virtual scene. The game play is that the user should move his arm towards the fruits, grasp one fruit at a time, then move the arms towards his mouth and release the fruit. Microsoft©Kinect's face tracking feature is used to make the fruit to disappear only if the user release the fruit near to the mouth. The score is updated for each correct eat action and the hand gesture angles are captured.

The game formulation of Hold and Release Object scenario is defined as follows.

States: a state description specifies the location of each of the five objects like banana and orange fruits placed in the plate object on the table object in the virtual space of user's AR environment.

Operators: Move hand towards plate, pick object, bend hand to reach mouth, release object.

Goal test: goal configuration for each repetition is to move the hands towards the plate, pick the fruit object from plate one by one, move the hands towards mouth and release the object to make it disappear within the time limit. A typical instance of Eating Activity game scenario with initial and goal state is shown in Figure 7.

Path cost: for each fruit object being correctly eaten within the time limit, the score is set to be 10.

7. RANGE OF MOTION (ROM) OF UPPER LIMB

The upper limb provides various degrees of freedom, a wide range of motion (ROM) at the joints and coordinated movement across multiple joints.[4] Upper limb rehabilitation is performed using set of ROM training. The description of ROM of upper Limb[d] supported by ADL Rehab Game play system is given in Table III.

[d]ROM of upper limb: http://www.exrx.net/Articulations/Shoulder.html, http://www.innerbody.com/, http://www.eatonhand.com/joi, last accessed 15 May 2017.

Table III. Description of ROM of upper limb.

ROM of upper limb	Description
Shoulder flexion	Moving the arm upward and downward to the front of the body
Shoulder abduction/adduction	Moving the arm upward and downward to the side i.e., far away/near to the midline of the body.
Shoulder medial rotation	Rotary movement around the longitudinal axis of the bone toward the midline of the body, turning the upper arm inward.
Elbow flexion	Moving forearm towards shoulder thereby decreasing the angle between arm and forearm
Elbow extension	Moving forearm away from shoulder thereby increasing the angle between arm and forearm
Wrist flexion	Decrease in angle between forearm and hand
Wrist extension	Increase in angle between forearm and hand
Hand grip	Co-ordinated movement of fingers for closing the fist.
Hand release	Co-ordinated movement of fingers for opening the fist.

Table IV. Mapping of game scenarios to ROM of upper limb.

Game scenarios	ROM of upper limb
Reach and touch objects	Shoulder flexion, shoulder abduction/adduction, wrist flexion, wrist extension
Hold and release objects	Shoulder flexion, shoulder abduction/adduction, hand grip, hand release, elbow flexion, elbow extension
Eating activity	Shoulder medial rotation, hand grip, hand release, elbow flexion, elbow extension

All the game scenarios in the system are designed to support the co-ordinated motion of shoulder, elbow and wrist of both left and right upper limbs. Table IV show the mapping of game scenarios to the ROM of upper limb that are mainly concentrated and have impact by undertaking the training of each scenario by the patient.

8. CONCLUSION

The project has aimed to improve the motivation of the patient while they perform different exercises training using gamification approach. All the feedback and technical data captured and recorded by the system help the physiotherapist in the monitoring and assessment process. The paper explains about the designs and development of the ADL Rehab Game Play system with AR based automated rehabilitation for the stroke survivors, those who are with impaired upper limbs by using Microsoft©Kinect and Unity 3D gaming engine. The infrastructure is made to design for the future development of the human study using new updated technologies that will help to maintain

and improve the mobility of the patients so that they can enjoy the therapy.

There are different therapies based on the computer games that the patients could be able to play to improve their score in terms of health. It is easy to start with the games, the patients hands are only tracked to get started and the calibration of the range of the motion takes place. These processes help the stroke survivors to play the games without facing any difficulties. Capturing 3D data by a game of Microsoft©Kinect sensor during the exercise session of the patients helps the therapist to access as much information about the patient's health by using the information system which is web or mobile based. The work is still going on to improve the implementation and the designs of the information system which takes the responsibility to manage data in the remote session by the physiotherapist or rehabilitation specialist.

References

1. J. Ojeda, F. Moreno, E. Ramırez, and O. Rodrıguez, Gesture-gross recognition of upper limbs to physical rehabilitation, *Proceedings of the International Congress of Numerical Methods in Engineering and Applied Sciences (CIMENICS)* **(2014)**, pp. PI7–12.
2. H. M. Feys, W. J. De Weerdt, B. E. Selz, and G. A. Cox Steck, *Controlled Multicenter Trial Stroke* 29, 785 **(1998)**.
3. K. Kyung, Y. M. Kim, and E. K. Kim, *Journal of Physical Therapy Science* 26, 417 **(2014)**.
4. D. H. Gates, L. S. Walters, J. Cowley, J. M. Wilken, and L. Resnik, *Am. J. Occup. Ther.* 70, 7001350010p1 **(2016)**.
5. G. C. Burdea, D. Cioi, J. Martin, D. Fensterheim, and M. Holenski, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* i8, 505 **(2010)**.
6. G. Kwakkel, B. J. Kollen, J. van der Grond, and A. J. H. Prevo, *Stroke* 34, 2181 **(2003)**.
7. Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh, and Mo Nours Arab, *International Journal on Smart Sensing and Intelligent Systems* 1, 137 **(2008)**.
8. J. Grudin, A moving target: The evolution of human-computer interaction, Human-Computer Interaction Handbook, edited by J. Jacko, 3rd edn., Taylor & Francis, CRC Press, Boca Raton, London, New York **(2012)**.
9. Merians, D. Jack, R. Boian, M. Tremaine, G. Burdea, S. Adamovich, M. Recce, and H. Poizner, *Phys. Ther.* 82, 898 **(2002)**.
10. M. McNeill, et al., Immersive virtual reality for upper limb rehabilitation following stroke, *Systems, IEEE International Conference on Man and Cybernetics* **(2004)**, ISBN: 0-7803-8566-7.
11. R. Lloréns, M. Alcañiz, C. Colomer, and M. D. Navarro, *Technologies in the Behavioral, Social and Neurosciences* 181, 108 **(2012)**.
12. M. Pedraza-Hueso, S. Martín-Calzón, F. J. Díaz-Pernas, and M. Martínez-Zarzuela, Rehabilitation using Kinect-based games and virtual reality, *International Conference on Virtual and Augmented Reality in Education, Procedia Computer Science* **(2015)**, Vol. 75, pp. 161–168.
13. S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd edn., Prentice Hall, Englewood Cliffs, New Jersey, USA **(2009)**.
14. M. Cameirão, S. Bermúdez, E. Duarte, and P. Verschure, *Restorative Neurology and Neuroscience* 29, 287 **(2011)**.
15. V. Szücs and C. Sik Lanyi, Abilities and limitations of assistive technologies in post-stroke therapy based on virtual/augmented reality, assistive technology: From research to practice, *12th European AAATE Conference*, IOS Press, Portugal **(2013)**, pp. 1087–1091.

Received: 6 October 2017. Accepted: 22 October 2017.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1452–1460, 2018

RESEARCH ARTICLE

MHD Flow and Heat Transfer of Non-Newtonian Nanofluids Over a Nonlinear Stretching Sheet

M. R. Krishnamurthy[1], K. Ganesh Kumar[1], B. J. Gireesha[1], and N. G. Rudraswamy[2,*]

[1] Department of Studies and Research in Mathematics, Kuvempu University, Shankaraghatta 577451, Shimoga, Karnataka, India
[2] Department of Mathematics, Sahyadri Science College (Autonomous), Shimoga 577201, Karnataka, India

This paper deals with the study of MHD boundary layer flow and heat transfer of non-Newtonian nanofluids past a nonlinear stretching sheet with the effects of nonlinear thermal radiation, chemical reaction, non-uniform heat source/sink, viscous dissipation and convective boundary condition. Here we have considered the different types of non-Newtonian nanofluid models namely, Casson nanofluid, Sisko nanofluid, Micropolar nanofluid and Carreau nanofluid. Similarity transformations are implemented to the governing partial differential equations to obtain coupled nonlinear ordinary differential equations. To check the behavior of these different types models we have employ as an efficient numerical method, Runge-Kutta-Fehlberg fourth-fifth order along with Shooting technique is used. The influence of several emerging physical parameters on velocity, temperature and concentration profiles in both linear and nonlinear stretching sheet in the presence of linear and nonlinear thermal radiation are studied and analyzed through graphs and tables in detail.

Keywords: Nanoparticle, Non-Newtonian Fluids, Nonlinear Thermal Radiation, Convective Boundary Condition, Numerical Solution.

1. INTRODUCTION

Heat transfer through convective boundary conditions has a great importance in thermal energy storage gas turbines, nuclear turbines and chemical reactions. Makinde and Aziz[1] numerically reported the boundary layer flow induced in a nanofluid due to a linearly stretching sheet with a convective boundary condition. Nield and Kuznetsov[2] and Kuznetsov and Nield[3] investigated the natural convective boundary layer flow of a nanofluid employing Buongiorno model. Das[4] analyzed the flow and convective heat transfer of nanofluid over a permeable stretching sheet with slip condition. Rahman and Eltayeb[5] have presented the flow and radiative heat transfer of nanofluid past a nonlinear stretching surface with convective boundary conditions.

Different models of non-Newtonian fluids based on their diverse flow behaviors have been proposed by the researchers. Casson fluid is classified as the most popular non-Newtonian fluid which has several applications in food processing, metallurgy, drilling operations and bioengineering operations. Nadeem et al.[6] investigated the steady flow of Casson fluid in the presence of nanoparticles and considered the magnetohydrodynamics effect within the fluid and convective condition along the surface. The mixed convection MHD flow of Casson nanofluid over a nonlinear permeable stretching sheet with viscous dissipation was considered by Prabhakar and Shanker.[7] Recently, many authors are studied the boundary layer flow and heat transfer of non-Newtonian fluid under various effects.[8–12]

The Sisko fluid model is one of the most important fluid model portraying the flow in the power-law and upper Newtonian regions. Aforementioned model is most appropriate to describe the flow behavior of fluids in high shear rate regions. Several recent studies on the Sisko nanofluid flow and heat transfer over a nonlinear stretching sheet considered with different types of effects have been reported by Refs. [13–15]. The Carreau fluid, holds constitutive relationship at both low and high shear rates. Due to this fact, it has achieved wider acceptance at present. Various aspects of the flow and heat transfer of Carreau fluid in the presence of nanoparticle over a stretching sheet have been explored by many researchers.[16–22] The Micropolar fluid is also one of the important non-Newtonian fluid, theory of Micropolar fluid was first proposed by Eringen.[23] The comprehensive references on this topic can be found in the some review papers, for example Refs. [24–29].

The present chapter deals with the study of MHD boundary layer flow and heat transfer of non-Newtonian nanofluids past a nonlinear stretching sheet under the influence of nonlinear thermal radiation, chemical reaction, non-uniform heat source/sink, viscous dissipation and

*Author to whom correspondence should be addressed.

1546-1955/2018/15/1452/009 doi:10.1166/jctn.2018.7377

convective boundary condition. Here the Casson nanofluid model is considered and results obtained are compared with Sisko nanofluid, Micropolar nanofluid and Carreau nanofluid. In particular, the boundary layer flow and heat transfer of Casson nanofluid past a nonlinearly stretching sheet is considered with various effects. Dimensionless expressions for velocity and temperature are solved numerically using Runge-Kutta-Fehlberg-45 order method with the help of symbolic algebraic software MAPLE. The presented plots illustrate the behavior of pertinent parameters on velocity and temperature. Numerical values of skin friction coefficient and local Nusselt number for different types of non-Newtonian fluids with nanoparticles are tabulated and are analyzed.

2. MATHEMATICAL FORMULATION

A steady flow of Casson fluid with nanoparticles caused by nonlinearly stretching sheet through porous medium is considered. The x-axis is taken along the direction of stretching sheet and y-axis is perpendicular to the surface. The sheet is stretched with the nonlinear velocity of the form $U_w = ax^n$ where $a > 0$, $n \geq 0$ are constants, which is shown in Figure 1. The flow is subjected to a variable magnetic field of strength $B(x) = B_0 x^{(n-1)/2}$. The electric field is absent whereas the induced magnetic field is neglected by assuming low magnetic Reynolds number. The effects of chemical reaction, nonlinear thermal radiation, viscous dissipation and non-uniform heat source/sink are considered in the energy equation. Furthermore, the temperature of the sheet is described by a function $T_w(x) = T_\infty + A_b x^n$ where $A_b > 0$ is constant and T_∞ is the ambient fluid temperature.

The continuity, momentum, energy and concentration equations for a steady boundary layer flow of Casson nanofluid are as follows,

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \tag{1}$$

$$u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = \nu_f\left(1 + \frac{1}{\beta_c}\right)\frac{\partial^2 u}{\partial y^2} - \frac{\sigma B^2}{\rho_f}u - \frac{\nu_f}{k'}u \tag{2}$$

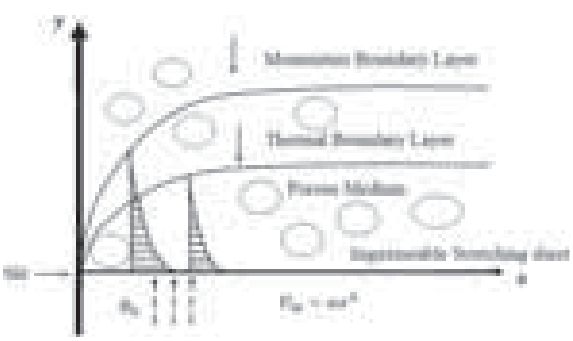


Fig. 1. Schematic representation of boundary layer flow.

$$u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \alpha\frac{\partial^2 T}{\partial y^2} + \tau\left[D_B\frac{\partial C}{\partial y}\frac{\partial T}{\partial y} + \frac{D_T}{T_\infty}\left(\frac{\partial T}{\partial y}\right)^2\right] - \frac{1}{(\rho c)_f}\frac{\partial q_r}{\partial y} + q''' + \mu_f\left(\frac{\partial u}{\partial y}\right)^2 \tag{3}$$

$$u\frac{\partial C}{\partial x} + v\frac{\partial C}{\partial y} = D_B\frac{\partial^2 C}{\partial y^2} + \frac{D_T}{T_\infty}\frac{\partial^2 T}{\partial y^2} - k_0(C - C_\infty) \tag{4}$$

where u and v are the velocity components along the x and y directions respectively, ν_f is kinematic viscosity of the nanofluid, σ is electrical conductivity, ρ_f is fluid density, ρ_p is density of nanoparticles, k' is the permeability of the porous medium, β_c is Casson fluid parameter, $\alpha = k/(\rho c)_f$ is nanofluid thermal diffusivity, D_B is Brownian diffusion coefficient, D_T is thermophoretic diffusion coefficient, $\tau = (\rho c)_p/(\rho c)_f$ is ratio of effective heat capacity of the nanoparticle material to the effective heat capacity of the fluid, μ_f is dynamic viscosity, q_r is radiative heat flux, q''' is non-uniform heat source/sink and k_0 is chemical reaction coefficient.

The boundary conditions in the present problem are

$$\begin{gathered} u = U_w = ax^n, \quad v = 0, \quad -k_f\frac{\partial T}{\partial y} = h_f(T_f - T), \\ C = C_w \quad \text{at } y = 0, \\ u = 0, \quad T = T_\infty, \quad C = C_\infty \quad \text{as } y \to \infty \end{gathered} \tag{5}$$

$$q_r = -\frac{4\sigma^*}{3k^*}\frac{\partial T^4}{\partial y} = -\frac{16\sigma^*}{3k^*}T^3\frac{\partial T}{\partial y} \tag{6}$$

$$q''' = \left(\frac{k_{nf}U_w(x)}{x\nu_f}\right)\left[A_1(T_f - T_\infty)f'(\eta) + B_1(T - T_\infty)\right] \tag{7}$$

Using Rosseland approximation for radiation and space and temperature dependent internal heat generation/absorption (non-uniform heat source/sink), given as in Eqs. (6) and (7), the Eq. (3) will becomes

$$u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \frac{\partial}{\partial y}\left[\left(\alpha + \frac{16\sigma^* T^3}{3k^*(\rho c)_f}\right)\frac{\partial T}{\partial y}\right] + \tau\left[D_B\frac{\partial C}{\partial y}\frac{\partial T}{\partial y} + \frac{D_T}{T_\infty}\left(\frac{\partial T}{\partial y}\right)^2\right] + q''' + \mu_f\left(\frac{\partial u}{\partial y}\right)^2 \tag{8}$$

Introduce the similarity transformations as,

$$\begin{gathered} u = ax^n f'(\eta), \\ v = -ax^{(n-1)/2}\sqrt{\frac{\nu_f}{a}}\left(\frac{n+1}{2}f(\eta) + \frac{n-1}{2}\eta f'(\eta)\right), \\ \eta = \sqrt{\frac{a}{\nu_f}}x^{(n-1)/2}y, \quad T = T_\infty(1 + (\theta_w - 1)\theta(\eta)) \\ \text{and} \quad \phi(\eta) = \frac{C - C_\infty}{C_w - C_\infty} \end{gathered} \tag{9}$$

where $\theta_w = T_f/T_\infty$, $\theta_w > 1$ is the temperature ratio parameter. With the help of aforementioned transformations, Eq. (1) is identically satisfied and Eqs. (2), (4) and (8) will take the following forms;

$$\left(1+\frac{1}{\beta_c}\right)f''' + \left(\frac{n+1}{2}\right)ff'' - nf'^2 - Qf' - k_p f' = 0 \tag{10}$$

$$\left[\left(1+Nr(1+(\theta_w-1)\theta)^3\right)\theta'\right]' + Pr\left(\frac{n+1}{2}\right)f\theta' - Prnf'\theta + PrNb\phi'\theta' + PrNt(\theta')^2 + A_1 f' + B_1\theta + PrEc(f'')^2 = 0 \tag{11}$$

$$\phi'' + \left(\frac{n+1}{2}\right)Scf\phi' + \frac{Nt}{Nb}\theta'' - \gamma\phi = 0 \tag{12}$$

The corresponding boundary conditions are;

$$f(\eta)=0, \quad f'(\eta)=1, \quad \theta'(\eta) = -B_i(1-\theta(0)),$$
$$\phi(\eta)=1 \quad \text{at } \eta=0,$$
$$f'(\eta)\to 0, \quad \theta(\eta)\to 0, \quad \phi(\eta)\to 0 \quad \text{as } \eta\to\infty \tag{13}$$

where f, θ and ϕ are functions of η and prime denotes derivatives with respect to η. $\beta_c = \mu_B(\sqrt{2\pi_c}/p_y)$ is Casson fluid parameter, μ_B is plastic dynamic viscosity of Casson fluid, $Q = \sigma B_0^2/\rho_f a$ is magnetic parameter called Hartmann number, $Nr = 16\sigma^* T_\infty^3/(3kk^*)$ is radiation parameter, $Nb = \tau D_B(C_w - C_\infty)/\nu_f$ is Brownian motion parameter, $Nt = \tau D_T(T_f - T_\infty)/\nu_f T_\infty$ is thermophoresis parameter, $Sc = \nu_f/D_B$ is Schmidth number, $Pr = \nu_f/\alpha$ is Prandtl number and $\gamma = k_0 Sc/a$ is chemical reaction parameter.

The skin friction coefficient (C_{fx}), local Nusselt number (Nu_x) and Local Sherwood number (Sh_x) are given by,

$$C_{fx} = \frac{\tau_w}{\rho U_w^2}, \quad Nu_x = \frac{U_w q_w}{ak(T_f - T_\infty)} \quad \text{and} \quad Sh_x = \frac{U_w q_m}{aD_B(C_w - C_\infty)} \tag{14}$$

where the shear stress along the stretching surface τ_w, the surface heat flux q_w and the surface mass flux q_m are given by

$$\tau_w = \mu_B\left(1+\frac{1}{\beta_c}\right)\left(\frac{\partial u}{\partial y}\right)_{y=0}, \quad q_w = -k\left(\frac{\partial T}{\partial y} + (q_r)_w\right)_{y=0} \quad \text{and} \quad q_m = -k\left(\frac{\partial C}{\partial y}\right)_{y=0} \tag{15}$$

Substituting the values of τ_w and q_w into the Eq. (14) we have

$$\sqrt{Re_x}C_{fx} = \left(1+\frac{1}{\beta_c}\right)f''(0), \quad \frac{Nu_x}{\sqrt{Re_x}} = -(1+Nr\theta_w^3)\theta'(0) \quad \text{and} \quad \frac{Sh_x}{\sqrt{Re_x}} = -\phi'(0) \tag{16}$$

here $Re_x = U_w^2/(a\nu_f)$ is the local Reynolds number.

3. NUMERICAL PROCEDURE

The system of nonlinear differential equations (10)–(12) along with the boundary condition (13) are solved numerically using fourth-fifth order Runge-Kutta-Fehlberg method with the help of symbolic algebraic software MAPLE. The absolute convergence criteria were taken as 10^{-6} in this method. It is most important to choose the appropriate finite values of η_∞. The asymptotic boundary conditions at η_∞ were replaced by η_6 in accordance with standard practice in the boundary layer analysis. The choice of $\eta_\infty = 6$ ensures that all numerical solutions approach the asymptotic values correctly.

4. RESULTS AND DISCUSSION

In order to analyze the results, numerical computation has been carried out using the method described in the previous section for various values of magnetic parameter (Q), Casson fluid parameter (β_c), nonlinear stretching parameter (n), Prandtl number (Pr), Schmidt number (Sc), Brownian motion parameter (Nb), thermophoresis parameter (Nt), thermal radiation parameter (Nr), temperature ratio parameter (θ_w), chemical reaction parameter (γ) and heat source/sink parameter (A_1 and B_1). For illustrations of the results, numerical values are plotted in Figures 2–23 and Tables II, III shows the values of skin friction coefficient, Nusselt number and Sherwood number for different

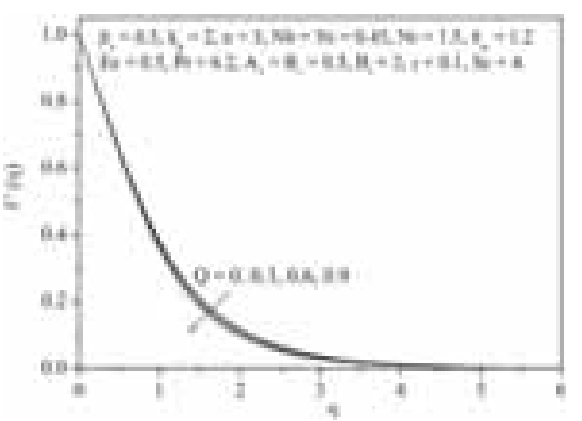


Fig. 2. Velocity field for different values of the magnetic parameter (Q).

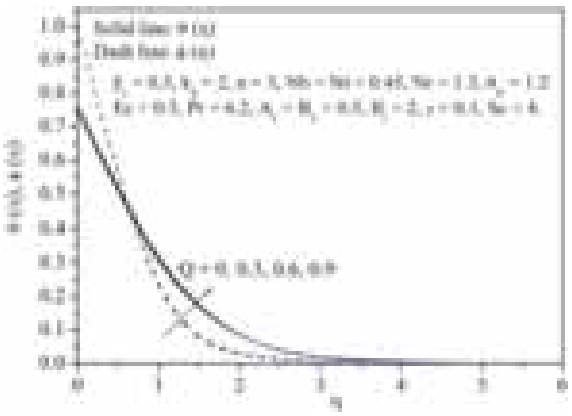


Fig. 3. Temperature and concentration field for different values of the magnetic parameter (Q).

types of non-Newtonian nanofluids with various values of the pertinent parameters in the presence of temperature ratio parameter and nonlinear stretching parameter.

Figures 2 and 3 respectively shows the effect of magnetic parameter (Q) on velocity, temperature and concentration profiles. In these figures it is observed that, with increasing values of Q, the velocity is found to be decreases. This is due to the fact that the applied magnetic field normal to the flow direction induces the drag in terms of Lorentz force which provides resistance to flow. The temperature and concentration profile increases with increasing values of Q, which is shown in Figure 3.

The Casson fluid parameter (β_c) on velocity, temperature and concentration distribution has been respectively depicted in Figures 4 and 5. From these figures, we observed that the velocity field decreases when β_c increases. An increase in β_c leads to an increase in plastic dynamic viscosity that creates resistance in the flow of fluid and a decrease in fluid velocity is observed. Figure 4

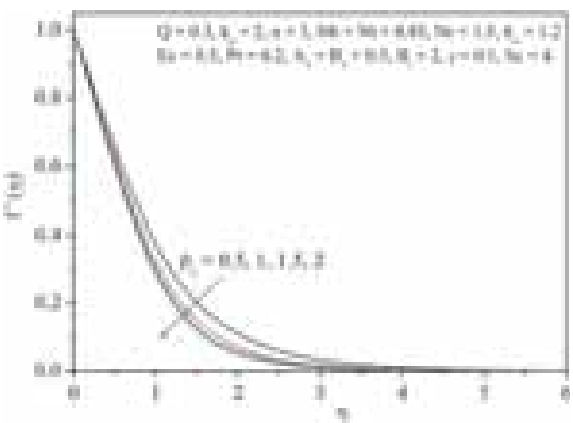


Fig. 4. Velocity field for different values of the Casson fluid parameter (β_c).

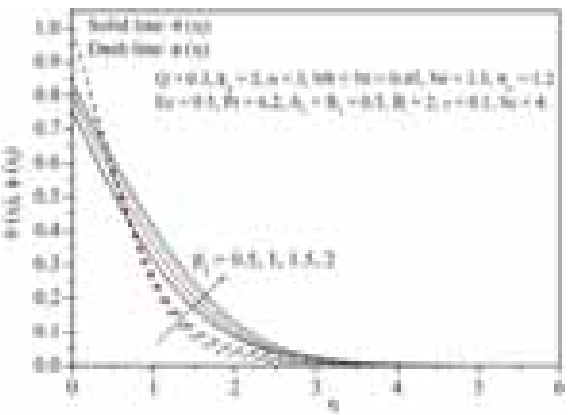


Fig. 5. Temperature and concentration field for different values of the Casson fluid parameter (β_c).

shows that β_c has quite opposite effect on the velocity compared to temperature and concentration profiles.

Figures 6 and 7 illustrates the effect of permeability parameter (k_p) on velocity, temperature and concentration

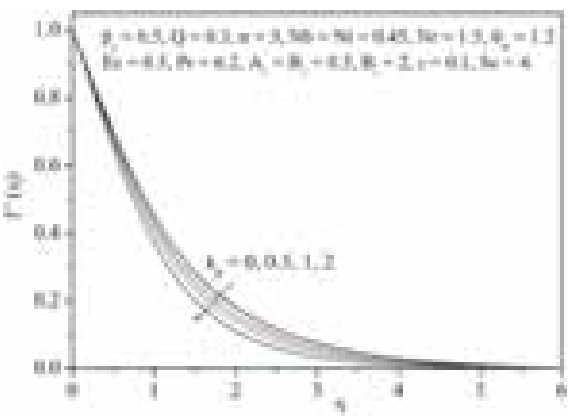


Fig. 6. Velocity field for different values of the permeability parameter (k_p).

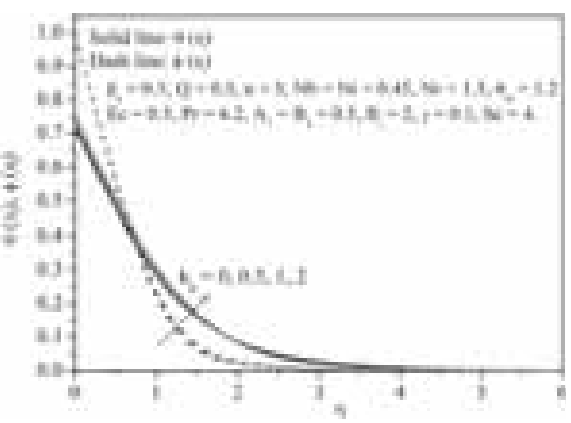


Fig. 7. Temperature and concentration field for different values of the permeability parameter (k_p).

RESEARCH ARTICLE

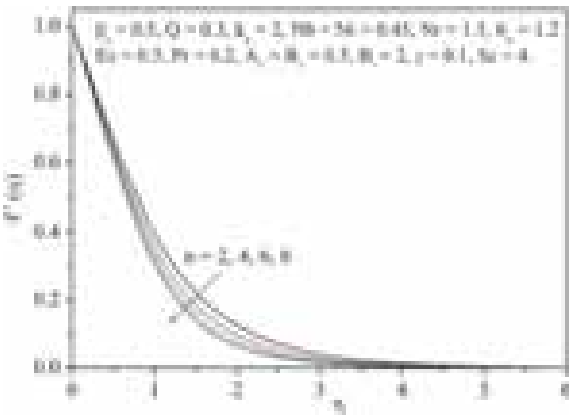


Fig. 8. Velocity field for different values of the nonlinear stretching parameter (n).

profiles. It is quite clear that there is a decrease in velocity with increase in permeable parameter. This is due to the fact that, as the permeability parameter increases, the resistance of the porous medium increases, which causes

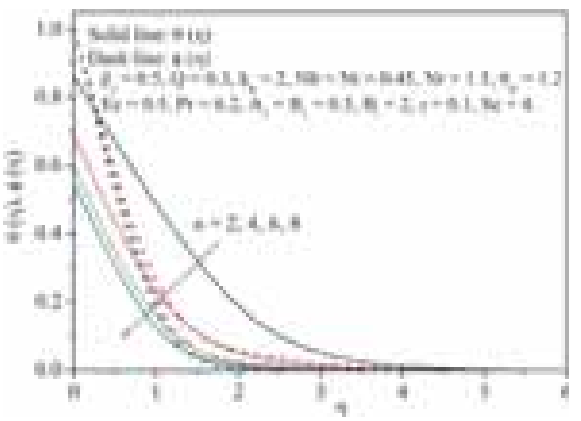


Fig. 9. Temperature and concentration field for different values of the nonlinear stretching parameter (n).

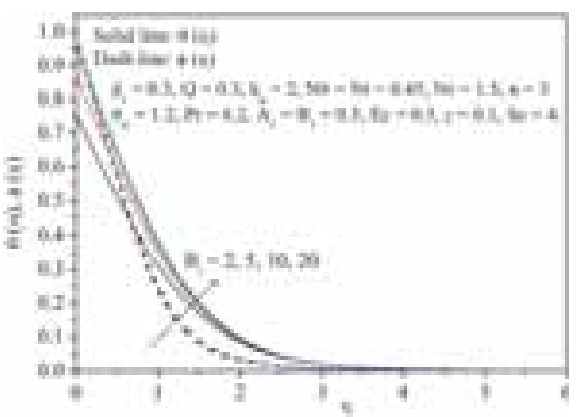


Fig. 10. Temperature and concentration field for different values of the Biot number (B_i).

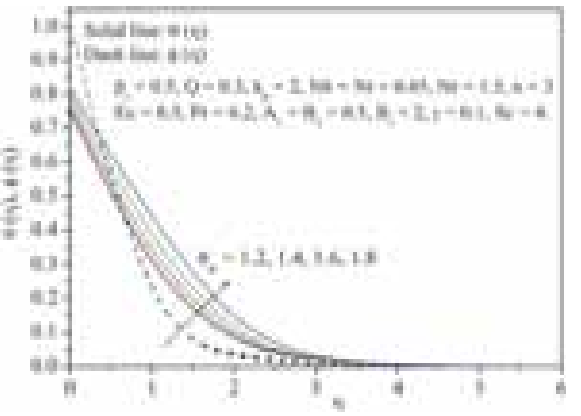


Fig. 11. Temperature and concentration field for different values of the temperature ratio parameter (θ_w).

the fluid to decelerate. Also it is observed that, the effect of values of permeable parameter contributes to the thickening of thermal boundary layer. This is evident from the fact that, the porous medium opposes the fluid motion.

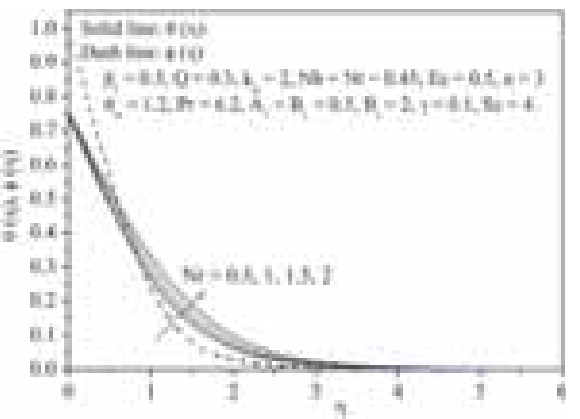


Fig. 12. Temperature and concentration field for different values of the thermal radiation parameter (Nr).

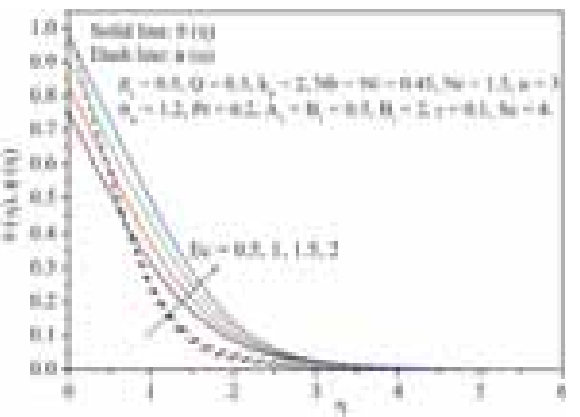


Fig. 13. Temperature and concentration field for different values of the Eckert number (Ec).

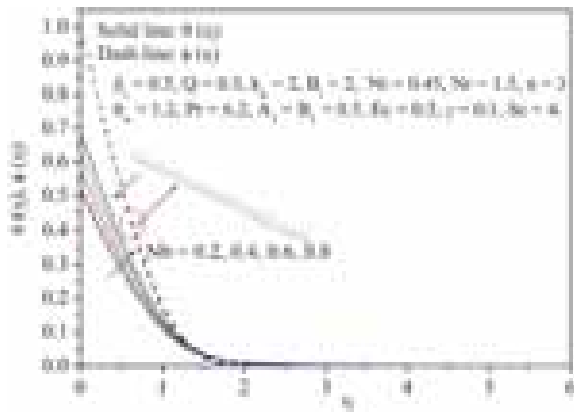


Fig. 14. Temperature and concentration field for different values of the Brownian motion parameter (Nb).

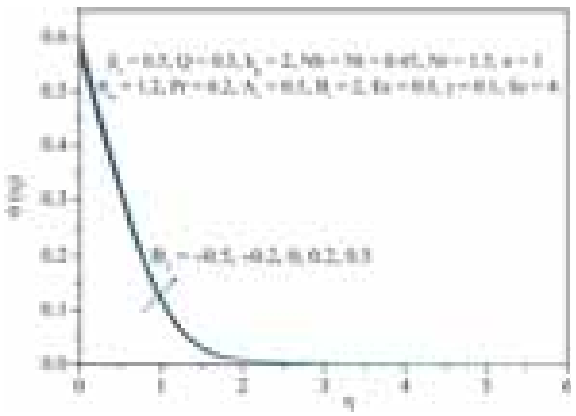


Fig. 17. Temperature field for different values of the heat sink parameter (B_1).

The resistance offered to the flow is responsible in enhancing the temperature and concentration.

Figures 8 and 9 depict the velocity, temperature and the concentration profiles for different values of nonlinearly stretching parameter (n). The fluid velocity is found to decrease with increasing n, while the temperature and concentration also decreases in this case. The effect of nonlinear stretching parameter is significant when n is low.

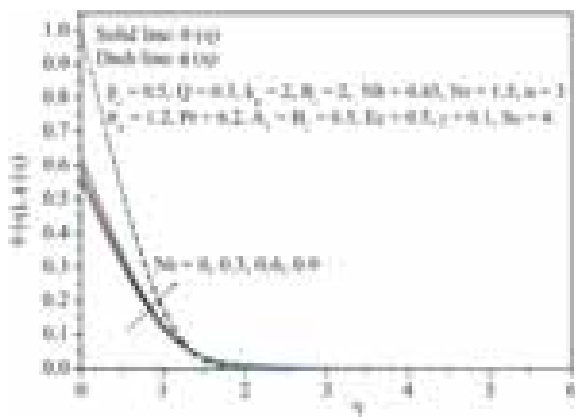


Fig. 15. Temperature and concentration field for different values of the thermophoresis parameter (Nt).

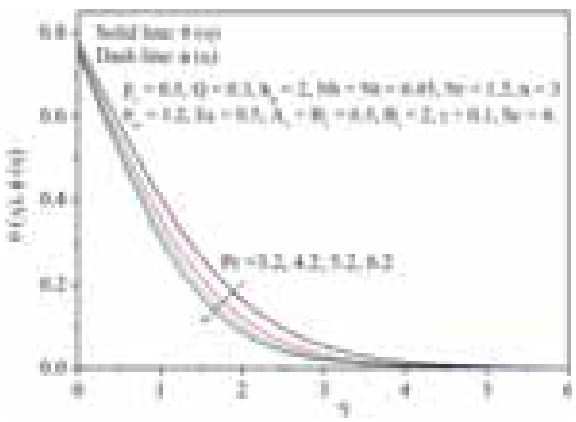


Fig. 18. Temperature field for different values of the Prandl number (Pr).

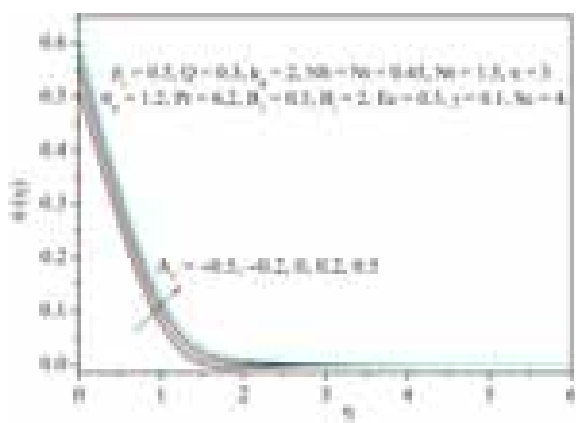


Fig. 16. Temperature field for different values of the heat source parameter (A_1).

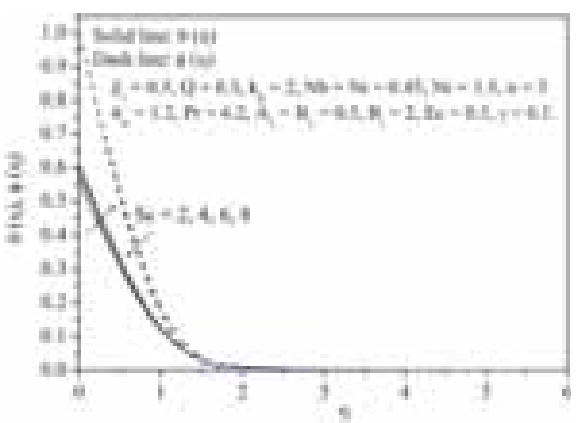


Fig. 19. Temperature field for different values of the Schmidth number (Sc).

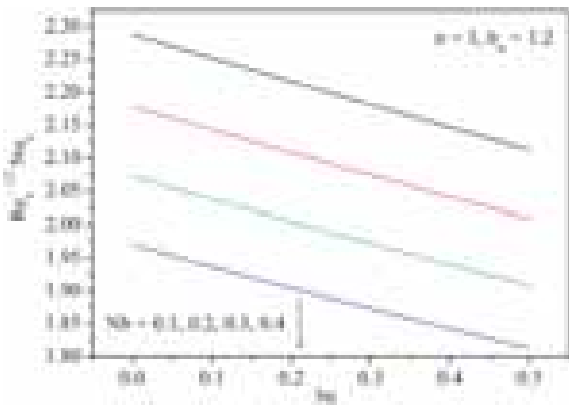


Fig. 20. Nusselt number and Sherwood number for different values of Brownian motion parameter (Nb) and thermophoresis parameter (Nt).

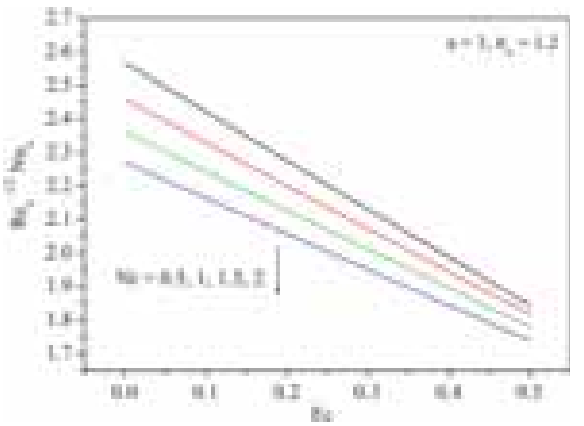


Fig. 22. Nusselt number and Sherwood number for different values of thermal radiation parameter (Nr) and Eckert number (Ec).

But the decrease in the velocity profile is negligible for a large n.

The variation of temperature $\theta(\eta)$ and mass concentration $\phi(\eta)$ with an increase in Biot number (B_i) can be seen in Figure 10. Here a gradual increase in B_i results in the larger convection at the stretching sheet which rises the temperature.

Figure 11 illustrates the effect of temperature ratio parameter (θ_w) on $\theta(\eta)$ and $\phi(\eta)$ curves. It is noticed that increasing the temperature ratio parameter θ_w increases the thermal state of the fluid, resulting in increases in temperature and concentration profiles. Figure 12 shows that the larger radiation parameter (Nr) corresponds to a increase in the temperature and concentration profiles. Increase in radiation parameter provides more heat to fluid that depicts an enhancement in the temperature and thermal boundary layer thickness. Figure 13 shows that as the Eckert number (Ec) increases the thermal boundary layer thickness. This phenomenon occurs because increase in the Eckert number means there will be increase of viscous dissipation that is the reason for temperature enhancement in the boundary layer.

Figure 14 illustrate the impact of the Brownian motion parameter (Nb) on the temperature and concentration distributions. Physically, the Brownian motion is stronger in the case of smaller nanoparticles which corresponds to the larger Nb and converse is the situation for smaller values of Nb. In thermal conduction nanoparticle's motion plays a pivotal role. Due to the intensified chaotic motion of the nanoparticles the kinetic energy of the particles is enhanced which as a result enhances the temperature of the nanofluid. Hence increasing Nb firmly raises temperature values throughout the regime. Away from the surface the larger values of Nb stifle the diffusion of the nanoparticles in the fluid regime which reduces the concentration distribution.

Figure 15 shows the impact of thermophoresis parameter (Nt) on the thermal and concentration boundary layer. A stronger thermophoretic force drives the particles from the hot sheet towards the ambient fluid and thus affects

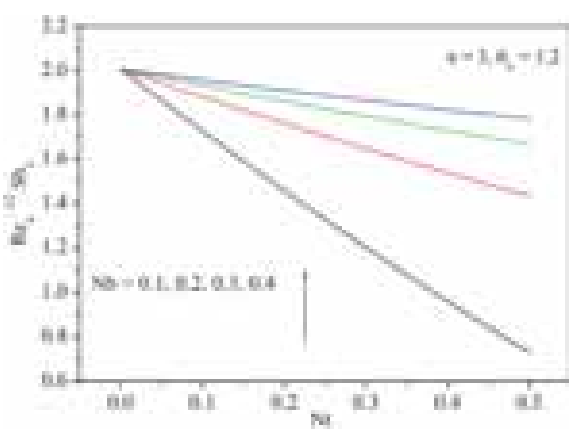


Fig. 21. Nusselt number and Sherwood number for different values of Brownian motion parameter (Nb) and thermophoresis parameter (Nt).

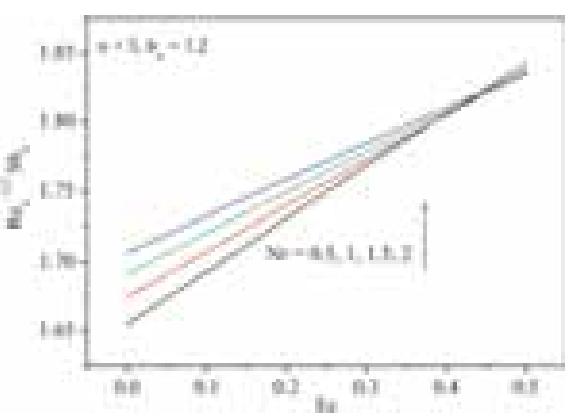


Fig. 23. Nusselt number and Sherwood number for different values of thermal radiation parameter (Nr) and Eckert number (Ec).

Table I. Numerical values of skin friction coefficient ($\sqrt{Re_x}C_{fx}$) and Nusselt number ($Nu_x/\sqrt{Re_x}$) for different types of non-Newtonian nanofluids for the different values of Q, k_p and β_c in the presence of $\theta_w = 1.2$ and $n = 3$ when $Nb = Nt = 0.45$, $Pr = 6.2$, $A_1 = B_1 = 0.5$, $Ec = 0.5$, $\gamma = 0.1$. Here, β_c is Casson fluid parameter, A_c is material parameter of Sisko fluid and λ_c is the material parameter of Carreau fluid.

			Casson nanofluid		Sisko nanofluid		Careau nanofluid	
Q	k_p	β_c, A_c, λ_c	$\sqrt{Re_x}C_{fx}$	$Nu_x/\sqrt{Re_x}$	$\sqrt{Re_x}C_{fx}$	$Nu_x/\sqrt{Re_x}$	$\sqrt{Re_x}C_{fx}$	$Nu_x/\sqrt{Re_x}$
0			4.8473	2.9661	1.0675	0.2601	0.9169	0.3599
0.3			4.9634	2.9458	1.2355	0.1382	1.1462	0.1790
0.6			5.0761	2.9257	1.3948	0.2862	1.3294	0.1862
	0		4.1058	3.0877	1.0675	0.2600	0.9169	0.3590
	1		4.5614	3.0148	1.5954	0.1303	1.5316	0.1745
	2		4.9634	2.9458	2.0567	0.3799	1.9186	0.5842
		0.5	4.9634	2.9458	1.3427	0.6379	1.0465	0.2230
		1	4.2844	2.6693	1.5588	0.3110	1.1021	0.3800
		2	3.8861	2.3847	1.9548	0.5901	1.2300	0.5290

larger extent of the fluid. Due to this reason, thermal boundary layer grows with Nt. An increase in Nt shifts the nanoparticles away from the sheet and forms a particle free layer in the vicinity of the sheet.

Figures 16 and 17, illustrates the temperature profiles $\theta(\eta)$ versus η for different values of A_1 and B_1. From these plots, it is evident that increase in A_1 [$A_1 > 0$ (heat source)], results in the enhancement of temperature, whereas $A_1 < 0$ (heat sink) leads to decrease in the thermal boundary layer. It is due to fact that for $A_1 > 0$, the thermal boundary layer generates the energy and obviously energy is absorbed for decreasing values of A_1 resulting in temperature dropping significantly near the boundary layer. Similar effect is observed form the Figure 17 for the effect of temperature-dependent heat source/sink parameter B_1.

The impact of generalized Prandtl number (Pr) on the temperature ($\theta(\eta)$) distribution is show in the Figure 18. Physically, enhancing the Prandtl number results in reduction in thermal diffusivity. It can be seen that the temperature and nanoparticle concentration profiles decrease with increase in the generalized Prandtl number.

Figure 19 depicts the influence of Schmidt number (Sc) on the temperature and concentration distribution. Schmidt number is analogous of Pr for mass transfer and relates the viscous or hydrodynamic boundary layer to nanoparticle volume fraction boundary layer. Increasing Sc corresponds to stronger viscous diffusion compared to mass diffusion which enhances molecular motions and their interactions and hence temperature rises. The variation in temperature $\theta(\eta)$ with increase in Sc is only accounted in the vicinity of the stretching sheet and nanoparticle volume fraction $\phi(\eta)$ decreases with increase in Sc.

Figures 20–23 shows the effect of Nusselt and Sherwood number for different values of Nb with Nt and Nr with Ec. Here, the Nusselt number decrease and Sherwood number increases respectively, for Nb and Nr with increase in Nt and Ec.

Table II. Values of Nusselt number ($Nu_x/\sqrt{Re_x}$) and Sherwood number ($Sh_x/\sqrt{Re_x}$) for different types of non-Newtonian nanofluids for the different values of pertinent parameters in the presence of $\theta_w = 1.2$ and $n = 3$ when $Pr = 6.2$, $A_1 = B_1 = 0.5$.

						Casson nanofluid		Sisko nanofluid		Careau nanofluid	
Ec	B_i	Nr	Nb	Nt	γ	$-\theta'(0)$	$-\phi'(0)$	$-\theta'(0)$	$-\phi'(0)$	$-\theta'(0)$	$-\phi'(0)$
0.5						2.9458	8.1390	0.6370	2.3049	0.2230	2.5484
1						2.3031	8.3307	0.9436	2.4916	0.2342	2.6200
2						1.0618	8.6971	0.7650	2.8233	0.1160	2.7647
	2					2.9458	8.1390	0.6370	2.3049	0.2230	2.5484
	5					4.1213	7.8826	0.7140	2.3040	0.2390	2.5485
	10					4.6818	7.7634	0.7443	2.3037	0.2444	2.5486
		0.5				1.5748	8.1280	0.3810	2.4000	0.5661	2.6498
		1				2.2812	8.1286	0.1910	2.3403	0.7550	2.5894
		2				3.5699	8.1533	0.9502	2.2820	0.3712	2.5193
			0.2			3.4669	7.0050	0.9721	2.4746	0.4210	2.6684
			0.4			3.0497	8.0268	0.7022	2.3219	0.2587	2.5617
			0.6			2.6402	8.3633	0.4441	2.2707	0.1310	2.5194
				0		3.1789	8.7450	0.1390	2.1581	0.4920	2.3872
				0.3		3.0249	8.3257	0.8871	2.2542	0.3007	2.4922
				0.6		2.8656	7.9678	0.3881	2.3573	0.1442	2.6053
					−0.1	2.9468	8.1102	0.6321	2.2517	0.2220	2.5059
					0	2.9463	8.1246	0.6352	2.2785	0.2231	2.5272
					0.1	2.9458	8.1390	0.6371	2.3049	0.2330	2.5484

Tables I and II shows the values of skin friction coefficient, Nusselt number and Sherwood number for different types of non-Newtonian nanofluids for the fixed values of temperature ratio parameter (θ_w) and nonlinear stretching parameter (n) by varying the values of pertinent parameters. From these tables, it is observed that compared to other non-Newtonian fluids Casson fluid has high thermal conductivity in the presence of nanoparticles.

5. CONCLUSION

The effects of viscous dissipation, non-uniform heat source sink and chemical reaction on MHD flow and nonlinear radiative heat transfer of Casson nanofluid induced by nonlinearly stretching surface considered with convective boundary condition is reported in this chapter. The computations to the nonlinear analysis are successfully performed using Runge-Kutta-Fehlberg-45 order method with the help of symbolic algebraic software MAPLE. The key results of this study are summarized as below:

- The velocity and momentum boundary layer thickness are decreasing functions and both the profiles of temperature and nanoparticle volume fraction are found to increase upon increasing the Casson fluid parameter (β_c).
- Increasing the values of magnetic parameter (Q) and permeability parameter (k_p), the velocity profile decreases and both temperature and concentration profiles increases.
- Nonlinear Rosseland approximation enhances the thermal boundary layer thickness.
- Chemical reaction parameter (γ) reduces the concentration distribution.
- For increasing values of Brownian motion parameter (Nb), reduces the nanoparticle volume fraction and increases the temperature distribution.
- Both temperature and nanoparticle volume fraction are increasing functions of thermophoresis parameter (Nt).
- Thermal radiation parameter (Nr) does not shown significant difference in the friction factor. But it helps to boost the temperature profiles of the flow.
- Temperature profile increases for increasing the values of Eckert number (Ec), heat source/sink parameter (A_1 and B_1), temperature ratio parameter (θ_w) and Biot number (B_i).
- The temperature profile reduces for an increase in the values of Schmidt number (Sc) and Prandtl number (Pr).
- The Casson fluid has more thermal conductivity when compared to other non-Newtonian fluids in the presence of nanoparticles.

References

1. O. D. Makinde and A. Aziz, *Int. J. Therm. Sci.* 50, 1326 **(2011)**.
2. D. A. Nield and A. V. Kuznetsov, *Int. J. of Heat and Mass Transfer* 52, 5792 **(2009)**.
3. A. V. Kuznetsov and D. A. Nield, *Int. J. of Heat and Mass Transfer* 49, 243 **(2010)**.
4. K. Das, *J. of the Egyptian Mathematical Society* 23, 451 **(2015)**.
5. M. M. Rahman and I. A. Eltayeb, *Meccanica* 48, 601 **(2013)**.
6. S. Nadeem, R. Ul Haq, and N. S. Akbar, *IEEE Transactions on Nanotechnology* 13, 109 **(2014)**.
7. B. Prabhakar and B. Shanker, *Journal of Applied Mathematics and Physics* 3, 1580 **(2015)**.
8. N. G. Rudraswamy, K. G. Kumar, B. J. Gireesha, and R. S. R. Gorla, *J. of Nanoengineering and Nanomanufacturing* 6, 278 **(2017)**.
9. K. G. Kumar, N. G. Rudraswamy, B. J. Gireesha, and R. S. R. Gorla, *Nonlinear Engineering* 6, 207 **(2017)**.
10. N. G. Rudraswamy, K. G. Kumar, B. J. Gireesha, and R. S. R. Gorla, *J. Nanofluids* 6, 300 **(2017)**.
11. K. G. Kumar, B. J. Gireesha, M. R. Krishnamurthy, and N. G. Rudraswamy, *Results in Physics* 7, 3031 **(2017)**.
12. K. G. Kumar, B. J. Gireesha, S. Manjunath, and N. G. Rudraswamy, *Int. J. of Mechanical and Materials Engineering* 12, 18 **(2017)**.
13. M. Khan and R. Malik, *AIP Advances* 5, 127202 **(2015)**.
14. G. V. Ramanaiah, M. S. Babu, and M. Lavanya, *Int. Refereed J. Engi. Sci. (IRJES)* 5, 17 **(2016)**.
15. M. Khan, R. Malik, and M. Hussain, *AIP Advances* 6, 055315 **(2016)**.
16. N. Ali and T. Hayat, *Applied Mathematics and Computation* 193, 535 **(2007)**.
17. M. S. Tshehla, *Int. J. Phys. Sci.* 6, 3896 **(2011)**.
18. B. I. Olajuwon, *Thermal Science* 15, S241 **(2011)**.
19. K. G. Kumar, B. J. Gireesha, S. Manjunath, and N. G. Rudraswamy, *Results in Physics* 7, 3976 **(2017)**.
20. T. Hayat, S. Asad, M. Mustafa, and A. Alsaedi, *Applied Mathematics and Computation* 246, 12 **(2014)**.
21. N. S. Akbar, S. Nadeem, Rizwan Ul Haq, and S. Ye, *Ain Shams Engineering Journal* 5, 1233 **(2014)**.
22. Masood Khan, Hashim, and Ali Saleh Alshomrani, *Plos One* 11, e0157180 **(2016)**.
23. A. C. Eringen, *J. Math Mech.* 16, 1 **(1966)**.
24. O. D. Makinde, K. G. Kumar, S. Manjunatha, and B. J. Gireesha, *Defect and Diffusion Forum* 378, 125 **(2017)**.
25. A. A. Joneidi, D. D. Ganji, and M. Babaelahi, *Int. Commu. in Heat and Mass Transfer* 36, 1082 **(2009)**.
26. R. A. Mohamed and S. M. AboDahab, *Int. J. Therm. Sci.* 48, 1800 **(2009)**.
27. A. Rehman and S. Nadeem, *Chinese Physics Letters* 29, 124701 **(2012)**.
28. S. Nadeem, S. Masood, R. Mehmood, and M. A. Sadiq, *Plos One* 10, e0124016 **(2015)**.
29. K. Das and P. R. Duari, *Int. J. of Applied and Compu. Mathematics* 3, 3229 **(2016)**.

Received: 24 October 2017. Accepted: 26 November 2017.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1461–1470, 2018

Magnetotactic Bacteria Moment Migration Optimization Algorithm for Generators Real-Power Rescheduling in Deregulated Power System

Jagadeeswar Reddy Chintam[1,*], V. Geetha[2], and D. Mary[3]

[1] Research Scholar, Department of EEE, Government College of Technology, Anna University, Coimbatore 641013, Tamilnadu, India
[2] Associate Professor, Department of EEE, Government College of Technology, Anna University, Coimbatore 641013, Tamilnadu, India
[3] Retired Professor, Department of EEE, Government College of Technology, Anna University, Coimbatore 641013, Tamilnadu, India

RESEARCH ARTICLE

In deregulated competitive electricity scenario, the transmission network congestion is the most important challenging issue to control and operate the transmission network within reliable limits. This work introduces a novel approach to magnetotactic bacteria moment migration optimization algorithm (MBMMOA) for real-power rescheduling of generators with congestion management (CM). The MBMMOA is a kind of new bionic optimization algorithm with moments of magnetosomes along the magnetic field of lines of the earth as a base. The suggested algorithm effectively applied on small as well as large IEEE-bus Network Topologies. In this work, voltage, line and loading limits considered for transmission network safety. The proposed techniques give superior results compared than the other techniques such as SA, RSM, PSO, SA etc. for the optimal power flow issues.

Keywords: Deregulation, Congestion Management, Optimal Power Flow, Generator Rescheduling, Magnetotactic Bacteria Moment Migration.

1. INTRODUCTION

In the current scenario, the electric power industry has made various changes in the operating, controlling and management. Generation, transmission, and distribution companies are working as self-regulating entities after deregulation. The increasing demand for electricity is giving rise to the market participants for trading in the electricity market. The transaction requires, energy transferred from the sellers to buyers. In Class, all proposed transactions are unable to be accommodated simultaneously due to the constraints of physical, voltage and stability limits of transmission network then the network is said congested.[1]

The CM issues in the deregulated environment overcome through the conventional and new evolutionary optimization techniques. The gradient method, nonlinear and LPP method, etc. the conventional techniques are given in Refs. [1–4]. To find the global minimum is impossible by the conventional optimization techniques. However, we have the solution for finding global minimum by the evolutionary optimization techniques such as differential evolution genetic algorithm (DEGA),[5] tabu search optimization technique,[6] DE optimization technique,[7] modified DE optimization technique,[8] cuckoo optimization technique,[9] ABC optimization technique,[10] HIS optimization technique,[11] TLO technique,[12] ant lion optimization technique,[13] hybrid evolutionary firefly optimization technique[14] had been proposed. In Ref. [15], fuzzy adaptive bacterial foraging optimization technique used to find the optimal power flow (OPF) issue of the transmission lines through generators active power re-dispatch. PSO and RCGA methods are used for changing of real-power of generators by rescheduling for CM in the deregulated environment.[16,17] The MBMMOA approach[18] is used for the application of CM issues by generator real-power rescheduling. It effectively minimizes the losses, congestion cost and overloading of transmission lines.

2. MATHEMATICAL CALCULATION

The main objective function of the proposed work has reduced the cost of congestion and at the same time satisfying the transmission network constraints without violation. The CM issue is solving by the method of generator real-power rescheduling. The cost associated with a change in real power output of the generator depends on the generating company's (Generating Companies.) price

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1461/010 doi:10.1166/jctn.2018.7378

bids. Therefore, the objective of the problem may be stated as in Eq. (1).[16]

$$\text{Minimize } F_c = \sum_{j \in N_g} (a_k \Delta P_{gj}^{+} + b_k \Delta P_{gj}^{-}) \quad \$/h \tag{1}$$

Where F_c is total congestion cost, a_k and b_k are the coefficient of the generator, ΔP_{gj}^{+} and ΔP_{gj}^{-} shows the change in incremental/decremental price bid offered by the Generating Companies ($/MWh).

The equality and inequality constraints have taken for the optimization problem as given below.

2.1. Equality Constraints

The real and reactive power flow equations stated by Eqs. (2)–(5).[19]

$$P_{gk} - P_{dk} = \sum_j |V_j||V_k||Y_{kj}| \cos(\delta_k - \delta_j - \theta_{kj}); \quad j = 1, 2, \ldots, N_b \tag{2}$$

$$Q_{gk} - Q_{dk} = \sum_j |V_j||V_k||Y_{kj}| \sin(\delta_k - \delta_j - \theta_{kj}); \quad j = 1, 2, \ldots N_b \tag{3}$$

$$P_{gk} = P_{gk}^{c} + \Delta P_{gk}^{+} - \Delta P_{gk}^{-}; \quad k = 1, 2, \ldots N_g \tag{4}$$

$$P_{dj} = P_{dj}^{c}; \quad j = 1, 2, \ldots N_d \tag{5}$$

Where P_{gk} and Q_{gk} represents the generated active and reactive powers at bus k. P_{dk} and Q_{dk} are the load active and reactive powers at bus k. V_j and V_k represents the voltages at bus jth and kth. δ_j and δ_k represents the bus voltage angle of bus jth and kth. θ_{kj} is the admittance angle of the lines linked between kth and jth. N_g, N_b, N_d represents the number of generators, buses and loads, respectively. P_{gk}^{c} and P_{dj}^{c} represents the active power produced by generator k and the active power used by load bus jth, respectively. The Eqs. (2) and (3) represent the active and reactive power balances at each node whereas Eqs. (4) and (5) represent the final power as a function of market clearing price.

2.2. Inequality Constraints

The details of the inequality constraints of operating and physical limits of all the generators, transmission lines, transformers are stated by Eqs. (6)–(10).[19]

$$P_{gk}^{\min} \le P_{gk} \le P_{gk}^{\max}, \quad \forall k \in N_g \tag{6}$$

$$Q_{gk}^{\min} \le Q_{gk} \le Q_{gk}^{\max}, \quad \forall k \in N_g \tag{7}$$

$$(P_{gk} - P_{gk}^{\min}) = \Delta P_{gk}^{\min} \le \Delta P_{gk} \le \Delta P_{gk}^{\max} = (P_{gk}^{\max} - P_{gk}) \tag{8}$$

$$V_n^{\min} \le V_n \le V_n^{\max}, \quad \forall k \in N_l \tag{9}$$

$$P_{ij} \le P_{ij}^{\max} \tag{10}$$

Where the superscripts min and max represent the minimum and the maximum values of the respective variables and N_l represents the number of lines.

2.3. MBMMOA

Mo had suggested, an optimization algorithm called Magnetotactic Bacteria moment migration Optimization Algorithm (MBMMOA). It is a nature-inspired method of by the magnetotactic bacteria.[21] It is a bacteria of a different kind with many micro magnetic particles in their bodies. The magnetotactic particles called as magnetosomes. The MBMMA has three main operators including moment generation, moment migration, and moment replacement.[22]

2.3.1. Initialize Population

The initial population filled with N number of randomly generated n-dimensional real-valued vectors (i.e., cells of an MTB). Let $X_i^0 = (x_{i1}^0, x_{i2}^0, \ldots, x_{in}^0)$ represents the ith cell (for generation index $t = 0$) initialized randomly. Subsequently, each cell is generated as given under:

$$x_{ij}^0 = x_{\min j} + \text{rand}(0, 1) \times (x_{\max j,} - x_{\min j,}) \tag{11}$$

Where $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, n$. ¼1; 2; …; N, j¼1; 2; …; n. $x_{\max j}$ and $x_{\min j}$ are upper and lower bounds for the dimension j, respectively. rand(0,1) is a random number from the uniform distribution between 0 and 1.

2.3.2. Interaction Distance

In the algorithm, each solution is seen as a cell containing a magnetosome chain. Before getting the interaction energy of cells, the distance $d_{i,r}$ of two cells x_i and x_r calculated as follows:

$$d_{i,r} = x_i - x_r \tag{12}$$

Thus, a distance matrix $D = [d_{1,r}, d_{2,r}, \ldots d_{i,r}, \ldots, d_{n,r}]'$ is got, where r is a randomly chosen integer in [1, N]. N is the size of cell population.

2.3.3. Moments Generation

Based on the distances between cells, the interaction energy e_t between two cells is termed as

$$e_t = \left(\frac{d_{ij}(t)}{1 + c_1 \times D_{ir} + c_2 \times d_{pq}(t)} \right)^3 \tag{13}$$

where t is the current generation number, c_1 and c_2 are constants, d_{ij} is one element of distance matrix D; d_{pq} is a randomly selected element from D; P and r are random integers in [1, N]; $q \in [1, J]$ stands for one randomly selected dimension; J is the dimension of a cell and $d_{i,r}$ stands for the Euclidean distance between two cells x_i, and x_r.

After obtaining the interaction energy, the moments are generated as follows:

$$m_i(t) = \frac{e_i(t)}{B} \tag{14}$$

Where B is a constant.

Then the total moment of a cell is regulated as follows:

$$x_{i,j}(t) = x_{i,j}(t) + m_{r,q}(t) \times \text{rand} \tag{15}$$

Where $m_{r,q}$ is randomly selected element from m_i, rand is a random number in interval (0, 1).

2.3.4. Moments Migration

After the moment's generation, the moment's migration realized as given below:

If rand > 0.5, the moments in the cell migrate as below:

$$x_{ij}(t+1) = x_{ij}(t) \tag{16}$$

Otherwise

$$x_{ij}(t+1) = x_{ij}(t) + (x_{\text{best, q}}(t) - x_{i,q}(t)) \tag{17}$$

Where $x_{\text{best, q}}$ is the qth dimension of the best individual in the current generation.

2.3.5. Moments Replacement

After the moment's migration, some worse moments replaced in the following way:

$$x_i(t+1) = m_{r,q}(t) \times ((\text{rand}(1,J) - 1) \times \text{rand}(1,J)) \tag{18}$$

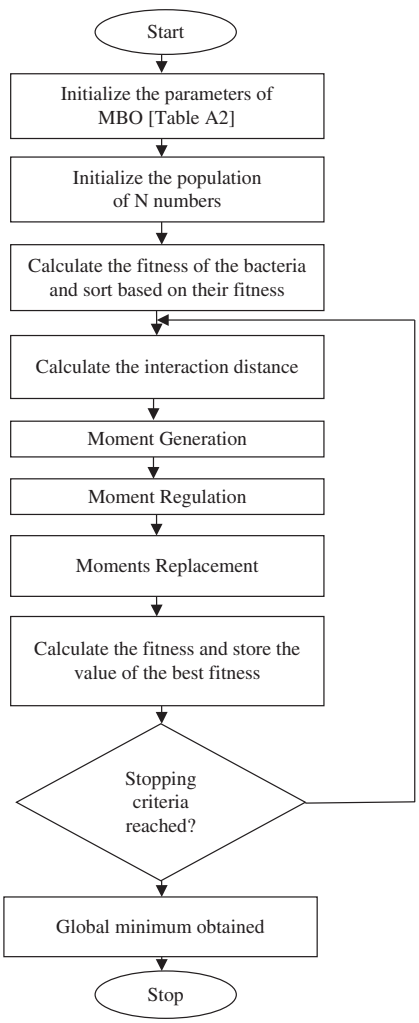


Fig. 1. Flowchart of the MBMMOA.

Where $m_{r,q}$ is the qth dimension of m_r. r is a randomly integer in $[1, N]$; $q \in [1, J]$ stands for one randomly selected dimension. rand $(1, J)$ is a random vector with J dimensions.

2.4. A Pseudo-Code of MBMMA is as Follows

I. Data Structures: Define the simple bounds and algorithm parameters.
II. Initialization: Randomly create the initial population in the search space.
III. While stop criteria is not met
For $i = 1{:}N$
Compute interaction distance according to Eq. (12)
end
for $i = 1{:}N$
Compute moments generation according to Eqs. (13)–(15)
end
Sort the population as per to fitness
For $i = 1{:}N$
Compute moments migration as per to Eqs. (16) and (17)
end
Sort the population according to fitness
For $i = 1{:}N/2$
Perform moments replacement according to Eq. (18)
end
End while

The detail of MBMMOA algorithm flow to solving the congestion issue in power system shown as a flowchart in Figure 1.

3. PROPOSED MBMMOA ALGORITHM FOR CM

In this work, the dimension is the number of generators participating in the CM problem. The penalty functions added to the objective function to construct the fitness function from transferred inequality constraints. The equality constraints and the reactive-power inequality constraints efficiently managed by Newton-Raphson power flow while the real-power inequality constraints are dealing with during the iteration process. Other inequality constraints of line power flow and load bus voltage considered as quadratic penalty functions. The fitness function of CM problem analyzed by the following:

$$\text{Minimize } F = F_c + PF_1 \times \sum_{i=1}^{\text{ol}} \left(P_{ij} - P_{ij}^{\max}\right) + PF_2 \times \sum_{j=1}^{\text{VBL}} \left((\Delta V_j)_j^2 + P_{ij}^{\max}\right) + PF_3 \times (\Delta P_g)^2 \tag{19}$$

Where

$$\Delta V_j = \begin{cases} \left(V_j^{\min} - V_j\right); & \text{if } V_j \le V_j^{\min} \\ \left(V_j - V_j^{\max}\right); & \text{if } V_j \le V_j^{\max} \end{cases} \tag{20}$$

$$\Delta P_g = \begin{cases} \left(P_g^{\min} - P_g\right); & \text{if } P_g \le P_g^{\min} \\ \left(P_g - P_g^{\max}\right); & \text{if } P_g \le P_g^{\max} \end{cases} \quad (21)$$

Here, F is the fitness function, ol represent the line overloading, VBL represents the line limit and PF_i $(i = 1, 2, 3)$ represents the penalty factor taken as 10,000 throughout the simulation process.[10]

Computational Procedure for MBMMOA Algorithm for CM

The following procedure applied to CM problem related to the above discussion.

Step 1. *Initialization*: The parameters of the algorithm initialized as in Table A2. The initial populations of N number of individuals randomly chosen between the lower and upper limits of generation. Each of these solutions corresponds to distinct generation rescheduling.

Step 2. *Fitness function evaluation*: The fitness of the population is evaluated using Eq. (1) and checked for adherence to constraint Eqs. (2) to (10).

Step 3. *Interaction distance calculation*: The interaction distances between magnetosomes of cells calculated using Eq. (12).

Step 4. *Moment generation*: Based on the interaction distance, the interaction energy calculated by using Eq. (13). After obtaining energy, the moments calculated using Eq. (14).

Step 5. *Moment regulation*: The moment of the cell regulated based on their fitness. A random number compared with mp and based on the outcome either Eq. (16) carried out or Eq. (17) carried out.

Step 6. *MTS Moments replacement*: The worst populations replaced with newer solutions based on Eq. (18).

4. NUMERICAL TEST RESULTS AND DISCUSSION

The performance of the MBMMOA evaluated in modified small IEEE-30, 57 bus Network Topologies and large IEEE-118 bus Network Topology.[16, 20] The Generating Companies submitted price bids and parameters of the proposed algorithm shown in appendix part of Tables A1 and A2. The total congestion cost and real-power losses have been evaluated to the all the test Classes of the Network Topology. However, the results of the MBMMOA compared to other popular current optimization algorithms such as SA, RSM, PSO, FA, EP, RCGA, PSO, HPSO and DE.

The details of the simulated Classes with different Network Topologies of the present work shown in Table I. The Transmission line overloading created by the any one or combination of (i) reducing the standard line

Table I. Details of the test classes with adopted network topologies.

Network topology	Test class	System contingency
Modified IEEE-30	1A	Line outage between 1–2
	1B	Line outage between 1–7 with load increment at all buses by 50%
Modified IEEE 57	2A	Reduced the line capacity between 5–6 and 6–12 from 200 to 175 MW and from 50 to 35 MW
	2B	Reduced the line capacity between 2–3 from 85 to 20 MW
IEEE 118	3	Line outage between 8–5 with increased load at the bus 11–20 by 57%

Table II. Details of the congested line flow of the IEEE-bus network topologies with different classes.

Test classes	Over flow lines	Line flow (MW) Before CM	Line flow (MW) After CM	Particular of the line limit, (MW)
1A	1–7	147.57	130	130
	7–8	140.23	123.55	130
1B	1–2	314.01	130	130
	2–8	97.86	62.46	65
	2–9	103.66	64.95	65
2A	5–6	188.69	171.57	175
	6–12	49.53	16.76	35
2B	2–3	36.60	17.92	20
3	16–17	209.24	97.71	175
	30–17	580.29	497.66	500
	8–30	363.52	143.68	175

The details of the simulated Classes with different Network Topologies of the present work shown in Table I. The Transmission line overloading created by the any one or combination of (i) reducing the standard line

Table III. Simulation results with a comparison of different algorithms for class-1A.

Parameters	SA[16]	RSM[16]	PSO[16]	FA[21]	MBMMOA [Proposed]
TC, $/h	719.86	716.25	538.95	511.8737	446.83
ΔPG1, MW	−9.076	−8.808	−8.61	−8.7783	−8.59617
ΔPG2, MW	3.133	2.647	10.4	15.0008	6.34873
ΔPG3, MW	3.234	2.953	3.03	0.1068	2.37012
ΔPG4, MW	2.968	3.063	0.02	0.0653	0.0525
ΔPG5, MW	2.954	2.913	0.85	0.1734	0.11077
ΔPG6, MW	2.443	2.952	−0.01	−0.618	1.27425
TG, MW	23.809	23.33	22.93	24.7425	18.75048

Notes: TG—Total generation rescheduling; TC—Total congestion cost.

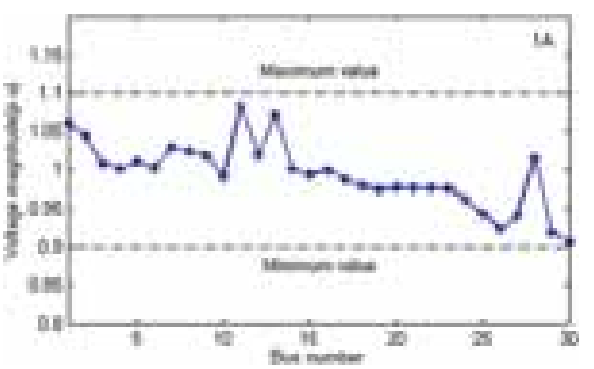


Fig. 2. Voltage profile at different buses for modified IEEE-30 topology with class-1A.

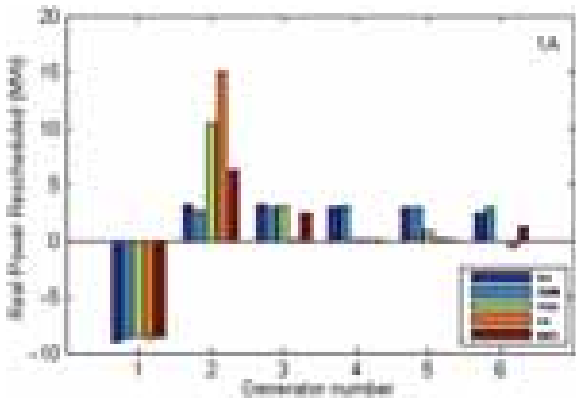


Fig. 3. Comparison of generators real-power rescheduling for modified IEEE-30 topology with class-1A.

limit, (ii) increase the line loading, (iii) outage of the line for contingency purpose. The particulars of line flow at the congested lines before and after CM shown in Table II.

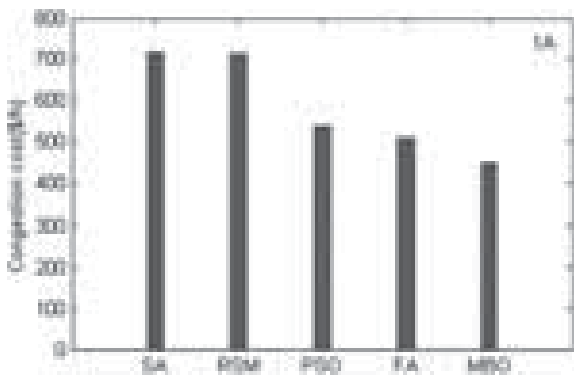


Fig. 4. Cost obtained by different algorithms for modified IEEE 30 topology with class-1A.

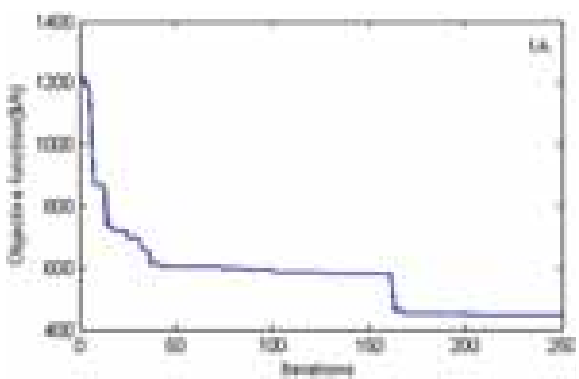


Fig. 5. Convergence profile for modified IEEE-30 bus topology with class-1A.

Table IV. Comparison of simulation results IEEE 30-bus network topology with class-1B.

Parameters	SA[16]	RSM[16]	PSO[16]	FA[21]	MBMMOA [Proposed]
TC, $/h	6068.7	5988	5335.5	5304.4	5234.64
ΔPG1, MW	NL	NL	NL	−8.5798	−9.00284
ΔPG2, MW	NL	NL	NL	75.9954	63.07442
ΔPG3, MW	NL	NL	NL	0.0575	37.92637
ΔPG4, MW	NL	NL	NL	42.9944	5.05114
ΔPG5, MW	NL	NL	NL	23.8325	26.22973
ΔPG6, MW	NL	NL	NL	16.5144	19.75705
TG, MW	164.53	164.5	168	167.974	161.04156

Notes: TG—Total generation rescheduling; TC—total congestion cost; NL—not given in literature.

4.1. Modified IEEE-30 Bus Network Topology

In this bus system, contains six generator buses, forty-one transmission lines, and twenty-four load buses. The total amount of active is 283.4 MW and reactive power is 126.2 MVAR. Contingencies manually created in Network Topology by making line outage and increased the load. The details of Class 1A and Class 1B shown in Table I.

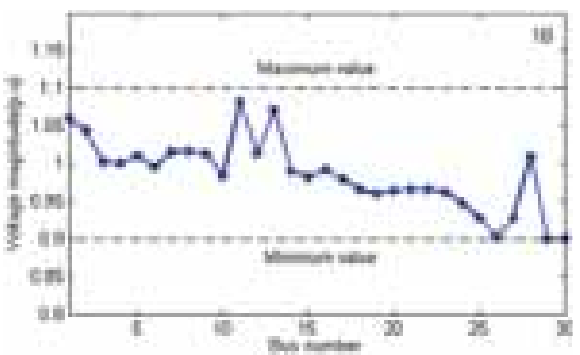


Fig. 6. Voltage profile at different buses for modified IEEE-30 bus topology with class-1B.

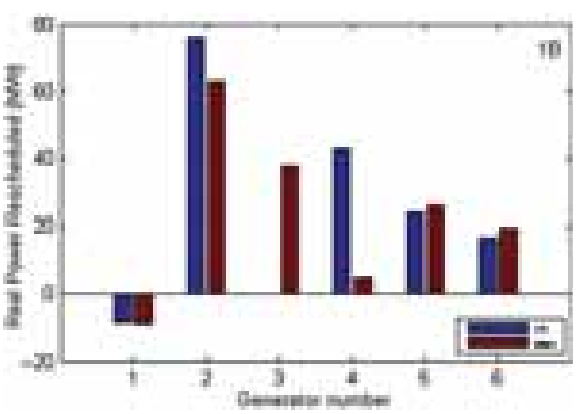


Fig. 7. Comparison of generators real-power rescheduling for modified IEEE-30 topology with class-1B.

RESEARCH ARTICLE

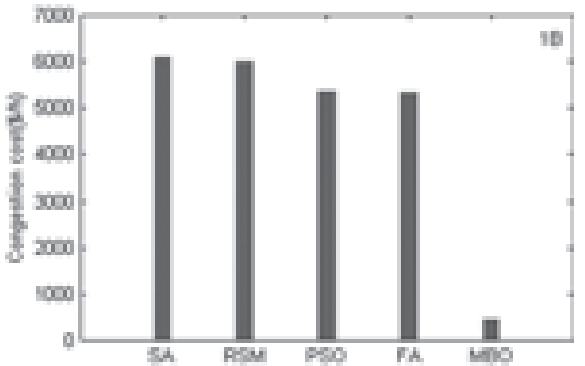


Fig. 8. Cost obtained by different algorithms for modified IEEE 30 topology with class-1B.

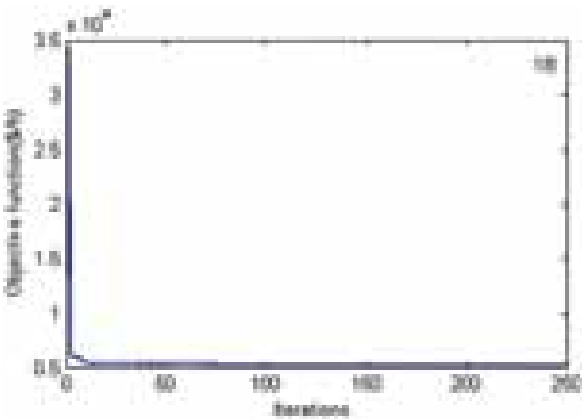


Fig. 9. Convergence profile for modified IEEE-30 bus network topology with class-1B.

In Class-1A, congestion manually formed by disconnect between lines of the buses 1–2. Due to line disconnection, the congestion occurs in lines among the buses 1–7 and 7–8. However, in those line flows have 147.57 MW and 140.23 MW instead of 130 MW before the CM shows in Table II. Hence, line overloading issue overcomes through the generators active-power rescheduling

Table V. Simulation results with comparison of different algorithms for class-2A.

Parameters	SA[16]	RSM[16]	PSO[16]	FA[21]	MBMMOA [Proposed]
TC, $/h	7114.3	7967.1	6951.9	6050.1	5637.04
ΔPG1, MW	74.499	59.268	23.13	5.6351	−0.05437
ΔPG2, MW	0	0	12.44	2.523	−11.72790
ΔPG3, MW	−1.515	37.452	7.49	0.5098	−5.81154
ΔPG4, MW	9.952	−47.39	−5.38	0.107	−57.78428
ΔPG5, MW	−85.92	−52.12	−81.21	−39.1514	−35.83866
ΔPG6, MW	0	0	0	−35.1122	−13.08276
ΔPG7, MW	0	0	39.03	62.1938	−26.02349
TG, MW	168.78	196.23	171.87	145.227	144.015

Notes: TG—Total generation rescheduling; TC—Total congestion cost.

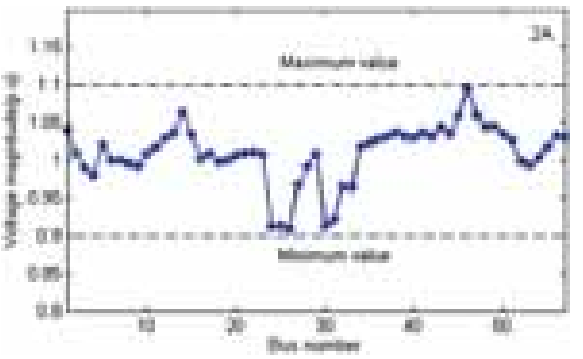


Fig. 10. Voltage profile at different buses for modified IEEE-30 bus topology with class-2A.

scheme. The obtained results by the proposed MBMMOA approach compared with various other optimization methods outlined in the literature shown in Table III. The proposed technique provides minimum congestion cost is 446.83 $/h without line overloading. Before the CM, the loss of total system is 16.132 MW and the same after CM

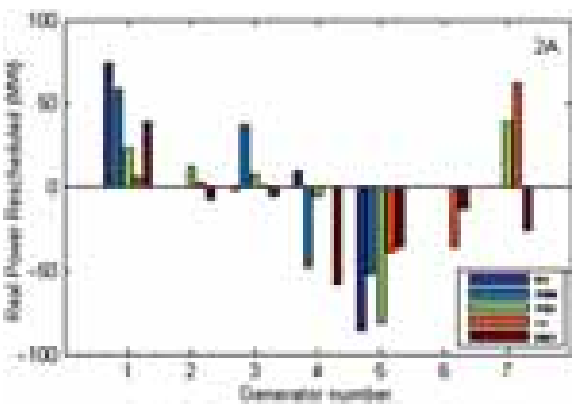


Fig. 11. Comparison of generators real-power rescheduling for modified IEEE 57-bus topology with class-2A.

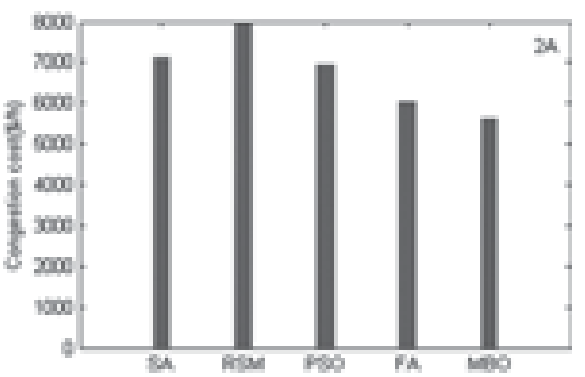


Fig. 12. Cost obtained by different algorithms for modified IEEE 57 topology with class-2A.

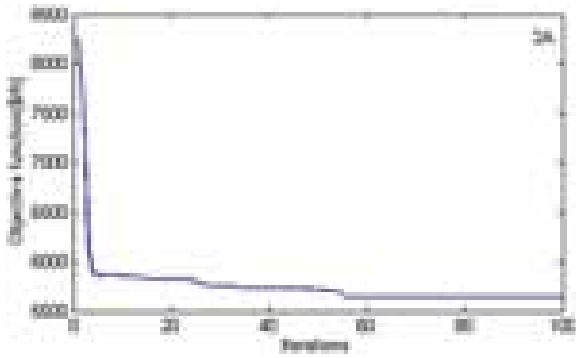


Fig. 13. Convergence profile for modified IEEE-57 bus topology with class-2A.

minimized to 12.713 MW. The voltage magnitude at different busses after CM is within the acceptable limit as shown in Figure 2. A comparative graphic illustration of active-power rescheduled by the PSO, RSM, SA and FA shown in Figure 3. Figure 4 gives the congestion cost. The convergence profile for the Class 1A is shown in Figure 5.

In Class 1B, the overloading in the line created by line outage between the buses 1–7 and load increased at all the

Table VI. Simulation results with comparison of different algorithms for class-2B.

Variables	SA[16]	RSM[16]	PSO[16]	FA[21]	MBMMOA [Proposed]
TC, $/h	4072.9	3717.9	3117.6	2618.1	2045.9
ΔPG1, MW	NL	NL	NL	0.3704	3.41323
ΔPG2, MW	NL	NL	NL	−27.5084	0.01204
ΔPG3, MW	NL	NL	NL	31.6294	13.75648
ΔPG4, MW	NL	NL	NL	0.3308	−0.19024
ΔPG5, MW	NL	NL	NL	−2.2549	−6.4417
ΔPG6, MW	NL	NL	NL	−1.9354	−3.18413
ΔPG7, MW	NL	NL	NL	−0.5101	21.17732
TG, MW	97.88	89.32	76.314	64.5393	48.17514

Notes: TC—Total congestion cost; TG—Total generation rescheduling; NL—Not given in literature.

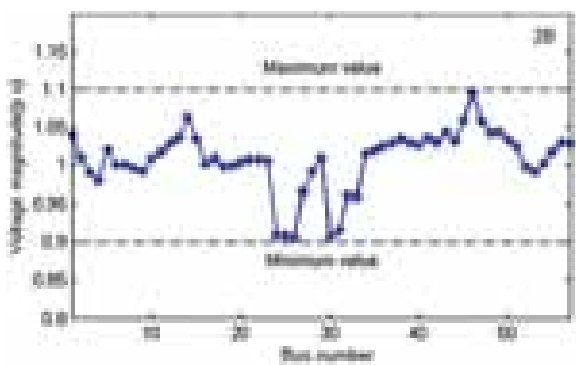


Fig. 14. Voltage profile at different buses for modified IEEE-30 bus topology with class-2B.

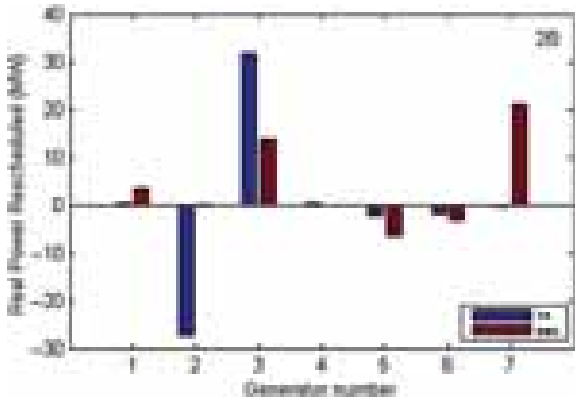


Fig. 15. Comparison of generators real-power rescheduling for modified IEEE 57-bus topology with class-2B.

buses by 50%. This considered contingency causes congestion in the lines among buses 1–2, 2–8 and 2–9 with power flow of 314.01 MW, 97.86 MW and 103.66 MW respectively, which are beyond the limits of their maximum power flow as shown in Table II. The total violation of the power is 255.53 MW owing to line congestion. The CM problem overcomes by the generator rescheduling

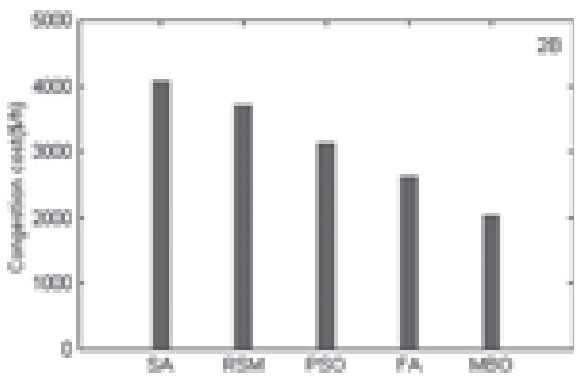


Fig. 16. Cost obtained by different algorithms for modified IEEE 57 topology with class 2B.

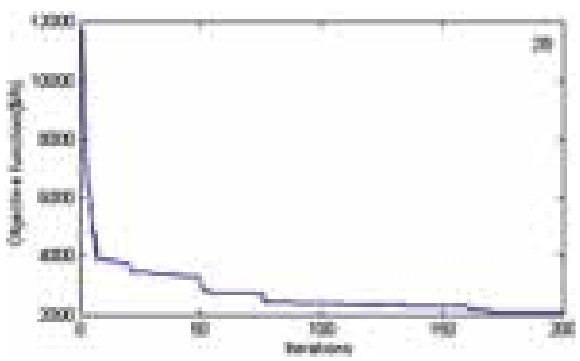


Fig. 17. Convergence profile for modified IEEE-57 bus topology with class-2B.

through the MBMMOA. The retrieved outcome compared with other literature[10, 22] while adopting SA, RSM, PSO, FA as presented in Table IV. The total congestion cost for CM is perceptibly less through the proposed MBMMOA approach compared with other approaches. Moreover, loss of the total system minimized to 14.48 MW instead of the 37.245 MW. After CM the voltage magnitude of at all the busses within the acceptable limit as illustrated in Figure 6. The change in active power of the generators and the comparative cost as offered by proposed approach,

Table VII. Simulation results with comparison of different algorithms for class-3.

Parameters	EP[20]	RCGA[20]	PSO[20]	HPSO[20]	DE[20]	MBMMOA [Proposed]
TC $/hr	42,886	17,693	17,742	17,365	16,080	12688.58
ΔPG1, MW	38.11	26.274	24.502	11.59	6.4705	6.46203
ΔPG2, MW	0	0	0	0	0	3.50642
ΔPG3, MW	63.06	31.422	1.0259	6.568	12.295	4.80864
ΔPG4, MW	−7.48	11.349	1.3148	−0.046	−0.056	11.29861
ΔPG5, MW	−200.6	−219.3	−21.82	−207.7	−208.1	−204.63063
ΔPG6, MW	48.58	101.42	105.32	80.126	105.32	8.59674
ΔPG7, MW	0	0	0	0	0	22.22707
ΔPG8, MW	49.17	0.0029	16.815	6.0059	1.9946	0.73921
ΔPG9, MW	64.056	8.0582	23.665	−0.239	1.2884	6.74496
ΔPG10, MW	0	0	0	0	0	2.34232
ΔPG11, MW	121.51	114.69	121.51	121.51	49.697	−0.52296
ΔPG12, MW	25.918	116.31	116.31	108.06	113.29	1.06161
ΔPG13, MW	85.37	7.4155	3.0092	10.752	0.8125	0.62038
ΔPG14, MW	32.857	−0.0019	0.0177	−0.536	2.1836	4.27972
ΔPG15, MW	31.616	28.095	8.8665	11.254	89.669	9.90095
ΔPG16, MW	39.071	1.7508	7.1475	0.0731	0.7819	0.27315
ΔPG17, MW	37.552	0.0004	0.4916	−0.394	−0.041	0.53489
ΔPG18, MW	0	0	0	0	0	1.4922
ΔPG19, MW	0	0	0	0	0	0.41466
ΔPG20, MW	89.001	1.7478	6.0792	4.3096	−0.089	−0.34584
ΔPG21, MW	96.075	18.528	19.359	77.661	102.85	−0.98802
ΔPG22, MW	93.828	33.426	36.798	72.901	7.3786	0.562
ΔPG23, MW	0	0	0	0	0	0.63867
ΔPG24, MW	0	0	0	0	0	0.26978
ΔPG25, MW	0	0	0	0	0	1.49653
ΔPG26, MW	0	0	0	0	0	−5.90213
ΔPG27, MW	0	0	0	0	0	0.36978
ΔPG28, MW	0	0	0	0	0	−118.19786
ΔPG29, MW	0	0	0	0	0	−6.55699
ΔPG30, MW	−450.9	0	10.702	0.6525	4.904	−4.91548
ΔPG31, MW	0	0	0	0	0	1.0129
ΔPG32, MW	0	0	0	0	0	0.4424
ΔPG33, MW	0	0	0	0	0	0.19373
ΔPG34, MW	0	0	0	0	0	0.87663
ΔPG35, MW	0	0	0	0	0	0.13969
ΔPG36, MW	0	0	0	0	0	0.18389
ΔPG37, MW	0	0	0	0	0	−3.84998
ΔPG38, MW	0	0	0	0	0	0.57649
ΔPG39, MW	0	0	0	0	0	−0.68255
ΔPG40, MW	0	0	0	0	0	−76.44461
ΔPG41, MW	0	0	0	0	0	0.20388
ΔPG42, MW	0	0	0	0	0	0.28521
ΔPG43, MW	0	0	0	0	0	1.2264
ΔPG44, MW	0	0	0	0	0	0.12652
ΔPG45, MW	0	0	0	0	0	−1.45549
ΔPG46, MW	0	0	0	0	0	−3.78854
ΔPG47, MW	0	0	0	0	0	2.29267
ΔPG48, MW	0	0	0	0	0	1.24258
ΔPG49, MW	0	0	0	0	0	0.16326
ΔPG50, MW	0	0	0	0	0	0.19049
ΔPG51, MW	0	0	0	0	0	0.02529
ΔPG52, MW	0	0	0	0	0	0.44904
ΔPG53, MW	0	0	0	0	0	0.73394
ΔPG54, MW	0	0	0	0	0	0.97093
TG	1574.9	719.8998	717.7839	720.4816	707.3281	528.25734

Notes: TC—Total congestion cost; TG—Total generator real power rescheduling.

given in Figures 7 and 8. The Graphical representation of the convergence profile with the number of iterations for Class 1B shown in Figure 9.

4.2. Modified IEEE-57 Bus Network Topology

In this network topology contains 7 generators, 50 loads, and 80 interconnected lines.[16] The total active is 1250.8 MW and reactive power is 336 MVAR. To validate the proposed approach with Class 2A and Class 2B shown in Table I.

In class-2A, the line limit fixed at 175 MW between the line of buses 5–6 and 35 MW for the line between the buses 6–12 instead, actual limits of the line are 200 MW and 50 MW respectively to generating the contingency in the lines shown in Table I. Due to this line flow exceeds their actual line limit between the lines of the buses 5–6 and 6–12, the total violated power becomes a 28.22 MW. The projected MBMMOA effectively solve the line power flow issue with generator real power rescheduling. The obtained results compared with SA, RSM, PSO, and FA as showed in Table V. Figure 10 shows the voltage profile of at all buses within an acceptable limit. A graphical view comparison of the generator active power rescheduling and cost given by the SA, RSM, PSO, and FA illustrated in Figures 11 and 12. From the Table V and Figure 12 shows the total congestion cost offered through the proposed MBMMOA, is 5637.04 $/h only. In addition, the total system loss reduced from 69.647 MW to 25.517 MW. Figure 13 gives the profile convergence of proposed approach.

In class-2B, the line limit reduced from 85 MW to 20 MW for creating contingency in the line between bus 2 and 3 shown in Table II. The details of the line flow data before and after CM shown in Table II. Details of the arrived outcome by MBMMOA compared with the methods listed in Table VI. It clearly has shown the cost offering only 2045.9 $/h for the introduced approach is lowest among all other methods.[10] Figure 14 expose the acceptable voltage profile at all the buses. The comparative change in real-power and costs offered by algorithms are

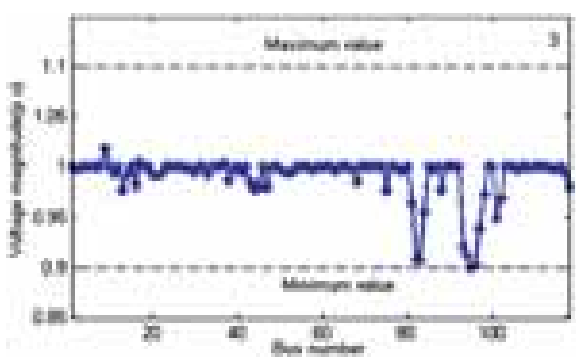


Fig. 18. Voltage profile at different buses for IEEE-118 bus topology with class-3.

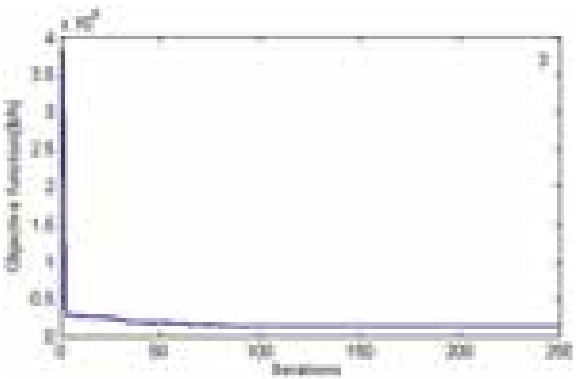


Fig. 19. Convergence profile for modified IEEE-57 bus topology with class-3.

displayed in Figures 15 and 16. After CM, the system loss comes down to 28.128 MW instead of 78.23 MW at the time of congestion. Figure 17 shows the fitness function of the convergence.

4.3. IEEE 118 Bus Topology with class-3

In this system, having the fifty-four generator buses, sixty-four load buses, and one-eighty-six transmission lines.

Class-3 outage of the line between 8–5 at the same timeline between the buses 11–20 increase in load at 1.57 times for creating contingency shown in Table I. The particulars of congested line flow with their limits given in Table II. Details of the total cost of congestion and total change of real power rescheduling obtained through the MBMMOA algorithm compared with various recent algorithms shown in Table VII. Hence, it shows that line overloading issue completely solved through the proposed algorithm with total generator real power rescheduling. The total congestion cost, when it compared with other different algorithms reported in Ref. [23] RCGA, DE, HPSO, PSO, and EP is the lowest price as 12688.58 $/h. Before the CM, the loss of the total system is 277.301 MW while the same is after CM minimized to 230.193 MW. Figures 18 and 19 illustrates that voltage and convergence profile. By evaluation of this Network Topology through the MBMMOA gives the best solution for the OPF.

5. CONCLUSION

In this work, exhibits a new technique for the solution of the CM issue. MBMMOA effectively enforced to reduce the cost of the generator rescheduling as well as eliminated the congestion problem. The contingencies created manually in transmission network by line outage, sudden load variation and limiting standard line limits considered in this study. The introduced technique applied at the modified IEEE-30, 57 and IEEE-118 bus topology. The attained test results compared with SA, RSM, PSO, FA, EP, RCGA, PSO, HPSO and DE. It analyzed that the proposed MBMMOA effectively relieves line overloading,

Congestion cost and reduces real power losses. The cost of rescheduling is much lower than the other approaches given in the literature. From all the considered simulated Class, it analysed that MBMMOA is a potential tool to resolve optimization issues, providing most economical, reliable and secure working conditions. MBMMOA suggested as an effective optimization tool for some other engineering applications.

APPENDIX

Table A1. Price bids are submitted by the generation companies for modified IEEE 30, IEEE 57-bus and IEEE-118 bus test topology.[19]

Bus number	Increment (a_k), \$/MWh	Decrement (b_k), \$/MWh	Bus number	Increment (a_k), \$/MWh	Decrement (b_k) \$/MWh
Modified IEEE-30 network topology					
1	22	18	4	43	37
2	21	19	5	43	35
3	42	38	6	41	39
Modified IEEE-57 network topology					
1	44	41	5	42	39
2	43	39	6	44	40
3	42	38	7	44	41
4	43	37			
IEEE-118 network topology					
1	40	38	65	25	17
4	43	35	66	26	18
6	41	38	69	28	15
8	44	39	70	43	39
10	22	17	72	47	38
12	23	18	73	44	36
15	45	35	74	43	39
18	41	38	76	44	36
19	44	36	77	47	38
24	45	35	80	23	15
25	27	18	85	43	39
26	23	18	87	25	17
27	41	39	89	44	36
31	45	35	90	42	38
32	24	18	91	41	38
34	43	37	92	44	36
36	44	36	99	43	37
40	43	35	100	42	36
42	47	38	103	24	17
46	45	35	104	44	38
49	25	18	105	43	39
54	32	30	107	42	38
55	42	36	110	43	37
56	43	37	111	22	18
59	23	17	112	42	38
61	26	16	113	47	33
62	44	36	116	45	35

Table A2. Details of MBOMMA parameters taken for optimization.[19]

Algorithm	Parameters and values
MBOMMA	Population size = 40; $c_1 = 20$; $c_2 = 0.003$; mp = 0.6; $B = 1$;

References

1. O. Alsac and B. Stott, *IEEE Trans. Power Appar. Syst.* PAS-93, 745 **(1974)**.
2. O. Alsac and B. Stott, *IEEE Trans. Power Appar. Syst.* PAS-93, 745 **(1974)**.
3. R. C. Burchett, H. H. Happ, and D. R. Vierath, *IEEE Trans. Power Appar. Syst.* PAS-103, 3267 **(1984)**.
4. Kalyanmoy Deb, Optimization for Engineering Design: Algorithms and Examples, PHI **(2009)**.
5. M. Todorovski and D. Rajicic, *IEEE Trans. Power Syst.* 21, 480 **(2006)**.
6. M. A. Abido, *Electr. Power Components Syst.* 30, 469 **(2002)**.
7. A. A. Abou El Ela, M. A. Abido, and S. R. Spea, *Electr. Power Syst. Res.* 80, 878 **(2010)**.
8. S. Sayah and K. Zehar, *Energy Convers. Manag.* 49, 3036 **(2008)**.
9. R. Rajabioun, *Appl. Soft Comput. J.* 11, 5508 **(2011)**.
10. C. Sumpavakup, I. Srikun, and S. Chusanapiputt, A solution to the optimal power flow using artificial bee colony algorithm, *2010 International Conference on Power System Technology* **(2010)**, pp. 1–5.
11. N. Sinsuphan, U. Leeton, and T. Kulworawanichpong, *Appl. Soft Comput. J.* 13, 2364 **(2013)**.
12. M. Ghasemi, S. Ghavidel, M. Gitizadeh, and E. Akbari, *Int. J. Electr. Power Energy Syst.* 65, 375 **(2015)**.
13. V. Mukherjee and S. Verma, *IET Gener. Transm. Distrib.* 10, 2548 **(2016)**.
14. J. R. Cintam, D. Mary, P. Thanigaimani, and P. Salomipuspharaj, *Int. J. Appl. Eng. Res.* 10 **(2015)**.
15. C. Venkaiah and D. M. Vinod Kumar, *Appl. Soft Comput. J.* 11, 4921 **(2011)**.
16. N. K. Sujatha Balaraman, *Journal of Electrical Systems* 7, 54 **(2011)**.
17. S. Balaraman, Congestion management in deregulated power system using real coded genetic algorithm. 2, 6681 **(2010)**.
18. H. Mo, L. Liu, and M. Geng, A new magnetotactic bacteria optimization algorithm, *2014 Int. Conf. Swarm Intell.* **(2014)**, Vol. 14, pp. 103–114.
19. J. R. Chintam and M. Daniel, *Energies* 11, 1 **(2018)**.
20. Sujatha Balaraman, Applications of Evolutionary Algorithms and Neural Network for Congestion Management in Power Systems, Anna University, India **(2011)**.
21. S. Verma and V. Mukherjee, *Eng. Sci. Technol. an Int. J.* 19, 1254 **(2016)**.

Received: 12 December 2017. Accepted: 1 January 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1471–1479, 2018

RESEARCH ARTICLE

Kriging and Latin Hypercube Sampling Assisted Simulation Optimization in Optimal Design of PID Controller for Speed Control of DC Motor

Amir Parnianifard[1,*], A. S. Azfanizam[1], M. K. A. Ariffin[1], M. I. S. Ismail[1], Mohammad Reza Maghami[2], and Chandima Gomes[2]

[1] *Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[2] *Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*

Investigating less computationally expensive methods for achieving optimal tuning of proportional, integral, and derivative gains in PID controller also has become a main challenging topic in the world of control systems. This paper aims to propose a new less-expensive method to the optimal design of PID controller. For this purpose, Kriging metamodel is used with Latin Hypercube Sampling (LHS) as a common experimental design method in the class of space filling design. Kriging can interpolate over whole design space and assisted to investigate global optimum point. A numerical case in the tuning of PID controller for linear speed control of DC motor is served to show the applicability and superiority of proposed method compared to two existing methods such as traditional Zeigler-Nichols method and Taguchi-Grey Relational Analysis (Taguchi-GRA).

Keywords: Kriging, PID Controller, DC Motor, Simulation Optimization, Metamodel, Computer Experiments, Latin Hypercube Sampling.

1. INTRODUCTION

Nowadays, one of the main challenges in engineering design and control system is investigating optimal design for proportional, integral, and derivative gain parameters in PID controller. There are different ways of optimizing control systems via a mathematical model, simulation model, or metamodel. Notably, choosing the best way depends on the amount of complexity in the system, see Figure 1. The main goals of simulation can be defined as first what-if study of model or sensitivity analysis and second optimization and validation of model.[1] The essential benefit of simulation is its ability to cover complex processes, either deterministic or random while eliminating mathematical sophistication.[2] Simulation optimization techniques are classified into model-based optimization methods and metamodel-based optimization methods.[3] In model-based, the simulation running is not expensive and model output can be used directly in optimization. Model-based investigate the optimal point by running original simulation model, while metamodel uses the approximation of the simulation model. Many large scales and detailed simulation models in a complex system, may be expensive to run in terms of time-consuming, computational cost, and resources.

This paper focuses on the optimal designing of PID controller for linear speed control of DC motor. Different model-based optimization methods have been proposed in literature in tuning of PID controller for speed control of DC motor such as metaheuristic methods,[4] Genetic Algorithms (GAs),[5] Particle Swarm Optimization (PSO)[6] and Taguchi method.[7] One of the main difficulty which arises from existing methods for tuning of PID controller is computational complexity that often is time-consuming to find optimal design of the controller.[8]

If the complex system not costly worth it to be derived physically for tuning PID controller, computer experiments through simulation model can be performed to extract approximation model instead of the real model. In such a case metamodel-based optimization methods (also called surrogate-based techniques) can be appropriately used to reduce the computational complexity in designing of PID controller.[9] Several main advantages have been mentioned by Ref. [10] for applying metamodel in optimization. There are (1) improving the efficiency of optimization,

*Author to whom correspondence should be addressed.

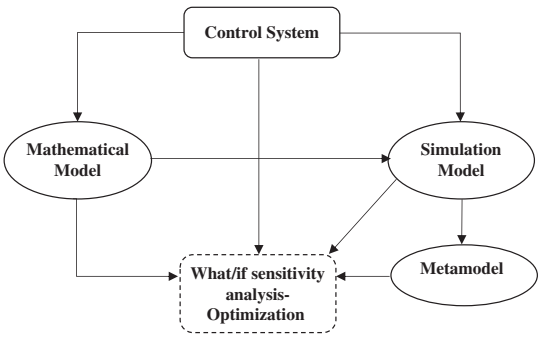


Fig. 1. Different ways for optimizing a control system based on mathematical, simulation, or metamodel.

(2) supporting parallel computation, (3) cover sensitivity analysis of input variables and gain better insights from the problem, and (4) handle both types of variables included discrete and continuous.

Yet, metamodel-based optimization methods have not been attended enough in control of DC motors. For more instance, for two common metamodels including polynomial regression and Kriging, we investigated relevant studies by an advanced search among metadata in IEEE Xplore and SCOPUS database as two common databases in electrical and control engineering. The search results have shown just 7 documents with keywords ("DC motor" AND "Kriging") and there was no result with keywords ("DC motor" AND "Kriging" AND "PID controller") and ("PID controller" AND "DC motor "AND "polynomial regression"). In Ref. [11] classical metamodel as response surface methodology have used to obtain better performance for rotor pole shape. In Refs. [12 and 13] Kriging have applied to design torque control of brushless DC motor used in electric vehicles. Kriging is a popular method for metamodeling with simulation data,[14] that can be employed in different control system problems. The main advantages of Kriging over some classical methods like different orders of polynomial regression is exact interpolating of Kriging, i.e., the predicted values and simulated output in the set of observed input values are exactly equal.[15] However, in the current research, the Kriging metamodel is used for global interpolation of sample points to obtain tuning of PID controller with at least number of simulation running.

The accuracy of the approximated metamodel is strongly depended on designing appropriate sample points. These sample points (also called simulation experiments) can be taken from Design Of Experiments (DOE) techniques (also known as sampling design methods) to achieve maximum information about underlying input/output relationship with at least number of sampling points. Here in order to construct Kriging, the Latin Hypercube Sampling (LHS) in the class of space filling methods is applied as the most popular method for designing of simulation experiments.[16]

The proposed methodology in this paper is contributed as follows. This model reduces the computational complexity in term of a number of iterations in simulation model compared to model-based optimization methods. Other contribution arises due to different performance criterions included Integral-Square-Error (ISE), Integral-Absolute-error (IAE), Integral-Time-Square-Error (ITSE), and Integral-Time-Absolute-Error (ITAE) are compared in the same model. The proposed procedure has multidisciplinary application and can be used generally for tuning of PID controllers in most types of control systems.

The rest of this paper is organized as follows. In Section 2, metamodel particularly Kriging assisted simulation optimization are explained. PID controller for speed control of DC motor is described in Section 3. A numerical case study is provided in Section 4. Finally, this paper is concluded in Section 5.

2. METAMODEL BASED SIMULATION OPTIMIZATION (MBSO)

The process of investigating the best value of input variables from among all possibilities is Simulation Optimization (SO), also known as optimization via simulation or simulation-based optimization. The objective of SO is obtaining the optimum value for output while minimizing the resource spent. In reality, there are many different methods in SO, see Refs. [17, 18]. Nowadays, SO methods based on Design and Analysis of Computer Experiments (DACE) have been commonly interested in displaying the performance of the complex system.[19] The essential benefit of DACE method is its ability to cover complex processes, in order to avoid intensive computational complexity, which might squander time and resource for estimating model's parameters.

Metamodels (also called surrogate models) are the basic tools to a simple approximation of input-output in the simulation model. Metamodeling has utilized variety statistical and mathematical approach for interpreting parameters and their relationship in original models.[20] In metamodeling techniques, the system behavior is evaluated by running a simulation model for a fixed period of time which called computer experiments (also called simulation experiments).[18] In this context, we use Kriging metamodel as most common global approximation model and its well-known experimental designs which called LHS.

2.1. Latin Hypercube Sampling (LHS)

LHS was first introduced by Ref. [21]. It is a strategy to generate random sample points, while guarantee all portions of the design space is depicted. Generally, LHS is intended to develop results in SO.[19] LHS has been commonly defined for designing computer experiments based on the space-filling concept.[22, 23] In general, for n input variables, m sample points are produced randomly into m intervals or scenarios (with equal probability). For the

particular instance the LHS design for $m = 4$, $n = 2$ is shown in Figure 2. The LHS strategy proceeds as follows:
i. In LHS, each input range divided into m subranges (integer) with equal probability magnitudes, and numbered from 1 to m.
ii. In the second step, LHS place all m intervals by random value between lower and upper bounds relevant to each interval, since each integers $1, 2, \ldots, m$ appears exactly once in each row and each column of the design space. Note that, within each cell of design, the exact input value can be sampled by any distribution, e.g., uniform, Gaussian or etc.

Undoubtedly the simplest and most popular computer experiment designs use LHS, see Refs. [16, 24]. In this context for our proposed method, the LHS is performed three times by different usage included (i) training sample points, (ii) validating sample points, (iii) search sample points. Training sample points are run by the original simulation model, and resulted input/output data apply to fit metamodel. Validating sample points are run with both original simulation model and metamodel, and extracted data apply to validate metamodel. Search sample points are run just by metamodel, and extracted data apply for the global search to find optimal point.

2.2. Kriging

Since Daniel G. Krige[25] has addressed the geostatistics in 1951, today Kriging models have been used as a widespread global approximation technique that applied widely in engineering problems.[26, 27] Kriging is an interpolation method which could cover deterministic data, and it is highly flexible due to ability in employing a various range of correlation functions. The higher accuracy of Kriging models than the other alternatives such as response surface modeling are confirmed in the literature.[28] In a Kriging model, a combination of a polynomial model and realization of a stationary point are assumed by the form of:

$$y = f(X) + Z(X) + \varepsilon \tag{1}$$

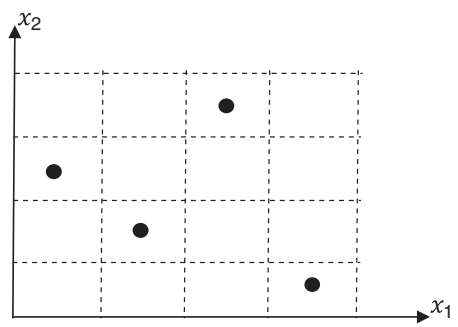


Fig. 2. An example for LHS design with two input factors, and four intervals.

Where in $f(X) = \sum_{j=0}^{k} \hat{\beta}_j f_j(X)$, the polynomial terms of $f_j(X)$ are typically first or second order response surface approach and coefficients $\hat{\beta}_j$ are regression parameters ($j = 0, 1, \ldots, k$). The term ε describes approximation error, and the term $Z(X)$ represents realization of a stochastic process, which most time normally distributed Gaussian random process with zero mean, variance σ^2 and non-zero covariance. The correlation function of $Z(X)$ is defined by:

$$Cov[Z(x_i), Z(x_p)] = \sigma^2 R(x_i, x_p) \tag{2}$$

Where σ^2 is process variance and $R(x_i, x_p)$ is the correlation function, and can be chosen from different functions which proposed in literature. Due to tuning the correlation function with sample data, the Kriging is extremely flexible to capture nonlinear treatment of model. Among literature, some studies have been found which sufficiently describe Kriging metamodel methodology, see Refs. [26, 29].

2.3. MBSO Algorithm

The MBSO algorithm is sketched in some steps as below:
➢ *Step* 1: Using LHS method to design computer experiments (simulation runs). If k is number of input variables, in order to construct Kriging metamodel, the size of training sample points (n_t) compute by $n_t = 10k$.[19] This types of samples points also called training data which used to estimate the metamodel.
➢ *Step* 2: Run simulation model for each input computation and compute the output Total running of simulation is n_t.
➢ *Step* 3: Fit Kriging metamodel over collected outputs. Here, in order to construct Kriging surrogate models, we use the Design and Analysis of Computer Experiments (DACE) MATLAB Kriging toolbox.[30]
➢ *Step* 4: Validate Kriging metamodel that fit in Step 3 by applying standardized residuals error. For this purpose, first design n_v sample points as validate experiments via LHS. Then run simulation and calculate relevant outputs (y_s) for each points ($s = 1, 2, \ldots, n_v$) So, for each input combination compute standardized residuals as below:

$$D_s = \frac{y_s - \hat{y}_s}{\sqrt{1/n_v \sum_{s=1}^{n_v} (y_s - \hat{y}_s)^2}} \tag{3}$$

where $\hat{y}_s$ is predicted output for sth input combination by Kriging that fit over training sample points. While most of the standardized residuals should be in the interval $-3 \leq D_s \leq 3$, and any observation outside of this interval (outlier) is potentially unacceptable respect to its observed simulation output.[31]
➢ *Step* 5: Optimize model by employing Kriging metamodel. In general, one the simplest approach of optimization method to search optimal solution by

RESEARCH ARTICLE

employing surrogate models instead of an expensive original model is the "grid search" method. In this approach, the whole design space is studied through evaluating of model components consist of objectives and constraints set over equally spaced grid points.[32] Also, the grid points can be replaced by random points which spread over whole design space (Monte Carlo random search) or some common space filling methods (e.g., LHS method). In this searching method, there is a chance to find the global optimum solution, due to the searching is not bounded to local valley.

3. PID CONTROLLER FOR SPEED CONTROL OF DC MOTOR

DC motors have been widely employed as actuating sources in industrial application. DC motor systems with their electromechanical components are widely employed in various industrial applications due to their advantages such as easy speed control, position control, and wide adjustability interval. A DC motor is constructed with two sub-systems. The first sub-system is electrical part and consists of armature inductance, armature resistance and the magnetic flux of the stator. Figure 3 shows the simplified electrical circuit of DC motor and the mechanical model of a rotor. The dynamic behavior of DC motor based on the mathematical model is expressed by:

$$J\frac{d\omega(t)}{dt} + B\omega(t) = T_m(t) \tag{4}$$

$$L\frac{di_a(t)}{dt} + Ri_a(t) = v_a(t) - v_b(t) \tag{5}$$

$$v_b = K_e\omega(t) \tag{6}$$

$$T_m(t) = K_t i_a(t) \tag{7}$$

PID controller employs three terms of proportional, integral and derivative to reduce an error signal (e) to be more close to zero. The output of a PID controller, equal to the control input to the plant, in the time-domain is as follows:

$$u(t) = K_p e(t) + K_i \int_0^t e(t)\,dt + K_d\frac{de}{dt} \tag{8}$$

The appropriate function of PID strongly depends on allocated magnitude to each gain of K_p, K_i, and K_d. In order to evaluate the performance of PID controller, four common creations are used. Note that All creations

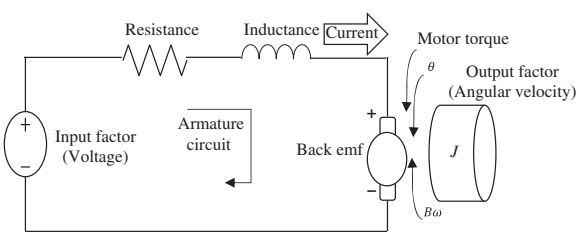


Fig. 3. An electrical and mechanical components of DC motor.

Table I. Physical parameters in the DC motor model.

Parameter	Title	Value	Unit
v_a	Input voltage	–	Voltage
R	Armature resistance	2.45	Ohm
L	Armature inductance	0.035	mH
J	Moment of inertia	0.022	Kgm^2
K_t	Motor torque constant	0.015	Nm/Amp
K_e	Back emf constant (electromotive force constant)	1.2	Vrad/s
B	Motor viscous friction	0.0005	N.m.s
ω	Angular velocity of the shaft	–	Rad/s

are associated with all time-response parameters included rise time, steady state error, settling time, and overshoot.[33] The mathematically features of performance creations in time domain are defined as below:

$$\text{Min } ISE = \int_0^\infty e(t)^2\,dt \tag{9}$$

$$\text{Min } IAE = \int_0^\infty |e(t)|\,dt \tag{10}$$

$$\text{Min } ITSE = \int_0^\infty t\cdot e(t)^2\,dt \tag{11}$$

$$\text{Min } ITAE = \int_0^\infty t\cdot |e(t)|\,dt \tag{12}$$

In above equations $e(t) = (y(t) - S_p)$, where $y(t)$ is response (output) in time t and S_p is desired set point for

Table II. Training sample points and performance outputs.

	Design variables			Performance output			
s	K_p	K_i	K_d	ITAE	ITSE	ISE	IAE
1	1.06	0.92	0.82	858.26	167.04	133.44	264.10
2	1.33	0.82	0.87	684.30	139.21	124.82	241.23
3	1.10	0.81	0.23	613.60	132.54	123.75	233.59
4	1.16	1.18	0.73	759.20	140.66	117.88	241.83
5	1.82	1.09	0.60	417.64	79.15	92.75	183.76
6	1.77	0.88	0.46	393.35	85.20	98.03	187.18
7	1.48	1.13	0.60	547.59	102.08	103.55	208.07
8	1.98	0.94	0.88	416.85	83.61	96.25	187.51
9	1.70	0.99	0.74	486.06	94.11	101.56	199.77
10	1.10	1.10	0.54	732.99	138.81	119.44	240.65
11	1.24	1.02	0.37	590.73	114.85	111.55	219.77
12	1.93	0.96	0.78	415.73	82.53	95.58	186.59
13	1.40	0.93	0.42	523.35	105.75	108.77	210.54
14	1.35	1.19	0.50	579.64	107.53	105.14	212.91
15	1.56	1.08	0.55	497.08	94.25	100.73	200.09
16	1.50	0.90	0.20	425.20	91.24	101.97	194.08
17	1.85	0.98	1.00	490.80	93.90	101.18	199.91
18	1.96	0.83	0.35	316.03	76.47	92.17	174.32
19	1.59	1.15	0.32	432.64	81.83	93.90	186.56
20	1.77	0.89	0.68	442.67	91.98	101.55	195.69
21	1.72	1.00	0.97	533.77	100.66	104.14	207.12
22	1.20	1.04	0.91	792.84	149.00	123.62	249.79
23	1.03	1.05	0.93	935.39	176.63	132.94	270.42
24	1.63	1.06	0.48	452.72	87.11	97.54	192.29
25	1.66	0.85	0.80	507.51	104.89	108.48	209.13
26	1.43	0.95	0.39	504.65	101.55	106.50	206.44
27	1.46	1.14	0.26	465.71	88.61	97.63	193.78
28	1.18	1.12	0.63	712.13	133.56	116.79	236.36
29	1.29	0.86	0.28	534.49	113.17	113.69	216.50
30	1.87	1.17	0.66	419.56	77.24	90.74	182.02

response of model (i.e., here the rotational speed is defined as an output in the DC motor system).

4. NUMERICAL CASE STUDY

In this section in order to show the applicability and flexibility of proposed method, we serve a numerical example to tuning PID controller for linear speed control of DC motor. In Ref. [34] the Grey Relational Analysis (GRA) has used when combined with Taguchi approach on the orthogonal array to investigate optimal gain parameters. Here, we use the same example with Ref. [34] to show the analytical comparison between proposed method, Taguchi-GRA and also classical tuning by Ziegler-Nichols method. So, the relevant load parameters in Eqs. (4)–(7) are defined in Table I.[35] The simulation model is modeled in MATLAB®/Simulink. We define four different response for the model, see Eqs. (9)–(12). In order to compare all performance creations in the same time domain, we evaluate the performance of PID controller with all creations ISE, IAE, ITSE, and ITAE in $0 \leq t \leq 30$ second.

Tuning the model with classic Ziegler-Nichols method gives the PID gain parameters as $K_p = 1.5865$, $K_i = 1.0018$, $K_d \in 0.62811$ (i.e., traditional Ziegler-Nichols

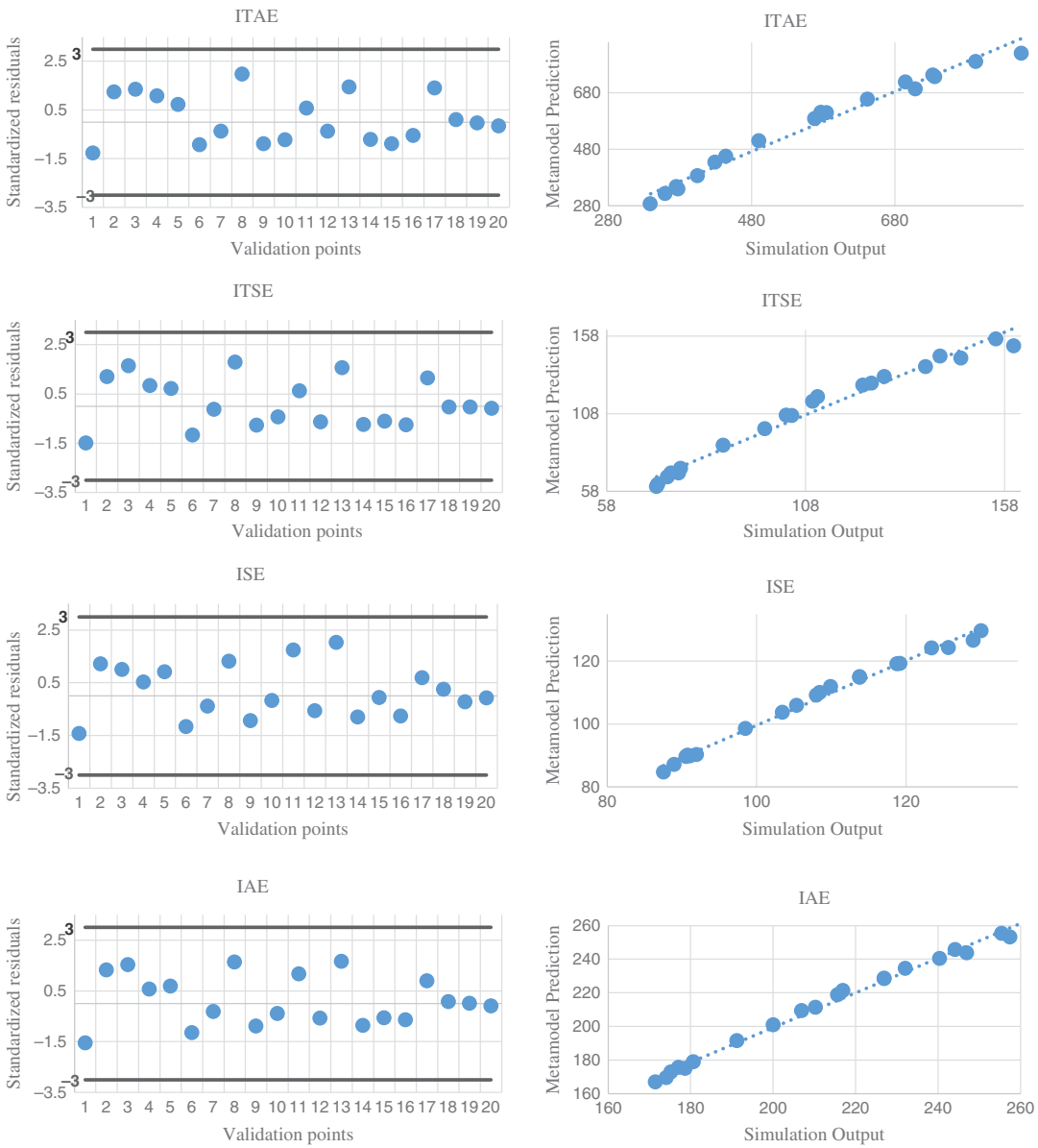


Fig. 4. Standardized residuals and scatterplot for each ITAE, ITSE, ISE, and IAE kriging metamodel.

tuning is done by MATLAB SISO design tool), so according to magnitude for each gain parameters, design ranges for K_p, K_i, and K_d are defined as below:

$$K_p \in [1, 2], \quad K_i \in [0.8, 1.2], \quad K_d \in [0.2, 1]$$

In order to design required simulation experiments, LHS is conducted with three ($k = 3$) input parameters (K_p, K_i, and K_d). As mentioned in Section 2.3, for $k = 3$, in order to construct Kriging metamodel, the size of training

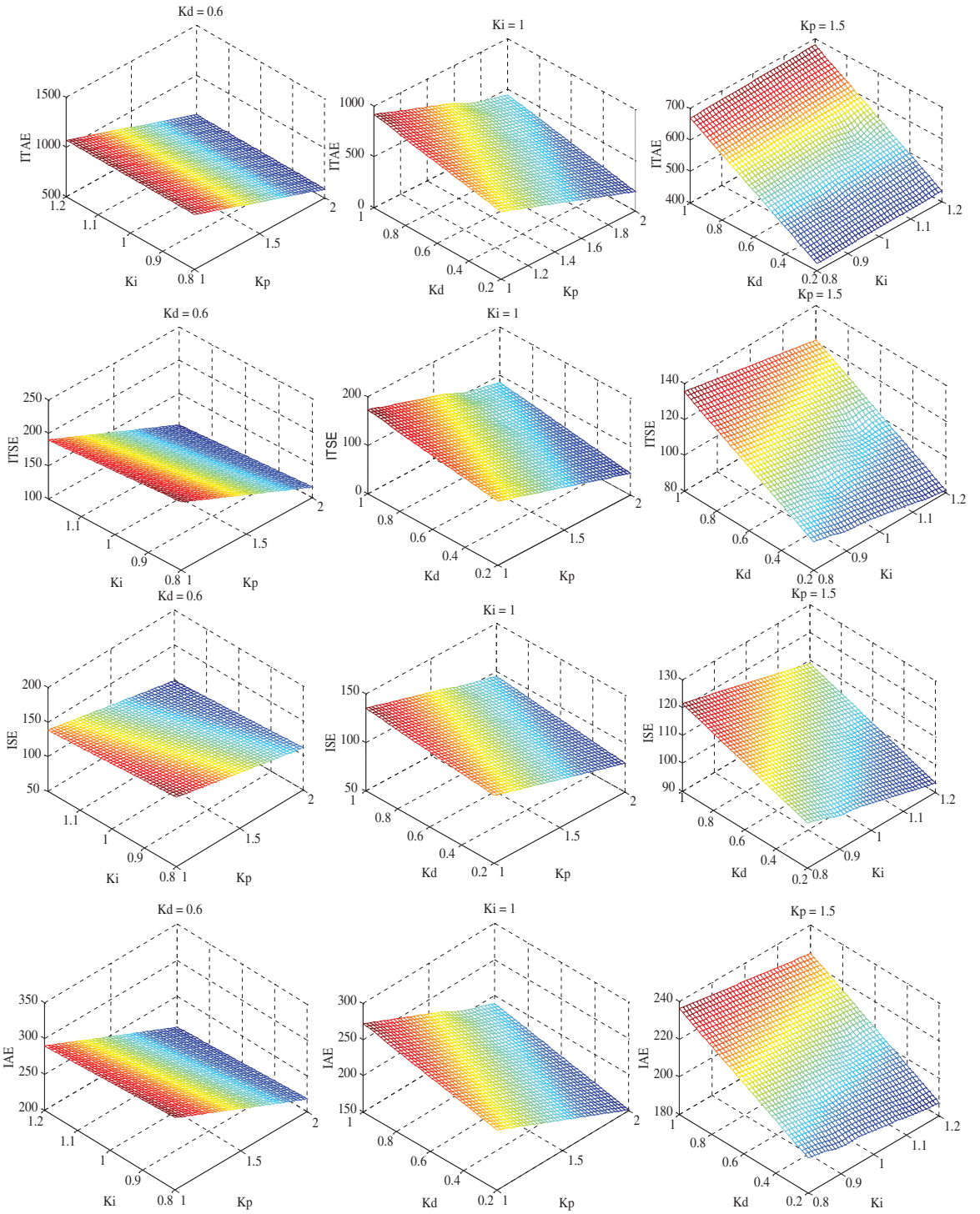


Fig. 5. Surface plot of four kriging metamodels over ITAE, ITSE, ISE, and IAE.

sample points needs to be at least $n_t = 10k = 30$. In the next step, the original simulation model is run for each training sample point and compute the relevant output. Here for the same input combination, four outputs included ITAE, ITSE, ISE, and IAE are computed. Table II depicts the relevant data among training sample points and each performance criterion.

For each criterion of ITAE, ITSE, ISE, and IAE the separate Kriging metamodel is fitted over input/output set of data. In order to validate metamodels, we again apply LHS to design 20 validate sample points over design spaces. So, for each of validating points the original simulation is run and obtain simulation output. Also for the same validate points, the Kriging metamodel is performed to predict the relevant output. Figure 4 illustrates results of standardized residuals validation and scatterplot for predicted output over simulation output. As can be seen from the validation results, we can accept all four metamodels due to all standardized results in each metamodel are being in $[-3, +3]$ and eyeball scaling on scatterplots confirm their validity Figure 5 plots Kriging metamodels based on two gain parameters when the third gain parameter limited in the center of relevant design range.

In the final step, we have implemented the grid search method by producing 10,000 sample points over design space. In order to guarantee global search, the LHS as common space filling design method is used to define search sample points. Among all sample points, the optimum point is investigated when according to each Kriging metamodel, the objective is the minimization of ITAE, ITSE, ISE, and IAE. Here the grid search method has computed the same optimum point for ITSE, ISE, and IAE. System performance for two optimum points which obtain by Kriging over ITAE and Kriging over ITSE, ISE, and IAE are compared by Taguchi-GRA result[34] and tradition Ziegler-Nichols method, see Table III In the current example, the optimization model is constructed by minimizing performance criterion as single-objective of the model (minimization of ITAE, ITSE, ISE, or IAE separately). Results confirmed that for each criterion, in the presented optimal point, the single objective of the model is satisfied, see Figure 6. Figure 7 shows the step responses of DC motor speed with PID gain parameters which gained respectively from traditional Ziegler-Nichols method, Taguchi-GRA, Kriging (ITAE), and Kriging (ITSE, ISE, IAE) method.

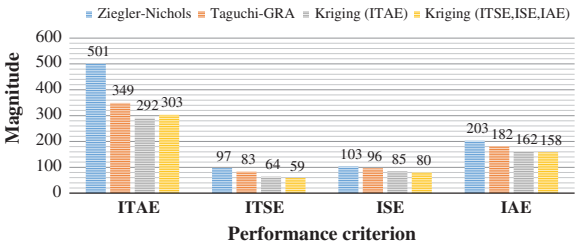


Fig. 6. Comparison of methods by minimization of performance criterion as single-objective.

Investigating less computationally cost methods (less time consuming) for achieving optimal tuning of proportional, integral, and derivative gains in a PID controller have also become an important challenging topic in the world of control systems, that in the current paper is attended as the main goal. Different methods have been proposed in the literature for tuning a PID controller in linear speed control of a DC motor, that most of them are model-based algorithms such as metaheuristic techniques, Genetic Algorithms (GAs), Particle Swarm Optimization (PSO), and Grey Relational Analysis (GRA). A number of iteration in model-based methods strongly depends on problem size and significantly is more than a number of computer experiments in metamodeling.[36] This is cause that in metamodeling, the time for optimizing problem in term of simulation running is much less than model-based algorithms. For instance, in the current paper, in tuning of PID controller for linear speed control of DC motor, just 30 simulation run is conducted to construct Kriging metamodel over obtained input/output data, and 20 simulation run is directed for check validity of metamodel. Results have confirmed the superiority of MBSO with combining Kriging and LHS method compared with traditional Ziegler-Nichols method and Taguchi-GRA method. It is clear from the final results that Taguchi-GRA obtains better results in settling time and overshoot performance, but worse result than Kriging method in rise time and totally in all performance parameters of ITAE, ITSE, ISE, and IAE. In order to reduce computational cost, the MBSO method tries to approximate and analyze a model with a minimum number of simulation experiments.

Note that, metamodel-based optimization methods by employing modern metamodels like Kriging derive global search over whole design space. Most experimental based

Table III. Optimal points derived through different tuning methods.

	Gain parameters			Performance						
Tuning method	K_p	K_i	K_d	ITAE	ITSE	ISE	IAE	Rise time (Sec.)	Settling time (Sec.)	Overshoot (%)
Ziegler-Nichols	1.6	1.0	0.6	501	97	103	203	3.21	10.7	4.82
Taguchi-GRA	2.0	0.8	0.6	349	83	96	182	3.9	6.24	0.767
Kriging (ITAE)	2.0	1.0	0.2	292	64	85	162	3.05	8.32	2.27
Kriging (ITSE, ISE, IAE)	2.0	1.2	0.2	303	59	80	158	2.59	8.9	4.77

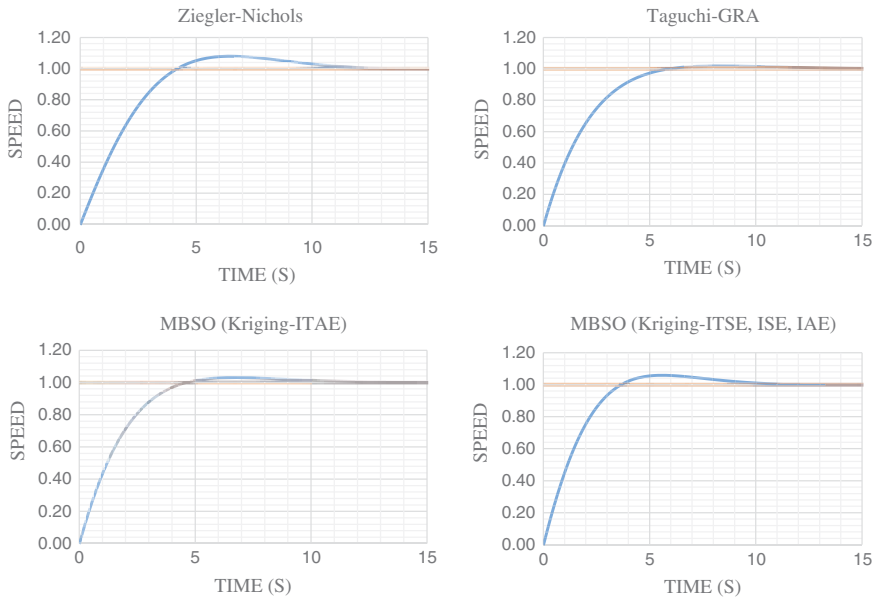


Fig. 7. Step responses with four different PID gain parameters obtain with Ziegler-Nichols, Taguchi-GRA, and MBSO.

methods like Taguchi-GRA search design space over discrete values, and limited its search just over designed sample points (i.e., one sample point with the best output is selected as an optimum point), but Kriging can interpolate over whole design space and not just limited over sample points. Note that, the grid search approach is used by employing Kriging in order to investigate global optimal point in the model.

5. CONCLUSION

The main goal of the current paper is to determine optimal PID gain parameters with at least computational costs. In the complex system, running simulation experiments most times are time-consuming, so design and analyzing simulation model need to be done with a minimum number of simulation experiments. We served an example for tuning of PID controller for linear speed control of DC motor. PID gain parameters included K_p, K_i, and K_d are considered as input variables and four different performance criterions included ITAE, ITSE, ISE, and IAE separately attended as an output of system. Here, we used Kriging metamodel to globally approximate the simulation model. Sample points or experiments are designed by LHS method to achieve filling of whole design space. Grid search method can find the optimal point by global search over whole design space. The evidence from this study confirm the superiority of proposed method compared with traditional methods such as Zeigler-Nichols and Taguchi-GRA. It is recommended that further research be undertaken in the following areas:

- Performing MBSO method with other types of metamodel such as different orders of polynomial regression or Radial Basis Function (RBF).
- Construct system by attending uncertainty in load parameters or under load disturbances.
- Instead of minimizing performance criterion as a single-objective of the model, use simultaneously minimizing settling time, overshoot, and rise time in the multi-objective model.
- Different types of input signals in PID controller such as sinusoidal, ramp, and so on need to be attended with MBSO method.

References

1. W. C. M. van Beers and J. P. C. Kleijnen, *Journal of the Operational Research Society* 54, 255 **(2003)**.
2. G. Figueira and B. Almada-Lobo, *Simulation Modelling Practice and Theory* 46, 118 **(2014)**.
3. F. A. C. Viana, T. W. Simpson, V. Balabanov, and V. Toropov, *AIAA Journal* 52, 670 **(2014)**.
4. M. M. Sabir and J. A. Khan, *Advances in Artificial Neural Systems* 2014, 1 **(2014)**.
5. A. Rubaai, M. J. Castro-Sitiriche, and A. R. Ofoli, *IEEE Transactions on Industry Applications* 44, 1977 **(2008)**.
6. R. G. Kanojiya and P. M. Meshram, Optimal tuning of PI controller for speed control of DC motor drive using particle swarm optimization, *2012 International Conference on Advances in Power Conversion and Energy Technologies (APCET)*, (Dc) **(2012)**, pp. 1–6.
7. G. Dewantoro, Robust fine-tuned PID controller using Taguchi method for regulating DC motor speed, *2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE)* **(2015)**, pp. 173–178.
8. H. P. Huang, J. C. Jeng, C. H. Chiang, and W. Pan, *Journal of Process Control* 13, 769 **(2003)**.
9. M. F. N. Shah, M. A. Zainal, A. Faruq, and S. S. Abdullah, Metamodeling approach for PID controller optimization in an evaporator process, *2011 4th International Conference on Modeling, Simulation and Applied Optimization, (ICMSAO 2011)* **(2011)**, pp. 5–8.

10. G. Wang and S. Shan, *Journal of Mechanical Design* 129, 370 **(2007)**.

11. K.-Y. Hwang, S.-B. Rhee, B.-Y. Yang, and B.-I. Kwon, *IEEE Transactions on Magnetics* 43, 1833 **(2007)**.

12. X. Liu, C. Ma, M. Li, and M. Xu, A kriging assisted direct torque control of brushless DC motor for electric vehicles, *2011 Seventh International Conference on Natural Computation (ICNC)*, IEEE **(2011)**, Vol. 3, pp. 1705–1710.

13. X. Liu, M. Li, and M. Xu, *International Journal of Automotive Technology* 17, 153 **(2016)**.

14. J. Havinga, A. H. van den Boogaard, and G. Klaseboer, *Structural and Multidisciplinary Optimization* 55, 1345 **(2017)**.

15. J. P. C. Kleijnen, *European Journal of Operational Research* 164, 287 **(2005)**.

16. W. C. M. Van Beers and J. P. C. Kleijnen, Kriging interpolation in simulation: A survey, *Proceedings of the 2004 Winter Simulation Conference* **(2004)**, Vol. 1, pp. 107–115.

17. G. Dellino and C. Meloni, Uncertainty Management in Simulation-Optimization of Complex Systems, Springer, Boston, MA **(2015)**.

18. J. P. C. Kleijnen, Simulation Optimization Through Regression or Kriging Metamodels.

19. J. P. C. Kleijnen, Design and Analysis of Simulation Experiments, 2nd edn., Springer, Cham **(2015)**.

20. J. P. C. Kleijnen, *European Journal of Operational Research* 256, 1 **(2017)**.

21. M. D. McKay, R. J. Beckman, and W. J. Conover, *Technometrics* 21, 239 **(1979)**.

22. T. Bartz-Beielstein, C. Jung, and M. Zaefferer, *Uncertainty Management in Simulation-Optimization of Complex Systems* 59 **(2015)**.

23. E. Del Castillo, Process Optimization A Statistical Approach **(2007)**, p. 480.

24. F. A. C. Viana, *Quality and Reliability Engineering International* 32, 1975 **(2016)**.

25. D. G. Krige, *Journal of the Southern African Institute of Mining and Metallurgy* 52, 119 **(1951)**.

26. J. P. C. Kleijnen, *European Journal of Operational Research* 192, 707 **(2009)**.

27. A. Parnianifard, A. Azfanizam, M. Ariffin, M. Ismail, and N. Ebrahim, *Decision Science Letters* 8, 17 **(2019)**.

28. R. R. Barton, Simulation optimization using metamodels, *Proceedings of the 2009 Winter Simulation Conference* **(2009)**, pp. 230–238.

29. T. W. Simpson, T. M. Mauery, J. Korte, and F. Mistree, *AIAA Journal* 39, 2233 **(2001)**.

30. S. N. Lophaven, J. Søndergaard, and H. B. Nielsen, DACE A Matlab Kriging Toolbox, IMM Informatiocs and Mathematical Modelling **(2002)**, pp. 1–28.

31. R. Myers, D. C. Montgomery, and M. C. Anderson-Cook, Response Surface Methodology: Process and Product Optimization Using Designed Experiments, 4th edn., John Wiley & Sons **(2016)**.

32. F. Jurecka, Robust design optimization based on metamodeling techniques, Ph.D. Thesis **(2007)**.

33. K. J. Astrom and T. Hagglund, PID Controllers: Theory, Design, and Tuning, Isa Research Triangle Park, NC **(1995)**, Vol. 2.

34. G. Dewantoro, *International Journal of Power Electronics and Drive Systems (IJPEDS)* 7, 734 **(2016)**.

35. N. Thomas and P. Poongodi, Position control of DC motor using genetic algorithm based PID controller, *Proceedings of the World Congress on Engineering* **(2009)**, Vol. 2, pp. 1–3.

36. M. N. Ab Wahab, S. Nefti-Meziani, and A. Atyabi, *Plos One* 10, e0122827 **(2015)**.

Received: 20 December 2017. Accepted: 14 January 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1480–1483, 2018

RESEARCH ARTICLE

HEMT Based 6T Static Random Access Memory Cell Design for High Speed Applications

Balwant Raj[1,*] and Sukhleen Bindra Narang[2]

[1] *UIET, Panjab University SSG Regional Centre Hoshiarpur, Punjab, India*
[2] *Dept. of Electronics Technology, Guru Nanak Dev University, Amritsar, Punjab, India*

In this paper 6T SRAM Cell has been designed using HEMT device. Due to high mobility of HEMT, it is more suitable candidate for high speed memories design. Read delay and write delay have been calculated for SRAM cell. The results of proposed HEMT based SRAM cell design are compared and with reported data for the validation of design approach. Based on the analysis of SRAM cell, read delay for 5 GHz was 38.82 ps and the write delay was 9.82 ps. The read and write operations are also verified using TCAD tool for the proposed design.

Keywords: SRAM, Delay, HEMT, TCAD Tool and Simulations.

1. INTRODUCTION

Memories are made of many different types and technologies. Semiconductor memory has the property of random access, which means that it takes the same amount of time to access any memory location, so data can be efficiently accessed in any random order.[1] Semiconductor memory has much faster access times than other types of data storage; a byte of data can be written to or read from semiconductor memory within a few nanoseconds or picoseconds, while access time for rotating storage such as hard disks is in the range of milliseconds.[2] The rapid growth in the requirement of semiconductor memories, there has been a number of technologies and types of memories that have emerged.[3,4] GaAs memories have high speeds and offers an alternative to Silicon based Memories. GaAs devices provide the fastest switching elements available today, not counting superconducting devices which operate at cryogenic temperatures.[5] However, GaAs circuits are preferred to superconductor circuits because of their compatibility with ECL and other silicon logic families and their ability to operate at room temperature instead of at cryogenic temperatures. This work presents the calculation of read/write delay and analysis for HEMT based SRAM cell design.

2. BASIC SRAM CELL DESIGN

A Basic 6T SRAM cell is shown in Figure 1. Transistors M_1, M_2 and M_3, M_4 are used to store 1 bit and transistors M_5 and M_6 are the access transistors and connect to the bit line and word line, these transistors are used for reading and writing operations. Static Random Access memory (SRAM) is able to read and write data contents as long as the power supply voltage is provided. SRAM is a type of semiconductor memory which uses bistable latching circuitry to store each bit. SRAM is different from DRAM which must be periodically refreshed to maintain its data while SRAM exhibits data remains but it is still volatile in the conventional sense. SRAM is faster and more reliable than most common DRAM. The delay and power of SRAM's have been reduced over the year.[6]

SRAM is also used in personal computers, routers and peripheral equipment's, workstations, internal CPU catches and external Burst mode SRAM catches, CPU resister files, hard disk buffers, router buffers, etc. LCD screens and digital printers are also used static SRAM to hold the image displayed or to be printed. Majority of the conventional SRAM's are based on MOSFET technology. Currently many companies are manufacturing SRAM in the GaAs technology which offers very low static power dissipation, superior noise margin and high operational speed.

3. SRAM CELL OPERATION

This section introduces the read/write operation described. HEMT based SRAM cell with separate read and write lines to achieve high speed and low power consumption. As SRAM cells are getting faster and faster, the need for handling large data at very high speed in microprocessor

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1480/004 doi:10.1166/jctn.2018.7380

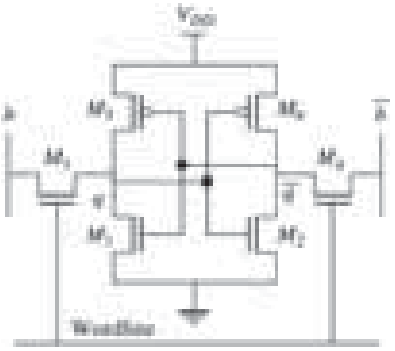


Fig. 1. A basic SRAM cell.

and other electronics applications increases. To access data at such high speed, gets more difficult to accomplish with conventional design of SRAM using MOSFET. HEMT based SRAM cell can overcome this challenge. The read and write operation of HEMT based SRAM cell is as given below.

3.1. Read Operation

In read operation, word line is high (WL = 1) and access transistors are 'ON.' The bit line and bit line bar are read by a sense amplifier. In order to reduce the power consumption of the non-selected cells, the following reading mechanism is used. In SRAM cell, the latch is only required to provide storage and is not responsible for driving the read lines. Therefore, it can be made to operate with a very low standby current, resulting in low power consumption and minimum area.[7]

3.2. Write Operation

The write operation is similar to that in a conventional six transistor SRAM cell. Data is placed on the bit write line, and the write word line is raised. The cross-coupled transistors force the internal nodes to change to appropriate voltage levels, maintaining the state of the cell. Writing a logic value into the storage nodes of the latch is accomplished by the application of the word line voltage and forcing one of the write lines to logic 1. During the read operation, the potentials of both write lines are kept low to prevent an accidental write into the cell.[7]

4. PROPOSED HEMT SRAM CELL DESIGN

In order to minimize the access time, the pull-up and pull-down delays of the circuits should be small. These points are considered in the proposed HEMT SRAM cell design. The proposed high-speed HEMT based SRAM cell is shown in Figure 2. The SRAM cell consists of four *n*-HEMT, two *p*-HEMT. Source–gate back biasing in the *p*-HEMTs, M_1 and M_2 are used as a subthreshold current reduction circuit in order to reduce the power

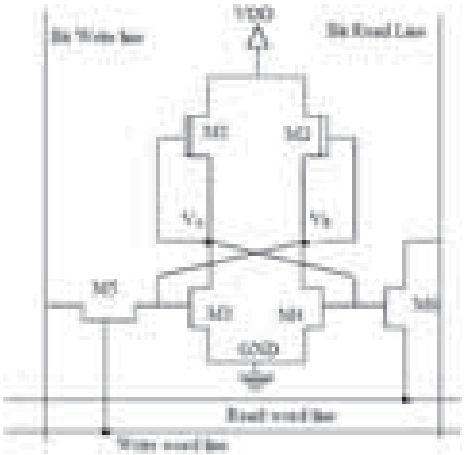


Fig. 2. Proposed HEMT based SRAM cell.

dissipation of the cell. The latch formed by the cross-coupled transistors M_3 and M_4 provides a robust storage element with reduced static power dissipation. Transistor M_5 implements one write-only port, while transistor M_6 acts as a read-only port.

5. RESULTS AND DISCUSSION

We evaluated the reading and writing operation of SRAM cell along with propagation delay. The simulations of SRAM cell have been done using TCAD tool.[10] The simulation results of SRAM cell with various frequencies for both read and write operations have been carried out. Also we have calculated the read delay and write delay in various read and write operations.

5.1. Validation of Read Operation

For the read operation the bit lines and word lines are charged to V_{DD}. At the storing node, logic '1' will pull the voltage on the corresponding bit lines up to a high (not V_{DD} because of the voltage drop across the *n*HEMT access transistor) voltage level. The sense amplifier will detect which bit line is at high voltage and which bit line is at ground.

5.2. Validation of Write Operation

For the write operation, in order to store logic '1' to the cell, *b* is charged to V_{DD} and bbar is charged to ground and wise verse for storing logic '0'. Then the word line is switched to V_{DD} to turn ON the *n*HEMT access transistor. When the access transistors are turned ON, the values of the bit lines are written into Node A and Node B. The node which storing the logic '1' will not go to full V_{DD} because of voltage drops across the *n*HEMT access transistor. After the write operation the word line voltage is reset to ground to turn off the *n*HEMT access transistor.

RESEARCH ARTICLE

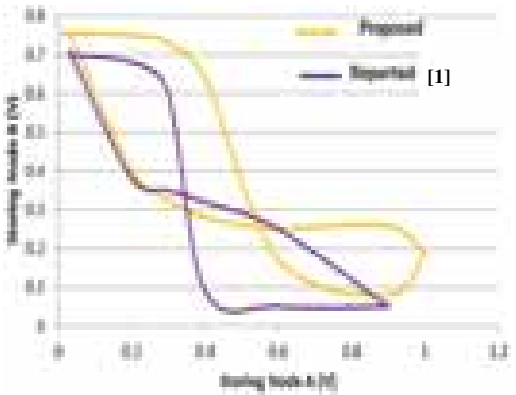


Fig. 3. Comparison of read operation.[1]

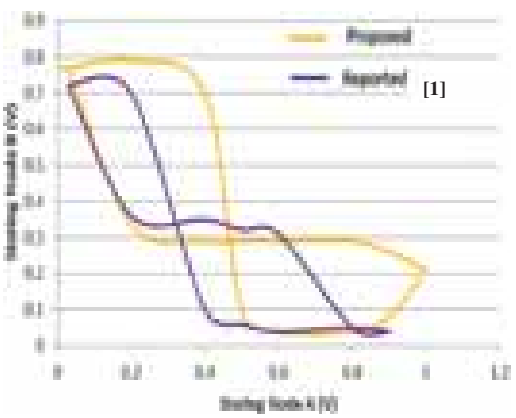


Fig. 4. Comparison write operation.[1]

5.3. Delay Analysis

The access time and performance of 6T SRAM cell depends on read and write delay. Access time is the time required to obtain a signal at the output of sense amplifier after the signal switches at the inputs. Access time is the time required to perform a single read or write operation of SRAM cell.

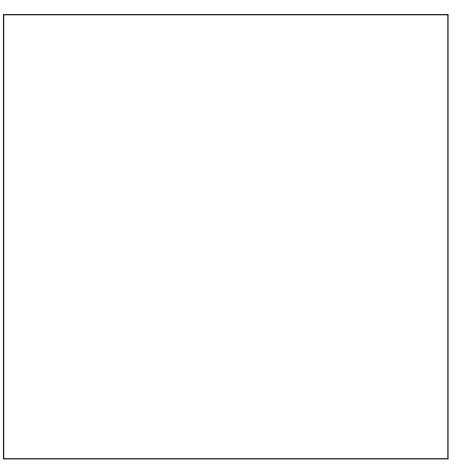


Fig. 5. WSNM of HEMT SRAM cell with 5 GHZ, 20 GHZ and 50 GHZ frequencies.

Table I. Delay calculation of HEMT SRAM cell.

S. no.	Delay type	$f = 5$ GHZ
1.	Write delay (ps)	9.82
2.	Read delay (ps)	38.82

5.3.1. Read Delay

For a 6T SRAM cell, the read delay T_d can be approximated as[6]

$$T_d = \frac{C_{BL}\Delta V}{I_{Read}}$$

Where C_{BL} is the bitline capacitance, ΔV is the input voltage difference required for the sense-amp and I_{Read} is the read current.

With $C_{BL} = 10$fF, $\Delta V = 0.8V_{DD}$, delay in read operation for various frequencies is calculated in Table I.

5.3.2. Write Delay

For a 6T SRAM cell, the read delay T_d can be approximated as[9]

$$T_d = \frac{C_{BL}\Delta V}{I_{write}}$$

Where C_{BL} is the bitline capacitance, ΔV is the input voltage difference required for the sense-amp and I_{Write} is the write current.

With $C_{BL} = 10$ff, $\Delta V = 0.8V_{DD}$, delay in write operation for various frequencies is calculated in Table I.

6. CONCLUSION

A high speed HEMT based SRAM design has been proposed in this work. The SRAM cell design and its read/write operation validation have been carried out. We compared our results with reported data for validation of our design approach. Further the delay analysis of Static Random Access Memory (SRAM) cell have been carried out for evaluation of read/write delay. The HEMT based SRAM cell was design and simulated in TCAD tool. The proposed HEMT SRAM cell will be useful for scientific and research community working for high performance memories design of future nanoscale circuits.

References

1. A. Bernel, R. P. Ribas, and A. Guyot, GaAs MESFET SRAM using a new high memory cell, *5th European Gallium Arsenide and Related III–V Compounds Application Symposium*, Bologna, Italy, September **(1997)**.
2. CMOS Memory Circuits by Tegze P. Haraszti.
3. B. Raj, A. K. Saxena, and S. Dasgupta, *Microelectronics International, UK* 26, 53 **(2009)**.
4. B. Raj, A. K. Saxena, and S. Dasgupta, *IEEE Circuits and System Magazine* 11, 38 **(2011)**.

5. R. J. Trew, *IEEE Transactions on Electron Devices* 52, 638 **(2005)**.

6. S. Notomi, Y. Awano, M. Kosugi, T. Nagata, K. Kosemura, M. Ono, N. Kobayashi, H. Ishiwari, K. Odani, T. Mimura, and M. Abe, A high-speed 1 K 4-bit static RAM using 0.5-m-gate HEMT, *Proc. GaAs IC Symp.* **(1987)**, pp. 177–180.

7. Oh. Saeroonter, J. Park, S. S. Wong, and H.-S. P. Wong, Modeling and analysis of III–V logic FETs for devices and circuits: Sub-22 nm technology III–V SRAM cell design, *11th Int'l Symposium on Quality Electronic Design*, IEEE **(2010)**, 978-1-4244-6455-5/10/$26.00.

8. E. Bushesri, V. Bratov, A. Thiede, V. Staroselsky, and D. Clark, *IEEE Electronics Letters* 31 **(1995)**.

9. Area-Performance Trade-Offs in Sub-Threshold SRAM Designs by George Cramer and Ping-Chen Huang.

10. ATLAS User Manual, SILVACO, January **(2013)**.

11. Oh. Saeroonter and H. S. Philip, *IEEE Electron Device Letters* 32 **(2011)**.

12. B. Raj, Quantum Mechanical Potential Modeling of FinFET, Book Towards Quantum FinFET, Springer **(2014)**, Vol. 17, pp. 81–97, ISBN 978-3-319-02021-1.

13. B. Raj, Jatin Mitra, Deepak Kumar Bihani, V. Rangharajan, A. K. Saxena, and S. Dasgupta, *Journal of Low Power Electronics (JOLPE)* 7, 163 **(2011)**.

14. V. K. Sharma, M. Pattanaik, and B. Raj, *International Journal of Electronics* 102, 200.

15. S. K. Vishvakarma, V. Agrawal, B. Raj, S. Dasgupta, and A. K. Saxena, *Journal of Computational and Theoretical Nanoscience* 4, 1144 **(2007)**.

16. B. Raj, A. K. Saxena, and S. Dasgupta, *Elsevier's Journal of Material Science in Semiconductor Processing, Elsevier* 16, 1131 **(2013)**.

Received: 23 January 2018. Accepted: 25 February 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1484–1489, 2018

RESEARCH ARTICLE

Density Functional Theory Study of Gold Doped Armchair Graphene Nanoribbons with Vacancies

Lotfi Benchallal[1], Slimane Haffad[1], Lyes Lamiri[1], Fouad Boubenider[2], Hachemi Zitoune[3], Badis Kahouadji[1], and Madani Samah[1,*]

[1] *Université A/Mira, Route Targua Ouzemour 06000 Bejaia, Algérie*
[2] *Laboratoire de physique des matériaux, Université Houari Boumediene, BP 32, El Allia 16111, Bab Ezzouar Alger, Algérie*
[3] *Université Mohand Oulhadj, Rue Drissi Yahia 10000 Bouira, Algérie*

Systematic investigation of electronic, magnetic of vacancies-defected AGNRs and substitution by gold (Au) atoms are done in the framework of DFT. Remarkably, one vacancy-AGNR with a semiconductor behavior exhibits a high magnetic moment equal to 1.65 μ_B. Substitution of the missed carbon atom by Au one causes suggestive changes in the electronic spin and reduces magnetic moment to a value of 0.98 μ_B and still remains semiconductor. Divacancies (2V) and four vacancies (4V) AGNRs are nonmagnetics and semiconductors a with relatively large band gaps (0.40 and 0.62 eV). Substitutions with two and three gold atoms respectively leaves unchanged the behavior of the first structure and transform the last one to metallic like.

Keywords: DFT, Vacancies, AGNR, Magnetic, Gold.

1. INTRODUCTION

Graphene as a two-dimensional structure is still the area of intense and interesting experimental and theoretical research.[1,2] It opens new possibilities for research in the field of nanoelectronics and several other fields.[3–8] Such as applications requires the control of the magnetism and electronic properties of these structures by controlling the effects of defects and impurities. Several works has been devoted to the study of defects and different types of impurities in these materials.[9–17] Studies have shown that graphene properties can be effectively altered by transition metal adsorption.[18] A detailed analysis of the nanoribbon properties doped with gold was presented by Lit et al.[19] Many works has been committed to the investigation of imperfections and diverse kinds impurities influences in these materials.[9–17] Studies have demonstrated that graphene properties can be successfully modified by metal atoms adsorption.[18] Exhaustive study of the nanoribbon properties doped with gold was given by Lit et al.[19] Wu et al.[20] have revealed that news and different properties are demonstrated when gold atoms are adsorbed on graphene on different configurations.

Understanding atomic-scale interactions between metal atoms and graphene seems to be more than necessary in the aim to optimize properties and fabricating of nanoelectronic devices. Recently, Akola and Hakkinen[21] using first principles calculations, have investigated and study the effects of gold (Au) atom and clusters on the defected graphitic surface.

This work is devoted to study the structural, electronic and magnetic properties of Au substituted gold atom in defected AGNRs in the framework of density functional theory (DFT) Initially pristine AGNRs is non-magnetic turn them magnetic in the presence of vacancies. in nature. Doping these structure with gold atoms changes their drastically their magnetization.

2. COMPUTATIONAL DETAILS

Our DFT calculations are performed using the Siesta code[22–24] based on localized numerical atomic orbitals (NAOs).[23,25] Perdew–Burke–Ernzerhof[26] GGA (PBE-GGA) functional is chosen in all our calculations. Interactions between periodic images of the defective graphene layer are avoided by enlarging the dimensions of the supercell perpendicular to the graphene plane until 15 Å. For all investigated supercells, $1 \times 16 \times 1$ Monkhorst–Pack[27] sampling for k-points is used for the unit cell of graphene. Molecular dynamic of atomic relaxation for the atomic coordinates is optimized until forces were smaller than 0.01 eV Å^{-1}. Troullier–Martins norm-conserving

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1484/006 doi:10.1166/jctn.2018.7381

pseudopotentials[28] including nonlinear core corrections for the metal atoms are used. The mesh cutoff used to calculate the Hartree and exchange-correlation contribution to the total energy and Hamiltonian equals 400 Ry. For the calculation of the magnetic and electronic properties a double-ζ polarized (DZP)[23, 25] basis set was used.

3. RESULTS AND DISCUSSION

Missing lattice carbon atom is the simplest to create a defect in our material. Single vacancies (SV) in graphene have been observed by many autors.[29–31] On Figure 1(A), the SV is characterized by the saturation of two of the three dangling bonds toward the missing atom and the third one remains because of geometrical reasons. This situation causes the formation of three pentagons-rings.

Before molecular dynamic relaxation, a single Au atom is placed at the vacancy location. Relaxation process remains the gold atom in the plane of the AGNR as founded in previous studies.[21,32–38]

Au atom is bonded to three nearest carbon atoms with bond length about 1.85 Å less than values obtained elsewhere (2.01 Å,[39] 2.1 Å,[40] 2.44[41]). Au atom makes the junction with three distorted hexagons where the C–C distance is between 1.36 Å and 1.48 Å leading for a lattice distortion. Different distortions of graphene are induced by different metal adatoms. Liu et al.[42–45] determined that the distortion patterns inside the graphene layer resulting from the adsorption of group I–IV metals are very comparable. It seems from their results that the distortion generated by Cu is much larger than that by Ag and Au adatoms. They found that the distortion inside the graphene plane of the second shell carbon atoms is larger than that of the first shell carbon atoms.

The calculated binding energy is about 1.1 eV per atom. This value confirms that he vacancy site to be more stable[46] in case of adsorption in good agreement with observation made in case of Au-adsorbed AGNRs.[47] Substitution of Au at various positions yields lower energies compared with other transition metals like Ti, Mn.[48] As reported by Refs. [49, 50], no apparent change or distortion within the optimized geometry of the investigated ribbons is found after Au substitution or adsorption.

Double vacancies (DV) can be formed by the coalescence of two SVs or by removing two neighboring atoms. Kotakoski et al. have reported such defect in electron microscopy experiments.[51] This case is shown on Figure 1(B). A fully reconstructed DV [(5-8-5) defect] appear rather than four hexagons in perfect graphene. Our results indicate that the formation energy of a DV is comparable to that of SV, in good agreement with previous results.[52, 53] Substituttion of the two missing carbon atoms by two gold atoms, leads after structural relaxation to the structure mentioned on Figure 2. We note that the graphene nanoribbon is distorted. One of the two atoms still in the plane of graphene structure and a second atom move out off. The Au–Au distance is equal to 2.95 Å. This value is greater that one obtained with Au dimer on graphene.[42]

Remove more than two carbon atoms form larger complex defects. It has been demonstrate that structures with even number of missing carbon atoms are more favorable than odd number ones[51] with more unsaturated bonds. The formation of large holes has indeed been observed in electron microscopy experiments.[54]

In our case, the four vacancies AGNR, exhibit a hole in the plane of the structure. Interatomic distances are between 1.36 and 1.48 Å and all carbon atoms remains in the plane. Three Au atoms are inserted in this hole and after relaxation the ribbon is distorted. Two Au atoms still in the distorted graphene plane and the third go out off. The three Au atoms form an equilateral triangle with an Au–Au distance equal to 2.62 Å. Martin Amf et al.[55] have obtained an isosceles triangle with larger distances.

It is well known that charge transfer between host and guest atoms is the modification in the electronic structure. The gold substitution in AGNR induces mainly the attractive interaction between the d-orbitals of gold atoms and the π-orbitals of the sp^2 hybridized carbon.[56] From the band structures, the semiconducting nature of one Au atoms and two Au atoms doped AGNR is obvious. Interesting informations are revealed at the Fermi level reveals about Au substitution. AGNRs give rise to exceptional flat bands as shown in Figure 2. These bands play an important role in the study of strongly correlated phenomena.[57] From Figure 2, we see that flat bands vary slightly with Au substitution. The variations across Fermi level are due to the difference in amount of charge transfer.

In this work, the band gaps change slightly from 0.11 eV in case of 1V-AGNR to a high band gap value of 0.15 eV in case of one Au atom at the hole site. The increase of band gap by substitution of carbon atoms by Au one is also observed by earlier studies.[58] Absorbing Au at the hole site decreases the magnetic moment of the structure. 1V-AGNR with magnetic moment of 1.65 μ_B, loses a part

Table I. Formation energies, band gaps, magnetic moment and nature of the materials for AGNRs with one vacancy, two vacancies and four vacancies and AGNRs doped with one, two and three gold atoms.

Structures	E_g (eV)	M (μ_B)	Formation energy (eV)	Nature of material
1 vacancy (1V)	0.00	1.65	0.22	Metal
1V + 1Au	0.15	−0.98	1.1	Semiconductor
2 vacancies (2V)	0.33	0.00	0.37	Semiconductor
2V + 2Au	0.41	0.00	0.93	Semiconductor
4 vacancies (4V)	0.46	2.00	0.77	Semiconductor
4V + 3Au	0.00	0.00	0.92	Metal

RESEARCH ARTICLE

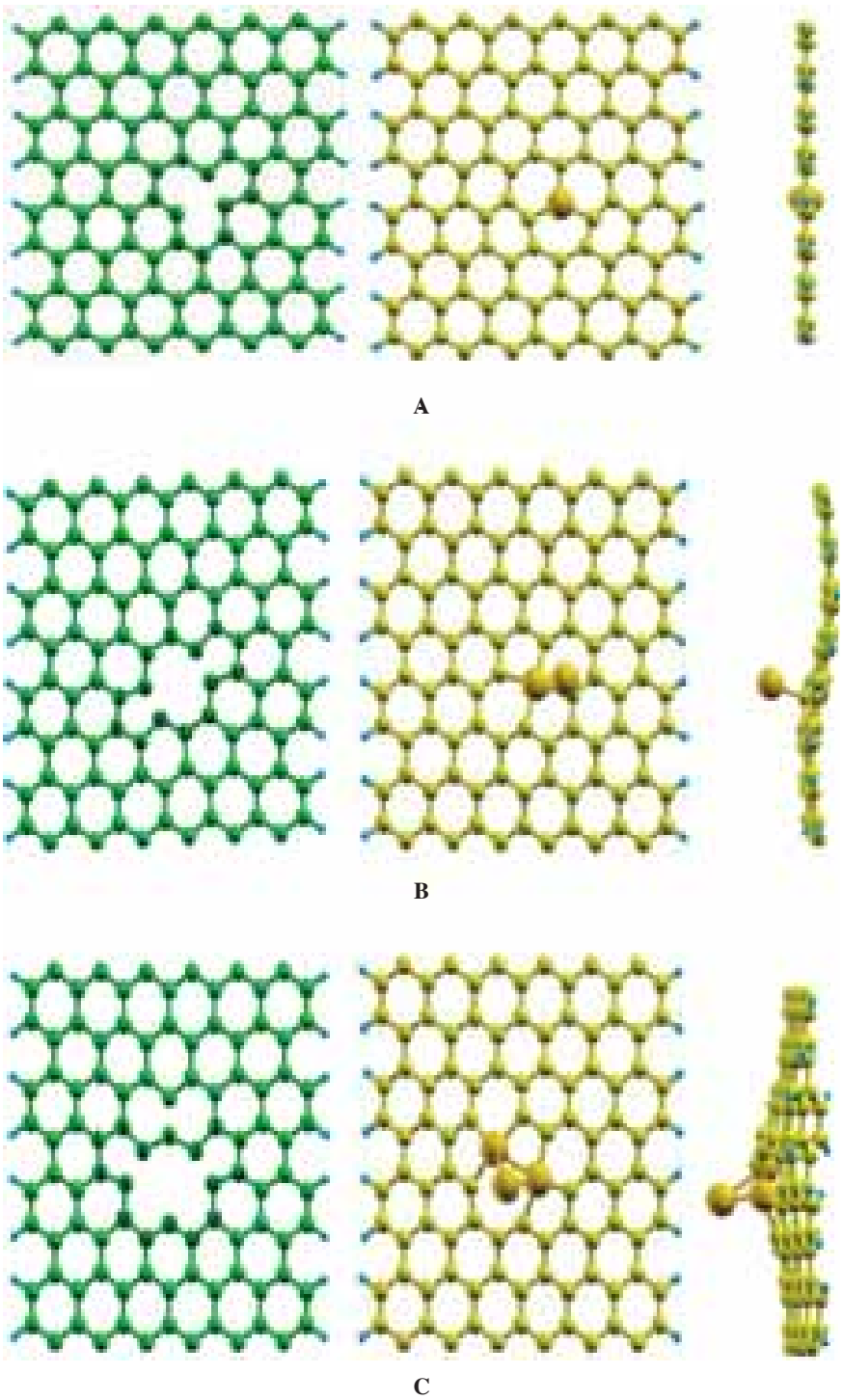


Fig. 1. Relaxed atomic structures of AGNR doped with: (A) One Au atom, (B) two Au atoms and (C) three Au atoms.

of its magnetization when an Au atom is inserted in the hole. The final moment equals $-0.98\ \mu_B$ (see Table I). The Au atom carry $0.14\ \mu_B$, sensitively equal to the value obtained by Santos et al.[40]

Spin dependent total density of states (TDOS) of the Au-adsorbed vacancy site is depicted on Figure 3. The PDOS calculations (note mentioned in this article) of Au-adsorbed at vacancy site divulges strong interactions

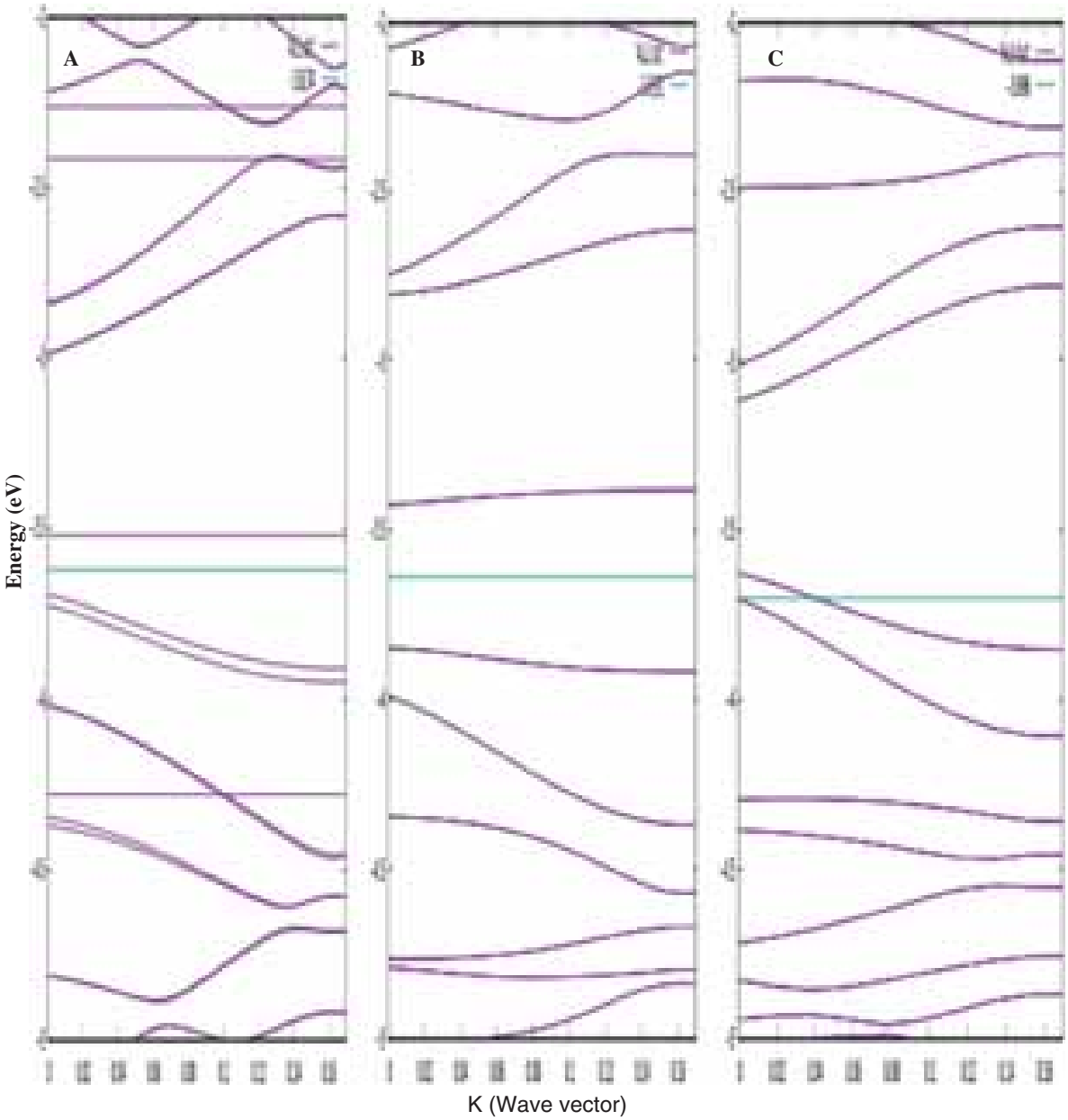


Fig. 2. Bands structure of AGNR doped with: (A) One Au atom, (B) two Au atoms and (C) three Au atoms.

between the electronic states of Au 5*d* and neighboring C atoms. We observe near the Fermi level, that the states due to Au-*s* orbitals are symmetric on spin polarization in contrast to spin independent calculations done by Srivastava et al.[58] The contribution from the metal atom is negligible and major part results from carbon. We can explain these results by the fact that the lower energy position of the *d* shell in the case of the noble metals.

Essentially, the spin is composed by the Au atom and its three nearest neighbor's carbon atoms leading for a consequent contribution from the relatively localized carbon *sp* lobes in the description of the defect states near the Fermi level for these impurities.[40]

Three Au atoms substitution in the 4-vacancies AGNR exhibits a metallic nature. Srivastava et al. have obtained that substituting Au at both edges of the ribbon reduces the band gap and turning the GNR into a metallic one.[58, 59] The metallicity is attributed to the delocalized 3*d* electrons of Au atom, which occupies the energy states near the Fermi level. The 4V-AGNR initially with 2.98 μ_B becomes nonmagnetic when three gold atoms are inserted in (see Table I).

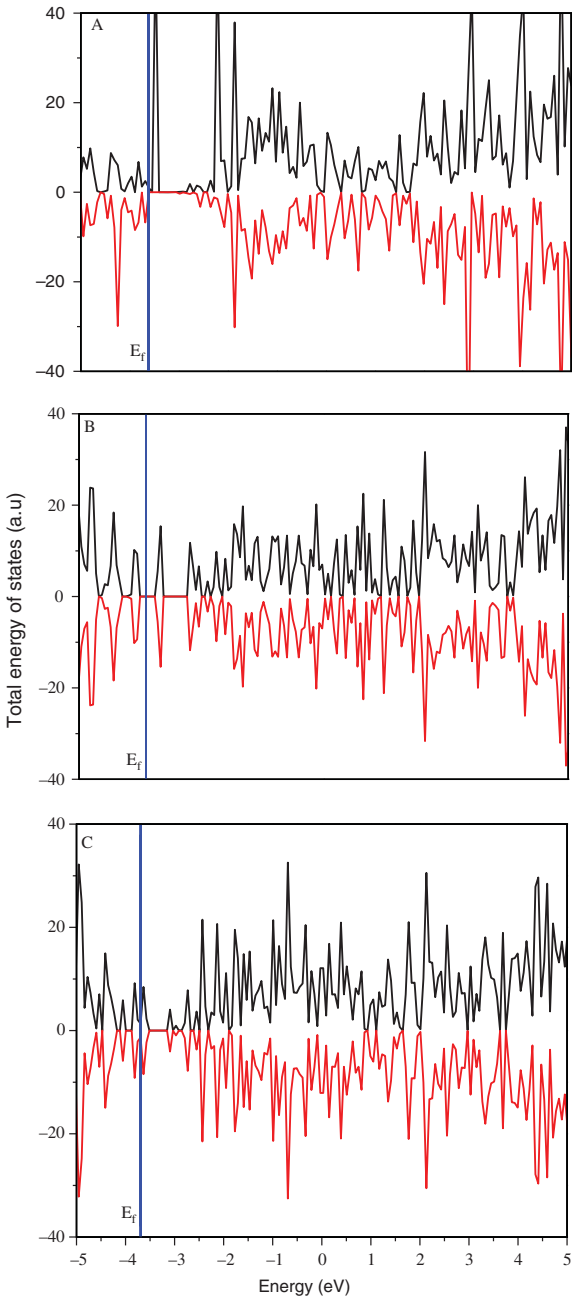


Fig. 3. Total density of states of AGNR doped with: (A) One Au atom, (B) two Au atoms and (C) three Au atoms.

4. CONCLUSION

Ab initio pseudopotential method combined with the generalized gradient approximation has been employed to study the structural, electronic and magnetic properties of AGNRs with one, two and four vacancies and doped with one, two and three gold (Au) atoms. Substitution of the missed carbon atom by Au one induce significant changes in the electronic spin and reduces magnetic moment to a value of 0.98 μ_B and still remains semiconductor. Divacancies (2V) and four vacancies (4V) AGNRs are nonmagnetics and semiconductors a with relatively large band gaps (0.40 and 0.62 eV). Substitutions with two and three gold atoms respectively leaves unchanged the behavior of the first structure and transform the last one to metallic like.

References

1. A. K. Geim and K. S. Novoselov, *Nat. Mater*. 6, 183 (**2007**).
2. A. H. Castro Neto, F. Guinea, N. M. Peres, K. S. Novoselov, and A. K. Geim, *Rev. Mod. Phys*. 81, 109 (**2009**).
3. P. Avouris, Z. Chen, and V. Perebeinos, *Nat. Nanotechnol.* 2, 605 (**2007**).
4. L. E. Hueso, J. M. Pruneda, V. Ferrari, G. Burnell, J. P. Valdés-Herrera, B. D. Simons, R. B. Littlewood, E. Artacho, A. Fert, and N. D. Mathur, *Nature* 445, 410 (**2007**).
5. B. Trauzettel, D. V. Bulaev, D. Loss, and G. Burkard, *Nat. Phys.* 3, 192 (**2007**).
6. N. Tombros, C. Jozsa, M. Popinciuc, H. T. Jonkman, and B. J. van Wees, *Nature* 448, 571 (**2007**).
7. O. V. Yazyev, *Nano Lett.* 8, 1011 (**2008**).
8. J. Fischer, B. Trauzettel, and D. Loss, *Phys. Rev. B* 80, 155401 (**2009**).
9. P. Esquinazi, D. Spemann, D. Höhne, A. Setzer, K. H. Han, and T. Butz, *Phys. Rev. Lett.* 91, 227201 (**2003**).
10. P. O. Lehtinen, A. S. Foster, A. Ayuela, A. V. Krasheninnikov, K. Nordlund, and R. Nieminen, *Phys. Rev. Lett.* 91, 017202 (**2003**).
11. P. O. Lehtinen, A. S. Foster, Y. Ma, A. V. Krasheninnikov, and R. Nieminen, *Phys. Rev. Lett.* 93, 187202 (**2004**).
12. T. Makarova and F. Palacios (eds.), Carbon Based Magnetism, Elsevier, Amsterdam (**2006**).
13. A. V. Krasheninnikov and F. Banhart, *Nat. Mater.* 6, 723 (**2007**).
14. B. Uchoa, V. N. Kotov, N. M. R. Peres, and A. H. Castro Neto, *Phys. Rev. Lett.* 101, 026805 (**2008**).
15. O. V. Yazyev, *Phys. Rev. Lett.* 101, 037203 (**2008**).
16. J. J. Palacios, J. Fernández-Rossier, and L. Brey, *Phys. Rev. B* 77, 195428 (**2008**).
17. J. Fernández-Rossier and J. J. Palacios, *Phys. Rev. Lett.* 99, 177204 (**2007**).
18. C. Cao, M. Wu, J. Jiang, and H.-P. Cheng, *Phys. Rev. B* 81, 205424 (**2010**).
19. J. Lit, M. P. Boneschanscher, D. Vanmaekelbergh, Mari Ijäs, A. Uppstu, M. Ervasti, A. Harju, P. Liljeroth, and I. Swart, *Nat. Commun.* 4, 1 (**2013**).
20. Y. Wu, W. Jiang, Y. Ren, W. Cai, W. H. Lee, H. Li, R. D. Piner, C. W. Pope, Y. Hao, H. Ji, J. Kang, and R. S. Ruoff, *Small* 8, 3129 (**2012**).
21. J. Akola and H. Hakkinen, *Phys. Rev. B* 74, 165404 (**2006**).
22. D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler, *Int. J. Quantum Chem.* 65, 453 (**1997**).
23. J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, *J. Phys.:Condens. Matter* 14, 2745 (**2002**).
24. D. Sánchez-Portal, P. Ordejón, and E. Canadell, Principles and Applications of Density functional Theory in Inorganic Chemistry II, Structure and Bonding Series, Springer, Berlin (**2004**), Vol. 113, pp. 103–170
25. J. Junquera, O. Paz, D. Sánchez-Portal, and E. Artacho, *Phys. Rev. B* 64, 235111 (**2001**).
26. J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* 77, 3865 (**1996**).
27. H. J. Monkhosrt and J. D. Pack, *Phys. Rev. B* 13, 5188 (**1976**).
28. N. Troullier and J. L. Martins, *Phys. Rev. B* 43, 1993 (**1991**).

29. M. H. Gass, U. Bangert, A. L. Bleloch, P. Wang, R. R. Nair, and A. K. Geim, *Nat. Nanotechnol.* 3, 676 (**2008**).
30. J. C. Meyer, C. Kisielowski, R. Erni, M. D. Rossell, M. F. Crommie, and A. Zettl, *Nano Lett.* 8, 3582 (**2008**).
31. M. M. Ugeda, I. Brihuega, F. Guinea, and J. M. Go mez-Rodrıguez, *Phys. Rev. Lett.* 104, 096804 (**2010**).
32. A. V. Krasheninnikov, P. O. Lehtinen, A. S. Foster, P. Pyykko, and R. M. Nieminen, *Phys. Rev. Lett.* 102, 126807 (**2009**).
33. P. Jensen, X. Blase, and P. Ordejón, *Surf. Sci.* 564, 173 (**2004**).
34. P. Jensen and X. Blase, *Phys. Rev. B* 70, 165402 (**2004**).
35. G. M. Wang, J. J. BelBruno, S. D. Kenny, and R. Smith, *Phys. Rev. B* 69, 195412 (**2004**).
36. G. M. Wang, J. J. BelBruno, S. D. Kenny, and R. Smith, *Surf. Sci.* 576, 107 (**2005**).
37. S. Malola, H. Häkkinen, and P. Koskinen, *Appl. Phys. Lett.* 94, 043106 (**2009**).
38. Y. Gan, L. Sun, and F. Banhart, *Small* 4, 587 (**2008**).
39. W. Zhang, L. Sun, Z. Xu, A. V. Krasheninnikov, P. Huai, Z. Zhu, and F. Banhart, *Phys. Rev. B* 81, 125425 (**2010**).
40. E. J. G. Santos, A. Ayuela, and D. Sánchez-Portal, *New Journal of Physics* 12, 053012 (**2010**).
41. R. Varns and P. Strange, *J. Phys.: Condens. Matter* 20, 225005 (**2008**).
42. X. Liu, C. Z. Wang, Y. X. Yao, W. C. Lu, M. Hupalo, M. C. Tringides, and K. M. Ho, *Phys. Rev. B* 83, 235411:1 (**2001**).
43. X. Liu, C. Z. Wang, M. Hupalo, Y. X. Yao, M. C. Tringides, W. C. Lu, and K. M. Ho, *Phys. Rev. B* 82, 245408:1 (**2010**).
44. M. Hupalo, X. Liu,, C. Z. Wang, W. C. Lu, Y. X. Yao, K. M. Ho, and M. C. Tringides, *Adv. Mater.* 23, 2082 (**2011**).
45. X. Liu, C. Z. Wang, M. Hupalo, W. C. Lu, M. C. Tringides, Y. X. Yao, and K. M. Ho, *Phys. Chem. Chem. Phys.* 14, 9157 (**2012**).
46. S. Gupta, G. Kaur, and K. Dharamvir, *AIP Conf. Proc.* 1447, 371 (**2012**).
47. H. Sevinçli, M. Topsakal, E. Durgun, and S. Ciraci, *Phys. Rev. B* 77, 195434 (**2008**).
48. S. S. Chauhan, P. Srivastava, and A. K. Shrivastava, *Solid State Commun.* 154, 69 (**2013**).
49. X. Peng-Yang, Z. Gui-Lin, L. Yong-An, W. Jian-Guo, and L. Xiao-Nian, *J. Acta Phys. Chim. Sin.* 28, 2 (**2012**).
50. K. T. Chan, J. B. Neaton, and M. L. Cohen, *Phys. Rev. B* 77, 235430 (**2008**).
51. J. Kotakoski, A. V. Krasheninnikov, and K. Nordlund, *Phys. Rev. B* 74, 245420 (**2006**).
52. A. V. Krasheninnikov, P. O. Lehtinen, A. S. Foster, and R. M. Nieminen, *Chem. Phys. Lett.* 418, 132 (**2006**).
53. A. A. El-Barbary, R. H. Telling, C. P. Ewels, M. I. Heggie, and P. R. Briddon, *Phys. Rev. B* 68, 144107 (**2003**).
54. C. O. Girit, J. C. Meyer, R. Erni, M. D. Rossell, C. Kisielowski, L. Yang, C.-H. Park, M. F. Crommie, M. L. Cohen, and S. G. Louie. *Science* 323, 1705 (**2009**).
55. M. Amft, B. Sanyal, O. Eriksson, and N. V. Skorodumova, *J. Phys.: Condens. Matter* 23, 205301 (**2011**).
56. R. S. Sundaram, M. Steiner, H.-Y. Chiu, M. Engel, A. A. Bol, R. Krupke, M. Burghard, K. Kern, and P. Avouris, *Nano Lett.* 11, 3833 (**2011**).
57. Y.-C. Lee and H.-H. Lin, *J. Phys. Conf. Ser.* 150, 042110 (**2009**).
58. P. Srivastava, S. Dhar, and N. K. Jaiswal, *Phys. Lett. A* 379, 835 (**2015**).
59. P. Srivastava, S. Dhar, and N. K. Jaiswal, *Appl. Phys. A* 117, 1997 (**2014**).

Received: 13 February 2018. Accepted: 11 March 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1490–1494, 2018

RESEARCH ARTICLE

State Dependent Riccati Equation Tracking Control for a Two Link Robot

Seyyed Mohammad Hosseini Rostami[1,*] and Mansoureh Ghazaani[2]

[1] Department of Electrical and Computer Engineering, Shiraz University of Technology, Shiraz, 71557-13876, Iran
[2] Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, 1477893855, Iran

The optimal control of nonlinear systems cannot be done similar to the methods of linear systems. One of the significant viewpoint's optimal control of nonlinear systems is using the SDRE method. This method in addition to stability creates a proper and robust function in a wide range of nonlinear systems. Nonlinear optimal controllers SDRE is extended linear optimal control LQR. Strategy SDRE state-dependent Riccati equation has been introduced in the past decade. This strategy is a very effective algorithm for nonlinear feedback analysis whose states are nonlinear, however, which is a flexible idea through state dependent weight matrices. This method involves factoring (i.e., parameterization) of nonlinear dynamics to the state vector and produces a valuable function matrix which is dependent on their states. Tracking control of multiple link robot path in recent years has received a lot of attention. It is expected to get a good performance because the robot control multiple links have a mixed system with a high degree nonlinear, connections, and dynamic behavior varies with time. At the end of this research, the result is that tracking control of the two link robot using state-dependent Riccati equation (SDRE) compared to LQR method has better performance.

Keywords: SDRE, Optimal Control, Nonlinear Control, Tracking Control.

1. INTRODUCTION

The robot is the latest technology achievements and includes machine subtle, sensitive and accurate, much faster than a human and do things automatically. Among the important features of robots is that are remote programmability and controllability; Therefore, they can be used in environments that humans are not able to work there. Using robots in the industry to improve the performance of dynamic systems, quality improvement and falling product prices, increased production percentage and greater ease. Furthermore, according to the recent developments in technology, the use of robots, as well as the human need for high-speed and high-quality robots, the idea of using light robots is suggested. Hence, an SDRE regulator[1,2] is used throughout this paper. SDRE technique developed as a design method, which provides a systematic and effective design of nonlinear controllers, filters, and observers. Because of its versatile features, SDRE is broadly used for different cases.[3–9] There are many articles in the field of tracking robot two-link.[10–13] This paper is organized as follows: a description of the SDRE formulation is given in Sections 2 and 3. Then, motion equations of the robot are presented in Section 4 and SDRE formulation of the robot is derived, and simulation results have reported in the Section 5. Finally, in Section 6, some comments and statements are obtained from this research.

2. NOTATION

The notation used throughout the paper is stated below.

R Weighting matrix inputs
Q Weighting matrix of states
G Matrix final cost
J Cost function
u Control signal.

3. SDRE NONLINEAR REGULATION

3.1. Problem Formulation

Consider a nonlinear system that is autonomous, full-state observable, and input affine, presented in the form of

$$\dot{x}(t) = f(x) + B(x)u(t), \quad x(0) = x_0 \tag{1}$$

Where $x \in R^n$ is the state vector, $u \in R^m$ is the input vector, and $t \in [0, \infty)$, with functions $f: R^n \to R^n$ and $B: R^n \to R^{n\times m}$, and $B(x) \neq 0 \; \forall \, x$. Without any loss of generality, the origin is assumed to be an equilibrium point,

*Author to whom correspondence should be addressed.

1546-1955/2018/15/1490/005 doi:10.1166/jctn.2018.7382

such that $f(0) = 0$. In this context, the minimization of the infinite-time performance criterion is as follows:

$$J(x_0, u(\cdot)) = \frac{1}{2}\int \{x^T(t)Q(x)x(t) + u^T(t)R(x)u(t)\}\, dt \quad (2)$$

Consider above equation which is nonquadratic in x but quadratic in u. The state and input weighting matrices are assumed state dependent such that $Q\colon R^n \to R^{n\times n}$ and $R\colon R^n \to R^{n\times m}$. These design parameters satisfy $Q(x) \geq 0$ and $R(x) > 0$ for all x.[14]

3.2. SDRE Nonlinear Regulator Problem

The SDRE methodology uses extended linearization as the key design concept in formulating the nonlinear optimal control problem. The linear control synthesis method for this case is the LQR synthesis method. SDRE feedback control is an "extended linearization control method" that provides a similar approach to the nonlinear regulation problem for the input-affine system (1) with cost functional (2). The nonlinear system of (1) can be rewritten in the linear structure using SDC form:

$$\dot{x}(t) = A(x)x(t) + B(x)u(t) \quad (3)$$

SDC form is not unique and should be chosen in such a way that $\{Q^{1/2}(x), A(x)\}$ and $\{A(x), B(x)\}$ are point wise observable and point wise stabilizable, respectively.

Where $G(x)$ is obtained from the solution of algebraic SDRE:

$$G(x)A(x) + A^T(x)G(x) + Q(x) - G(x)B(x)R^{-1}B^T(x)G(x) = 0 \quad (4)$$

to get $G(x) \geq 0$.

Where $G(x)$ is the unique, symmetric, positive-definite solution of the algebraic State Dependent Riccati Equation. As a result of this, the nonlinear control is obtained as:

$$u(t) = -R^{-1}(x)B^T(x)G(x)x(t) \quad (5)$$

The response obtained is locally asymptotically stable and is optimal.[15]

4. PROBLEM FORMULATION OF TWO LINK ROBOT

Consider a class of uncertain affine nonlinear MIMO systems as follows:[16]

$$\dot{x} = f(x) + g(x)u + d, \quad y = v(x) \quad (6)$$

In which $x \in R^n$ is the state vector, $u \in R^m$ is the input vector and $y \in R^m$ is the output vector of the system. $f(x), g(x) = [g_1(x), \ldots, g_m(x)]$ and $g_i(x)$. Field vectors are smooth in the $D \subset R^n$ range. $v(x)$ is a relatively smooth output vector function, which is $v(x) = [v_1, \ldots, v_m]^T$. d is a vector function representing an uncertainty system; This uncertainty system can include: parametric changes of the system, unmodified dynamics and external disturbances.

$$\dot{z}(t) = F(t)z(t), \quad z(t_0) = z_0, \quad \tilde{y}(t) = V(t)z(t) \quad (7)$$

Where $z \in R^P$ is the state vector and $\tilde{y} \in R^m$ is the output vector. $F(t)$ and $V(t)$ are time-varying matrices of appropriate dimensions. Suppose that the pair $[F(t), V(t)]$ is observable. Our objective is to design a controller so that the output y from the system (8) can track the ecosystem's output $\tilde{y}$ asymptotically, some given performance criterion is minimized, and the system can exhibit global robustness to uncertainties. Ignoring the uncertainties, the nominal system of the uncertain affine nonlinear system (8) is as follows:

$$\dot{x} = f(x) + g(x)u, \quad y = v(x) \quad (8)$$

4.1. SDRE Nonlinear Tracking Problem

Consider an infinite-time nonlinear tracking problem. Minimize the following nonlinear cost functional nonquadratic in x but quadratic in u:

$$J = \frac{1}{2}\int_0^\infty \{E^T(t)Q(x)E(t) + u^T(t)R(x)u(t)\}\, dt \quad (9)$$

$$E(t) = \xi(t) - y(t) \quad (10)$$

$y(t)$ is output and $\xi(t)$ is desired output. With SDC is obtained:

$$\begin{cases} \dot{x}(t) = A(x)x(t) + B(x)u(t) \\ y(t) = V(x)x(t) \end{cases} \quad (11)$$

By solving the SDRE equation, the following equation is obtained:

$$V^T(x)Q(x)V(x) + G(x)A(x) + A^T(x)G(x) - G(x)B(x)R^{-1}B^T(x)G(x) = 0 \quad (12)$$

Then $s(x)$ is calculated as follows:

$$s(x) = ([A(x) - B(x)R^{-1} \times B^T(x)G(x)]^T)^{-1} \times V^T(x)Q(x)\xi(t) \quad (13)$$

Finally, the control signal is calculated as follows:

$$u(t) = -R^{-1}B^T(x)[G(x)x(t) - s(x)] \quad (14)$$

5. DYNAMIC EQUATIONS OF TWO-LINK ROBOT

Consider the dynamic structure of a two-link robot in Figure 1.[1] In Figure 1 L_1 and L_2 show the robot arms,

RESEARCH ARTICLE

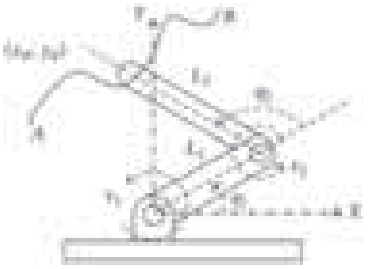


Fig. 1. Dynamic structure of the two-link robot.

τ_1 and τ_2 show the torque of the actuator, q_1 and q_2 show the rotation angle of the robot links in sequence, and $A-B$ track the path using (x_d, y_d).

Equations of the two-link robot with regard to Figure 1 are given as the following.

$$M(q)\ddot{q}+C(q,\dot{q})\dot{q}+N(q)+d=\tau \tag{15}$$

Where $q=[q_1 \;\; q_2]^T$ is the displacement vector, $\tau=[\tau_1,\tau_2]$ is the torque applied to the link and d is the system's uncertainty. $M(q)$, $C(q)$, $N(q)$ and g are defined by the following equations:

$$\begin{aligned} M(q)&=\begin{bmatrix}0.1+0.01\cos(q_2) & 0.01\sin(q_2)\\ 0.01\sin(q_2) & 0.1\end{bmatrix},\\ C(q,\dot{q})&=\begin{bmatrix}-0.005\sin(q_2)\dot{q}_2 & 0.005\cos(q_2)\dot{q}_2\\ 0.005\cos(q_2)\dot{q}_2 & 0\end{bmatrix},\\ N(q)&=\begin{bmatrix}0.01g\cos(q_1+q_2)\\ 0.01g\cos(q_1+q_2)\end{bmatrix}, \quad g=9.8,\\ d&=\begin{bmatrix}0.2 & 0\\ 0 & 0.2\end{bmatrix}\end{aligned} \tag{16}$$

Our goal is to design a robust optimal tracker controller so that q_1, q_2, $\dot{q}_1$ and $\dot{q}_2$ can trace $\tilde{y}_1$, $\tilde{y}_2$, $\dot{\tilde{y}}_1$ and $\dot{\tilde{y}}_2$ respectively, so must be define a proper function benchmark function to minimize it, and the system can resistant to uncertainties.

Select a state vector as follows:

$$\xi=\begin{bmatrix}\xi_1 & \xi_2 & \xi_3 & \xi_4\end{bmatrix}^T=\begin{bmatrix}q_1 & \dot{q}_1 & q_2 & \dot{q}_2\end{bmatrix}^T \tag{17}$$

Then define the following equation:

$$E=\xi-\tilde{y}=\begin{bmatrix}q_1-\tilde{y}_1 & \dot{q}_1-\dot{\tilde{y}}_1 & q_2-\tilde{y}_2 & \dot{q}_2-\dot{\tilde{y}}_2\end{bmatrix}^T \tag{18}$$

According to (15), τ is input and consider $N(q)$ to zero.

$$q=[q_1,q_2]^T, \quad \begin{cases}x_1=q_1\\ x_2=\dot{q}_1\\ x_3=q_2\\ x_4=\dot{q}_2\end{cases} \tag{19}$$

According to (15) and (19), is obtained (20).

$$\ddot{q}+\frac{C}{M}\dot{q}=\frac{1}{M}(\tau-d)\rightarrow\ddot{q}+a\dot{q}=b \tag{20}$$

$$\begin{cases}\dot{x}_1=\dot{q}_1=x_2\\ \dot{x}_2=\ddot{q}_1\\ \dot{x}_3=\dot{q}_2=x_4\\ \dot{x}_4=\ddot{q}_2\end{cases} \tag{21}$$

$$\begin{bmatrix}\ddot{q}_1\\ \ddot{q}_2\end{bmatrix}+\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}\dot{q}_1\\ \dot{q}_2\end{bmatrix}=\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}\tau_1\\ \tau_2\end{bmatrix} \tag{22}$$

$$\begin{cases}\ddot{q}_1=-\begin{bmatrix}a_{11} & a_{12}\end{bmatrix}\begin{bmatrix}\dot{q}_1\\ \dot{q}_2\end{bmatrix}+\begin{bmatrix}b_{11} & b_{12}\end{bmatrix}\begin{bmatrix}\tau_1\\ \tau_2\end{bmatrix}\\ \ddot{q}_2=-\begin{bmatrix}a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}\dot{q}_1\\ \dot{q}_2\end{bmatrix}+\begin{bmatrix}b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}\tau_1\\ \tau_2\end{bmatrix}\end{cases} \tag{23}$$

From (21) and (23) is obtained (24).

$$\begin{aligned}\dot{X}=AX+BU=&\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & a_{11} & 0 & a_{12}\\ 0 & 0 & 0 & 1\\ 0 & a_{21} & 0 & a_{22}\end{bmatrix}X\\ &+\begin{bmatrix}0 & 0\\ b_{11} & b_{12}\\ 0 & 0\\ b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}\tau_1\\ \tau_2\end{bmatrix}\end{aligned} \tag{24}$$

6. SIMULATION RESULTS AND THEIR INTERPRETATIONS PARAMETERIZATION

In this section, the effectiveness of the proposed method is verified by simulating the robot model.

Figure 2 shows control inputs applied to the robot with minimal effort to achieve our goal. After reaching the goal, the control effort becomes zero. q_1 and q_2 are the first and second link move that are entered as the desired values inputs and want to be traced and according to Figures 3 and 4, they are tracked. According to Figure 5 is shown that with minimum control effort, system state variables converge to zero. The purpose of this method is to find control inputs which by applying it to the system (17),

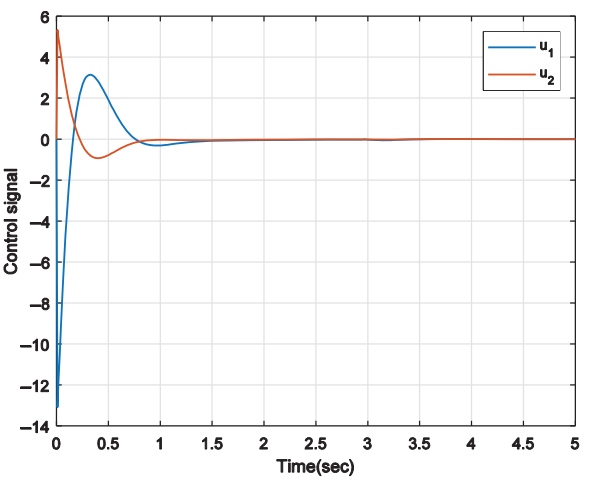


Fig. 2. Control inputs applied to the links in SDRE simulation.

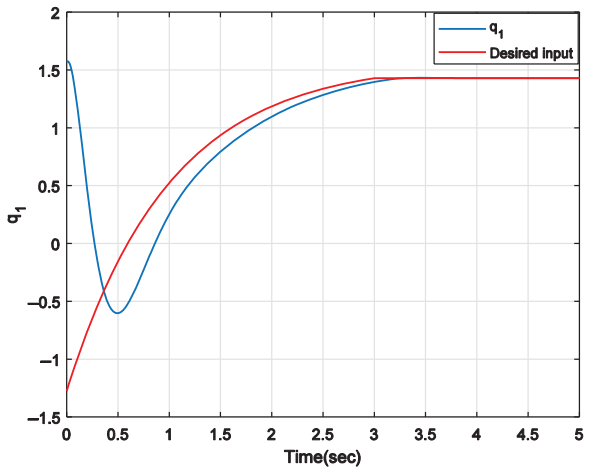


Fig. 3. The response tracking positioning q_1 for the first link with SDRE method.

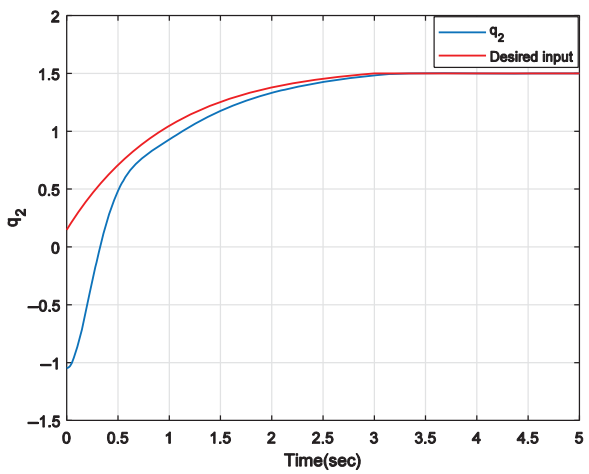


Fig. 4. The response tracking positioning q_2 for the second link with SDRE method.

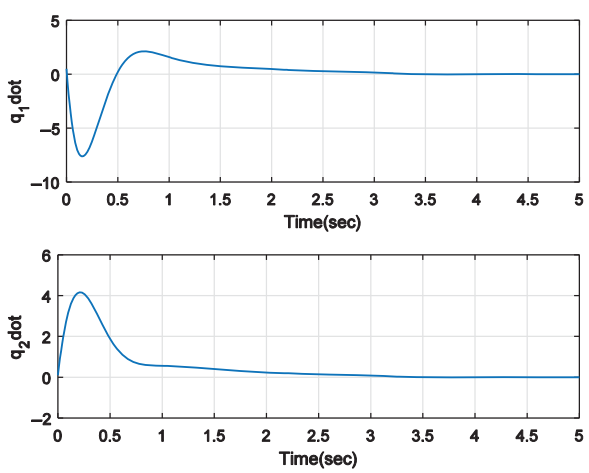


Fig. 5. The response tracking speed $\dot{q}_1$ for the first link and speed $\dot{q}_2$ for the second link with SDRE method.

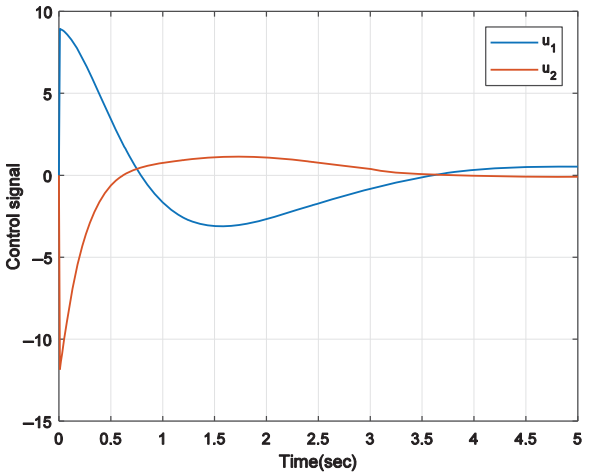


Fig. 6. Control inputs applied to the links in LQR simulation.

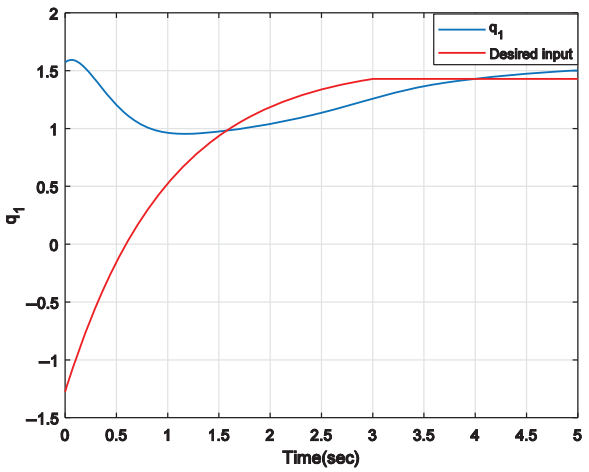


Fig. 7. The response tracking positioning q_1 for the first link with LQR method.

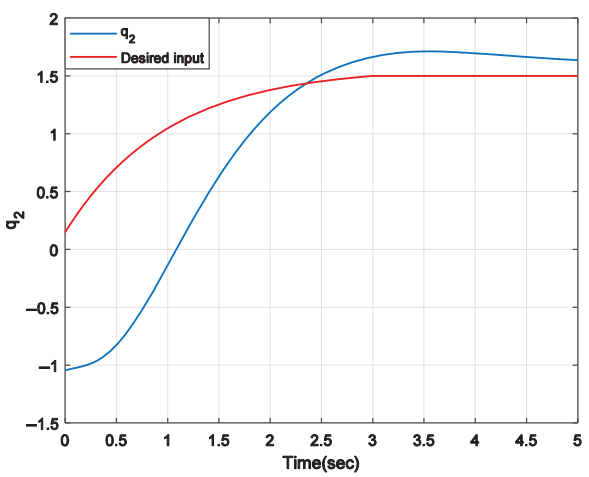


Fig. 8. The response tracking positioning q_2 for the second link with LQR method.

RESEARCH ARTICLE

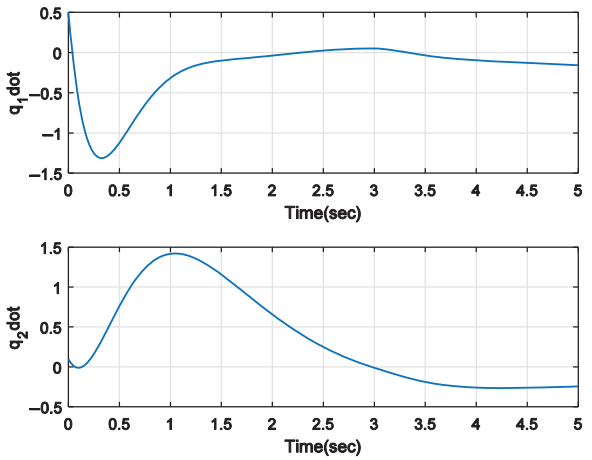


Fig. 9. The response tracking speed $\dot{q}_1$ for the first link and speed $\dot{q}_2$ for the second link with LQR method.

while stabilizing the system and satisfy the constraints, the defined cost function (15) become minimized, and, moreover, and when the control effort become zero, the state variable also gets zero; the simulation figures shows that our SDRE method has good performance.

Figures 7 and 8 obtained of the LQR method in which a nonlinear system has been linearized around the vector $X_v = [1, 1, 1, 1, 1]^T$ and initial conditions. Figure 6 shows the control inputs cannot become zero because according to Figures 7 and 8; the robot cannot trace the path. Also, Figure 9, shows the state variables of system do not converge to zero. So the result is that with SDRE method where the matrix depends on the state the answer is better.

7. CONCLUSIONS

In this paper, SDRE approach for the two-link robot was introduced for solving the path tracking problem. Given that the equation of motion of a two-link robot is nonlinear, so using linear control methods does not give a satisfactory answer. To solve this problem is used the Suboptimum methods. SDRE method is used for this paper as a suboptimum method. In the end, from the simulation results is concluded that the proposed solution SDRE method is better than the linear LQR method. I proposed that in the future, use a robust controller, such as a sliding mode method in combination with SDRE method for Disturbance rejection.

References

1. R. Iraji and S. M. Ghadami, *Science International* 27, 1183 **(2015)**.
2. Q. Jia, Y. Liu, G. Chen, and H. Sun, *The Open Mechanical Engineering Journal* 9, 992 **(2015)**.
3. F. Sağlam and Y. S. Ünlüsoy, *Journal of Automation and Control Research* 1, 1 **(2014)**.
4. A. Khamis and D. S. Naidu, *International Journal of Control* 1, 1 **(2014)**.
5. S. R. Nekoo, *ISA Transactions* 52, 285 **(2013)**.
6. M. H. Korayem and S. R. Nekoo, *ISA Transactions* 57, 117 **(2015)**.
7. Y. L. Kuo and T. L. Wu, *J. Comput. Theoret. Nanosci.* 13, 1013 **(2016)**.
8. J. Slávka and S. Ján, Application of the state-dependent Riccati equation method in nonlinear control design for inverted pendulum systems, *IEEE 11th International Symposium on Intelligent Systems and Informatics* **(2013)**, pp. 26–28.
9. F. Liccardo, S. Strano, and M. Terzo, Optimal control using state-dependent Riccati equation (SDRE) for a hydraulic actuator, *World Congress on Engineering* **(2013)**, pp. 3–5.
10. A. Wang, H. Yu, and S. Cang, *Journal of the Franklin Institute* 354, 1 **(2016)**.
11. A. H. Korayem, M. I. Rahagi, H. Babaee, and M. H. Korayem, *Robotica* 35, 119 **(2017)**.
12. A. T. Azar, J. Kumar, V. Kumar, and K. P. S. Rana, Control of a two link planar electrically-driven rigid robotic manipulator using fractional order SOFC, *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics* **(2017)**, pp. 57–68.
13. K. Lochan, B. K. Roy, and B. Subudhi, *Robotics and Autonomous Systems* 97, 1 **(2017)**.
14. Ç. Tayfun, State-dependent riccati equation (SDRE) control: A survey, *Proceedings of the 17th World Congress the International Federation of Automatic Control*, Seoul, Korea **(2008)**, pp. 6–11.
15. P. K. Menon, T. Lam, L. S. Crawford, and V. H. L. Cheng, Real-time computational methods for SDRE nonlinear control of missiles, *American Control Conference* **(2002)**, pp. 1–17.
16. H. Phang and X. Yang, *Hindawi Publishing Corporation Journal of Applied Mathematics* 2013, 1 **(2013)**.

Received: 13 March 2018. Accepted: 8 April 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1495–1510, 2018

RESEARCH ARTICLE

Linear Instability of Water–Oil Electrohydrodynamic Nanofluid Layers: Analytical and Numerical Study

Galal M. Moatimid and Mohamed A. Hassan*

Department of Mathematics, Faculty of Education, Ain Shams University, Roxy, Cairo, 11311, Egypt

The linear instability of two electrically conducting nanofluid layers (Water-Mineral oil) is investigated. Two cases are considered, the first case of the Water Al_2O_3-Oil CuO with its physical values for the nanofluid parameters. The second case considers Water Al_2O_3-Oil TiO_2 nanofluid layers. The nanofluid model includes both the Brownian motion and thermophoresis effects which are discussed due to the nanoparticels volume fraction percent and the temperature difference at the basic state. All properties of the nanofluids are considered variables. Firstly the stability is discussed analytically by obtaining the solutions of the linearized equations of motion with the equations of heat and nanoparticles volume fraction with employing the normal modes technique. The continuity of the normal stress tensor at the interface is applied to produce an implicit dispersion relation between the growth rate and the wave number. Secondly, the stability analysis is studied numerically depending on the relation of the energy growth function with time. The numerical results support the analytical stability results. The streamlines contours are presented graphically in the two nanofluid layers. The results illustrate that the Water Al_2O_3-Oil CuO interface is more stable than Water Al_2O_3-Oil TiO_2 nanofluid interface. Also, the increase of the temperature difference destabilizes the interface due to the increase of particles motion. The numerical results illustrate that the greater energy growth function occurs at the Oil TiO_2 layer which is the most unstable nanofluid layer.

Keywords: Interfacial Linear Instability, Nanofluid, Electrohydrodynamic, Normal Modes Analysis, Numerical Analysis.

1. INTRODUCTION

Interfacial instability of two fluid layers has many physical applications. The physical applications include the flow of air over oceans and seas, the stability of the solar wind-magnetosphere interface, oil and water exploration industry, dynamo generation of cosmic magnetism and the interaction between two streams of different velocities in the solar wind. Hsieh[1] established the earliest work in studying the interfacial instability with heat and mass transfer. Many other different physical effects such as the electric field, the flow through porous structures and the non-Newtonian property were studied. The electrohydrodynamic linear stability of finitely conducting fluid through porous media with mass and heat transfer was investigated by Moatimid and Obied Allah.[2] The two-layer instability of electrified Maxwell fluid with the heat and mass transfer was studied by Zakaria et al.[3] It was found that the Reynolds number has a destabilizing role in the stability criteria. Moatimid[4] studied the nonlinear instability of two dielectric viscoelastic fluids.

Most of the previous studies of the two-layer instability considered the Newtonian or non-Newtonian fluids only. Recently, there is a subclass of fluids which has many industrial and physical applications that is the nanofluids. Nanofluids consist of a base fluid (Newtonian or non-Newtonian) with nanoparticles (1–100 nm). The electric chip cooling and coolants for advanced nuclear systems are examples of the industrial applications of the nanofluids. In studying the instability, especially for the nanofluids, there exists an interaction between the thermal instability and the hydrodynamic perturbations. So, the thermal effects must be considered in the instability study. Manosh et al.[5] studied the free convective boundary layer flow instability over a vertical heated flat plate, with respect to two-dimensional wave disturbances. Lobov and Lyubimova[6] studied the stability of Soret-induced flow in a vertical layer. Nield and Kuznetsov[7] studied the linear stability analysis for the Newtonian nanofluid horizontal layer. In the previous articles, the study considers the flow of one layer nanofluid. To the best of the authors' knowledge, there are not enough studies which consider the stability of two superposed nanofluid layers.

*Author to whom correspondence should be addressed.

Physically, the heat and mass transfer across the interface, especially for the nanofluid layers, plays an essential role in determining the flow field. Especially, in the boiling process, the motion of the film and bubbles is mainly dependent on the effect of mass and heat transfer. Many types of contacting equipment, e.g., boilers, condensers, evaporators, gas absorbers, pipelines, chemical reactors and nuclear reactors are examples for the instability which includes the heat and mass transfer processes between two superposed fluid layers. Moatimid[8] studied the instability of two rotating magnetic fluid columns in the presence of mass and heat transfer. It is observed that the mass and heat transfer coefficients play an important role in the nonlinear stability of the system. The study of stability with heat and mass transfer was considered by many authors. The viscous potential theory was used to study the nonlinear Kelvin-Helmholtz instability of the interface between two viscous fluids with considering the mass and heat transfer across the interface by Awasthi et al.[9] The authors deduced that the heat and mass transfer has a destabilizing effect. Tiwari et al.[10] established a linear analysis for the capillary instability with heat and mass transfer across the interface. This study showed that the axial magnetic field, the heat, and mass transfer have a stabilizing effect. The problem of thermal instability and heat transfer of viscoelastic fluids was discussed by Jun et al.[11] In all previous references, also many others, the instability of two nanofluid layers was not considered.

Ferromagnetic liquids are homogenous colloidal suspensions of ferromagnetic particles in a light liquid. This construction of fluid is recently studied as the nanofluids. Magnetic fluids are a special category of smart nanomaterials, in particular magnetically controllable nanofluids.[12] These nanomaterials manifest simultaneously fluid and magnetic properties. So, in studying the nanofluid, which contains nano-metallic suspended particles, it is convenient to consider the electromagnetic properties of the nanofluid. Zakaria and Gamiel[13] studied the temporal stability properties of a falling liquid film down an inclined plane in the presence of a constant electromagnetic field. It is found that the electric field destabilizes the film flow whereas the magnetic field stabilizes it. The electrohydrodynamic instability of the interface between two viscous fluids with different electrical properties, subjected to a normal electric field, was studied by Ozen et al.[14] The authors deduced that the electric field has a dual effect. Moatimid[15] studied the stability of two rotating magnetic fluid columns with mass and heat transfer. The method of multiple scales was applied to determine the necessary and sufficient conditions for the stability. Also, Moatimid and Obied Allah[2] deduced that the mass and heat transfer parameter, as well as Darcy's coefficients, play a stabilizing role in the stability picture. The thermophoresis and Brownian motion effects on enhancement the nanofluid heat transfer were studied by Haddad et al.[16] The natural convection heat transfer of Water-CuO nanofluid was considered using the Rayleighe Bénard problem.

In the present study, two important physical factors that affect the interfacial instability are combined. The first one is the heat and mass transfer across the interface due to the Brownian motion and thermophoresis of the nanoparticles. The second factor is the electromagnetic property that arises due to the existence of the metallic suspended nanoparticles in the fluid layers. So, the instability of two superposed electrically conducting nanofluid layers (Water–Oil) with nanoparticles Al_2O_3, CuO and TiO_2 is studied. The two layers are subjected to a longitudinal external electric field with constant strength. This problem finds its importance in the engineering applications especially in petroleum applications and in the nanofluid solar pond.

The study is organized as follows: In Section 2, the basic equations that govern the motion with its physical nanofluid physical parameters are presented. The interfacial conditions together with the linear boundary conditions are presented in Section 3. The solutions of the linear equations of motion, heat and mass transfer are presented in Section 4. The numerical study of the instability due to the growth of the energy function is presented in Section 5. The results of the present study due to the analytical and numerical studies of the instability are discussed in Section 6. Also, the discussion of the streamlines contours is presented in Section 7. Finally, concluding remarks for the important results are presented in Section 8.

2. FORMULATION OF THE PROBLEM

Consider the flow of two parallel viscous nanofluid layers. The lower layer is Water Al_2O_3 nanofluid, while the upper layer is the mineral Oil with nanoparticles of CuO or TiO_2. The statically stable situation is considered, so the upper fluid (mineral oil) is lighter than the lower one (water layer). The subscripts (1) and (2) refer to the lower and upper fluid, respectively. The two fluids are incompressible and have uniform densities and viscosities ρ_{f1}, μ_{f1}, ρ_{f2}, μ_{f2} with the nanoparticles with densities ρ_{p1}, ρ_{p2} for the lower and the upper layers, respectively. The interface between the two nanofluids is initially flat located at $y = 0$. Also, it is considered that the two fluids are initially quiescent with zero streaming velocities. The two nanofluid layers are electrically conducting with dielectric constants ε_1 and ε_2. The gravitational acceleration g acts vertically downward. The fluids are influenced by the longitudinal external electric field of intensity E_0. The bounding surfaces are taken as $y = -h_1$ and $y = h_2$ which are kept at the uniform temperatures T_{10} and T_{20}, and constant nanoparticle fraction ϕ_{10} and ϕ_{20}, respectively. This physical situation is shown in Figure 1.

The interface function becomes after making a small disturbance,

$$F(x, y, t) = y - \eta(x, t) \tag{1}$$

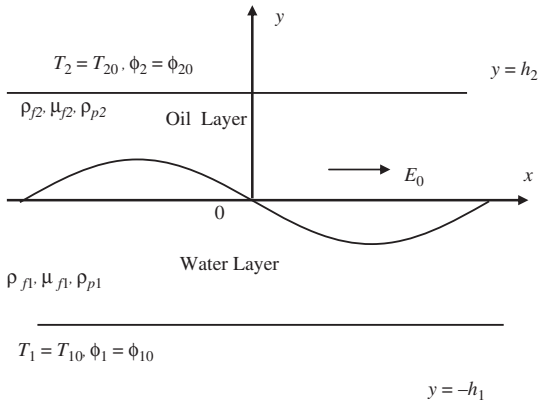


Fig. 1. Physical model and coordinates.

So, the unit outward normal vector in the linear form is given by

$$\hat{n} = -\frac{\partial \eta}{\partial x} e_x + e_y \tag{2}$$

where, e_x and e_y are the unit vectors along the x and y directions, respectively.

The basic equations of the system are as follows:[16]

1—The continuity equation:

$$\frac{\partial u_j}{\partial x} + \frac{\partial v_j}{\partial y} = 0 \tag{3}$$

2—The momentum equations:

$$\rho_{nfj}\left(\frac{\partial u_j}{\partial t} + u_j\frac{\partial u_j}{\partial x} + v_j\frac{\partial u_j}{\partial y}\right) = -\frac{\partial P_j}{\partial x} + \frac{\partial}{\partial x}\left(2\mu_{nfj}\frac{\partial u_j}{\partial x}\right) + \frac{\partial}{\partial y}\left(\mu_{nfj}\left(\frac{\partial v_j}{\partial x} + \frac{\partial u_j}{\partial y}\right)\right) \tag{4}$$

$$\rho_{nfj}\left(\frac{\partial v_j}{\partial t} + u_j\frac{\partial v_j}{\partial x} + v_j\frac{\partial v_j}{\partial y}\right) = -\frac{\partial P_j}{\partial y} + \frac{\partial}{\partial y}\left(2\mu_{nfj}\frac{\partial v_j}{\partial y}\right) + \frac{\partial}{\partial x}\left(\mu_{nfj}\left(\frac{\partial v_j}{\partial x} + \frac{\partial u_j}{\partial y}\right)\right) - \rho_{nfj} g \tag{5}$$

3—The heat transfer equation:

$$(\rho c_p)_{nfj}\left(\frac{\partial T_j}{\partial t} + u_j\frac{\partial T_j}{\partial x} + v_j\frac{\partial T_j}{\partial y}\right) = \frac{\partial}{\partial x}\left(k_{nfj}\frac{\partial T_j}{\partial x}\right) + \frac{\partial}{\partial y}\left(k_{nfj}\frac{\partial T_j}{\partial y}\right)(\rho c_p)_{pj} \times \left(D_{Bj}\left(\frac{\partial T_j}{\partial x}\frac{\partial C_j}{\partial x} + \frac{\partial T_j}{\partial y}\frac{\partial C_j}{\partial y}\right) + \frac{D_{Tj}}{T_b}\left(\left(\frac{\partial T_j}{\partial x}\right)^2 + \left(\frac{\partial T_j}{\partial y}\right)^2\right)\right) \tag{6}$$

4—The equation of the volumetric nanoparticle fraction:

$$\frac{\partial \phi_j}{\partial t} + u_j\frac{\partial \phi_j}{\partial x} + v_j\frac{\partial \phi_j}{\partial y} = \frac{\partial}{\partial x}\left(D_{Bj}\frac{\partial \phi_j}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_{Bj}\frac{\partial \phi_j}{\partial x}\right) + \frac{\partial}{\partial x}\left(\frac{D_{Tj}}{T_b}\frac{\partial T_j}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{D_{Tj}}{T_b}\frac{\partial T_j}{\partial y}\right) \tag{7}$$

where, the suffix $j = 1, 2$ refers to the equations and the physical properties of the first and second mediums, respectively. ρ_{pj} is the particle density, ρ_{nfj} is nanofluid density, T_j is the fluid temperature, ϕ_j is the nanoparticle volume fraction, D_{Bj} is the Brownian diffusion coefficient, P_j is the pressure, D_{Tj} is the thermophoretic diffusion coefficient, $c_{p_{nfj}}$ and c_{pj} is the specific heat of the nanofluid and nanoparticle, respectively, k_{nfj} is the nanofluid thermal conductivity, T_b is the fluids temperature at initial basic state (which is assumed to be the same temperature in both layers at the initial time $t = 0$), $\underline{V} = ue_x + ve_y$ is the velocity vector components in the x and y directions, and t is the time.

Since the two fluids are considered to be electrically conducting. Also, it is assumed that there are no volume charges are presented in the bulk of the fluids and at the interface in the equilibrium state. Also, the quasi-static approximation is valid for the model at hand.[17] So, Maxwell's equations are

$$\nabla \times \underline{E} = 0, \quad \nabla \cdot \varepsilon \underline{E} = 0 \tag{8}$$

The electric field is a curl-free vector, because there are no surface charges, having an electric scalar potential Ω such that

$$\underline{E}_j = E_0 e_x - \nabla \Omega_j \tag{9}$$

So, the electric potential Ω_j satisfies Laplace's equation

$$\nabla^2 \Omega_j = 0 \tag{10}$$

The effective density of the nanofluid ρ_{nf} is given by[16, 18]

$$\rho_{nfj} = \lfloor \phi_j \rho_{pj} + (1 - \phi_j)\rho_{fj}(1 - \beta_j(T_j - T_b)) \rfloor \tag{11}$$

Where, β_j is the thermal volumetric expansion in the two media ($j = 1, 2$).

The nanofluid viscosity according to Crocione model.[19]

$$\mu_{nfj} = \mu_{fj}\frac{1}{1 - 34.87(d_p/d_{fj})^{-0.3}\phi^{1.03}} \tag{12}$$

Where μ_{fj} is the fluid viscosity in the two layers, d_p is the nanoparticle diameter and d_{fj} is the fluid particle diameter.

Also, the thermal conductivity of the nanofluid is given by[19]

$$k_{nfj} = k_{fj}\left(1 + 4.4 Re_j^{0.4} Pr_j^{0.66}\left(\frac{T}{T_{fr}}\right)^{10}\left(\frac{k_p}{k_{fj}}\right)^{0.03}\phi^{0.66}\right) \tag{13}$$

RESEARCH ARTICLE

Table I. Physical properties water–oil with different nanoparticles.[20–23]

Property	Mineral oil (paraffin oil)	Water	Al_2O_3	CuO	TiO_2
ρ (kg/m^3)	800	997.1	3970	6500	4250
c_p (J/kg·K)	2000	4179	765	535.6	686.2
k (W/m·K)	0.15	0.613	40	20	8.9538
β (1/K)	0.1	0.000267	0.000024	0.000051	0.000024
μ (kg/m·s)	0.1	0.000855	–	–	–
ε (–)	1.6	80.4	9.9	18.1	100

Where T_{fr} is the temperature of the freezing point of the base fluid, k_p is the thermal conductivity of the nanoparticles. The Reynolds is $Re=(\rho_f k_B T)/(3\pi\mu_f^2 l_f)$ and the Prandtl number is $Pr=\mu_f/(\rho_f\alpha_f)$. Where, $k_B=1.3807\times 10^{-23}$ J/K is the Boltzmann constant, $l_f=0.17$ nm is the mean free path of the fluid particles[16] and α_f is the thermal diffusivity of the fluid. Also, the heat capacitance of the nanofluid $(\rho c_p)_{nfj}$ is given as

$$(\rho c_p)_{nfj}=\lfloor(1-\phi_j)(\rho c_p)_{fj}+\phi_j(\rho c_p)_p\rfloor \tag{14}$$

The Brownian diffusion coefficient D_{Bj} and the thermophoretic diffusion coefficient D_{Tj} are defined as

$$D_{Bj}=\frac{k_B T}{3\pi\mu_f d_p} \tag{15}$$

$$D_{Tj}=\left(\frac{\mu_f}{\rho_f}\right)\left(\frac{0.26k_f}{2k_f+k_p}\right)\phi \tag{16}$$

The physical values of the previous parameters for the Water–Oil system with different nanoparticles (Al_2O_3, CuO and TiO_2) are given in Table I.

The previous physical values in Table I will be considered during the graphical representation in the discussion section to study the effect of different types on nanoparticles on the stability behavior.

The boundary conditions for the problem will be displayed in the following section.

3. THE BOUNDARY AND INTERFACIAL CONDITIONS

The boundary conditions must be well defined in order to solve the system of the previous governing Eqs. (3)–(7). The conditions of the problem are classified into four categories. The first one is the kinematic conditions at the upper plate $y=h_2$ and at the lower plate $y=-h_1$ and at the interface $y=\eta(x,t)$, where η is the disturbed interface height measured with respect to the initial level $y=0$. The second considers the thermal and volumetric fraction conditions. The third is Maxwell's conditions of the electric field. The last one is the stress balance conditions at the interface. These boundary constraints may be summarized as follows:

(i) *The Kinematics Boundary Conditions*

The kinematic boundary conditions, express the fact that the interface has always comprised the same fluid particles. Therefore, the function $\eta(x,t)$ satisfies

$$\left\|\rho_{nfj}\left(\frac{\partial\eta}{\partial t}+\underline{V}_j\cdot\nabla(y-\eta)\right)\right\|=0 \tag{17}$$

where $\|f\|=f_2-f_1$ and the notation denotes to the subtraction of the function f at the upper and lower fluid layers, respectively.

The tangential components of the velocity vector are continuous at the interface

$$\hat{t}\cdot\|u_j\|=0,\quad \text{at } y=\eta(x,t) \tag{18}$$

Also, the normal components of the velocity vector are continuous at the interface

$$\hat{n}\cdot\|v_j\|=0\quad \text{at } y=\eta(x,t) \tag{19}$$

Where $\hat{n}$ is the unit outward normal vector at the interface which is defined in Eq. (2) and $\hat{t}=e_x+(\partial\eta/\partial x)e_y$ is the unit tangent vector.

Since the lower and the upper plates are rigid; the velocity components have zero values in it. So, the boundary conditions at the boundaries are given as

$$u_1=0,\quad v_1=0,\quad \text{at } y=-h_1 \tag{20}$$

$$u_2=0,\quad v_2=0,\quad \text{at } y=h_2 \tag{21}$$

(ii) *The Thermal and Volumetric Fraction Boundary Conditions*

The temperature and the normal heat flux are continuous across the interface. So, the thermal interfacial conditions may be written as

$$\|T_j\|=0,\quad \left\|k_{nfj}\frac{\partial T_j}{\partial y}\right\|=0\quad \text{at } y=\eta(x,t) \tag{22}$$

Also, the temperatures at the plats are constants as given in the problem formulation

$$T_1=T_{10},\quad T_2=T_{20},\quad \text{at } y=-h_1 \text{ and } y=h_2 \tag{23}$$

Similar boundary and interfacial conditions are held for the volumetric fraction, so

$$\|\phi_j\|=0,\quad \left\|D_{Bj}\frac{\partial\phi_j}{\partial y}\right\|=0\quad \text{at } y=\eta(x,t) \tag{24}$$

$$\phi_1=\phi_{10},\quad \phi_2=\phi_{20},\quad \text{at } y=-h_1 \text{ and } y=h_2 \tag{25}$$

(iii) *The electric Potential Conditions* (*Maxwell's conditions*)

The normal electric displacement must be continuous across the interface, so the interfacial condition for the normal electric field displacement yields

$$-\frac{\partial\eta}{\partial x}\left\|\varepsilon_j\left(E_0-\frac{\partial\Omega_j}{\partial x}\right)\right\|-\left\|\varepsilon_j\frac{\partial\Omega_j}{\partial y}\right\|=0 \tag{26}$$

Also, the tangential electric filed components are continuous at the interface

$$\frac{\partial \eta}{\partial x}\left\| \frac{\partial \Omega_j}{\partial y} \right\| + \left\| \frac{\partial \Omega_j}{\partial x} \right\| = 0 \tag{27}$$

Since the bounding plates are earthed the electric potential at the boundaries satisfies

$$\frac{\partial \Omega_1}{\partial y} = 0 \quad \text{and} \quad \frac{\partial \Omega_2}{\partial y} = 0, \quad \text{at } y = -h_1 \text{ and } y = h_2 \tag{28}$$

(iv) *The Balance of the Stress Tensor Components*

The continuity of the tangential and normal components of the stress tensor must be satisfied at the interface. So, the continuity of the tangential stress tensor yields

$$\hat{n} \cdot \| -P_j I + \tau_{mn}^{(j)} \| \cdot \hat{t} = 0, \quad \text{at } y = \eta(x,t) \tag{29}$$

Where $\tau_{mn}^{(j)}$ is the viscous stress tensor which is defined as

$$\tau_{mn}^{(j)} = \mu_{nfj}\left(\frac{\partial u_m^j}{\partial x_n} + \frac{\partial u_n^j}{\partial x_m} \right) \tag{30}$$

Meanwhile, the continuity of the normal stress tensor across the interface requires

$$\begin{aligned} &\rho_{nf1}(\underline{V}_1 \cdot \nabla F)\left(\frac{\partial F}{\partial t} + \underline{V}_1 \cdot \nabla F \right) \\ &= \rho_{nf2}(\underline{V}_2 \cdot \nabla F)\left(\frac{\partial F}{\partial t} + \underline{V}_2 \cdot \nabla F \right) + \left(\| P_j - 2\mu_{nfj} \hat{n} \cdot \nabla \otimes \underline{V}_j \cdot \hat{n} \| \right. \\ &\left. - \frac{1}{2} \| \varepsilon_j (E_n^2 - E_t^2)_j \| + \sigma \nabla \cdot \hat{n} \right) |\nabla F|^2, \quad \text{at } y = \eta(x,t) \end{aligned} \tag{31}$$

Where σ is the surface tension parameter which is considered to be constant and the product sign "$\otimes$" refers to the multiplication of differential operators.

4. THE ANALYTICAL STABILITY ANALYSIS

In this section, the basic equations of the system will be solved. The solution will be based on the normal modes analysis. So, the functions of the problem take the forms

$$\left.\begin{aligned} &\eta(x,t) = A e^{ikx+\omega t} + c.c., \\ &T_j(x,y,t) = T_{bj} + T_j^+(y) e^{ikx+\omega t} + c.c., \\ &\phi_j(x,y,t) = \phi_{bj} + \phi_j^+(y) e^{ikx+\omega t} + c.c., \\ &\psi_j(x,y,t) = \psi_j^+(y) e^{ikx+\omega t} + c.c., \\ &\Omega_j(x,y,t) = \Omega_j^+(y) e^{ikx+\omega t} + c.c. \end{aligned}\right\} \tag{32}$$

Where T_b, ϕ_b and P_b represent the initial basic values of the temperature, volumetric fraction, and the pressure, respectively. The notation "$c.c.$" refers to the complex conjugate of the functions. A is the initial amplitude of the applied perturbation, k is the wave number of the disturbance, which is assumed to be real and positive, $\omega = \omega_r + i\omega_i$ is the growth rate of the wave motion which determines the stability behavior of the system. Physically, when $\omega_r > 0$, the disturbance grows and the system becomes unstable, when $\omega_r < 0$, the disturbance decays and the system become stable and when $\omega_r = 0$ the system is neutrally stable.

Also, the stream function of the velocity ψ_j satisfies the continuity equation by the assumptions

$$u_j = \frac{\partial \psi_j}{\partial y} \quad \text{and} \quad v_j = -\frac{\partial \psi_j}{\partial x} \tag{33}$$

The two fluid layers are assumed to be at rest in the basic state, so the stream functions are zeros in the basic state of the two fluid layers.

Before solving the equations of the problem, it is convenient to rewrite the basic equations in dimensionless forms. This can be done by special choice of a characteristic mass, length and time. Consider the following dimensionless forms depending on the characteristic length $= h = (h_1 + h_2)$, which represents the distance between the two plates, the characteristic time $= h^2/\alpha_{f0}$ (where α_{f0} is the fluid thermal diffusivity at the reference temperature T_b) and the characteristic mass $= \sigma h/g$. The other dimensionless quantities are given by

$$\left.\begin{aligned} &u^* = u\frac{h}{\alpha_{f0}}, \quad v^* = v\frac{h}{\alpha_{f0}}, \quad t^* = t\frac{\alpha_{f0}}{h^2}, \\ &\rho^* = \frac{\rho_{nf}}{\rho_f}, \quad \mu^* = \frac{\mu_{nf}}{\mu_f}, \quad x^* = x\frac{1}{h}, \\ &y^* = y\frac{1}{h}, \quad E_0^{*2} = E_0^2\frac{h}{\sigma} \quad k^* = kh, \\ &\omega^* = \omega\frac{h^2}{\alpha_{f0}}, \quad T^* = \frac{T - T_b}{\Delta T}, \quad \phi^* = \frac{\phi}{\phi_b} \end{aligned}\right\} \tag{34}$$

The nanofluid physical parameters Eqs. (11)–(16) are substituted into the system of Eqs. (3)–(7) and the resulting equations are linearized. So, the resulting linear equations contain these parameters in its basic forms. In other words, the nanofluid parameters Eqs. (11)–(16) will depend on the basic values of the volume fraction ϕ_b and the basic temperature T_b. So, the stability will be studied due to the variability of basic value the volume fraction ϕ_b and the basic temperature T_b. Hence, the basic Eqs. (4)–(7) in the linearized dimensionless form, after neglecting the pressure between Eqs. (4) and (5), may be written as

$$\frac{\rho_j^*}{Pr_j}\frac{\partial}{\partial t^*}(\nabla^2 \psi_j^*) = \mu_j^*(\nabla^4 \psi_j^*) + R_{Nj}\frac{\partial \phi_j^*}{\partial x} - R_{aj}\frac{\partial T_j^*}{\partial x} \tag{35}$$

$$\frac{\partial T_j^*}{\partial t^*} = \alpha_{fj}^*\left(\frac{\partial^2 T_j^*}{\partial x^{*2}} + \frac{\partial^2 T_j^*}{\partial y^{*2}} \right) \tag{36}$$

$$\frac{\partial \phi_j^*}{\partial t^*} = \frac{1}{Le_j}\left(\frac{\partial^2 \phi_j^*}{\partial x^{*2}} + \frac{\partial^2 \phi_j^*}{\partial y^{*2}} \right) + \frac{N_{Aj}}{Le_j}\left(\frac{\partial^2 T_j^*}{\partial x^{*2}} + \frac{\partial^2 T_j^*}{\partial y^{*2}} \right) \tag{37}$$

Where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplace's operator.

It is important to note that the previous Eqs. (35)–(37) are in the linear dimensionless form. So, the nonlinear terms are omitted. Also, the nanofluid density in Eq. (11) is applied for the dilute nanofluid suspension (small value of the nanoparticle volume fraction ϕ) and for small temperature. Therefore, the term $\rho_f\ \beta_j\ \phi_j\ T_j$ can be excluded from Eq. (11). It follows the contribution of the term ρg in Eq. (5) yields the last two linear (first order) terms in Eq. (21).[18] The asterisk refers to the dimensionless quantity, from now on the asterisks will be omitted for simplicity. The other dimensionless parameters are defined as:

$$R_{Nj}=\frac{(\rho_{pj}-\rho_{fj})gh^3\phi_b}{\alpha_{f0}\mu_f},\ \text{is the concentration Rayleigh number.}$$

$$\alpha^*_{fj}=\frac{\alpha_{fnj}}{\alpha_{f0}},\ \text{is the thermal diffusivity ratio.}$$

$$R_{aj}=\frac{\rho_{fj}\beta_j g h^3\Delta T}{\alpha_{f0}\mu_f},\ \text{is the thermal Rayleigh number.}$$

$$Le_j=\frac{\alpha_{f0}}{D_{Bj}},\ \text{is the Lewis number.}$$

$$N_{Aj}=\frac{D_{Tj}\Delta T}{D_{Bj}T_b\phi_b},\ \text{is the modified diffusivity ratio.}$$

To study the stability of the present problem, the general solution of the governing Eqs. (35)–(37) with the boundary and interface conditions is obtained. The solution technique starts by finding the Eigenfunctions of the heat distributions $T_j^+(y)$ from Eq. (36), the subscript plus sign will be omitted for simplicity. To do this, the normal modes form in Eq. (32) is applied with Eq. (36). Then the heat equation becomes

$$\frac{d^2T_j}{dy^2}-\left(\frac{\omega}{\alpha_{fj}}+k^2\right)T_j=0 \tag{38}$$

The general solution for the heat equation with applying the thermal boundary conditions Eqs. (22)–(23) may be written as:

$$T_1(y)=A_1e^{m_1y}+B_1e^{-m_1y} \tag{39}$$

$$T_2(y)=A_2e^{m_2y}+B_2e^{-m_2y} \tag{40}$$

Where, the constants A_1, A_2, B_1, B_2, m_1 and m_2 (also, all coming constants) are defined in the Appendix.

By the same manner, the volumetric fraction equation (Eq. (37)) takes the form

$$\frac{d^2\phi_j}{dy^2}-(k^2+\omega Le_j)\phi_j=N_{Aj}(k^2-m_j^2)T_j \tag{41}$$

The general solution of Eq. (41) can be determined, with the help of Eqs. (39), (40) and with applying the volume fraction condition Eqs. (24), (25), in the form

$$\phi_1(y)=A_3e^{m_3y}+B_3e^{-m_3y}+Q_1T_1 \tag{42}$$

$$\phi_2(y)=A_4e^{m_4y}+B_4e^{-m_4y}+Q_2T_2 \tag{43}$$

The normal modes form in Eq. (32) is applied with Eq. (35) to give

$$\frac{d^4\psi_j}{dy^4}-(k^2+n_j^2)\frac{d^2\psi_j}{dy^2}+k^2n_j^2\psi_j=\frac{ik}{\mu_j}(R_{Nj}\phi_j-R_{aj}T_j) \tag{44}$$

It is evident to mention here that, the previous equation (Eq. (44)) is the same that obtained in Ref. [3] but for $R_N, R_a=0$. This means that the Brownian motion and the thermophoresis effects were neglected in Ref. [3]. So, the present study is a generalization of the previous case.

The general solution of the stream functions in Eq. (44) takes the form

$$\psi_1(y)=A_5e^{ky}+B_5e^{-ky}+A_6e^{n_1y}+B_6e^{-n_1y}+A_7e^{m_3y}+B_7e^{-m_3y}+A_8e^{m_1y}+B_8e^{-m_1y} \tag{45}$$

$$\psi_2(y)=A_9e^{ky}+B_9e^{-ky}+A_{10}e^{n_2y}+B_{10}e^{-n_2y}+A_{11}e^{m_4y}+B_{11}e^{-m_4y}+A_{12}e^{m_2y}+B_{12}e^{-m_2y} \tag{46}$$

The previous solution in Eqs. (45) and (46) contains eight arbitrary constants A_5, A_6, A_9, A_{10}, B_5, B_6, B_9, B_{10}. Meanwhile, the other constants are known and defined in the Appendix. So, we need eight conditions to determine the general form of the stream function. These conditions are the boundary conditions Eqs. (20), (21), the interface conditions Eqs. (18), (19) and the tangential stress balance condition Eq. (29). These conditions, in the linear form, become

$$\begin{aligned}
&\frac{\partial\psi_1}{\partial y}=0,\quad \psi_1=0,\quad \text{at } y=-h_1,\\
&\frac{\partial\psi_2}{\partial y}=0,\quad \psi_2=0,\quad \text{at } y=h_2,\\
&\frac{\partial\psi_1}{\partial y}=\frac{\partial\psi_2}{\partial y},\quad \psi_1=\psi_2,\quad \text{at } y=\eta,\\
&\rho_1(\omega A+ik\psi_1)=\rho_2(\omega A+ik\psi_2),\quad \text{at } y=\eta,\\
&\mu_1\left(\frac{d^2\psi_1}{dy^2}+k^2\psi_1\right)=\mu_2\left(\frac{d^2\psi_2}{dy^2}+k^2\psi_2\right),\quad \text{at } y=\eta
\end{aligned} \tag{47}$$

After using the previous conditions in Eq. (47), the general solutions for the stream function Eqs. (45), (46) are completely determined and the constants become well defined.

To study the stability, the electric potential function Ω_j must be defined. By employing the normal modes form in Eq. (32) with the Laplace's Eq. (10), and according to the boundary and interface electric conditions Eqs. (26)–(28) the general solution of the electric potential becomes

$$\Omega_1(y)=\frac{iAE_0(\varepsilon_2-\varepsilon_1)\sec hkh_1}{\varepsilon_2\tanh kh_2+\varepsilon_1\tanh kh_1}\cosh k(y+h_1) \tag{48}$$

$$\Omega_1(y)=\frac{iAE_0(\varepsilon_2-\varepsilon_1)\sec hkh_2}{\varepsilon_2\tanh kh_2+\varepsilon_1\tanh kh_1}\cosh k(y-h_2) \tag{49}$$

On the other hand, in view of Eqs. (4) and (5) the perturbed pressure may be determined from the following equation

$$P_j(y) = \left(ik\mu_j + \frac{i\rho_j\omega}{k}\right)\frac{d\psi_j}{dy} - \frac{i\mu_j}{k}\frac{d^3\psi_j}{dy^3} \tag{50}$$

The solution of the stream functions ψ_1 and ψ_2 are used in determining the perturbed pressure functions $P_1(y)$ and $P_2(y)$ from Eq. (50) for the first and second medium, respectively. Finally, the dispersion relation is presented by the normal stress balance condition Eq. (31). This condition may be written in the linear form as follows

$$P_2 - P_1 + 2ik\mu_2\frac{\partial\psi_2}{\partial y} - 2ik\mu_1\frac{\partial\psi_1}{\partial y} - \varepsilon_2 E_0\frac{\partial\Omega_2}{\partial x} + \varepsilon_1 E_0\frac{\partial\Omega_1}{\partial x} - \frac{\partial^2\eta}{\partial x^2} = 0 \tag{51}$$

So, the solution of the stream functions ψ_1 and ψ_2 in Eqs. (45), (46), the perturbed pressure P_1 and P_2 from Eq. (50), the electric potential functions Ω_1, Ω_2 from Eqs. (48) and (49) and the disturbed interface function η from Eq. (32) are used with the normal stress balance condition Eq. (51) to obtain the transcendental dispersion relation in the form

$$D(\omega, k) = 0 \tag{52}$$

The previous relation Eq. (52) represents the linear dispersion equation for the surface waves propagating through viscous nanofluids, which is a transcendental function. The roots of this dispersion relation will be determined numerically with the help of the *Mathematica* program. This dispersion relation is used to describe the relationship between perturbed growth rate and the wave number. As stated before, each negative value of the real part of ω (or the negative real root of ω) represents the stable state of the interface. Meanwhile, if the real root has a positive sign, the disturbances will grow with time which corresponds to the unstable state. This will discuss in details graphically in the next section.

5. THE NUMERICAL STABILITY ANALYSIS

The advantage of the numerical solution is that this solution is suitable for any form of the initial conditions for the perturbed functions $T_j(x, y, t)$, $\phi_j(x, y, t)$ and $\psi_j(x, y, t)$. So, the dependence of time is not restricted in the exponential form likewise the normal modes analysis. The stability analysis depends on the energy growth rate function $G(t)$ that will be obtained by the numerical solution. The departure of the curve indicates the unstable state.[24]

The numerical solutions of the system of partial differential equations (35)–(37), will be determined based on following form for the perturbed functions:

$$F(x, y, t) = \bar{F}(y, t)e^{ikx} \tag{53}$$

Where $\bar{F}(y, t)$ refers to the perturbed state of the functions $\bar{T}_j(y, t)$, $\bar{\phi}_j(y, t)$ and $\bar{\psi}_j(y, t)$. Then, the dependence on time is not restricted in the exponential form. So, Eqs. (35)–(37) will be transferred into the system of partial differential equations that depends on (y, t) only as:

$$\frac{\rho_j}{Pr_j}\frac{\partial}{\partial t}\left(-k^2\bar{\psi}_j + \frac{\partial^2\bar{\psi}_j}{\partial y^2}\right) = \mu_j\left(k^4\bar{\psi}_j - 2k^2\frac{\partial^2\bar{\psi}_j}{\partial y^2} + \frac{\partial^4\bar{\psi}_j}{\partial y^4}\right) + ikR_{Nj}\bar{\phi}_j - ikR_{aj}\bar{T}_j \tag{54}$$

$$\frac{\partial\bar{T}_j}{\partial t} = \alpha_{fj}\left(-k^2\bar{T}_j + \frac{\partial^2\bar{T}_j}{\partial y^2}\right) \tag{55}$$

$$\frac{\partial\bar{\phi}_j}{\partial t} = \frac{1}{Le_j}\left(-k^2\bar{\phi}_j + \frac{\partial^2\bar{\phi}_j}{\partial y^2}\right) + \frac{N_{Aj}}{Le_j}\left(-k^2T_j + \frac{\partial^2T_j}{\partial y^2}\right) \tag{56}$$

The finite difference method is applied to solve this system numerically. The partial derivatives of the stream functions $\bar{\psi}(y, t)$, $\bar{T}_j(y, t)$ and $\bar{\phi}_j(y, t)$ are obtained depending on the central finite difference. The resulting finite difference equations are solved for the range $-h_1 \to 0$ in the first medium (Water layer) and for the range $0 \to h_2$ in the second medium (Oil layer). The time step is taken as $\Delta t = 0.005$ and for the distance $\Delta y = 0.1$. The numerical solution yields the numerical values for the perturbed function $\bar{\psi}_j(y, t)$. Then the energy density function $Eng_j(t)$ is computed through both layers according to the form[24]

$$Eng_1(t) = \int_{-h_1}^{0}(|u_1(t)|^2 + |v_1(t)|^2)\,dy \tag{57}$$

$$Eng_2(t) = \int_{0}^{h_2}(|u_2(t)|^2 + |v_2(t)|^2)\,dy \tag{58}$$

and the energy growth function $G_j(t)$ (for the media $j = 1, 2$) can be defined as

$$G_j(t) = \frac{Eng_j(t)}{Eng_j(0)} \tag{59}$$

The stability analysis is discussed due to the graphical relation between the energy growth function $G_j(t)$ and the time. The departure of the growth function with time indicates the unstable behavior. The initial state for the perturbed functions is assumed to be quiescent or constant for all perturbed functions.

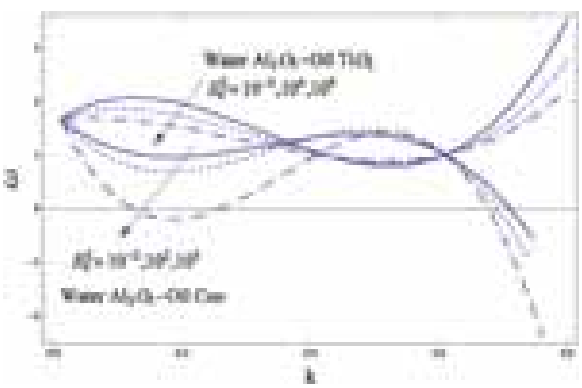


Fig. 2. Stability diagram for a system having the particulars: $T_1 = 300$, $T_2 = 350$, $h_1 = 5$, $h_2 = 3$, $C_1 = 0.01$, $C_2 = 0.03$, $\phi_{b1} = 0.01$, $\phi_{b2} = 0.03$.

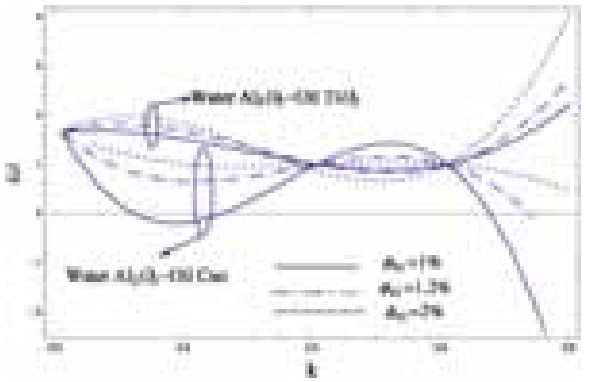


Fig. 3. Stability diagram for a system having the particulars: $T_{10} = 300$, $T_{20} = 350$, $h_1 = 5$, $h_2 = 3$, $C_{10} = 0.01$, $C_{20} = 0.03$, $E_0^2 = 10^5, 10^9$, $\phi_{b2} = 0.03$.

6. RESULTS AND DISCUSSION

6.1. For the Analytical Study by Normal Modes Method

This section is devoted to studying the stabilizing effect of the different parameters, especially the physical nanofluid parameters. The nanofluid parameters (11–16) depend, in its linear form, on the basic value of the volume fraction percent ϕ_b and the basic temperature value of the nanofluid layer T_b. So, we confined our attention to study the effects of these parameters beside the electric field intensity E_0. The instability of the interface depends on the relation between the growth rate ω and the wave number k. In all figures, the dispersion relation represents the stable state for the negative real roots of ω.

As seen in Figure 2, the stability curve damping corresponds to the increase of both the wave number k and the electric field intensity E_0^2. This means that the system is more stable for highly longitudinal electric field intensity and the electric field has a stabilizing effect. This is an early result which is obtained by many authors.[25, 26] Also, this figure illustrates that the Water Al_2O_3-Oil CuO is more stable than the Water Al_2O_3-Oil TiO_2. This can be explained due to the diversity of the nanofluid densities. From Table I it is clear that the density of the Oil CuO is greater than the Oil TiO_2 layer. So, we can conclude that increasing the density of the second medium stabilizes the interface. This means that the enhancement of density stabilizes the interface.

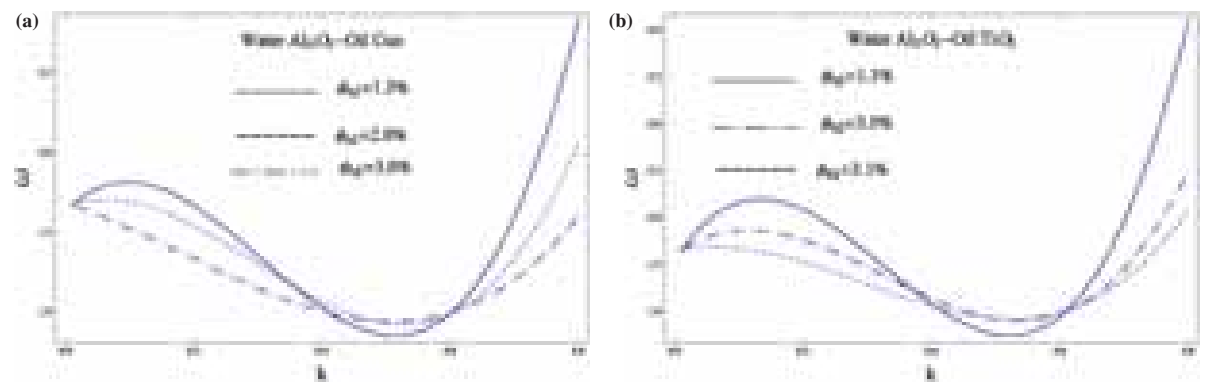


Fig. 4. (a) Stability diagram for a system having the particulars: $T_{10} = 300$, $T_{20} = 350$, $h_1 = 5$, $h_2 = 3$, $C_{10} = 0.01$, $C_{20} = 0.03$, $E_0^2 = 10^5$, $\phi_{b1} = 0.03$. (b) Stability diagram for a system having the particulars: $T_{10} = 300$, $T_{20} = 350$, $h_1 = 5$, $h_2 = 3$, $C_{10} = 0.01$, $C_{20} = 0.03$, $E_0^2 = 10^9$, $\phi_{b1} = 0.03$.

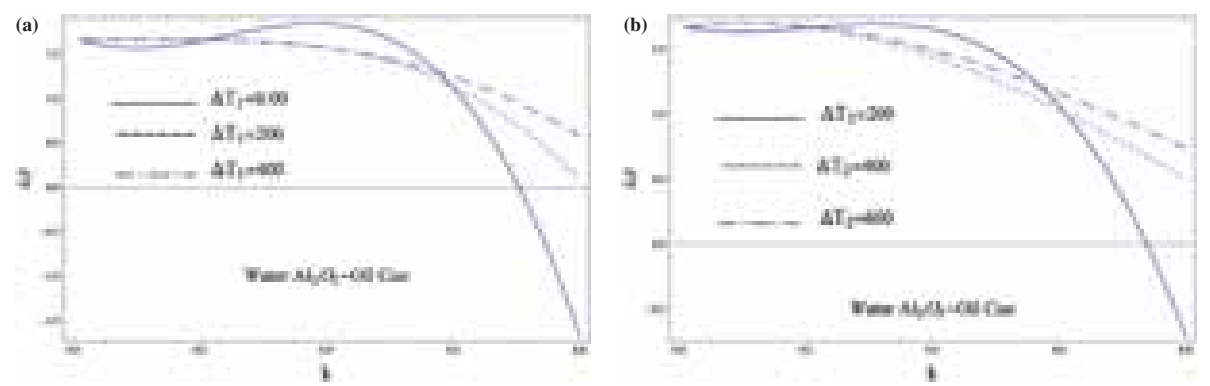


Fig. 5. (a) Stability diagram for a system having the particulars: $\Delta T_2 = 0$, $h_1 = 5$, $h_2 = 3$, $C_{10} = 0.01$, $C_{20} = 0.03$, $E_0^2 = 10^5$, $\phi_{b2} = 0.01$. (b) Stability diagram for a system having the particulars: $\Delta T_2 = 200$, $h_1 = 5$, $h_2 = 3$, $C_{10} = 0.01$, $C_{20} = 0.03$, $E_0^2 = 10^5$, $\phi_{b1} = 0.03$, $\phi_{b2} = 0.01$.

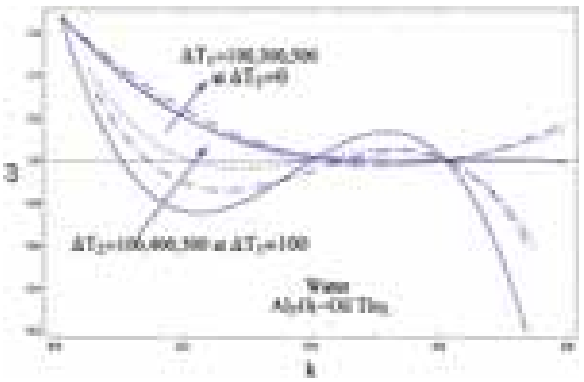


Fig. 6. Stability diagram for a system having the particulars: $h_1 = 5$, $h_2 = 3$, $C_{10} = 0.01$, $C_{20} = 0.03$, $E_0^2 = 10^9$, $\phi_{b1} = 0.03$, $\phi_{b2} = 0.031$.

Meanwhile Figure 3 depicts that the increase of the nanoparticle volume fraction percent $\phi_{b1} = 1, 1.2, 2\%$ destabilizes the interface for both Water–Oil systems. Increasing nanoparticles volume fraction percent in the base Water layer produces an increase in the Brownian motion. The Brownian motion is the random motion of particles suspended in a fluid (liquid or gas) resulting from their collision with the fast-moving atoms or molecules in the gas or liquid.[27] So, the existence of the nanoparticles increases the motion in the fluid medium which in turn destabilizes the interface. This physical situation implies that the nanoparticle volume fraction percent in the basic initial state has a destabilizing effect in both nanofluid Water–Oil layers as shown in Figure 3. But, the opposite effect is observed for the nanoparticle volume fraction percent in the Oil layer $\phi_{b2} = 1.3, 2, 3\%$ in Figures 4 and 5. This result seems to be in contradiction with the previous deduced result of Figure 3 for volume fraction percent of the Water layer. But, there is another important factor in studying the physics of the nanofluid.[27] This phenomenon is the aggregation of the nanoparticles. This means that for a higher density of the colloid particles ($\phi_{b2} = 1.3, 2, 3\%$ in the Oil layer) the particles agglomerate which reduces the Brownian motion. So, the increase of the nanoparticles density (the densities of the CuO TiO_2 nanoparticles are greater than the density of Al_2O_3 due to Table I) causes

Table II. Values of the growth rate ω at different values of the wave number k.

The nanofluid layers	The wave number k	The growth rate ω
Water Al_2O_3-Oil CuO	0.3	0.234349
	0.6	1.09642
	0.9	−9.76878
	1.2	−46.6343
Water Al_2O_3-Oil TiO_2	0.3	1.3782
	0.6	0.935476
	0.9	7.40256
	1.2	27.9265

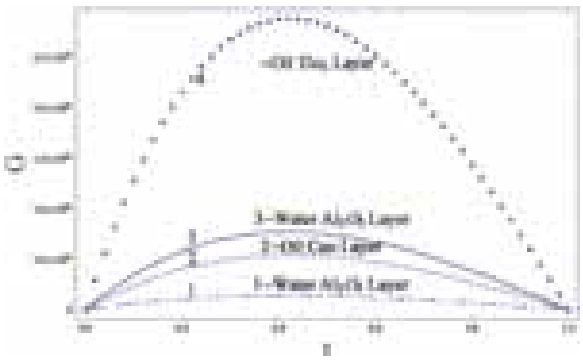


Fig. 7. Stability diagram for the energy distribution function with: $k = 3$, $\phi_{b1} = 0.08$, $\phi_{b2} = 0.05$, $C_{10} = 0.01$, $C_{20} = 0.03$, $T_{10} = 300$, $T_{20} = 350$, $h_1 = 5$, $h_2 = 5$, $E_0^2 = 10^5 - 10^9$.

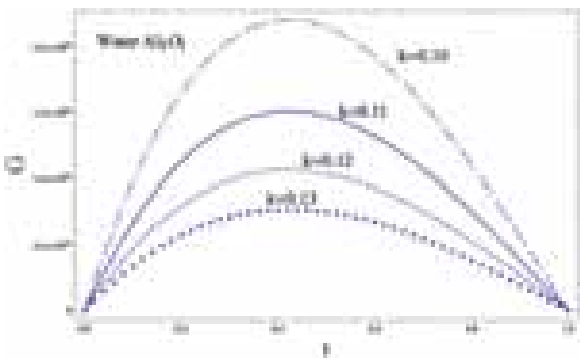


Fig. 8. Stability diagram for the energy distribution function with: $\phi_{b1} = 0.08$, $\phi_{b2} = 0.05$, $C_{10} = 0.01$, $C_{20} = 0.03$, $T_{10} = 300$, $T_{20} = 350$, $h_1 = 5$, $h_2 = 5$, $E_0^2 = 10^5$.

a reduction of the random movement of the particles and stabilizes the system as illustrated in Figures 4 and 5.

The second important factor in studying nanofluids is the thermophoresis effect. Thermophoresis is a phenomenon observed in mixtures of mobile particles

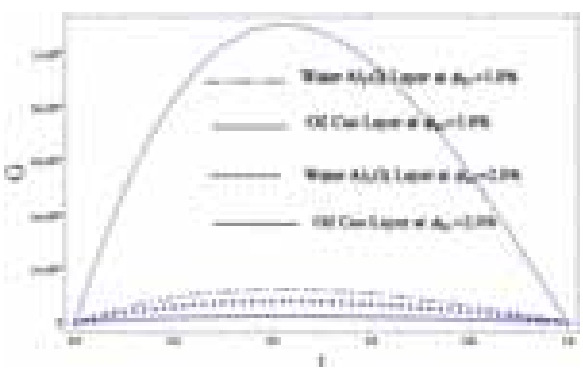


Fig. 9. Stability diagram for the energy distribution function with: $\phi_{b2} = 0.03$, $C_{10} = 0.01$, $C_{20} = 0.03$, $T_{10} = 300$, $T_{20} = 350$, $h_1 = 5$, $h_2 = 5$, $\delta = 0.05$, $E_0^2 = 10^5$.

RESEARCH ARTICLE

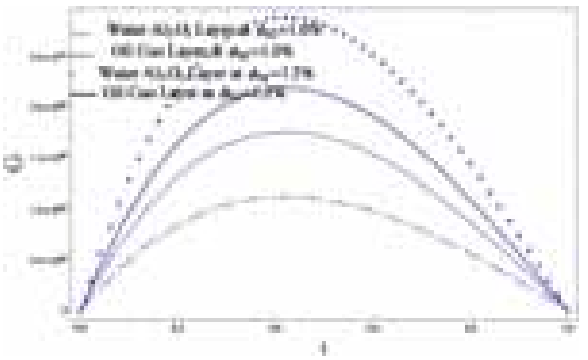


Fig. 10. Stability diagram for the energy distribution function with: $\phi_{b1} = 0.03$, $C_{10} = 0.01$, $C_{20} = 0.03$, $T_{10} = 300$, $T_{20} = 350$, $h_1 = 5$, $h_2 = 5$, $\delta = 0.05$, $E_0^2 = 10^5$.

where the different particle types exhibit different responses to the force of a temperature gradient.[27] The thermophoretic can be determined due to the temperature difference $\Delta T_j = T_{bj} - T_{0j}$. This difference is measured between the temperature of the fluid layer in the basic initial state T_{bj} and the temperature of the boundary of this layer T_{0j}. The results obtained experimentally by Chieruzzi et al.[28] show that the addition of 1.0 wt.% (wt% means weight percent, i.e., weight of solute/weight of solvent) of nanoparticles to the base salt increases the specific heat of 15% to 57% in the solid phase and of 1% to 22% in the liquid phase. Increasing the specific heat of the nanofluid means a lack of heating capacity. In other words, increasing the temperature gradient in the nanofluid tends to evaporate the fluid easily and provokes

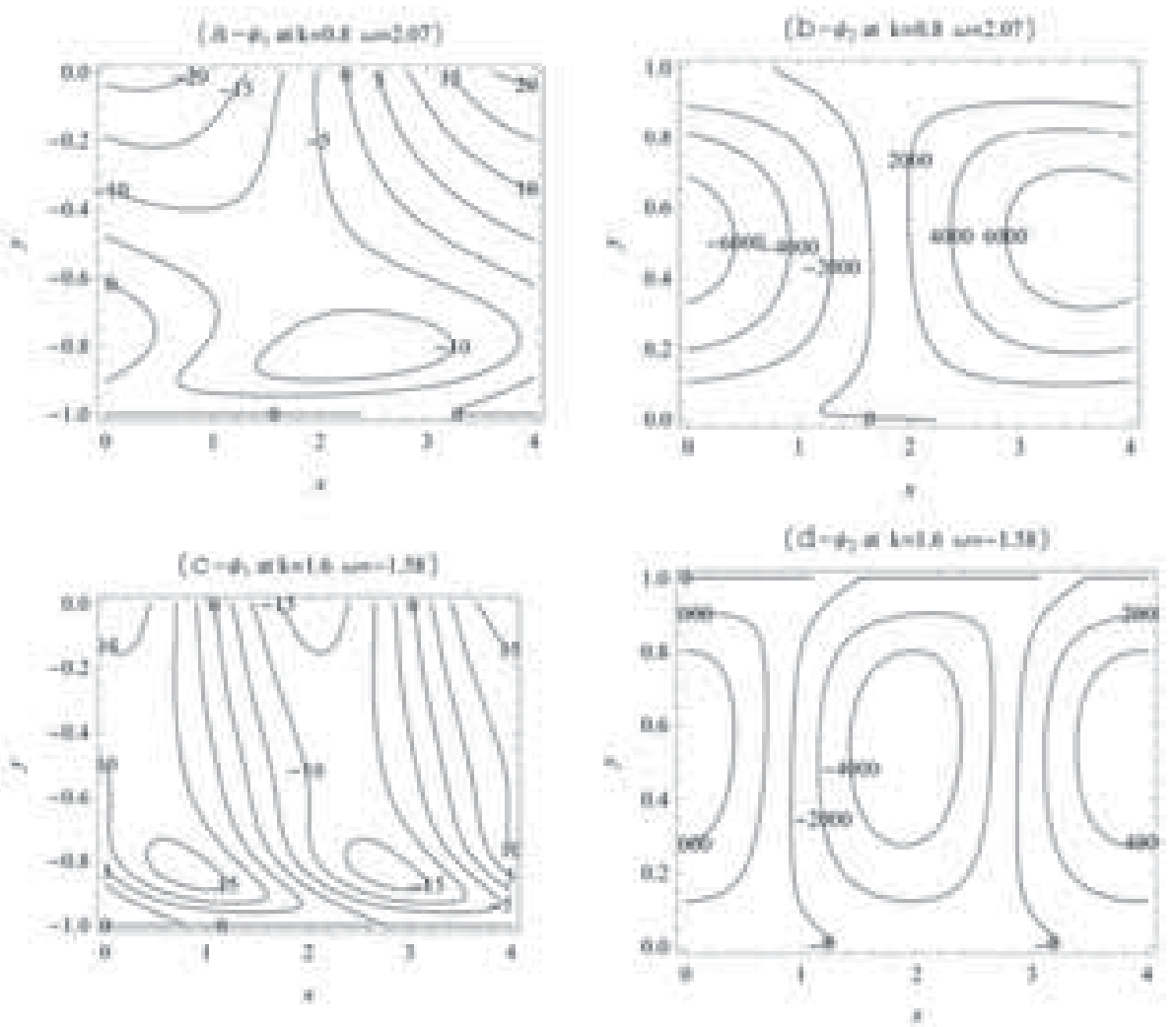


Fig. 11. Streamlines contours for a system having the particulars: $\rho_1 = 1$, $\rho_2 = 0.001$, $h_1 = 1$, $h_2 = 1$, $\mu_1 = 0.01$, $\mu_2 = 1 \times 10^{-5}$, $G_1 = 50$, $G_2 = 60$, $R_{N1} = 25$, $R_{N2} = 30$, $R_{a1} = 20$, $R_{a2} = 50$, $t = 0.002$, $\alpha_{f2} = 20$, $\alpha_{f1} = 30$, $C_{10} = 10$, $C_{20} = 30$, $T_{10} = 1$, $T_{20} = 10$, $D_B = 30$, $K_f = 20$, $Le_1 = 20$, $Le_2 = 30$, $N_{A1} = 10$, $N_{A2} = 20$, $E_0 = 1 \times 10^{-4}$.

higher instabilities of the fluid particles. So, the increase of the temperature difference ΔT_j increases the motion of the nanoparticles and consequently destabilizes the interface. This occurs in both systems of Water Al_2O_3-Oil CuO and Water Al_2O_3-Oil TiO_2 for both differences ΔT_1 and ΔT_2 as shown in Figures 5(a)–(b) and 6.

According to the normal modes assumption Eq. (32), the fluid layer becomes stable for the negative real values of the growth rate ω, which indicates decay with time (stability) and viseversa for the positive values. Table II illustrates that the system is unstable for small values of the wave number ($k \leq 0.6$) and the instability occurs for large wave number ($k \geq 0.9$) for the Water Al_2O_3-Oil CuO nanofluid layers. Meanwhile, the Water Al_2O_3-Oil TiO_2 interface is unstable through this range. This result supports the previous result of Figure 2 and also concludes that the wave number has a stabilizing effect.

6.2. For the Numerical Study by the Finite Difference Method

In this subsection the graphical representation of the stability analysis due to the numerical solution is presented and discussed. The stability behavior is discussed due to the energy growth rate function $G(t)$ that was obtained by the numerical solution. The relation between $G(t)$ and the time is presented graphically. The departure of the energy growth function represents the unstable state.

Figure 7 illustrates that the most unstable layer is the Oil TiO_2 which coincides the highest energy curve. This result is identical to the previous result of Figure 3.

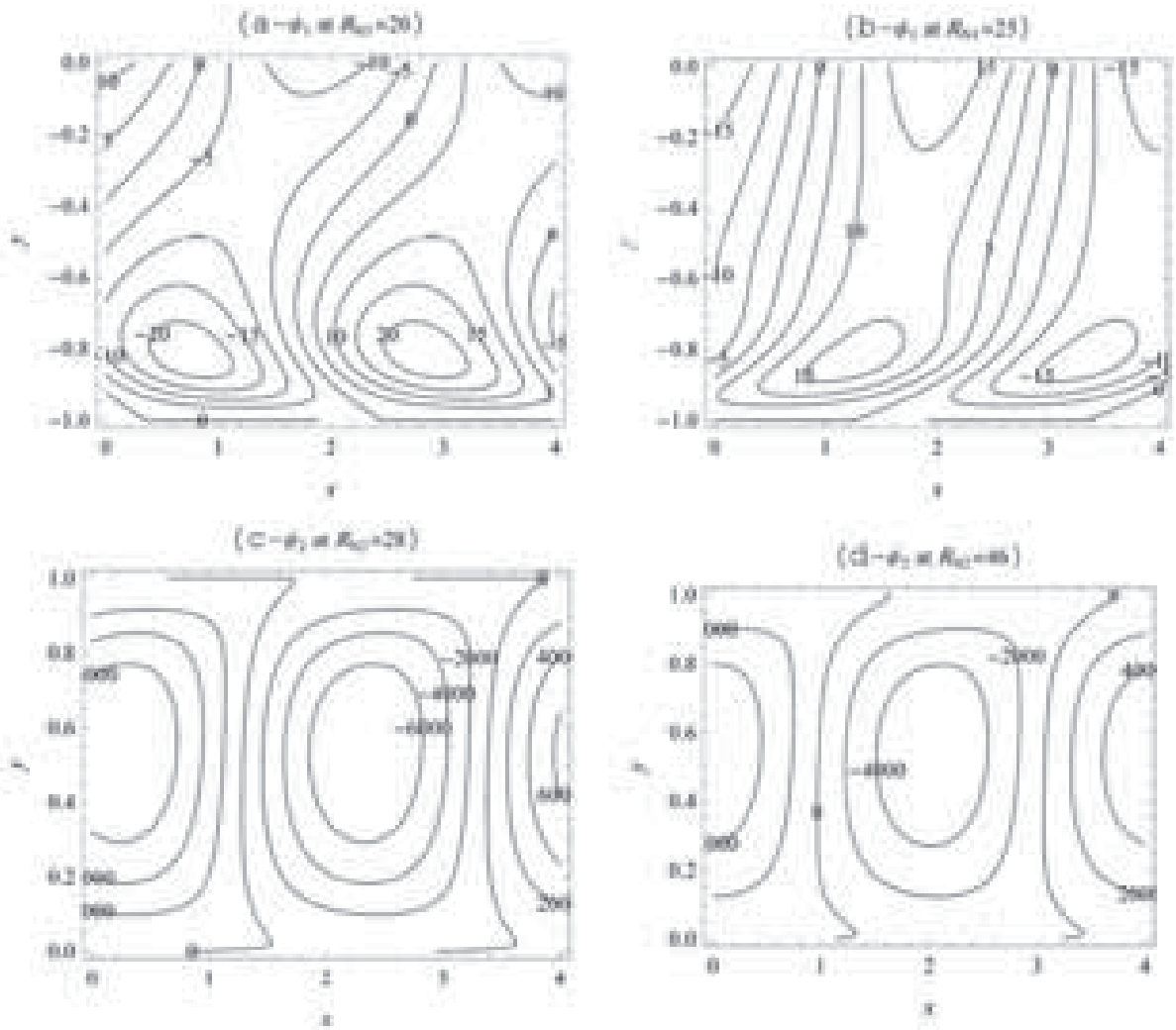


Fig. 12. Streamlines contours for a system having the particulars: $\rho_1 = 1$, $\rho_2 = 0.001$, $h_1 = 1$, $h_2 = 1$, $\mu_1 = 0.01$, $\mu_2 = 1 \times 10^{-5}$, $G_1 = 50$, $G_2 = 60$, $R_{a1} = 30$, $R_{a2} = 50$, $t = 0.002$, $\alpha_{f2} = 20$, $\alpha_{f1} = 30$, $C_{10} = 10$, $C_{20} = 30$, $T_{10} = 1$, $T_{20} = 10$, $D_B = 30$, $K_f = 20$, $Le_1 = 20$, $Le_2 = 30$, $N_{A1} = 10$, $N_{A2} = 20$, $E_0 = 1 \times 10^{-4}$.

This means that the Water Al_2O_3-Oil CuO is more stable than the Water Al_2O_3-Oil TiO_2.

Figure 8 shows that the energy growth rate $G(t)$ decreases with the increase of the wave number k which stabilizes the fluid layer. This figure illustrates this effect of the Water Al_2O_3 layer. This numerical deduction coincides with the previous analytical results that were deduced according to the tabulated results in Table II.

The result of Figure 9 supports the previous result of Figure 3. That is the volume fraction percent for the first layer ϕ_{b1} in the basic state have a destabilizing effect. This occurs due to the enhancement of the energy growth rate function $G(t)$ with the increase of the volume fraction percent ϕ_{b1} from 1% to 2%. This effect is more evidence for the Water Al_2O_3 layer. Also, the volume fraction percent for the second layer ϕ_{b2} have a destabilizing effect due to increasing the Brownian motion which in turn enhances the energy distribution and destabilizes the layer as seen in Figure 10.

7. STREAMLINES CONTOURS

In this section, the effects of different parameters, especially the nanofluid parameters, on the streamlines contours are discussed. To do this, we will not adhere to the values of the constants in Table I, but we will look at the default values to study the effect of all dimensionless parameters. Figures 11(a)–(b) represent the effect of the concentration Rayleigh number R_{N1} on the contour of the stream function ψ_1 in the first medium. These figures show that the flow consists of two cells (nodes or closed

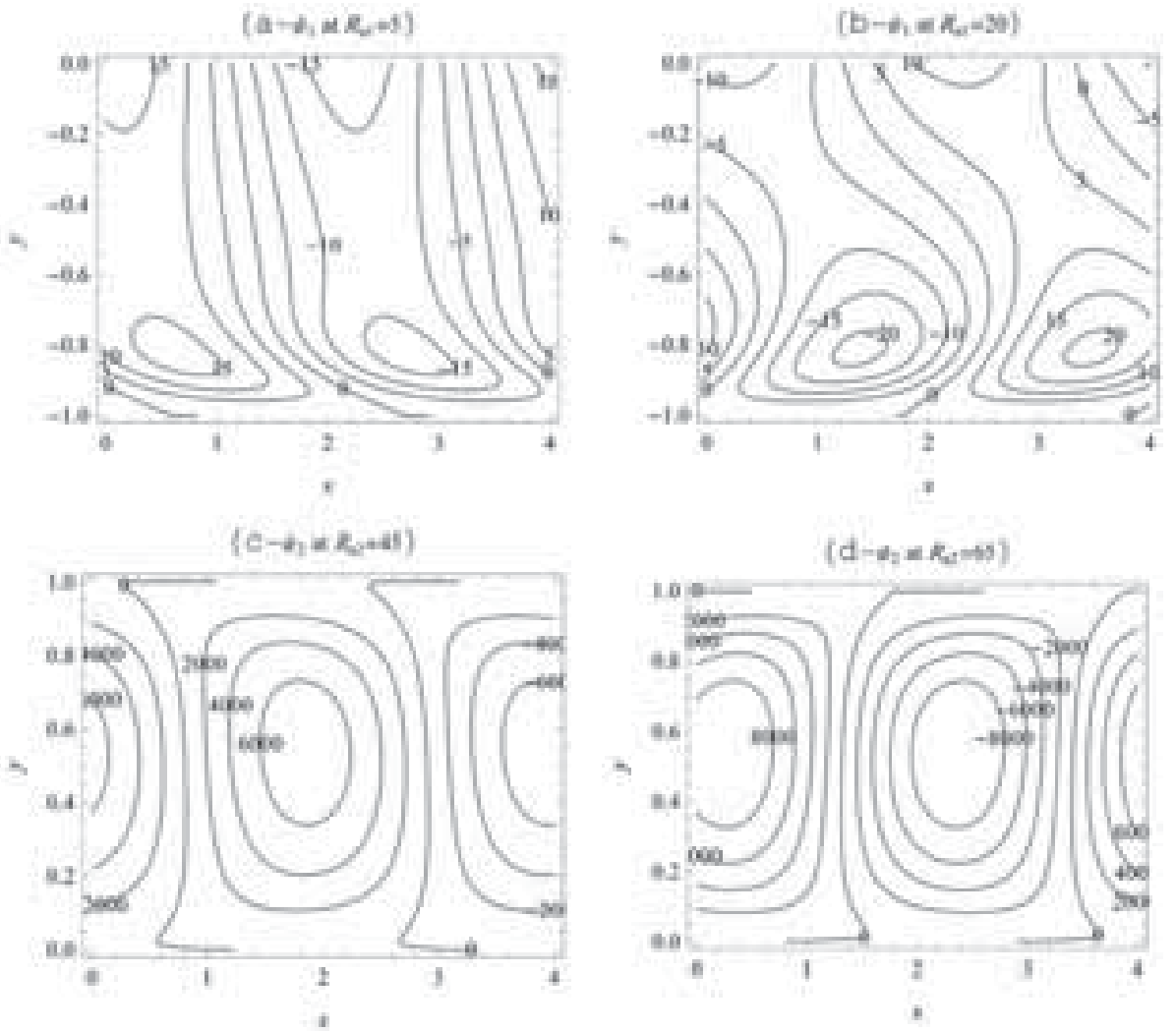


Fig. 13. Streamlines contours for a system having the particulars: $\rho_1 = 1$, $\rho_2 = 0.001$, $h_1 = 1$, $h_2 = 1$, $\mu_1 = 0.01$, $\mu_2 = 1 \times 10^{-5}$, $G_1 = 50$, $G_2 = 60$, $R_{N1} = 25$, $R_{N2} = 40$, $k = 1.5$, $t = 0.002$, $\alpha_{f2} = 20$, $\alpha_{f1} = 30$, $C_{10} = 10$, $C_{20} = 30$, $T_{10} = 1$, $T_{20} = 10$, $D_B = 30$, $K_f = 20$, $Le_1 = 20$, $Le_2 = 30$, $N_{A1} = 10$, $N_{A2} = 20$, $E_0 = 1 \times 10^{-4}$.

contours), one of them with positive circulation and the other one has a negative circulation. When the value of R_{N1} increases from 20 to 25 the central contours move towards the lower plate and decrease in size. It is evident to notice that the previous situation in the stable state where the real part of the growth rate, that was obtained during the calculations, is $\omega = -1.6204$. Similar trend is observed for the increasing of R_{N2}, in Figures 11(c)–(d), except that the contours position is still at the center of the second medium.

The number of cells and the density of the streamlines increase with decreasing the thermal Rayleigh number R_{a1} in the first medium as shown in Figures 12(a)–(b). This is due to the stabilizing effect of this parameter. Meanwhile, the number of cells remains the same in the second medium as shown in Figures 12(c)–(d).

Figures 13(a)–(d), examine the streamlines contours in cases of stable/unstable states. The stable state is corresponding to high values of the wave number k and negative values of the growth rate ω. Figures 13(a)–(c) represent the streamlines contours for the stream function ψ_1 in the first medium. Meanwhile, parts (a–d) represent ψ_2 in the second medium. Also, the parts (a–b) describe the unstable state ($k < 1$ and $\omega > 0$), while parts (c–d) describe the stable state. It is observed that the streamlines density and the cells in the stable state are greater than the unstable state in both media. Also, the cells of the lighter fluid (the second upper medium which has a small density and viscosity) are greater than the other medium. This is logically valid; whereas the lighter fluid has a higher movement of its particles which in turn gives more contours.

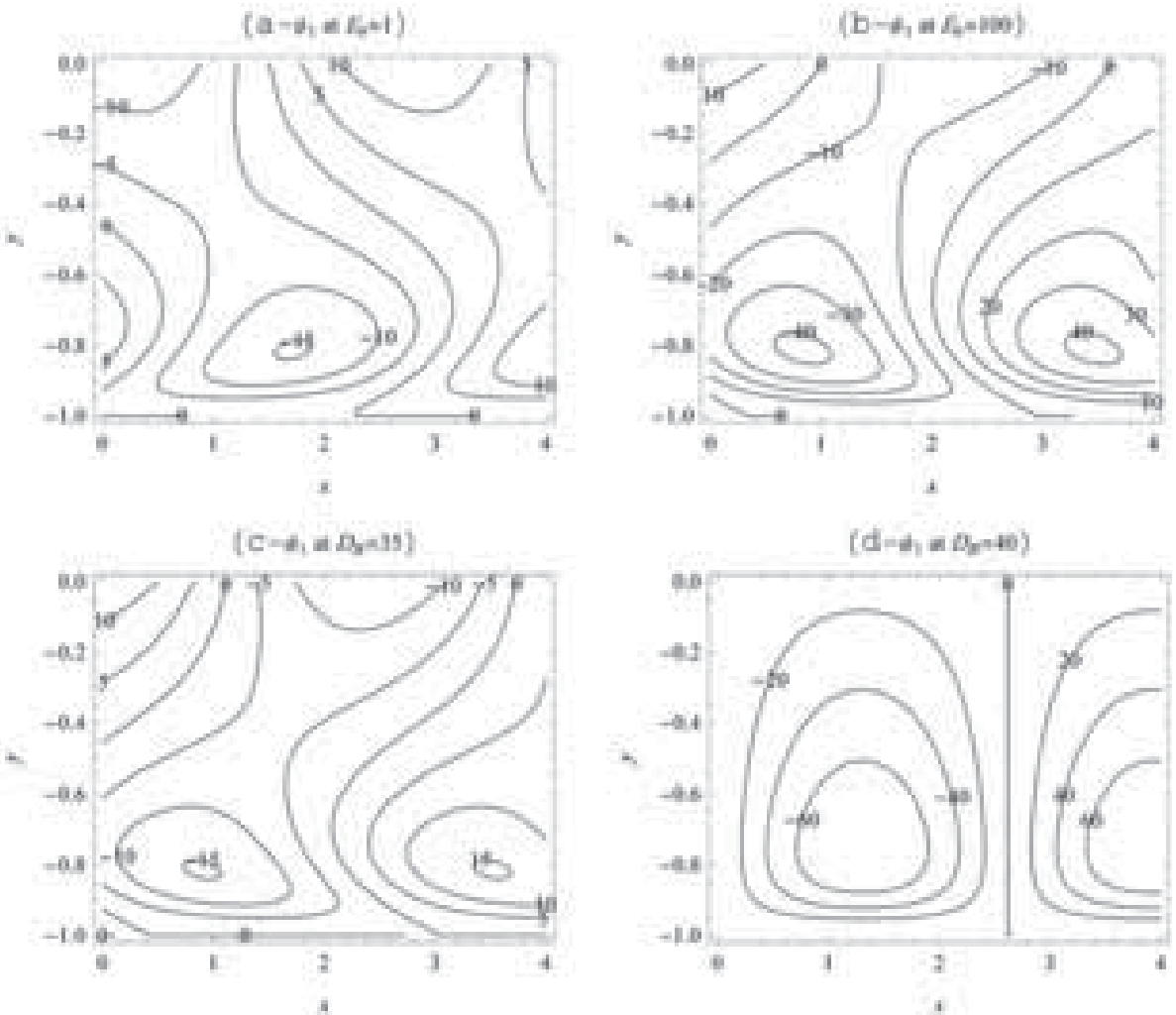


Fig. 14. Streamlines contours for a system having the particulars: $\rho_1 = 1$, $\rho_2 = 0.001$, $h_1 = 1$, $h_2 = 1$, $\mu_1 = 0.01$, $\mu_2 = 1 \times 10^{-5}$, $G_1 = 50$, $G_2 = 60$, $R_{N1} = 25$, $R_{N2} = 30$, $R_{a1} = 20$, $R_{a2} = 50$, $t = 0.002$, $\alpha_{f2} = 20$, $\alpha_{f1} = 30$, $C_{10} = 10$, $C_{20} = 30$, $T_{10} = 1$, $T_{20} = 10$, $k = 1.2$, $K_f = 20$, $Le_1 = 20$, $Le_2 = 30$, $N_{A1} = 10$, $N_{A2} = 20$.

The nanoparticles movement increases, according to the Brownian motion. This motion is represented by the Brownian diffusion ratio D_B. So, increasing the values of D_B corresponding to a higher movement of the particles, this leads to an increase in the cells and the streamlines. This phenomenon is illustrated in Figures 14(c)–(d). There is a considerable increase in the cell numbers in the first medium due to the variation of D_B from 35 to 50. The similar effect is observed for the increase in the electric field intensity E_0 as shown in Figures 14(a)–(b). The cell number increase to two cells in the first medium according to the increase in E_0. Also, this behavior is related to the stabilizing effect of the electric field intensity.

8. CONCLUSION

In this paper, the interfacial stability of two nanofluid layers (Water-Mineral Oil) is examined. The analytical solutions for the velocity, temperature, nanoparticles volume fraction and the electric field potential are obtained. Firstly, the dispersion relation is obtained as a transcendental relation between the wave number and the growth rate. Secondly, the problem is studied numerically by obtaining the numerical solution and the energy growth function. In the numerical solution the dependence on time for all functions is unknown. Numerical calculations are performed to study the stability of the interface according to the nanoparticle volume fraction percent and the temperature difference. The streamlines contours are represented graphically for the important nanofluid parameters. The main results of our study are discussed in the discussion section which can be summarized in the following points:

1—The longitudinal electric field has a stabilizing effect. This is a common result in the previous literature.[25, 26]

2—The Brownian motion describes the movement of the nanoparticles. This motion enhances the instability of the interface according to the increase of the nanoparticle volume fraction percent.

3—The Water Al_2O_3-Oil CuO is more stable than the Water Al_2O_3-Oil TiO_2 due to the increase of the nanofluid density.

4—The temperature difference, between the basic temperature of the layer and the temperature of the ambient wall, evaporate the fluid easily and provoke higher instabilities. So, the temperature difference destabilizes the interface. This result was previously obtained experimentally in Ref. [28].

5—The contours of the streamlines increase due to the Brownian motion and the electric field intensity.

APPENDIX

• The constants A_1, A_2, B_1, B_2, m_1 and m_2 that appear in temperature solutions Eqs. (39), (40), may be listed as follows

$$m_1=\sqrt{\frac{\omega}{\alpha_{f1}}+k^2},\quad m_2=\sqrt{\frac{\omega}{\alpha_{f2}}+k^2},$$

$$A_2=\left(\left[T_1-T_2e^{m_2h_2}\left(\cosh m_1h_1+\frac{k_fm_2}{m_1}\sinh m_1h_1\right)\right]\right)\Bigg/\left(\left[\cosh m_1h_1(1-e^{2m_2h_2})-\frac{k_fm_2}{m_1}\sinh m_1h_1(1+e^{2m_2h_2})\right]\right),$$

$$A_1=\frac{1}{2}(A_2+B_2)+\frac{k_fm_2}{2m_1}(A_2-B_2),$$

$$B_2=e^{m_2h_2}T_2-A_2e^{2m_2h_2},$$

$$B_1=\frac{1}{2}(A_2+B_2)-\frac{k_fm_2}{2m_1}(A_2-B_2)$$

• The constants $A_3, A_4, B_3, B_4, m_3, m_4, Q_1$ and Q_2 that appear in the volume fraction solutions Eqs. (42), (43), may be listed as follows

$$m_3=\sqrt{\omega Le_1+k^2},\quad m_4=\sqrt{\omega Le_2+k^2},$$

$$Q_1=\frac{N_{A1}(k^2-m_1^2)}{m_1^2-m_3^2},\quad Q_2=\frac{N_{A2}(k^2-m_2^2)}{m_2^2-m_4^2},$$

$$Q_3=\cosh m_3h_1[Q_2(A_2+B_2)-Q_1(A_1+B_1)]-\frac{D_B}{m_3}\sinh m_3h_1[Q_2(A_2m_2-B_2m_2)]+\frac{Q_1}{m_3}\sinh m_3h_1[A_1m_1-B_1m_1]+Q_1[A_1e^{-m_1h_1}+B_1e^{m_1h_1}],$$

$$B_4=\left(C_1-Q_3-\left(\cosh m_3h_1-\frac{D_Bm_4}{m_3}\sinh m_3h_1\right)\times[C_2e^{-m_4h_2}-Q_2e^{-m_4h_2}(A_2e^{m_2h_2}+B_2e^{-m_2h_2})]\right)\Bigg/\left(\cosh m_3h_1+\frac{D_Bm_4}{m_3}\sinh m_3h_1-e^{-2m_4h_2}\times\left(\cosh m_3h_1-\frac{D_Bm_4}{m_3}\sinh m_3h_1\right)\right),$$

$$A_4=C_2e^{-m_4h_2}-B_4e^{-2m_4h_2}-Q_2e^{-m_4h_2}\times(A_2e^{m_2h_2}+B_2e^{-m_2h_2}),$$

$$A_3=\frac{1}{2}(A_4+B_4+Q_2(A_2+B_2)-Q_1(A_1+B_1))+\frac{D_B}{2m_3}(A_4m_4-B_4m_4+Q_2(A_2m_2-B_2m_2))-\frac{Q_1}{2m_3}(A_1m_1-B_1m_1),$$

$$B_3=\frac{1}{2}(A_4+B_4+Q_2(A_2+B_2)-Q_1(A_1+B_1))$$
$$-\frac{D_B}{2m_3}(A_4m_4-B_4m_4+Q_2(A_2m_2-B_2m_2))$$
$$+\frac{Q_1}{2m_3}(A_1m_1-B_1m_1)$$

• The constants A_5-A_{12}, B_5-B_{12}, n_1 and n_2 that appear in the stream functions solutions Eqs. (44), (46), may be listed as follows

$$n_1=\sqrt{\frac{\rho_1\omega}{\mu_1 Pr_1}+k^2},$$
$$n_2=\sqrt{\frac{\rho_2\omega}{\mu_2 Pr_2}+k^2},$$
$$A_7=\frac{ik}{\mu_1}\left[\frac{R_{N1}A_3}{m_3^4-(k^2+n_1^2)m_3^2+k^2n_1^2}\right],$$
$$B_7=\frac{ik}{\mu_1}\left[\frac{R_{N1}B_3}{m_3^4-(k^2+n_1^2)m_3^2+k^2n_1^2}\right],$$
$$A_8=\frac{ik}{\mu_1}\left[\frac{Q_1R_{N1}A_1-R_{a1}A_1}{m_1^4-(k^2+n_1^2)m_1^2+k^2n_1^2}\right],$$
$$B_8=\frac{ik}{\mu_1}\left[\frac{Q_1R_{N1}B_1-R_{a1}B_1}{m_1^4-(k^2+n_1^2)m_1^2+k^2n_1^2}\right],$$
$$A_{11}=\frac{ik}{\mu_2}\left[\frac{R_{N2}A_4}{m_4^4-(k^2+n_2^2)m_4^2+k^2n_2^2}\right],$$
$$B_{11}=\frac{ik}{\mu_2}\left[\frac{R_{N2}B_4}{m_4^4-(k^2+n_2^2)m_4^2+k^2n_2^2}\right],$$
$$A_{12}=\frac{ik}{\mu_2}\left[\frac{Q_2R_{N2}A_2-R_{a2}A_2}{m_2^4-(k^2+n_2^2)m_2^2+k^2n_2^2}\right],$$
$$B_{12}=\frac{ik}{\mu_2}\left[\frac{Q_2R_{N2}B_2-R_{a2}B_2}{m_2^4-(k^2+n_2^2)m_2^2+k^2n_2^2}\right]$$

The other eight constants $A_5, B_5, A_6, B_6, A_9, B_9, A_{10}, B_{10}$ can be obtained by solving the following eight equations by any mathematical method (like Cramer's method)

$$A_5ke^{-kh_1}-B_5ke^{kh_1}+A_6n_1e^{-n_1h_1}-B_6n_1e^{n_1h_1}+A_7m_3e^{-m_3h_1}$$
$$-B_7m_3e^{m_3h_1}+A_8m_1e^{-m_1h_1}-B_8m_1e^{m_1h_1}=0,$$
$$A_5e^{-kh_1}+B_5e^{kh_1}+A_6e^{-n_1h_1}+B_6e^{n_1h_1}+A_7e^{-m_3h_1}$$
$$+B_7e^{m_3h_1}+A_8e^{-m_1h_1}+B_8e^{m_1h_1}=0,$$
$$A_9ke^{kh_2}-B_9ke^{-kh_2}+A_{10}n_2e^{n_2h_2}-B_{10}n_2e^{-n_2h_2}$$
$$+A_{11}m_4e^{m_4h_2}-B_{11}m_4e^{-m_4h_2}+A_{12}m_2e^{m_2h_2}$$
$$-B_{12}m_2e^{-m_2h_2}=0,$$
$$A_9e^{kh_2}+B_9e^{-kh_2}+A_{10}e^{n_2h_2}+B_{10}e^{-n_2h_2}+A_{11}e^{m_4h_2}$$
$$+B_{11}e^{-m_4h_2}+A_{12}e^{m_2h_2}+B_{12}e^{-m_2h_2}=0,$$
$$A_5+B_5+A_6+B_6+A_7+B_7+A_8+B_8$$
$$-(A_9+B_9+A_{10}+B_{10}+A_{11}+B_{11}+A_{12}+B_{12})=0,$$
$$A_5k-B_5k+A_6n_1-B_6n_1+A_7m_3-B_7m_3+A_8m_1$$
$$-B_8m_1-(A_9k-B_9k+A_{10}n_2-B_{10}n_2+A_{11}m_4$$
$$-B_{11}m_4+A_{12}m_2-B_{12}m_2)=0$$
$$\rho_1\omega A+ik\rho_1(A_5+B_5+A_6+B_6+A_7+B_7+A_8+B_8)$$
$$-\rho_2\omega A-ik\rho_2(A_9+B_9+A_{10}+B_{10}+A_{11}+B_{11}$$
$$+A_{12}+B_{12})=0,$$
$$\mu_1(A_5k^2+B_5k^2+A_6n_1^2+B_6n_1^2+A_7m_3^2+B_7m_3^2+A_8m_1^2$$
$$+B_8m_1^2)+\mu_1k^2(A_5+B_5+A_6+B_6+A_7+B_7$$
$$+A_8+B_8)-\mu_2(A_9k^2+B_9k^2+A_{10}n_2^2+B_{10}n_2^2+A_{11}m_4^2$$
$$+B_{11}m_4^2+A_{12}m_2^2+B_{12}m_2^2)-\mu_2k^2(A_9+B_9+A_{10}$$
$$+B_{10}+A_{11}+B_{11}+A_{12}+B_{12})=0$$

• The constants ℓ_1 and ℓ_2 that appear with the dispersion relation Eq. (52), may be listed as follows

$$\ell_1=ik\mu_1+\frac{i\rho_1\omega}{k},\quad \ell_2=ik\mu_2+\frac{i\rho_2\omega}{k}$$
$$D(\omega,k)=A_9(k\ell_2+i\mu_2k^2)+B_9(-k\ell_2-i\mu_2k^2)$$
$$+A_{10}\left(n_2\ell_2-\frac{i\mu_2}{k}n_2^3+2ik\mu_2n_2\right)$$
$$+B_{10}\left(-n_2\ell_2+\frac{i\mu_2}{k}n_2^3-2ik\mu_2n_2\right)$$
$$+A_{11}\left(m_4\ell_2-\frac{i\mu_2}{k}m_4^3+2ik\mu_2m_4\right)$$
$$+B_{11}\left(-m_4\ell_2+\frac{i\mu_2}{k}m_4^3-2ik\mu_2m_4\right)$$
$$+A_{12}\left(m_2\ell_2-\frac{i\mu_2}{k}m_2^3+2ik\mu_2m_2\right)$$
$$+B_{12}\left(-m_2\ell_2+\frac{i\mu_2}{k}m_2^3-2ik\mu_2m_2\right)$$
$$+A_5(-k\ell_1-ik^2\mu_1)+B_5(k\ell_1+ik^2\mu_1)$$
$$+A_6\left(-n_1\ell_1+\frac{i\mu_1}{k}n_1^3-2ik\mu_1n_1\right)$$
$$+B_6\left(n_1\ell_1-\frac{i\mu_1}{k}n_1^3+2ik\mu_1n_1\right)$$
$$+A_7\left(-m_3\ell_1+\frac{i\mu_1}{k}m_3^3-2ik\mu_1m_3\right)$$
$$+B_7\left(m_3\ell_1-\frac{i\mu_1}{k}m_3^3+2ik\mu_1m_3\right)$$
$$+A_8\left(-m_1\ell_1+\frac{i\mu_1}{k}m_1^3-2ik\mu_1m_1\right)$$
$$+B_8\left(m_1\ell_1-\frac{i\mu_1}{k}m_1^3+2ik\mu_1m_1\right)$$
$$+\frac{AE_0^2k(\varepsilon_2-\varepsilon_1)^2}{\varepsilon_2\tanh kh_2+\varepsilon_1\tanh kh_1}+Ak^2$$

References

1. D. Y. Hsieh, *Trans. ASME Series D, J. Basic Engng.* 74, 156 **(1972)**.
2. G. M. Moatimid and M. H. Obied Allah, *Appl. Math. Modeling* 34, 3118 **(2010)**.
3. K. Zakaria, M. A. Sirwah, and S. A. Alkharashi, *Commun. Theor. Phys.* 55, 1077 **(2011)**.
4. G. M. Moatimid, *Can. J. Phys.* 82, 1109 **(2004)**.
5. C. P. Manosh, D. A. S. Rees, and M. Wilson, *Int. J. Heat Mass Trans.* 48, 809 **(2005)**.
6. N. I. Lobov and T. P. Lyubimova, *Comp. Ren. Mec.* 341, 393 **(2013)**.
7. D. A. Nield and A. V. Kuznetsov, *Europ. J. Mech. B/Fluids* 29, 217 **(2010)**.
8. G. M. Moatimid, *J. Colloid Int. Sci.* 250, 108 **(2002)**.
9. M. K. Awasthi, R. Asthana, and Z. Uddin, *Int. Comm. Heat Mass Trans.* 71, 216 **(2016)**.
10. D. K. Tiwari, M. K. Awasthi, and G. S. Agrawal, *Ain Shams Engng. J.* 6, 1113 **(2015)**.
11. N. Jun, S. Zai-Hong, and T. Wen-Chang, *J. Hydrodynamic* 27, 809 **(2015)**.
12. S. Odenbach, Ferrofluids: Magnetically Controllable Fluids and Their Applications, Lecture Notes in Physics, Springer-Verlag, Berlin, Heidelberg **(2002)**, p. 253.
13. K. Zakaria and Y. Gamiel, *Int. J. Non-Linear Mech.* 48, 37 **(2013)**.
14. O. Ozena, N. Aubry, D. T. Papageorgiou, and P. G. Petropoulos, *Electrochim. Acta* 51, 5316 **(2006)**.
15. G. M. Moatimid, *J. Colloid Int. Sci.* 250, 108 **(2002)**.
16. Z. Haddad, E. Abu-Nada, H. F. Oztop, and A. Mataoui, *Int. J. Ther. Sci.* 57, 152 **(2012)**.
17. J. R. Melcher, Field Coupled Surface Waves, MIT Press, Cambridge, Mass **(1963)**; J. R. Melcher, Continuum Electrodynamics, MIT Press, Cambridge, Mass **(1981)**.
18. D. A. Nield and A. V. Kuznetsov, *Europ. J. Mech. B/Fluids* 29, 217 **(2010)**.
19. M. Corcione, *Energy Conv. Manage.* 52, 789 **(2011)**.
20. Websites: www.clippercontrols.com/dielectric-constants, www .engineeringtoolbox.com/liquid-dielectric constants.
21. M. A. Al-Nimr and A. M. A. Al-Dafaie, *Energy* 68, 318 **(2014)**.
22. M. Rafiq, Y. Lv, and C. Li, *J. Nanomaterials* 2016, ID 8371560, 23 **(2016)**.
23. H. A. Pakravan and M. Yaghoubi, *Int. J. Ther. Sci.* 50, 394 **(2011)**.
24. W. O. Criminale, T. L. Jackson, D. G. Lasseigne, and R. D. Joslin, *J. Fluid Mech.* 339, 55 **(1997)**.
25. A. E. K. Elcoot and G. M. Moatimid, *Physica A* 343, 15 **(2004)**.
26. J. R. Melcher, Continuum Electrodynamics, MIT Press, Cambridge, Mass **(1981)**.
27. https://en.wikipedia.org. "Brownian motion" and do the same for other physical definitions.
28. M. Chieruzzi, G. F. Cerritelli, A. Miliozzi, and J. M. Kenny, *Nanoscale Res. Lett.* 8, 448 **(2013)**.

Received: 15 March 2018. Accepted: 8 April 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1511–1515, 2018

Effective of Nanoparticles in Fluids Flow and Radiative Heat Transfer Over a Stretching Sheet

H. A. El-Dawy

The High Institute of Engineering and Technology, Tod Luxor, 85842, Egypt

In this paper, we study the behavior of fluids flow on stretching sheets also change in heat transfer under there exist nanoparticles Cuo in fluid and what happens in the velocity and heat transfer with the boundary layer flow with thermal radiate induced with velocity u and heat transfer T. We search for solutions and solve the problem to a double of (ODE) including four parameters the radiation NR, the ratio temperature. Magnetic field M and the nanoparticle Φ. The final results for velocity, temperature and heat transfer characteristics are presented in both graphical and tabular.

Keywords: Nanofluid, Ratio Temperature, Radiation, Magnetic Field.

RESEARCH ARTICLE

1. INTRODUCTION

In this paper we will study the effects of nanoparticle on behavior of fluid flow and radiate of heat transfer on stretching surface, recently many studies study fluid flow on many shape tube stretching sheet and also many studies dependent on nanofluid in our search we study flow fluid over a plate moving directly as Sakiadis (2005)[1] gave examples of nuclear power plants, gas turbines, satellites, but we study this properties and effect nanoparticle and magnetic field. The boundary layer behavior on a moving continuous surface is an important type of flow occurring in a number of engineering processes. Aerodynamic extrusion of plastic sheets, cooling of an infinite metallic plate in a cooling bath, the boundary layer along a liquid film in condensation processes and polymer sheet or filament extruded continuously from a dye Numerical results for hydro-magnetic mixed convection flow over a permeable non-isothermal wedge were reported by Prasad et al.[2] Hossain et al.[3] studied the effect of nanofluid using the (Range kutta) on natural convection flow of an optically thick viscous incompressible flow, and further, Hossain et al.[4] analyzed the effect of variable viscosity on this type of flow. On the other hand, Raptis and Perdikis[5] used a linearized form of the aforesaid Roseland approximation in view to analyze the steady flow of a viscoelastic fluid flow an unmoving surface. These simplifications permit an easier analysis and many investigations. Hydrodynamics, non-Newtonian fluids, porous media, etc. constitute analytical or numerical attempts which have been made in the recent past (Misra and Sinha, 2013).[6] The problem of steady micropolar nanofluid flow past a stretching sheet has been devised by many (see, made a comprehensive review of the heat transfer in nanofluids. Very recently, unsteady fluid flow with Ref. [7] or without (Bachok et al., 2011) thermal radiation effects also analyzed. One of the objectives of the present paper is to extend the investigation of Cortell (2008c)[8] to analyze the Sakiadis flow generated by a sheet stretched with a velocity which is assumed to be proportional to the $x\,1/3$ quantity, x is the distance from the slit. We assume appropriate boundary conditions for the energy equation that may assure the existence of similarity solutions (i.e., constant temperature at the surface) when radiative nonlinear heat transfer is studied. Very recently, Rahman and Eltayeb,[9] Pantokratoras and Fang[10] used the Rosseland[11] diffusion approximation in studying radiative nonlinear heat transfer in different geometries. Assuming sufficiently small temperature differences between the plate and the ambient fluid. And using of nanofluids flow El-Dawy et al.[12]

2. FLOW ANALYSIS

Let us consider the flow of an incompressible viscous Nanofluid past a stretching sheet coinciding with the plane $y = 0$, the flow is confined to $y > 0$. Two equal and opposite forces are applied along the x-axis so that the wall is stretched keeping the origin fixed. Absorbing-emitting but non-scattering medium. Use is made of usual notation and then we can express the basic equations describing the conservation of mass and momentum in the boundary layer as

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \tag{1}$$

$$\rho_{\mathrm{nf}}\left(u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y}\right) = \mu_{\mathrm{nf}}\frac{\partial^2 \bar{u}}{\partial \bar{y}^2} - \sigma B_0^2 \bar{u} + g\beta_{nf}(T - T) \tag{2}$$

 1546-1955/2018/15/1511/005 doi:10.1166/jctn.2018.7384

Table I. Physical properties.

Physical properties	Fluid phase	Cu
C_p (j/kg k)	4179	385
ρ (kg/m^3)	997.1	400
k (w/m k)	0.613	8933
β (k^{-1})	21×10^{-5}	1.67×10^{-5}

Here x and y are the Cartesian coordinates measured along the plate and normal to it, u and v are the velocity components along x and y axes, respectively, T is the temperature of the nanofluid, φ is the nanoparticle volume fraction, μ_f and μ_s are dynamic viscosity of the fluid and solid B_0 the strength of the magnetic field, σ is electrical conductivity of the fluid the dynamic viscosity of the base fluid, ρ_f and ρ_s are the densities of the fluid and of the solid fractions, μ_{nf} is the viscosity of the nanofluid. ρ_{nf} is density of nanofluids, β_{nf} thermal expansion coefficient nanofluids and $\propto_{\mathrm{nf}}$ is the thermal diffusivity of the nanofluids, which are given by, see Oztop and Abu-Nada[13]

$$\frac{K_{\mathrm{nf}}}{K_f} = \frac{(K_s + 2K_f) - 2\varphi(K_s - K_f)}{(K_s + 2K_f) + \varphi(K_s - K_f)}$$

$$\mu_{\mathrm{nf}} = \frac{\mu_f}{(1-\varphi)^{2.5}}$$

$$\alpha_{\mathrm{nf}} = \frac{K_{\mathrm{nf}}}{(\rho C_p)_{\mathrm{nf}}} \quad (\rho C_p)_{\mathrm{nf}} = (1-\varphi)(\rho C_p)_f + \varphi(\rho C_p)$$

$$\beta_{\mathrm{nf}} = (1-\varphi)\beta_f + \varphi\beta_s$$

$$u_w = \frac{v}{L^{4/3}} x^{1/3} \quad v = 0 \text{ at } y \to 0 \quad u \to 0 \text{ as } y \to \infty \tag{3}$$

where L is a characteristic length.

Now we introduce the following non-dimensional variables.

$$\eta = y\frac{x^{-1/3}}{L^{2/3}} \quad u = \frac{v}{L^{4/3}} x^{1/3} f'(\eta)$$
$$v = -\frac{v}{L^{2/3}} x^{-1/3} \frac{[2f - \eta f']}{3} \tag{4}$$

It is clear that u and v satisfy the equation of continuity (i.e., Eq. (1)).

By substituting these new variables in Eq. (2) we get

$$3\mu_{\mathrm{nf}} \frac{d^3 f}{d\eta^3} + (\rho)_{\mathrm{nf}} \left(2f\frac{d^2 f}{d\eta^2} - \left(\frac{df}{d\eta}\right)^2 \right) - M\frac{df}{d\eta} + \beta_{\eta f}\frac{dT}{d\eta} = 0 \tag{5}$$

similarity variable η and f is the dimensionless stream function.

The boundary conditions (3) can be written as

$$f = 0 \quad f' = 1 \quad \text{at } \eta = 0 \quad f' \to 0 \quad \text{as } \eta \to \infty \tag{6}$$

$$\mu_{\mathrm{nf}} f''' + \rho_{\mathrm{nf}}(f \cdot f'' - f' \cdot f') - M \cdot f' + \beta_{\eta f} \cdot \theta = 0 \tag{7}$$

3. HEAT TRANSFER ANALYSIS

By using usual boundary layer approximations, the equation of the energy for temperature T in the presence of thermal radiation is given by

$$u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \propto_{\eta f} \frac{\partial^2 T}{\partial y^2} - \frac{1}{(\rho Cp)_{\eta f}} \frac{\partial q_r}{\partial y} \tag{8}$$

where $\propto_{\eta f}$ is the thermal diffusivity, Cp_{nf} is the specific heat of the nanofluid at constant pressure and q_r is the radiative heat flux in y-direction.

$$T = T_w \quad \text{at } y = 0 \quad T \to T_\infty \quad \text{As } y \to \infty \tag{9}$$

with $T_w > T_\infty$, and T_∞ is the fluid temperature far away from the surface, T_w is the temperature at the wall.

Using the Rosseland approximation for radiation,[11] the radiative heat flux is simplified as

$$q_r = -\frac{4\sigma^*}{k^*}\frac{\partial T^4}{\partial y} \tag{10}$$

Where σ^* and k^* are the Stefan-Boltzmann constant and the mean absorption coefficient, respectively.
from Eq. (10) we get

$$q_r = -\frac{16\sigma^*}{3k^*} T^3 \frac{dT}{dy} \tag{11}$$

Where T is the temperature across the boundary layer. We have supposed T as x-independent, and in view to Eq. (11), Eq. (8) reduces to by defining the non-dimensional temperature $\theta(\eta)$ as

$$\theta(\eta) = \frac{T - T_\infty}{T_w - T_\infty} \tag{12}$$

We obtain $T = T_\infty(1 + (R-1)\theta)$ With

$$R = \frac{T_w}{T_\infty} \tag{13}$$

R is being the temperature ratio parameter. Taking into account the above we get

$$-\frac{\partial q_r}{\partial y} = \frac{16\sigma^* T_\infty^3}{3k^*} \frac{x^{-2/3}}{L^{4/3}} [T_\infty[1 + \theta(R-1)]^3 \cdot (R-1)\theta''$$
$$+ 3[1 + \theta(R-1)]^2 (\theta')^2] \tag{14}$$

Substituting Eqs. (4) and (14) into Eq. (8) and after some algebraone can write

$$u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \propto_{\eta f} \frac{16\sigma^* T_\infty^3}{3k^*} \frac{x^{-2/3}}{L^{4/3}} T_\infty((R-1)\theta'')$$
$$+ \frac{1}{(\rho Cp)_{\eta f}} \frac{16\sigma^* T_\infty^3}{3k^*} \frac{x^{-2/3}}{L^{4/3}}$$
$$\times [T_\infty[1 + \theta(R-1)]^3 \cdot (R-1)\theta''$$
$$+ 3[1 + \theta(R-1)]^2 (\theta')^2] \tag{15}$$

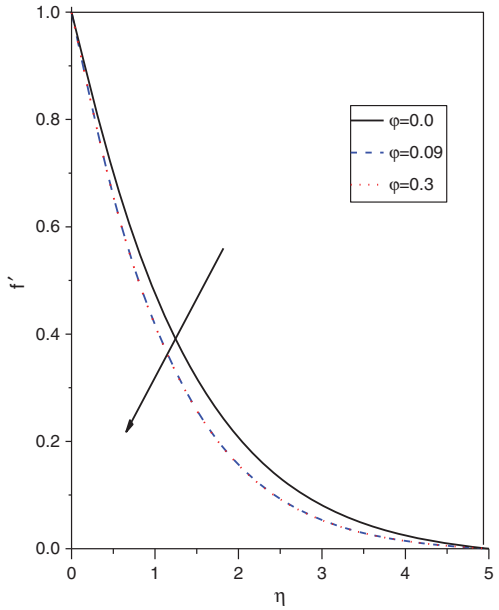


Fig. 1. Represent effective of volume particle φ on velocity f'.

$$-\frac{2}{3}\nu_{\eta f} f\theta' = \frac{k_{\eta f}}{(\rho Cp)_{\eta f}}\mathrm{NR}((R-1)\theta'') + \frac{1}{(\rho Cp)_{\eta f}}\mathrm{NR}\cdot[[1+\theta(R-1)]^3\cdot(R-1)]\theta'' + 3[1+\theta(R-1)]^2(\theta')^2] \quad (16)$$

This table show the effective of variation of nanoparticle on the skin fraction and Nusselt number with

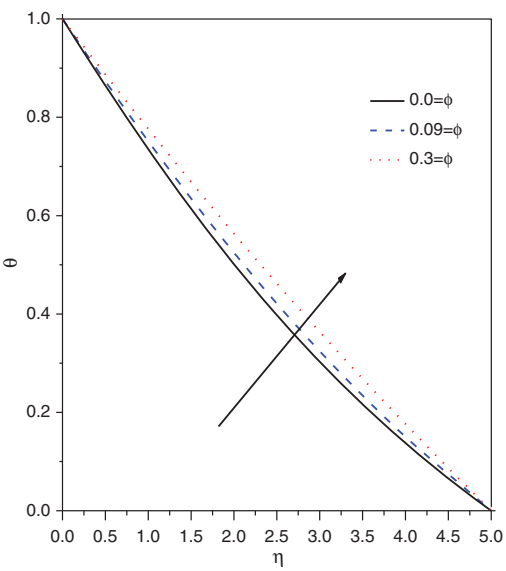


Fig. 2. Show the effective of volume particle φ on temperature θ.

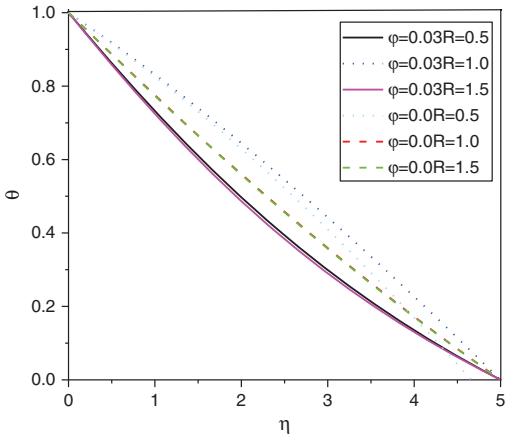


Fig. 3. Show the effective of ratio temperature θ_w on temperature θ.

variation of ratio temperature.

∅	θ_w	$-f''(0)$	$-\theta'(0)$
0.0	0.0	0.8061	0.2743
	0.3	0.8067	0.2784
	0.5	0.8072	0.2876
0.05	0.0	0.8238	0.2658
	0.3	0.8243	0.27
	0.5	0.8248	0.278
0.1	0.0	0.8324	0.2597
	0.3	0.8328	0.2634
	0.5	0.8333	0.2703

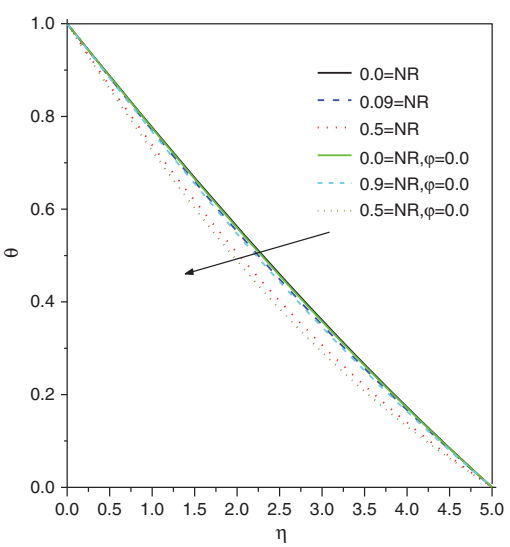


Fig. 4. Show the effective of radiation parameter NR on temperature θ.

4. RESULTS AND DISCUSSION

From Figure 1 we see effective of nanoparticles Φ on the velocity f' of fluids flow when nanoparticles increasing the velocity decreasing when magnetic field $M = 1.0$, parameter radiation $\mathrm{NR} = 1.0$ and ratio temperature $R = 1.5$.

From Figure 2 we see effective of nanoparticles Φ on the temperature θ of fluids flow when nanoparticles increasing the temperature increasing when magnetic field $M = 1.0$, parameter radiation $\mathrm{NR} = 1.0$ and ratio temperature $R = 1.5$.

From Figure 3 we see effective of ratio temperature R on the temperature θ of fluids flow when ratio temperature θ_w increasing the temperature increasing when magnetic field $M = 1.0$, parameter radiation $\mathrm{NR} = 1.0$ and ratio temperature $\Phi = 0.05$.

From Figure 4 we see effective of parameter radiation NR on the temperature θ of fluids flow when parameter radiation NR increasing the temperature decreasing when magnetic field $M = 1.0$, nanoparticles $\Phi = 0.05$ and ratio temperature $R = 1.5$.

From Figure 5 we see effective of magnetic parameter M on the temperature θ of fluids flow when

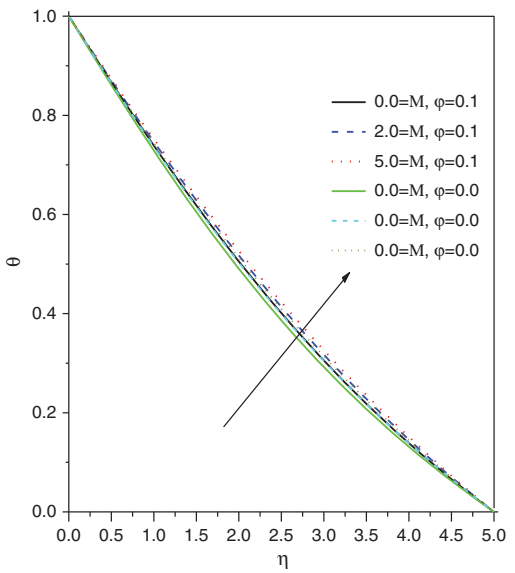


Fig. 5. Represent parameter magnetic field M on temperature θ.

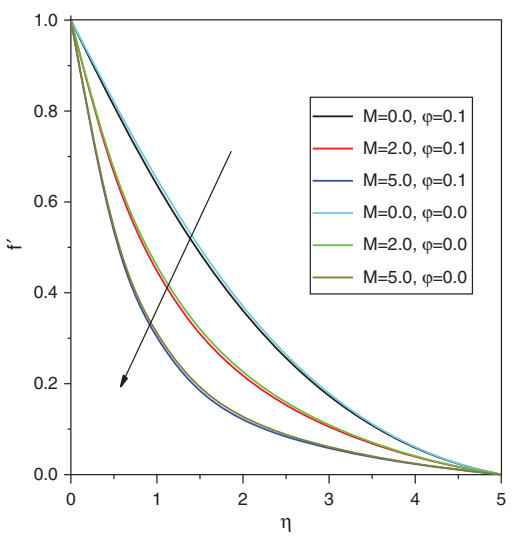


Fig. 6. Represent parameter magnetic field M on velocity f'.

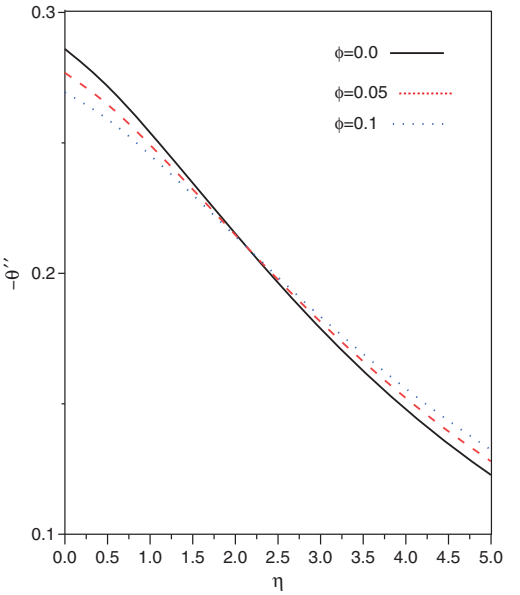


Fig. 7. The effective of nano-particles on Nusselt number.

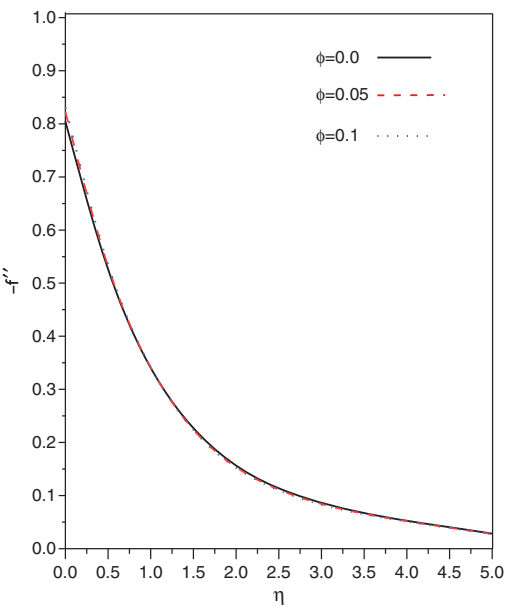


Fig. 8. The effective of nano-particle on skin fraction.

magnetic parameter M increasing the temperature increasing when radiation parameter NR = 1.0, nanoparticles $\Phi = 0.05$ and ratio temperature $R = 1.5$.

From Figure 6 we see effective of magnetic parameter M on the velocity f' of fluids flow when magnetic parameter M increasing the velocity f' decreasing when radiation parameter NR = 1.0, nanoparticles $\Phi = 0.05$ and ratio temperature $R = 1.5$.

5. CONCLUSION

We deduce from this result the nano particles and thermal radiation effects on flow influenced by a nonlinearly sheet was studied numerically. The effects of various physical parameters like NR, M, R and φ on heat transfer phenomena have been studied. It should also be concluded that in contrast to the linear Rosseland diffusion approximation, when use is made of the nonlinear one, the problem is also governed by the newly nanoparticles $\varnothing$. Similarity solutions for the case of stretching materials which have a wide variety of technical and environmental applications were found for all the aforementioned dimensionless physical parameters. The results presented indicate quite clearly that φ on.

NOMENCLATURE

Cp	specific heat at constant temperature
$(\rho Cp)_{nf}$	heat capacitance of the nanofluid
g	acceleration due to gravity
k	thermal conductivity
k_{nf}	thermal conductivity of the nanofluid
T	fluid temperature
T_∞	fluid temperature of the ambient fluid
u, v	velocity components along x and y directions
U_∞	free stream velocity.

Greek Letters

α_{nf}	thermal diffusivity of the nanofluid
β	thermal expansion coefficient
ϕ	nanoparticle volume fraction
η	similarity variable
θ	dimensionless temperature
ρ	density
ρ_{nf}	density of the nanofluid
ψ	stream function.

Subscripts

f	fluid fraction
s	solid fraction
nf	nanofluid fraction.

References

1. B. C. Sakiadis, *American Institute of Chemical Engineers* 7, 26 **(1961)**.
2. K. V. Prasad, P. S. Datti, and K. Vajravelu, *Journal of King Saud University Science* 25, 313 **(2013)**.
3. M. A. Hossain, M. A. Alim, and D. A. S. Rees, *Int. J. Heat Mass Transfer* 42, 181 **(1999)**.
4. M. A. Hossain, K. Khanafer, and K. Vafai, *International Journal of Thermal Sciences* 40, 115 **(2001)**.
5. A. Raptis and C. Perdikis, *Zeits. Angew. Mathematik und Mechanik* 78, 277 **(1998)**.
6. V. S. Arpaci, *Int. J. Heat Mass Transfer* 11, 871 **(1968)**.
7. K.-L. Hsiao, *Applied Mathematics and Information Sciences* 6, 59S **(2012)**.
8. R. Cortell, *Chinese Physics Letters* 25, 1340 **(2008a)**.
9. M. M. Rahman and I. A. Eltayeb, *Meccanica* 48, 601 **(2013)**.
10. A. Pantokratoras and T. Fang, *Physica Scripta* 87, 015703 **(2013)**.
11. S. Rosseland, Springer Verlag, Berlin **(1931)**, pp. 41–44.
12. H. A. El-Dawy, A. A. Mohammedien, and R. Gorla, *J. Nanofluids* 2, 1 **(2013)**.
13. H. F. Oztop and E. Abu-Nada, *Int. J. Heat Mass Transfer* 29, 1326 **(2008)**.

Received: 23 March 2018. Accepted: 8 April 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1516–1527, 2018

RESEARCH ARTICLE

Charge Density and Electrical Characteristics of 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) Molecular Nanowire by Quantum Chemical Study

Subramanian Palanisamy[1], Ponnusamy Srinivasan[2,*],
Aruputharaj David Stephen[3], and Karuppannan Selvaraju[1,*]

[1] Department of Physics, Kandaswami Kandar's College, Velur, Namakkal (Dt) 638182, India
[2] Department of Physics, Chikkaiah Naicker College, Erode, Erode (Dt) 638004, India
[3] Department of Physics, Sri Shakthi Institute of Engineering and Technology, Coimbatore 641062, India

The present investigation of structural, charge density and electrical characteristics of Au and thiol substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecular wire by using quantum chemical calculations has been carried out with density functional theory (DFT). The various applied electric fields (0.00–0.26 VÅ^{-1}) altered the geometrical parameters and the corresponding electrostatic and transport properties of the molecule has been analyzed. The variations in the atomic charges (MPA, NPA) of the molecule for the various applied electric fields have been compared. The HOMO-LUMO gap of the molecule for zero bias is 2.19 eV, as the field increases this gap decrease to 0.27 eV. The ESP shows the potential difference between charges accumulated of the molecule for various applied electric fields. To increase the applied electric field, the dipole moment of the molecule increases from 3.14 Debye to 49.17 Debye.

Keywords: Molecular Wire, Density Functional Theory, HLG, ESP, Dipole Moment.

1. INTRODUCTION

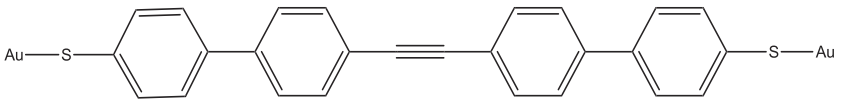
Molecular Nano electronics is the emerging field in electronics; it expands energy from the knowledge of molecular chemistry and nanophysics. Mostly, it's deals with design and fabrication of nano-electronic devices based on molecules.[1,2] Interestingly, a typical nano-electronic device consists of metal–molecule–metal junction formed between macroscopic electrodes and a single organic molecule or a monolayer of molecules. The most important part of the research is to predict the conductance of the molecule. The theoretical and the experimental investigations provide adequate input for the design of highly conducting molecules using to fabricate nano-electronic devices. The molecular conductance has been measured experimentally, by using scanning tunneling microscopy (STM), atomic force microscopy (AFM), and monolayer evaporation techniques[3–6] and so on. Past years, quantum chemical calculations have been employed to determine the electron transport characteristics of the molecules using density functional theory (DFT) combined with non-equilibrium Green's function technique.[7,8] It is well known that the length variation plays an important role in predicting the molecular conductance of the molecular wires.[9] Specifically, we construct the extended molecule by including a gold atom between the molecule and the Au surfaces, as suggested by Lang et al.,[10] to include the interaction between the molecule and gold electrodes into a molecule system. In this theoretical research study broadly, we report the computational details, structural properties, atomic charges, molecular orbital analysis, molecular electrostatic potential and molecular dipole moment of the 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule (Fig. 1).

2. COMPUTATIONAL DETAILS

The 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule has been carried out by Density Functional Theory (DFT) has using Gaussian09 program package.[11–13] In this investigation, DBPE molecule was undertaken (Fig. 1) for the zero field and applied field of five biasing steps (0.05, 0.10, 0.15, 0.21 and 0.26 VÅ^{-1}). A combination of Becke's three parameters exchange function and Lee, Yang and Parr gradient-corrected correlation function (B3LYP hybrid function) is applied for whole DFT calculation along with LANL2DZ (Los Alamos National

*Authors to whom correspondence should be addressed.

1546-1955/2018/15/1516/012 doi:10.1166/jctn.2018.7385



1,2-di([1,1'-biphenyl]-4-yl)ethyne

Fig. 1. Schematic diagram of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE).

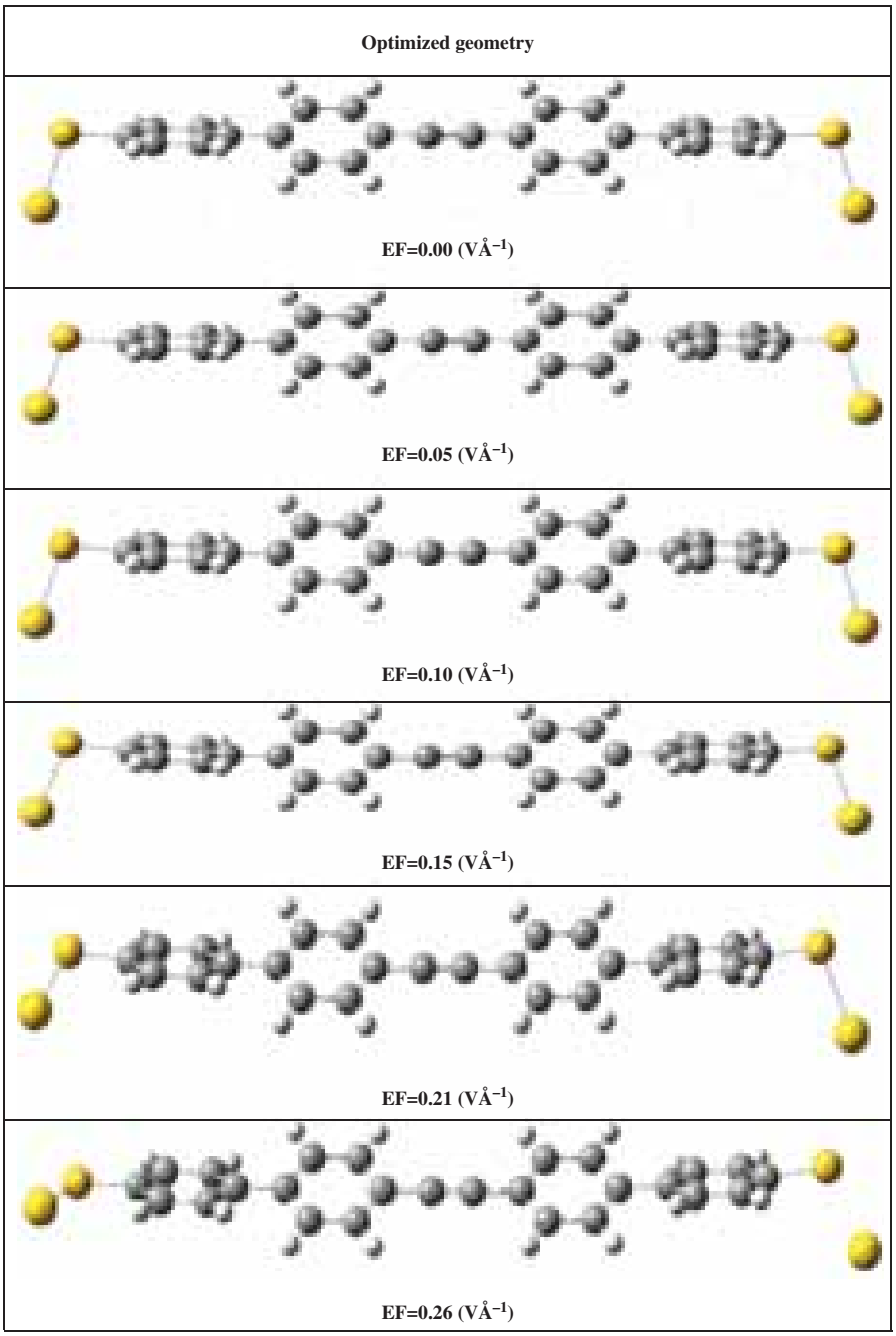


Fig. 2. Optimized geometry of 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecular wire for the zero and various applied EFs.

Table I. Bond lengths (Å) of the Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for the zero and various applied EFs.

Bonds	0.00	0.05	0.10	0.15	0.21	0.26
C–C Bonds						
Ring-I						
C(1)–C(2)	1.412	1.412	1.412	1.412	1.413	1.413
C(2)–C(3)	1.402	1.402	1.402	1.402	1.402	1.402
C(3)–C(4)	1.418	1.418	1.419	1.419	1.419	1.419
C(5)–C(4)	1.418	1.418	1.418	1.418	1.419	1.419
C(5)–C(6)	1.402	1.402	1.402	1.402	1.402	1.401
C(6)–C(1)	1.412	1.412	1.412	1.412	1.412	1.412
Ring-II						
C(7)–C(8)	1.417	1.418	1.418	1.418	1.418	1.418
C(7)–C(12)	1.418	1.418	1.418	1.418	1.418	1.418
C(8)–C(9)	1.400	1.400	1.400	1.400	1.400	1.400
C(9)–C(10)	1.420	1.420	1.420	1.420	1.420	1.420
C(10)–C(11)	1.419	1.420	1.420	1.420	1.420	1.421
C(11)–C(12)	1.400	1.400	1.399	1.399	1.399	1.399
Ring connector						
C(4)–C(7)	1.487	1.487	1.487	1.486	1.486	1.486
C(10)–C(13)	1.431	1.431	1.430	1.430	1.430	1.429
C(13)–C(14)	1.229	1.229	1.229	1.229	1.229	1.229
C(14)–C(15)	1.431	1.431	1.431	1.431	1.431	1.431
C(18)–C(21)	1.487	1.487	1.487	1.487	1.487	1.487
Ring-III						
C(15)–C(16)	1.420	1.419	1.419	1.419	1.419	1.419
C(15)–C(20)	1.419	1.420	1.420	1.420	1.420	1.420
C(16)–C(17)	1.400	1.400	1.400	1.401	1.401	1.400
C(17)–C(18)	1.418	1.417	1.417	1.417	1.417	1.418
C(18)–C(19)	1.418	1.417	1.418	1.418	1.418	1.418
C(19)–C(20)	1.400	1.400	1.400	1.399	1.399	1.399
Ring-IV						
C(21)–C(26)	1.418	1.417	1.417	1.416	1.416	1.416
C(21)–C(22)	1.418	1.418	1.418	1.418	1.418	1.418
C(22)–C(23)	1.402	1.402	1.402	1.402	1.402	1.402
C(23)–C(24)	1.412	1.412	1.412	1.412	1.412	1.412
C(24)–C(25)	1.412	1.412	1.412	1.411	1.411	1.412
C(25)–C(26)	1.402	1.403	1.403	1.404	1.404	1.404
C–H bonds						
C(2)–H(2)	1.086	1.086	1.086	1.086	1.086	1.086
C(3)–H(3)	1.087	1.087	1.087	1.087	1.086	1.086
C(5)–H(5)	1.087	1.087	1.087	1.087	1.087	1.087
C(6)–H(6)	1.086	1.086	1.086	1.087	1.087	1.087
C(8)–H(8)	1.087	1.087	1.087	1.087	1.087	1.087
C(9)–H(9)	1.087	1.086	1.086	1.086	1.086	1.086
C(11)–H(11)	1.087	1.087	1.087	1.087	1.087	1.087
C(12)–H(12)	1.087	1.087	1.087	1.087	1.087	1.088
C(16)–H(16)	1.087	1.087	1.086	1.086	1.086	1.086
C(17)–H(17)	1.087	1.087	1.087	1.087	1.087	1.087
C(19)–H(19)	1.087	1.087	1.087	1.087	1.087	1.087
C(20)–H(20)	1.087	1.087	1.087	1.087	1.087	1.087
C(22)–H(22)	1.087	1.087	1.087	1.087	1.087	1.087
C(23)–H(23)	1.086	1.086	1.086	1.086	1.086	1.086
C(25)–H(25)	1.086	1.086	1.086	1.086	1.087	1.087
C(26)–H(26)	1.087	1.087	1.087	1.087	1.088	1.088
Terminal bonds						
S(1)–C(1)	1.837	1.837	1.837	1.836	1.836	1.835
S(2)–C(24)	1.838	1.838	1.839	1.840	1.841	1.841
Au(1)–S(1)	2.400	2.400	2.401	2.401	2.402	2.403
Au(2)–S(2)	2.400	2.399	2.398	2.395	2.395	2.392

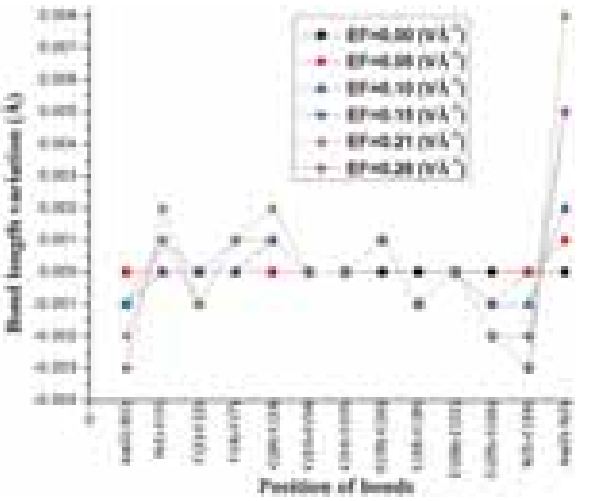


Fig. 3. Variation of bond lengths for various applied EFs with reference to zero EF.

Laboratory of Double Zeta) basis set, which provide effective core potential and the detailed description of the effect of heavy metal atoms in the molecule.[14–17] All these quantum chemical calculations were performed for various levels of applied electric field to study the structural properties (bond lengths, bond angles and torsion angles), bond topological (electron density, Laplacian of electron density, bond ellipticity and the eigen values) and the electrostatic properties of the molecule. Using Bader's theory of "Atoms in Molecules" (AIM) the

Table II. Bond angles (°) of the Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for the zero and various applied EFs.

Bonds	0.00	0.05	0.10	0.15	0.21	0.26
S–C–C bonds						
S(1)–C(1)–C(2)	120.3	120.6	120.7	120.9	121.0	121.1
S(1)–C(1)–C(6)	120.2	120.0	119.9	119.7	119.7	119.6
S(2)–C(24)–C(23)	120.3	120.7	120.9	121.7	121.9	123.0
S(2)–C(24)–C(25)	120.4	119.9	119.7	118.9	118.8	117.8
Au–S–C bonds						
Au(1)–S(1)–C(1)	103.5	103.3	103.3	103.3	103.3	103.4
Au(2)–S(2)–C(24)	103.2	103.5	103.3	103.7	103.7	103.9

Table III. Torsion angles (°) of the Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for the zero and various applied EFs.

Bonds	0.00	0.05	0.10	0.15	0.21	0.26
S–C–C–C bonds						
S(1)–C(1)–C(2)–C(3)	175.4	175.6	175.7	176.0	176.1	176.2
S(2)–C(24)–C(25)–C(26)	175.9	174.9	175.0	174.1	174.1	174.5
Au–S–C–C bonds						
Au(1)–S(1)–C(1)–C(2)	91.3	87.1	85.8	82.8	82.2	80.6
Au(2)–S(2)–C(24)–C(25)	90.9	99.8	102.4	115.8	118.5	135.7

bond topological and the electrostatic properties have been calculated from the AIMPAC software.[18] The deformation density of the molecules has been plotted using the software *wfn2plots* and XD packages.[19] The electrostatic potential of the molecule has been plotted with Gview[20] to visualize the isosurface of positive and negative ESP regions of the molecule. The GaussSum program has been used to determine the density of states (DOS) at various levels of applied EFs.[21]

Table IV. MPA atomic charges (e) of the molecule for various applied EFs ($V Å^{-1}$).

Atoms	0.00	0.05	0.10	0.15	0.21	0.26
C-Atoms						
C(1)	−0.221	−0.221	−0.219	−0.218	−0.216	−0.214
C(2)	−0.197	−0.192	−0.190	−0.186	−0.185	−0.183
C(3)	−0.378	−0.377	−0.376	−0.375	−0.375	−0.374
C(4)	0.358	0.358	0.358	0.358	0.358	0.358
C(5)	−0.378	−0.378	−0.380	−0.381	−0.382	−0.383
C(6)	−0.200	−0.206	−0.209	−0.214	−0.216	−0.220
C(7)	0.339	0.339	0.339	0.338	0.338	0.337
C(8)	−0.380	−0.378	−0.377	−0.376	−0.375	−0.373
C(9)	−0.405	−0.404	−0.402	−0.401	−0.400	−0.398
C(10)	0.412	0.412	0.413	0.412	0.413	0.414
C(11)	−0.405	−0.407	−0.409	−0.411	−0.413	−0.415
C(12)	−0.380	−0.380	−0.381	−0.381	−0.382	−0.382
C(13)	−0.112	−0.115	−0.117	−0.119	−0.121	−0.124
C(14)	−0.113	−0.110	−0.108	−0.105	−0.102	−0.100
C(15)	0.412	0.411	0.410	0.408	0.407	0.407
C(16)	−0.406	−0.403	−0.400	−0.398	−0.396	−0.395
C(17)	−0.378	−0.380	−0.380	−0.380	−0.379	−0.378
C(18)	0.339	0.339	0.339	0.338	0.338	0.338
C(19)	−0.379	−0.382	−0.384	−0.386	−0.387	−0.389
C(20)	−0.406	−0.406	−0.407	−0.408	−0.409	−0.411
C(21)	0.359	0.357	0.357	0.354	0.354	0.353
C(22)	−0.377	−0.379	−0.378	−0.381	−0.380	−0.387
C(23)	−0.200	−0.187	−0.183	−0.162	−0.158	−0.134
C(24)	−0.223	−0.221	−0.221	−0.213	−0.209	−0.184
C(25)	−0.197	−0.208	−0.212	−0.227	−0.231	−0.246
C(26)	−0.376	−0.380	−0.382	−0.385	−0.386	−0.388
H-Atoms						
H(2)	0.247	0.248	0.250	0.251	0.253	0.255
H(3)	0.236	0.239	0.242	0.246	0.249	0.252
H(5)	0.236	0.233	0.231	0.229	0.227	0.224
H(6)	0.247	0.244	0.241	0.237	0.233	0.229
H(8)	0.234	0.235	0.237	0.238	0.239	0.241
H(9)	0.242	0.245	0.249	0.252	0.255	0.258
H(11)	0.242	0.240	0.238	0.236	0.234	0.232
H(12)	0.234	0.230	0.227	0.224	0.221	0.217
H(16)	0.242	0.244	0.246	0.249	0.251	0.253
H(17)	0.233	0.238	0.241	0.245	0.249	0.251
H(19)	0.233	0.233	0.232	0.231	0.229	0.227
H(20)	0.242	0.239	0.235	0.231	0.228	0.224
H(22)	0.235	0.239	0.241	0.244	0.246	0.247
H(23)	0.247	0.249	0.251	0.250	0.252	0.248
H(25)	0.247	0.246	0.244	0.242	0.239	0.233
H(26)	0.235	0.233	0.230	0.228	0.225	0.220
S-Atoms						
S(1)	0.084	0.081	0.076	0.073	0.067	0.062
S(2)	0.085	0.082	0.080	0.069	0.063	0.043
Au-Atoms						
Au(1)	−0.054	−0.057	−0.060	−0.064	−0.068	−0.072
Au(2)	−0.054	−0.043	−0.032	−0.014	0.001	0.025

3. RESULTS AND DISCUSSION

3.1. Structural Properties

Figure 2 depicts the optimized geometry of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecular wire for the zero and the maximum applied EF 0.26 $V Å^{-1}$. The 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecular wire has a four benzene rings and the Au atoms are attached at the ends of the molecule through thiol atoms. The thiol atom forms good link between the conjugated 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) and the Au atom.[22] Generally, the conducting molecules are very sensitive to the applied external electric field. Often, the field also alters the conformation of the molecules, which leads to actual change in the transport properties of the molecules.[23] Hence, to understand the structural stability of the molecules over the range of applied field, it is essential to compare its zero field geometry with the applied electric fields. For the zero field, the C–C bond distances of the molecule are vary from 1.400 Å to 1.420 Å. As the field increases (0.00 to 0.26 $V Å^{-1}$), these distances are slightly modified and the value is ∼1.421 Å. The maximum observed variation is 0.024 Å. However, the C–H bond lengths are ∼1.087 Å, almost remains same for the various applied EFs (Table I). Initially, the zero field distances of S(1)–C(1) and S(2)–C(24) bonds are found to be unequal and the values are 1.837 Å to 1.838 Å, respectively.

The maximum applied field (0.26 $V Å^{-1}$), the S(1)–C(1) bond distance in the L-end decreases from 1.837 Å to 1.835 Å. Whereas, the value of the S(2)–C(24) bond distance in the R-end also increases from 1.838 Å to 1.841 Å

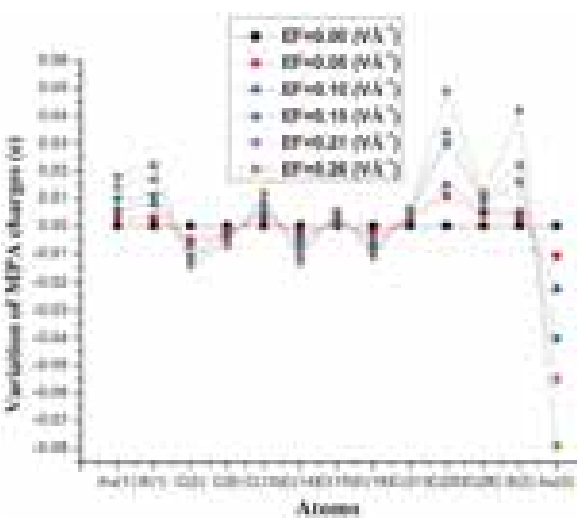


Fig. 4. Variation of MPA charges of the molecule for the zero and various applied EFs.

RESEARCH ARTICLE

(Fig. 2). The variation in the L-end (0.002 Å) is smaller than the R-end (0.003 Å) this values are matched with reported theoretical values.[24–26] Especially, the zero field terminal bonds of Au–S bonds at either ends are found to be equal and these value is 2.400 Å respectively. When the maximum applied field (0.26 VÅ^{-1}), the L-end of the Au(1)–S(1) bond length is slightly increases to 2.403 Å; whereas, the value of the R-end of Au(2)–S(2) bond length is decreases to 2.392 Å; also the variations in both ends are unequal. For the maximum applied field (0.26 VÅ^{-1}), the bond distance various from 0.003 and 0.008 Å to L-end and R-end. This small difference is attributed to the applied field lengthening the Au–S bond through by shrinking the S–C bond distance in the wire (Fig. 3) and the distances of S–C and Au–S bonds are very close to the previously reported theoretical values.[24]

Table V. NPA atomic charges (e) of the molecule for various applied EFs (VÅ^{-1}).

Atoms	0.00	0.05	0.10	0.15	0.21	0.26
C-Atoms						
C(1)	−0.163	−0.164	−0.164	−0.165	−0.166	−0.167
C(2)	−0.194	−0.193	−0.192	−0.191	−0.189	−0.188
C(3)	−0.200	−0.199	−0.197	−0.195	−0.194	−0.192
C(4)	−0.035	−0.033	−0.032	−0.030	−0.028	−0.026
C(5)	−0.200	−0.202	−0.204	−0.205	−0.207	−0.209
C(6)	−0.194	−0.196	−0.198	−0.201	−0.203	−0.205
C(7)	−0.042	−0.045	−0.047	−0.049	−0.052	−0.054
C(8)	−0.201	−0.199	−0.197	−0.196	−0.194	−0.193
C(9)	−0.168	−0.167	−0.166	−0.165	−0.163	−0.163
C(10)	−0.105	−0.103	−0.101	−0.098	−0.096	−0.093
C(11)	−0.168	−0.170	−0.172	−0.174	−0.175	−0.178
C(12)	−0.201	−0.202	−0.204	−0.205	−0.207	−0.208
C(13)	−0.002	−0.006	−0.010	−0.015	−0.019	−0.023
C(14)	−0.002	0.002	0.006	0.010	0.015	0.019
C(15)	−0.105	−0.108	−0.110	−0.113	−0.115	−0.119
C(16)	−0.168	−0.166	−0.164	−0.162	−0.160	−0.158
C(17)	−0.200	−0.200	−0.198	−0.198	−0.196	−0.196
C(18)	−0.042	−0.040	−0.037	−0.034	−0.031	−0.027
C(19)	−0.200	−0.203	−0.205	−0.207	−0.209	−0.211
C(20)	−0.168	−0.169	−0.170	−0.172	−0.173	−0.174
C(21)	−0.035	−0.038	−0.041	−0.046	−0.049	−0.058
C(22)	−0.200	−0.199	−0.197	−0.194	−0.192	−0.188
C(23)	−0.194	−0.193	−0.192	−0.194	−0.193	−0.199
C(24)	−0.163	−0.161	−0.159	−0.155	−0.152	−0.145
C(25)	−0.194	−0.197	−0.199	−0.205	−0.208	−0.218
C(26)	−0.200	−0.202	−0.203	−0.204	−0.205	−0.203
H-Atoms						
H(2)	0.235	0.237	0.238	0.239	0.241	0.242
H(3)	0.222	0.224	0.227	0.229	0.231	0.233
H(5)	0.222	0.221	0.219	0.218	0.216	0.214
H(6)	0.235	0.233	0.230	0.228	0.225	0.223
H(8)	0.222	0.223	0.224	0.225	0.226	0.227
H(9)	0.227	0.229	0.232	0.234	0.236	0.238
H(11)	0.227	0.226	0.224	0.223	0.222	0.220
H(12)	0.222	0.220	0.218	0.216	0.213	0.211
H(16)	0.227	0.229	0.230	0.232	0.234	0.235
H(17)	0.221	0.224	0.226	0.229	0.231	0.233
H(19)	0.222	0.221	0.220	0.219	0.217	0.216
H(20)	0.227	0.225	0.223	0.220	0.218	0.216
H(22)	0.222	0.224	0.226	0.227	0.229	0.230
H(23)	0.235	0.237	0.239	0.240	0.241	0.239
H(25)	0.235	0.234	0.232	0.230	0.228	0.225
H(26)	0.222	0.220	0.218	0.216	0.214	0.212
S-Atoms						
S(1)	−0.135	−0.136	−0.138	−0.140	−0.143	−0.145
S(2)	−0.134	−0.139	−0.144	−0.154	−0.161	−0.169
Au-Atoms						
Au(1)	0.195	0.192	0.189	0.184	0.180	0.176
Au(2)	0.195	0.208	0.221	0.247	0.263	0.301

However, the electric field made a significant variation in the selected bond angles and torsion angles of the terminal S–C–C and Au–S–C bonds. For the zero field distances of S–C–C bonds are found to be unequal and the values are 120.2° to 120.4° (Table II). As the field increases, the distance in the left-end increases ∼121.1° and the value of the same bond in the right-end also increases ∼123.0°. The zero field distances of Au–S–C bonds are found to be unequal and the values from 103.2 to 103.5°. As the field increases, the bond distance in the left-end decreases to (∼103.4°), whereas the value of the same bond in the right-end also increases (∼103.9°). Similarly, for the zero fields, the torsion angles for thiol linked S–C–C–C bonds are 175.4°. Interestingly, as the field increases this angle (L-end) increases from 175.4° to 176.2°. In the other R-end, the torsion angle decreases from 175.9° to 174.5°. The zero field torsion angles of Au–S–C–C bonds L-end is 91.3°; as the field increases, the maximum applied field (0.26 VÅ^{-1}) this angle large decreases to 80.6°, whereas, the R-end of torsion angles increases from 90.9° to 135.7° (Table III). When compared with the R-end, the variation in the terminal group of the L-end has been varied significantly, which indicates that this group is highly sensitive to the positive than the negative fields.

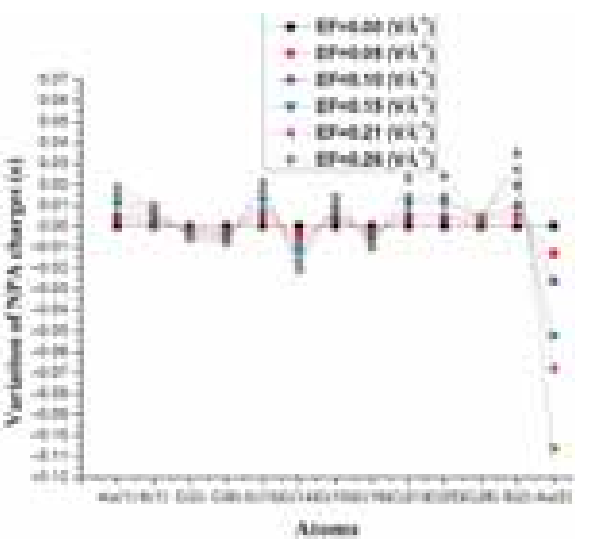


Fig. 5. Variation of NPA charges of the molecule for the zero and various applied EFs.

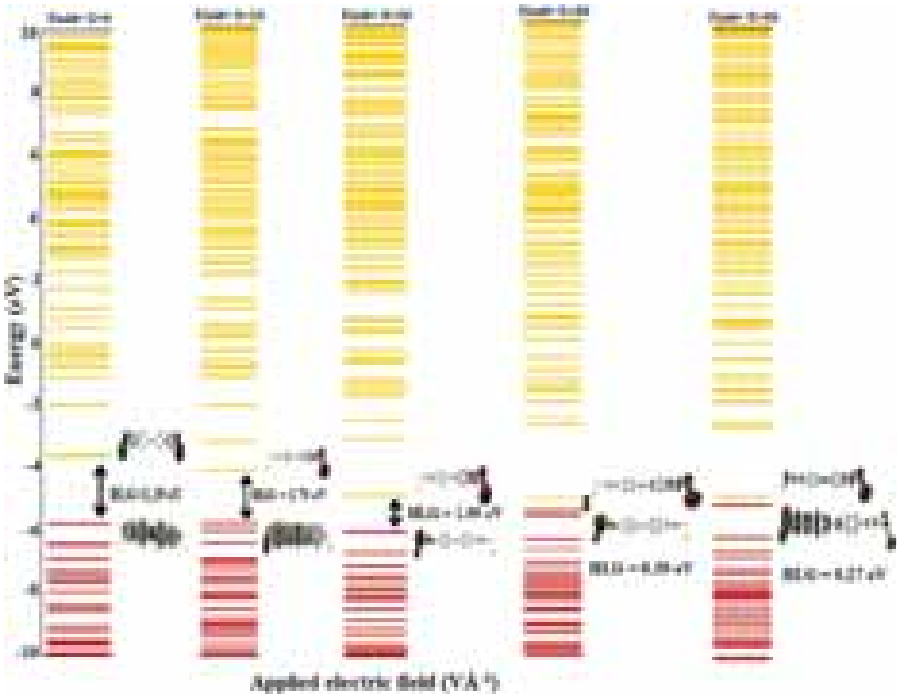


Fig. 6. The MO energy level diagrams of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for zero and various applied EFs.

3.2. Atomic Charges

The scheme of point charge distribution of molecules plays a major role on understanding the chemical reactivity and electrostatic potential.[27–30] In generally, Mullikan population analysis (MPA) and Natural population analysis (NPA)[27] methods are used to calculate the atomic charges of molecules.[31] Both methods are basis set dependent, in which a small change in basis set can cause a large change in the calculated net charges. However, in NPA model the charges are estimated on the basis of natural orbital population, it provides better charge distribution. Further, both models predict almost negative charge for all C-atoms in the molecule. The MPA charges of all C-atoms for zero field from −0.406e to 0.412e. When the applied field increases, the charges almost increase up to 0.414e. The maximum observed variations in all C-atoms are 0.049e. The linker atoms on either ends [S(1) and S(2)] possess same MPA charge (0.084e to 0.085e) for zero field, as the field increases this positive value decreases to 0.062e for S(1), whereas the charge of S(2) decreases to 0.043e. The maximum variations in the both ends are 0.042e. For the zero field, the charge of Au atom at either ends are found to be equal (−0.054e). As the field increases, the charge of Au atom in the L-end (negative value) increases to −0.072e, similarly, for the Au present at the R-end, the value increases to 0.025e. The maximum variations in the both ends are 0.079e. The MPA charges for the zero and various applied EFs of the molecule are presented in Table IV. The variations of MPA charges for various EFs with reference to zero field are plotted as in Figure 4.

For the zero field, the NPA charge for all C-atoms are found positive and negative values; the charges vary for the increase of field. When the applied field increases from 0.00 VÅ^{-1} to 0.26 VÅ^{-1}, the charge increases to 0.019e. The linker atoms on either ends [S(1) and S(2)] possess same NPA charge (−0.134 to −0.135e) for zero field, as the field increases, this negative value

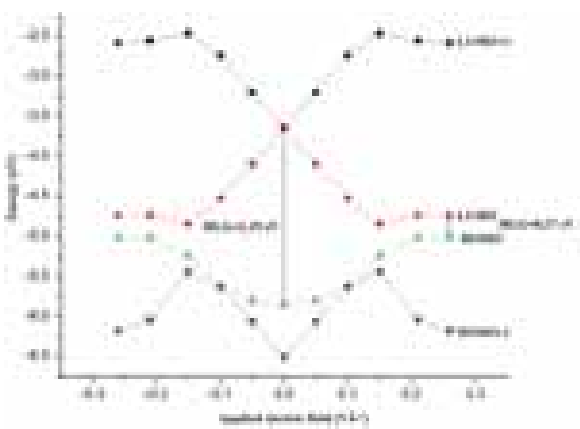


Fig. 7. Energy level diagram of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) based molecule for the zero and various applied EFs.

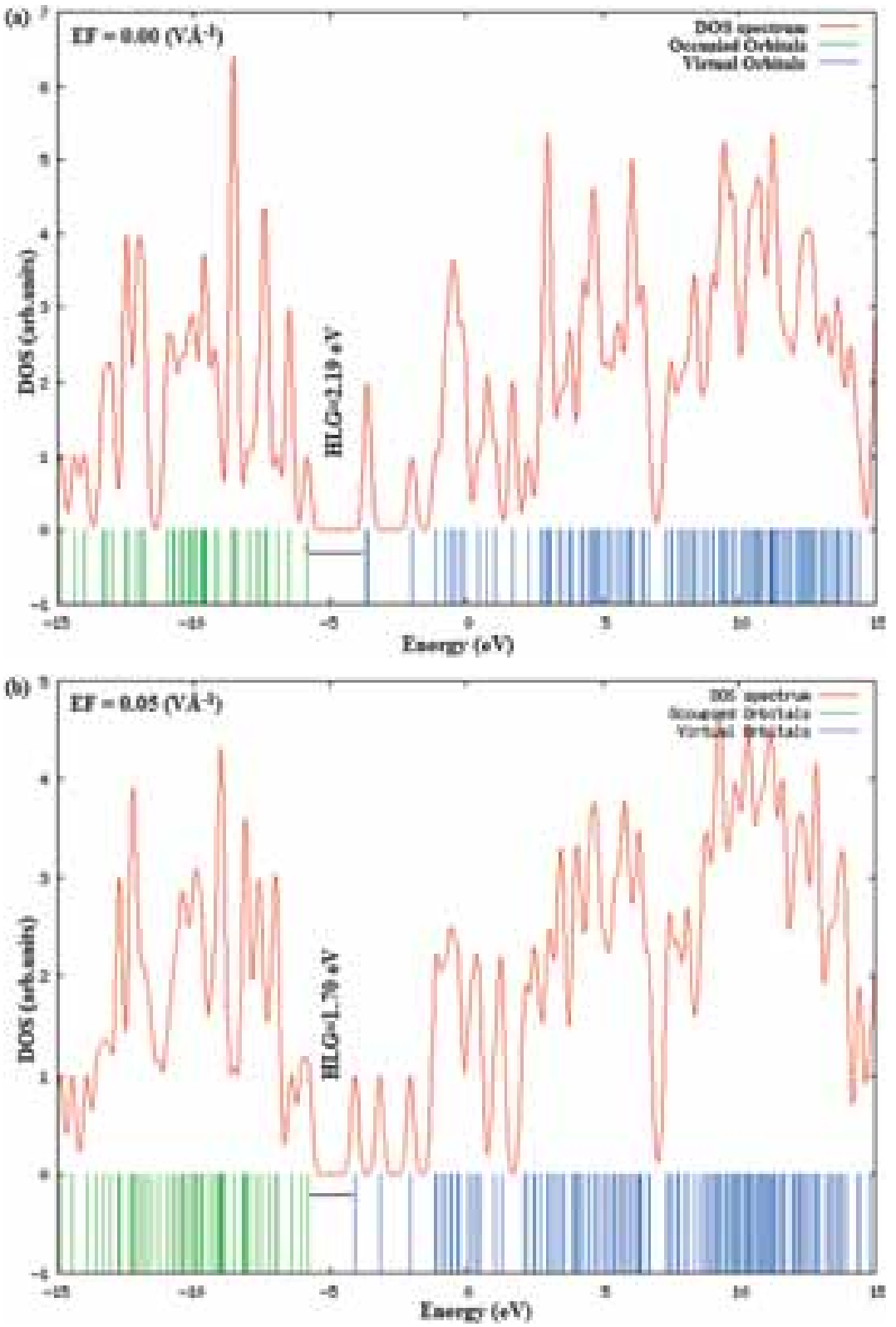


Fig. 8. Continued.

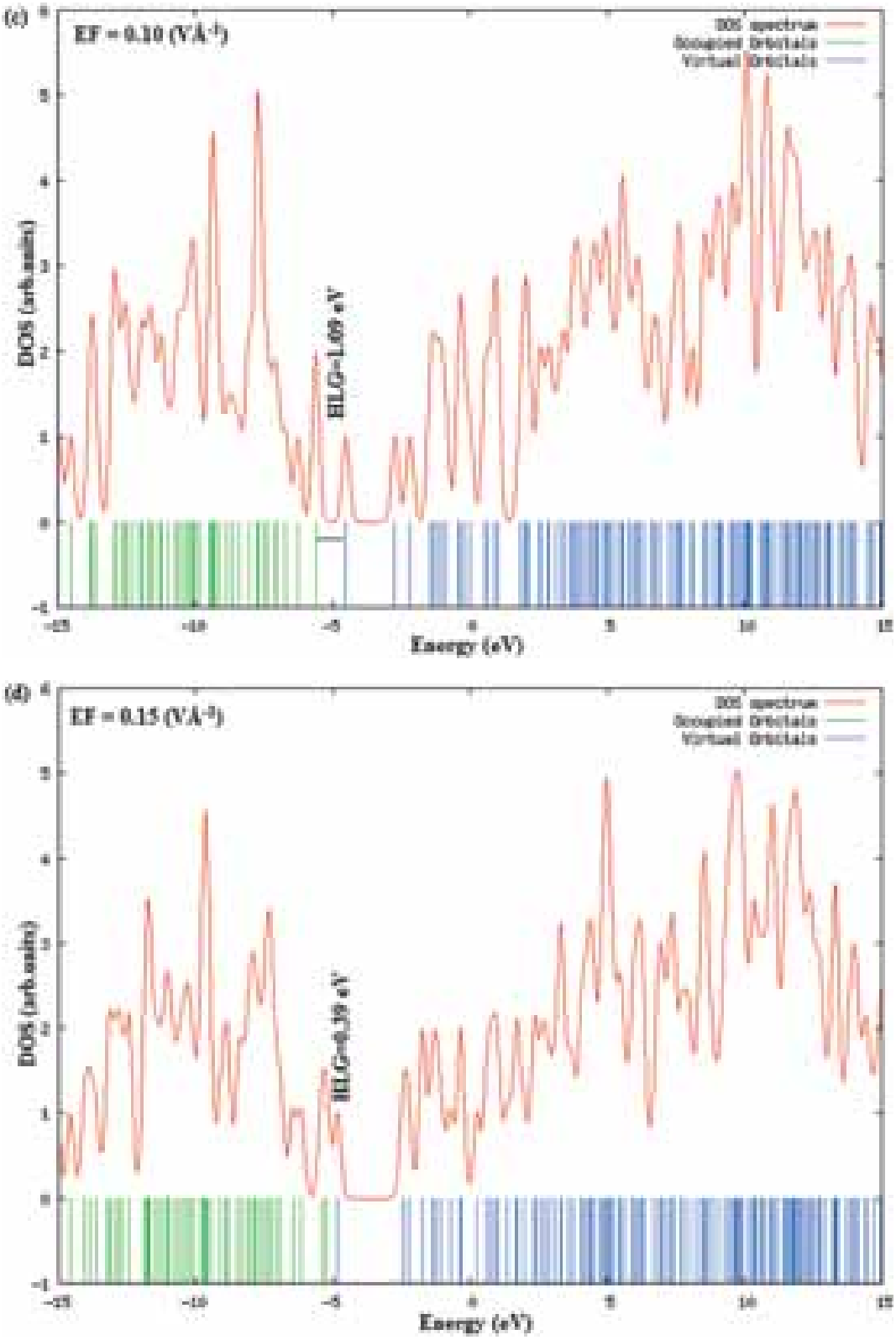


Fig. 8. Continued.

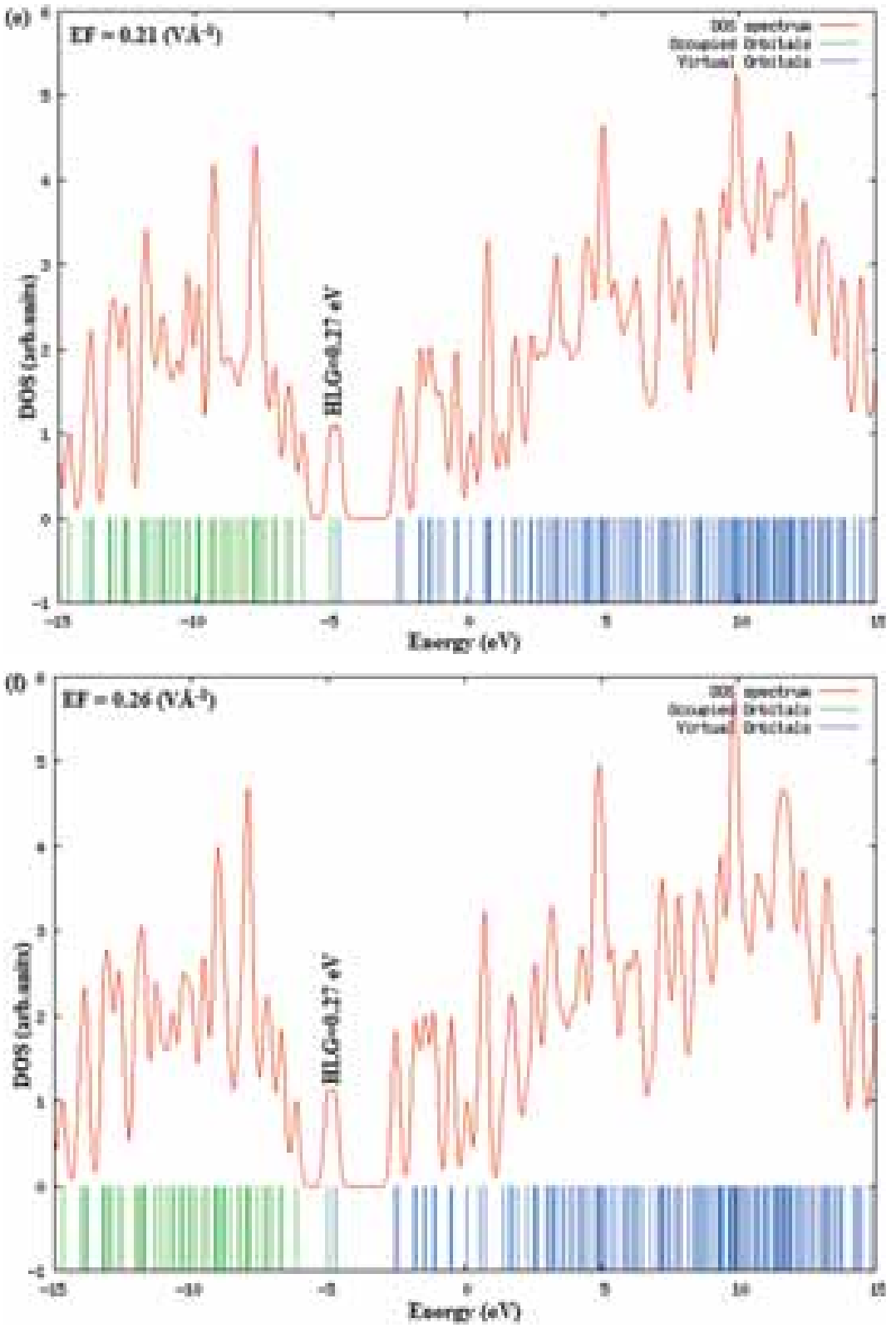


Fig. 8. The density of states (DOS) for the zero and various applied EFs.

increases to −0.145e for S(1), whereas the charge of S(2) increases to −0.169e. The maximum variations in the both ends are 0.035e. For the zero field, the charge of Au atom at either ends are found to be equal (0.195e). As the field increases, the charge of Au atom in the L-end decreases to 0.176e, similarly, for the Au present at the R-end, the value increases to 0.301e. The maximum variations in the both ends are 0.106e. The large variation of charges in the terminal atoms have been observed for various applied EFs. The difference of NPA charge distribution for zero and various applied EF is listed in Table V. The variations of NPA charges for various EFs with reference to zero field are plotted as in Figure 5.

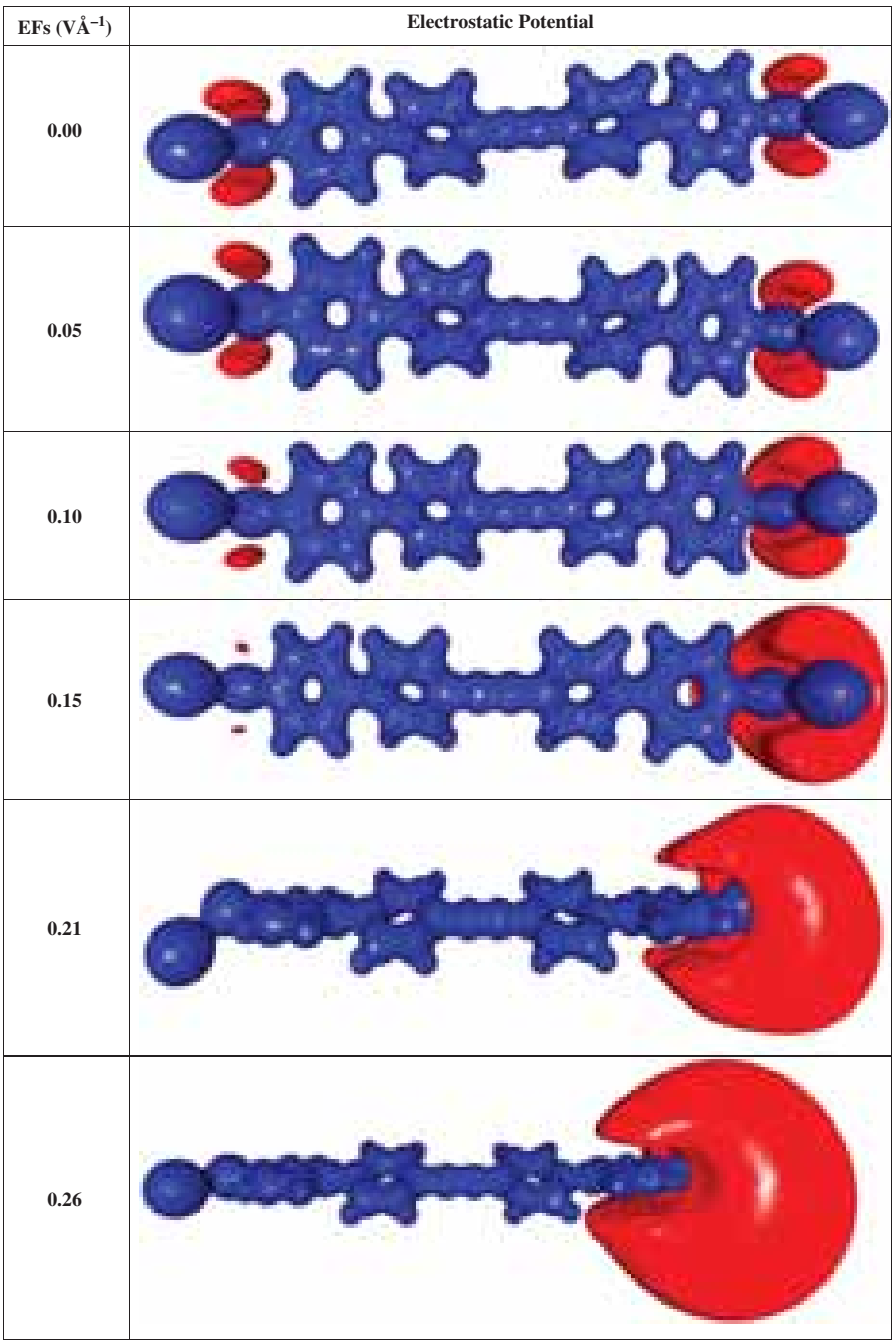


Fig. 9. Molecular electrostatic potential of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for zero and various applied electric fields (0.00–0.26 VÅ^{-1}). Blue: positive potential (0.05e Å^{-1}), Red: negative potential (−0.05e Å^{-1}).

3.3. Molecular Orbital Analysis

The electrostatic transport process through the molecule, it is based on the coupling between the molecule and the two electrodes and the energy gap (HLG) between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of the molecule. Molecular orbitals (MOs) are created, when the atomic orbitals are brought together. Each molecular orbital has a calculated energy level; it is assumed that the electrons first will occupy the lowest energy level Molecular orbitals.[32–34] Hence, it is necessary to examine the variations in HLG and molecular orbital energy levels[35,36] for the applied EFs. Figure 6 represents the energy level diagrams of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for zero and various applied EFs. For the applied field (0.00–0.26 V$Å^{-1}$), the HLG decreases from 2.19 eV to 0.27 eV. Figure 7 shows the energy levels of the molecule for various applied EFs.

The density of states (DOS) shows here also conforms, when the field various the HLG in which the green lines indicate the HOMO and the blue lines indicates LUMO; the decrease of HLG as shown in Figures 8(a)–(f). Here, the hybridization of the molecular level with that of the gold atom broadens the density of states spectrum peaks. At the zero field, the DOS peaks are in minima, indicates the discrete molecular level with HLG, 2.19 eV, further increase of field to 0.26 V$Å^{-1}$, both HOMO and LUMO levels approach each other and their gap decreases to 0.27 eV. The energy level variations of the molecule are almost symmetric and hence the conductivity of the molecule is irrespective of the direction of external EFs. The large decrease of HLG facilitates large electron conduction through the molecule.[37]

3.4. Molecular Electrostatic Potential

The Electrostatic Potential shows the 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for the zero and maximum applied EF (0.26 V$Å^{-1}$) is also shown in Figure 9. This ESP of this molecule which is identifies the charged region of the molecules.[38] For the zero and various applied fields, in which the Blue are indicates positive potential (0.05e $Å^{-1}$), and the Red are indicates negative potential (−0.05e $Å^{-1}$) of the molecule. The Au–S bond regions exhibit high negative ESP. It shows that the charged regions (red) molecules are opposing contributions from the nuclei and the electrons.[39,40] For the zero field, the negative ESP is concentrated around the S-atoms, which are present at either ends of the 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule with S-atoms and the rest of the molecule carries positive ESP. For the increase of applied electric field from (0.26 V$Å^{-1}$), the negative ESP at the L-end of the molecule gradually decreases for each applied electric fields and it disappears, while at the same R-end gradually increases and finally spreads around the right end edges of the molecule

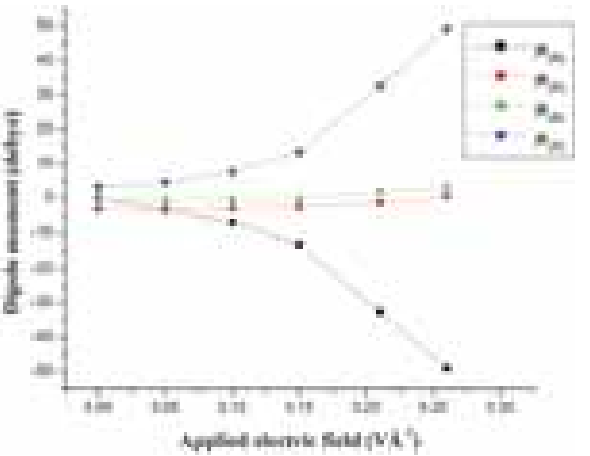


Fig. 10. Molecular dipole moment of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for the zero and various applied EFs.

for the applied maximum two fields (0.21 V$Å^{-1}$ and 0.26 V$Å^{-1}$). This shows that the applied electric field increases the charges appear to be carried slowly from left-end to right-end.

3.5. Molecular Dipole Moment

The natures of chemical bonds are not completely intersection in the optimized molecules. The dipole moment can be calculated that include the electronic correlation effects shows the traditional methods[41] of the 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule. The applied electric fields, the dipole moment of the molecules are measured.[42,43] Hence, the ability of electron transport through the molecule, to determine the dipole moment for various applied EFs. The dipole moment of the molecule can be calculated from zero and the maximum applied electric fields. The molecular dipole moments for zero fields are 3.14 Debye. The molecule becomes highly polarized for the maximum applied electric fields (0.26 V$Å^{-1}$); the resultant maximum dipole moments are 49.17 Debye. For interestingly, the higher dipole moment values are good agreement and electrical conductivity of the molecules for reported values.[44] The large variations of Au and S substituted 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule for the zero and various applied EFs are plotted as shown in Figure 10.

4. CONCLUSIONS

The present research computational study on 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecular wire describes the structural parameters the electrical characteristics for zero and various external applied fields. For the applied field, the Au–S bond length at the L-end is found slightly longer than the R-end; this difference may be due to the gold atom at the left end experiences stronger electric

field than the same at the R-end. As the field increases, the structural observed variations are small; however, in most cases it is found to be systematic and almost uniform. When the applied field increases, the hybridization of molecular energy levels broaden the DOS and decreases the HLG from 2.19 eV to 0.27 eV, the large band gap energy decreases at the high electric field, facilitates to have been electrical conductivity is very high of this molecule. The applied electric fields from zero to 0.26 $V Å^{-1}$, which polarizes the molecule, in the effect of dipole moment increases from 3.14 D to 49.17 D. Predominantly, the terminal groups of the 1,2-di([1,1′-biphenyl]-4-yl)ethyne (DBPE) molecule are found to be more sensitive to applied electric field in the comparison of other molecular region. The structural confirmation, distribution of atomic charges, molecular orbital analysis, electrostatic properties and molecular dipole moment of the molecule in the study may support to design several kinds of new charge transfer molecules.

References

1. F. Chen, J. Hihath, Z. Huang, X. Li, and N. J. Tao, *Annu. Rev. Phys. Chem.* 58, 535 **(2007)**.
2. G. Cuniberti, G. Fagas, and K. Richter, *Lect. Notes Phys*. 680, 1 **(2005)**.
3. C. Xue and F.-T. Luo, *Tetrahedron* 60, 6285 **(2004)**.
4. D. J. Wold and C. D. Frisbie, *J. Am. Chem. Soc.* 122, 2970 **(2000)**.
5. A. Aviram, C. Joachim, and M. Pomerantz, *Chem. Phys. Lett.* 146, 490 **(1988)**.
6. J. Lee, G. Lientschnig, F. Wiertz, M. Struijk, R. Janssen, R. Egberink, D. N. Reinhoudt, P. Hadley, and C. Dekker, *Nano Lett.* 3, 113 **(2003)**.
7. M. Di Ventra, S. T. Pantelides, and N. D. Lang, *Phys. Rev. Lett.* 84, 979 **(2000)**.
8. M. A. R. Xue, *Phys. Rev. B* 68, 115406 **(2003)**.
9. C. Bombis, F. Ample, L. Lafferentz, H. Yu, S. Hecht, C. Joachim, and L. Grill, *Angew. Chem. Int. Ed*. 48, 9966 **(2009)**.
10. M. Ventra, Di. S. T. Pantelides, and N. D. Lang, *Phys. Rev. Lett.* 84, 979 **(2000)**.
11. J. M. Seminario and P. Politzer, Modern Density Functional Theory: A Tool For Chemistry, 1st edn., Elesvier, New York **(1995)**, Vol. 2.
12. W. R. Wadt and P. J. Hay, *J. Chem. Phys.* 284, 82 **(1985)**.
13. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian Inc., Pittsburgh, PA **(2003)**.
14. A. D. Becke, *J. Chem. Phys*. 98, 5648 **(1993)**.
15. P. J. Hay and W. R. Wadt, *J. Chem. Phys.* 82, 270 **(1985)**.
16. E. Papajak, J. Zheng, X. Xu, R. L. Hannah, and D. G. Truhlar, *J. Chem. Theory Comput.* 7, 3027 **(2011)**.
17. A. C. Tsipis, A. G. Orpen, and J. N. Harvey, *Dalton Trans.* 17, 2849 **(2005)**.
18. F. W. Biegler-Konig, R. F. W. Bader, and T. H. J. Tang, *J. Comput. Chem.* 3, 317 **(1982)**.
19. A. Volkov, P. Macchi, L. J. Farrugia, C. Gatti, P. Mallinson, T. Richter, and T. Koritsanszky, XD2006—A computer program for multipole refinement, topological analysis of charge densities and evaluation of intermolecular energies from experimental or theoretical structure factors, Version 5.33. State University of New York at Buffalo, NY; University di Milano, Italy; University of Glasgow, UK; CNR-ISTM Milano, Italy; Free University of Berlin, Germany **(2007)**.
20. A. Frish, E. Lecn, A. B. Nielson, A. Holder, A. J. Roy Dennigton, and T. A. Keith, Gaussian Incorporated, Pittsburgh PA **(2003)**.
21. N. O. Boyle, GaussSum, Revision 2.1, http://GaussSum.sf.net.
22. R. Ahmadi and M. Darvish Ganji, *JUST* 33, 33 **(2007)**.
23. M. Ventra, Di. N. D. Lang, and S. T. Pantelides, *J. Chem. Phys.* 281, 189 **(2002)**.
24. M. Kondo, T. Tada, and K. Yoshizawa, *J. Phys. Chem. A* 108, 9143 **(2004)**.
25. K. Yoshizawa, T. Tada, and A. Staykov, *J. Am. Chem. Soc.* 130, 9406 **(2008)**.
26. H. Basch and A. M. Ratner, *J. Chem. Phys.* 119, 11926 **(2003)**.
27. C. Campana, B. Mussard, and T. K. Woo, *J. Chem. Theory Comput.* 5, 2866 **(2009)**.
28. R. S. Mulliken, *J. Chem. Phys*. 23, 1833 **(1955)**.
29. F. Martin and H. Zipse, *J. Comput. Chem.* 26, 97 **(2005)**.
30. Y. Ye, M. Zhang, and J. Zhao, *J. Mol. Strut: Theochem*. 822, 12 **(2007)**.
31. R. F. W. Bader, Atoms in Molecules—A Quantum Theory, Oxford University Press, Oxford, U.K. **(1990)**.
32. W. B. Davis, W. A. Svec, M. A. Ratner, and M. R. Wasielewski, *Nature* 60, 396 **(1998)**.
33. J. Li, J. K. Tomfohr, and O. F. Sankey, *Physica E: Low-Dimensional Systems and Nanostructures* 19, 133 **(2003)**.
34. D. Rai, H. Joshi, A. D. Kulkarni, S. P. Gejji, and R. K. Pathak, *J. Phys. Chem. A* 111, 9111 **(2007)**.
35. R. J. Magyar, S. Tretiak, Y. Gao, H. L. Wang, and A. P. Shreve, *Chem. Phys. Lett.* 401, 149 **(2005)**.
36. J. Chen, W. Wang, M. A. Reed, A. M. Rawlett, D. W. Price, and J. M. Tour, *Appl. Phys. Lett.* 77, 1224 **(2000)**.
37. J. Q. Lu, J. Wu, H. Chen, W. Duan, B. L. Gu, and Y. Kawazoe, *Physics Letters A* 323, 154 **(2004)**.
38. I. G. Csizmadia, Theory and Practice of MO Calculations on Organic Molecules. I. Elsevier, Amsterdam **(1976)**.
39. J. S. Murray and K. Sen, Theoretical and Computational Chemistry, University of New Orleans, New Orleans, LA, USA **(1996)**, Vol. 3.
40. H. R. Petrucci, S. W. Harwood, F. G. Herring, and D. M. Jeffrey, General Chemistry: Principles and Modern Application and Basic Media Pack, 9th edn., Pearson Edu. Inc., New Jersey **(2007)**.
41. G. S. Maciel and E. Garcia, *Chem. Phys. Lett.* 409, 29 **(2005)**.
42. J. Mazurkiewicz and P. Tomasik, *Natural Science* 4, 276 **(2012)**.
43. D. Farmanzadeh and Z. Ashtiani, *Struct. Chem.* 21, 691 **(2010)**.
44. D. Rai, A. D. Kulkarni, S. P. Gejji, and R. K. Pathak, *J. Chem. Phys.* 128, 0343101 **(2008)**.

Received: 19 April 2018. Accepted: 18 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1528–1529, 2018

A SPECIAL SECTION

A Special Section on the First International Conference on Cyber Physical Systems for Next Generation Computing (CPSNGC2018), Sathyamangalam, Tamilnadu, India, 1–3 March, 2018

The First International Conference on Cyber Physical Systems for Next Generation Computing (CPSNGC2018) organized by department of Information Technology, Computer Science and Engineering successfully held at Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India during March 2018. The conference was scheduled over three days from Thursday 1st March to Saturday 3rd March 2018. The primary objective of the organizers was to host an comprehensive and affordable conference designed to reach as wide an spectators as possible. The delegate fee was established at an affordable rate for full delegates and for students which covered attendance of the full conference programme, morning and afternoon tea, and lunch. The ability to allow the low registration fee was facilitated by the generous support of the conference sponsors and we are extremely grateful to the sponsors such as IEEE (Institute of Electrical and Electronics Engineers), DRDO (Defense Research and Development Organization), ISTE (Indian Society for Technical Education), ICTACT (ICT Academy of Tamilnadu) and Wireline Solution India Private Limited.

The programme started each day at 8.30 am and finished at 17.00 pm. Parallel sessions were scheduled for the presentation of papers with session chairs from respective domain knowledge. On Friday and Saturday evenings, poster sessions were held to encourage the young researchers. Keynote speakers Dr. IR. Vinesh Thiruchelvam, Asia Pacific University of Technology and Innovation, Kualalumpur, Malaysia, Mr. Emanuel Kolb, ABB Private Limited, Germany and Dr. D. Sudharsan, ABB corporate research, Bangalore, India delivered the special key note related to the conference theme for 30 minutes. It was very much inspiring to find that all accepted participants in the field agreed to be present at the conference, even when offered only a 15-minute slot. The conference proved to be exceptionally popular, attracting 120 participants from 10 nations; and the papers contained in this special edition provide a measure of the diversity of subjects presented and discussed.

Cyber Physical Systems has been a topic of intense research interest over the past 5 years, especially with respect to studies utilizing Internet of Things, Artificial Intelligence, Cloud Computing and Data Analytics. The range of subjects covered in this special edition which includes prediction and forecasting techniques of data analytics, application of various machine learning algorithms, cyber physical systems and cloud security, wireless sensor networks, applications using Internet of Things and various image processing techniques.

In conclusion, Cyber physical systems remain an active field of research to enhance the networking of software and physical components in a smart way to improve the quality of our life. This conference has been built on the aims of the potential researcher to showcase their innovative ideas and a platform for the younger researchers in the field to present and discuss the findings.

Guest Editors
A. Bazila Banu
Department of Computer Science and Engineering
Bannari Amman Institute of Technology
Sathyamangalam, Erode District
Tamilnadu, India
D. Rajesh Kumar, T. Gopalakrishnan
Department of Information Technology
Bannari Amman Institute of Technology
Sathyamangalam, Erode District, Tamilnadu, India

 1546-1955/2018/15/1528/002 doi:10.1166/jctn.2018.7373

ABOUT THE GUEST EDITORS

A. Bazila Banu is an Associate Professor in the Department of Computer Science and Engineering at Bannari Amman Institute of Technology, India. She received her Ph.D. degree in Information and Communication Engineering at Anna University, India in 2015. She holds 16 years of professional experience includes academic and software Industry. She is an advisory board member for International Association of Data Science, Asia Pacific University of Technology and Innovation, Kualalumpur, Malaysia. She published 11 articles in National and International journals. Her Research interest includes Big Data and Data Analytics.

D. Rajesh Kumar is an Assistant Professor in the Department of Information Technology at Bannari Amman Institute of Technology, India. He holds a Ph.D. degree in Information and Communication Engineering by the Anna University, India. He has presented papers at conferences both National and International, published articles and papers in various journals, and contributed a chapter to the book Cognitive Data Science Methods and Models over IoT, Springer International Publishing. His research and publication interests include Wireless Sensor Networks and Cloud Computing. He is an Expert Advisory Panel Member of Texas Instruments Inc USA.

T. Gopalakrishnan is an Assistant Professor (Senior Grade) in the Department of Information Technology at Bannari Amman Institute of Technology, India. He received his Ph.D. degree in Information and Communication Engineering at Anna University, India in 2017. He holds 11 years of professional experience includes academic and R&D activity. He has published 19 articles in National and International journals. His Research interest includes Information Retrieval, Data Analytics and E-Learning.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1530–1533, 2018

RESEARCH ARTICLE

Level-Wised Directed Acyclic Graph Scheduling on Cloud Resources

C. K. Karthika Devi[1,*] and K. Kousalya[2]

[1] PG Scholar, School of Computer Science and Engineering, Kongu Engineering College, Perundurai 638052, India
[2] School of Computer Science and Engineering, Kongu Engineering College, Perundurai 638052, India

Cloud computing is a distributed computing model which enables developers to automatically deploy applications during task allocation and storage distribution. Cloud computing tends to share a pool of virtualized computer resources and equipments of computation, storage and information. In cloud computing environment, task scheduling is a key problem which needs to consider various factors restricting the user tasks and is responsible for selecting the most suitable resources of cloud computing for user task. Task scheduling algorithm is an NP complete problem which places an important role in cloud computing. The task can be dependent or independent. In the case of independent task, any task can be executed in any order. But in the case of dependent task predecessor task should be executed before the successor. In some cases it may not happen, where the successor task is executed first before the predecessor. This leads to increase in the execution time. To overcome these problems, Level-wised DAG scheduling and Enhanced Level-wised DAG scheduling are introduced. For this DAG are scheduled level by level. As soon as the predecessor task gets completed, the successor begins its scheduling. As the result of these scheduling methods, the makespan time gets significantly reduced as compared with others.

Keywords: Cloud Computing, DAG Scheduling, Level-Wised DAG Scheduling, Enhanced Level-Wised DAG Scheduling, Makespan Time.

1. INTRODUCTION

A heterogeneous system is outlined as a variety of various system resources, which may be native or geographically distributed that are utilized to execute computationally intensive applications. The potency of executing parallel applications on heterogeneous systems critically depends on the ways used to schedule the tasks of a parallel application. The target is to reduce the overall completion time or makespan. The task scheduling problem for heterogeneous systems is additional complicated than the different execution rates among processors and possibly different communication rates among different processors. A popular representation of an application is the directed acyclic graph (DAG), which includes the characteristics of an application program such as the execution time of tasks, the data size to communication between tasks and task dependences. The DAG scheduling drawback has been shown to be NP-complete even for the homogeneous case; so, analysis efforts during this field have been mainly targeted on getting low-complexity heuristics that manufacture good schedules. Within the literature, one of the most effective list-based heuristics is that the Heterogeneous Earliest finish Time (HEFT).[10]

2. LITERATURE REVIEW

Arabnejad and Barbosa[1] proposed a completely unique list-based scheduling algorithm referred as Predict Earliest Finish Time (PEFT) for heterogeneous computing systems. Algorithm is merely supported an Optimal Cost Table (OCT) that's used to rank tasks and for processor choice. The analysis and experiments supported at random generated graphs with numerous characteristics and graphs of real-world applications.

Bajaj and Agrawal[2] proposed a Task duplication-based scheduling formula for Network of Heterogeneous systems (TANH) which provides best results for applications described by Directed Acyclic Graphs (DAGs) provided an easy set of conditions on task computation and network communication time can be satisfied. To produce substantial improvement over existing work on the Task Duplication-based Scheduling algorithm (TDS) is shown.

Bittencourt and Madeira[3] deals with the problem of resources to request from a public cloud based on the current demand and on resources costs. HCOC: The Hybrid Cloud Optimized Cost scheduling algorithm. HCOC

*Author to whom correspondence should be addressed.

1546-1955/2018/15/1530/004 doi:10.1166/jctn.2018.7334

decides which resources should be leased from the public cloud and aggregated to the private cloud to provide sufficient processing power to execute a workflow within a given execution time. Presents extensive experimental and simulation results which show that HCOC can reduce costs while achieving the established desired execution time.

Canon and Jeannot[4] proposed completely different metrics and show a way to correlate each another. In addition, completely different methods for minimizing the makespan while increasing the robustness: from an evolutionary meta heuristic (best solutions however longer computation time) to additional easy heuristics creating approximations (medium quality solutions however quick computation time). Compare these completely different approaches by experimentation and show different approximations of the Pareto front for this bi-criteria problem.

Cordeiro et al.[5] proposed GGen—a unified and standard implementation of classical task graph generation methods used in the scheduling domain. It also provides an in-depth analysis of each generation method, emphasizing important graph properties that may influence scheduling algorithms.

Mao and Humphrey[7] projected an approach whereby the fundamental computing parts are virtual machines (VMs) of varied sizes/costs, jobs are such as workflows. The goal is to make sure all jobs are finished inside their deadlines at minimum financial price. It dynamically allocate/deallocate VMs and scheduling tasks on the foremost cost-efficient instances.

Raju et al.,[8] proposed a new strategy, OTA (Optimal Time Algorithm), which improves the effectiveness of speculative execution significantly. OTA do not consider the difference between the execution time of tasks on the same processors, they may form clusters of tasks that are not similar to each other.

Rodriguez et al.,[9] proposed a resource provisioning and scheduling strategy for scientific workflows on Infrastructure as a Service (IaaS) Clouds. It present an algorithm based on the meta-heuristic optimization technique, Particle Swarm Optimization (PSO).

Topcuouglu et al.,[10] projected a two novel scheduling rule for a bounded range for heterogeneous processors with an objective to at the same time meet high performance and quick planning time, that are known as Heterogeneous Earliest-Finish-Time (HEFT) algorithm and the Critical-Path-on-a-Processor (CPOP) algorithm.

3. PROBLEM STATEMENT

A scheduling system model is composed of heterogeneous computing environment. A DAG as $G = \{V, E\}$. $V = \{v_1, v_2, v_3 \ldots v_n\}$ where a vertex v_i represents a corresponding task t_i which will be executed consecutive on one computer. $E = \{e_1, e_2, e_3, \ldots, e_m\}$ is the precedence relationship between tasks because of data dependency. Therefore, a task cannot begin before its precedent task is completed. Given a task dependency edge $v_i \rightarrow v_j$, task t_i is that the parent task or predecessor of task t_j, and likewise task t_j is that the child task or successor of task t_i. E is set of communication cost between task.

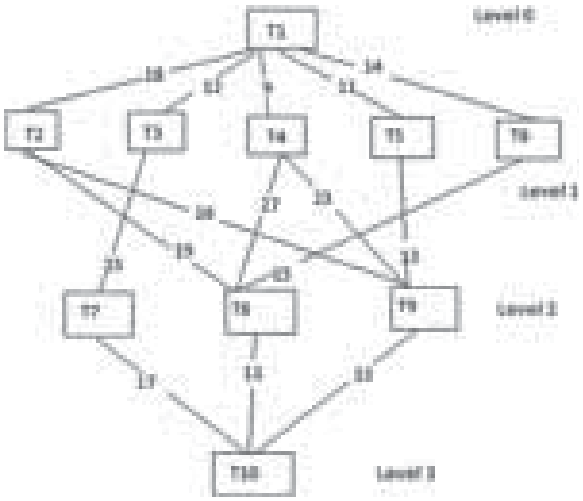


Fig. 1. Represents DAG jobs.

Table I. Estimation of matrix.

TASK	P1	P2	P3
T1	14	16	9
T2	13	19	18
T3	11	13	19
T4	13	8	7
T5	12	13	10
T6	13	16	9
T7	7	15	11
T8	5	11	14
T9	18	12	20
T10	21	7	16

In Figure 1 represents DAG jobs. Entry task is the first task in a DAG with no predecessor. Exit task is the last task in a DAG with no successor. Due to the heterogeneous environment in cloud where each computing node could have different capacities, here used a estimation time matrix $X[i, j]$ to indicate the running time of task t_i on the VM instance type j. The estimation time matrix is shown in Table I. p represents the number of processor.

4. SYSTEM IMPLEMENTATION

4.1. Level-Wised DAG Scheduling

The entry task is scheduled with the processor having the lowest estimation time in the multiprocessor environment. In the considered environment the processor $p3$ has the lowest estimation time for the task T1. So the task T1 is scheduled to the processor P3. After scheduling the task T1, the task in the first level of the DAG to be scheduled is selected by evaluating the parameters: estimation time of each task in the processor p, communication cost between

RESEARCH ARTICLE

the scheduled task and the task in first level, each processor starting time. The sum of these parameters represents the computation time of each task that is to be computed for each task in the first level in DAG.

The task to which the computation time is low is scheduled next. Now the processor starting time is changed with the task scheduled to it. Hence it is to be updated to the task that is to be scheduled further. Therefore the computation time for each task in the first level in the DAG is to be recomputed with the updated processor starting time. There can be one or more communication links in the DAG between tasks from one level to another level. In this case if there is one communication link from the task in first level to the task in the second level, the task in the second level is scheduled once the task in the first level is scheduled. Since the second level task is dependent only on one task in the first level. These are shown in Figure 2.

The task in the second level that are having more than one communication link has to be evaluated with the computation time. The computation time comprises of the parameters: the highest scheduled time for the task in the first level to the task in the second level, the highest communication link to the task and the estimation time of each processor. The sum of these parameters represents the computation time for the task in the second level in the DAG. The processor starting time has to be updated to the other task in the second level for the evaluation of the computation time once the task with the lowest computation time is scheduled. The process is repeated to all the level of the DAG.

4.2. Enhanced Level-Wised DAG Scheduling

In the Level-wised DAG scheduling algorithm, the makespan time is massive. In order to reduce the makespan time of the Level-wised DAG scheduling algorithm, the enhanced level-wised DAG scheduling algorithm is proposed. In enhanced level-wised DAG scheduling, the task T1 is scheduled to which the processor estimation time is low. The first level tasks are scheduled same as that of the DAG scheduling algorithm. The second level task that is to be scheduled is different from that of the DAG scheduling algorithm. If the task in the second level has more than one communication link, the sum of the communication link value and the ending time of the task in the first level that is linked to the task in the second level is computed for each task linked. The maximum value obtained from the sum is identified and it is used for the scheduling process.

The maximum value obtained for the specified task that is linked is selected and that value is added to the

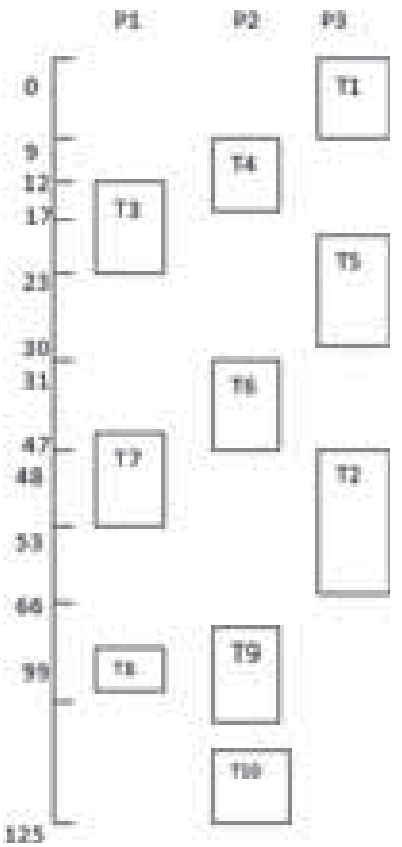


Fig. 2. Level-wised DAG scheduling.

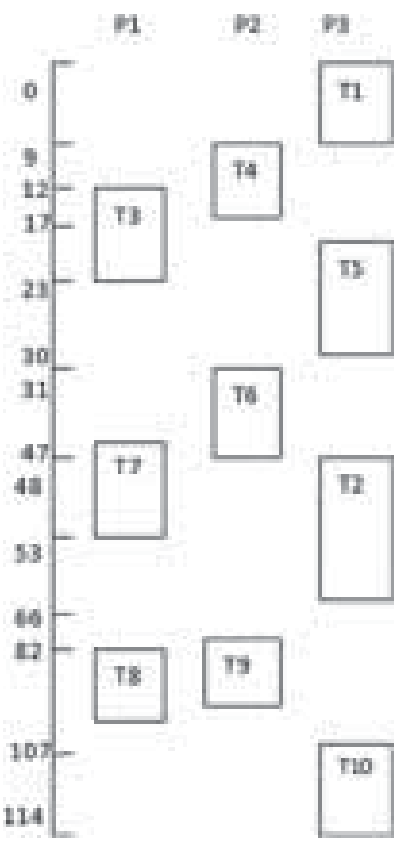


Fig. 3. Enhanced level-wised DAG scheduling.

estimation time of each processor of the task. The task scheduled to the processor with the lowest value obtained by the sum of the estimation time, selected communication link and the selected ending time of the task. The process is repeated until all the task are scheduled. The enhanced level-wised DAG scheduling is shown in Figure 3.

5. RESULT AND ANALYSIS

5.1. Performance Comparison

The Table II shows the comparison of makespan time of each scheduling algorithm. To compare the overall results of different algorithm with respect to their total makespan time, then presented in-depth analysis to investigate the reasons of our improvements.

The comparison of various algorithms makespan time is explained briefly below:

As analyzed from the Figure 4, PETS[13] has higher makespan time as compare to other algorithms. Enhanced Level-wised DAG scheduling has the minimum makespan time as compared to the other algorithm. It reduces 9.2% as compared to DAG scheduling.

Table II. Comparison of makespan time.

Algorithm	Makespan time
PETS	147
HCPT	142
HEFT	133
Lookahead	127
Level-wised DAG scheduling	125
Enhanced level-wised DAG scheduling	114

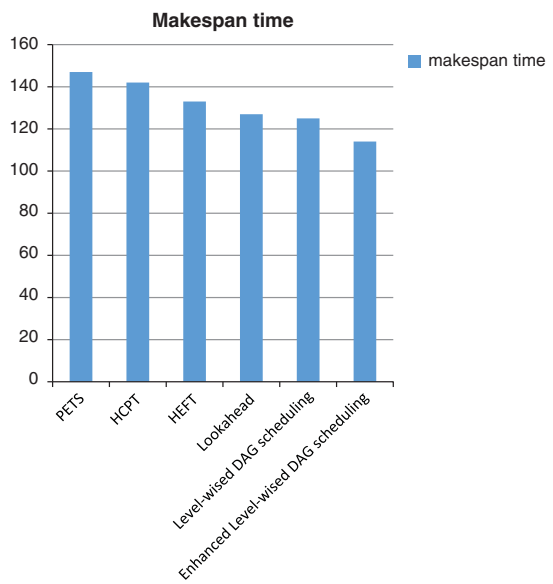


Fig. 4. Comparison of makespan time.

As conclude that Enhanced Level-wised DAG scheduling has minimize the makespan time as compare to the other scheduling algorithm.

6. CONCLUSION

At present there is no dependent and independent task. When the given task is dependent, one task needs the resources of other task. It leads to rescheduling of the task, so it may increase the total makespan time. This leads user to wait for long time to get task completed.

To overcome the drawback in the existing system a new Directed Acyclic Graph scheduling algorithm is proposed. It is for scheduling dependent task. In dependent task one task depends another task. So, using DAG, the scheduling is based on how they are scheduled. Most DAG scheduling algorithm schedules the task to decrease the makespan time which is not in order. In the proposed system the task are scheduled orderly according to their priority basis.

The future work is to consider the load balancing and to reduce the computation cost in the Directed Acyclic Graph.

Acknowledgment: I take immense pleasure to express my heartfelt thanks to my guide, Dr. K. Kousalya M.E., Ph.D., for her valuable ideas and suggestions, which have been very helpful in the project. I am grateful to all the faculty members of the Computer Science and Engineering Department, of Kongu Engineering college for their support.

References

1. H. Arabnejad and J. Barbosa, *IEEE Trans. Parallel Distrib. Syst.* 25, 682 **(2014)**.
2. R. Bajaj and D. Agrawal. *IEEE Trans. Parallel Distrib. Syst.* 15, 107 **(2004)**.
3. L. Bittencourt and E. Madeira, *J. Internet Serv. Appl.* 2, 207 **(2011)**.
4. L. Canon and E. Jeannot, *IEEE Trans. Parallel Distrib. Syst.* 21, 532 **(2010)**.
5. D. Cordeiro, G. Mounié, S. Perarnau, D. Trystram, J. M. Vincent, and F. Wagner, Random graph generation for scheduling simulations, *Proceedings of the 3rd International ICST Conference on Simulation Tools and Techniques* **(2010)**, pp. 60:1–60:10.
6. E. Ilavarasan and P. Thambidurai, *J. Computer Sciences* 3, 94 **(2007)**.
7. M. Mao and M. Humphrey, Auto-scaling to minimize cost and meet application deadlines in cloud workflows, *ACM/IEEE Conf. Supercomput.* **(2011)**, Vol. 49, p. 12.
8. R. Raju, J. Amudhavel, M. Pavithra, S. Anuja, and B. Abinaya, A heuristic fault tolerant MapReduce framework for minimizing makespan in hybrid cloud environment, *International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE)* **(2014)**, pp. 1–4.
9. M. A. Rodriguez and R. Buyya, *IEEE Trans. Cloud Comput.* 2, 222 **(2014)**.
10. H. Topcuouglu, S. Hariri, and M.-Y. Wu, *IEEE Trans. Parallel Distrib. Syst.* 13, 260 **(2002)**.

Received: 20 April 2018. Accepted: 16 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1534–1539, 2018

RESEARCH ARTICLE

A Design and Implementation of Routing Protocols for Low Power Lossy Networks

P. S. Dinesh[1,*] and C. Palanisamy[2]

[1] Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamil Nadu, India
[2] Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamil Nadu, India

Routing protocol for low-power and lossy networks (RPL) is the standard IPv6 based routing protocol for low power, lossy Networks (LLNs) proposed by IETF. It is proposed for networks with characteristics like small packet size, lossy links, low bandwidth, low data rate and low power resources. RPL is a single path routing protocol and the existing objective functions do not support creation of multiple routing paths between source and destination. Multipath routing can be used to achieve multi-fold objectives, including higher reliability, increased throughput, fault tolerance, congestion mitigation and hole avoidance. In this paper, M-RPL a multi-path extension of RPL is proposed that aims to provide temporary multipath routing during congestion over a path. Congestion is primarily detected by a forwarding node by monitoring packet delivery ratio and is mitigated by providing partially disjoint multipath routing. Detailed simulation analysis of M-RPL against RPL shows that using multi-path RPL reduces congestion and increases the overall throughput. Hence M-RPL is suitable for supporting high data rates as compared to single path RPL.

Keywords: Low Power and Lossy Networks, Multipath Routing, RPL, Congestion Control, Partially Disjoint Paths.

1. INTRODUCTION

Wireless sensor network (WSN) is a class of low power networks that are comprised of distributed and self-governing sensor nodes. These sensor nodes are used to sense the physical environment and they supportively forward the data in a multi-hop fashion to the sink node (data collector). Today, WSNs are used in various military applications, health care applications and industrial monitoring.

Low power lossy network (LLN) is a terminology that refers to networks that are comprised of miniature sized battery powered devices having lossy wireless connectivity. Examples of LLNs include wireless sensor networks and wireless personal area networks (WPAN). Design of routing protocols that can select reliable and low delay paths for providing high data rates is an uphill task in LLNs because of poor link connectivity and resource constrained nature of devices. Internet engineering task force (IETF) in this regard established a research group termed routing over low power lossy network (ROLL) for the design of a routing protocol suitable for LLNs. The objective of ROLL is to design a sophisticated routing protocol that establishes reliable paths, promptly reacts to link failure, utilize minimum energy and reduces computational cost for routing. The reason for the formation of new routing protocol by ROLL is that the existing routing protocols are not deemed suitable for LLNs.[1] The ROLL working group published RPL protocol for low power lossy network in 2008.[2]

RPL is a gradient-based proactive routing protocol with bidirectional links that builds directed acyclic graphs (DAGs) based on routing metrics and constraints.[3] RPL can create one or more destination oriented DAG (DODAG) for every root (sink) node within the network. DODAG root is the main root node which constructs the complete DODAG. To build DODAG, root node first multicast DODAG information object (DIO) with initial rank value 1. Rank defines the individual node positions within the respective DODAG.[2] Apart from rank value, the DIO contains information about objective function, IDs, routing cost, related metrics and network information.[2] Different objective functions are proposed in RPL such as objective function zero (OF0)[4] and the minimum rank with hysteresis objective function (MRHOF)[5] for the construction of DODAG. RPL strives to minimize the cost for reaching the root (sink) from any node in the LLN using an object function.[1] Neighbors of root node receive DIO message and use this information to update their rank, join DODAG and choose preferred parent by sending feedback to the

*Author to whom correspondence should be addressed.

1546-1955/2018/15/1534/006 doi:10.1166/jctn.2018.7335

root. The best preferred parent is used by a child node for routing based on expected transmission count (ETX) and energy. The DIO messages are periodically multi-casted by nodes for topology maintenance. Periodic feedback is also uni-casted by child nodes to their respective parent node using destination advertisement object (DAO) to maintain point-to-multipoint and point-to-point connectivity.[2]

RPL provide single path routing and does not support multi-path routing. Single-path routing is achieved with minimum computational complexity and resource utilization but it reduces the achievable output of the network.[6, 7] In high traffic load when an active path is congested and fails to transmit the data then finding multiple paths for data forwarding can increase throughput. Multipath routing strives to find several paths from a source node to destination. Multipath routing improves reliability, provides fault-tolerant routing and reduces congestion. Existing literature[8, 9] has identified the need for supporting multipath routing in RPL. Despite few research efforts in this domain, providing multipath routing using RPL to support high data rate is still an open research area.

In this paper, multi-path RPL (M-RPL) is proposed that aims to provide temporary multiple paths during congestion over a path. In MRPL, data forwarding nodes on the routing path detect congestion using packet delivery ratio (PDR) and if PDR decreases below a certain threshold the predecessors of congested node forwards packets on multiple paths to the sink. The remaining part of this paper is organized as follows; related works are discussed in Section 2. M-RPL protocol is presented in Section 3. Simulation analysis and protocol performance is analyzed in Section 4. The last section concludes this paper along with future directions.

2. RELATED WORKS

RPL is relatively a new addition proposed for LLNs and in existing literature performance enhancement of RPL using multipath routing techniques is not been investigated in detail. Considerable research efforts[10, 11] are carried out on non-RPL based reactive routing protocols especially for WSNs. However, in this section we survey only the existing works that are related to RPL.

LB-RPL[9] is a load balanced extension of RPL protocol that addresses the issue of packet loss due to buffer limitations. It aims to achieve balanced workload distribution in the network. When a node detects that the network load is increasing and the buffer size has exceeded a threshold then it holds its DIO messages. DIO holding means that is delays the transmission of its DIO message. Greater the level of congestion higher will be the hold time. When the child of the congested node does not receive a DIO message from its parent then it assumes a link failure. Therefore, it switches to next available parent from the parent list. As a result, the routing path changes and load over the parent node decreases. However, mitigation of congestion using DIO holding is slow to react as DIO intervals are usually long (multiple minutes). Also, by holding DIO the child nodes do not immediately switch their parent.

In Ref. [12], DAG-based multipath routing protocol for mobile sinks is proposed. This protocol allows a node to find alternative paths in case of link failure. The alternative link or secondary link is found using reactive on-demand approach. When a link failure occurs node requests another sink to build a new DODAG and establishes new path with that sink. Main purpose of this protocol[12] is to establish an alternative path or redundant path with multiple sinks, so that the alternate path can be used in case the optimal path is unavailable.

Multipath opportunistic RPL routing protocol[8] is designed to work over IEEE 802.15.4. The objective of this protocol is to provide QoS using multiple paths. It increases packet delivery ratio to minimize overhead and energy consumption. Multipath opportunistic RPL[8] uses a cross layer design approach in which RPL interacts with the underlying link layer. It uses link quality estimations and IEEE 802.15.4 is tailored to provide such information, so that better links are used in DAG formation. A node only forwards higher-priority delay sensitive data through multiple paths to increase packet delivery ratio.

3. NETWORK MODEL

In this section basic network assumption, definition and network characteristic are discussed. M-RPL is suitable for both random and grid topologies. Also, any layered underlying medium access control (MAC) protocol can be used with M-RPL. Nodes in M-RPL act as a router as well as data originator. RPL uses a single path between sources to destination for data forwarding but M-RPL modifies the existing RPL to achieve multipath routing. It is assumed that multiple paths exist for data forwarding within the network. This is a common assumption as low power device networks such as WSN have high node density. This allows WSNs to have fault-tolerance, higher connectivity and better network coverage. Terminologies of parent and child nodes are frequently used in a DODAG established by RPL. Considering Figure 1 node 10 is the parent node for node 7, 11 and 21 whereas it is the child node for node 6. The parent child relation is subject to network conditions depends upon objective function, link quality and mobility of nodes within the network.

The basic aim of M-RPL algorithm is to provide temporary multipath routing during congestion over a path. Congestion occurs when the arrival rate of packets exceeds the forwarding rate of packets at a node. This results in packet drops at the congested node. Congestion can easily occur in event based network, where a sudden impulse of event information increases the traffic load over certain nodes within the network. As shown in Figure 1, the converge cast nature of upstream traffic coming from multiple

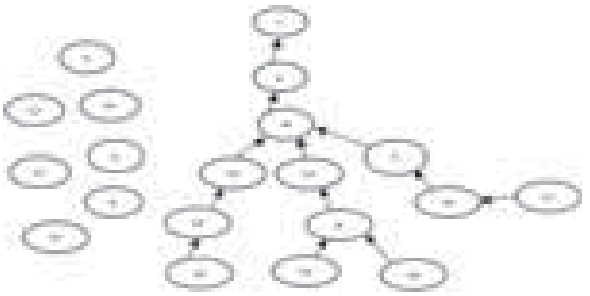


Fig. 1. Multi-hop converge cast data forwarding in sensor networks.

sources to a single destination results can result in congestion over certain nodes e.g., node 10. Congestion adversely effects the operation of the routing protocol and the overall performance of the network decreases.

RPL is a proactive routing protocol. RPL is a network layer protocol that obeys the layered principle of protocol design. It provides services only to the layer above and uses the services of lower layer. M-RPL algorithm does not change the nature of RPL and operates on the network layer only. During DODAG formation every node strives to connect with the DODAG through the best parent node. Therefore for sending data to the sink every node has a path to the sink through its parent node. M-RPL uses DIO control message to send congestion notification (CN) to their child nodes. In RPL DIO messages are multi-casted after periodic intervals using trickle timer algorithm. Trickle timer expires quickly when the root node starts DAG construction and it gradually increases as the topology stabilizes. The trickle algorithm allows nodes in a lossy and shared medium to exchange information in a highly robust, energy efficient, simple, and scalable manner.[13]

The metric used for congestion detection in M-RPL is PDR. It is the ratio of total number of packets received by the total number of packets sent at the destination. PDR can be calculated both at the parent node and at the child node. In the previous case, parent requires information related to numbers of packet sent by individual child nodes to estimate the PDR. But in the latter case, child node requires information regarding the number of packets received by the parent to estimate its PDR. M-RPL uses PDR at the parent node for congestion detection.

4. MULTIPATH RPL (M-RPL)

In this section, the operation of M-RPL is explained in length. Operation of M-RPL is divided into two main parts, congestion detection and congestion mitigation. In M-RPL congestion is detected on any forwarding node whereas mitigation of congestion is performed by introducing multipath routing at nodes prior to the congested node.

4.1. Congestion Detection

Algorithm of congestion detection proposed by M-RPL is shown in Figure 2. The congestion detection algorithm is triggered on the reception of any incoming packet at a node. PDR is calculated by all the nodes between source and sink that are involved in the routing of information. In order to calculate the PDR a node must have the information of expected data rate of each child node. In M-RPL a child node uses DAO messages for communicating its current forwarding rate to the parent node. M-RPL uses average PDR at a node during a certain congestion interval (CI) to detect congestion. The length of CI depicts the decision time for the detection of congestion. Very small CI means frequent congestion decisions which is not be appropriate when number of traffic flows passing through a node are rapidly changing. Large value of CI means more delay in congestion decisions that can result in increased packet drops.

If the PDR exceeds a certain congestion threshold then a notification message is sent to the child node. Congestion notifications (CN) are sent to the child node by its parent node using periodic DIO message. Hence the overhead of extra transmission of packet is avoided by M-RPL. Numbers of reserved bits are available in the DIO packet header and one of these bits carries the CN. The length of DIO intervals is based on trickle timer that increases as the network topology stabilizes. During stable and idle network conditions DIO interval can be as large several minutes whereas during changes in network topology the DIO interval can be as low as few seconds. The length of DIO and CI are not aligned with each other. If CN is to be transmitted after the expiry of CI then the node checks the status of its trickle timer. DIO messages are generated after the expiry of trickle timer. If trickle timer is about to expire that is the remaining time is less than half CI then the DIO will carry the CN. Otherwise, an emergency DIO (EDIO) is transmitted immediately to notify congestion. In case of M-RPL, EDIO is the only extra message incorporated in

Algorithm to detect congestion at any node using PDR

1. **FOR EACH** packet received **DO**
2. **IF** pkt_dest_address **NOT EQUAL TO** current_node
3. Pkt_counter++
4. **END IF**
5. **IF** CI Expired **THEN**
6. PDR = $C_{\text{Pkt\_counter}}/C_{\text{expected\_packets}}$
7. $C_{\text{Pkt\_counter}} = 0$; (reset variables)
8. **IF** PDR < Cong_TH **THEN**
9. Send PDR to child nodes
10. **END IF**
11. **END IF**
12. **END**

Pkt_counter: to count the pkts received
Cong_TH: congestion threshold

Fig. 2. Congestion detection algorithm based on PDR.

RPL to notify congestion immediately when normal periodic DIO message is not available at that time.

4.2. Congestion Mitigation

Congestion mitigation is triggered once a child node receives a DIO message containing congestion notification. The child node receiving this message starts multipath routing by splitting their forwarding rate into half. The child node forwards one packet to its original parent who is congested and forwards the next packet to any other node from the parent list maintained in parent table (PT). Therefore, during congestion mitigation a node drops its forwarding rate to the congested node to half. The rest of the data is forwarded through alternate path using any other parent node. Thus the child nodes strive to reduce forwarding packets to the congested node and this helps in the reduction of the congestion at their parent node.

It is possible that a child node can select another congested neighboring node for temporary forwarding of data when its original parent has sent CN. In order to avoid forwarding packets to congested neighboring nodes, each node also saves the CNs of their neighbors in their PT. This does not cost extra communicational or computational overhead as nodes in RPL listen to their neighbors DIO for topology maintenance.

Congestion mitigation by splitting of information over two routing paths is performed on the immediate child of the congested node. But this can work effectively, if the child's neighbors are not congested because in wireless medium access all nodes are sharing the same medium. As a result in M-RPL, the CN can be further multi-casted towards the source nodes. If the status of 50% of the neighbor nodes in the PT nodes is congested then a node receiving CN from its parent will rely this message further towards source node. Congestion notification is communicated by the congested node after every CI. If a child node does not hear CN from its parent node then its stops splitting of the forwarding traffic.

5. PERFORMANCE EVALUATION

In this section the performance of M-RPL and RPL is analyzed in a random topology. We have simulated M-RPL and RPL in Cooja[14] simulator in Contiki operating system.[14] The sensor nodes are randomly deployed in 100×100 m. Figure 3 shows the simulation topology, the sink node 1 is placed on the corner of the network. The circle at the bottom of Figure 3 shows the event reporting nodes that are at multi-hop distance from the sink. Background traffic is generated from different nodes that are uniformly within the sensor field. The details of network parameters used in the simulation analysis are illustrated in Table I. In Ref. [9] very low data rate of two packets per 2.5 to 5 minutes is used but is not suitable for high data rate applications.

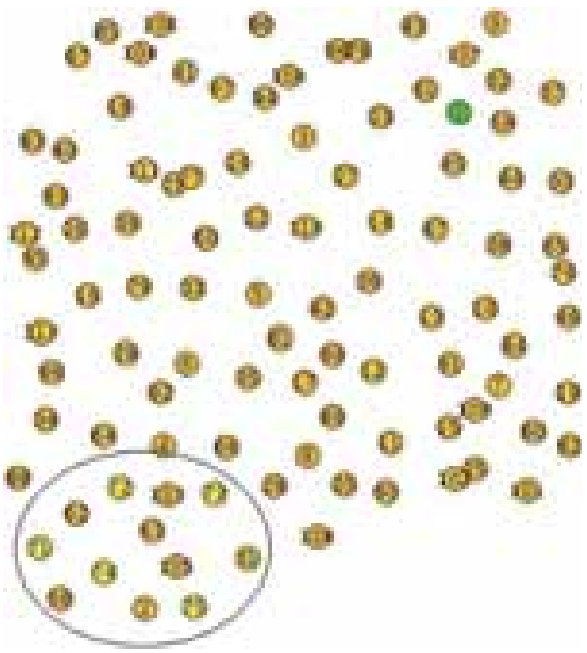


Fig. 3. Sensor network topology used for simulation analysis.

Therefore, we have high data rates ranging from 2 packets per second to 1 packet per two seconds. The performance of RPL and M-RPL are evaluated against throughput, end-to-end latency and energy consumption. All simulations are run for 10 times and averaged results are presented.

The throughput observed at the sinkis shown in Figure 4. Five source nodes are generating one packet per second from the event region for 500 seconds. Packets are dropped and congestion occurs as traffics converge on certain nodes near the sink. Throughput of RPL is persistently lower than M-RPL because the single paths used by

Table I. Network parameter used in simulation analysis.

Parameter	Value
Network layer	RPL/M-RPL
MAC layer	802.15.4
Topology	Random
Simulation time	600 s
Event reporting	500 s
Objective Function (OF)	RPL-mrhof
DIO minimum interval	4 sec
DIO maximum interval	17.5 min
Data intervals (single packet per interval)	0.5, 1, 1.5, 2 sec
Number of source nodes	2–12
TX range	20 m
Interference range	35 m
Packet size (excluding header)	50 Bytes
Tx power	0.0174 W
Rx power	0.0188 W

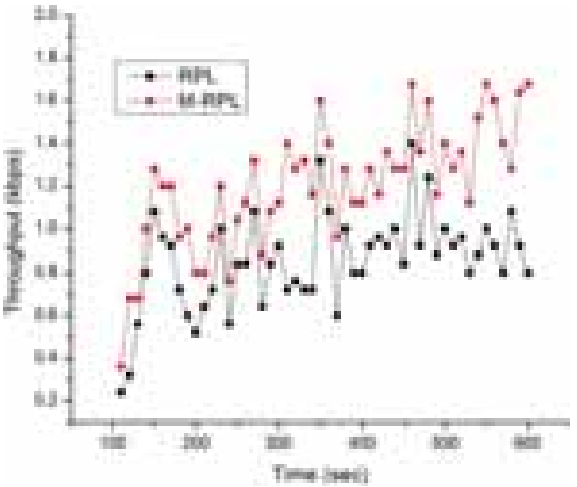


Fig. 4. Average throughput of M-RPL and RPL.

RPL get congested. On the other hand, M-RPL splits traffic near the congested node over different paths resulting in decrease in congestion and increase in throughput.

Average throughput observed at sink node for M-RPL and RPL under different packet intervals of 0.5, 1, 1.5 and 2 seconds is shown in Figure 5. Lower packet interval means higher data rate whereas higher packet generation interval correspond to lower data rates. Throughput of M-RPL is significantly better than RPL in Figure 5 at all packet generation interval because of multi-path routing. It is evident from Figure 5 that as the data rate is decreased (1 pkts per two sec) the performance of RPL gets better because congestion is not severe.

Average throughput using different source nodes is shown in Figure 6. Throughput of both the protocols increases with the increase in number of source nodes

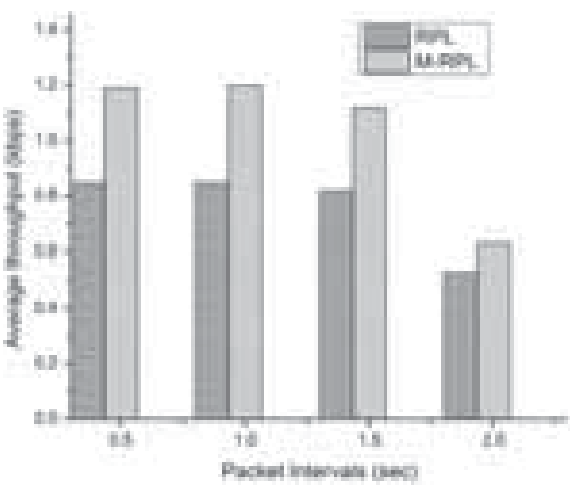


Fig. 5. Average throughput of RPL and M-RPL under different packet generation intervals.

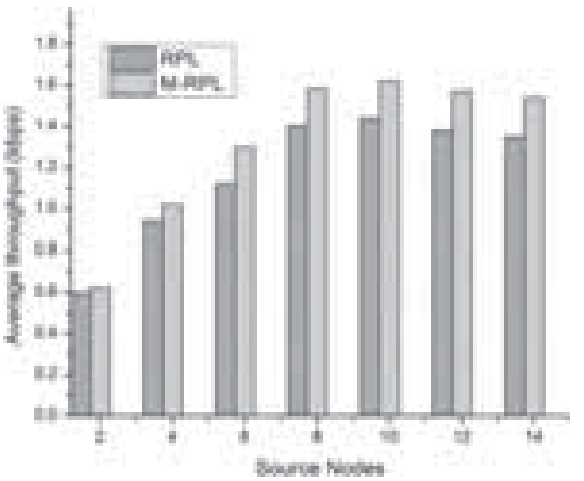


Fig. 6. Average throughput observed at sink using different number of source nodes.

but M-RPL out performs RPL in all scenarios shown in Figure 6. When the number of source nodes are 12 and 14 the throughput of M-RPL is also affected and it tends to decrease instead of increasing. This is because the network gets congested and this type of congestion is persistent. It is difficult to remove congestion even through multipath routing as alternate paths are either not available or are congested.

The average end-to-end delay of RPL and M-RPL is shown in Figure 7. The latency of both the protocol is high because the devices have very low duty cycle therefore nodes spent most of their time in sleep mode compared to active mode. It is noticeable that the delay of M-RPL is similar to RPL initially.

This is because M-RPL initially establishes a single path for data routing as defined by RPL. However, after

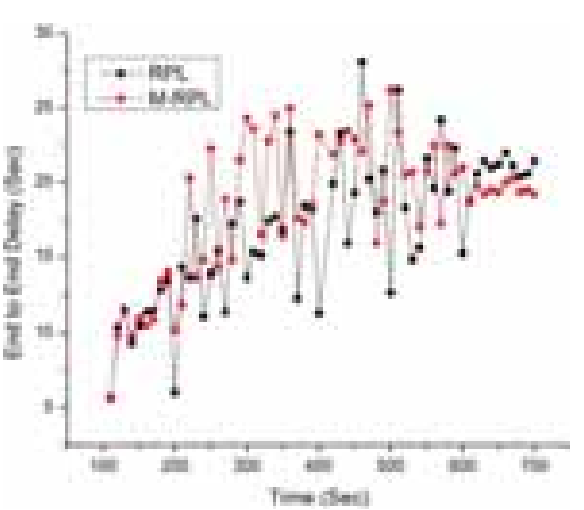


Fig. 7. Average end-to-end delay of M-RPL and RPL.

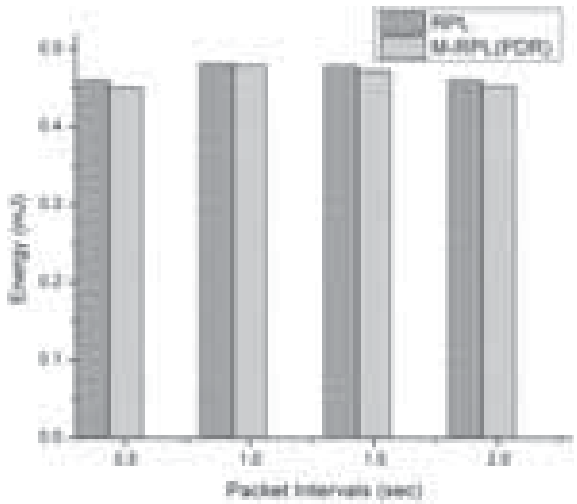


Fig. 8. Average per-bit energy consumption of RPL and M-RPL at different packet generation intervals.

the detection of congestion at any congested node data splitting start resulting in the creation of multiple routing paths. At the start of data splitting the latency of M-RPL is slightly greater than RPL because of the introduction of multiple paths. It is also noticeable that the delay of RPL decreases and becomes similar or less than M-RPL. This is because data is continuously reported and congestion becomes persistent that forces RPL nodes to change their parents. This result in change of network topology and results in increased delay. Thus as data is continuously reported the end-to-end delay of M-RPL stabilizes and becomes less than RPL.

Average per-bit energy consumption of M-RPL and RPL at different packet generation intervals is shown in Figure 8. Per-bit energy is calculated by dividing total energy expenditure by total number of bits transmitted/received within simulation time. RPL has higher packet drop ratio resulting in wastage of energy on unsuccessful transmissions. Despite the fact that M-RPL creates more than one path by splitting of traffic flows as congested node it has lower per bit energy consumption than RPL. M-RPL has higher packet delivery ratio and less packets are dropped resulting in decrease of energy wastage.

6. CONCLUSION

RPL is the standard protocol recommended by IETF for IPv6 based LLNs. The current standard of RPL supports single path routing. In this paper, multipath extension of RPL is proposed termed M-RPL. The proposed protocol using PDR detects congestion on a node in the routing path and starts multipath routing at the predecessor of congested node. Multiple paths are created by splitting forwarding rate on both preferred parent (congested node) and alternate parent available in RPL. Detailed simulation analysis shows that M-RPL is capable of supporting higher data rates as compared to RPL. It is also noticeable that M-RPL extends RPL and does require significant changes in the original protocol.

References

1. N. Accettura, L. Grieco, G. Boggia, and P. Camarda, Performance analysis of the RPL routing protocol, *IEEE International Conference on Mechatronics (ICM)*, Turkey, April **(2011)**, pp. 767–772.
2. T. Winter and P. Thubert, RFC 6550: RPL: IPv6 routing protocol for low-power and lossy networks, Internet Engineering Task Force (IETF) Request For Comments, March **(2010)**.
3. J. P. Vasseur, M. Kim, K. Pister, N. Dejean, and D. Barthel, *Internet Engineering Task Force (IETF) Request for Comments* **(2012)**.
4. P. Thubert, *Internet Engineering Task Force (IETF) Request for Comments* **(2012)**.
5. O. Gnawali and P. Levis, *Internet Engineering Task Force (IETF) Request for Comments* **(2012)**.
6. D. Son, B. Krishnamachari, and J. Heidemann, Experimental study of concurrent transmission in wireless sensor networks. *Proc. of the 4th International Conference on Embedded Networked Sensor Systems*, USA, October **(200)**, pp.237–250.
7. J. Kang, Y. Zhang, and B. Nath, End-to-end channel capacity measurement for congestion control in sensor networks, *Proc. of the 2nd International Workshop on Sensor and Actor Network Protocols and Applications* (*SANPA*), USA, August **(2004)**.
8. B. Pavkovi, F. Theoleyre, and A. Duda, Multipath opportunistic RPL routing over IEEE 802.15.4, *Proc. of the 14th ACM International Conference on Modeling, Analysis and Simulation*, USA, July **(2011)**, pp. 179–186.
9. X. Liu, J. Guo, G. Bhatti, P. Orlik, and K. Parsons, Load balanced routing for low power and lossy networks, *Wireless Communications and Networking Conference (WCNC)*, China, April **(2013)**, pp. 2238–2243.
10. M. Radi, B. Dezfouli, K. Bakar, and M. Lee, *Sensors Journal* 12, 650 **(2012)**.
11. M. Masdari and M. Tanabi, *International Journal of Future Generation Communication and Networking* 6, 181 **(2013)**.
12. G. Xu and G. Lu, Multipath routing protocol for DAG-based WSNs with mobile sinks, *Proc. of the 2nd International Conference on Computer Science and Electronics Engineering* (*ICCSEE*), China, March **(2013)**.
13. P. Levis, T. Clausen, J. Hui, O. Gnawali, and J. Ko, *Internet Engineering Task Force* (*IETF*) *Request for Comments* **(2011)**.
14. Contiki: The Open Source OS for the Internet of Things, Contiki OS. [Online], Available: http://www.contiki-os.org/index.html [Accessed: March-2014].
15. T. Ming-Hao, Y. Ren-Lai, L. Shu-Jiang, and W. Xiang-Dong, Multipath routing protocol with load balancing in WSN considering interference, *Proc. of IEEE Industrial Electronics and Applications Conference (ICIEA)*, Singapore, June **(2011)**, pp. 1062–1067.

Received: 16 April 2018. Accepted: 14 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1540–1544, 2018

RESEARCH ARTICLE

Machine Vision Based Congregation Management System Using myRIO

R. Prashanna Rangan*, K. Sethuramalingam, and P. Sairam Santhosh

UG Scholar, Department of Mechatronics Engineering, Kongu Engineering College, Erode 638060, India

In the present times, taking of attendance is mandatory not only in educational institutions, but also in MNC firms so as to ensure the current state of the respective persons. In many companies, a person's salary is provided based on his working hours and hence, it is an inevitable work that has to be taken out to check the individual's current status. With the advancement of technology, human life has become more sophisticated. This paper aims in developing a smart sensor based approach for taking the attendance. This project utilizes a CMOS camera, by which the face of every individual is recognized in comparison to the database and attendance is taken in seconds. This work can be very useful in taking the attendance more precisely and at a rate much faster than the other. The accuracy is quite high when compared with other technologies, as it recognizes the whole face of the individual and hence there is no possibility for making a false attendance.

Keywords: LabVIEW, Facial Recognition, Template Matching, myRIO, Automated Attendance, Digital Image Processing, Web Camera.

1. INTRODUCTION

1.1. Motivation for the Project

In the recent years, taking of attendance has become mandatory not only in educational institutions, but also in companies, so as to ensure whether the respective persons are doing their duty properly. As it is a vital aspect, by which a person gets his salary in a firm, it is important to take proper attendance. The conventional method of taking the attendance consists of a note book, where there is a list of all the members belonging to that specific area. A person takes the responsibility to readout all the names and marks with respect to the presence of the individual as present or absent. But, practically it is impossible to check for each and every persons, when there is a huge stack of crowd and there is also a possibility of false attendance. Hence it is mandatory to rev up the system with the modern equipments.

1.2. Existing Technology

The existing technologies either used a RFID (Radio Frequency Identity) tag or they used a finger print system, Retinal Scan or Voice recognition systems to detect the persons and mark the attendance. They has its own advantages and disadvantages.

Reference [1] explains the usage of RFID based attendance management system, which makes use of a RFID reader and RFID based ID cards, distributed to all students. Similar to the above said system, Ref. [2] also consists of RFID readers and tags, in addition to it, setup of an Infrared based motion sensor to take up the counts of the students who entered the class and it compares the value with the RFID read values to get the actual count. It sends the alert message to the lecturer as well as to the management of the institution. The prototype, as explained by Ref. [3] states the usage of same RFID based attendance system, which further utilizes the usage of a JAVA application at the backend to manage the attendance.

System[4] made use of the Digital Image Processing technique to identify the faces of the individual and marked attendance based on the time they sat inside. It made use of algorithms to search and compare each and every faces inside the room. It showed a 78% accuracy rate. System[5] is similar to the system[4] in many aspects. In the work done in Ref. [6] overcomes the disadvantage of Ref. [4] by taking a short video clip of the classroom continuously and marking the attendance. It uses various algorithms to improve the accuracy rate.

1.3. Demerits of Existing Technologies

Biometric systems ensures that none other than the person can mark their attendance, but it lacked when the persons gets himself wounded on the respected finger. In the same way when the person forgets to take his RFID tag or if it is lost the person won't be able to get his/her attendance.

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1540/005 doi:10.1166/jctn.2018.7336

The biometric and RFID detectors can also occasionally fails which also leads to mismatch of attendance. The RFID tags also has another disadvantage that, there is a possibility for misleading of attendance by taking someone's ID's. Hence many of the above said technologies proved to be not sufficient to work efficiently under these circumstances.

Since the methods[1–3] makes use of RFID tags to identify the individuals, it proves to be insufficient and time consuming process to mark attendance. Infrared based motion sensor has also a possibility of making a double count or leave a count, when two or more persons enters simultaneously.

Though the system[4] proved to be useful, it has many complex codings and used complex algorithms to match the faces. It made use of a camera, only inside the class room. The camera captures the images of all present inside the lecture room and marks the attendance. But this technology is not suited for all the times, say for an example if a person leans down, he will eventually be marked absent. System[5] is similar to the system[4] in many aspects. It made used of the same algorithms to train the faces and compare them with the real time acquired image. System[6] creates a database, in which the attendance data will be available to all the authorities, but is non-changeable. The created database is also a size-consuming one. The database is tough to create, as it has many complex codings and algorithms. Hence if any problem occurs, it is not very easy to find and solve the errors.

1.4. Proposed System

In common, all the above said technologies used a textual-based programming language and utilized complex algorithms to detect the faces and mark the attendance. Hence to overcome these kind of limitations, this paper proposes a technology, which utilizes graphical way of programming, which eliminates the usage of algorithms.

In the proposed method, the attendance is taken with the aid of a pair of cameras. As soon as the person's face is recognized by the camera, the image is compared with the database and if it matches he/she gets a present in their attendance for the hour. This technology is very useful to avoid fake attendance and its accuracy is quite high compared to other technologies.

The proposed method finds to be useful not only to Educational Institutions, but also to the Industries and Software Firms, who takes attendance regularly to have an eye on their employees.

2. COMPONENTS USED AND ITS WORKING

2.1. LabVIEW Workbench

LabVIEW is a platform, which has the Graphical User Interface. The programs made here are Graphical oriented and there is no text-based programming. It eliminates the usage of the tough and complex algorithms. It consists of a front panel and a block diagram, as shown in Figure 1. The front panel resembles the industrial panel cluster and hence the name. Block diagram is the area in which the logical programming is made by hot-wiring method. There are many hard-wares that can be linked with LabVIEW and it provides a more easy solution to complex problems.

Fig. 1. LabVIEW front panel and block diagram.

2.2. NI myRIO

myRIO, as shown in Figure 2 is a stand-alone hardware developed by National Instruments.

The myRIO-1900 is a tool which can be used to implement multiple design concepts with one reconfigurable I/O (RIO) device, as shown in Figures 3 and 4.

Featuring I/O on both sides of the device in the form of MXP and MSP connectors, it includes 10 analog inputs, six analog outputs, 40 digital I/O lines, WiFi, LEDs, a push button, an onboard accelerometer, a Xilinx FPGA, and a dual-core ARM Cortex-A9 processor. It can be programmed by either LabVIEW or C. This onboard WiFi enabled device helps in developing remote embedded applications.

2.3. USB Web Cam

A USB web cam is as shown in the Figure 5. The USB web cam is used as an eye to monitor the class room. The True-Color technology of this webcam controls the amount of exposure for bright, colorful pictures. With 360-degree

Fig. 2. NI myRIO 1900.

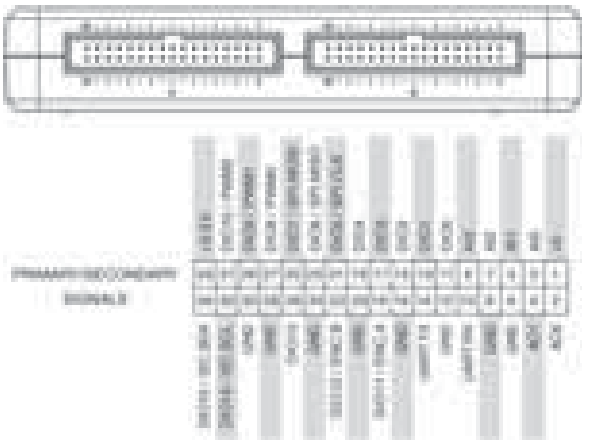


Fig. 3. NI myRIO MXP connections.

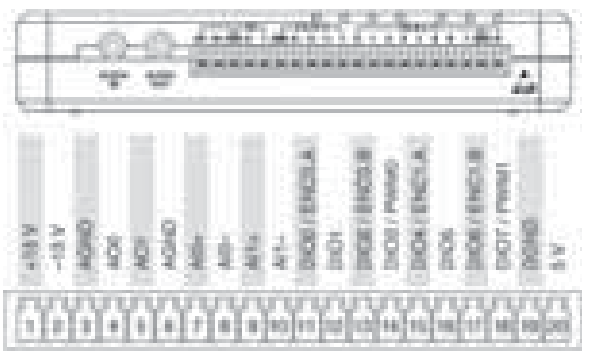


Fig. 4. NI myRIO MSP connections.

Fig. 5. USB web cam.

lens rotation, this webcam provides all-around view for complete coverage.

3. WORKING

In these recent years, taking of attendance has become mandatory due to lack of safety aspects and increased malpractices in the company has become a major threat and is the reason behind the arise of many problems. Hence it is mandatory to check the current status of each and every individual that work for the specific firm/studies in the specific institution and to ensure their presence or absence. The software is first trained by using the NI Vision assist, as shown in Figure 6. The images are logged to a specific location in the local disk of the computer so that in case of a security breach, the intruder's face can be acquired with ease. It is as shown in the Figure 7. Template matching algorithm is as shown in Figure 8. The simulation is made first using the webcam connected to the LabVIEW PC and the results are acquired as shown in the Figure 9. The camera is controlled with the help of an NI myRIO-1900 controller, as illustrated in Figure 10. As soon as the camera finds a person on the floor it captures his/her image and compares it with the pre stored database images for that specific person. In this project Digital Image Processing plays a vital role in scanning each and every person, who may cross the floor with the help of a CCTV camera. The reason digital image processing has been used is that it has the ability to categorize the objects.

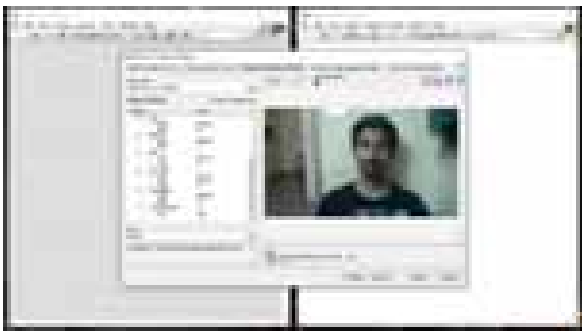


Fig. 6. Training the software.

A camera is placed at the entrance of the room. When a person enters the room, his face is captured and gets compared with the database. If the image suits with the database, his/her current status will be marked accordingly and if it doesn't suits with the pre stored data, the captured image is forwarded to the security head stating that there's a security breach and some strangers have entered the firm/institution.

Fig. 7. Logging the images.

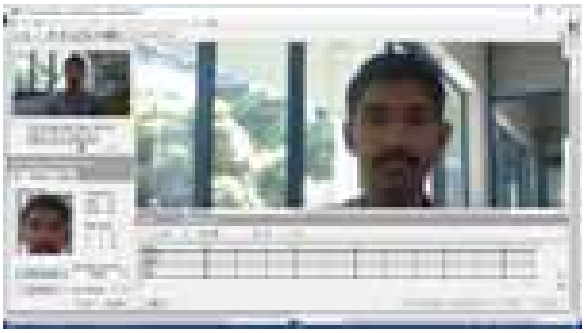


Fig. 8. Template matching.

Fig. 9. Working simulation.

Fig. 10. Working prototype.

Another camera is placed inside the room, which monitors each and every individual's state and if anyone goes out of the room it changes that particular person's current status for that hour, after a time interval of 10 minutes.

This process provides a two-way checking system for the persons and ensures safety by intimating the security officials whenever an unknown person is found.

4. FEASIBILITY STUDY

The feasibility study is the study of factors influencing the success of the product, for the purpose of resource determination. Providing the solution for the project is more possible in an economic way. As it involves only one time investment for the components and it costs around Rs. 80,000. Hence it is economically feasible for any firm for these applications. The system is also operationally feasible, as it has the ease of user interface. The system uses simple components which are readily available in the market and it is easy to work with them. Hence the system also finds to be technically feasible.

5. MERITS AND DEMERITS IN IMPLEMENTING THE TECHNOLOGY

5.1. Merits

- Reliable and cost efficient.
- Checking of persons in a comparably easier and safer way.
- The automated attendance counter technology provides easy way to check for persons and counts efficiently in areas where there is a dense population.
- High levels of accuracy regarding the attendance of a person is ensured.
- Promotes greener solution and is a part of smart city concept.
- Reduces probability of damage to the instruments.

5.2. Demerits

- Accuracy in finding the faces of persons in high dust environment depends mainly on light visibility available in the room.
- The camera may occasionally fail and hence faulty signals could be sent to the microcontroller board.

6. FUTURE SCOPE

The same setup with a little further modification can be used for industries. The template matching algorithm proved to be useful for object tracking and comparison in various manufacturing industries. Especially in production areas, which requires very high accuracy and ease of usage can implement this technology.

7. CONCLUSION

It will be a great achievement when a camera automatically takes attendance with high degree of reliability and accuracy. If the automated attendance counter technology exists in future, then there might not be a chance for false attendance to persons and the time required will be extremely reduced.

If it does not exist, then the automated systems will become a daydream and go in a path, which will be of many problems, including high installation and maintenance costs for any other technology similar to this. Moreover, there will be no assurance for cent percent accuracy. The main motive of this project is to detect and recognize the person and to provide attendance to the persons at a rate faster than any other.

References

1. Ankita Agarwal and Ashish Bansal, *International Journal of Information and Computation Technology* 3, 131 **(2013)**.
2. G. Chandhiny and T. Aravinth, *International Journal of Scientific and Engineering Research* 6, 23 **(2015)**.
3. Elima Hussain, Priyanka Dugar, Vaskar Deka, and Abdul Hannan, *International Journal of Computer Applications* 30 **(2014)**.
4. E. T. Shin, W. J. Chew, and F. Choong, *Journal of Engineering Science and Technology* 45 **(2015)**.
5. G. R. Pooja, M. Poornima, S. Palakshi, Bhanu Prakash Varma, and A. N. Krishna, *International Journal of Avanced Networking and Applications* 363 **(2016)**.
6. Aziza Ahmedi and Suvarna Nandyal, *The International Journal of Engineering and Science* 4, 1 **(2015)**.

Received: 16 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1545–1550, 2018

RESEARCH ARTICLE

QR Code Based Shopping with Secure Checkout for Smartphones

A. Banu[1,*], K. Ganagavalli[1], and G. Ramsundar[2]

[1] *Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam 638401, India*
[2] *Department of Computer Science and Engineering, KPR Institute of Engineering and Technology, Arasur, Coimbatore 641407, India*

As the number of mobile devices continuous to grow, more and more online and offline retailers are incorporating mobile commerce into their business strategy. More retailers are bringing their online shop to offline world using QR codes. QR codes is a cell phone reliable bar code that can store website URL, phone number, email address and much more. A virtual store shows products on some printed materials, on PDA's or Digital display, or on product itself. As the movement of using QR codes in shopping is still in its infancy. Android is a growing technology which have started to fulfill need with lots of application to make things handy. QR Codes can connect the users to the information quickly and easily. In this paper, we explore how QR codes can be used in shopping. The low technical barrier of creating and reading QR codes allows innovative shop owners to incorporate them into their shopping endeavors. The operations to retrieve or store QR codes are incredibly simple and quick, and with mobile devices. The consumer of our application can able to purchase products through the QR Code and stored that product in the cart. To increase the security in online payment, two levels authentication is provided during the transaction process using PAKE protocol at level one and OTP at level two. For wide accessibility, the whole system is developed using open source and freely-available software.

Keywords: QR Code, Shopping, Smartphones, E-Commerce Transaction Security, ElGamal Encryption.

1. INTRODUCTION

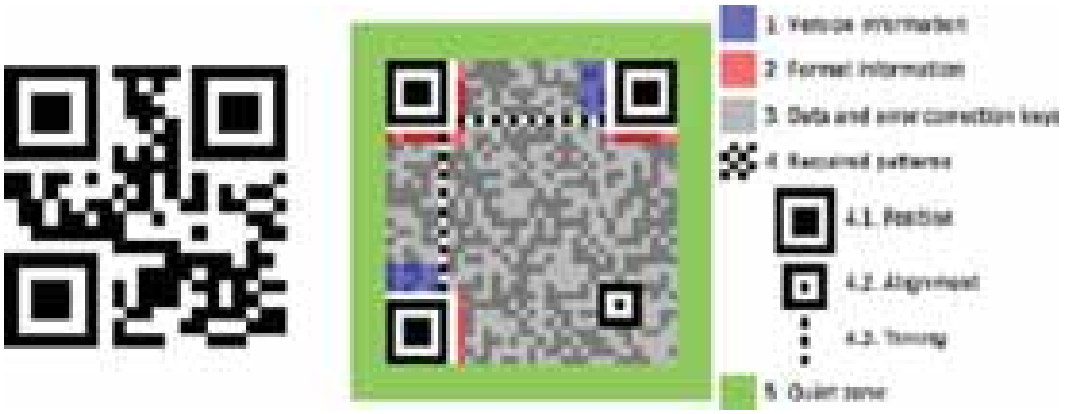
QR codes are square images composed of modules. A module is the minimum square element and the basic information unit. The number of modules depends on the specific version of the QR code used and is directly related to the amount of information stored. It goes from 21×21 modules (version 1) to 177×177 modules (version 40). QR codes display several patterns that allow for fast identification and decoding: finder patterns, timing patterns, and alignment patterns. These patterns and the general representation of QR code are illustrated in Figure 1. The reader is referred to Ref. [1] for detailed information. The data cell, which is shaded in gray, is the area code that stores the actual data.

With the improvement of living standards, most big shopping malls experienced reconstruction and scale expansion. Currently, shopping malls are bigger with more abundant goods and more variety of wares. People are in the pursuit of high quality consumer goods. At the same time, the pursuit of efficient shopping is gradually revealed. Various factors like increase in population, long working hours, wide variety of products, different geographical locations and has made shopping an unpleasant and time consuming task. To solve the difficulty in customer shopping, QR code based shopping system is proposed.

QR code based shopping is the electronic version of shopping. The aim of this project is to do online retail using a mobile phone and QR code. With the improvement of living standards, shops are bigger after constructing with more abundant goods and more variety of wares. Therefore, building a simple, fast and convenient shopping application system has become a mutual concern of merchants and customers. In view of the mobile phone has become a popular consumer products, a simple optimization method was given to design shopping system run on smart phones, with the help of QR code generation and recognition technology.

With the guidance of QR codes the user can easily get the product details and add that product in the cart. The admin of our application will create the QR code based on the product details and allow the users for scanning. During the online purchase of the carted product, a secure

*Author to whom correspondence should be addressed.



(a) (b)

Fig. 1. (a) A typical QR code (b) structure of a QR code.

way to check out the purchase is also designed using two step authentication. After the verification of with two step authentication, the user can able to complete the purchase.

The rest of the paper is structured as follows. Section 2 describes some related work how the QR code are used in android applications, Section 3 shows the system overview and implement work, in Section 4 security analysis of the work is done during checkout process and conclusion are drawn in the last session.

2. RELATED WORK

RedLaser[4] is an application with the option to scan all major retail barcodes (including UPC, EAN, UPC-E, and EAN-8) and QR codes. It allows the user to search for prices from thousands of online retailers for millions of products and compare the prices, find the nearest locations, and even check out books at the library. It provides the product descriptions, reviews, and nutritional information to the user and finds deals and coupons from many merchants. Popular products scanned by other RedLaser users can also be viewed. Also allows the user to create personalized QR codes with user's info, contacts, website's URL, location and more with the powerful QR code engine. Keyword, Voice, and Image recognition search option is provided to find the best deal, even when user doesn't have a barcode handy.

ShopSavvy[5] helps to find the best online and local prices, reliable product info, and provides a way to shop anywhere the user want. Allows the user to Scan or Search to find the lowest prices guaranteed and also set price alerts. It provides the user to create and share lists of favorite products with user's favorite people. Users can scan a QR code to obtain URLs, contact information, calendar events and other data stored within the Quick Response codes. It also allows individuals as well as businesses to generate their own QR codes for display on signage, business cards, advertisements and other various vehicles for generating awareness. It provides Sharing, saving and printing of QR codes, On-screen display, history and Searching of scanned codes.

The Micello[6] in an application that allows mapping the indoor world with the same level of accuracy and precision as outdoor maps. Provides accurate maps that tell exactly where the user is in within a location. Allows the user to search and find what is present in the user's surroundings. Micello maps also include intricate navigation which offers routing from point to point within the indoor map. In addition, the network offers appropriate level to level navigation via the proper structure (such as escalators, stairs, etc). However the application utilizes NFC technology and QR Codes to facilitate the process of finding its current location.

PAKE[9] is method to establish a secret key between two communicating parties based upon their knowledge of a password. Established secret key can be used to secure exchange of messages such that an unauthorized party can obtain no information regarding the messages exchanged without the knowledge of the secret key. This means that if an attacker hacks the server data, still will not be possible to masquerade as the client unless they perform a brute force search for the password.

3. SYSTEM OVERVIEW AND IMPLEMENTATION

The process in shopping has been simplified through the proposed system. Customers scan the QR code of the object to be purchased and update their carts, once the products are chosen, online payment is made as shown in Figure 2. The good are delivered to the customer's door step. At the back end product information maintenance and security in online transaction is provided.

3.1. QR Code Design

Each product has to be identified uniquely and hence QR codes are made use to satisfy this purpose. Initially each

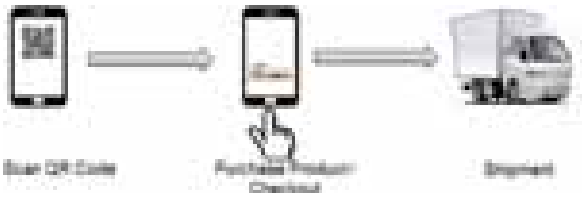


Fig. 2. QR code shopping overview.

product is assigned with its own QR code which is the responsibility of the Shopkeeper. QR codes when processed reveal the respective product id.

A sequence of steps are involved in the generation of QR code. Initially, the product id is given into the encoding engine and the unique QR code is generated. The various stages to be undergone by each of the product id to generate the QR code are Ref. [3] as follows Data Analysis, Data Encoding, Error Correction Coding, Structure Final Message, Module Placement in Matrix, Data Masking, Format and Version Information.

1. *Data Analysis* Is the first stage in generating the QR code where the mode in which the information to be encoded is determined. Numeric, alphanumeric, byte, or kanji are modes available for information encoding.
2. *Data Encoding* the encoding mode is chosen. The result of this step is a string of bits that is split up into data code words. Each of the data code are 8 bits long.
3. *Error Correction Coding* is used by QR codes. Once after the string of data bits are created which represents the PID, those bits are used to generate error correction code words using a process called Reed-Solomon error correction.
4. *Structure Final Message* Data and error correction codewords that are generated in the previous steps must now be arranged in structured order. For large QR codes, the data and error correction codewords are generated in blocks, and the blocks are interleaved based on the QR code specification.
5. *In, Module Placement in Matrix* the codewords are arranged in matrix in a specific way. The patterns that are common to all QR codes are placed as boxes at three corners.
6. *Data Masking* is used to break up patterns in the data area that might confuse a scanner, such as large blank areas or misleading features that look like the locator marks. The mask patterns are defined on a grid that is repeated as necessary to cover the whole symbol.

Fig. 3. QR code design.

(a) (b)

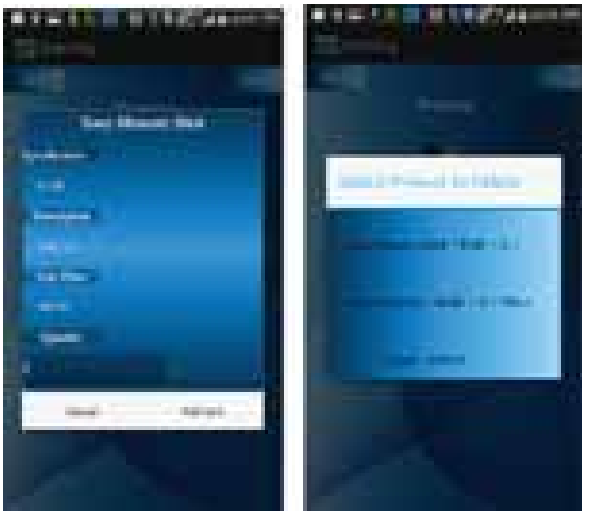
Fig. 4. (a) Home screen (b) new product update.

7. *Format and Version Information* is the final step in which the format and version information (if necessary) is added to the QR code. It is done by adding pixels in particular areas of the code that were left blank in previous steps.

In the application that is developed navigation through a few screens leading to the screen to facilitating the shopkeeper to create the QR code for the concerned new product that is to be launched. Initially the Home Screen is displayed on clicking the application which contains three major options namely Purchase, Product info and Account info. In order to proceed to the creation of the QR code, product info option is chosen. The login screen is launched on click of Product info in which only the admin/retail owner are provided the access. This screen checks the credentials of the retail owner. When the admin is successfully logged in, the Product info page is launched where the options to add new product info like type, description, price, avail quantity and its description are provided. On addition of new product, the product info gets updated in database and a unique QR code is generated for that product. Also the option to delete a product or update the information of already available product is provided, when done it is updated in the database.

3.2. Purchase Product

To purchase a particular product the QR code of the product is needed in order to obtain the product details. The QR code is scanned by using the user's device such as android smartphone. The recognized QR code is converted and compared with the cloud data. The appropriate product details that match with the converted QR code are retrieved. The product is added to cart for purchase based on the customer's interest.



(a) (b)

Fig. 5. (a) Product information (b) cart details.

On clicking of Purchase in home screen which provides the following options,

1. *Shopping*: When a product is to be purchased the QR code is needed and for that QR code are scanned for the specific product information.

ZXing (Zebra Crossing)[7] is an open-source, multi format 1D/2D barcode image processing library implemented in java. The purpose of the library is to use the build in camera on mobile phones to decode the barcodes on the device. The library is available on Varity pf platforms, among which Android. And for implementing the generation of QR code and for decoding information from QR code this ZXing package is used.

2. *Manage cart*: The series of products that are selected by the customer in order to purchase are maintained in the cart. The product information of the selected products in cart is also provided to the user.

3. *The checkout process* is the last step in completing the transaction and making a sale. For the payment the user must provide payment information like a transaction ID and password. For shipping of the product the both billing and shipping addresses and contact information are collected from the user. Once the order is confirmed then the user receives some form of order confirmation.

The next step is to fetch the information of the scanned product from the database. Initially, QR code of the product specifies the particular product uniquely. The scanned QR code is used to obtain the details of the product. Based on the QR code provided by the user the information of the appropriate product is retrieved from the database and provided to the user. When the user decides to proceed with the billing, he can view the products that are in his shopping cart by clicking the manage cart option. The series of products are listed to the user based on the user's interest

(a) (b)

Fig. 6. (a) First step authentication using PAKE protocol (b) Second step verification using OTP.

and the selected products are maintained in the cart. Also the details about each selected product is provided.

3.3. Checkout Process

In this module, we create an account for the user for making online purchasing. If the user wants to purchase a product, the user has to enter the transaction id and password. After that a One Time Password is send to the user mobile for verification. After this authentication process get completes, the user can able to purchase that product.

The security of the system is built upon cryptography algorithm namely ElGamal encryption Scheme. Which consists of three phase's namely key generation, encryption and decryption as follows.

1. *Key generation*: A cyclic group of G of large prime order q with a generator is chosen. Then the decryption key x ia randomly chosen from Z_q^* and the encryption key y is calculated as $y = g^x$. G, g and y are published as the public parameters of the encryption scheme.

2. *Encryption*: Upon input of a plain text message m, it chooses an integer r randomly from and outputs the cipher text $C = E(m, y) = (A, B) = (g^r, m \cdot y^r)$

Fig. 7. Initialization process.

Fig. 8. Registration process.

3. *Decryption*: On input of a cipher text (A, B) and the decryption key x, it outputs the plain text message $m = D(C, x) = B/A^x$.

3.4. Client–Server Authentication Using PAKE

In order to authenticate client to server and server to client during payment PAKE is proposed. The system has three phases namely initialization, registration, authentication and key exchange. Initialization is where the public parameters required for registration and authentication are established and published. Preceding to authentication each client chooses a password (*pwd*) and generates password authentication information (*Reg*) for server. After successful registration client only remembers only password for authentication. Finally during the key exchange phase the client establish different secret key with the server.

(a) *Initialization*:

In initialization module, the server publish the public parameter of the system. The server chooses cyclic group G of large prime order q with a generator g_1. Server randomly chooses $s1$ and $s2$ from G and calculate $g_2 = g_1^{s1s2}$. The server also chooses a hash function H.

(b) *Registration*:

Prior to authentication it is necessary that every client C register with the server. The client C chooses the password *pwd*. Then the C randomly chooses the decryption key x from G and calculate encryption key $y = g_1^x$ for server. Then the client encrypt the password *pwd* with y as $(A, B) = (g^a, g_2^{pwd} \cdot y^a)$. Where, $a = H\ (pwd)$ mod q. The client also finds another private parameter $b = H(pwd) \oplus a$. At last the client C delivers the password authentication information to the server. And the registration process is described in the Figure 8.

(c) *Authentication and key exchange*:

Assuming that the server received the password authentication information of a client C during the registration. Now the client and the server can mutually authenticate each other and can generate a secret key. The authentication and key exchange process shown in the Figure 9.

Fig. 9. Authentication and key exchange process.

4. SECURITY ANALYSIS

4.1. Security Analysis Against Passive Attack

Passive attack does not attempt to interfere with the stored data like listening to the transactions that are happening over. There are three messages exchanged between the client and server for key exchange and authentication. The first message from the client to the sever contains (UN, R, A). The first attribute does not contribute to any kind of passive attack. R is given by $(g_1^r \cdot g_2^{-pwd})$ which is difficult to factorize and said to be a discrete logarithmic or NP hard problem. The third parameter is A is hard to resolve as it is dependent on R. The next message is from the server to the client stating the hash value of the key generated through SHA512. Due to the usage of secure hash function, the data is preserved. Thus a passive adversary could obtain no useful information.

4.2. Security Analysis Against Active Attack

Active attack is performed by any attacker to damage or modify the information. Let us assume that the attacker has tracked the first message. He modifies the fields in it by guessing the formula for R. Certain parameters are retrieved from the public database and the other parameters are given random values. This modified message is sent to the server were the key value is generated in server without validating the client request using the value of B obtained from the private database. The server also computes the values of H and $R1$ which is sent to the client. The client generates the key which mismatches with that of the one generated in the server since the parameter B is a private data which also need the original password for computing. Therefore the server authentication fails at the client side.

5. CONCLUSION

From customers perspective the degree of complexity in buying any material is low as they can order for any product without visiting the store. As a result eliminating the time spent for purchasing. From seller's point of view, there is no need to establish the shop eliminating the cost for it. The existing retailers can attract their customers

RESEARCH ARTICLE

by providing the QR codes in their advertisements. The protocol for key exchange and client-server authentication is provided using ElGamal encryption scheme which makes the system secured against passive and active attack.

References

1. Denso Wave: About QR Code, Retrieved January **(2014)**, http://www.qrcode.com/en.
2. International Standard, ISO/IEC 18004. Information technology-Automatic identification and data capture techniques-QR Code 2005 bar code symbology specification, 2nd edn. **(2006)**.
3. Carolyn: QR Code Tutorial, Retrieved January **(2014)**, http://www.thonky.com/qr-code-tutorial/.
4. http://redlaser.com/application/, on April **(2014)**.
5. http://shopsavvy.com/downloads, on April **(2014)**.
6. http://www.micello.com/products/maps, on April **(2014)**.
7. ZXing: ZXing ("Zebra Crossing"), Retrieved January **(2014)**, https://github.com/zxing/zxing.
8. D. Munoz-Mejias, I. Gonzalez-Diaz, and F. Diaz-de-Maria, *IEEE Transactions on Consumer Electronics* 57, 1320 **(2011)**.
9. X. Yi, S. Ling, and H. Wang, *IEEE Transactions on Parallel and Distributed Systems* 24, 1773 **(2013)**.
10. William Stallings, Cryptography and Network Security: Principles and Practice, Second edn., Prentice-Hall, New Jersey **(1995)**.
11. Ed Burnette Hello Android: Introducing Google's Mobile Development Platform, Pragmatic Bookshelf, Third Edition, Andrew Hunt The Pragmatic Programmers, Raleigh, NC **(2010)**.
12. Herb Schildt, Java the Complete Reference, Eight edn., McGraw-Hill Osborne, USA **(2011)**.

Received: 17 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1551–1554, 2018

Convergence of Internet of Things, Big Data and Cloud—An Enhanced State Information Architecture

A. Bazila Banu* and Ravi Shankar Mishra

Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam 638401, India

From last few decades, Industries are using automated machines to do any work. Automated industries need to build Machine to Machine (M2M) model by using intelligent sensors and controllers. All the transactions in M2M creates a huge number of data in the form of text, video and speech, which is collectively called as Big Data. However data gathered from all the machines are not used efficiently and thrown without analysis. Also the corporate usage of the data was not understood clearly. Based on the data gathered from Internet of Things (IoT), Big Data use case should be designed. Otherwise data gathered from various devices will not be utilized properly for analysis and decision making process. Few challenges present in Big Data maintenance can be resolved by accommodating cloud for backup and recovery. In this paper, the authors proposed Enhanced State Information Architecture (ESIA) which act as an integrated framework for IoT, Big Data and cloud compared to future state information architecture (FSIA).

Keywords: Big Data, IoT, Cloud, Enhanced State Information Architecture.

RESEARCH ARTICLE

1. INTRODUCTION

Today, due to the development of different technologies like web technologies, social media, mobile sensing devices and automated machines the amount of data increases like exponential curve of mathematics. For example twitter possess 70M tweets per day, thereby generating 8TB daily. Big Data is a general term to express the huge amount of data which is so voluminous and complex that traditional data processing and application software are inadequate to deal with them.[1] The data are available in both form structured and unstructured. Big data is important because we can use these data to various purpose like analysis, decision making, forecasting and prediction. It is more useful for business purpose to predict the market on daily basis in order to take decisions more confidentially and accurately based of the results. The more data we collect the more accurate result we get and able to optimize business processes. Big data not only help us to achieve the business objectives but also acting as special aid to drive customer needs. Apart One of the biggest source of big data is IoT devices, which we can operate from far away using internet. Other sources as depicted in Figure 1 like healthcare, Finance, telecommunication, video surveillance and social media contributes a major volume. In present days the volume of big data is used in multiple ways to grow business and to know the world.[2]

1.1. Definition and Characteristics of Big Data

The term big data is used to describe the growth and the availability of huge amount of structured and unstructured data. Big data which are beyond the ability of commonly used software tools to create, manage, and process data within a suitable time.[2]

There are four main characteristics that defines the big data

1. *Data Volume*: The word Big Data itself defines the volume of data. At present the data existing is in petabytes and is supposed to increase in zettabytes in near future. Data Volume measure the amount of data available in an organization.[3] The size of the data determines the value and potential insight and whether it can actually be considered big data or not.
2. *Data Variety*: Data variety describes structural variation of data like text, audio, video, image etc. It also describes the data type it contains along with variety in which it represents.[1]
3. *Data Velocity*: This means how fast the data is being produced and how fast the data needs to be processed to meet the demands and the challenges which lie ahead in the path of growth and development.[4] This characteristics is not being related to the speed of data incoming data but also the speed at which data low and aggregated.
4. *Data Value*: Data value measures the usefulness of data in making decisions. Data science is exploratory and useful in getting to know data, but "analytic science"

*Author to whom correspondence should be addressed.

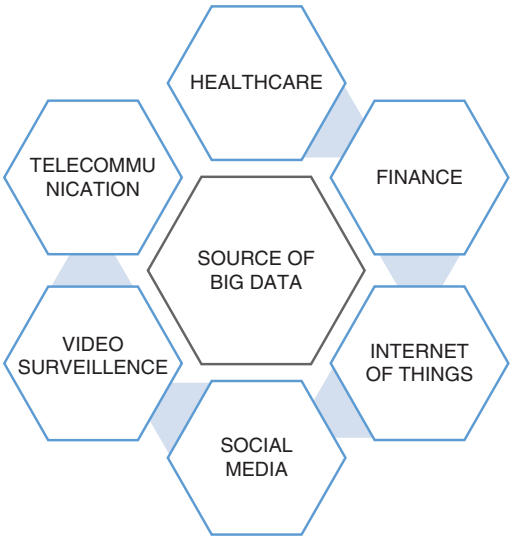


Fig. 1. Showing source of big data.

encompasses the predictive power of big data. User can run certain queries against the data stored and can deduct important results from the filtered data obtained and can also rank it according to the dimension they require. These reports help people to find the business trend according to change the business strategies.[3]

5. *Data veracity*: Big Data Veracity refers to the biases, noise and abnormality in Data. The data quantity vary greatly thereby affecting the accurate analysis.

According to Keyur et al.,[5,6] there are seven characteristics of big data. (1) Intelligence (2) Connectivity (3) Dynamic Nature (4) Enormous Scale (5) Sensing (6) Heterogeneity (7) Security. As we need these type of technology to do our work faster with high accuracy but using technology is not sufficient. We have to develop some good tools and software to handle these technologies.

1.2. Brief Review of IoT

The concept of "Internet of Things" acronym IoT, was given by the member of the Radio Frequency Identification (RFID) development community in 1999. IoT is the network of physical device. It is the network of all the devices like internet is the network of computers. IoT is a new revolution of the Internet. In IoT devices are connected in such a way that it can communicate and share with each other according to some protocol in order to smart reorganizations. We are using this technology in various field like in agriculture for sensing soil moisture and nutrient, healthcare to check the blood coagulation time and many more things and these day every organization is working on smart city project using IoT. Using IoT in most of the field need appropriate tools to handle the big volume of data.

1.3. Brief Review of Cloud

Cloud is the technology which allow us to store and access data and program over the internet instead of our computer's hard drive. Some of the examples of cloud includes Google Drive, Apple iCloud, Amazon Cloud Drive, Dropbox, Sugar Sync etc. According to the researcher,[13] we can classify cloud on the basis of our need like (a) Public Cloud—This can be accessed by any subscriber (b) Private Cloud—Established for an organization in such a way that the people of the that organization will only be able to access (c) Community Cloud—We can share this cloud between two or more organization which have same cloud requirements (d) Hybrid Cloud—Combination of any public, private or Community. Cloud technology has too many advantages like it may reduce the expense for storage of any organization, one can access the data anywhere in that way it reduces the chance of data failure i.e., retrieve the data anytime if lost. Usage of cloud helps to monitor the entire Infrastructure of an organization.

2. LITERATURE SURVEY

Opportunities always follow some challenges. In the similar way the new technologies gives us Opportunities to do our work more efficiently and accurately in less period of time. Technology like IoT allow us to handle things using internet from far away. There are so many issues and challenges are present for each and every technologies like lack of proper software or tools, internet speed etc. In which data related challenge is more prominent and need more attention because every day the rate of generating data is increasing. Challenges in data management depends on its Volume, Variety, Combining Multiple Data Sets, Velocity, Veracity, Quality, Availability, Discovery, Extensiveness, Personally recognizable Information, assertiveness, and Processing.

Researcher,[7] suggest that the challenges are mainly related to characteristics of big data. The volume of data is increasing but we have not proper tool to handle this. Also the data complexity is increasing. We are getting data in the form of raw, semi-structured and unstructured from different source. Here, the challenge is how to handle the variety of data? We are getting high voluminous and complex data with high speed but we have not proper technology to handle this overwhelming data flow.

Researcher,[8] suggest that the challenges are not only related to characteristics but due to storage of large volume of input/output. Due to the lack of adequate Information architecture huge amount of data is not utilized effectively. This challenge can be resolved by the researchers to overcome computational complexity, uncertainty, and inconsistencies in an effective manner.

Researcher[9,10] analyzed about the Data security and Privacy. Organization or companies are using user's personal information to perform illegal transactions. Big Data

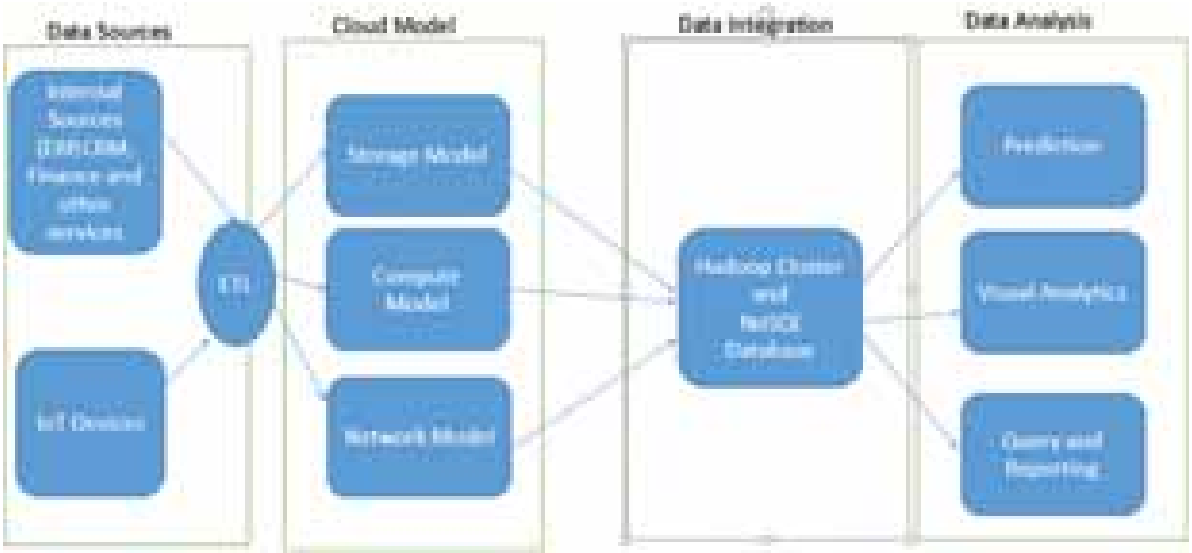


Fig. 2. Enhanced state information architecture.

Working Group at the Cloud Security Alliance Organization categorized the Data Security and Privacy challenge into four category (a) Infrastructure Security (b) Data Privacy (c) Integrity and Reactive Security (d) Data Management. To overcome these challenges, authors proposed Enhanced State Information Architecture by comparing the Future State Information Architecture.

Beyond these challenges, backup and recovery is also an important task to be handled for Big Data. NoSql (Not only SQL) database provides a mechanism for Key-Value, Document, Columnar and Graph storage. Researcher[11] analyzed the usage of NoSql database for handling big Data and identified solutions for multi-site deployment of NoSql databases with cloud storage.

3. ARCHITECTURE

3.1. Future State Information Architecture (FSIA)

FSIA proposed by the researchers consists of three layers such as data sources, data staging and data management. NoSql Database is used for integrating data sources and data management layer.[12] Although the architecture possess support for data integration, activities like data backup and recovery remains challenging for business purpose. Hence the authors proposes a new framework known as Enhanced State Information Architecture (ESIA).

3.2. Enhanced State Information Architecture (ESIA)

The main objective of ESIA is to integrate cloud for storage and retrieval of large volume of data. Among the different types of cloud capacity model. Hence the enhanced framework will act as an integrated platform for Big data, IoT and Cloud for data management and resource utilization. Figure 2 depicts the ESIA framework.

3.3. ESIA Framework Consist of the Following Layers

(i) *Data Sources and ETL*:

Data gathered from well-known in-house systems such as Enterprise Resource Planning, Customer Relationship Management and from IoT devices the corporate are then transformed to ETL (Extraction, Transaction and Loading) operation. Here the data will preprocessed for storage and retrieval process. Data formatting and cleaning will be accomplished in the data to eliminate redundant data.

(ii) *Cloud Model*

Business authorities should decide the cloud model based on their requirement to create a baseline of server, storage and network infrastructure. Jiang et al. proposed self-adaptive cloud capacity model. Business requirements for cloud capacity planning varies with respect to cost, resource usage and other factors. However suitable cloud model should be selected based on the demand forecasting techniques suggested by the authors.[14]

(iii) *Data Integration model*

To handle large volume of data, Hadoop cluster can be used as front end and NoSql database can be used as backend.

(iv) *Data Analysis*

Fourth layer comprises of series of data analysis such as prediction, query reporting and visual analytics. Different types of data analysis is possible due to the NoSql database.

4. DISCUSSION AND CONCLUSION

Overall the ESIA framework provides strong capacity management for handling big data obtained from various sources and IoT devices. Few challenges like low ACID

RESEARCH ARTICLE

(Atomicity, Consistency, Isolation and Durability) property can be resolved by accessing the data through cloud for further processing. This may avoid data loss. Moreover the data received from the IoT devices like sensors can be analyzed without errors with in stipulated time frame.

References

1. R. Stackowiak, A. Licht, V. Mantha, and L. Nagode, Big Data and the Internet of Things Enterprise Information Architecture for a New Age, 1st edn., A Press, Berkeley, CA **(2015)**, ISBN 978-1-4842-0986-8.
2. Jai Prakash Verma, Smita Agrawal, Bankim Patel, and Atul Patel, *International Journal on Soft Computing, Artificial Intelligence and Applications* 5, 41 **(2016)**.
3. U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, *J. Bus. Res.* 70, 263 **(2017)**.
4. M. Padmavalli, *IOSR Journal of Computer Engineering (IOSR-JCE)* 18, 13 **(2016)**.
5. K. K. Patel and S. M. Patel, *International Journal of Engineering Science and Computing* 6, 5 **(2016)**.
6. Ahmed Bader, Hakim Ghazzai, Abdullah Kadri, and Mohamed-Slim Alouini, *IEEE* 4, 3257 **(2016)**.
7. Barnaghi, P. Barnaghi, A. Sheth, and C. Henson, *IEEE Intell. Syst* 67 **(2013)**.
8. D. P. Acharjya and P. Kauser Ahmed, *(IJACSA) International Journal of Advanced Computer Science and Applications* 7, 511 **(2016)**.
9. J. Moreno, M. A. Serrano, and E. Fernández-Medina, *Future Internet* 8, 4 **(2016)**.
10. Atish Kathpal and Priya Sehgal, BARNS: Towards building backup and recovery for NoSQL databases, *Proceedings of the 9th USENIX Conference on Hot Topics in Storage and File Systems*, USENIX Association Berkeley, CA, USA **(2011)**, p. 3.
11. Vidushi Jain and Aviral Upadhyay, *International Journal of Computer Applications* 167, 10 **(2017)**.
12. K. Grolinger, W. A. Higashino, A. Tiwari et al., *J. Cloud Comp.* 2, 22 **(2013)**.
13. Alexa Huth and James Cebula, Basics of Cloud Computing, United States Computer Emergency Readiness Team **(2011)**.
14. Y. Jiang, C.-S. Perng, T. Li, and R. Chang, Self-adaptive cloud capacity planning, *2012 IEEE Ninth International Conference on Services Computing* **(2012)**, p. 73.

Received: 17 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1555–1561, 2018

Robust Sybil Attack Detection Mechanism for Social Networks

A. Bharathi, D. Yazhini Priyanka*, P. Padmapriya, and K. Priyadharsnee

Department of Information Technology, Bannari Amman Institute of Technology, 638402, Tamil Nadu, India

An online social network (OSN) is a platform to build social networks or social relations among people. Nowadays, OSN are commonly used by billions of users to interact, and they are key platforms for content and opinion dissemination, social and professional networking, recommendations, scouting, alerting, and political campaigns. This paper mainly focus on recommended systems (RSs) in OSN. Accuracy and robustness are important performance metrices that characterize customized information or suggestions provided by RSs. One of the challenge is faced by OSNs, Sybil attack in which a single malicious user introduces multiple bogus identities and pretends to be multiple and real users in the network. The proposed work is to enhance the robustness and accuracy of the recommendation system. Sybil-Resist and influence level enabled Recommendation system (LainRec) techniques are used in our proposed work which overcomes the Sybil attack in social network.

Keywords: Sybil Attack, Online Social Network, Link Analysis, Recommender System.

RESEARCH ARTICLE

1. INTRODUCTION

1.1. Sybil Defense Mechanism in Online Social Network

The network is a combination of Sybil and non-Sybil regions. In such network to discover malicious nodes following assumption is made. Even if an attacker creates numerous identities in a social network, it won't be able to set up massive social links with legitimate nodes. Thus, it is weakly connected to the rest of the network but can connect to other Sybils.

1.2. Sybil Attack

In an Online Social Network[2] not all users are honest user. A malicious user creates fake identities (Sybils) and suppresses the opinions of honest users. An attack using such Sybils is known as Sybil attack. However the emerging IoT is vulnerable to Sybil attacks where attackers can manipulate fake identities or abuse pseudo identities to compromise the effectiveness of the systems. In the presence of Sybil attacks, the IoT system may generate wrong reports, and user 8 might receive spam and lose their privacy. From a recent report in 2012, a substantial number of user accounts are confirmed as take or Sybil accounts in online social networks (OSN), totally 76 million (7.2%) in facebook, and 20 million fake accounts created in twitter per week. Recent research efforts have been focused on studying Sybil attacks and how to detect and defend them.

A Sybil attack[6] is also used by companies that increase the Google Page Rank rating of the pages of their customers, and has been used to link particular search terms to unexpected results for political commentary. Reputation systems are a common target for Sybil attacks including real-world systems like eBay. The vulnerability of a reputation system to a Sybil attack depends on how cheaply the identities can be created. An entity on a peer-to-peer network is a piece of software, and it has access to local resources. By using an identity, an entity can advertise itself on the peer-to peer network and also can mapping to many identities.

1.3. Communication Models of Sybil Attack

Sybil attacks consists of four communication models such as Direct/Indirect communication, simultaneous/Non-simultaneous communication. Let us discuss the each communication model briefly.

In direct communication, The Sybil node communicates directly with the honest node. When an honest node broadcasts a message in the network, the Sybil node which shares a link with that node will listen to this message. In Indirect communication, there isn't any honest node can be reached by the Sybil node. On the contrast, one or some Sybil nodes declare that they could reach the honest region. So they will use a Sybil node which

*Author to whom correspondence should be addressed.

has a relationship with an honest node to transmit their messages.

In simultaneous communication, in some cases, an adversary can fabricate many Sybil identities. For example, if the identity of a node is a 32 bit integer, then the adversary can assign a 32 bit value to the Sybil node as its identity. But if there is a mechanism that can verify the identity of the Sybil node, like naming space, then the adversary can't fabricate fake identities. So, in that case, the adversary needs to assign a legal identity to the Sybil node. This kind of identity theft is hard to be detected if the adversary invalidates the original node. Simultaneous communication means the adversary let its all Sybil identities take part in one network communication. So if one node only be allowed to use its identity once, the adversary can recycle some of its Sybil identities to make that Sybil node seems like many honest identities.

And the last communication model is Non-simultaneous communication model. If the attacker uses some of the Sybil nodes only in a specified time period and other Sybil nodes in another time period, then it will look like the regular action of the honest nodes. Although Sybil attacks could be alleviated by using a trusted central authority by issuing credentials of the real users or requiring some kinds of payments, it is hard to achieve because it is difficult or impossible to establish a single entity that all the users of the website are willing to trust. Some other ways, like binding every user to IP address or requiring users to solve a simple puzzle, cannot make effective protections, because the adversary is able to harvest IP addresses which is quite different to each other, thereby it is difficult to filter them. So need to build up a scheme to analysis the characteristics of the Sybil attack and defend it by some computer-based algorithms.

2. RELATED WORK

Melville et al.[4] introduce Recommendation systems can be attacked in various ways, and the ultimate attack form is reached with a Sybil attack, where the attacker creates a potentially unlimited number of Sybil identities to vote. This paper presents DSybil, a novel defense for diminishing the influence of Sybil identities in recommendation systems. Mobasher et al.[5, 12] introduce how to build the recommendation system using collaborative filtering approach. Collaborative filtering is the most popular approach to building recommender systems and has been successfully employed in many applications. This approach assumes a social network among users and makes recommendations for a user based on the ratings of the users that have direct or indirect social relations with the given user. As one of their major benefits, social network based approaches have been shown to reduce the problems with cold start users.

Gibbons[8] introduce how fully self-organized mobile adhoc networks (MANETs) are weaken by a Sybil user. Due to the complex nature of MANETs and its resource constraint nodes, there has always been a need to develop lightweight security solutions. Since MANETs require a unique, distinct, and persistent identity per node in order for their security protocols to be viable, Sybil attacks pose a serious threat to such networks. A Sybil attacker can either create more than one identity on a single physical device in order to launch a coordinated attack on the network or can switch identities in order to weaken the detection process, thereby promoting lack of accountability in the network. This paper introduce a lightweight scheme to detect the new identities of Sybil attackers without using centralized trusted third party or any extra hardware, such as directional antennae or a geographical positioning system. Kuan Zhang et al.[15] shows that the emerging Internet-of-Things (IoT) are vulnerable to Sybil attacks where attackers can manipulate fake identities or abuse pseudo identities to compromise the effectiveness of the IoT and even disseminate spam. Social network based Sybil detection (SNSD) mechanism is used to mitigate the Sybil attack. SNSD schemes usually partition Sybil nodes and honest ones into two parts: Sybil region and honest one, these would be viewed as a graph partitioning problem. In this unknown node is ranked according to its social connections with the known trusted nodes.

Guojun Wang et al.[16] introduces how to address the sybil and passive attacks in E-Commerce applications. Peer to peer (P2P) e-commerce applications exist at the edge of the Internet with vulnerabilities to passive and active attacks.

This paper exploits the neighbor similarity trust relationship (SybilTrust) to address Sybil attack. SybilTrust that can be used to identify and protect honest peers from Sybil attack. The Sybil peers can have their trust cancelled and dismissed from a group. Based on the group infrastructure in P2P e-commerce, each neighbor is connected to the peers by the success of the transactions it makes or the trust evaluation level. A peer can only be recognized as a neighbor depending on whether or not trust level is sustained over a threshold value. SybilTrust enables neighbor peers to carry recommendation identifiers among the peers in a group. This ensures that the group detection algorithms to identify Sybil attack peers to be efficient and scalable in large P2P e-commerce networks.

In Ref. [17] NodeXL is used to explore the social network graph with graph metric calculation. The NodeXL template displays graphs using a custom Windows Presentation Foundation (WPF) control that can be reused in custom applications. In fact, the template is just an application wrapper around a set of reusable, prebuilt class libraries.

3. PROPOSED WORK

There are two possible models for the implementation of the Sybil defense mechanism based on the link analysis

and influence level update. To mitigate the consequences of Sybil attacks, techniques can be divided into two sub areas; exploiting social network characteristics and analyzing link structures. SybilResist scheme is used to identify the Sybil attack. In this Sybil-Resist scheme based on the random walk theory to determine the Sybil region and defend the Sybil attack.

Figure 1 shows the architecture of our proposed work, First analyse the rank matrix for set of users on set of items. After that, draw the weighted rating graph for each and every user with each item. Our proposed LainRec design is used to analyse the link structure and relationship between the users in online social networks. LainRec exploits the properties of link analysis and influence levels iteratively.

3.1. Influence Initialization Bipartite Graph

The main purpose of LainRec[18] is to iteratively disseminate influence levels. In order to exploit this idea, view a user-item rating matrix as a bipartite graph since the rating matrices consist of two basic sets; user and item sets.

Figure 2 shows the rating matrix of users and items. Let $R_{m,n}$, be the rating matrix (containing m users and n items), the rating score from user i on item j, the set of users containing m users and the set of items containing n items, respectively. There are two phases to generating a weighted bipartite graph. Bipartite graph converts the rating matrix into two disjoint sets; user and item sets.

(1) User item matrix converted into a weighted rating graph (G) in Figure 3.

$$G = \begin{cases} (U, V, E) \rightarrow R_{m,n} \text{ such that} \\ U_i \rightarrow \text{user } i, \quad V_j \rightarrow \text{item } j \end{cases} \tag{1}$$

(2) Weighted rating graph is transformed into weighted opinion graph (G'). To represent the link weight as

Fig. 2. Rating matrix.

three opinions (positive, negative and neutral) according to Ri, j.

Link opinion function defined by using,

$$f'(Ei,j) = \begin{cases} +1, & \text{if user gives positive recommendation} \\ 0, & \text{if user gives neutral recommendation} \\ -1, & \text{if user gives negative recommendation} \end{cases} \tag{2}$$

3.2. Sybil-Resist Scheme

In this phase[7] social network will be taken as a graph G which consists of edges E and Vertices V. The majority of V is honest users which are also called as honest nodes. There are also several adversaries in the social network and each of them has a number of Sybil identities. Each Sybil identity is regarded as a Sybil node in graph G. If two identities are friends, then there is an edge between those two nodes which represents their relationship. The edges are considered as undirected. All the honest nodes consist of an honest region while all the Sybil nodes consist of a Sybil region, a Sybil region is controlled by the malicious user, and thus there can be any number of edges among the Sybil nodes.

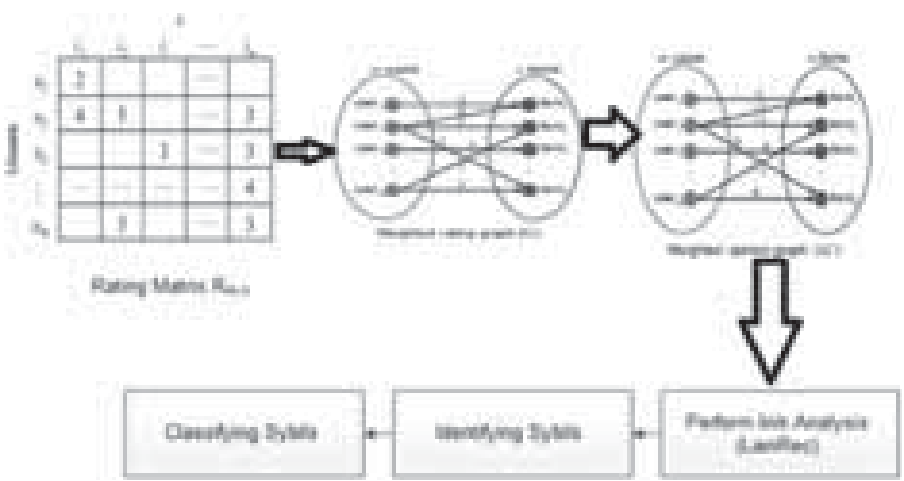


Fig. 1. Architecture of proposed model.

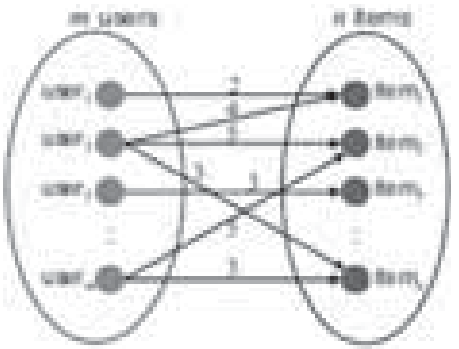


Fig. 3. Weighted rating graph.

The proposed algorithm consists of two main parts: Pre-processing algorithm which is used to pre-process the topology and get some characteristic values of the honest node; and a Sybil identification algorithm to identify the suspect nodes in the social network.

3.3. Algorithm for Pre-Processing

1. Preprocessing $(G, h)J = \{\}$
2. for $i = 1$ to f
3. Perform a random walk with length $lS = \log n$ starting from h $J = J$ U {the ending node of the random walk}
4. End
5. $l = l\min$
6. while $l \Leftarrow l\max$
7. for $i = J(1)$ to J(last)
8. Perform R random walks with length l starting from node i
9. Get ni as the number of nodes with frequency no smaller than t end
10. Out = {1, mean($\{ni: I \in J\}$), stdDeviation ($\{ni: I \in J\}$)
11. $l = l +$ interval
12. end

In this algorithm, social network will be treated as a social graph $G(V, E)$. Random walks play an important role in this scheme. A random walk on a graph refers to a movement that a node randomly selects one of its neighbors to move. If node i has the degree di, then the probability that the node will make a movement to one of its neighbor is $1/di$. This program first performs f random walks with length $ls = \log n$ starting from the known node h. The random walk starts from the minimum length, and adds an interval in every loop until. For each l, the algorithm performs R random walks starting from every node in J. Then we count the frequency of every node and select all the nodes whose frequency is no smaller than a threshold t (a constant with value of 5).

3.4. Algorithm for Sybil Identification

The second part is to determine whether a suspect node u is a Sybil node or not. The pseudo code is given as follows.

1. Sybil Identification (G, u, output of Alg. 1)
2. $l = l0$
3. while $l \Leftarrow l\max$
4. perform R random walks with length l starting from u
5. $m =$ the number of nodes whose frequency is no smaller than t
6. let the tuple corresponding to length l in the output of Alg. 1 {1, mean, stdDeviation}
7. If mean-$m >$ stdDeviation $* \alpha$ then
8. Output u is a Sybil node End
9. $l = l * 2$ End
10. Output u is an honest node

Sybil identification algorithm makes a comparison between the number of nodes whose frequency is bigger than or equal to t and the mean value in tuple {l, mean, stdDeviation} in Algorithm 1. And if the inequality mean-$m >$ stdDeviation $* \alpha$ is valid, output is Sybil node and end this algorithm. Otherwise, double the value of l and do the loop again until l is larger than the upper limit of l, $l\max$. And if the inequality mean-$m >$ stdDeviation $* \alpha$ is still false, suppose that the node is an honest node. To set the value of t, we perform R random walks starting from h with length $l\max$, the number of nodes with frequency bigger than or equal to t should be larger than $|V|/2$. This defending scheme adopt adaptive test model to identify the suspect node in order to guarantee that it can fit all sizes of Sybil regions.

4. IMPLEMENTATION

The simulation of social network graph is implemented using NodeXL. NodeXL is a free and open-source network analysis and visualization software package. NodeXL is a set of prebuilt class libraries using a custom Windows Presentation Foundation control additionally. NET assemblies can be developed as "plug-ins" to import data from outside data providers. Currently implemented data providers for NodeXL include Facebook, Twitter, and Wikipedia.

4.1. Data Import

NodeXL imports UCINet and GraphML files, as well as Excel spreadsheets containing edge lists or adjacency matrices, into NodeXL workbooks. NodeXL also allows quick collection of social media data via a set of import tools which can collect network data from e-mail, Twitter, YouTube, and Flickr. NodeXL requests the user's permission before collecting any personal data and focuses on the collection of publicly available data, such as Twitter statuses and follows relationships for users who have made their accounts public. These features allow NodeXL users

Table I. Overall graph metrices.

Parameters	Value
Vertices	363
Unique edges	450
Edges with duplicates	0
Total edges	450
Self-loops	0
Reciprocated edge ratio	Na
Connected components	1
Single-vertex connected components	0
Maximum vertices in a connected component	363
Maximum edges in a connected component	450
Maximum geodestic distance (diameter)	4
Average geodestic	2.594298
Graph density	0.00684900232866079
Modularity	0.387311

to instantly get working on relevant social media data and integrate aspects of social media data collection and analysis into one tool.

4.2. Data Representation

NodeXL workbooks contain four worksheets: Edges, Vertices, Groups, and Overall Metrics. The relevant data about entities in the graph and relationships between them are located in the appropriate worksheet in row format. For example, the edges worksheet contains a minimum of two columns, and each row has a minimum of two elements corresponding to the two vertices that make up an edge in the graph. Graph metrics and edge and vertex visual properties appear as additional columns in the respective worksheets. This representation allows the user to leverage the Excel spreadsheet to quickly edit existing node properties and to generate new ones, for instance by applying Excel formulas to existing columns.

4.3. Graph Analysis

NodeXL contains a library of commonly used graph metrics: centrality, clustering coefficient, and diameter. NodeXL differentiates between directed and undirected networks. NodeXL implements a variety of community detection algorithms to allow the user to automatically discover clusters in their social networks.

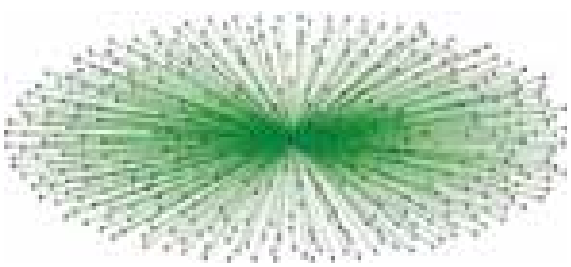


Fig. 4. Visualization of social graph.

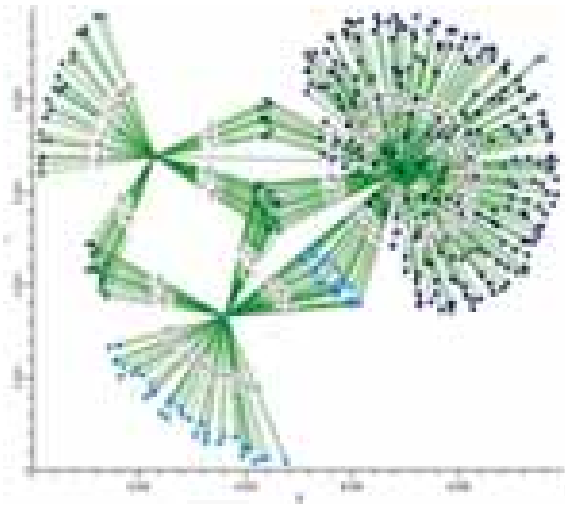


Fig. 5. Visualization of opinion graph.

4.4. Description of Data Set

MovieLens data sets were collected by the Group Lens Research Project at the University of Minnesota. This data set consists of 1,000,209 ratings of approximately 3,900 movies made by 6,040 MovieLens users. The data was collected through the MovieLens web site (movielens.umn.edu) during the seven month period from September period from September 19th, 1997 through April 22nd, 1998.

4.5. Overall Graph Metrices

Table I shows the general information of the topology.

Fig. 6. Pre-processing topology.

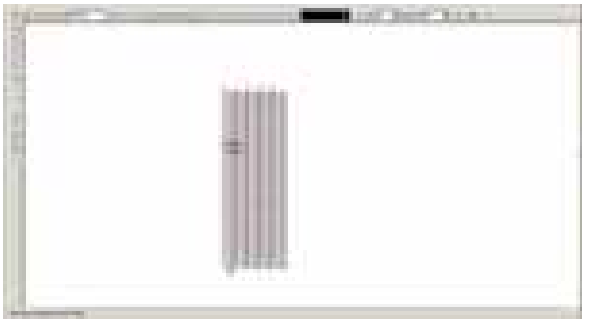


Fig. 7. Node deployment.

4.6. Construction of Social Graph

NodeXL visualization tool is used to generate social graph using MovieLens data set. The tool supports multiple social network data providers that import graph data. Figure 4 shows the construction of social graph. Social graph representing the relationship between the user and items.

Figure 5 shows the construction of opinion graph. Link weight as three opinions (Positive, negative and neutral).

Figure 6 shows the preprocessing topology. Preprocessing is used to get the characteristics value of the topology. Social graph and honest node these two parameters are taken as inputs of the pre-processing algorithm. Random walks play an important role in this scheme. A random walk on a graph refers to a movement that a node randomly selects one of its neighbors to move. The random walk starts from the minimum length, and adds an interval in every loop until.

Figure 7 shows the deployment of node. NodeXL trace file will be given to input of NetAnim which consist of 380 nodes.

Figure 8 shows the detection of Sybil node. Red nodes represent honest nodes while blue nodes represent Sybil nodes. Algorithm makes a comparison between the number of nodes whose frequency is bigger than or equal to t and the mean value in tuple $\{l$, mean, stdDeviation$\}$ in Algorithm 1. If the inequality mean-$m >$ stdDeviation $* \alpha$ is valid, u is a Sybil node.

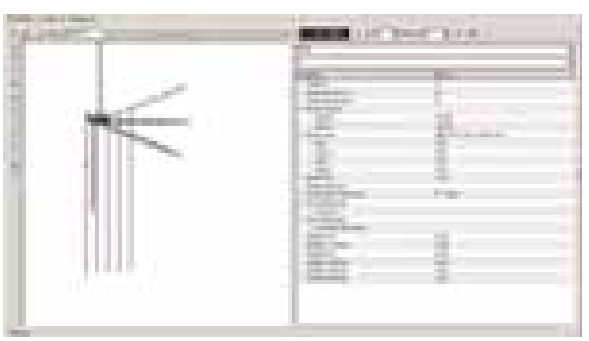


Fig. 8. Detection of Sybil node.

Fig. 9. Detection of sybil region.

Figure 9 shows the detection of Sybil region to detect the Sybil region among the detected nodes. Sybil region detection Algorithm takes a social topology $G(V, E)$ and a known Sybil node s as inputs. Sybil region as a sub graph of G which doesn't have a small cut, and it consists of all the Sybil nodes.

5. CONCLUSION

The main goal of RSs is to provide personalized recommendations suited to a user's taste after analyzing patterns of user interest in products. Sybil attack is an attack where malicious users suppress opinions of honest users by creating fake identities called Sybils. Sybil attacks are prevalent in wireless sensor networks, online social networks, Peer to Peer networks and mobile networks. Sybil attack proves to be a threat to the recommendation system reducing its robustness. LainRec method has been introduced which uses link analysis and iteratively updating influence level by transforming a recommendation matrix into a bipartite graph. User and item influence level updating function are introduced to verify Sybil users using the bipartite graph. LainRec iteratively update the influence level until there is no additional change. Social rating graph is constructed using nodeXL visualization tool. After that Sybil-Resist scheme based on the random walk theory is introduced to identify the Sybil node. Simulations are performed using Network Simulator-3. MovieLens dataset comprising of 1,000,209 anonymous ratings of approximately 3,900 movies made by 6,040 MovieLens users out of 1048,576 ratings are used for proposed system.

References

1. B. Levine, C. Shields, and N. Margolin, A survey of solutions to the Sybil attack, Technical Report, University of Massachusetts Amherst, Amherst, MA **(2006)**.
2. B. Mobasher, R. Burke, R. Bhaumik, and C. Williams, *ACM Trans. Internet Technol.* 7, 23 **(2007)**.
3. G. Danezis and P. Mittal, SybilInfer: Detecting Sybil nodes using social networks, *Proc. of ISOC NDSS* **(2009)**.
4. G. Linden, B. Smith, and J. York, *IEEE Internet Comput.* 7, 76 **(2003)**.

5. G. Noh, H. Oh, K.-H. Lee, and C.-K. Kim, *Journal of Communications and Networks* 17 **(2015)**.

6. G. Wang, F. Musau, S. Guo, and Muhammad Bashir Abdullahi, *IEEE Transactions on Parallel and Distributed Systems* 26, 824 **(2015)**.

7. H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman, *IEEE-ACM Trans. Netw.* 16, 576 **(2008)**.

8. H. Yu and P. B. Gibbons, *IEEE-ACM Trans. Networks* 18, 885 **(2010)**.

9. H. Yu, P. B. Gibbons, and M. Kaminsky, Toward an optimal social network defense against Sybil attacks, *Proc. of ACM PODC* **(2007)**, pp. 376–377.

10. H. Yu and C. Shi, DSybil: Optimal Sybil-Resistance for Recommendation Sytems **(2004)**, pp. 259–268.

11. J. Douceur and J. S. Donath, The sybil attack, *Proc. of ACM IPDPS* **(2002)**, p. 251260.

12. K. Zhang, X. Liang, R. Lu, and X. S. Shen, *IEEE Internet of Things* 1, 372 **(2014)**.

13. M. Jamali and M. Ester, A matrix factorization technique with trust propagation for recommendation in social networks, *Proc. 4th ACM Conf. Recommender Syst.* **(2010)**, pp. 135–142.

14. P. Melville and V. Sindhwani, Encyclopedia of Machine Learning, Springer, Boston, MA **(2011)**, pp. 829–838.

15. P. Melville, R. J. Mooney, and R. Nagarajan, Content-boosted collaborative filtering for improved recommendations, *Proc. Nat. Conf. Artificial Intell.* **(2002)**, pp. 187–192.

16. W. Wei, F. Xu, C. C. Tan, and Q. Li, SybilDefender: Defend Against Sybil Attacks in Large Social Networks, *Proc. IEEE INFOCOM* **(2012)**, pp. 1951–1959.

17. W. Ma, S.-Z. Hu, Q. Dai, and T.-T. Wang, Sybil-resist: A new protocol for sybil attack defense in social Network, *Proc. of ACM PODC* **(2007)**, pp. 376–377.

Received: 18 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1562–1567, 2018

RESEARCH ARTICLE

Location Identification of Sensor Nodes Using Dynamic Path Planning Mobile Nodes in Wireless Sensor Networks

P. Purusothaman*, M. Gunasekaran, and B. Gopalakrishnan

Department of Information Technology, Bannari Amman Institute of Technology, 638401, Tamil Nadu, India

Localization is a crucial issue in wireless sensor networks. The location accuracy of sensor node increases with the number of beacons points which can provide reference information for localization. However, implementation cost is increased if more number of beacons is used in wireless sensor networks. Accuracy, cost and scalability are three main factors to be considered when designing localization method with high energy efficiency. Developing a Range-free localization technique with a single mobile beacon equipped with GPS offers coarse positioning accuracy with low computational expenses. This paper presents a D* SQUARE algorithm which gives better accuracy with a single anchor node. The path of a mobile anchor node is considered in such a way to optimize the errors and get best location information considering the presence of obstacles. The location of all the sensors are Called using single mobile beacon and the mechanism is implemented in NS2.

Keywords: Mobile Anchor Node, D* SQUARE, Beacons, Sensors, GPS.

1. INTRODUCTION

Wireless sensors are gadgets which are utilized to detect some physical amounts like temperature, weight, mugginess and so on.[1] The detected information is exchanged to the information sink. The exchange of information frame sensor to sink is by multi bounce component. In numerous applications vitality source substitution is unrealistic. So multi-jump components are received to spare vitality of the sensors since they are low controlled gadgets. Hence the life of every sensor can be expanded.

A large portion of the application like woodland fire location, territory observing and so on requires area data of the sensors. Fundamentally the limitation component can be isolated into run based and go free. In range based plan, the area data is more exact than range free yet it is intricate and costly. Keeping in minimum the end goal to actualize this plot extra equipment is required.[2] The areas are figured from the separation or precise information's between thee hubs. With a specific end goal to acquire this data infra-red, ultrasonic or X-beams gadgets are utilized. As a result go based plan is more minimum boggling and costly likewise the framework require extra types of gear to get these sort of signs that make the gadget massive. Some range based plans are AOA, TDOA, TOA, and RSSI.[11, 13] In range free plan portable hub is equipped with GPS recipient. The hub moves arbitrarily through the zone of intrigue and intermittently communicates a reference point message. The guide has the present arrange of the versatile hub. From this data the area of different hubs are resolved. A few mobile hub developments in view of particular calculation are depicted in Refs. [3–8].

The description of this paper is as follows, Chapter 2 explains the related papers in Localization, Chapter 3 explain the Wireless Sensor Networks Design, Implementation and calculation behind the Signal Point Selection. Chapter 4 is the proposed work which compares Dynamic Square Trajectory Algorithm with the Path Planning Algorithm. Chapter 5 shows the result analysis followed by conclusion.

2. RELATED WORK

In the localization algorithms introduced by Sichitiu and Ramadurai,[6] the sensor hubs decide their locations in light of the reference point messages transmitted by a versatile mobile sensors. In any case, for this situation, the sensor hubs evaluate their positions by applying the RSSI system to the guide message.

Galstyan et al.[7] proposed a dispersed confinement plot in view of a solitary versatile stay, in which radio availability imperatives were forced so as to minimize the vulnerability in the restriction comes about.

Zhong and He[8] proposed different probabilistic limitation plans in view of a solitary versatile guide, and demonstrated that non-parametric probabilistic estimation

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1562/006 doi:10.1166/jctn.2018.7340

systems yield more strong and precise restriction comes about than parametric estimation techniques.

Koutsonikolas et al.[9] proposed three way path planning for the portable sensors hub in the restriction way introduced by Sichitiu and Ramadurai, in particular SCAN, DOUBLE SCAN and HILBERT. In SCAN, the portable sensors stay along a standard path measurement (e.g., the *x*-pivot or *y*-hub bearing), and the separation between two neighboring path of the hub direction characterizes the determination of the direction. SCAN is straightforward and gives uniform scope to the whole system. Be that as it may, the co linearity of the reference points corrupts the exactness of the limitation comes about. In DOUBLE SCAN, the co linearity issue is settled by driving the mobile sensors in both the *x*-and the *y*-bearings. In any case, while this technique enhances the confinement execution of the sensor hubs, the way length is increased with that of SCAN, and along these lines the message overhead increments as needs be. In HILBERT, the portable grapple hub is driven in a bended direction with an end goal that the sensor hubs can build non-collinear guide focuses and the total way length is lessened which reduces the message overhead.

Huang and Zaruba[10] exhibited two further way arranging plans assigned as CIRCLES and S-CURVES, separately, to avoid the co linearity issue natural in the limitation strategy. In CIRCLES, the portable sensor takes path in concentric round directions focused at the inside purpose of the sending range, while in S-CURVES, the stay takes path in S-molded bend as opposed to a straightforward straight line as in the SCAN technique.

Hu and Evans presented a way arranging plan for a versatile grapple hub in light of the trilateration restriction scheme.[11] The stay hub moves as per an equilateral triangle direction and communicates its present position data in the detecting zone. In the wake of accepting three position data, every sensor hub can evaluate itself area in light of trilateration count. The separation between the sensor hub and the stay hub can be measured in light of Received Signal Strength (RSS).

Liao and Lee et al. introduced a versatile way making arrangements for the portable grapple node.[12] The stay hub moving as indicated by a consistent triangle with the length of its radio range first communicates three reference point messages. The sensor hubs that don't have the foggiest idea about their own particular positions then demand the stay hub to convey more guide messages. The grapple hub in this manner chooses its direction on the premise of the got ask for messages.

Yun et al.[13] has proposed an appropriated restriction location selection in light of Received Signal Strength Indicator (RSSI) by the three anchor messages. The separation from three portable anchors is resolved and the unknown hubs can process its area by utilizing roundabout triangulation idea. In light of the utilizing basic concept of RSSI Messages comes with respect to the confinement overheads and power utilization, they have proposed and checked that the triangulation approach performs moderately much superior to anything two anchor and single anchor methodologies.

Chan et al.[14] has proposed a novel approach of determining situating the hubs in a WSN with the utilization of the node to node distance measurement techniques like, Received Signal Strength (RSS) or Time of Arrival (ToA) estimations. They have utilized three location aware messages for subspace calculation of situating the hubs in a WSN.

Wang and Jin[15] has proposed an enhanced Approximate Point-in-Triangulation Test (APIT) calculation for area estimation in WSN which performs best when contrasted with the first APIT calculation Technique, in certain special scenarios like Irregular radio pattern, random node placement, with very low communication overhead.

RESEARCH ARTICLE

3. WIRELESS SENSOR NETWORK DESIGN

WS Network is a ad hoc system of portable nodes associated by remote connections which connects forms a network with nodes which are in its vicinity. Every node in this system is allowed to move autonomously in any direction, and while moving along these lines change its connections to different static sensor nodes regularly and broadcasts its messages. Each must forward the location messages irrelevant to its own particular to the routing device. The essential test in building a WS Network is preparing every node to constantly keep up the data required to legitimately course activity. Such systems may work independent from anyone else or might be associated with the bigger Internet. WSN is a sort of adhoc based networking systems. A new simulation method is created using the parameters defined in tcl with mobility pattern included to the mobile anchor nodes which performs the following operations:

- Topology introduction
- Execution
- Agent
- Link Creation.

The Topology introduction sets up field settings inside parcels utilized by the whole simulation. The cron scheduler runs the network simulator in timely defined way and can be supported by other schedulers. The agent is identifying the sink node where all the location estimation messages get transferred to make the calculation. Link creator is used to network protocol implementation strategy which denoted the topology used in implementation like AODV, DSR or DSDV routing implementation protocol.

3.1. Implementation

The implementation plan was motivated by the perpendicular bisector of a chord guess. The Figure 1 portrays that

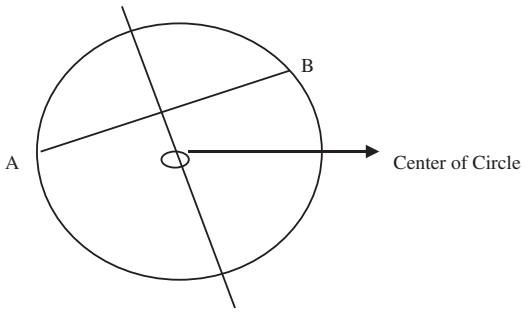


Fig. 1. Perpendicular bisector of a chord conjecture.

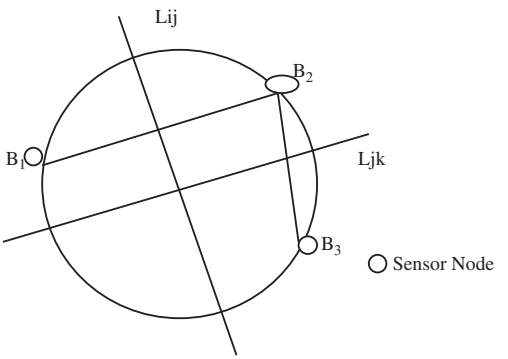


Fig. 2. Perpendicular bisector chord.

the perpendicular bisector of any chord passed through the center of the circle. As appeared in Figure 1, the chord AB is a straight line joining 2 points on the circumference of the circle.

3.2. Signal Point Selection

In the implement, two chords are established by selecting two end points. A message containing the node (id, location, timestamp) is broadcasted by each portable anchor node and it is received by the sensor node, which are in its list. Each sensor node keeps up an arrangement of reference point focuses and a guest list. The reference point is created which is endpoint on the sensor node's correspondence circle. The maintained list stores both the portable stays whose messages have been received by the sensor node and their timestamp. The *n*th guide point in the sensor is spoken to as (id, location, time stamp) and the nth passage in the guest rundown can be recorded as (id, lifetime). At the point when a sensor node gets a reference point message from a versatile grapple point, the node will check whether the stay point is in its guest list. If not, a guide point will be included and the stay point with a predefined lifetime will be embedded in the guest list. Something else, the reference point message will be overlooked and the lifetime of the portable stay point will be developed. A mobile node entry in the maintained list is removed when the life time of the message is expired based on its time stamp, the last signal point is considered as the reference point.

3.3. Sensor Location Estimation

After three beacons are acquired, two unique chords can be created. As appeared in Figure 2 the arrangement of chose signal focuses is are listed as B_1, B_2, B_3 and their areas are in two dimensional coordinates are mentioned as (xi, yi), (xj, yj), (xk, yk), and Two chords randomly picked B_1B_2, B_1B_3, and, are created on the reference point of beacons. Consider that lines Lij and Ljk are the comparing opposite bisectors of the chords B_1B_2, B_1B_3 and B_2B_4, individually. In this manner, by simple logarithmic computation, the conditions of two lines Lij and Ljk can be exhibited as follows:

$$\begin{aligned} &\{Lij:\ a12x + b12y = C12 \\ &\{Ljk:\ a23x + b23y = C23 \end{aligned} \tag{1}$$

On the off chance that three signal focuses are gotten on the communication circle of a sensor node, it takes after that the portable anchor node must go through the sensor nodes circle no less than two events.

3.4. Path Planning Algorithm

In the predefined path plan proposed in this review, the distance between two progressive vertical portions of the anchor direction (i.e., the determination of the stay trajectory) is indicated as $R - X$, where R is the communication radius of the anchor node and X is set in the range $0 < X \leq R/3$. This is based that if X is greater than $R/3$, $R - X$ will be smaller than $2R/3$. Consequently, the separation between four progressive vertical segments is not exactly the distance across of the correspondence circle (i.e., $2R$). Therefore, the mobile anchor will pass through the sensor nodes circle more than three circumstances. At the end of the day, expanding the value of X may acquire excess guide focuses.

On the other hand, diminishing the estimation of X may bring about the chord length to fall beneath the threshold value. Therefore, a careful decision of X is required. The dimensions on the boundary are carefully chosen as it is very difficult to find out the location of the sensor node that are situated close to the boundary regions, as appeared in Figure 3. By extending the detecting field, and picking a proper estimation of X, the proposed way arranging plan guarantees that the portable anchor node goes through the radius circle of every sensor node either a few circumstances. As appeared in Figure 3, the aggregate way length D is given as in Eq. (2).

$$\begin{aligned} D = (L+2R) * (([L+2R]/(R-X)) + 1) + (R-X) \\ * ((L+2R)/(R-X)) \end{aligned} \tag{2}$$

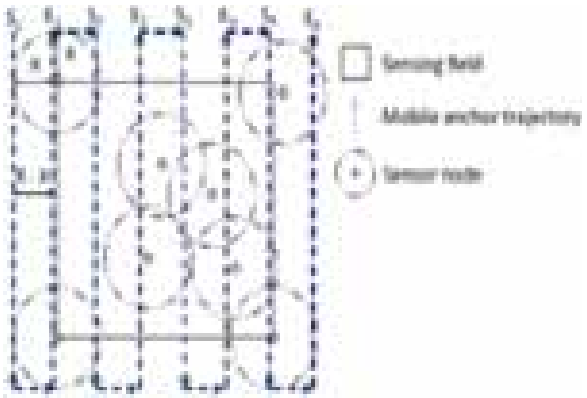


Fig. 3. Path planning algorithm.

3.5. Dynamic Square Trajectory Algorithm

A mobile anchor follows squares path in both the X and Y Direction such that the distance between the consecutive X and Y axis is $2R$, R is the communication radius of the Mobile Anchor Node. The total distance travelled by the mobile anchor covering the whole radius is $(Len/2R)-1$ where *Len* is the Length of a largest square and the Square size decreases by $2R$ is calculated using Eq. (3).

$$\text{Total distance} = 4[Len + (Len-2R) + (Len-4R) + \cdots + (Len-2(Len/2R-1)R)$$
$$\text{Distance } D = (Len/R+2)Len \tag{3}$$

Considering the presence of obstacles D* Lite Algorithm is applied in the square. The Vertexes of the Area to be localized are already added to the Queue. The Mobile Anchor node starts from the center position mentioned in the Figure 4 as A and move towards the next location added in the Queue. In case of obstacles encountered in the specified Path between L and M as per Figure 4, the D* Lite Algorithm dynamically changes the path and reaches the next specified (X, Y) coordinate mentioned in the Queue. The D* Square Travels in both X and Y direction, which reduces the error in both the X and Y direction.

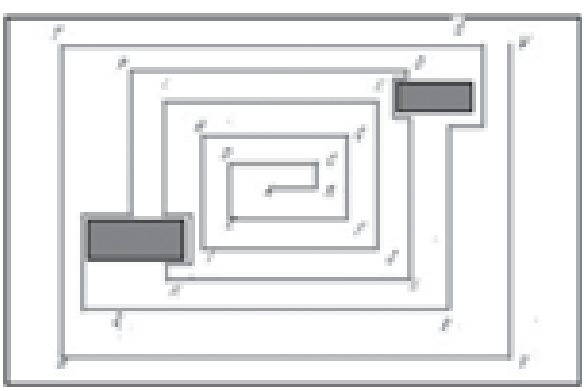


Fig. 4. Dynamic square path.

The D* Lite Algorithm Applied is given below,

```
procedure CalKey(s) return [minimum(g(s), rightside(s)) +
h(start, s) + mtrs;minimum(g(s), rightside(s))];
procedure Initialize( )
node = ∅; mtrs = 0;
for all s ∈ S rightside(s) = g(s) = ∞;
rightside(s_goal) = 0; U.Insert(s_goal, [h(start, s_goal); 0]);

procedure UpdateVertex(u)
if (g(u)! = rightside(u) AND u ∈ node) node.Update(u,
   CalKey(u));
else if (g(u) ! = rightside(u) AND u/∈ node)
  node.Insert(u,CalKey(u));
else if (g(u) = rightside(u) AND u ∈ node) node.
   Remove(u);

procedure ShortestPath( )
while (node.TopKey( ) < CalKey(start) OR rightside
  (start) > g(start))
u = node.Top( ); k_old = node.TopKey( ); k_new =
  CalKey(u);
 if(k_old < k_new) node.Update(u, k_new);
 else if (g(u) > rightside(u)) g(u) = rightside(u);
node.Remove(u);
for all s ∈ Pred(u)
if (s ! = s_goal) rightside(s) = minimum(rightside(s),
  c(s, u) + g(u));
UpdateVertex(s);
 else g_old = g(u);
 g(u) = ∞;
 for all s ∈ Pred(u) ∪ {u}
 if (rightside(s) = c(s, u) + g_old)
 if (s ! = s_goal) rightside(s) = minimum s′ ∈ Succ(s)
(c(s, s′) + g(s′));
 UpdateVertex(s);

procedure Main( )
 last = start; Initialize( ); ShortestPath( );
 while (start ! = s_goal)
/*if (g(start) = ∞) then there is no known path*/
 start = argminimum s′ ∈ Succ(start)(c(start, s′) + g(s′));
 Move to start;
 Scan graph for changed edge costs;
 if any edge costs changed
 mtrs = mtrs + h(last, start);
 last = start;
 for all directed edges (u, v) with changed edge costs
 c_old = c(u, v);
 Update the edge cost c(u, v);
 if (c_old > c(u, v))
 if (u ! = s_goal) rightside(u) = minimum(rightside(u),
c(u, v) + g(v));
 else if (rightside(u) = c_old + g(v))
```

if (u ! = s_goal) rightside(u) = minimum s′ ∈ Succ(u) (c(u, s′) + g(s′));
UpdateVertex(u);
ShortestPath( )

4. PERFORMANCE EVALUATION

The Performance metrics of Dynamic Path planning based localization technique is implemented on Network Simulator (version 2) widely known as NS2 and its metrics are measured in terms of error estimation, % of Nodes localized with respect to time and No. of Nodes.

4.1. Simulation Model

The values of the parameters used for simulation are as shown in Table I are simulated using NS2.[16]

(i) *Error estimation analysis*

The difference between the actual location of a sensor node and the Calculated location of all the nodes is known as estimation error. The average is taken considering the number of sensor nodes as 25, 50, 75, 100 and plotted. Referring to the Figure 5, the Error Estimation Analysis of D∗SQUARE is reduced to half of SCAN.

$$NLE = \frac{SQRT((X_{Act} - X_{Loca})^2 + (Y_{Act} - Y_{Loca})^2)}{\text{Area}} \quad (4)$$

(ii) *Time versus % of node localized*

Time taken to localize the sensor nodes is Calculated. Also the no. of nodes localized at fixed time interval is considered. Here the time interval taken for calculation is 20 secs. Since the algorithms changes its direction both the axis at equal intervals the algorithms gets localized quickly. Referring to the Figure 6, the time taken to localize sensor nodes using D* SQUARE is 15% more efficient and quicker than SCAN.

(iii) *Sensors nodes versus % of nodes localized*

With the increase in deployment of sensor nodes, the % of nodes localized is calculated. The D* Square algorithm best identifies the location of the node, that are situated near the boundary region. Figure 7 depicts better node localization in D* SQUARE Compared with SCAN.

Table I. Simulation parameters.

Parameters	Value
Topology size	1000 m × 1000 m
No. of sensor nodes	25, 50, 75,100
No. of anchor nodes	1
MAC layer	IEEE 802.11
Simulation Time	100 sec
Traffic source	Constant bit rate (CBR)
Node placement	Random waypoint
Packet size	512 bytes
Transmit power	360 mW
Transmission range	50 m
Routing protocol	Adhoc on-demand distance vector (AODV)
Initial energy	5.1 Joules
Speed	10 m/sec

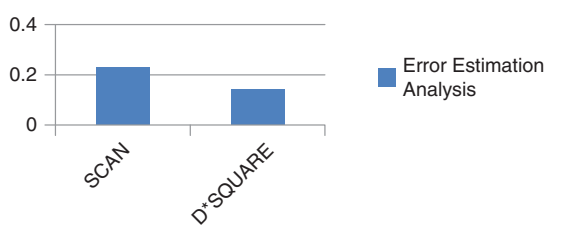


Fig. 5. Error estimation analysis.

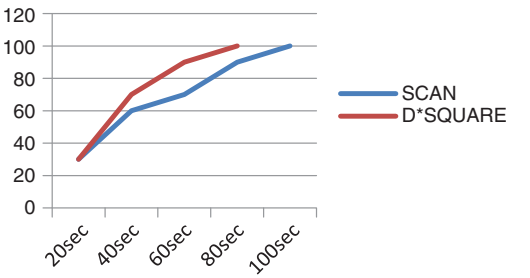


Fig. 6. Time versus % of nodes localized.

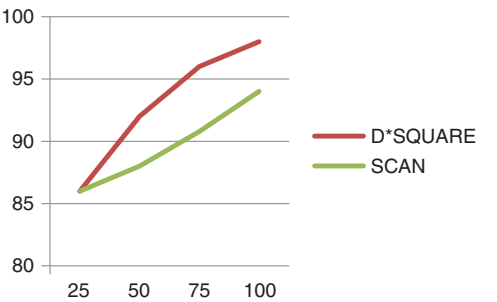


Fig. 7. Sensors nodes versus % of nodes localized.

5. CONCLUSION

In WSNs the localization is a emerging research as it is an unsolvable problem. By incorporating the Received Signal Strength Indicator Method to calculate the distance estimation, and D* square algorithm to find out the location of the sensor node by moving the sensor node in a predefined path, the Dynamic Path Planning algorithm works better and localizes the sensor node quickly with better accuracy. This work can be further enhanced by applying Reinforcement learning algorithms for the path planning procedures to decide how effectively to move the anchor node in a predefined path. By applying reinforcement algorithm, we dynamically change the direction of the mobile anchor based on the feedback of the Received Signal Strength.

References

1. Imane Benkhelifa and Samira Moussaoui, Appl: Anchor path planning based localization for wirelesssensor networks, *Mosharaka International Conference on Communications, Computers and Applications*, MIC-CCA2011.

2. S. Divya and P. Purusothaman, *International Journal of Computer Science and Network Security* 15, 71 **(2015)**.

3. H. Chenji and R. Stoleru, *IEEE Transactions on Mobile Computing* 12, 1094 **(2010)**.

4. J. Xiao, L. Ren, and J. Tan, Research of TDOA based self-localization approach in wireless sensor network, *Proceedings of the IEEE Int. Conf. on Intelligent Robots and Systems*, October **(2006)**, pp. 2035–2040.

5. P. Purusothaman and S. Sivakumar, *International Journal of Engineering Research and Technology* 2, 1323 **(2013)**.

6. V. Ramadurai and M. L. Sichitiu, Localization in wireless sensor-networks: A probabilistic approach, *Proc. of the 2003 International Conference on Wireless Networks (ICWN 2003)*, Las Vegas, NV, USA, June **(2003)**, pp. 300–305.

7. A. Galstyan, B. Krishnamachari, K. Lerman, and S. Pattem, Distributed online localization in sensor networks using a moving target, *Proc. of the International Symposium on Information Processing in Sensor Networks*, April **(2004)**.

8. Z. Zhong and T. He, Achieving range-free localization beyond connectivity, SenSys **(2009)**.

9. D. Koutsonikolas, S. M. Das, and Y. C. Hu, Path planning of mobile landmarks for localization in wireless sensor Networks, *Proc. of IEEE Distributed Computing Systems Workshops* **(2006)**, pp. 86–86.

10. G. V. Zaruba, M. Huber, and F. A. Karmangar, Monte carlo sampling based in-home location tracking with minimumimal RF infrastructure requirements, *Proc. of IEEE GLOBECOM'04* **(2004)**, Vol. 6, pp. 3624–3629.

11. L. Hu and D. Evans, Localization for mobile sensor networks, MobiCom **(2004)**.

12. W. H. Liao, Y. C. Lee, and S. P. Kedia, *IET Communications* **(2010)**.

13. S. Yun, J. Lee, W. Chung, and E. Kim, Centroid Localization Method in Wireless Sensor Networks Using TSK Fuzzy Modeling **(2007)**.

14. F. K. W. Chan, H. C. So, and W.-K. Ma, A novel subspace approach for wireless sensor network positioning with range measurements, *IEEE Conference on Acoustics, Speech and Signal Processing* **(2007)**, Vol. 2, pp. 1037–1040.

15. J. Z. Wang and H. Jin, Improvement on APIT localization algorithm for Wireless Sensor networks, *IEEE International Conference on Network Security, Wireless Communication and Trusted Computing* **(2011)**, pp. 190–195.

16. NS-2 Simulator, http://www.isi.edu/nsnam/ns/.

Received: 18 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1568–1572, 2018

RESEARCH ARTICLE

Privacy Preserving Public Auditing for Cloud Storage Using Elliptic Curve Digital Signature

K. S. Arvind[1,*], R. Manimegalai[2], and K. S. Suganya[3]

[1] *Research Scholar, Anna University, Chennai 600025, India*
[2] *Peelamedu Samanaidu Govindasamy College of Technology, Coimbatore 641004, India*
[3] *Bannari Amman Institute of Technology, Erode 638401, India*

Cloud Computing provides ubiquitous and infinite data services for cloud users. Among all the services, cloud data storage service has gained more importance in recent years. These services allow the cloud data owners (CDO) to store their data remotely and access them at their own convenience. With increase in popularity, there are also some concerns regarding their security and privacy of cloud owner's data. Data integrity is one of the important concerns regarding the cloud data storage services. For which, many public auditing schemes have been proposed over the years. These public auditing schemes employed nontraditional cryptographic approaches for ensuring the integrity and privacy preserving of the Cloud owner data. They stated that the traditional cryptographic approaches when employed for public auditing in Cloud storage will increase the computational overhead for the entities and resources. In our paper, we propose the use of traditional but lightweight cryptographic and hashing techniques to achieve the data integrity and confidentiality. Firstly the cloud data owners will employ the lightweight AES encryption algorithm for encrypting their data and storing it in the cloud server with Cloud Service Provider (CSP). Secondly CDO makes use of SHA-3 for generating the hash value on the encrypted data and signing it using Elliptic Curve Digital Signature. Third Party Auditors (TPA) on receiving the audit request form CDO will request the Cloud Service Provider (CSP) for the encrypted data. Finally, TPA will recompute the hash values and sign and compare them with the signature of CDO. This Public auditing protocol also employs SiteKey authentication for mutual authentication between CDO and TPA. The Security analysis and evaluation of the Proposed Public Auditing protocol ensures that the system is efficient and secure.

Keywords: Third Party Auditors (TPA), Cloud Data Owner (CDO), Cloud Service Provider (CSP), Public Auditing, Cloud Storage, Elliptic Curve Digital Signature, SiteKey Authentication.

1. INTRODUCTION

Cloud storage is one of the most important services because of the growing nature of data and technology. With the growth of eminent technologies like Ubiquitous Computing, Big data, 5 G and so on, the importance of cloud storage has reached its pinnacle. More and more users have started storing their data in various cloud applications like Google drive, Dropbox, iCloud, Microsoft SkyDrive and Amazon S3. But also there have been many breaches and outages relating to Google drive, Dropbox etc. in the recent years such as Gooligan attack, etc. According to recent reports from IBM Security services, the cloud data breach has increased by 12.3 percent in 2017.

There are many factors which contribute to the security and privacy concerns regarding Cloud storage. Firstly, the cloud users cannot entirely trust the cloud service providers. Secondly, the cloud service provider may hide the data loss events. Finally, there may be reason for data corruption due to cloud server failure and unavailability. Consequently, there are also concerns about the integrity of the cloud user's data.

Many data auditing schemes for cloud data storage have been proposed over the years. Reference [1] have proposed a scheme called Provable Data Possession (PDP) which allows the integrity verification without retrieving the original data using homomorphic verifiable tags. Reference [9] proposed the mutual verifiable Provable Data Possession which involved independent TPA. The proposed scheme used Elliptic Curve Cryptography for generating

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1568/005 doi:10.1166/jctn.2018.7341

homomorphic authenticators. (Jules et al., 2007) presented a scheme called Proof of Retrievability (POR) using symmetric-key cryptography, using pseudo random sampling technique and efficient error-coding. (Shacham and Waters, 2008) proposed the use of BLS Signatures which allows public verifiability. They also proposed a second scheme which builds on Pseudo Random Function and homomorphic authenticators that are constructed based on BLS signatures.

In order to prevent this, a Third Party Auditor (TPA) was involved for checking the integrity of the data and ensures their data correctness. This scheme is known as Public Auditing. This process of involving a Third Party Auditor reduced the workload of the cloud user and computational overhead and ensures the integrity of the data (Liu et al., 2014).[3–13] All these schemes have different aspects on cloud data auditing.

Reference [3] was one of the first few to achieve privacy preserving for Cloud Storage using Third Party Auditor. They introduced the concept of public key based homomorphic linear authenticator which enables the TPA to perform public auditing without retrieving the original data. This was made possible by the algebraic properties of the homomorphic linear authenticator. In spite of their resourcefulness, their work lacked in the data dynamics which allowed the data operations such as insertion, deletion, modification to be performed on the data. This scheme was also revised by Ref. [7] which included implementation on Amazon EC2 Public cloud. Reference [4] have achieved this data dynamics along privacy preserving using Merkle Hash Tree for block tag authentication. The scheme also supported multiple auditing simultaneously. Similarly, dynamic hash table was proposed by (Tien et al., 2015) which combined the use of the mentioned data structure for efficient public auditing and also ensures privacy preserving.

Furthermore, the use of ring signature to verify the correctness of the shared data was proposed by Ref. [5]. This method efficiently addressed the problem of preserving the identity privacy of Cloud users by utilizing ring signatures to construct homomorphic authenticators. This scheme efficiently addressed the integrity of the shared data and also allows multiple auditing to be performed simultaneously. But there is problem of proving the data freshness in the cloud Storage in the discussed ORUTA scheme.

There have been many concerns on the integrity of Third Party Auditor (TPA) which have been addressed by (Liu et al., 2014). This scheme involved additional authentication between the Cloud Users and TPA. The other important factor is the privacy of the identity of the Cloud user. This concern was also addressed by Ref. [8] using blinding technique. The scheme also addressed the problem of multiple audit requests by TPA. Reference [10] proposed a modification by using polynomial based authentication tags and proxy tag update techniques which allows multiple user to modify the data and involves batch integrity checking technique. Further, this scheme also allows user revocation to the cloud. Reference [12] proposed a new scheme for providing security for keys thereby ensuring key resilience in the privacy preserving public auditing.

In this paper, we overcome the many difficulties faced by the public auditing of cloud storage services. Firstly we make use of AES to encrypt the data before storing the data with Cloud Service Provider. This ensures that data cannot be tampered both by Cloud Service Provider and Third Party Auditor. Secondly, we propose the use of Secure Hash Algorithm 3 which will be used to generate hash values on the encrypted data at both Cloud data Owner and Third Part Auditor sides. We then recommend the use of Elliptic Curve Digital Signature (ECDSA) for signing the hash value which will be shared with the TPA. TPA on receiving the signature will store it. When the Cloud Data owner request the audit request of their encrypted data, the TPA sends the Auditing Challenge to the Cloud Service Provider. The Cloud Service Provider will in turn provides the encrypted data which the TPA uses to recomputed the hash value using SHA-3 and then sign it using ECDSA. Finally the TPA will verify the data correctness and integrity by comparing the signatures. This addresses all the threats which will be discussed in threat models in Section 2. And also since we use Public Key ECDSA for signing the data, we can ensure the identity privacy of the Cloud Data Owner from unauthorized user. This process will be detailed in the Section 3. Section 4 will address the security analysis of the proposed system and provide an insight on the security and privacy of the public auditing protocol in Cloud storage.

2. PROBLEM DEFINITION

In this section, we will discuss about the general entities involved in public auditing, threats faced by the entities and the design intents of this paper.

2.1. System Model

In common, there are three entities involved in the process of Public auditing namely Cloud Data Owner (CDO), Cloud Service Provider (CSP) and Third Party Auditor (TPA) as in Figure 1. The Cloud Data Owner (CDO) uploads their encrypted data in the cloud storage and their role involves requesting auditing delegation to the Third Party Auditor (TPA). Secondly, the Cloud Service Provider provides the requested encrypted data to the authorized Third Party Auditor to ensure the data correctness and data integrity of the encrypted data. Finally, the TPA has the following roles namely (1) They will authorize the audit request from CDO, (2) They will ensure mutual authentication between TPA and CDO, (3) They will request the audit data from the CSP and performs data integrity check

Fig. 1. Three entities involved in public auditing.

without decrypting the data. This process will be further explained in detail in Section 3.

2.2. Threat Models

As discussed above, there are three entities involved in public auditing namely Cloud Data Owner (CDO), Cloud Service Provider (CSP) and Third Party Auditor (TPA). There are possibilities for threats from Cloud Service Provider and Third Party Auditor to the Cloud Data Owner which is discussed below.

(A) *Cloud Service Provider*: The following threat scenarios are possible considering cloud storage by Cloud Service Provider.

(1) The data stored in the cloud storage can be modified by other users of the same cloud or external intruders with malicious intent.

(2) The cloud service provider may intentionally modify the data to save storage space.

(3) The cloud service provider may hide the data loss to the data owner due to system failure or other intrusion activities.

(B) *Third Party Auditor*: The following threat scenarios are possible considering Third Party Auditor.

(1) TPA is expected to a reliable and trustworthy entity. The malicious insiders from TPA should not be able to access the data and also the identity of the data owner.

(2) TPA must also ensure their integrity by mutually authenticating themselves with the Cloud Data owner and Cloud Service Provider to prevent masquerade attack.

2.3. Design Intents

The public auditing for Cloud storage requires the system meets the following design intents.

1. *Public Verifiability*: The Third Party Auditor should be able to make an auditing request to the Cloud Service Provider without any hassle and they need not download the entire data form the Cloud service provider resulting in fast auditing.

2. *Data Correctness*: The data stored in the Cloud Service Provider need to be intact and if there is any slight modification then the auditing proof should be rejected.

3. *Privacy Preserving*: The data should not be leaked both by CSP and TPA.

4. *Multiple Auditing*: The TPA should be able to handle multiple audit requests simultaneously without any data leakage.

5. *Identity Privacy*: The identity of the CDO should not be revealed explicitly to the TPA and external hackers.

3. PROPOSED SYSTEMS

There is ongoing demand for the proper public auditing of cloud storage. The process of appointing a third party auditor to check the integrity of the data in cloud storage by the cloud data owner is called Public Auditing. In our proposed system, we make use of following security techniques namely Advanced Encryption Standard (AES) for encrypting the data, Secure Hash Algorithm (SHA-3) for generating the hash value, Elliptic Curve Digital Signature for Public Verifiability and SiteKey authentication for mutual verification which is explained in detail in the Figure 2. The process of our public auditing has been divided into following phases namely

(1) Setup Phase: The data owner first splits the files into blocks and then encrypts the data using a lightweight Advanced Encryption Standard (AES) proposed by Ref. [14] which makes use optimal Substitution boxes using Composite fields and the data is stored in the Cloud storage.

(2) Hashing Phase: The data owner then employs the Secure Hash Algorithm-3 and generates the hash value for each block which is then concatenated to form a single hash value using XOR operation.

(3) Signature Generation Phase: The data owner then signs the concatenated hash value using the ECDSA Signature generation as in Figure 3. Then the Signature is sent to the Third Party Auditor after mutually authenticating themselves using SITEKEY authentication as in Ref. [16]. The Third Party Auditor then stores the signature with the corresponding public key for future computations.

(4) Signature Verification Phase: On receiving an audit request from CDO, then TPA request the encrypted data from CSP. With the received block of data, hashing is performed using (SHA-3) to get the hash value which is concatenated and signed using the same parameters involved in Figure 3. Then the stored signature and computed signature is compared and if found same, then the data is unmodified and integrity of the data is ensured. The same should be reported to the Cloud Data Owner.

To achieve identity privacy and also ensuring the mutual authenticity between the CDO and TPA, we make use of SiteKey Visual Cryptography which makes use of image to authenticate the Cloud Data Owner and Third Party Auditor. To achieve mutual authentication, the original image

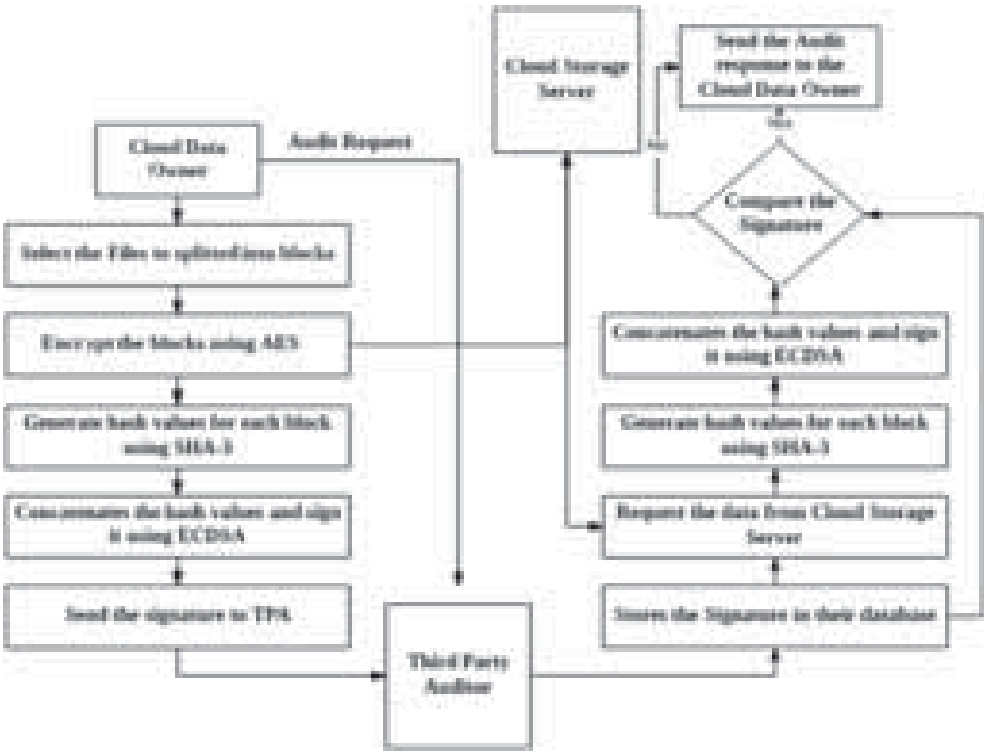


Fig. 2. Flow diagram of the proposed public auditing system.

is divided into two image share namely Share CDO and Share TPA which is available with CDO and TPA respectively. Upon login, the CDO will provide their image Share CDO which will be overlaid with the Share TPA of Third Party Auditor which will reveal the original image.

4. SECURITY ANALYSIS AND EVALUATION

The security aspects of the proposed public auditing method discussed in the previous section are studied in this section.

ECDSA Key Generation:	ECDSA Signature Generation:	ECDSA Signature Verification:
The entity Cloud Data Owners (CDO) key pair is related with the following set of Elliptic curve domain parameters $CDO_D = (p, FR, a, b, G, n, h)$. E is an elliptic curve defined over GF_p, and Q is a point of prime order n in $E(GF_p)$, p is a prime. Then the CDO does the following: 1. Randomly select an integer x in the interval $[1, n-1]$. 2. Compute $x = xQ$. 3. CDO's public key is x, CDO's private key is d.	To sign a message m, the entity Cloud Data Owners (CDO) with domain parameters $CDO_D = (q, FR, a, b, G, n, h)$ does the following: 1. Select a random or pseudo-random integer k in the interval $[1, n-1]$. 2. Compute $kP = x_1, y_1$ and $r = x_1 \bmod n$ (where x_1 is regarded as an integer between 0 and $q-1$). If $r = 0$ then go back to step 1. 3. Compute $k^{-1} \bmod n$. 4. Compute $s = k^{-1} \{h(m) + dr\} \bmod n$, where h is the Secure Hash Algorithm (SHA-3). If $s = 0$, then go back to step 1. 5. The signature for the message m is the pair of integers (r, s).	To verify Cloud Data Owner's signature (r, s) on m, Third Party Auditor (TPA) obtains an authenticated copy of CDO's domain parameters $D = (q, FR, a, b, G, n, h)$ and public key x and does the following: 1. Verify that r and s are integers in the interval $[1, n-1]$. 2. Compute $w = s^{-1} \bmod n$ and $h(m)$. 3. Compute $u1 = h(m)w \bmod n$ and $u2 = rw \bmod n$. 4. Compute $u1P + u2Q = (x0, y0)$ and $v = x0 \bmod n$. 5. Accept the signature if and only if $v = r$.

Fig. 3. Elliptic curve digital signature algorithm.

4.1. Integrity

The proposed system recommended the use of Secure Hash Algorithm-3 for generating hash value which does not involve the use of public key or homomorphic tags and thereby increases the performance of the system and the SHA-3 as it involves optimal asymmetric encryption padding. And also, since their announcements by NIST, there are no known attacks against SHA-3. And computation overhead of SHA is low when compared with MAC and homomorphic authenticators.

4.2. Confidentiality

The cloud data owner used the AES-128 bit encryption algorithm to encrypt the data before storing it in the cloud storage. We have employed a light weight AES algorithm which has optimal S boxes and composite fields ensuring the improved performance so that this system can also be opted for mobile cloud.

4.3. Authentication

To address the concern of identity privacy and ensure that the entities are trustworthy in public auditing, the proposed system makes use of Visual Cryptography for SiteKey authentication in addition to the regular user identification and passphrase. It is difficult for the hacker to regenerate the image share of TPA and masquerade as Third Party Auditor.

4.4. Batch Auditing

Since the TPA is an independent entity, it can be employed by many CDO at the same time for making audit request. Also the proposed system makes use of lightweight SHA-3 and ECDSA implementation which will reduce the computation cost, thereby relieving the unnecessary resources of TPA. Thus the TPA can handle multiple audit request at the same time without facing any functionality overhead.

5. CONCLUSION

In spite of employing nontraditional cryptographic approaches for public auditing of cloud storage, they lag in one or more design objectives like privacy preserving, data confidentiality and data correctness. Our proposed system makes use of traditional cryptographic approaches like Symmetric Key encryption such as AES, Hashing technique like SHA-3, Public Key Digital Signature Algorithm such as Elliptic Curve Digital Signature Algorithm for providing Confidentiality, integrity and identity privacy respectively. To ensure mutual authentication, we have proposed the use of Visual Cryptography algorithm for SiteKey authentication. In future, we will be looking in the data dynamics of the data in cloud storage and implementing the same in the OPENSTACK cloud.

References

1. G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, L. Kissner, Z. N. J. Peterson, and D. X. Song, Provable data possession at untrusted stores, *Proceedings of ACM Conference on Computer and Communications Security 2007* **(2007)**, pp. 598–609.
2. G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. N. J. Peterson, and D. Song, *ACM Trans. Inf. Syst. Secur.* 14, 1 **(2011)**.
3. C. Wang, Q. Wang, K. Ren, and W. Lou, Privacy-preserving public auditing for data storage security in cloud computing, *Proceedings of the IEEE INFOCOM 2010* **(2010)**, pp. 525–533.
4. Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, *IEEE Trans. Parallel Distrib. Syst.* 22, 847 **(2011)**.
5. B. Wang, B. Li, and H. Li, Oruta: Privacy-preserving public auditing for shared data in the cloud, *Proceedings of the IEEE Cloud 2012* June **(2012)**, pp. 295–302.
6. K. Ren, C. Wang, and Q. Wang, *IEEE Internet Comput.* 16, 69 **(2012)**.
7. C. Wang, S. Chow, Q. Wang, K. Ren, and W. Lou, *IEEE Trans. Comput.* 62, 362 **(2013)**.
8. S. G. Worku, C. Xu, J. Zhao, and X. He, *Comput. Electr. Eng.* 40, 1703 **(2014)**.
9. Y. Ren, J. Shen, J. Wang, J. Han, and S. Lee, *J. Internet Technol.* 16, 317 **(2015)**.
10. J. Yuan and S. Yu, Efficient public integrity checking for cloud data sharing with multi-user modification, *Proceedings of the IEEE INFOCOM 2014*, April–May **(2014)**, pp. 2121–2129.
11. H. Tian, Y. Chen, C. Chang, H. Jiang, Y. Huang, Y. Chen, and J. Liu, *IEEE Trans. Serv. Comput.* 1 **(2015)**.
12. J. Yu, K. Ren, C. Wang, and V. Varadharajan, *IEEE Trans. Inf. Forensics Secur.* 10, 1167 **(2015)**.
13. G. Yang, J. Yu, W. Shen, Q. Su, F. Zhang, and R. Hao, *J. Syst. Softw.* 113, 130 **(2016)**.
14. Mehran Mozaffari-Kermani and Arash Reyhani-Masoleh, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 19, 85 **(2011)**.
15. M. Rao, T. Newe, and I. Grout, Secure hash algorithm-3 (SHA-3) implementation on Xilinx FPGAs, Suitable for IoT applications, *Proceedings of the 8th International Conference on Sensing Technology* **(2014)**, pp. 352–357.
16. K. S. Arvind and R. Manimegalai, *International Journal of Applied Engineering Research* 143 **(2015)**.

Received: 19 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1573–1576, 2018

XML Keyword Search Using Breadth-First Search

R. Vinothsaravanan* and C. Palanisamy

Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamil Nadu, India

Keyword search is the one of the most reliable way of information retrieval over XML data. Most widely used XML query processing techniques are Lowest Common Ancestor (LCA), Smallest LCA (SLCA) and Exclusive LCA (ELCA). If the semantic constraints are applied to LCA, SLCA or ELCA, sometimes it returns NULL or missing elements and undesired results due to inefficiency of existing methods are the common-ancestor-repetition (CAR) and visiting-useless-nodes (VUN) problems. This paper proposes a generic Breadth-First Search (BFS) strategy to answer a given keyword query semantics to overcome the CAR problem.

Keywords: XML, Keyword Search, BFS.

RESEARCH ARTICLE

1. INTRODUCTION

Searching XML data can be effectively attained through query language such as XPath and XQuery, if the schema information is available. Since the heterogeneity and the hierarchical structure of XML documents, extracting information from the XML data often becomes intractable.

Consider the following fragment of Research XML database,

```
<Researcher>
    <SIG Head> Jack </SIG Head>
    <Research Field> IOT </Research Field>
    <Person>
        <Name> Jack </Name>
        <Conference>
            <Author> Jack </Author>
            <Title>
            ...IOT...internet...cloud
            ...
            </Title>
            <Abstract>
            ...IOT...internet...cloud
            ...
            </Abstract>
        </Conference>
        <Journal>
            <Author> Jack </Author>
            <Title>
            ...IOT...internet...cloud
            ...
            </Title>
            <Abstract>
            ...IOT...internet...cloud
            ...
            </Abstract>
        </Journal >
    </Person>
    <Person>
        <Name> Tim </Name>
        <Book>
            <Author> Tim, Jack
            </Author>
            <Title>...IOT...
            </Title>
            <Abstract>
            ...IOT...
            </Abstract>
            <Ref> Tim, IOT </Ref>
        </Book>
</Researcher>
```

This XML document can be modeled as labeled tree as in Figure 1. Assume that user execute two queries in the XML document, $Q1 = \{\text{Jack, IOT}\}$ and $Q2 = \{\text{Tim, IOT}\}$. To execute the query two primitive operations are implemented to get desired results. Operation 1 test the document order of two nodes and operation 2 compute the Common Ancestor of two nodes. For a given keyword query $Q1, \ldots, Qn$, numerous semantics have been defined to get desired output, among this primitive semantics are Lowest Common Ancestor (LCA), Subset of Lowest Common Ancestor (SLCA) and Exclusive Lowest Common Ancestor (ELCA). Certainly, a system maintaining many query semantics will help users to find

*Author to whom correspondence should be addressed.

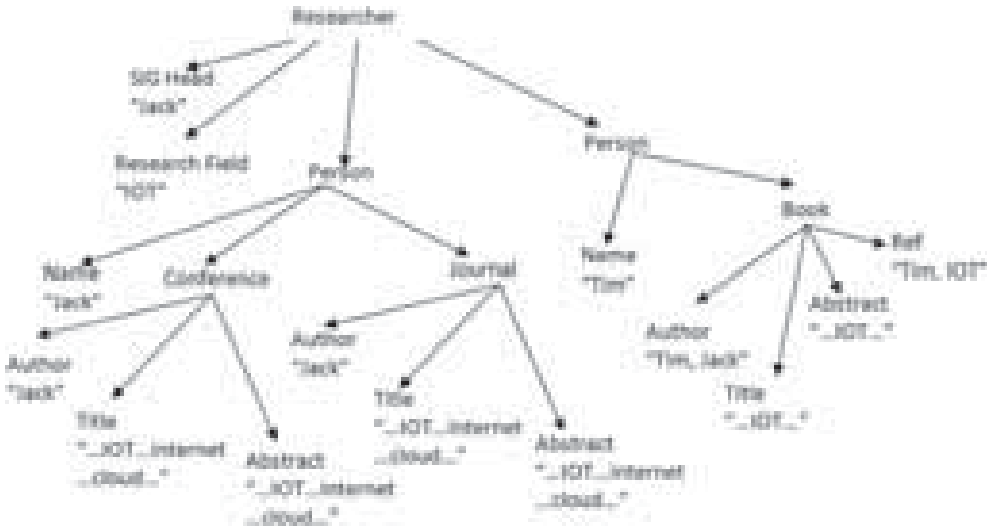


Fig. 1. A labeled tree of XML documents.

more valuable results, as some query semantics will not work on all the situations. Existing method to compute LCA/SLCA/ELCA are still ineffective due to redundant computation.

In some existing system LCA based semantics like SLCA and ELCA are represented and computed with Dewey labels/ID as primitive operating unit and visit the Common Ancestor (CA) nodes in top-down approach i.e., depth-First left-to-right order as in Figure 2. Either one of the primitive operations visit every CAs which are involved once. Two common problems arise in the LCA based computations are Common-Ancestor-Repetition (CAR) and Visiting Useless Nodes (VUN). The CAR problem is based on primitive operation 1 and 2, which takes every Dewey Label as the primitive operating unit.

A labeled tree node is classified in to Common Ancestor (CAs), useless nodes (UNs) and Auxiliary nodes (ANs). Common Ancestor (CAs), every keyword of the query contains in its subtree, such as node 1 in Figure 2 for query $Q2$. UNs, which is not a child of any CAs, for $Q2$, 6, 8, 9, 10, 12, 13, 15, 16, 17, 18, 19 and 20th nodes in Figure 2. ANs, are a child node of few CAs. For $Q2$, ANs are 3, 4 and 14th nodes in Figure 2.

To address the CAR problem, there are some strategies like generic Top-down XML keyword query processing,

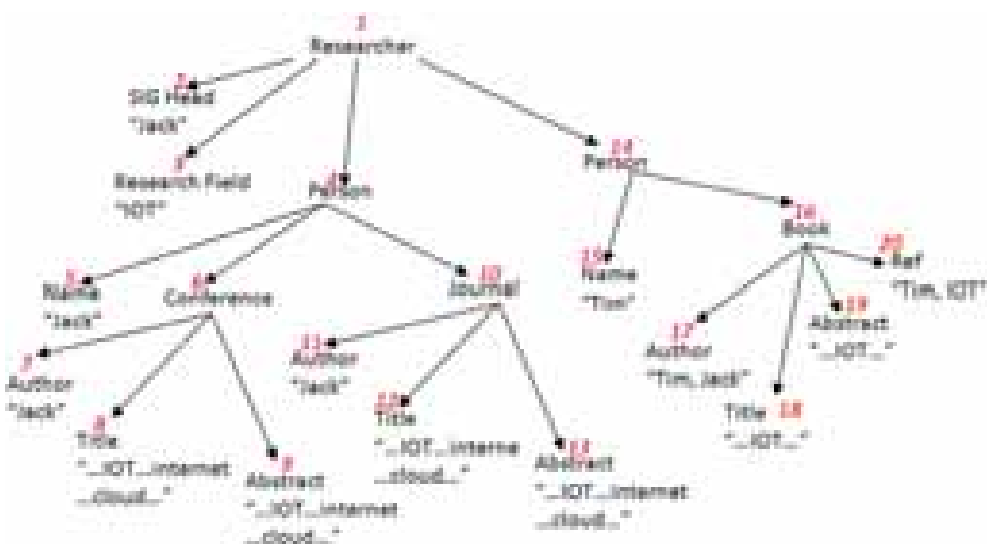


Fig. 2. A Dewey labeled tree of XML documents.

FWDELCA and BWDELCA. By traditional inverted list and labeling scheme independent, the VUN problems are addressed. This paper propose XML keyword search using Breath-First Search to address the CAR problem.

2. RELATED WORK

2.1. Data Model

XML documents are modeled as a labeled order tree as in Figure 2, where nodes represent attributes or elements, while edges represent direct nesting relationship between nodes. Every text data is taken as the text value of the attribute or element node. In Figure 2, "Jack" is the text value of node 5. Node v contains keyword k or v is a keyword node. Each node v is represented with its ID, based on its DeweyID, label can be concatenation of all IDs on the path from root node to v. e.g., the DeweyID label of node 5 is 1, 4, 5.

2.2. Query Semantics

Given a query $Q = \{k1, k2, \ldots, kn\}$ and XML document XD, the set of CA nodes of Q on XD is $\mathrm{CA}(Q) = \{v \mid v \supset$ keyword of Q at least once$\}$. e.g., for Query $Q1$ and XD, CA nodes are 1, 4, 6, 10, 14 and 16th nodes. Given a query $Q = \{k1, k2, \ldots, kn\}$ and XML document XD, the set of LCA nodes of Q on XD is $\mathrm{LCA}(Q) = \{v \mid v = \mathrm{LCA}(v1, v2, \ldots, vn), vi \in \mathrm{Li}\ (1 \le i \le n)$. LCA nodes of Query 1 on XD are 1, 4, 6, 10 and 16th nodes. The set of SLCAs of Q on XD is $\mathrm{SLCA}(Q) = \{v \mid v \in \mathrm{LCA}(Q)$ and there is no u that is not belongs to $\mathrm{LCA}(Q)\}$. SLCA nodes of $Q1$ on XD are 1, 6, 10 and 16th nodes. The set of ELCAs of Q on XD is $\mathrm{ELCA}(Q) = \{v \mid \exists v1 \in L1, \ldots, vn \in Ln(v = \mathrm{LCA}(v1, \ldots, vn)\text{^}\forall I \in [1, n]\}$. ELCA nodes of $Q1$ on XD are 1, 6, 10 and 16th nodes.

2.2.1. Algorithms

JDewey-E computes ELCA by performing set intersection operation on all lists of each tree depth from the leaf to the root. After finding the set of common nodes, it recursively delete all ancestor nodes in higher levels. JDewey-E suffers from the CAR and VUN problem. FwdELCA and BwdELCA implement inverted lists of node IDs, i.e., IDList, as the basis of ELCA computation. These algorithm avoid CAR problem but it suffer from VUN problem. A top-down XML query processing algorithms avoid CAR and VUN problem but the space complexity will be more when the depth of the XML document is high.

3. PROPOSED WORK

XML documents are modeled as labeled tree, where labels are in Breadth-First search (BFS) order as in Figure 3. If the labels are in BFS order, finding the CAs and LCA based semantics are easy. A labeled tree node is classified in to Common Ancestor (CAs), useless nodes (UNs) and Auxiliary nodes (ANs). Common Ancestor (CAs), every keyword of the query contains in its subtree, such as node 1 in Figure 3 for query $Q2$. UNs, which is not a child of any CAs, for $Q2$, 7, 8, 9, 10, 12, 13, 15, 16, 17, 18, 19 and 20th nodes in Figure 3. ANs, are a child node of few CAs. For $Q2$, ANs are 3, 4 and 5th nodes in Figure 3.

Next find the Inverted DeweyID label list. Inverted DeweyID label list of "Jack" is

1	2	3	4	5
1	1	1	1	1
2	4	4	4	5
	6	7	8	10
		11	14	17

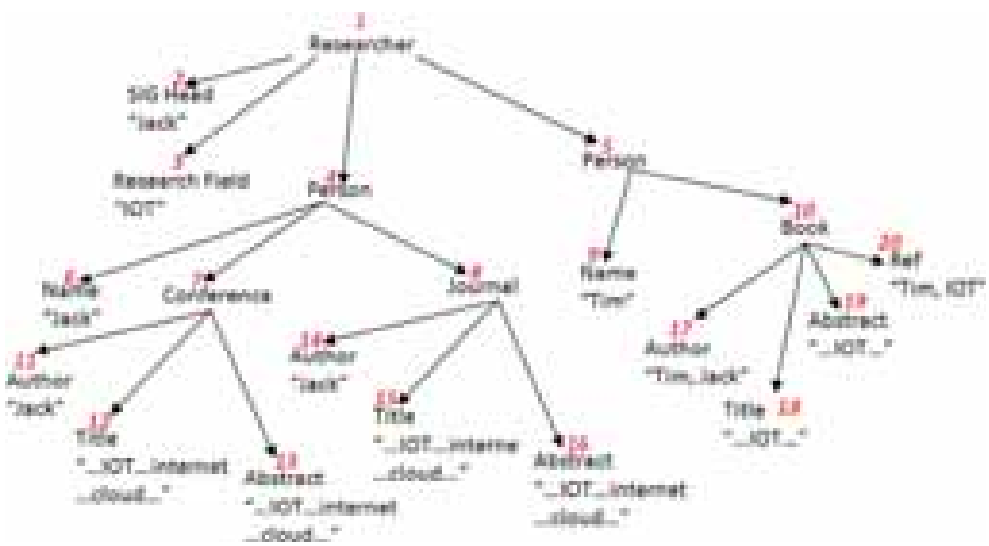


Fig. 3. A Dewey labeled tree of XML documents using BFS order.

Inverted DeweyID label list of "IOT" is

1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1
3	4	4	4	4	5	5	5
	7	7	8	8	10	10	10
	12	13	15	16	18	19	20

By using this inverted DeweyID labels, system can track the CAs and LCA based semantics. And also finding CARs and classifying UNs and ANs are done using inverted DeweyID. For the keyword "Jack," tracking of keyword search is.

Keyword node	Path
2	1
6	$4 \rightarrow 1$
11	$7 \rightarrow 4 \rightarrow 1$
14	$8 \rightarrow 4 \rightarrow 1$
17	$10 \rightarrow 5 \rightarrow 1$

Each and every keyword, system will give detailed path information. Using this path, identifying CAs and CARs can be done. ANs of the keyword "Jack" are 4 and 5th nodes. UNs of the keyword "Jack" are 7, 8 and 10th node. After classifying UNs and ANs of every keyword, Visiting-Useless-Nodes (VUN) can be avoided.

3.1. BFSCA-Algorithm

BFSCA($Q = \{k1, k2, k3, \ldots, kn\}$)

1. Initialize v as root node and points first DeweyID labels.
2. Label the nodes based upon the Breadth-First search order.
3. If all nodes are processed then return the XML documents with its label.
4. Find the inverted DeweyID of every keyword present in the XML document XD.
5. Identify the CAs and LCA based semantics based upon inverted DeweyID labels.
6. Classify the ANs and UNs from DeweyID labels.

This Algorithm recursively gets all CA nodes in a Breath-First order. For each CA node v, it finds out the number of occurrences of each query keyword in its subtree. Using the binary search operation, child CA nodes by intersecting v's child lists can be identified. As each CA node is visited only once, our method does not suffer from the CAR problem.

4. CONCLUSION

Most of the Existing XML keyword search algorithms are inefficient, while solving CAR and VUN problem. Some of the algorithm like generic top-down query processing strategy that visits all CA nodes only once and thus avoids CAR problem and VUN problem. Proposed algorithm also visits all CA nodes only once. So it avoids CAR problem and VUN problem. In this proposed algorithm, as the level of XML document is increased, the space complexity also increased. If the algorithm is applied with hash index or B-Tree index, the performance will be same as generic top-down query processing algorithm.

References

1. J. Zhou, W. Wang, Z. Chen, J. X. Yu, X. Tang, Y. Lu, and Y. Li, *IEEE Transactions on Knowledge and Data Engineering* 28, 1340 **(2016)**.
2. T. N. Le and T. W. Ling, *SIGMOD Record* 45, 17 **(2016)**.
3. J. Li, C. Liu, and J. X. Yu, *IEEE Trans. Knowl. Data Eng.* 27, 660 **(2015)**.
4. J. Zhou, Z. Bao, W. Wang, J. Zhao, and X. Meng, *Int. J. Very Large Data Bases* 23, 25 **(2014)**.
5. Z. Liu and Y. Chen, *J. World Wide Web* 14, 671 **(2011)**.
6. J. Zhou, Z. Bao, W. Wang, T. W. Ling, Z. Chen, X. Lin, and J. Guo, Fast SLCA and ELCA computation for XML keyword queries based on set intersection, *Proc. 28th Int. Conf. Data Eng.* **(2012)**, pp. 905–916.
7. J. Li, C. Liu, R. Zhou, and W. Wang, Suggestion of promising result types for XML keyword search, EDBT **(2010)**.
8. Z. Liu and Y. Chen, *J. Proc. Very Large Data Bases Endowment* 1, 921 **(2008)**.
9. Y. Xu and Y. Papakonstantinou, Efficient LCA based keyword search in XML data, *Proc. 11th Int. Conf. Extending Database Techn.: Adv. Database Technol.* **(2008)**, pp. 535–546.

Received: 19 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1577–1581, 2018

Dispersion and Modal Characteristics Analysis of 2D Rectangular Plasmonic Waveguide Structure

N. Hemadevi* and A. Arunya Revathi

Electrical and Electronics Engineering, Alagappa Chettiar Government College of Engineering and Technology, Karaikudi 630003, TamilNadu, India

A new investigation of rectangular plasmonic waveguide and its properties of the modes that are supported by 2-D structure. The Dispersion characteristic analysis of the rectangular waveguide is identified by changing the cladding material and also changing the geometrical profile of the waveguide. Propagation loss is reduced due to the presence of the air space in the core region of the waveguide. Dispersion analysis is done for every mode of the transmission and the effect of varying the shape parameters on the cutoff in the mode dispersion is studied. The closed profile of the metallic waveguides can provide the ultra-compact mode confinement. The Trapezoidal shape waveguide that causes the most significant shift in the cutoff region. The specific modal size, the fundamental mode of the standard rectangular plasmonic waveguide has a large propagation length compared with the corresponding modes of these plasmonic waveguides. This design could progress and assist potential for the enlargement of integrated photonic circuits.

Keywords: Plasmonic Waveguides, Metallic Waveguide, Extra Ordinary Transmission.

RESEARCH ARTICLE

1. INTRODUCTION

One of the most recent activities of modern optics is associated with the handling of light at the nanoscale. An optical system with plasmonic waveguides can allow scale down optical components through excitation of surface plasmon polariton or plasmon modes. Plasmonic Polariton is the study of the interaction between electromagnetic field, free electron in a metal and it is the energy distribution of the electromagnetic field. Many applications are developed by plasmonics.[1] These developments are recognised by the nanofabrication techniques.[2] Metal-dielectric structures display a varied range of novel physical phenomena which can be hired for numerous applications such as sensing, imaging and waveguiding. Plasmonic optical elements afford suitable potentials to incredulous the diffraction limit and they permit squeezing light to subwavelength[3] proportions. These are permitted miniaturizing devices to surge their functionalities in signal processing. The study of plasmonic waveguide is of a vital reputation for the design of nanoscale assemblies and circuits besides their feature solicitations. Optical waveguides guiding surface plasmon polaritons have been a subject of widespread studies and dissimilar types of metal-dielectric waveguide have been put forward theoretically and established experimentally. The simplest plasmonic waveguide is an interface between metal and insulator which ropes plasmonic polaritons.[4]

By providing the basis for sub-wavelength photonic devices, plasmonic waveguides could help diminish the size gap between electronic system. Application include in the sensing and EOT (Extra Ordinary Transmission) portent. Numerous types of plasmonic waveguide structures have been projected, including MIM (Metal-Insulator-Metal) plasmonic waveguide, groove, wedge, nanowires. IMI (Insulator-Metal-Insulator) and MIM plasmonic waveguide has a great interest because of the absolute ease of the trading process for such waveguide.[5] A meticulous mathematical study of MIM and IMI waveguide guided modes with symmetric and asymmetric cases where upper and lower cladding has different refractive indices.[6]

Long-range propagations and high spatial confinement[7] features of waveguide is done by the analysis of Ag–SiO_2–Ag MIM waveguide guides modes dispersion (see Fig. 1). Slight search has been conducted in the fiction to categorize the guided modes of the rectangular plasmonic waveguide with ample details and inclusiveness. Utmost studies for this waveguide type have been limited to a certain effect or a singularity at the precise wavelength band. Before that a rectangular plasmonic waveguide can reduced better abilities in contrast to MIM because, although all presentation of MIM waveguide

*Author to whom correspondence should be addressed.

Fig. 1. Schematic diagram of a 3-D MIM plasmonic waveguide.

exists for a single polarization but the rectangular waveguide can bid such proficiencies for both polarizations.

Dispersion faces and its applications are studied through numerical analysis of the presence of an x–y polarized guided modes in the rectangular waveguide. Sensitivity of rectangular waveguide is higher than MIM.[8] Energy transmission and sensing application have been recognized. Recognition of the dispersion curve for the modes in the different region can be oppressed in altered applications, such as sensing, enhanced transmission, slow and fast light and negative index of refraction solicitations.[9]

The objective is to perform a modal analysis and dispersion relation for the rectangular plasmonic waveguide. They are done by identifying all key guided modes. Due to the dispersion analysis, it is easy to find the dispersion curve factor for each key mode by varying the material profile. Dispersion curve can be subjugated in different application. Profiles changes from rectangular to one of imaginable trapezoidal plasmonic waveguide profiles. The study of the consequence varying the quantifiable on the key mode dispersion.

2. PLASMONIC WAVEGUIDE

2.1. Guided Mode Plasmonic Waveguide

The dispersion relation for the plasmonic waveguide is derived from the design of MIM waveguide. They have the pair of equations.

$$k_d a = \frac{-k_m \varepsilon_d}{k_d \varepsilon_m} \tag{1}$$

$$k_d a = \frac{-k_d \varepsilon_m}{k_m \varepsilon_d} \tag{2}$$

(or)

$$n_{\text{eff}} = \frac{k_{\text{spp}}}{k_0}$$

Where k_d and k_m defines the dielectric material components. ε_m and ε_d are the dielectric function (or) constants of metal and insulator. k_{spp} defines wave vector of surface plasmonic k_0 defines wave vector of incident light and n_{eff} defines the effective mode index.

The above Eqs. (1) and (2) is clear that direct change from propagation electromagnetic waves into surface plasmonic is impossible due to the mismatch of their wave vector. These are many ways to solve this problem. Such as Kretschmann/otto configuration.

These are two parity modes in the direction of propagation. They are even mode and odd mode. The electric field and magnetic field can be propagated through the medium by the direction when odd mode parity can be created by the Eq. (1) (i.e.,) odd mode means when Electric field of Z-direction is odd (Ez) whereas the other direction of x and y (Ex and Ey) are even on other hand Ex is even.

3. RESULTS AND DISCUSSION

3.1. Dimensional Analysis

Geometric structure for a rectangular plasmonic waveguide have the Metallic and dielectric regions. The dimensions used for the dielectric region is 250×120 (i.e.,) have a width of 250 nm and a height of 120 nm. A periodic boundary conditions are employed for the study of EOT phenomena. Dielectric center to the adjacent dielectric center with a hole array of width 470 nm and height 280 nm in the x–y directions.

In this paper, finite element method (Comsol) is used for analysis of numerical method of a rectangular plasmonic waveguide. This finite Element Analysis based on the full vectorial wave equations mainly for the high index difference in the waveguide and consideration of polarization coupling being simulated.

The key guided modes in a rectangular plasmonic waveguide has been investigated in the dimension of 250×120 nm, where metallic and dielectric region can be chosen as silver and air (see Fig. 2). In palik's experiment value for the permittivity of silver have been used. There are four guided modes of lowest order have been identified using comsol. They are seem resemble the odd and even parity components of the vertical and horizontal boundaries. Guided modes refered the combination of vertical,

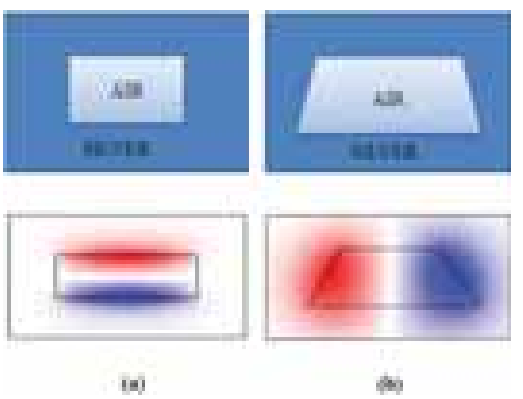


Fig. 2. (a) Schematic diagram of the studied rectangular plasmonic waveguide and electric field distribution of fundamental plasmonic mode. (b) Schematic diagram of the studied trapezoidal plasmonic waveguide and electric field distribution of fundamental plasmonic mode.

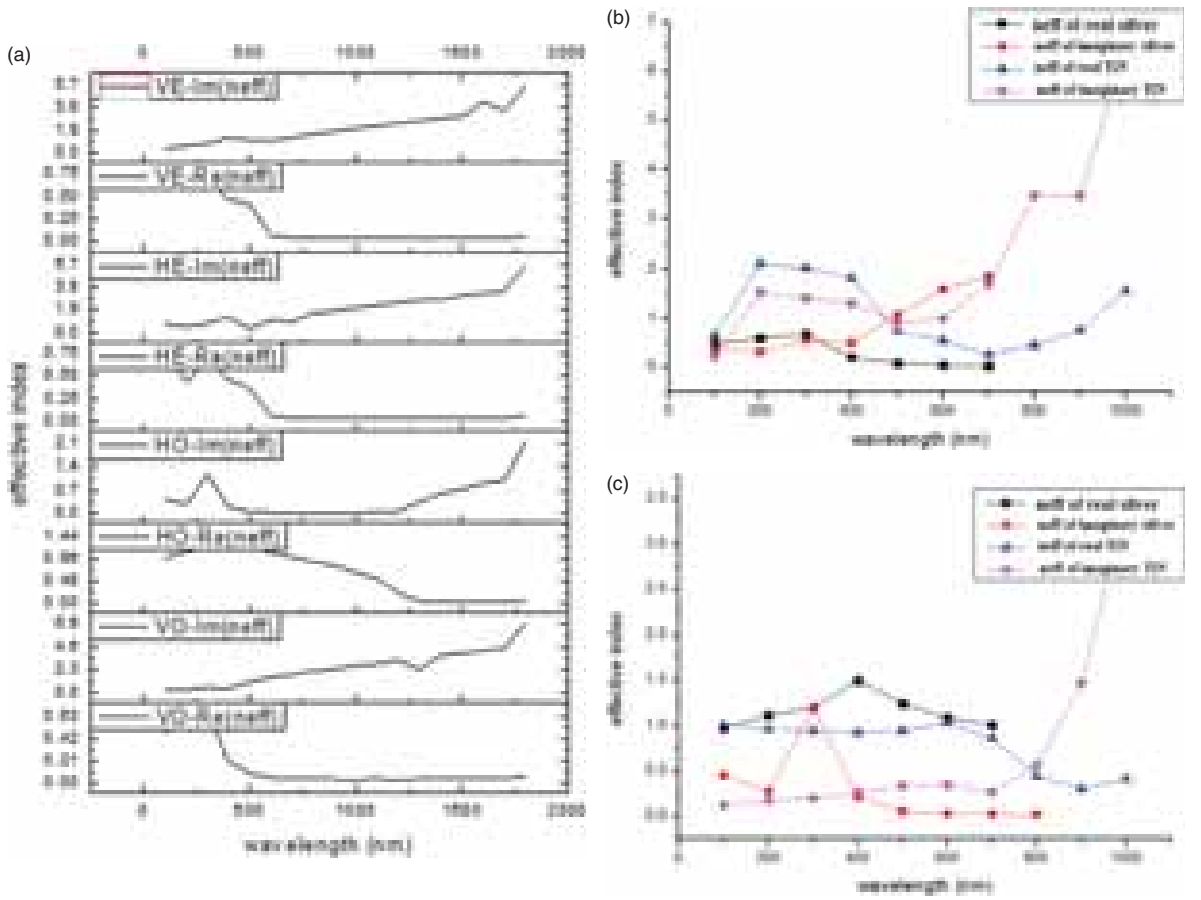


Fig. 3. (a) Curves for the four guided modes in a rectangular waveguide. Both real and imaginary part of the modes are represented. (b) Assessment between the relative of VO modes in silver versus TiN cladding. They directs VO peaks and cutoff suffer from red shift in the event of TiN when associated to silver. (c) Comparison between the relative of HO modes in silver versus TiN cladding. They directs HO peaks and cutoff suffer from red shift in the event of TiN when associated to silver.

horizontal, odd and even (see Fig. 3). This investigation of guided mode profile is mainly for getting the similar mode profile of MIM waveguide. Initially, investigation of mode profile starts form the lowest order of odd Ez components and low-order even Ex and Hy components and also VE, HO, HE mode profiles.

All these lowest-order mode profiles match what would be expected for the horizontal or vertical MIM waveguide. This analysis used to find the single rectangular waveguide encapsulated the two MIM waveguide benefits of

Table I. Propagation length analysis for the rectangular plasmonic waveguide.

	Mode	Propagation length (nm)	Wavelength (nm)
Rectangular plasmonic waveguide	HO	2780	600
	VO	110	250
	HE	62.118	360
	VE	78.24	450

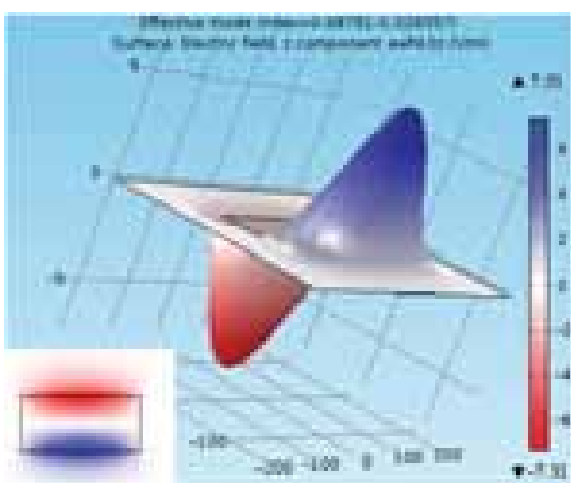


Fig. 4. The respective mode profile Hz is odd, whereas Ey and Hx are even in the lowest-order odd mode generated by the vertical (short) boundaries (VO) in the rectangular waveguide and its dispersion features.

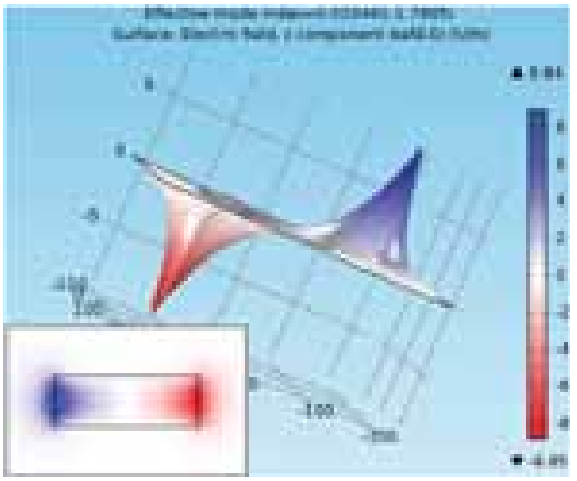


Fig. 5. The respective mode profiles. Ez is odd, whereas Ex and Hy are even, in the lowest-order odd mode generated by the vertical (short) boundaries (VO) in the rectangular waveguide and its dispersion features.

guided modes. This guided mode is similar at HO and VO mode. They are terms of spatial mode profile of rectangular waveguide. HO mode is compared the microwave guided mode of TE_{10} and VO is compared to the TM_{11} of microwave guided mode. There is a cutoff wavelength in the HO and VO as 760 nm and 440 nm. They indicating a red shift as well. When apply this mode for the possible applications, to analyse the dispersion of the rectangular waveguide.

The modal effective index is obtained by the ratio of the wavenumber of the waveguide mode. The propagation length analysis for the rectangular plasmonic waveguide is defined by the equation of $L_p = \lambda/\{4\pi Im(n_{\text{eff}})\}$. Finally, the maximum propagation length for the mode of HO,VO is defined by the above equation (see Table I).

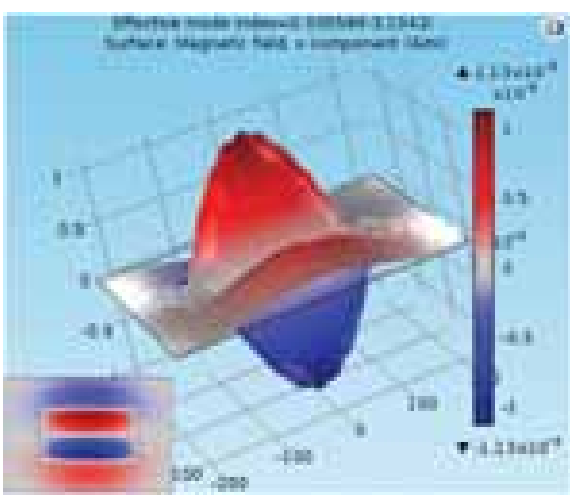


Fig. 6. Ez is even, whereas Ey and Hx are odd, in the lowest-order even mode generated by the horizontal (long) dielectric boundaries (HE) and its dispersion features.

3.2. Dispersion Relation

Mode profiles for the lowest-order modes (VO, VE, HO, and HE) in a rectangular metallic waveguide and dispersion characteristics are shown below (see Figs. 4–7). The dielectric filling is air and the metal is silver. The insets in all subfigures represent the x–y projection of their respective mode profiles. The mode profiles are generated at wavelength 650 nm, except for the HE mode, which has its component profile generated at 500 nm. Dispersion analysis for the plasmonic waveguide are done by using MATLAB from the below equation (see Fig. 7).

$$d = -\frac{\lambda}{c} * d^2 \frac{n_{\text{eff}}}{d\lambda^2}$$

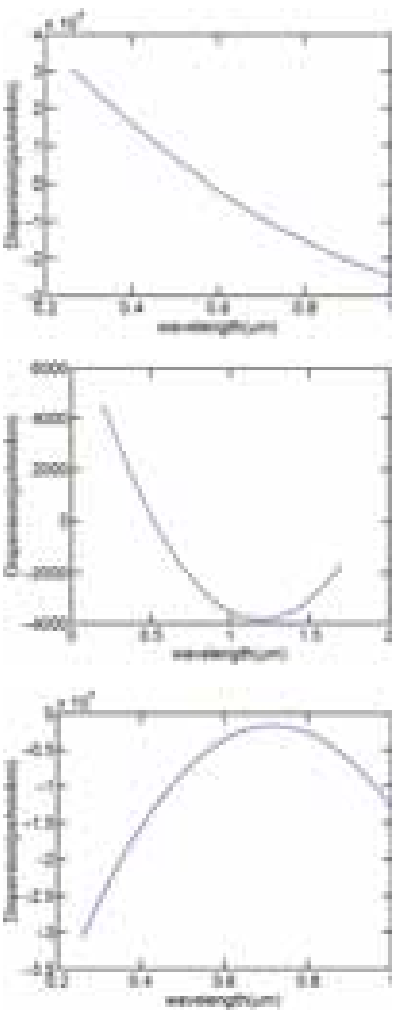


Fig. 7. Continued.

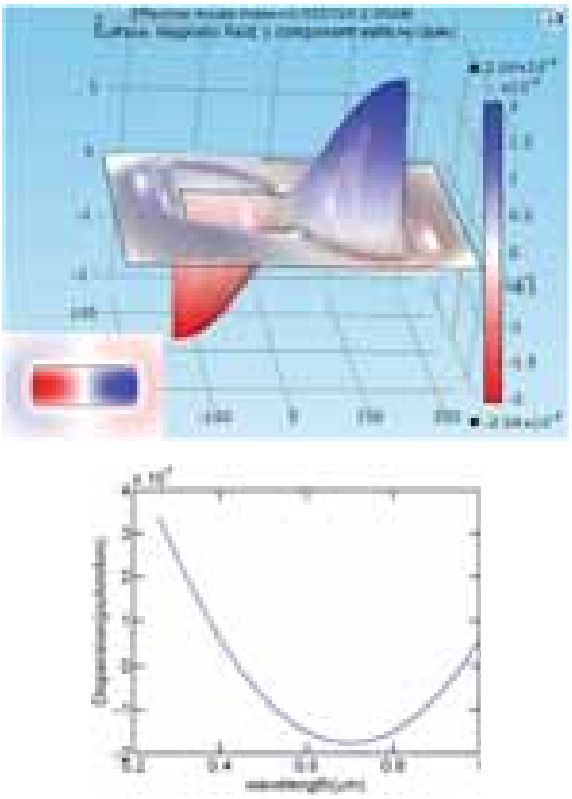


Fig. 7. *Ez* is even, whereas *Ex* and *Hy* are odd, in the lowest-order even mode generated by the vertical (short) dielectric region boundaries (VE) and its dispersion features.

Dispersion relation for the four key guided modes is shown in Figures 4–7 where dispersion is plotted against the wavelength. The graph explains that each key guide modes offer from the cutoff wavelength VO and VE-guided modes at around 440 nm and HO and HE at 760 nm.

4. CONCLUSION

The analysis of four low-order guided modes of the rectangular plasmonic waveguide was discussed. These mode analyses of the guided modes are related to MIM waveguides. The Maximum propagation length in rectangular structure HO and VO modes is identified with the wavelength range 650 nm, 250 nm, 360 nm, 450 nm. The cutoff of the HO mode occurs at a wavelength larger than that of VO mode due to the longer horizontal boundaries compared to the vertical boundaries. Each boundary occurs by the cutoff wavelength. They are controlled by controlling the boundary length. The dispersion can be exploited in applications such a sensing, filtering and also EOT.

References

1. S. A. Maier, Plasmonics: Fundamentals and Applications, Springer-Verlag, US **(2007)**.
2. H. Hu, X. Zeng, C. Tong, W. A. Anderson, Q. Gan, J. Deng, and S. Jiang, *J. Plasmonics* 8, 239 **(2013)**.
3. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* 424, 824 **(2003)**.
4. P. Berini, *Phys. Rev. B* 61, 10484 **(2000)**.
5. Y. S. Bian, Z. Zheng, X. Zhao, J. S. Zhu, and T. Zhou, *Opt. Express* 17, 21320 **(2009)**.
6. H. J. Lezec, J. A. Dionne, and H. A. Atwater, *Science* 316, 430 **(2007)**.
7. R. F. Oulton, V. J. Sorger, D. A. Genov, D. F. P. Pile, and X. Zhang, *Nat. Photon.* 2, 496 **(2008)**.
8. X. Kong, Z. Li, and J. Tian, *Opt. Express* 21, 9437 **(2013)**.
9. T. Yang and K. Crozier, *Opt. Express* 17, 1136 **(2009)**.
10. M. A. Swillam and A. S. Helmy, *Opt. Express* 18, 19831 **(2010)**.

Received: 19 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1582–1589, 2018

RESEARCH ARTICLE

Very Large Scale Integration Implementation of High Speed Median Filter with Low Power Switching Algorithm

T. Brindha[1,*], B. Vinoth Kumar[2], and K. N. Vijeya Kumar[3]

[1] PG Scholar, Department of EEE, Dr. Mahalingam College of Engineering and Technology, Tamil Nadu 560011, India
[2] Department of EEE, Dr. Mahalingam College of Engineering and Technology, Tamil Nadu 560011, India
[3] Department of ECE, Dr. Mahalingam College of Engineering and Technology, Tamil Nadu 560011, India

The main objective of this paper is to build an efficient hardware architecture for image denoising using spatial filtering techniques. In this design, high speed sorting, noise detection and pixel replacements are carried out with low power and delay utilization. A comparative study of standard median filter with the Low Power Switching Median Filter (LP-SWMF) has been done. Performance measurement of Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Image Enhancement Factor (IEF), Structural Similarity Index (SSIM), Mean Absolute Error (MAE) is done comparing the filters. The designs are modeled in Xilinx System Generator (XSG) and are implemented in SPARTAN-6 using hardware co-simulation.

Keywords: Standard Median Filter, Switching Median Filter, Low Power Switching Median Filter, System Generator, Performance Measurements.

1. INTRODUCTION

Digital image processing is a process of converting an image into digital data and processing them to obtain useful information. Nowadays digital image processing has become ineluctable in the fields of space exploration, surveillance, medicine authentication, automation, artistic effects (etc.), moreover digital image processing helps the people to retrieve data an interpret information[1] e.g., weather forecasting. The main source of noise in digital images arises during image acquisition (digitization) or during image transmission.[2] The performance of image sensor is affected by variety of reasons such as environmental condition during image acquisition or by the quality of the sensing element themselves. During the process of image retrieval, the image gets affected by addition of unwanted data that corrupts the quality of the image called as noises. Thus, the process of denoising is vital to remove the unwanted information's in the image. The process of denoising is illustrated in Figure 1. The image obtained from the source is taken as $I(X, Y)$, the process of degradation adds up noise to the original image named as $N(X, Y)$ and the noised image is sent to the noise filter where the noise is filtered using suitable filter and the output obtained is $O(X, Y)$. The noise is usually divided into Gaussian noise, the balanced noise and the impulse noise. The arisen impulse noises display as light and dark noise pixels under random distribution on the image. This not only corrupts true information of the image, but also seriously affects the visual effects of the image.

2. XILINX SYSTEM GENERATOR

The process is carried out in xilinx system generator (XSG). The process involves various stages of image preprocessing, central unit and post-processing unit as shown in the Figure 2.

2.1. Image Pre-Processing Unit

As the images are two-dimensional (2D) arrangement, to meet the hardware requirement the image should be preprocessed and given as one dimensional (1D) vector. The model based design used for image preprocessing. The model of pre-processing unit is shown in the Figure 3.

2.2. Central Processing Unit

All Xilinx blocks work on fixed point but the real-world signal (image, voice signal, etc.) are floating point so here the gateway in and gateway out blocks acts as translators for converting the real-world signal into the desired form.[23] Xilinx block sets works only with the presence of properly driven gates illustrated in the Figure 4.[3–5]

2.3. Image Post-Processing Unit

The image post processing blocks which are used to convert the image output back to floating point type are shown

*Author to whom correspondence should be addressed.

1546-1955/2018/15/1582/008 doi:10.1166/jctn.2018.7344

Fig. 1. Structure of denoising.

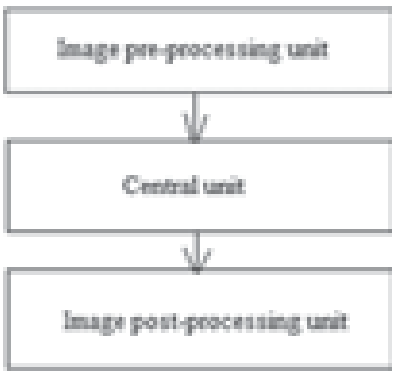


Fig. 2. Simulation procedure.

Fig. 3. Image pre-processing unit.

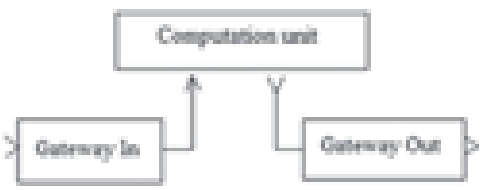


Fig. 4. Central processing unit.

Fig. 5. Image post-processing unit.

in Figure 5. For post processing, it uses a buffer block which converts scalar samples to frame output at lower sampling rate, followed by a 1D to 2D format signal block.

3. VARIOUS TYPES OF NOISES

3.1. Uniform Noise

The uniform noise cause by quantizing the pixels of image to a number of distinct levels is known as quantization noise.[6] It has approximately uniform distribution. In the uniform noise, the level of the gray values of the noise are uniformly distributed across a specified range. Uniform noise can be used to generate any different type of noise distribution. This noise is often used to degrade images for the evaluation of image restoration algorithms. This noise provides the most neutral or unbiased noise. This is derived in the equation following

$$P(Z) = \begin{cases} \dfrac{1}{b-a}, & \text{if } a \le Z \le b \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

The mean and variance are given by

$$\mu = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}$$

3.2. Gaussian Noise

This noise has a Probability Density Function (PDF) of the normal distribution. It is also known as Gaussian distribution. It is a major part of the read noise of an image sensor that is of the constant level of noise in the dark areas of the image. This is derived by the equation following

$$P(Z) = \sqrt{\frac{1}{2\pi\sigma^2} e^{(-(Z-\mu)^2)/(2\sigma^2)}} \tag{2}$$

Where Z = gray value, s = standard deviation and μ = mean.

3.3. Speckle Noise

Speckle noise is a glandular noise that inheritably exists in and degrades the quality of the active radar, synthetic aperture radar (SAR). Medical ultrasound and optic coherence tomography images. It is a conventional radar results from random fluctuations in the return signal from an object that is no bigger than a single image-processing element. It increases the mean gray level of a local area. This noise is a multiplicative noise.

$$P(Z) = \frac{Z^{\alpha-1} e^{-z/a}}{\alpha - 1! a^{\alpha}} \tag{3}$$

3.4. Salt and Pepper Noise

The salt-and-pepper noise are also called shot noise, impulse noise or spike noise that is usually caused by faulty memory locations, malfunctioning pixel elements in the camera sensors, or there can be timing errors in the process of digitization. In the salt and pepper noise there are only two possible values exists that is a and b and the probability of each is less than 0.2. If the numbers greater than his numbers the noise will swamp out image. For 8-bit image the typical value for 255 for salt-noise and pepper noise is 0. Salt and pepper noise is illustrated by

$$P(Z) = \begin{cases} P_a, & \text{for } Z = a \\ P_b, & \text{for } Z = b \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

3.5. Rayleigh Noise

Radar range and velocity images typically contain noise that can be modeled by the Rayleigh distribution. Rayleigh noise presents in radar range images. This is illustrated in the following equation

$$P(Z) = \begin{cases} \frac{2}{b}(Z-a)e^{(-(Z-a)^2)/b}, & \text{for } Z \geq a \\ 0, & \text{for } Z < a \end{cases} \quad (5)$$

Where $\mu = a + \sqrt{\pi b/4}$, $\sigma^2 = b(4-\pi)/4$.

3.6. Gamma Noise

The noise can be obtained by the low-pass filtering of laser based images. Gamma noise is generally seen in the laser based images. It obeys the gamma distribution. This is illustrated by the following expression

$$P(Z) = \begin{cases} \frac{a^b z^{b-1} e^{-az}}{(b-1)!}, & \text{for } Z \geq 0 \\ 0, & \text{for } g < 0 \end{cases} \quad (6)$$

Where, $\mu = b/a$, $\sigma^2 = b/a^2$.

4. MEDIAN FILTER

There are many denoising filters in existence, median is found to be the one of the standard denoising filter for impulse type of noises. It is mainly used for smoothing and denoising of digital images the main advantage of using median filter is that it preserves the edges of the image.[7] Edges protects the important information of the image thus edges are considered to be the important part an image. The median filter is a non-linear filter that works on the affected pixels of the image and process them to get an efficient result from them.

4.1. Impulsive Noise

Median filter works effectively on impulsive noises.[8] When an image gets affected by impulsive noise the pixels of the image is replaced by 0 or 255 which is also called as salt and pepper noise. The affected image is shown in the Figure 6 that explains how the pixels changes after the influence of noise. Figure 6(a) shows the pixels of unaffected image and Figure 6(b) shows the affected pixels.

4.2. Working of Median Filter

A standard median filter selects the number of pixel based on the window size, either $3*3$ or $5*5$. All the elements of the selected window are sorted and the center value is found which is called as median value. This value replaces the affected pixels. The working is further illustrated in the following Figures 7(a) and (b).

(a)

(b)

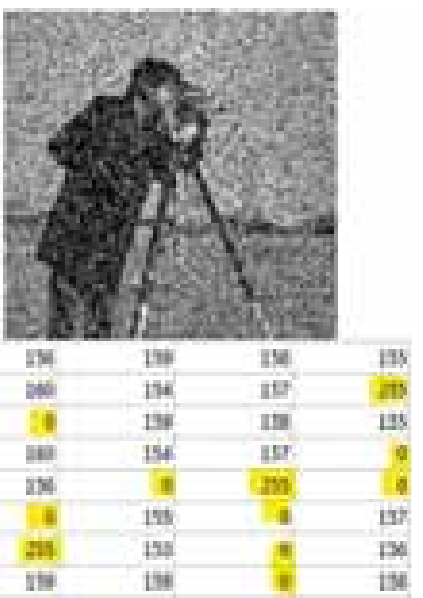


Fig. 6. (a) Original pixels (b) pixels when affected by salt and pepper noise.

4.3. Working of Switching Median Filter

Unlike standard median filter switching median does not replaces all the pixels. Switching median filter has a noise detection module that detects the affected pixel and replaces only the affected pixel. The structure of switching median filter is shown in the Figure 8.

The median computation unit includes the architecture for sorting, illustrated in the Figure 8, that includes seven comparators for finding the median value. The input unit connected to the window creation unit gives the elements to the median computation unit where the elements are

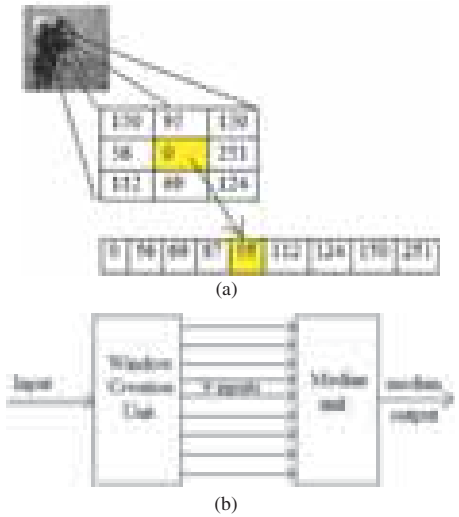


Fig. 7. (a) Median computation (b) functional block of median filter.

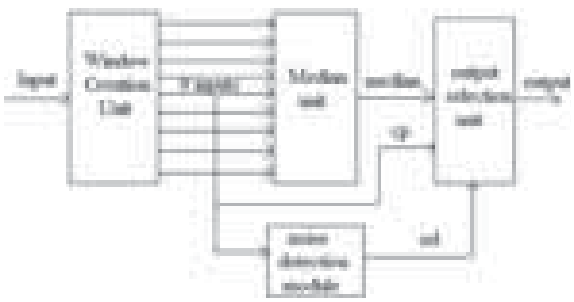


Fig. 8. Switching median computation.

sorted and the final median value is obtained. This process is repeated for number of iteration until all the windows are processed. The proposed architectures are designed in pipeline manner[8] to minimize computational time. Parallel processing[7] has also been done for process acceleration. A 3×3 or 5×5 window size can be selected for computation of the filter output. The drawback of median filter is overcome by switching median filter with a noise detection module as shown in the Figure 8.

The output selection unit holds a multiplexer with noise detection module as select line. The values of median computation module and the centre pixel value are detected using a noise detection module. This module checks whether the value is 0 or 255. The median computation module common for both standard median and switching median is shown in Figure 9.

Fig. 9. Median computation block.

5. PROPOSED MEDIAN FILTER WITH LOW POWER SWITCHING

The main drawback of median filter and switching median is that it computes median value for all the pixel value irrespective of affected pixel. This takes up more power for computation of unnecessary median value. In order to overcome this problem in median filter a noise detection module[9] with enable unit is introduced. The enable unit turns on only when the values at the input is detected to have noise. The neighbourhood pixels of the affected pixel is alone processed. This helps us to reduce calculation of all the median values. If the pixel is not affected then the median computation does not takes place. The existing centre value is left unchanged. The structure for low power[10] switching median filter is shown in Figure 10.

5.1. Simulation of Low Power Switching Median Filter (LP-SWMF)

The software simulation is done in Xilinx 14.5 tool of 4 GB RAM. The results have been tested on gray scale images of size $256 * 256$, the images taken for testing are as follow (See Fig. 11).

Fig. 10. Structure of low power switching median.

(a) (b) (c)

Fig. 11. Sample images. (a) Lena.png (b) Cameraman.tif (c) Bridge.jpeg.

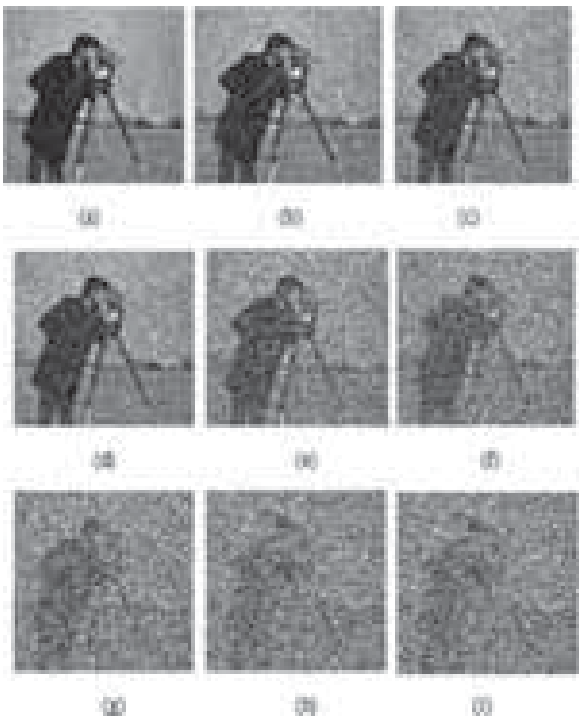


Fig. 12. Noised cameraman image. (a) 10% (b) 20% (c) 30% (d) 40% (e) 50% (f) 60% (g) 70% (h) 80% (i) 90%.

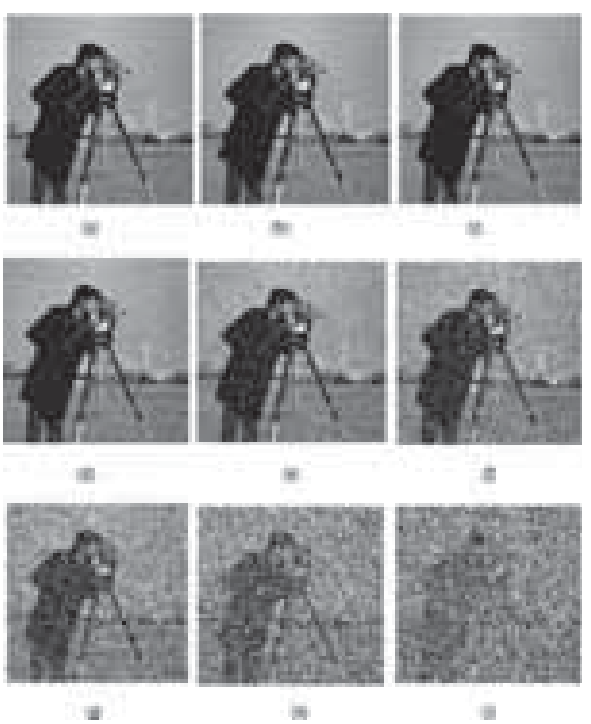


Fig. 13. Restored images using SMF.

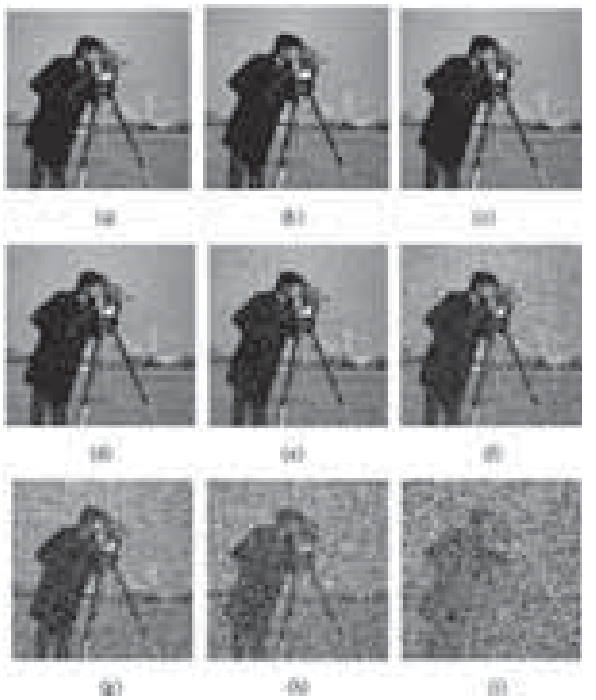


Fig. 14. Restored images using SWMF and LP-SWMF. (a) 10% (b) 20% (c) 30% (d) 40% (e) 50% (f) 60% (g) 70% (h) 80% (i) 90%.

The images with noise percentage ranging from 10 to 90 percentage are taken as test images and processed with standard median and switching median filters. There are various image quality parameters to detect the effect of filter over the image. The various parameters chosen are calculated and the results are analyzed for Standard Median Filter (SMF), Switching Median Filter (SWMF) and Low Power Switching Median Filter (LP-SWMF). The power and delay parameters are measured along with the device utilized. Denoising filters SMF and SWMF are applied for noised cameraman image and the results are analyzed. The results are shown in the Figures 12–14. Simulation setup of design is shown in the Figure 15. RTL schematic views of SMF, SWMF and LP-SWMF are shown in Figures 16(a)–(c) respectively that helps in better understanding of the designs.

6. QUALITY AND UTILIZATION PARAMETERS

6.1. MSE

MSE denotes Mean Square Error. Expression (7) is used to calculate MSE.

$$\mathrm{MSE} = \frac{1}{MN} \sum_{M=0}^{M-1} \sum_{N=0}^{N-1} |e(M,N)^2| \tag{7}$$

Where, M and N are the total number of pixels in the horizontal and the vertical direction in the image respectively

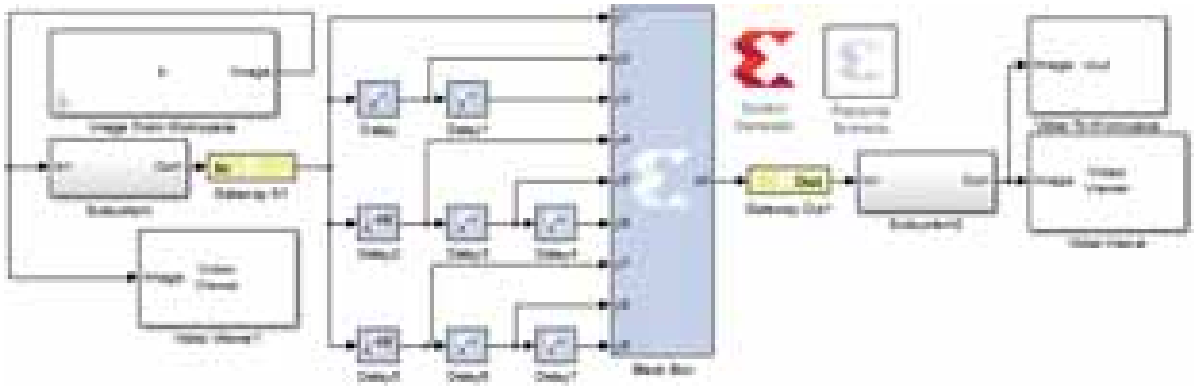


Fig. 15. Simulation setup in xilinx system generator.

(a)

(b)

(c)

Fig. 16. RTL schematic diagram (a) SMF (b) SWMF (c) LP-SWMF.

i.e., size of the image. $e(M, N)$ is the error difference between the original and the distorted images. This measures the average of squares of errors.

6.2. PSNR

PSNR denotes Peak Signal to Noise Ratio. It is defined as the proportion between the peak signal power to noise power calculated using the expression (8)

$$\mathrm{PSNR} = 10\log\frac{S^2}{\mathrm{MSE}} \tag{8}$$

Where, $s = 255$ for an 8-bit image. MSE denotes Mean Square Error.

6.3. MAE

Mean absolute error (MAE) is a measure of difference between two images. MAE is given by the expression (9)

$$\mathrm{MAE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|X_{ij} - Y_{ij}| \tag{9}$$

Where, X_{ij} is original image, Y_{ij} is restored image.

6.4. IEF

IEF is defined as Image Enhancement Factor. It is the ratio between the square of the noise power in a noisy image to the square of the residual noise present in the denoised image and is given by Eq. (10)

$$\mathrm{IEF} = \frac{\sum_i\sum_j(\eta(i,j) - Y(i,j))^2}{\sum_i\sum_j(\bar{Y}(i,j) - Y(i,j))^2} \tag{10}$$

Where, Y, $\bar{Y}$ and η is original, restored and corrupted image respectively.

Table I. Evaluation of MSE, PSNR, MAE, IEF, MSSIM for SMF, SWMF, LP-SWMF.

Noise (%)	Factor	MSE	PSNR	MAE	IEF	SSIM
10	SMF	22.632	34.584	1.77	89.247	0.9670
	SWMF	6.2541	40.169	0.3108	320.069	0.9934
	LPSWMF	6.2541	40.169	0.3108	320.069	0.9934
20	SMF	76.03	29.32	2.5076	52.75	0.9168
	SWMF	35.773	32.595	0.8392	112.191	0.9619
	LPSWMF	35.773	32.595	0.8392	112.191	0.9619
30	SMF	296.96	23.404	4.2785	20.359	0.7584
	SWMF	176.503	25.664	2.0954	34.304	0.8402
	LPSWMF	176.503	25.664	2.0954	34.304	0.8402
40	SMF	895.82	18.609	8.3888	9.0561	0.4780
	SWMF	590.020	20.422	5.1149	13.651	0.6005
	LPSWMF	590.020	20.422	5.1149	13.651	0.6005
50	SMF	2.055e3	15.002	15.8087	4.9019	0.2556
	SWMF	1.483e3	16.420	11.1086	6.8078	0.3345
	LPSWMF	1.483e3	16.420	11.1086	6.8078	0.3345
60	SMF	4.064e3	12.041	28.431	2.979	0.1191
	SWMF	3.121e3	13.187	21.687	3.869	0.1594
	LPSWMF	3.121e3	13.187	21.687	3.869	0.1594
70	SMF	6.947e3	9.7126	46.462	2.026	0.055
	SWMF	5.710e3	10.564	37.773	2.475	0.077
	LPSWMF	5.710e3	10.564	37.773	2.475	0.077
80	SMF	1.085e3	7.7742	70.4268	1.489	0.0288
	SWMF	9.383e3	8.4071	60.9038	1.719	0.0363
	LPSWMF	9.383e3	8.4071	60.9038	1.719	0.0363
90	SMF	1.518e4	6.3178	97.2535	1.193	0.0137
	SWMF	1.421e4	6.6050	90.6770	1.277	0.0163
	LPSWMF	1.421e4	6.6050	90.6770	1.277	0.0163

Table II. Calculation of power, delay and device utilization by the filters.

Parameter	SMF	SWMF	LPSWMF
Path delay (ns)	2.20	2.12	2.47
Static power (mW)	156.68	156.64	156.62
Dynamic power (mW)	19.91	12.66	9.88
Slices	160	156	186
FF's	304	304	304
LUT's	490	495	549
IOB's	16	16	16

6.5. SSIM

The structural similarity (SSIM) index is a method for predicting the perceived quality of Digital images. SSIM is used for measuring the similarity between two images. SSIM is a perception-based model that considers image degradation as perceived change in structural information, while also incorporating important perceptual phenomena, including both luminance masking and contrast masking terms and is given by Eq. (11)

$$\mathrm{SSIM}(X, Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{11}$$

Where, μ_x represents average of X; μ_y represents average of Y; σ_X^2 is the variance of X; σ_Y^2 is the variance of Y, σ_{XY} represents the co variance of X and Y.

Thus, the various quality parameter is tested and the results are calculated. The Table I shows the detail of all the quality parameters tested shown in part VII.

7. RESULTS AND DISCUSSION

The simulation has been tested on various percentage of noises ranging from 10% to 90%. The performance of the proposed system is measured using various parameters and are listed in Table I that proves that the proposed LPSWMF is equally good comparing the SWMF in its performance. Along with the quality parameters the device utilization is also calculated to know the power utilization of the designed architecture. Table II shows the results of static power, dynamic power, delay, slices, flip-flop's, LUT's. The hardware co-simulation is done using the spartan-6 and the device utilization are calculated. The power consumed by LPSWMF is comparatively low when compared to SMF by 50 percent and SWMF by 39 percent.

8. CONCLUSIONS

A low power consuming denoising filters (LP-SWMF) is designed. The existing standard median filter is replaced by Seven Comparator Sorting Architecture (SCSA) that eventually improves the speed of sorting. The proposed Low Power Switching Median Filter (LP-SWMF) works on denoising the image with low power utilization. From the analysis it is concluded that the proposed LP-SWMF reduces the power consumption of SMF by 39% and the SWMF by 50%. This shows that the proposed LP-SWMF is a better low power switching median filter.

References

1. R. C. Gonzalez and R. E. Woods, Digital Image Processing, Englewood Cliffs, Prentice-Hall, Eaglewoods **(2002)**.
2. G. Jiang, M. Luo, S. X. Chen, and K. Bai, An improved filtering algorithm for the CT image processing, *IEEE International Conference on Applied System Innovation*, IEEE **(2017)**.
3. P. Luciano, C. L. Sotiropoulou, and S. Gkaitatzis, *IEEE Transactions on Nuclear Science* 64, 1374 **(2016)**.
4. M. C. Hanumanth Raju, M. Ravi Shankar, D. R. Ramesh Babu, and S. B. Satish, *International Journal of Information and Electronics Engineering* 1, 1 **(2011)**.
5. E. Ineo Junior, L. M. Garces, A. J. Cabrera, and T. C. Pimento, *IEEE Latin Transactions* 14, 2120 **(2016)**.
6. W. Y. Han and J. C. Lin, *Electron. Lett.* 33, 124 **(1997)**.
7. Y. Zhang and S. Q. Zhengb, *VLSI Design* 11, 137 **(2000)**.
8. S. Nirmal Raj, S. Ashok, P. Bala Vengateswarlu, and G. Vishnuvardhan Rao, *Journal of Chemical and Pharmaceutical Sciences* 9 **(2016)**.
9. W. Luo and D. Dang, An efficient method for the removal of impulse noise, *IEEE International Conference on Image Processing* **(2006)**, Vol. 12, pp. 2601–2604.
10. W. Luo and D. Dang, *IEEE Signal Processing Letters* 14, 193 **(2017)**.
11. Haidi Ibrahim and Nicholas Sia Pik Kong, *IEEE Signal Processing Letters* 54, 1920 **(2008)**.
12. K.-L. Chung and Y.-K. Lin, *Elsevier Signal Processing* 63, 101 **(1997)**.
13. S. J. Ko and Y. H. Lee, *IEEE Transactions on Circuits and Systems* 38, 984 **(1991)**.
14. L. Eng and K. Ma, *IEEE Trans. Image Process* 10, 242 **(2017)**.
15. R. Hwang and R. A. Hadded, *IEEE Transaction on Image Processing* 4, 499 **(1995)**.
16. Mohammed Alareqi, R. Elgouri, M. Tarhda, K. Mateur, A. Zemmouri, A. Mezouari, and L. Hlou, Design and FPGA implementation of real-time hardware co-simulation for image enhancement in biomedical applications, IEEE **(2017)**.

Received: 19 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1590–1593, 2018

RESEARCH ARTICLE

Restaurant Visitor Time Series Forecasting Using Autoregressive Integrated Moving Average

G. Boomija*, A. Anandaraj, S. Nandhini, and S. Lavanya

Department of Computer Science, Narasu's Sarathy Institute of Technology, Salem 636305, Tamil Nadu, India

Many restaurants may not have proper prediction system for forecasting the visitors. The main aim of this paper is to forecast visitors to restaurants in Japan. There is a need to know how many visitors to expect each day. Forecasting helps in effective planning for purchase of ingredients and scheduling of staffs. The visitor forecasting is challenging because unpredictable factors do affect the arrival of visitors to restaurants. This paper builds forecast of visitors using regression techniques. In order to estimate the visitors, datasets of reservation and visitation of the restaurants are used. The time series data is analyzed using Box-Jenkins ARIMA models the residuals have the structure of AR (Auto Regressive). Forecasting of visitors will provide information that helps the restaurants to be much more successful in overcoming the troubles caused by factors. This also allows the restaurants to focus on creating an enjoyable dining experience for visitors.

Keywords: Forecasting, Regression Techniques, Time Series Data, Auto Regressive, ARIMA.

1. INTRODUCTION

There is a steady increase of restaurants in the 21st century and running a local restaurant is not an easy job. The number of visitors to a restaurant for a period of time is random. The advent of many a restaurants in a region marks the chance of the existing restaurants to lose scope. There is a predicament with expectation of visitors to a restaurant. It seems to be inconvenient for the successful run of the eatery in a locality thus, disturbing the economic attributes of the functioning market. Forecasting plays a major role with planning the total number of visitors using the time series data. The reservation of visitors are also a part of the prediction perspective. The challenges include in building a technically efficient model to forecast and this paper explores the appliance of time series forecasting model Autoregressive Integrated Moving Average (ARIMA). The ARIMA approach planned by Box and Jenkins is popular as it yields most accurate forecasting results. A study of time series data of visitors to restaurants in Japan over a period of 2016–2017 to create models.

2. RESTAURANTS AND FACTORS AFFECTING ITS PROCEEDS

Restaurants are public place that provide food on commercial basis. Restaurants offer food as a service in desire to satisfy the visitors. Prediction of these visitors helps the restaurants to improve its functional services and this paper focuses the purpose. Recruit Holdings has unique access to restaurant data. Hot Pepper Gourmet—a restaurant review service, AirREGl—a restaurant sales service and Restaurant Board—reservation log management software are all owned by Recruit Holdings. Using these data retrieved with assistance from websites, forecasting of visitors becomes evident. There are certain factors that might stop the restaurant's attendance. These factors include weather conditions, local competitions and new restaurants with few historical data. Weather is always unpredictable in Japan, in spite of the usual monsoons and summers. Local competition is undesirable based on location and new restaurants will have less past data which is insufficient to forecast visitors in future.

3. DATASETS INVOLVED IN PREPROCESSING

Time series datasets are available from the two sites, Hot Pepper Gourmet (hpg) and AirREGl (air). Through hpg the visitors can search restaurants and reserve online. While air is a reservation control and cash register system. The datasets of reservation and visitation for a year are mainly used. As it is a time series data, time series models are applied. The datasets used are reservation dataset of air and hpg followed by the information of restaurants where latitude and longitude is a part of locating the restaurants.

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1590/004 doi:10.1166/jctn.2018.7345

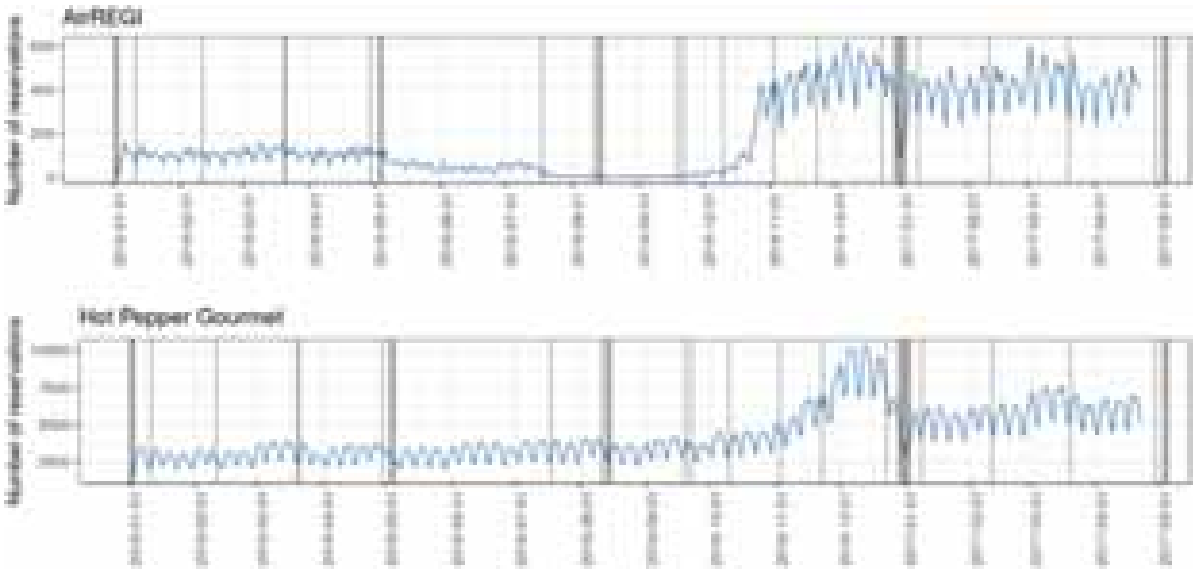


Fig. 1. Reservations made with both the sites.

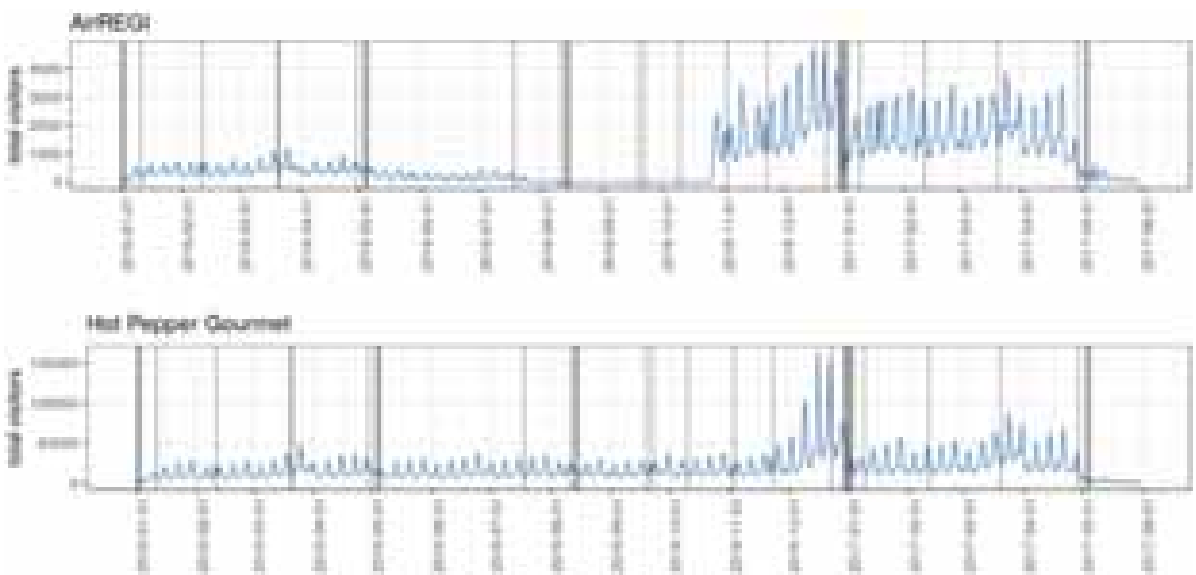


Fig. 2. Total number of visitors based on reservation.

A relation dataset is used, where the common restaurants in both the sites are filtered. Information of date is analyzed to withstand with holidays and Japan has a week's holiday called "The Golden Week." A summary of visits are analyzed as Explanatory Data along with holiday and reservations made. There is a comparison between restaurants based on the cuisine offered in both the sites.

Figure 1 portrays the number of reservation made based on the date in both the sites.

Figure 2 depicts the number of visitors who have arrived for the reservations made on that specific date for the restaurants at both the sites.

4. PURPOSE OF TIME SERIES IN PROCESSING

A time series is a structured dataset that can be annual, monthly, or in this case daily. The visit data and reserve data used here are time series data. Even date is analyzed for better results. Descriptive model is applied for making predictions. The visit data specifies the time of reservation and reserve data specifies the time when the reservation was made. This model helps in describing the trends, patterns and it uses the information in time series to forecast future values of that series. Making prediction about future is called as extrapolation. Forecasting

involves taking models that fit on historical data and use them to predict future. The purpose of time series analysis are to understand or model the stochastic mechanism that gives rise to an observed series and to predict or forecast the future values of a series based on the history of that series. Time series include four parts level, trend, seasonality and noise. Level is undesirable because it forms the baseline values and in this paper the baseline values are the data of date and time. The next important feature is seasonal variation, where seasons in Japan play a vital role. These are the parameters of time series considered in this paper.

5. FEATURE ENGINEERING FOR BETTER ESTIMATION

There is no source on how to do feature engineering and it depends on the problem. Feature engineering is a thoughtful creation of new input fields from existing. These input fields could result in better inferences. In this case of forecasting the number of visitors to a restaurant, there is an observation of weather as a factor. So, adding an extra feature of weather to the dataset can represent the result at its best. The inputs can be turned into a version which the model can understand. The data can be analyzed in two perspectives. One is to see the difference pattern between visitors on days of holidays and non-holidays and the restaurant visitor behavior. This varies from restaurant to restaurant and processing using these can add a feature to the result as well. In the datasets, we can even feature snowfall record in a particular region, pressure and humidity. Even some stations might have a record of solar radiation and cloud cover as a feature of weather. There will be some important features which can lead to accuracy in forecasting and some features that seem irrelevant. The relevant factors are featured in this paper.

From Figure 3 changes in graph can be seen. The gray lines indicate holidays and the number of visitors vary with different sites and the restaurants present in common with both the sites.

By Figure 4 the visitors for a particular week is considered and the number of visitors during Friday, Saturday and Sunday is high in comparison with the remaining days of the week.

6. ARIMA-A TIME SERIES MODEL FOR FORECASTING

Box and Jenkins developed a time series model ARIMA, which combines three types of processes. This model is used with time series data to understand the data or to predict future in series.

- AR: Auto regression, A model that uses the dependent relationship between an observation and some number of lagged observations.
- I: Integrated, The use of differencing raw observations to make the time series stationary.
- MA: Moving Average, A model that uses the dependency between an observation and residual error.

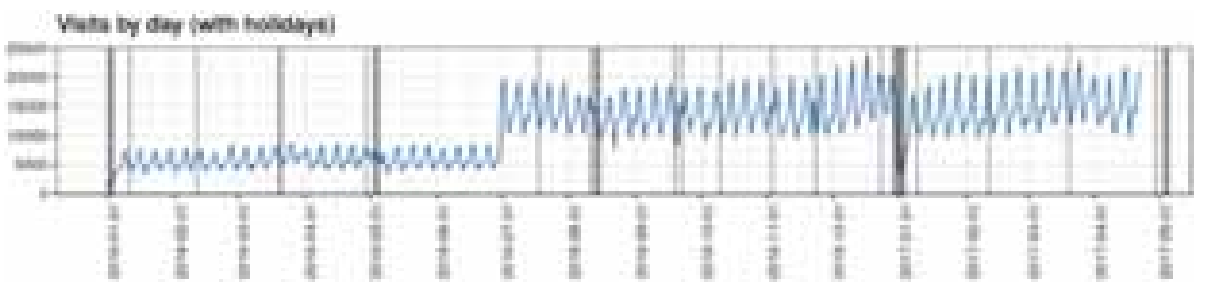


Fig. 3. Visitor forecast during holidays and non-holidays.

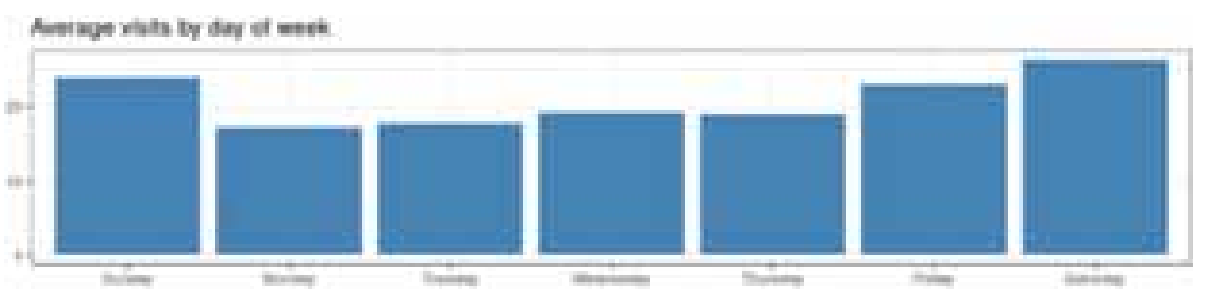


Fig. 4. Visitor forecast for a week.

ARIMA has three parameters

- p: Lag order.
- d: Degree of differencing.
- q: Order of moving average.

For building a stochastic model, three steps are involved as an iterative approach. Identification, Estimation and Diagnostic checking. ARIMA model gives better results for forecasting and this time series model is employed in forecasting the visitors to restaurants in Japan where p, q and d are integers greater than or equal to zero.

7. FORECASTING OF VISITORS USING ARIMA

The main technique being explored is ARIMA which is a forecasting technique that projects the future values of a series based entirely on its own inertia. Its main application is in the area of short term forecasting requiring at least 40 historical data points. The auto regressive part of the model determines the dependence on previous time points. The data is split into training set and testing set. The training period was different to testing period as there are different cuisine restaurants in this case. So, the train period is shortened and the test period too. Here period is weekly cycle and the visitors per cuisine and per restaurant is valuated. These data are taken for training to the ARIMA model. As ARIMA model uses (p, d, q) the series should be made stationary and choosing of order models with ACF and PACF for fitting the ARIMA to the state of train is fixed.

8. PERFORMANCE MEASURES USING RMSE AND MAPE

The prediction of visitor arrivals during testing is done and the measures of accuracy were calculated. Root mean square error and Mean absolute percentage error were used for forecasting error in time period and to measure accuracy of continuous variables. The average y value with given x value is predicted using regression. They can be positive or negative as the predicted value under or over estimates the actual value. Squaring the residuals, averaging the squares, and taking the square root gives the r.m.s error. Then use the r.m.s. error of the y values as a measure of the spread about the predicted y value. It's the square root of the average of squared differences between prediction and actual observation.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(X_{\mathrm{obs},i} - X_{\mathrm{model},i})^2}{n}}$$

MAPE measures the average magnitude errors in a set of predictions without directions. It is the average over the test sample of absolute difference between predictions and actual observations.

$$\mathrm{MAPE} = \sqrt{\frac{\sum_{i=1}^{n}(X_{\mathrm{obs},i} - X_{\mathrm{model},i})}{n}}$$

Both are used to express average and RMSE have interesting implication for large errors. RMSE is more useful than MAPE because it is steady and RMSE does not necessarily increase with the variance of the errors. RMSE increases with the variance of the frequency distribution of error magnitudes. Hence RMSE is better.

9. CONCLUSION

This paper explores a growing trend in the future of restaurants by forecasting the number of visitors each day. In this study the time series forecasting model ARIMA is functional. Based on the above an efficient prediction of visitors to come in future was analyzed for the profits of the restaurant. This prediction results in easy scheduling of ingredients and staffs for enhanced experience of visitors. Even offers can be offered during bad weather conditions for the successful running of the restaurants. The results of the study is revealed using metrics like RMSE and MAPE, where RMSE gives better accuracy in predicting the visitors. The information thus obtained from forecasting the visitors will act as an aid of support to the restaurants in focusing with the development of creating an enjoyable dining experience for the visitors. As far as the experimental results, the system thus proposed works well in forecasting visitors and is robust too.

RESEARCH ARTICLE

References

1. C. Shu and A. Tsang, Adjusting for the Chinese New Year: An Operational Approach, External Department Hong Kong Monetary Authority **(2005)**.
2. K. Torre, D. Delignieres, and L. Lemoine, *British Journal of Mathematical and Statistical* 60, 85 **(2007)**.
3. T. Trimbur, Seasonal heteroskedasticity in Census Bureau construction series, Statistical Research Division U.S. Census Bureau Washington DC. 20233–9100 **(2006)**.
4. F. L. Chu, *Tourism Management* 29, 79 **(2008)**.
5. Tourism Council of Thailand, http://www.thailandtourismcouncil.org/home.php **(2009)**.
6. L. Vaughan and E. Romero-Frías, *Journal of the American Society for Information Science and Technology* 65, 707 **(2013)**.
7. S. Vosen and T. Schmidt, Forecasting private consumption **(2011)**.
8. W. H. K. Tsui, H. O. Balli, A. Gilbey, and H. Gow, Forecasting of Hong Kong airport's passenger throughput **(2014)**.
9. J. J. Flores, M. Graff, and H. Rodriguez, Evolutive design of ARMA and ANN models for time series forecasting **(2012)**.
10. I. Khandelwal, R. Adhikari, and G. Verma, Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition **(2015)**.

Received: 19 April 2018. Accepted: 14 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1594–1597, 2018

RESEARCH ARTICLE

Performance Improvement in Birefringence of Photonic Crystal Fiber Using Full-Vectorial Finite Element Method

D. Rajeswari[1,*] and A. Arunya Revathi[2]

[1] Research Scholar, Alagappa Chettiar Government College of Engineering and Technology, Karaikudi 630003, Tamilnadu, India
[2] Alagappa Chettiar Government College of Engineering and Technology, Karaikudi 630003, Tamilnadu, India

The effect of ellipticity on birefringence in a rectangle Photonic Crystal Fiber (PCF) having elliptical air holes is investigated, with single and double defects using Full-Vectorial Finite Element Method (FV-FEM) simulation. It is reported that birefringence increased with the increase of ellipticity. In this paper, this analysis is extended to single and double defect structures, where three adjacent air holes are omitted vertically from the rectangle structure and two selected columns are omitted vertically from the same. This double defect structure is found to have more birefringence than the structure with single defect. It is found that birefringence due to single defect is more for lower value of ellipticity; however at higher value of ellipticity, birefringence due to double defect is more than the one that could be due to single defect. Influence of pitch on birefringence is also investigated.

Keywords: Photonic Crystal Fiber, Birefringence, Finite Element Method.

1. INTRODUCTION

Photonic crystal fibers have demonstrated to be advantageous in many aspects than the conventional optical fibers. PCF structure contains an array of air holes as part of its cladding which differentiates it from the conventional optical fibers,[1,2] which have only a core and cladding or a combined core and cladding in case of multimode graded index fiber. Many photonic devices can be designed by PCF, like the PCF polarization filter, PCF polarization splitters, PCF sensors, PCF WDM (Wavelength Division Multiplexing), etc.

To fill the holes in the PCFs has produced extensive attention and has been the focus of study for many years. A number of materials have been filled into the fibers such as the liquid crystal, ethanol, and metal of silver and gold, which achieved many novel properties.[3–5] It is possible to obtain PCFs with absolutely opposite properties by changing the geometric characteristics of the air-holes in the fiber cross section that means their dimension or position. The key difference between PCF and conventional fiber is the Photonic crystal fibers have an air-silica cross section, whereas normal optical fibers have glass cross-sections.[6]

According to their mechanism for confinement Photonic crystal fibers possess two modes of act (i) index guiding and (ii) photonic band gap fibers. In index guiding PCF, where light is confined in a higher refractive index region, the light is guided by total internal reflection between the solid core and cladding region. Instead, when the core has a refractive index lower than that of the cladding region, as in hollow-core fibers, it is necessary the incidence of the photonic band gap.[7] A PCF has a holey cladding region with an adjoining a solid core.[7–9]

The birefringence and dispersive possessions of Polarization maintaining photonic crystal fiber were analysed using Finite element method. By breaking the circular symmetry and implementing asymmetric defect structures such as dissimilar air hole diameter, varying the number of circular and elliptical air holes, high birefringent PCF can be designed. The number of air hole rings around the central core need to be increased for making the structure more narrowed at an exact wavelength.[8,10] One of the richest features of PCFs is their application as Polarization Maintaining Fiber (PMF). Different approaches have been investigated to realize PMPCFs, which is universally based on birefringence.

The significant point in realizing the birefringence is to abolish the symmetry of fiber structure and increase the effective index difference between the two orthogonal polarization modes. The structural symmetry can be demolished either by changing the airhole size near the core area[11] or by altering the shape of the airholes (elliptical airholes).[10] A high level of birefringence is often required, in many sensing applications and in desires where light is required to maintain a linear polarization state. The most common technique to break the PCF structural symmetry is by changing the airhole size along two orthogonal axes near the core region. This results in increasing the difference between the effective indexes

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1594/004 doi:10.1166/jctn.2018.7346

of the orthogonal polarization modes. To the best of our understanding, highest birefringence in PCF's is reported with elliptical airholes; however, elliptical airholes are very difficult to control during the fabrication process.[10] To maintain the linear polarization, high level of birefringence is required by reducing polarization mode dispersion. This is possible by breaking the circular symmetry and implementing asymmetric defect structures, such as dissimilar air hole diameter, varying the number of circular and elliptical air holes. Using these concepts a high birefringence PCF can be designed. PMFs have extensive application in optical communications.

PMFs are also required for high bit rate transmission systems as they can eliminate polarization mode dispersion (PMD) and also stabilize the practical performance of optical devices. Recently polarization maintaining (PM) PCFs have been fabricated and their transmission characteristics have been reported[12] using two different airhole diameters along two orthogonal axis neighboring the core region, which provides an effective index difference between two orthogonal polarization modes.[11] The birefringence of such PM PCFs can be further improved by tuning the size of airhole diameters near the core region. It is shown that birefringence for PMPCF can be of one order higher greatness than that for conventional PMFs and is stated to be in the order of 10^{-3}.[12] PM PCFs show better polarization maintaining features that steady the polarization state in the fiber.

In this paper, introducing the defect by eliminating elliptical air holes in the fiber structure and its birefringence property of the photonic crystal fiber is investigated by using full-vectorial finite element method. Another one important parameter of the PCF is pitch. It is also varied and the effect of pitch on birefringence can be determined.

2. DESIGN AND PRINCIPLE OF OPERATION

Rectangle structure PCF design is acknowledged to be the finest arrangement for obtaining high birefringence. The material used here is pure silica with the lattice constant of 1.2 μm. The two different structures are designed and compared its various parameters. The first structure proposes elliptical air holes with single defect, whereas the second comprises elliptical air holes with two adjacent selected air holes missing vertically.

The rectangle structures having with single and double defect are shown in Figures 1(a) and (b). It is seen that elliptical air holes are periodically arranged throughout the crystal structure and we introduce here the nonlinearity in air holes of photonic crystal fiber. Ellipticity ratio is $\eta = r_x/r_y$ where r_x and r_y are the half lengths of the major and minor axes of the elliptical holes. Proposed PCF has the following length of the semi-major and semi-minor axes as $r_x = 0.5$ μm, $r_y = 0.2$ μm. Pitch is assigned with value $\Lambda = 1.2$ μm. Sellmeier equation is used for the purpose of determination of silica material (background) dispersion.

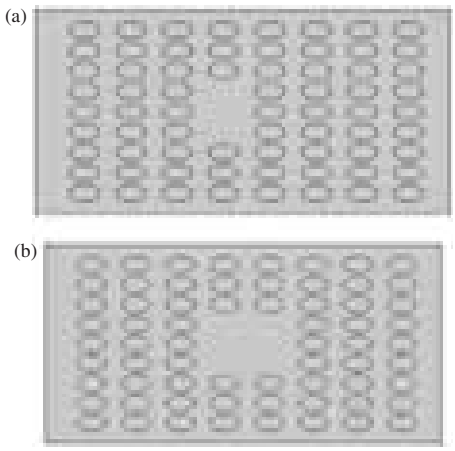


Fig. 1. Rectangle structure PCF with the core consisting of single defect and double defect.

This equation is given below

$$n(\lambda)^2 = 1 + \frac{A_1\lambda^2}{\lambda^2 - \lambda_1^2} + \frac{A_2\lambda^2}{\lambda^2 - \lambda_2^2} + \frac{A_3\lambda^2}{\lambda^2 - \lambda_3^2} \quad (1)$$

Here $A_1 = 0.6961663$, $A_2 = 0.4079426$, $A_3 = 0.897479$, $\lambda_1 = 0.068404$, $\lambda_2 = 0.1162414$, and $\lambda_3 = 9.896161$. All elliptical holes are filled with air which has the refractive index of 1.

There are various guiding mechanisms in two orthogonal directions, the two polarized (x and y) hybrid-guiding fundamental modes no longer degenerate, and then this kind of symmetric/asymmetric PCF has high birefringence. Birefringence can increase with the high refractive index of the filled material and the differences between diameter and the bandgap will shift to the longer wavelength.

$$B = |n_{eff}^x - n_{eff}^y| \quad (2)$$

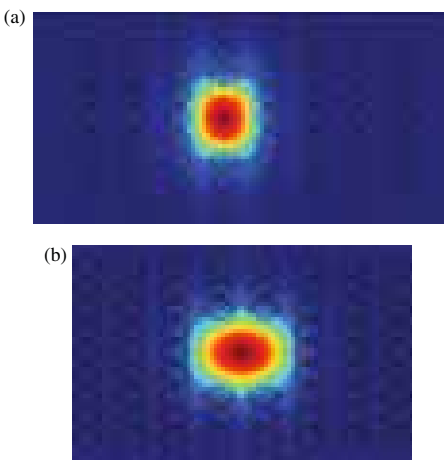


Fig. 2. Optical field distributions for the fundamental mode in the coupled fiber structure.

Where n_{eff}^{x} is the effective refractive index of the x-polarized fundamental mode and n_{eff}^{y} is the effective refractive index of the y-polarized fundamental mode.

3. SIMULATION RESULT ANALYSIS

We have carried out simulation to verify the effect of ellipticity on birefringence of the structures of single defect and double defect PCF using FEM method. Here we introduced the excitation at the center of the structure in both the cases. This birefringence PCF is formed by a regular rectangle lattice of elliptical air holes on silica, however the first structure is single defect (three missing air holes) PCF and second one is double defect (adjacent missing selected air holes) PCF.

The linear property in the birefringence versus wavelength plot in a PCF structure depends on the shape of the structure, such as pitch, refractive index (n) and ellipticity ratio ($\eta = r_x/r_y$) where r_x and r_y are the major and minor axes of elliptical holes are directed along x axes respectively.

Figure 2 shows the electric field distribution of the PCF. We have measured ellipticity in three different conditions such as $\eta = 1.5$, $\eta = 2$, and $\eta = 2.5$ which correspond to $x = 0.3$, $y = 0.2$; $x = 0.4$, $y = 0.2$; and $x = 0.5$, $y = 0.2$ respectively. The modal effective index difference x and y polarized fundamental modes are simulated at wavelength from 1.2 μm to 1.7 μm.

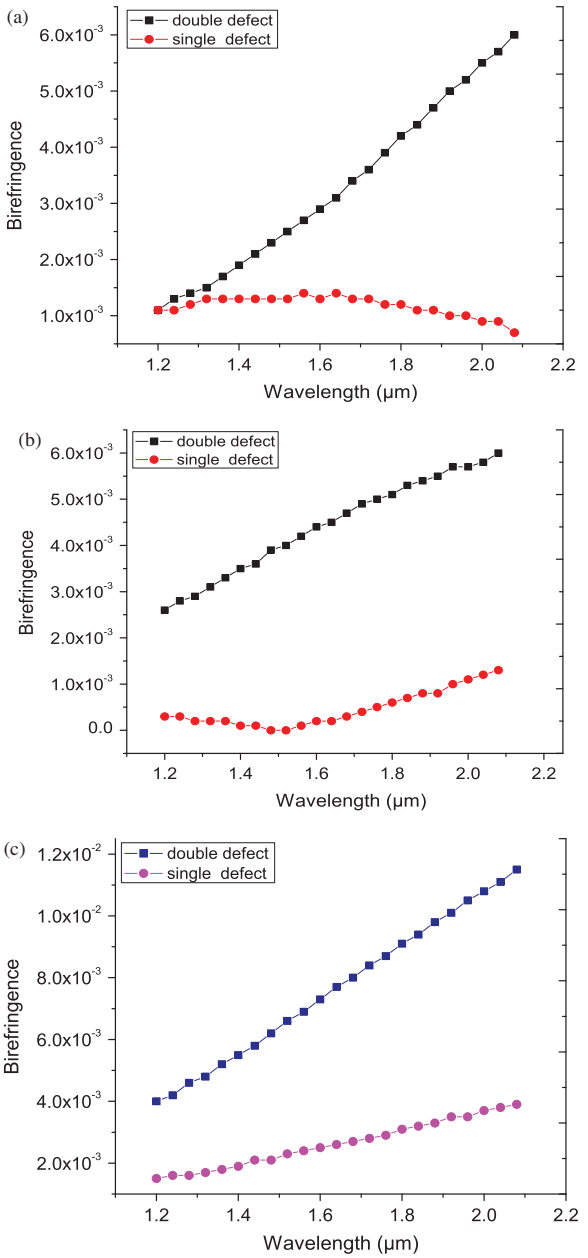


Fig. 3. Dependence of birefringence on wavelength for elliptical air hole PCFs with the lattice constant = 1.2 μm and ellipticity of (a) 1.5 (b) 2 (c) 2.5 for both single defect and double defect.

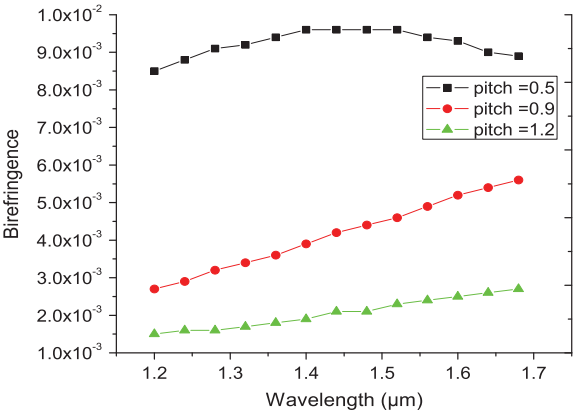


Fig. 4. Dependence of pitch on birefringence for the double defect structure.

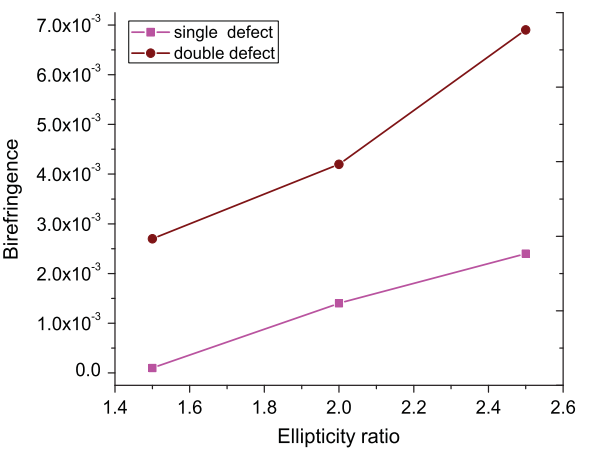


Fig. 5. Birefringence of the fundamental mode for both the defect elliptical hole PCFs with pitch = 1.2 μm at =1.55 μm.

Table I. Comparison of earlier reported birefringence PCF with the here proposed rectangle PCF with elliptical air holes.

Title	Author	Structure	Birefringence
Design of highly birefringent chalcogenide glass PCF: A simplest design (2011)	Bhawana Dabas et al.	Hexagonal	2.2×10^{-3} at 1.55 μm
Analysis of non-linear PCF for birefringence application using FDTD method (2014)	N. Muduli et al.	Hexagonal	30.63×10^{-5} at 1.55 μm
Highly birefringent non-linear PCF for optical sensing of analytes in aqueous solutions (2016)	Huseyin Ademgil et al.	Hexagonal	0.68×10^{-3} at 1.55 μm
A reflective temperature-insensitive all-fiber polarization mode interferometer and pressure sensing application based on PM-PCF (2016)	Zhen Yin et al.	Hexagonal	9.7×10^{-4} at 1.2 MPa
Our proposed PCF		Rectangle	9.2×10^{-3} at 1.55 μm

Figure 3(a) shows the variation of index difference of PCF structure with respect to wavelength of input light constant to ellipticity ($\eta = 1.5$) for both single defect and double defect. It is found that index difference varies almost linear with respect to wavelength of light for both the cases (single defect and double defect).

A strange variation however seen from the Figure 3 i.e., at lower wavelength, the index difference of double defect PCF is slightly greater than the single defect PCF. But for wavelengths greater than 1.5 μm, the variation of index difference of single defect PCF is faster than the double defect PCF. This variation leads to index difference of single defect PCF to be larger than the double defect PCF. It is also observed that the difference of index difference between single defect and double defect PCF increases at a faster rate with increase in wavelength. Further it is found that, the single defect PCF shows more birefringence behavior than the double defect PCF.

Simulations are also made to verify birefringence in PCF for ellipticity ratio ($\eta = 2$ and $\eta = 2.5$) result is shown in Figures 3(b) and (c). From the figure, the variation of index difference is nearly linear with respect to wavelength; however the index difference of single defect is always higher than the double defect PCF. While considering the ellipticity of $\eta = 1.5$, there is intersection between the plots at wavelength of 1.2 μm additionally the birefringence property of double defect is slightly greater than the single defect.

As birefringence is a function of structural configuration such as lattice constant, major and minor radius of air holes and refractive index of background material, we have also thorough the simulation using FEM method to find out the index difference with respect to ellipticity ($\eta = x/y$) at the wavelength of 1.5 μm. Figure 4 shows simulation result for the variation of index difference with respect to the lattice constant (pitch) at the wavelength of 1.5 μm. Figure 5 shows the ellipticity ratio effect on birefringence for the single and double defect structure. The difference of index difference between double defect and single defect PCF varies with respect to ratio of ellipticity. Comparison of earlier reported birefringence PCF with the proposed rectangle PCF is tabulated in Table I.

4. CONCLUSION

FEM simulation is used to investigate the effect of ellipticity and pitch on the birefringence in a rectangle photonic crystal fiber having elliptical air holes with single and double defect. It is found that the birefringence property in rectangle structure PCF depends on structural configuration, where ellipticity plays a vital role. The double defect structure is found to have more birefringence than the structure with single defect. At the higher values of ellipticity, birefringence due to the double defect is more than the single defect. These results appear to be important as the birefringence property can be controlled by suitably adjusting the structure and adjust the parameters in the air holes of the PCF.

References

1. J. C. Knight, *Nature* 424, 847 **(2003)**.
2. P. Russell, *Science* 299, 358 **(2003)**.
3. P. S. J. Russell, *J. Lightwave Technol.* 24, 4729 **(2006)**.
4. M. Piliarik and J. Homola, *Opt. Express* 17, 16505 **(2009)**.
5. M. A. Schmidt, L. N. P. Sempere, H. K. Tyagi, C. G. Poulton, and P. S. J. Russell, *Phys. Rev. B, Condens. Matter* 77, Article ID: 033417 **(2008)**.
6. Vinod Kumar Singh and S. S. Mishra, Designing of endlessly single-mode highly polarization maintaining birefringent photonic crystal fiber with low confinement loss at wavelength 1.55 μm, *Symposium on Photonics and Optoelectronics (SOPO)*, May **(2011)**.
7. Ritu Sharma, Vijay Janyani, and Anuradha Sharma, *International Journal of Computer Science and Emerging Technologies (E-ISSN: 2044-6004)* 2, 238 **(2011)**.
8. K.-Y. Yang, Y.-F. Chau, Y.-W. Huang, H.-Y. Yeh, and D. P. Tsai, *J. Appl. Phys.* 109, 093103 **(2011)**.
9. C. Wu, J. Li, X. Feng, B.-O. Guan, and H.-Y. Tam, *Journal of Lightwave Technology* 29, 943 **(2011)**.
10. Y. Yue, G. Y. Kai, Z. Wang, T. T. Sun, L. Jin, Y. F. Lu, C. Zhang, J. Liu, Y. Li, Y. Liu, S. Yuan, and X. Dong, *Opt. Lett.* 32, 469 **(2007)**.
11. A. Ortigosa-Blanch, J. C. Knight, W. J. Wadsworth, J. Arriaga, B. J. Mangan, T. A. Birks, and P. St. J. Russell, *Opt. Lett.* 25, 1325 **(2000)**.
12. M. Tanaka and M. Fujita, *Opt. Express* 9, 676 **(2001)**.

Received: 20 April 2018. Accepted: 28 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1598–1602, 2018

RESEARCH ARTICLE

Health Monitoring Framework for in Time Recognition of Pulmonary Embolism Using Internet of Things

D. Rajesh Kumar*, T. Akash Krishna, and Amitabh Wahi

Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam 638401, India

Great economy implies great social insurance. Be that as it may, current human services framework confront colossal difficulties, for example, maturing populaces, development of perpetual disease and utilization of well expensive vital health technologies. Early detection of various disease can help the economy to overcome vital health issues. Many diseases are completely curable if they are detected in the initial stage so that pre detection of diseases plays a crucial role in healthcare. Pulmonary embolism being a major health issue affects more than 5 million people worldwide. Essentially it is caused by long periods of latency, gravity makes blood stagnate in the most reduced areas of the body, which may prompt a blood coagulation. It could happen if individuals sitting for an extensive excursion or lying in bed recuperating from a sickness and particularly for individuals in coma state. By the progression of innovation, a conspicuous arrangement can be drawn out through the idea of Internet of Things. IoT being a developing proficient innovation, a gadget that can early identify Pulmonary embolism can be made out which unmistakably recognize the patient has a possibility of this illness.

Keywords: Early Detection, Pulmonary Embolism, Coma State, Internet of Things.

1. INTRODUCTION

Late investigations demonstrates that every year more than 5 million peoples are influenced by pulmonary embolism. Essential care is normally the primary purpose of contact for a patient. Essential care is regularly given by general expert's family specialists, dental specialists, drug specialists, maternity specialists, and so on. It is the place most safeguard wellbeing can be accomplished and where early analysis can be conceivable, which may avert more costly healing center treatment being required. Current economy being a beneficial one, they generally have less rate of work utilizing physical diligent work. This makes every person to stay stationary at a place or work under such conditions. Great Health implies mentally and physically a man should carry on with his life minus all potential limitations. Be that as it may, individuals who are not ready to accomplish great wellbeing must be minded more as their life is futile. Individuals typically in coma state don't comprehend the end result for their body for quite a while. Inward bleedings can happen of which pulmonary embolism a noteworthy issue is looked by individuals in coma state and also individuals whose way of life rely upon being stationary for a long time period. Early identification of pulmonary embolism can be illuminated through straightforward medicine and can be kept from getting complicated. Basically pulmonary embolism is a blockage of a supply route in the lungs by a substance that has moved from somewhere else in the body through the circulatory system. Manifestations may include shortness of breath, chest pain especially after breathing in, and coughing up. Signs of a PE incorporate low blood oxygen levels, quick breathing, and fast heart rate with a mild fever. Keeping in mind that the end goal to settle this is by utilizing the concept of IoT.

Internet of Things being a developing innovation, it act as prominent gateway for some advancements that rebuilds regular day to day existence. Sensors are driving a generally new tech slant in the Internet of Things. Basically, the "things" referenced in this clumsy trendy expression are machines implanted with sensors that accumulate, store and investigate information. Furthermore, since they're altogether connected to the Internet, they can transfer that information for additionally preparing, download refreshed programming and frequently be controlled from a far distance. Indeed, even with prickly issues approaching, these astute machines are as of now adjusting circles as differing as healthcare, city arranging, transportation and power age, farming and family administration. The gadgets themselves may be small scale, yet they're causing large scale moves by the way we live and work.

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1598/005 doi:10.1166/jctn.2018.7347

For such a condition, I thought of incorporating a gadget to early detect pulmonary embolism will screen the general movement by observing the heart rate, inhale rate, blood oxygen level and in addition body temperature. The gadget computes the activity of the individual and stores the acquired an incentive to the cloud server. The values are then moved to a mobile application which is arranged with a setting of indicating status of the disease. So by this the user can give satisfactory health care to individuals and thus detect pulmonary embolism at the initial stage.

2. EXISTING SYSTEM

As indicated by my revelation the current analytic apparatuses are of high cost and they are difficult to access from anyplace whenever. Ultrasound test, Computed tomography examine test, Lung Ventilation/Perfusion Scan, Pulmonary Angiography and through blood test we can analyze pneumonic embolism. As expressed these test are of high cost, so keeping in mind the end goal to identify PE rather from these test the gadget that will be made utilizing IoT must be costless and it ought to give precise possibility of event. Existing advancements in IoT demonstrates the likelihood of getting parameters such as heartbeat rate, ECG rate and additionally body temperature. As pulmonary embolism rely upon these variables it's conceivable to get the qualities utilizing these sensor innovation and identify the infection. Other than these parameters we can get blood oxygen level and in addition the inhale rate. Combining all these values and through constant analysis we can detect the possibility of Pulmonary Embolism.

The survey by Riazul Islam, Kwak, Humaun Kabir, Mahmud Hossain, and Kwak[1] they cited the potential outcomes of IoT innovation to be executed on healthcare. IoT being a creating innovation numerous underlying parameters in the human body can be recorded and in view of this we can ascertain the conceivable ailments that will happen based on these qualities. The paper likewise clarifies the conceivable strategies for putting away information that got from these sensors. Information putting away is a vital perspective as these information readings can be utilized for future determination. As innovation propels these parameters can likewise be perused from a wrist watch which later created to a smart watch. Security issues are additionally talked about in the paper with the goal that the information exchange ought to must be secured as they can additionally make data loss or some other entangled issues.

Natarajan, Prasath, and Kokila[2] they presented new conceivable parameters that can be recorded utilizing IoT. Pulse Oximeter which can give Blood Oxygen rate is another parameter we can get through IoT innovation. Glucose level monitoring is another strategy we can embrace utilizing IoT that gives the present glucose level. Along these lines headway in innovation regularly makes part of chances that can be executed in healthcare industry. Social insurance being an essential field every parameter engaged with it are important to the point that it analyze different illnesses from the underlying stage itself.

Park and Cho[3] clarifies the accuracy and the possibility of long term data monitoring in an IoT project. Development of innovation profoundly brings about difference in a few parameters, as in IoT there happens an adjustment in values we acquire from the sensors. Sensor perusing can cause a conspicuous change in IoT gadgets as a few elements are relying upon it. So in this paper they unmistakably expresses that there can be constrained measure of exactness to the readings we get in each cycle. Additionally the qualities we get from the sensor that we store in different server should be exact and they ought to likewise have a probability of utilization later on.

In Estimation by Nagadeepthi, Sushrutha Bhardwaj and Pushpalatha[4] they effectively acquainted an arrangement of calculation with locate the respiratory rate from ECG. By observing ECG rate we can draw breathe in rate from the readings we quantified by the ECG sensor and the advanced information is perused with the assistance of graphical UI software. The information is put away in an exhibit and the QRS tops every moment are distinguished and heart rate is computed. As these QRS crests comprise of respiratory data, a calculation will be connected onto the QRS information to locate the quantity of inclines every moment, which gives the respiratory rate. Subsequently the Heart Rate and Respiratory Rate every moment will be ascertained and shown continuous on PC.

In Ref. [5] cites the likelihood and the side effects of pulmonary embolism. Rapid change in pulse rate, blood oxygen level and in addition the breath rate joined by mild fever. Early recognition will be an aid to patients mostly in extreme coma and also individuals whose work culture makes them to be sit still for a drawn out stretch of time. So all things considered it is anything but difficult to identify the sickness as opposed to getting muddled. Other demonstrative strategies existing are of high cost and they are not doable to be utilized at all over the place.

Reference [11] specifies the importance to predict medical readings by the use of Naive Bayesian Classifier so that accurate results are obtained and is easy to implement. Bayesian methodologies are an in a general sense imperative system. Given the likelihood circulation, Bayes classifier can provably accomplish the ideal outcome. Bayesian technique depends on the likelihood hypothesis. Bayes Rule is connected here to ascertain the back from the earlier and the probability, in light of the fact that the latter two is by and large less demanding to be figured from a likelihood demonstrate. Measurements give a solid major foundation to measurement and assessment of results. Nonetheless, calculations in light of measurements should be adjusted and scaled before they are connected.

Reference [12] depicts the possibility of healthcare monitoring using Bayesian classifier so that it's easy to classify

and obtain a specific result. A few doctor's facilities utilize choice emotionally supportive systems, however they are to a great extent constrained. Clinical choices are regularly made in view of specialists' instinct and experience instead of on the information rich information covered up in the database. This training prompts undesirable predispositions, blunders and unreasonable medicinal costs which influences the nature of administration gave to patients.

3. PROPOSED SYSTEM

The principle point of this gadget is to give early identify pulmonary embolism in an individual. Patients in coma state will be profited more as they are obscure to the end result for their body or to individuals who are in paralyzed state. At the underlying stage there happens a change in indispensable body parameters such as pulse rate, inhale rate, blood oxygen level which is accompanied with a mild fever. The gadget will have a heartbeat rate sensor, ECG sensor, pulse oximetry sensor and additionally a body temperature sensor. The gadget is intended to have full usefulness that it ought to identify the likelihood of pulmonary embolism in people. The gadget is associated with an application where the recorded esteems from the sensor are shown, which later after figuring give correct outcome whether infection is available or not. In the middle of the gadget and application a cloud server is associated which will store the data from the sensor that can be utilized for future finding.

4. ARCHITECTURE

Figure 1 represents the overall architecture of the system.

1. *Pre-detection Device*—The gadget comprises of the considerable number of sensors that catches the required parameters. It likewise have yield ports that comprise of wired closures used to identify the parameter from the body as they are associated with body. The Screen joined will demonstrate the present preparing status.
2. *Processing Screen*—The screen will update the present status of the gadget as it take parameter perusing from the individual. On the off chance that there happens any perusing disappointment or any issue happens it is told on the screen.
3. *Power Button*—As a typical power button it plays out the ON and OFF function for the gadget. Push button control switch is utilized as it is anything and can spare power. The power button power up the whole device so that the respective sensors will perform their function.
4. *Data Transfer to the Cloud*—The data from the device is sent to a cloud server where data can be stored as well as transferred for further calculation to the mobile application attached to it. The transfer of data will be highly secured as no data loss will occur.
5. *Cloud Server*—The data transferred from the device is stored as well as transferred to the mobile application for further calculation. The values can be further obtained for future diagnosis.
6. *Data Transfer to the Mobile Application*—Data from the Cloud is further transferred to the mobile application. As in other end this data transfer will also be secure and no data loss will occur.
7. *Mobile Application*—Data from the cloud is exchanged to the Mobile Application. The application ought to be able to ascertain the sensor readings and yield the outcome as there happens any shot of pulmonary embolism for the particular individual under checking.

Fig. 1. Detailed architecture of the system.

4.1. Detailed Working Mechanism

Individual to be examined are checked with this gadget. The gadget will have 4 yield strings in which 3 are associated with two segments of heart and in addition to right half of the guts. The fourth port is associated with the forefinger in order to get the body temperature and blood oxygen level. The screen joined will demonstrates every single current status of the gadget and its operation. The gadget is presently changed to ON state by pressing the power button which makes the screen appended to the gadget to yield that machine is on working state and prepared for conclusion. The sensors joined in the gadget will note down the ECG rate, Blood oxygen level, pulse rate, inhale rate and in addition body temperature which are additionally exchanged to the mobile application associated with it. The values are first exchanged to a cloud server from that point to the versatile application. The cloud will go about as medium for information stockpiling and arbiter for the information exchange. The portable application can be either on android or on iOS application ought to have an interface that can figure the sensor values and give a conspicuous yield that the coveted individual has pulmonary embolism or not. Last outcome is yielded at application which can be additionally recorded for future purposes.

Readings from the sensor must be taken care of in a way that it need to associated with the conceivable readings which will cause the illness. Possible readings can be obtained in any manner such that each and every reading should have to be compared with the possibility index chart. The index chart contains possible readings of the factors that create pulmonary embolism.

Figure 2 represents the overall vital signs that pulmonary embolism shows if a person is affected by it. The IoT device created here will provide a fixed values to be

Index chart of vital signs	
Vital signs	Rate change
Pulse rate	>100 beats per min
Breathe rate	>20 breathe per min
Blood oxygen level	<40 mm Hg
Body temperature	>100 degree Celsius

Fig. 2. Index chart of vital signs.

compared with the possible index chart. Values after comparison only will depict a final result so that possibility of having disease can be accurate. Since it's a live detection procedure depicting with a possible data reading of a patient can be induced so that the whole procedure can be discussed well. A person approaches to test for Pulmonary Embolism in a way that he connected the device to his body and the device starts to analyze the parameters for the disease. After analyzing the readings are passed to cloud for storage, then from there it is then moved to application where we get the ultimate result. The result received will be gone through a series of comparison with the index table so that we can assure the possibility of disease is there or not.

Technically the readings are compared and given the result is by using a classification algorithm. The algorithm is deeply used for health prediction so that a possible outcome can be obtained. Naive Bayesian Classifier can be induced so that predicting the disease will be accurate by handling the obtained data from the device. In machine learning, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong independence assumptions between the features. Naive Bayes has been studied extensively since the 1950s. It was introduced under a different name into the text retrieval community in the early 1960s, and remains a popular method for text categorization, the problem of judging documents as belonging to one category or the other with word frequencies as the features. With appropriate pre-processing, it is competitive in this domain with more advanced methods including support vector machines. It also finds application in automatic medical diagnosis. Bayes Rule is connected here to figure the illness from the earlier esteems and the probability, in light of the fact that the later two is by and large less demanding to be computed from a likelihood display. Measurements give a solid central foundation to measurement and assessment of results. A sample reading by the device get processed by the Bayes algorithm and provides a sample decision.

Figure 3 represents the calculated graph obtained by observing the values from the sensor which is saved on the cloud. Naive Bayesian Classifier will sort out the obtained value from the device and then calculate the possibilities. Prescient examination being a key part of machine learning calculations, enables clients to settle on improved and

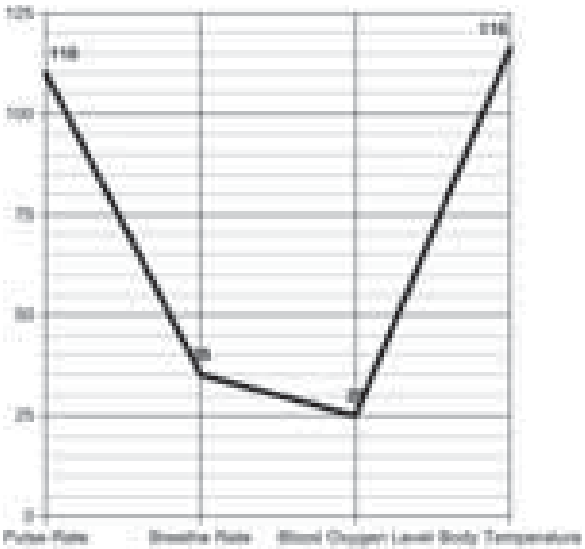


Fig. 3. Calculated graph.

administered choices. Visual Analytics is an instrument to cost-successfully sort the rich and quickly increasing information in the field of therapeutic research. It causes us adapt to the varying information in a sorted out way, which the human cerebrum would have the capacity to picture effortlessly. This would thusly give new creative and potential outcomes. This examination gives organized information as well as starts organized musings in the psyche of people. As professionals investigate certain bizarre circumstances, the procedure of visual examination process would give arranged applicable information identified with it. This would thus diminish the cost of keeping up gigantic measure of information.

5. CONCLUSION AND FUTURE WORKS

The primary favorable position drawn out from this gadget is to early recognize pulmonary embolism with the goal that further treatment can be embraced by the patients. More than 5 million individuals pass on consistently because of pulmonary embolism, in this manner it is very vital to identify it from the underlying stage itself. From the analysis chart we can close the conceivable general changes that happen in the key parameters for a person.

Pulmonary embolism is a typical complexity of hospitalization and adds to 5 to 10 percent of passing in hospitalized patients, making it one of the main sources of preventable healing center passing. In spite of it being a huge medical issue, the genuine occurrence of pulmonary embolism is questionable. The determination of pneumonic emboli can be troublesome and requires particular imaging methods that are not accessible in all doctor's facilities or human services settings. The evaluated rate of analyzed Pulmonary embolism is 71 to 117 for each 100,000 man years, however the genuine frequency is

probably going to be substantially more than this rate since ponders demonstrate that for each instance of analyzed, non-deadly pulmonary embolism, there are 2.5 instances of lethal pulmonary embolism analyzed simply after death. Different examinations have evaluated that more than five million individuals are influenced by pneumonic embolism every year, with 100,000 to 200,000 of these occasions being lethal. Over portion of all analyzed instances of aspiratory embolism on the planet happen in patients in healing centers or nursing homes. Pulmonary embolism has earned the notoriety of a quiet executioner on the grounds that not as much as half of patients who bite the dust of pulmonary embolism were determined to have the issue preceding passing.

Mortality from untreated pulmonary embolism was said to be 26%. This figure originates from a trial distributed in 1960 by Barrit and Jordan, which analyzed anticoagulation against fake treatment for the administration of pulmonary embolism. Barritt and Jordan played out their investigation in the Bristol Royal Infirmary in 1957. This examination is the main fake treatment controlled trial to look at the place of anticoagulants in the treatment of Pulmonary embolism, the consequences of which were convincing to the point that the trial has never been rehashed as to do as such would be viewed as unscrupulous. So, the revealed death rate of 26% in the fake treatment gather is most likely an exaggeration, given that the innovation of the day may have identified just extreme pulmonary embolism.

As technology advances healthcare becomes more reliable than it was before. Health Information Technology is one of the mechanism in with the use of data preparing including both hardware and software that arrangements with the capacity, recovery, sharing, and utilization of medicinal services data, wellbeing information, and learning for correspondence and choice making. Technology is an expansive idea that arrangements with an animal categories' use and information of devices and artworks, and how it influences an animal types' capacity to control and adjust to its condition. Be that as it may, a strict definition is slippery. "Innovation" can allude to material objects of utilization to mankind, for example, machines, equipment or utensils, however can likewise incorporate more extensive subjects, including frameworks, strategies for association, and methods. Informatics alludes to the study of data, the act of data handling, and the building of data frameworks. Informatics underlies the scholastic examination and professional utilization of figuring and correspondences innovation to human services, wellbeing instruction, and biomedical research. Wellbeing informatics alludes to the convergence of data science, software engineering, and medicinal services. Wellbeing informatics portrays the utilization and sharing of data inside the human services industry with commitments from software engineering, arithmetic, and brain science. It manages the assets, gadgets, and techniques required for upgrading the securing, stockpiling, recovery, and utilization of data in wellbeing and biomedicine. Wellbeing informatics devices incorporate PCs as well as clinical rules, formal medicinal phrasings, and data and correspondence frameworks. Therapeutic informatics, nursing informatics, general wellbeing informatics, drug store informatics, and translational bioinformatics are sub trains that illuminate wellbeing informatics from various disciplinary perspectives.

The gadget can likewise be additionally extended as innovation propels. Instead of observing crucial parameters as the innovation progresses we can make simple strategy to discover the ailment either by filtering the chest region. Primary maxim we can draw in from IoT is to make something that is helpful to common natives which makes their standard work less complex.

References

1. S. M. Riazul Islam, D. Kwak, M. D. Humaun Kabir, Mahmud Hossain, and K.-S. Kwak, The Internet of Things for Health Care: A Comprehensive Survey, IEEE **(2015)**.
2. K. Natarajan, B. Prasath, and P. Kokila, *Journal of Network Communications and Emerging Technologies* 2016, 6 **(2016)**.
3. D. Park and J. Cho, *The Scientific World Journal* 1 **(2014)**.
4. M. R. Nagadeepthi, M. Sushrutha Bhardwaj, and B. G. Pushpalatha, *International Journal of Electronics Signals and Systems* 2014, 1 **(2014)**.
5. F. Lavorini, V. DiBello, M. L. DeRimini, G. Lucignani, L. Marconi, G. Palareti, R. Pesavento, D. Prisco, M. Santini, N. Sverzellati, A. Palla, and M. Pistolesi, *MRM Journal* 8, 75 **(2013)**.
6. C. V. Pollack, D. Schreiber, S. Z. Goldhaber, D. Slattery, J. Fanikos, B. J. O'Neil, . . . , J. A. Kline, *Journal of the American College of Cardiology* 57, 700 **(2011)**.
7. A. Yogaraj, M. R. Ezilarasan, R. V. Anuroop, C. S. Sivanthiram, and Sunil Kumar Thakur, "IOT Based Smart Healthcare Monitoring System for Rural/Isolated Areas," IJPAM **(2017)**.
8. M. Miniati, C. Cenci, S. Monti, and D. Poli, *PLoS One* 7 **(2012)**.
9. B. Morici, *Journal of the American Academy of Physician Assistants* 27, 18 **(2014)**.
10. Q.-Y. Ji, M.-F. Wang, C.-M. Su, Q.-F. Yang, L.-F. Feng, L.-Y. Zhao, S.-Y. Fang, F.-H. Zhao, and W.-M. Li, *Scientific Reports* 7 **(2017)**.
11. R. Bhuvaneswari and K. Kalaiselvi, *International Journal of Computer Science and Telecommunication* 3 **(2014)**.
12. S. Indhumathi and G. Vijaybaskar, *International Journal of Advanced Research in Computer Engineering and Technology* 4 **(2015)**.
13. D. Rajesh Kumar, A. Shanmugam, C. Palanisamy, and A. M. Natarajan, *Asian Journal of Research in Social Sciences and Humanities* 6 **(2016)**.

Received: 20 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1603–1606, 2018

Sentiment Analysis on Multi Domain Reviews Using Machine Learning Methods

N. Keerthini[1,*] and P. C. D. Kalaivaani[2]

[1] *PG Scholar, Department of Computer Science and Engineering, Kongu Engineering College, Perundurai 638052, India*
[2] *Department of Computer Science and Engineering, Kongu Engineering College, Perundurai 638052, India*

Due to increase in usage of social network and online promoting sites, ton of reviews were obtained. The reviews will be collected and given input to the proposed approach for making qualitative decisions. The proposed approach classifies the reviews into either positive or negative polarity. Supervised machine learning strategies helps to classify these reviews with improved accuracy. In this proposed approach, four supervised machine learning algorithms such as Naive Bayes (NB), Maximum Entropy (ME), Stochastic Gradient Descent (SGD), and Support Vector Machine (SVM) have been used for classification of sentimental reviews. *N* grams and its combination were considered for classification. The accuracy of different strategies were examined so as to access their performance.

Keywords: Naive Bayes (NB), Support Vector Machine (SVM), Maximum Entropy (ME), Stochastic Gradient Descent (SGD).

RESEARCH ARTICLE

1. INTRODUCTION

Data mining is the method of analyzing or extracting hidden patterns of information, according to different views for categorization into useful information. It is collected and assembled in data warehouses for efficient business decision making. Data mining is referred to as knowledge discovery and data discovery.

Sentiment analysis, also called opinion mining, analyzes people's opinion towards entities like products, organizations, and their associated attributes. In order to derive significant data from people's sentiments, different machine learning techniques were applied.

There are primarily two kinds of machine learning techniques in sentiment analysis, i.e., the technique based on supervised and unsupervised learning. In supervised learning technique, the dataset is labeled and trained to get a reasonable output that helps in proper decision making. Unsupervised learning method does not need any label information to resolve the problem of processing of unlabeled data, clustering algorithms are used.

Sentiment analysis is observed to be applied in three different levels such as document level, sentence level, and aspect level. Document level classifies whether the document's opinion is positive, negative or neutral. Sentence level determine whether the sentence expresses any negative, positive or neutral opinion. Aspect level focuses on all expressions of sentiments present within given document.

2. RELATED WORK

Feldman[1] has proposed classification algorithm used to classify sentiment analysis methods. The vector representations of text can be used for performance-wise classification process. It is most competitive approach and discussed in recent advances in sentiment analysis. The analysis of the sentiment conveyed by a piece of natural language text is guided by the text's structure.[1]

Gautam and Yadav[2] has proposed Sentiment analysis of twitter data using machine learning approaches and semantic analysis. From pre-processed dataset, adjective were extracted that have some meaning which is called feature vector, then selected the feature vector list and thereafter applied machine learning based classification algorithms. The performance of classifier was measured in terms of recall, precision and accuracy.[2]

Zhang et al.[3] has proposed the classification of Chinese comments based on word2vec and SVM^{perf}. It is based on word2vec tool to cluster similar features and POS based feature selection approach to generate the training data. Word2vec tool adopts continuous bag of words model and continuous skip-gram model to learn the vector representation of words. SVM^{perf} is an implementation of SVM for

*Author to whom correspondence should be addressed.

multivariate performance measures that follows an alternative structural formulation of SVM optimization problem for binary classification.[3]

Niu et al.[4] have proposed a multi-view sentiment analysis (MVSA) dataset, including a collection of image-text pair with manual annotation collected from Twitter. Their approach of sentiment analysis will be classified into two components, i.e., lexicon based and statistic learning. In case of lexicon based analysis, a set of opinion words or phrases are considered which have predefined sentiment score. While in statistic learning, various machine learning techniques are used with dedicated textual features.[4]

3. PROBLEM DESCRIPTION

Most of the authors have used unigram approach to classify the reviews. Unigram provides comparatively better result, but fail in some cases. The review "The pen is not good," when analyzed using unigram method, provides the polarity of sentence as neutral with the presence of one positive polarity word 'good' and one negative polarity word 'not.' But when the statement is analyzed using bigram approach, it gives the polarity of sentence as negative due to the presence of words 'not good,' which is correct.

Machine learning algorithms work on the data represented as matrix of numbers. But the sentiment data are always in text format. So, it needs to be converted to number matrix. In order to convert the text data into matrix of numbers, TF-IDF have been applied. The rows of the matrix of numbers represents a particular text file where as its column represent each word or feature present in that respective file.

Different machine learning algorithms were proposed for the classification of IMDb and multidomain dataset using n-gram techniques such as Unigram, Bigram, Trigram, combination of unigram and bigram, bigram and trigram, and unigram and bigram and trigram. Four different machine learning techniques such as Naive Bayes, Maximum Entropy, Support Vector Machine, and Stochastic Gradient Descent were used for classification.

The performance of the machine leaning algorithms was evaluated using parameters like precision, recall, f-measure, and accuracy. The results obtained shows better accuracy.

4. PROPOSED METHODOLOGY

4.1. Dataset

The acl Internet Movie Database (IMDb) dataset is considered for sentiment analysis. It consists of 1000 positively labeled reviews, 1000 negatively labeled reviews. Also considered multidomain dataset, which consists of 1000 positively labeled and 1000 negatively labeled reviews.

4.2. Data Preprocessing

Data preprocessing involves transforming raw data into an understandable format. It is used to remove irrelevant information in reviews. The text reviews sometimes consist of absurd data, which need to be removed, before considered for classification. The usually identified absurd data are Stopwords, Numeric and Special characters. Figure 1 represents workflow of proposed system.

Assigning automatic descriptors to the given token is known as tagging. The descriptors are referred as tag. POS Tagger is used to assign part of speech to the words in the document. Parts of speech includes nouns, verbs, adverbs, adjectives, pronouns, conjunctions and their other subcategories.

N-gram is the continuous sequence of n items from a given sequence of text or speech. For unigram $N = 1$, a single word was considered. A pair of words was considered for bigram where $N = 2$. Trigram refers to set of words having count equal to three.

Term Frequency calculates how frequent a term appears in a document. It is possible that a term would occur much more times in long documents than shorter ones. Thus, the term frequency is divided by the document length or the total number of terms within the document.

$$\mathrm{TF}(t) = (\text{Number of times term } t \text{ occurs in a document}) / (\text{Total number of terms in the document}) \quad (1)$$

Inverse Document Frequency measures the rarity of a term in the whole corpus. If a term occurs in all documents of the collection, its idf is zero. But it is identified that certain terms, like "is," "of," and "that," could appear a lot of times but have little importance.

$$\mathrm{IDF}(t) = \log\_e (\text{Total number of documents} / \text{Number of documents with term } t \text{ in it}) \quad (2)$$

Term frequency-Inverse document frequency, is a numeric value which is used to calculate the importance of a word in a corpus based on how often it appears in that document and in a given collection of documents. The tf-idf of a term is the product of its tf and its idf.

$$\text{TF-IDF} = \mathrm{TF}(t) * \mathrm{IDF}(t) \quad (3)$$

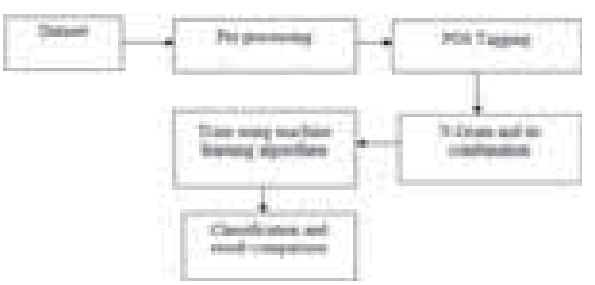


Fig. 1. Proposed methodology.

4.3. Machine Learning Based Methods

After the text reviews are converted to matrix of numbers, then those matrices were considered as input for the following four different supervised machine learning algorithms for classification.

4.3.1. Naive Bayes Method

Using probabilistic classifier and pattern learning, the set of documents were classified. It assumes that all features which are identified in the dataset were mutually independent of each other. It works based on Bayes rule. This model considers word frequency information in document for analysis, where a document is considered to be an ordered sequence of words. The probability of a word event is independent of word context and its position in the document. After estimating the parameters calculated from training document, classification process is carried out on text document by calculating posterior probability of each class and selecting the highest probable class.

4.3.2. Maximum Entropy Method

The training data are used to set constraint on conditional distribution. Each constraint is used to express characteristics of training the data. These constraints then are used for testing the data. In order to use ME, a set of features that is needed to be selected. For text classification purpose, word counts are considered as features and if a word occurs frequently in a class, the weight of word-class pair becomes higher in comparison to other pairs. These highest frequency word-class pairs are considered for classification purpose.

4.3.3. Stochastic Gradient Descent Method

SGD method is used when the training data size is mostly large in nature. Each iteration estimates the gradient on the basis of single randomly picked example. The convergence of SGD gets effected by the noisy approximation of the gradient. If learning rate decreases slowly, the parameter estimate decreases equally slowly; but if rate decreases too quickly, the parameter estimate takes significant amount of time to reach the optimum point.

4.3.4. Support Vector Machine Method

Data are analyzed and decision boundaries are defined by having hyper planes. In two category case, the hyper plane separates the document vector of one class from other classes, where the separation is maintained to be large as possible. Since SVM requires input in the form of a vector of numbers, the reviews of text file for classification need to be converted to numeric value. After the text file is converted to numeric vector, it may go through a scaling process, which helps to manage the vectors and keep them in the range of [1, 0].

Table I. Confusion matrix.

	Predicted positive	Predicted negative
Actual positive	TP	FN
Actual negative	FP	FN

5. EXPERIMENT AND RESULTS

Collections of movie-review documents labeled with respect to their overall sentiment polarity were taken for analysis. Dataset consists of 1000 positive and 1000 negative processed reviews. Based on true positive, true negative, false positive and false negative rates, accuracy were calculated by forming confusion matrix. It represents the values of TP, TN, FP, FN considering actual positive-predicted positive, actual negative-predicted negative, actual negative-predicted positive and actual positive-predicted negative respectively. Table I is the confusion matrix.

Accuracy is the measure of correct prediction of the classifier compared to the overall data points. It is the ratio of correct predictions to the total number of predictions made by the classifiers.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{4}$$

The confusion matrix obtained for unigram Naive bayes will be as follows,

Confusion Matrix

	Positive	Negative
Positive	53	22
Negative	36	39
Accuracy = 61.33		

Similarly, accuracy have been calculated for other machine learning strategies and *n* gram combinations. Figure 2 shows the accuracy results of movie review dataset. The results obtained shows that Support vector machine gives better performance over others. It also shows that unigram have better accuracy than other

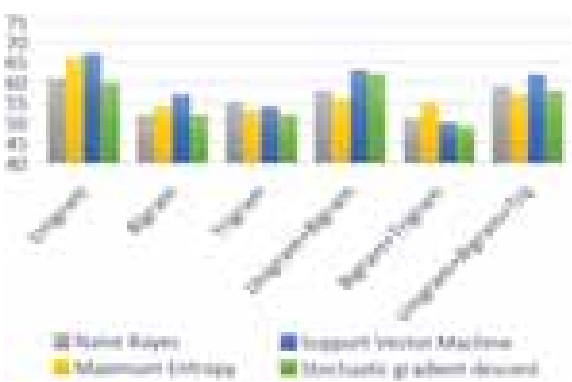


Fig. 2. Accuracy of movie review dataset.

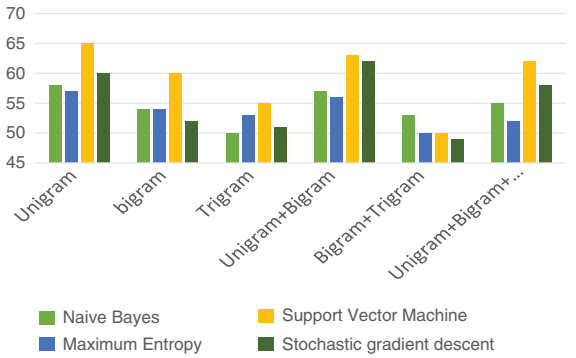


Fig. 3. Accuracy of book reviews dataset.

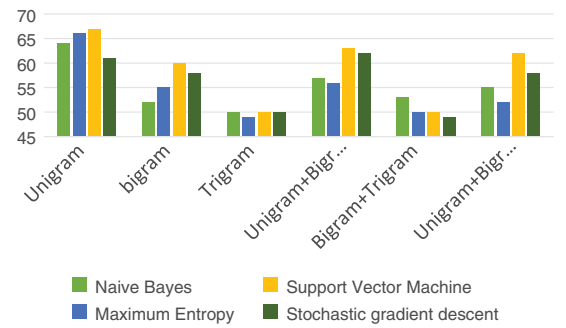


Fig. 4. Accuracy of electronic dataset.

n grams and its combinations. It also inferred that *n* gram combination have some better accuracy compared to bigram and trigrams.

Figure 3 represents the classification results of book reviews dataset. In this unigram and unigram + bigram shows better performance.

Figure 4 is the accuracy results of electronic products review dataset.

6. CONCLUSION

In this paper, four different algorithms such as Naive Bayes, maximum Entropy, Support Vector Machine and Stochastic Gradient Descent using *n*-gram approach like unigram, bigram, trigram, unigram + bigram, bigram + trigram and unigram + bigram + trigram were applied.

The proposed approach classifies the reviews into either positive or negative polarity. Hence it is able to guide the people properly by informing them about the disadvantage or good features of the products. It is observed that as the value of '*n*' in *n*-gram increases the classification accuracy decreases i.e., for unigram and bigram, the result obtained using the algorithm is remarkably better, but when trigram, four-gram, five-gram classification are carried out, the value of accuracy decreases.

Acknowledgment: I take immense pleasure to express my heartfelt thanks to my guide, P. C. D. Kalaivaani M.E., Assistant Professor (SLG) for her valuable ideas and suggestions, which have been very helpful in the project. I am grateful to all the faculty members of the Computer Science and Engineering Department of Kongu engineering, for their support.

References

1. R. Feldman, *Communications of the ACM* 56, 82 **(2013)**.
2. G. Gautam and D. Yadav, Sentiment analysis of twitter data using machine learning approaches and semantic analysis, *2014 Seventh International Conference on Contemporary Computing (IC3)*, IEEE **(2014)**, pp. 437–442.
3. D. Zhang, H. Xu, Z. Su, and Y. Xu, *Expert Systems with Applications* 42, 1857 **(2015)**.
4. T. Niu, S. Zhu, L. Pang, and A. El Saddik, Sentiment analysis on multi-view social data, MultiMedia Modeling, Springer International Publishing, Switzerland **(2016)**, pp. 15–27.
5. S. M. Liu and J.-H. Chen, *Expert Systems with Applications* 42, 1083 **(2015)**.
6. L. Bottou, Stochastic gradient descent tricks, Neural Networks: Tricks of the Trade, Springer **(2012)**, pp. 421–436.
7. T. Mikolov, K. Chen, G. Corrado, and J. Dean, Efficient estimation ofword representations in vector space, arXiv preprint arXiv:1301.3781 **(2013)**.
8. K. Mouthami, K. N. Devi, and V. M. Bhaskaran, Sentiment analysis andclassification based on textual reviews, *2013 International Conference on Information Communication and Embedded Systems (ICICES)*, IEEE **(2013)**, pp. 271–276.
9. R. Garreta and G. Moncecchi, Learning Scikit-Learn: Machine Learningin Python, Packt Publishing Ltd. **(2013)**.
10. B. Liu, *Synthesis Lectures on Human Language Technologies* 5, 1 **(2012)**.
11. B. Luo, J. Zeng, and J. Duan, *Expert Systems with Applications* 44, 138 **(2016)**.
12. J. A. Balazs and J. D. Velásquez, *Information Fusion* 27, 95 **(2016)**.
13. Montejo-Ráez, E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Urenã-López, *Computer Speech and Language* 28, 93 **(2014)**.
14. K. Ravi and V. Ravi, *Knowledge-Based Systems* 89, 14 **(2015)**.

Received: 20 April 2018. Accepted: 9 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1607–1614, 2018

Keyless Cryptosystem for Secure Primitive Pairing of Mobile Devices

Prakash Gopalakrishnan* and B. Uma Maheswari

Department of Computer Science and Engineering, Amrita School of Engineering, Bengaluru Amrita Vishwa Vidyapeetham 560035, India

Smartphone is a leading mobility platform of the present computing era. Nowadays, transmitting precious information securely between mobile phones is considerably a big challenging task. The mobile devices must get paired in such a manner that effective data exchange should take less time and energy consuming, thus overcoming the fraudulent attacks in the communication channel. The most well known technique of building communication protocols for setting up a secure transfer between such mobile devices is the Cryptographic protocols. To confront security by overcoming the dominance of the invader effectively, Cryptographic techniques are now embedded in wireless environment. Keyless Cryptosystem is one such method which is elaborated as Keyless Key Agreement Cryptography method. This review paper throws a lime light on the overview of Keyless Cryptosystem. Various protocols using this technique are compared and highlighted here. This paper further explains the cryptosystem used in primitive pairing of mobile devices in a secure manner.

Keywords: Keyless Cryptosystem, Key Agreement Protocol, ShaVe and ShaCK Protocol, SHOT Protocol.

RESEARCH ARTICLE

1. INTRODUCTION

Secure key establishment between two parties can be addressed normally using Public Key Infrastructure (PKI) or an Online Trusted Third Party. Otherwise, another most important technique could be is to use the Diffie-Hellman Key Exchange solution.[1] Unfortunately, these solutions cannot always be applied to resource-constrained devices operating in pervasive environments because of the lack of availability of either the public key or the trusted third party as required by Asymmetric Cryptography which led to the high computation and bandwidth overhead. A preliminary solution addressing the secure key establishment issue was provided in 1983 by Schneider and Alpern.[2] The notion of this work starts with these two security experts. They proposed a new cryptographic method where the security is highlighted by hiding the creator of the message (i.e.,) the identity of the sender is hidden. This novel technique is known as Keyless Cryptosystem.

Keyless Cryptography is a method in which the security of key exchange lies in the anonymity of the sender. This concise review provides a justification of the protocols that use this technique, in the real world communication environment. The protocol using Keyless Cryptosystem allows the sender and receiver to create a single random key such that eavesdroppers or even other legitimate users are restricted to get access of this random key. It is assumed that both sender and receiver have an anonymous channel, i.e., a channel that could perfectly hide the creator of the message. The details of implementation of various Keyless Cryptosystem Techniques are compared and presented in the following sections.

2. WHY KEYLESS CRYPTOSYSTEM?

Public Key Cryptosystem may be prone to impersonation where a successful attack can be staged by the adversary on a Certificate Authority (CA). In open multiuser environment, Public Key Cryptosystem is best suited; whereas in other wireless environments, Public Key Cryptosystem is hectic to be implemented effectively. For this purpose, there emerges yet another dimension of Cryptography—The Keyless Cryptosystem. The basic concept present in Keyless Cryptosystem algorithm is that both sender and receiver shall create a unique random key such that the wiretappers are avoided from getting the random key. The assumption is that the communication between the parties, takes place in an anonymous channel. This hides the creator of the message, which is the key idea behind Keyless Cryptosystem technique.

*Author to whom correspondence should be addressed.

3. RELATED WORKS

Several researches have undergone on the secure data exchange between any two paired devices and the process of establishing the shared secret key securely is still considered to be an ongoing dispute in the security era. In these papers, different categories of related work on secure data exchange mechanisms are discussed along with the briefings of the protocols that use Keyless Cryptosystem proposed by various authors.[1–4, 7, 14, 17] This survey also throws light on the secret key extraction by Keyless Cryptosystem from the pairing devices used in both static and dynamic environments,[5, 8–10] wireless networks and environments using sensor nodes[8, 9] and extraction of secret key from Channel Measurements.[19–21] This survey provides a detailed analysis of previous works on using cryptographic techniques without secret keys but with relevant security. It also brief about the protocols like Movement Based Pairing protocol, SHave and ShaCK protocol, BUMP protocol, SHOT protocol and COKE protocol using Keyless Cryptosystem proposed and implemented by various authors.

The prior research carried out by Yung,[3] was used in a kinetic procedure to collect accelerometers data during the shaking of the two devices coupled together. Shared secret comes from the accelerometers data that Mayrhofer and Gellersen[4] has proved to be correlated in their research work. Unfortunately, such an approach is effective for devices that can be shaken together only. As suggested by Zeng et al.,[5] establishing a new secret without preconfigured information and avoiding asymmetric cryptography is a challenging topic that has been previously undertaken by leveraging anonymous channels or adopting received signal strength key establishment protocols. The authors Lester et al.[6] and Castelluccia et al.,[7] explained how anonymous channels guarantee that an adversary cannot identify the actual signal source even if it is able to eavesdrop all the transmitted messages. Two peers can exchange the bits of a secret message, but it is not possible for an adversary to associate the current transmitted bit to the actual source. Such Cryptoless algorithms are particularly useful in pairing of resource-constrained devices. Such keyless procedures are briefed by Jana et al.,[8] Patwari et al.[9] Thus a gist of this study gives an overview of how Keyless Cryptosystem is used for secure key exchange in paired mobile devices.

Croft et al.[10] implemented the adaptive ranking based uncorrelated bit extraction (ARUBE) method to eliminate the non-reciprocities which are predictable in consequence of Wireless Sensors having the differed hardware characteristics in transceiver. It is noted that Ranking is considered to be robust where the order of measurements will be the same even when the measured values at various nodes are of a diverse scale. This Ranking method extracts the bit independent of fading distribution where it keep away from the disagreements that has been caused due to the differing transmit powers and Received Signal Strength Indicator (RSSI) circuit variations. These Ranking methodologies have initiated very low computational complexity and could adapt automatically the transmitter and receiver differences.

Wilhelm et al.[11] introduced a Key Generation Protocol in their work for Wireless Sensor Networks test bed. It was experimented and shown that the protocol could provide secrecy in real-world measurements up to 160 bits capability and with limited wireless channels. Depending upon the behavior of Wireless channels, Secret keys are successfully generated in more than 95% case scenarios which have been evaluated with MICAz sensor motes. The authors concluded that for a very small available number of wireless channels, this protocol could be applied to a variety of applications. For example to establish an authenticated channel, a part of device-pairing schemes has been used for deploying the initial keys.

Mukundan et al.[12] proposed a novel lightweight cryptographic hash function—Hash-One for meeting the requirements of constrained devices such as Wireless Sensors, RFID tags, and other embedded system devices. It has been noted that the security level in the existing devices like SPONGENT, QUARK, PHOTON, and GLUON are set to 80-bit standard with vast area (GE). Hash-One, on the other hand has achieved the same security within a limited area of 1006 GE for 160-bit pre-image resistance and 80-bit resistance in collision. For analyzing the security features, Hash-One considered the scope of Algebraic attack and Differential attack. Three different cryptographic randomness tests—Strict Avalanche Criterion (SAC) test, Coverage test and Coverage test are successfully applied in Hash-One thereby attaining the compactness in highly constrained devices.

Praveen and Sethumadhavan[13] use a set of elliptic curves in Identity Based Encryption (IBE) which is commonly used in the applications of pairing based cryptography. The main feature of elliptic curves is pairing which can be applied in lightweight memory devices. The authors improved the efficiency by implementing the MNT elliptic curves and effective computable endomorphism for faster pairing computations.

4. PRIMITIVE TECHNIQUES OF KEYLESS CRYPTOSYSTEM

4.1. Initial Protocol Scheme

Keyless Cryptosystem depends on the sender's anonymity. Bowen Alphen explained about the protocol that uses this technique of hiding the sender of the message and not the contents of message.[2] To distribute secret keys in computer networks, three ways of implementing the key distribution protocol are suggested by the author. They are

- Central Key Distribution Facility
- Broadcast Network
- Special Communication Channel.

According to the first scheme, the protocol assumes the presence of a trustworthy process where it preserves a database which upon receipt of the message will post them on the blackboard randomly. By reading the message, the intended parties that are communicating can determine their pair. It is impossible for the other parties to find out the message source just by network traffic monitoring.

In later schemes, the end-users are connected through ring, Ethernet or any other broadcast network where the assumption is that the message departs from the blackboard will move only with the destination address and finally a special two-wired communication channel is established to connect a pair of people. The drawback of this protocol is that even while transmitting data encapsulated, still it poses security vulnerabilities and this Keyless Cryptographic System was not considered as an effective method to protect the data. This technique is prone to active Wiretapper attack and thus does not provide adequate secure data storage.

Yung[1] suggested an extension of this scheme where the keys are distributed through a semi-anonymous channel. In the original scheme, the key were randomly generated by the protocol thus disallowing adversary from getting that value, but data was stored using fixed keys. But the authors of this paper implemented using a secure fixed key and a secure conferencing to distribute keys to a group of users. They also use mutual authentication method of preventing attack from active attacker. For mutual authentication, both the sender and receiver use a fixed bit—the authenticator, which is known to each of them. The authenticator is the first bit of fixed size packets that are transmitted between the users. Hence, the success rate of attacks is considerably reduced to half the size of packet having less probability of attack. For large number of packets, the success rate of attack on the protocol is trivial. The findings from the two papers suggest the primitive methods of Keyless Cryptosystem protocol.

5. KEYLESS CRYPTOSYSTEM PROTOCOL IN SECURE PAIRING OF SMART PHONES

5.1. ShaVe and ShaCK Protocol

Simultaneous mobile interactions which pair the devices must be done with at most security. Shaking the mobile devices is an easy way for the devices to be paired. Mayrhofer and HansGellersen[4] proposed a protocol technique that uses a secret key for mutual authentication which is considered as a frequent progressive approach i.e., the ShaVe and ShaCK Protocol in combination with the Keyless Cryptographic protocols has been used for secure authentication. The former Protocol ShaVe—Shaking for Verification is used for exchanging the keys through Diffie-Hellman algorithm followed by another key exchange that are compared with sensor data for verification of authenticity.

The latter ShaCK—Shacking to Construct a Key, is based on feature extraction method where the features are matched obtained from the sensor data using the accelerometers which is used to build a strong cryptographic key. This technique is used for dynamic interactions. The main idea introduced, is authentication by the process of shaking a pair of mobile devices. This shaking creates a movement limited channel where an accelerometer is used to sense the progress. The data collected is used for key generation, verification and authentication process.

Two authentication protocols are used. The protocols based on Diffie-Hellman algorithm[1] and Interlock protocols are used for verifying the key together with analyzing the coherence of accelerometer data. This is used for generating keys from sensor data streams where the main threat is only active middle man attack. ShaVE and ShaCk protocols prevent some of the major threats like devices getting accidentally paired by accidental joint movement. An adversary try to copy the movements of an authenticated user. Then it estimates the accelerations experienced by the attacked devices and gives its own time series analysis. This report is same as successful authentication by the actual device. Thus these threats are overcome by implementing this protocol. It is well known that these methods could be easily learnt and highly feasible to be implemented on smart mobile phones where it is significantly secure from the adversaries and are fully controlled over a Wireless channel.

5.2. Key Agreement Protocol—Movement Based Pairing

It is an optimized approach where the secret key is established at a low cost. This protocol describes how the communicating parties exchange packets with the source field set. The eavesdropper cannot get a clue of the sender of the intended packet, as the random packets agree on a secret bit. When devices are paired, they are moved to establish security via RSSI (Received Signal Strength Indicator). The main security must be established during pairing. When data is received from the sensor, it must be checked whether it is got from the appropriate device sensor. For this purpose, two sub protocols—Pairing Validation protocol and Key Exchange protocol are discussed.

Key agreement protocol proposed by this paper [7] discusses how to ensure security against middle man attacks and Denial of Service attacks. A Start message is initially sent such that parties communicating shall start their exchange of message which contains the key size and the address of source packet. Alpern's Keyless Cryptosystem is used in ad hoc environment where the Key poisoning attack is also defeated by these agreement protocols which are proved experimentally. This protocol does not depend on operations that are CPU constrained and the main focus of Keyless Cryptosystem is implemented based on these constraints. The transmission power of the paired

devices, distance, and angle between the analyzer of signal and source, and the relative terminal position are the four factors on which the strength of the signal received depends.

5.3. Protocol for Secure Pairing—iOS Application

Hacker et al. had used this ShaVe and ShaKe protocol[14] along with the introduction of new methods of random projections and hamming distance. The mobile devices simultaneously undertake a shared experience (hold physically the paired devices and perform shaking on them for a specific duration) with the data collected for use it to verify a session key that was exchanged previously. In their MA Thesis, it is explained with an iOS application for data collection through shaking via the accelerometer present in the device. The collected data sets provide an affirmative valid result that is compared with a threshold value. The main advantage is that the users need not remember any type of password or any complex movement because of the improvements of the existing protocols. These iOS apps which are running on iPhone collect movement data securely using accelerometer. Two methods that are examined here, can strongly categorize the devices as shaken independently or shaken together. Again the data sets return a positive valid result as above and thus the security is ensured. Shave and Shake methods can be used with various new techniques to form a secret key that is more effective and deceives the adversary.The flaws in accelerometer may make the adversary bring in an alternative and makes an attack. So, error in this device must be intimated to the user with a warning which shall be improved in future.

5.4. Bump Protocol

Bump is an iOS application with a smart matching algorithm called Bump Protocol[15] that is running on a server present in the cloud. Phones have sensors, which on bump with each other will send the information to the server. The information such as initial authentication followed by contacts, photos, audio, videos, etc. are exchanged using these apps. Thus, devices are paired with the phone that feels the same bump, as the users of these apps are the deciders to whom the information is to be shared. Security is in their hands, as physical touch of the phones will make the peers access the information.

As per the policies of BUMP technologies, they don't support PayPal Services.[16] A strong internet connection is a must while using this application or else there are chances for eavesdroppers to play. Bump apps also do not use Bluetooth as this application had lots of vulnerabilities and experiences when it was introduced. So to overcome it, changes were made in the matching algorithm to check for the location of information also. Because, passive wire tappers started to produce same force, as originally produced when a mobile bumps with another phone and pretends to be an authentic device to the server to steal the data. Bump is an application that works on cross platforms and hence attack on it could not be avoided, though it is the most evident application in this smart phone era.

Exchange of information between smart phones, is the foremost activity of the users. Many users tend to do multiple tasks with at most security through these mobile platforms. People use smart phones to interact with those whom they meet, to have a protected or sheltered communication where these interactions are initiated through the cryptographic key exchange techniques. Hence, many protocols have been developed to smooth the progress of this exchange. Regrettably, the protocols that provide guaranteed higher level of security are often suffered from usability issues. Those protocols that are easily accessible may not provide the fail-safe security from adversaries.

5.5. SHOT (Shake It On) Protocol

Work of Studer et al.[15] clearly shows the attacks initiated through induced delay against Bump through the Middle man attack. The summarized results show that for an optimal delay of just 2 seconds, seventy percent of the attacks would be achieved enough. It was well known by the attacker enforcing the Middle man attack about the time and location of bumping the smart phones together, along with the estimated force of the bump that was initiated for communication between the authorized users. Having these details, then the attacker could submit the related information obtained from the Bump to the concerned mobile phone's server. This related information spoofs the legitimate user where the data may transfer among unethical users.

Due to this situation, this unprotected algorithm which was widely used by Google play store was shutdown on 2014. Flock, a supporting application of Bump also was shut down on the same year and the team promised to bang back with a new application which is completely secure. The main drawback of Bump is inaccuracy in measurement of physical aspects of Bump protocol.[14] There is inability of the sender to control the identification used during confirmation for accessing services. There are server flexibility in accepting delayed bump requests and sensor accuracy. It means, the phone is unable to know its exact location or how hard the phone was bumped. This team has suggested an algorithm called SHOT (Shake It On)[15] that uses the phones' accelerometer and venerator to establish a human observable channel. This channel gives an open identification of the devices participating in the exchange. So, users can be notified of the attack. Security is ensured by using pre authentication of cryptographic keys. If still the attacker is injecting information, a hash function which is second-preimage resistant is used to verify it and receives a copy of public key. Working of this algorithm is quick, compared to Bump Protocol giving greater security.

5.6. COKE (Crypto-Less Over-the-Air Key Establishment) Protocol

Roberto Pietro and Oligeri[17] proposed a protocol that explains the way of extracting, a particular sequence of shared bits from the identical point of piece of information observed. These details are presented and handled in many research papers. The consumed signal power shall be used to extract the shared secrets from the two peered systems. It was seen from the authors' perspective that Multipath Fading which occurs had introduced the dissimilarities on the signal power received. However, the authors have suggested a few algorithms that would recover the errors and ultimately, agree on a secret key which is being shared. Here the accelerometer is commonly used to check whether the two peered systems are manipulated by the same user. Their algorithm is focused to combine together the peers, then to perform the shaking with peers and accumulating the values from the accelerometer and at the end to make a key agreement on the collected values.

As Key Exchange using Keyless Cryptography was proposed earlier by Alpern and Schneider,[2] they have initiated to generate and distribute the secret key through hidden source identity (transmitter) and not with the packet contents. Here each packet is taken along with only one secret key bit and the source field of the packet is hidden. Thus, the adversary could know the secret bit value, but would not know about the sender of the bit. Thus this keyless key agreement procedure in a wireless environment is narrated efficiently using this Crypto less Over-the-air Key Establishment protocol (COKE).

6. SECRET KEY EXTRACTION BY DEVICES USED IN THE ENVIRONMENT

6.1. MEMS Accelerometers

As the day to day need of the users increases, they carry more devices, without the ability for the devices to unobtrusively interact with one another. The user spends more effort on coordinating, rather than using, these devices. Lester et al.,[6] in their paper presented a method based on a coherence function. This function is a measure of linear correlation in the frequency domain which is used to monitor and analyze data in motion (i.e.,) data in movement recorded by MEMS accelerometers. As the researchers use inexpensive accelerometers, these devices are used to find out whether the two mobile devices are carried by the same person or not.

Sensors have a similar performance when compared to more expensive accelerometers. The results got on testing these MEMS accelerometers from a hefty test group proving the algorithm's strength and its capability to withstand time delays of real world. This device is tolerant and liberal to inter-device communication latencies, and it needs a little communication bandwidth. The methods of using this accelerometer are explained in this paper.

6.2. MAKE in Wireless Networks

MAKE—Multiple Antenna Key Generators. In Wireless Networks, Zeng et al.[5] proposed how to exploit multiple antenna diversity for generation of shared secret key. In their research work, fading in wireless channels is exposed to randomness. This is more in multiple-antenna devices. So Multiple-Antenna Key Generator (MAKE) was suggested to generate a secret key that is shared. Using this channel, eavesdropper cannot gain much more information about the valid legitimate channel. RSSI—Received Signal Strength Indicator measurements are used to generate the keys used for communication, when compared to system with single antenna system. MAKE increases the generation of bit rate. MOTO mode—Multiple One antenna To One antenna, SOTO mode—Single One antenna To One antenna are used to evaluate the performance of MAKE. The channel is probed faster at short interval to get high bit rate and to extract more randomness from the channels. The bit generation is limited when the time changes quickly in the channel. To generate random bits; the channel has to be investigated more.

7. COMPARATIVE STUDY OF PROTOCOLS USING KEYLESS CRYPTOSYSTEM

Various protocols of Keyless Cryptosystems that are highlighted in this paper, are compared with certain parameters like, the various types of attack for which the protocols taken for assessment are providing solutions, devices that are inbuilt by the system that use the outcomes that aid in efficient working of these protocols, data transmitted and its efficiency, technical feasibility, economic feasibility, environmental study and the security measures.

Analytical studies of these parameters are done on these procedures to find out the impacts of them. Thus this Table I shown below provides the details on, how the various protocols of Keyless Cryptosystem's novel technique behave. The protocols elaborated in this research which is taken for comparisons are, Keyless Agreement Movement based Pairing Protocol, ShaVe and ShaCk Protocol, Bump protocol, SHOT Protocol and COKE Protocol.

Mystica and Gopalakrishnan Prakash[18] revealed clearly that there are many recent protocols accessible in Keyless Cryptosystem. All the protocols provide adequate security features and mitigate the attacks based on their individual capabilities. BUMP protocols are well suited for securing iOS based Apple Smart phones. As this protocol is vulnerable to Eavesdropping, SHOT protocol came into the limelight which could overcome Meet-in-the-Middle Attack. Further advancements led to the introduction of COKE

protocol, i.e., the Crypto-less Over-the-air Key Establishment Protocol. This is considered to be the most prominent of present security era and used in recent times. This technique has leveraged the concept of "No Cryptography," but the exchange of plaintext message only. It has been eminent that the strength of this technique strongly relies on the adversary's difficulty in identifying correctly the transmission of just one-bit. This is indeed the sender of that bit, and not its value has been exchanged in plain text. Due to these low requirement capabilities, COKE can be specifically suited to resource constrained Wireless devices and Wireless embedded systems. This follows the scenario that smart phones are saving the energy at premium level and thus maintaining the sufficient security feature.

Table I. Comparative study of various protocols.

	Protocols				
Comparative parameters	Keyless agreement movement based pairing protocol	ShaVe and ShaCk protocol	BUMP protocol	SHOT protocol	COKE protocol
Solution to types of attack	i. Passive attack ii. Active attack iii. Middle man attack iv. Denial-of-service	i. Middle man attack ii. Offline guessing attack iii. Offline brute force iv. Partial message collision attacks		i. Man in the middle attack ii. Denial-of-service attack	i. Bit position guessed through the calculated signal strength received ii. Identification of radio source in the physical layer
Devices incorporated	i. CPU constrained devices (sensors)	i. Communication channel that are out of band. ii. Sensor along with minimum feedback (beeper or light emitting diode) used independently iii. UI exclusive devices like sensor nodes, portable hard drive accessories that are wireless. iv. Mobile phones and handhelds wireless accessories equipped with accelerators	i. Accelerometer readings. ii. Sensor readings	i. Vibrator-to-accelerometer ii. Vibrator class is used iii. Hardware sensor manager used to access accelerometer. iv. When accelerometer reading are more android hardware sensor event listener is used	i. Resource constrained devices (WNS, wireless embedded system)
Data transmitted	i. Infrared transmission of packets with 15 KB	i. Sensor data streams used to transmit accelerated data	i. Photos, videos, contacts, other files from a computer to a smart phone and vice versa transmitted via a web service ii. File size limited to 20 MB. Large files take a much longer time to transfer	i. SHOT server accessed through the internet. Hence used to transfer majority of data between mobile phones	i. Applicable to any type of data transmissions with the condition that transmissions are time synchronized
Technical feasibility	i. RSA with very low exponent specification ii. Keyless key agreement	i. One way hash function	i. Hold phones together and gently bump one phone against the second phone. Mobile phones feel the vibration when it is bumped for transfer of data.	i. Bouncy castle cryptographic operations ii. SHA-1 to create pre-authenticator and to verify the public key iii. 1024-bit RSA signatures	i. Cryptographic hash function—(SHA-1, SHA 512, etc.)

Table I. Continued.

Comparative parameters	Protocols				
	Keyless agreement movement based pairing protocol	ShaVe and ShaCk protocol	BUMP protocol	SHOT protocol	COKE protocol
Economic feasibility	i. Protocols used for pairing sensors—CPU constrained devices have limited memory. ii. Inexpensive	i. Sensor data used to process floating point notations that are expensive ii. CPU has to be embedded with the floating point unit's hardware iii Cryptographic operations makes the signal processing costly	i. This protocol runs on iOS platform effectively and is expensive	i. Significant cost according to the usability of the system	i. Inexpensive compared to other protocols
Security measures	i. Ethernet card (moving it at specific speed) is used to confuse the attacker	i. ShaCk message sizes constantly avoiding offline brute force attack and partial message collision attack ii. Candidate key part message changes with the duration of shake iii. To protect against offline attack, number of candidate key part messages have got shaken duration.	i. Security depends on bump server's ability to determine the mobile phones that are physically interacting, time location, force with which the two phones were physically bumped together. ii. Modify or delete the contents of USB storage/SD card. iii. Prevent devices from sleeping by consistent notification. iv. Periodically it connects and disconnect from Wi-Fi	i. Shot protocol is comparatively quicker than Bump ii. Shot provides reliable security	i. Strategy to deceive the adversary is experimentally elaborated ii. Attacks on devices running with protocol are less prone to MitM and DoS attack
Environmental study	i. Comfortable in places where capacity of the sensor is confined		i. Application specific to running on iOS based iPhones and android mobile phones ii. Cannot be used on mobile phones running on other platforms	i. MitM attack setup on BUMP is explained using Mac Book Pro with OS X 10.5 dummy net as an access point	i. Its best suited with radio environment having omni directional antenna.
Real time example		i. Nokia 5500 with Symbian application that starts automatically and make use of the Nokia sensor API to get the data from accelerometer	i. Apple iPhones with iOS and BUMP apps	i. Motorola DROID smartphone's with android version 2.2	

8. CONCLUSION

This work creates an impression on Keyless Cryptosystem, giving a brief explanation of various protocols and techniques. Using this new dimension of cryptography, high security is established in the wireless communication, making it adaptive to any scenario and providing high protection against data leakage. On the above stated comparative analysis, COKE is a distinct protocol of the recent

times, to challenge the adversary by its technique, focusing on establishing secrecy in single bit i.e., each packet carries only one bit of the secret key and the packet source field is hidden. In this way, the adversary knows the value of the secret bit; but she does not know the sender of that bit. This protocol is ideal for devices that use low power, thus economically feasible. Received signal power can be used to extract shared secrets between two peers. Limitation of this protocol is that, information can be retrieved with certainty, only by using a directive antenna; but usage of such antenna will pave way for attacks and the next limitation is, position guessing issue occurs, while establishing secret keys. Enhancement of this protocol might be overcome by future techniques.

References

1. W. Diffie and M. Hellman, *IEEE Trans. Inform. Theory* IT-22, 472 **(1976)**.
2. B. Alpern and F. B. Schneider, *Information Processing Letters* 16, 79 **(1983)**.
3. M. M. Yung, *Information Processing Letters* 21, 35 **(1985)**.
4. R. Mayrhofer and H. Gellersen, *IEEE Transactions on Mobile Computing* 8, 792 **(2009)**.
5. K. Zeng, D. Wu, A. Chan, and Prasant Mohapatra, *Proceedings of the 29th Conference on Information Communications* **(2010)**, pp. 1837–1845.
6. J. Lester, B. Hannaford, and G. Borriello, Lecture Notes in Computer Science Book Series **(2004)**, Vol. 3001, pp. 33–50.
7. C. Castelluccia and P. Mutaf, *Proceedings of the 3rd International Conference on Mobile Systems Applications and Services MobiSys* **(2005)**, pp. 51–64.
8. Suman Jana, Sriram Nandha Premnath, M. Clark, Sneha K. Kasera, Neal Patwari, and Srikanth V. Krishnamurthy, *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking* **(2009)**, pp. 321–332.
9. N. Patwari, J. Croft, Suman Jana, and Sneha Kumar Kasera, *IEEE Transactions on Mobile Computing* 9, 1 **(2010)**.
10. J. Croft, N. Patwari, and Sneha K. Kasera, *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks* **(2010)**, pp. 70–81.
11. M. Wilhelm, I. Martinovic, and J. B. Schmitt, *Proceedings of the 3rd ACM Conference on Wireless Network Security* **(2010)**, pp. 139–144.
12. Puliparambil Megha Mukundan, Sindhu Manayankath, Chungath Srinivasan, and Madathil Sethumadhavan, *IET Information Security* 10, 225 **(2016)**.
13. I. Praveen and M. Sethumadhavan, *Proceedings of the First ACM International Conference on Security of Internet of Things SecurIT* **(2012)**, pp. 145–149.
14. M. Hacker, M. Crovella, and L. Reyzin, Technical Report BUCS-TR-2012-011, Computer Science Department, Boston University **(2012)**.
15. A. Studer, T. Passaro, and L. Bauer, *Proceedings of the 27th Annual Computer Security Applications Conference* **(2011)**, pp. 333–342.
16. https://www.theverge.com/2014/1/1/5262480/google-acquired-bump-shutting-down-this-month.
17. R. D. Pietro and G. Oligeri, *IEEE Trans. Info. Forensics and Security* 8, 163 **(2013)**.
18. Fatima Joselyn Mystica and Gopalakrishnan Prakash, *International Journal of Science and Research (IJSR)* 4, 2 **(2015)**.
19. S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, *Proc. 14th ACM Int. Conf. Mobile Comput. Network. MobiCom* **(2008)**, pp. 128–139.
20. T.-H. Chou, A. M. Sayeed, and S. C. Draper, *Proc. 2009 IEEE Int. Symp. Inf. Theory* **(2009)**, pp. 2296–2300.
21. A. A. Hassan, W. E. Stark, J. E. Hershey, and S. Chennakeshu, *Digital Signal Processing* 6, 207 **(1996)**.

Received: 21 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1615–1624, 2018

RESEARCH ARTICLE

A Novel Battery and Super Capacitor Using IOT to Interface Renewable Energy Sources with Dual Direction and Its Application

S. Sankarananth[1,*], M. Karthiga[2], C. Karthikeyan[3], and M. Vinith Kumar[1]

[1] Department of EEE, Excel College of Engineering and Technology, Komarapalayam 637303, India
[2] Department of CSE, Bannari Amman Institute of Technology, Sathyamangalam 638401, India
[3] Department of EEE, K. S. Rangasamy College of Engineering, Tiruchengode 637215, India

A multi-input DC–DC help converter is proposed to get control from a few info sources and to supply the managed yield voltage to the heap from the sources. The information sources are the sustainable sources like sun based, wind, power module and vitality stockpiling frameworks like battery and super capacitor the IOT innovation was utilized as a part of this procedure to monitor and control. The battery used is bidirectional in its energy stream consequently the converter topology empowers the charging and releasing of the capacity component through information control sources. The task will be in the bidirectional route for the battery and the super capacitor. The control procedure proposed for this converter depends on stage move control and to manage the voltage to utilize PI controller. The proposed topology has the benefit of intrinsic bidirectional power stream, negligible number of transformation steps, high effectiveness and concentrated control can be executed.

Keywords: Hybrid Energy Storage System (HESS), Bidirectional DC–DC Converter (BDC), Battery, Super Capacitor, Solar, Wind, Fuel Cell Energy Storage System (ESS), Phase Shift (PS).

1. INTRODUCTION

There are five info ports the Bidirectional DC–DC converters (BDCs) effectively steps up or venturing down the voltage level between its information and yield alongside the ability of energy stream in both the bearings just or the battery and super capacitor. BDCs can be arranged into confined and non-secluded composes. Galvanic disengagement between multi-source frameworks is a prerequisite ordered by numerous models. Voltage coordinating if there should arise an occurrence of substantial voltage proportions between two sources, faculty wellbeing, clamor lessening, impedance blocking ability and right activity of security frameworks are the primary explanations for galvanic detachment which is required by numerous delicate DC loads.

Normal IBDC topology is Dual Active Bridge (DAB) topology in which full extension voltage nourished converters are utilized at both the finishes of the segregation transformer. Keeping in mind the end goal to interface different sustainable power sources or to interface various vitality stockpiling frameworks with the heap; multiport confined bidirectional dc–dc converters are finding their expanding applications. Hybrid Energy Storage System (HESS)—battery and super capacitor mix are finding their promising application in cross breed and power module controlled Electric Vehicles (EV), shrewd matrix, extensive scale wind and photovoltaic frameworks Uninterruptible Power Supplies (UPS) and so forth.[9] The ideal decision of development for this kind of HESS is triple dynamic full extension (TAB) topology[8–10] which is gotten from double dynamic scaffold (DAB) topology.[1]

In Ref. [5] diverse structures of multiport converters is introduced to interconnect various sources. Isolate dc–dc converters are utilized for singular sources. These converters are connected together at the DC transport and controlled autonomously. However, a downside of this structure lies in the way that it is innately mind boggling and has high cost because of various transformation stages. Multiport structure both the DC interface and the attractive coupling are joined to such an extent that entire framework is dealt with as single power converter in which concentrated control can be executed with insignificant number

*Author to whom correspondence should be addressed.

of transformation steps in light of the fact that here the framework assets are shared.

Three port triple half scaffold bidirectional dc–dc converter is displayed in Ref. [4] that interface super capacitor, energy component and load. In any case it has just bidirectional port subsequently it doesn't bolster regenerative load. Half scaffold setup has of the converter has restrictions from the view purpose of energy rating, since high recurrence of current needs to go through the half extension capacitors in full. More finished because of wide working extent at the battery or super capacitor port limits the wide zero voltage exchanging (ZVS) run. Three port current bolstered full extension bidirectional converters which has the benefit of lessened current swell at the contribution because of the nearness of info inductor is introduced in Ref. [8]. The disadvantage of this topology is that it is reasonable just for low power and low yield voltage applications. In spite of the fact that the channel necessities are less it builds the control many-sided quality of the circuit.

Three port bidirectional converter for half breed energy unit framework is introduced in Ref. [11] in which the full extensions work in square wave mode stage moved from each other. One of the downside of this converter is that the exchanging recurrence and the inductor esteems can't be resolved autonomously for a similar power level. An arrangement resounding three port converter with stack side diode connect is exhibited in Ref. [10] in which every one of the ports are voltage encouraged ports. ZVS is conceivable in this converter with the end goal that exchanging misfortune diminishes. In any case, the disadvantage of this converter is ZVS isn't feasible for stack side port and not reasonable for regenerative load. More finished to get estimations of inductors more than the spillage inductance of the transformer the changing recurrence should be decreased. With a specific end goal to beat these disadvantages an arrangement resounding triple dynamic full scaffold topology is proposed in this paper in which all the voltage bolstered ports are bidirectional that interfaces battery which goes about as essential stockpiling and super capacitor which goes about as transient stockpiling.

2. PROPOSED TRIPLE ACTIVE BRIDGE CONVERTER

2.1. Circuit Construction

The proposed triple dynamic extension bidirectional dc–dc converter circuit comprises of three dynamic full scaffolds as appeared in Figure 1, three winding transformer and two arrangement thunderous tanks framed by L1, C1 and L2, C2. Battery is utilized as Energy Storage System 1 and super capacitor is utilized as Energy Storage System 2. Port 1 is nourished by battery port and port 2 is bolstered by super capacitor. Port 3 is load port (Resistive load and DC motor load). Each port consists of full bridge circuit with four switches. The switches are realized using MOSFET enabling bidirectional current flow in all the ports. The three winding transformer is mostly a

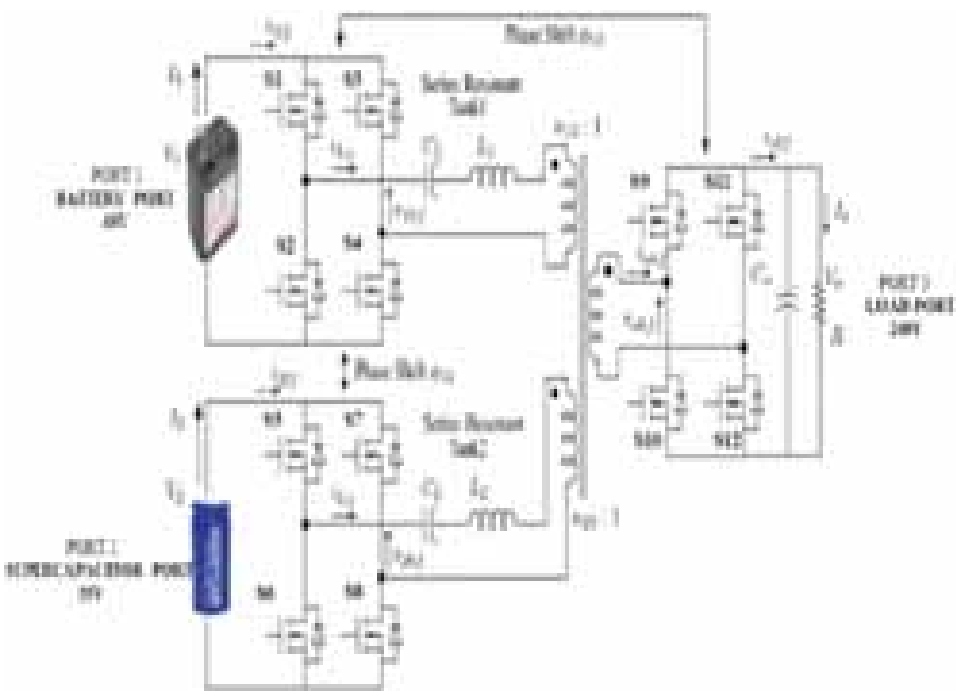


Fig. 1. Multiport isolated bidirectional dc–dc converter.

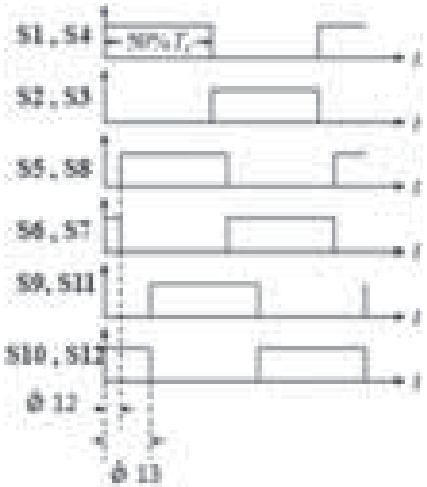


Fig. 2. Idealized key switching waveform for phase shift control.

step-up transformer. It has three functions (1) combines input DC sources in magnetic form, (2) provides electrical isolation, (3) step up voltage from Low Voltage Side (LVS) to High Voltage Side (HVS), (4) leakage inductance acts as energy transfer elements. The converter is operated at constant switching frequency FS above the resonant frequency F_r. Phase move control system is the customary control procedure which is generally actualized in BDCs converters. TAB BDCs utilizes the stage move φ between the voltage crosswise over the two sides of the separation transformer to control the exchange of energy through the components of the thunderous tank. When utilizing this control system all the switches of the converter are driven at half obligation cycle. The switches of port 1 M1, M4 and M2, M3 are controlled integrally. The switches of port 2 M5, M8 and M6, M7 are controlled correspondingly. Essentially the switches of port 3 M9, M12 and M10, M11 are controlled correspondingly. The two stage move control factors $\varphi 12$ and $\varphi 13$ considered. They control the stage move between the square wave yields of the dynamic extensions. The stage shifts $\varphi 12$ and $\varphi 13$ considered positive if Vohf slacks V1hf and V1hf slacks V2hf separately. The essential voltages V1hf and V2hf either (+V1 or −V1) and (+V1 or −V2) and the optional voltage Vohf is either (+V0/2n) or (−V0/2n). The admired key exchanging waveform for stage move control system is appeared in Figure 2.

3. PROPOSED BLOCK DIAGRAM

The multiport dc–dc converter for synchronously power transfer from multiple renewable energy sources uses only one power electronic switch in each input port connected to a source. The introduced converter does not use any controllable switch on the secondary side of the transformer.[2–4] The proposed converter has the least numbers of switches and thereby a lower cost. The newly introduced converter is applied for power management of a wind/solar hybrid generation systems, which consists of a WTG and two varied PV panels. The power generation from solar and wind energy are designed using perturbation and observation (P&O) MPPT algorithm, in which the WTG and PV panels can be controlled at the same time and extract the maximum power. The figure shows the block diagram of the introduced multiport DC–DC converter. It consists of PV Panels, Wind turbine generator, Boost converter, MPPT controller, High frequency transformer and an inverter.

4. MODES OF OPERATION

The task of the proposed TAB converter depends on the two noteworthy imperatives.

(1) When the port supplies energy to the heap i.e., when the port goes about as source it goes about as inverter (releasing mode).

(2) When the port sinks control i.e., charging mode it goes about as rectifier.

Here the methods of activities has three sort of modes, here they spoke to as mode A, mode B and mode C. One mode will demonstrates the task of the battery with the heap and another mode demonstrates the activity of super capacitor with the heap and mode C will clarifies about the task of the bidirectional activity of the battery and super capacitor with the heap.

4.1. Mode 1

In mode 1 both port 1 and port 2 supplies the heap port. Both the battery port and super capacitor port is in releasing mode as appeared in Figure 3. The stage move between port 1 and port 2 Φ is 00. The stage move amongst port 1 and port 3 Φ13 is 240 Port 1 goes about as inverter since it supplies control. MOSFET switches M1, M4 for positive mode or MOSFET switches M5, M8 for positive. Here the task clarifies about the working of the battery and

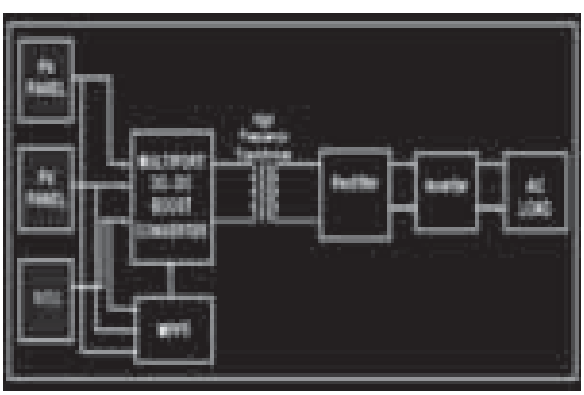


Fig. 3. Block diagram multiport DC–DC converter.

RESEARCH ARTICLE

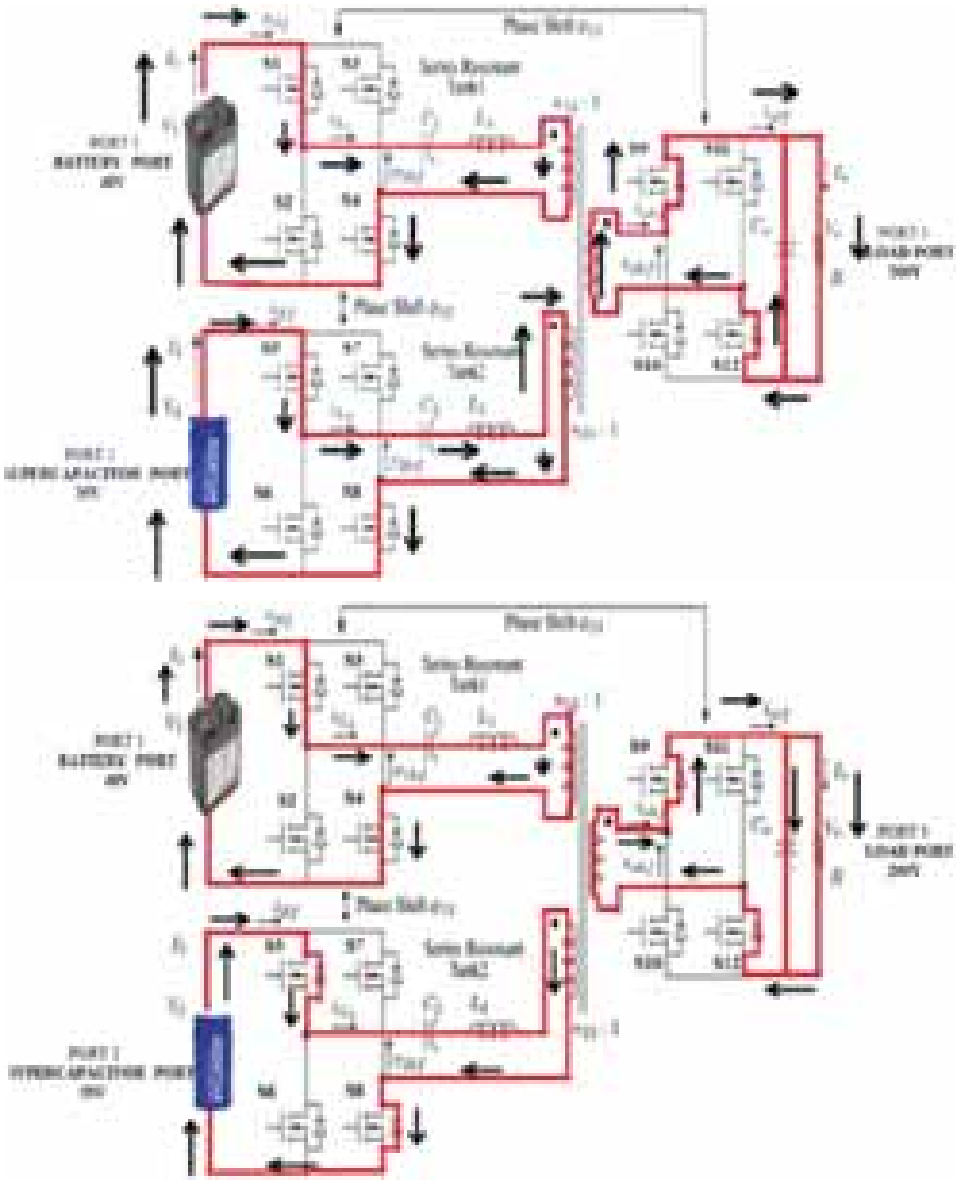


Fig. 4. Mode 1 and 2 circuit diagram.

the super capacitor in a bidirectional way. Other sources like the sun based vitality, wind energy and the fuel vitality won't have the bidirectional operations. So on account of this M2, M3 for negative mode conducts. Port 2 acts as inverter. Reason just the activities of battery and super capacitor bidirectional with the heap had clarified here. Mode or M6, M7 for negative mode conducts. Port 3 goes about as rectifier since it sinks power with the end goal that anti parallel diodes of MOSFET switches comes vigorously by the procedure of synchronous correction. Anti parallel diode D9, D12 conducts for positive mode or the anti parallel diode D10, D11 conducts for negative method of activity. Port 1, port 2 and port 3 control is sure when the converter works at mode 1.

4.2. Mode 2

In mode 2 the power provided by port 2 to the heap is zero. Port 1-Battery port supplies the port 2-Supercapacitor port and also supplies the R-stack as appeared in Figure 4 and consequently it goes about as inverter. The stage move between port 1 and port 2 Φ12 is 67.50. The stage move amongst port 1 and port 3 Φ13 is 51.50. Port 2 and port 3 goes about as rectifier since it is provided by the battery port. A since the port 2-Supercapacitor

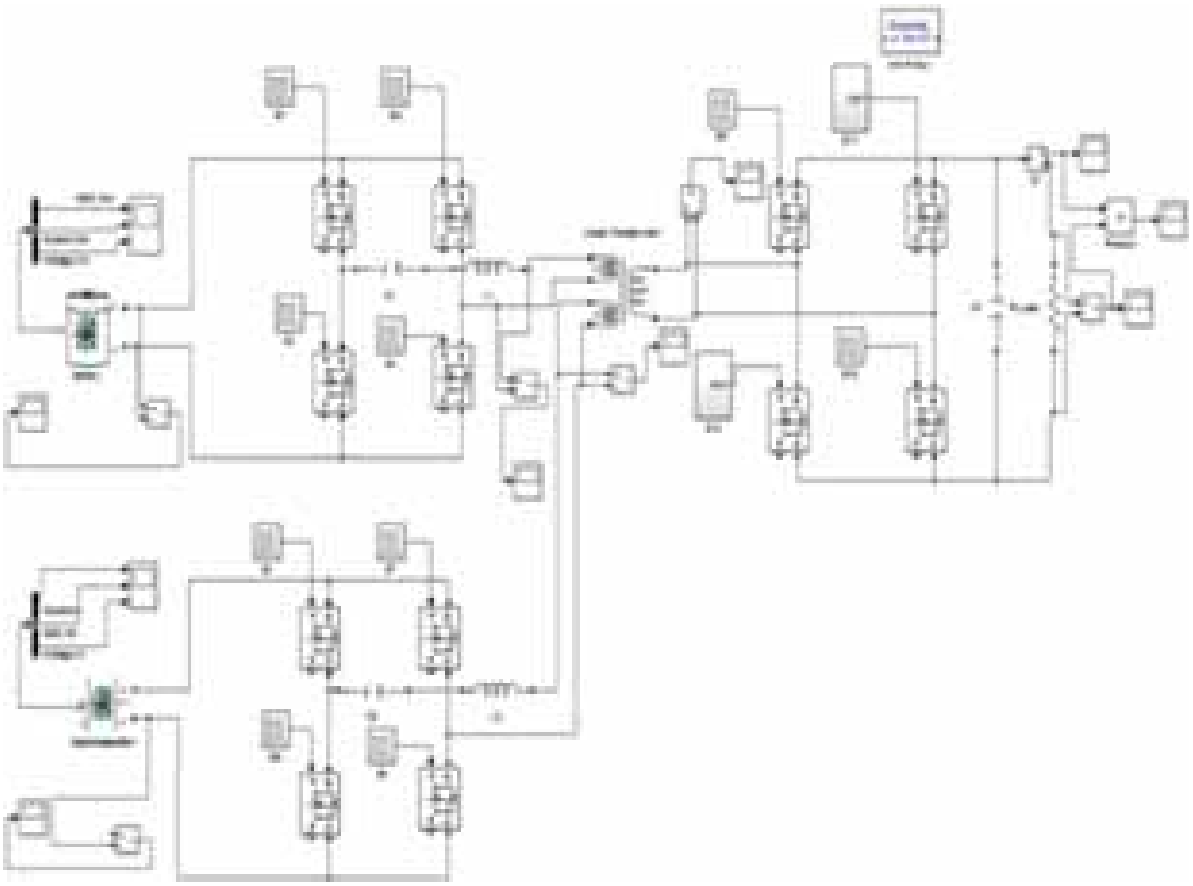


Fig. 5. Simulation diagram of multiport isolated bidirectional dc–dc converter with *R*-load for mode 1 and 2 operation.

port is in charging mode, the yield energy of port 2 is negative.

4.3. Mode 3

In mode 3 port 1 and 2 are charged by the DC-engine. Control is nourished from the regenerative load when the produced emf is more prominent then the terminal voltage. This power streams in the invert heading through port 3 to the port 1 and port 2 along these lines battery and super capacitor will get charged as appeared in Figure 5 and subsequently these ports will goes about as rectifier. The stage move between port 1 and port 2 Φ12 is 240. The stage move amongst port 1 and port 3 Φ13 is 00. Port 3 goes about as inverter.

5. ANALYSIS AND DESIGN

For design of proposed multiport isolated bidirectional dc–dc converter the frequency factor of 1.1 is chosen and the quality factor of 4 is chosen[2] for the design purpose and it is calculated from the following equations.

$$\text{Frequency ratio} = \frac{ws}{wi} = \frac{2 * \pi * fs}{1/\sqrt{LiCi}} \tag{1}$$

$$\text{Quality factor } Qi = \frac{Z1}{8 * R * n_{13}/\pi^2} \tag{2}$$

$$\text{Impedence } zi = \sqrt{\frac{Li}{Ci}} \tag{3}$$

The inductor values for series resonant tank is calculated from the following Eq. (7). Tables I–III provides the basic design specifications. The value of capacitor is

Table I. Converter design specifications.

Parameters	Values
Port 1-Battery port input voltage	48 Volts
Port 2-Supercapacitor port input voltage	35 Volts
Port 3-Load port output voltage	200 Volts
Switching frequency	100 KHZ
Turns ratio between port 1 and 3	0.24
Turns ratio between port 2 and 3	0.18
Inductance of series resonant tank 1	28.44 μH
Inductance of series resonant tank 2	15 μH
Capacitance of series resonant tank 1	0.1 μF
Capacitance of series resonant tank 2	0.22 μF
Output current	2.6 A
Output power	500 Watts
Resistance *R*	800

Table II. Supercapacitor cell specifications.

Parameters	Values
Equivalent capacitance of the super capacitor cell	50 Farads
Voltage rating of the super capacitor cell	2.7 Volts
Equivalent DC series resistance ESR	20 mA
Maximum surge voltage of a cell	2.85 Volts
Maximum leakage current	0.075 mA

Table III. Battery design specifications.

Parameters	Values
Battery voltage	48 Volts
Ampere–hour rating	12 Ah
Initial state of charge	90%

chosen such that frequency factor is 1.1.

$$Li = \frac{v_i * v_i * T}{162 * P} \tag{4}$$

Where T is the time period and p is the output power in watts. The output voltage, current and power is calculated from the following equation

$$V_0 = \frac{(V1/n_{13}) * \sin \varnothing_{13}}{Q1(F1 - 1/F1)} + \frac{(V2/n_{12}) * \sin(\varnothing_{12} - \varnothing_{13})}{Q1(F1 - 1/F1)} \tag{5}$$

$$I_0 = \frac{8}{\pi^2} * \frac{(V1/n_{13}) * \sin \varnothing_{13}}{Q1(F1 - 1/F1)} + \frac{8}{\pi^2} * \frac{(V2/n_{12}) * \sin(\varnothing_{12} - \varnothing_{13})}{Q1(F1 - 1/F1)} \tag{6}$$

If the load is resistive load then output power is calculated as

$$P_0 = I_0 * I_O * R \tag{7}$$

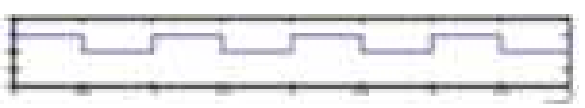


Fig. 6. Mode 1—Gate signal for switches S1, S4, S5, and S8.

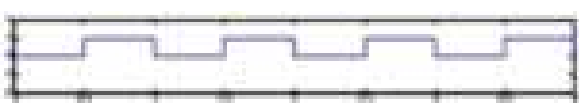


Fig. 7. Mode 1—Gate signal for switches S2, S3, S6, S7.

Fig. 8. Mode 1—Gate signal for switches S9, S12.

Fig. 9. Mode 1—Gate signal for switches S10, S11.

Fig. 10. Port 1—Battery port input voltage.

Degree of phase shift φ_{13} between port 1 and 2 and Φ_{12} between port 1 and 2 is calculated from the following equation.

$$\varnothing_{13} = \frac{\sin^{-1}(I_{1(\text{ref})} Z1(F1 - (1/F1))}{V_{0(\text{ref})} n_{13}(8/\pi^2)} \tag{8}$$

$$\varnothing_{12} = \frac{\varnothing_{13} - \sin^{-1}(I_{0(\text{ref})} - ((8/\pi^2) * n_{13} V_1(\sin \varnothing_{13})))}{((8/\pi^2) n_{23} V_2)/Z1(F1 - (1/F1))} \tag{9}$$

Super capacitor bank of 35 volt is modelled by connecting 13 super capacitor cells each of 2.5 volt. In order to increase the net energy capacity of the super capacitor bank 14 cells are connected in parallel. In order to increase the net energy density, the number of parallel connected cells should be higher than the number of series connected cells.[12] The total of 182 cells is used to model a super capacitor bank of 35 volts Lead acid battery of 48 volt with 12 ampere hour rating is modelled to act as dc source for port 1. The Gate signal for switches S10 and S11 is shown in Figures 9 and 10 depicts the battery input.

6. SIMULATION RESULTS

6.1. Open Loop Simulation Results

In Figure 5 the simulation circuit diagram of multiport isolated bidirectional dc–dc converter is shown with battery as input for port 1 of 48 V and super capacitor as a input for port 2 of 35 V and port 3 is the load port (R-load for mode 1 and 2 and DC motor load for mode-3) of 200 V. All the three ports are active bridges the switches are realized using MOSFETS. Simulation for mode 1 operation in which both battery and super capacitor supplies the resistive load. Load port output voltage and current for mode 1 operation is shown in Figures 11–13. In Figures 5–8,

Fig. 11. Port 2—Super capacitor port input voltage.

Fig. 12. Port 3—Load port input power.

Fig. 13. Port 3—Load port output current.

Fig. 14. Mode 2—Gate signal for switches S1, S4.

X-axis represents voltage in volts and *Y*-axis represents time in seconds.

Simulation result for mode 2 operation is shown above in which battery is in discharging mode since it supplies load and super capacitor. Load gets reduced by half as shown in Figure 21 since it is supplied by port 1 alone. State of charge of battery is decreasing as shown in Figure 18 and super capacitor is in charging mode as shown in Figure 19, state of charge of super capacitor is increasing from its initial value of 93%.

From the Figure 20 to 21 it is inferred that the super capacitor port sunks power so that the output power of port 2 is negative of the value of −189 watts. In Figures 14–17, *X*-axis represents voltage in volts and *Y*-axis represents

Fig. 15. Mode 2—Gate signal for switches S5, S8.

Fig. 16. Mode 2—Gate signal for switches S9, S10.

Fig. 17. Mode 2—Gate signal for switches S9, S10.

Fig. 18. Port 1—Battery state of stage.

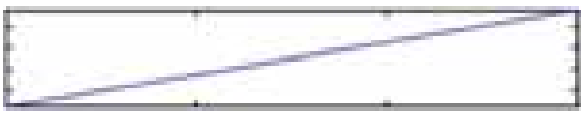


Fig. 19. Port 2—Super capacitor state of stage.

time in seconds. Gating signal for opposite pair of switches of each port are controlled complementarily. Simulation result for mode 3 operation is shown in which both the battery and super capacitor is charged by the DC motor load which acts as DC-generator in regenerative mode as shown in Figures 25 and 26 the state of charge of both battery and super capacitor is increasing. In regenerative mode power is negative so the torque is negative of −42 N-m as shown in Figure 27 and the speed is positive of 16 rad/sec as shown in Figure 28. From these results bidirectionality and regenerative operation is inferred. In Figures 22–24, *X*-axis represents voltage in volts and *Y*-axis represents time in seconds. Gating signal for opposite pair of switches of each port are controlled complementarily. In Closed loop control technique the gating signals are controlled with the help of Proportional plus Integral (PI) controller as shown in Figure 29. The output voltage from port 3 is compared with the reference voltage signal of 200 Volts. The error signal from the comparator is fed to the PI regulator where the error signal is regulated and fed to the input side. Such that despite of any changes in the input

Fig. 20. Port 2—Super capacitor port output power 1.

Fig. 21. Port 2—Super capacitor port output power 2.

Fig. 22. Port 3—Load port output power.

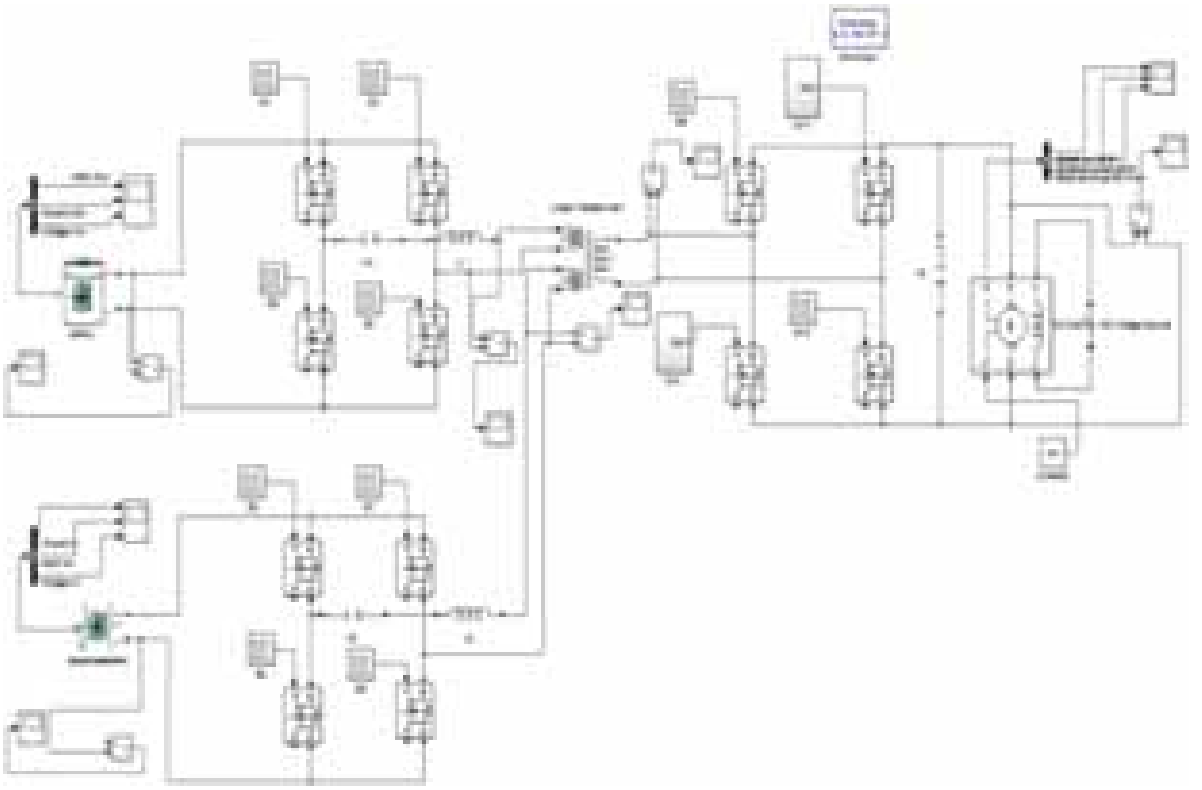


Fig. 23. Simulation circuit diagram with DC motor load for mode 3 operation.

Fig. 24. Mode 3—Gating signal for switches S1, S4, S9, S12.

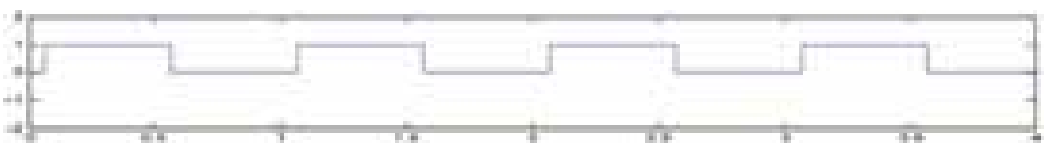


Fig. 25. Mode 3—Gating signal for switches S5, S8.

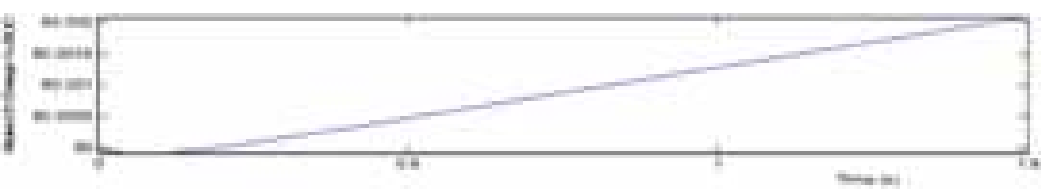


Fig. 26. Mode 3—Port 1—Battery port state of charge.

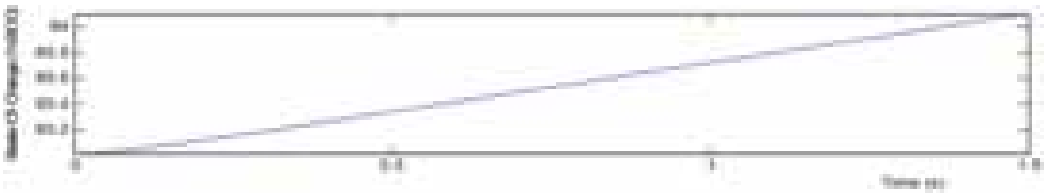


Fig. 27. Port 2—Supercapacitor port state of charge.

Fig. 28. DC motor speed.

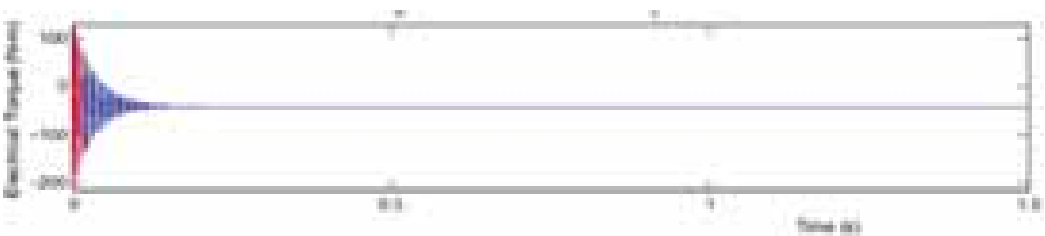


Fig. 29. DC motor torque.

side parameters the output side parameters is made constant. With phase shift control technique in closed loop simulation the output voltage obtained is 187 volts and output current is 2.3 amps as shown in Figures 30 and 31 but the expected voltage is 200 Volts and output current is 2.5 amps. This might be due to the drawbacks of phase shift control technique which is overcome in the pulse width modulation plus phase shift control technique which is extended future work.

Fig. 30. Port 3—Load port output voltage.

Fig. 31. Port 3—Load port output current.

7. CONCLUSION

A multiport dc–dc converter in which multiple ports are voltage fed ports is proposed in this paper to interface multiple energy storage systems such as battery and super capacitor. It is shown that from the analysis and simulation results the power flow between the ports can be controlled by the series resonance and phase shifting the square wave outputs of the three active bridges. With this technique there exists a drawback such as secondary side voltage overshoot due to voltage mismatch between the input and output side of the input ports. In order to overcome the drawback of phase shift control technique pulse width modulation plus phase shift control technique is to be implemented which is feature work.

References

1. Amari Mansour, Bacha Faouzi, Ghouili Jamel, and Elgharbi Ismahen, *International Journal of Hydrogen Energy* 39, 1580 **(2014)**.
2. B. S. Nathan and V. R. Ramanarayanan, Analaysis simulation and design of series resonant converter for high voltage applications, *Proceedings of IEEE International Conference on Industrial Technology*, December **(2000)**, Vol. 1, pp. 688–693.
3. C. Zhao and J. W. Kolar, A novel three phase three port UPS employing a signal high frequency isolation transformer, *Proc. IEEE. Power*

Electronics Specialists Conference (PESC), Aachen, Germany, June **(2005)**, pp. 4135–4141.
4. H. Tao, A. Kotsopoulos, J. L. Duarte, and M. A. M. Hendrix, Triple half-bridge bi-directional converter controlled by phase shift and PWM, *Proc. IEEE Appl. Power Electron. Conf. Expo (APEC)* **(2006)**, pp. 1256–1262.
5. H. Tao, A. Kotsopoulos, J. L. Duarte, and M. A. M. Hendrix, *IEEE Proc. Electr. Power Appln.* 153, 451 **(2006)**.
6. H. Al Atrash, F. Tian, and I. Batarseh, *IEEE Trans. Power Electron.* 22, 341 **(2007)**.
7. H. Tao, J. L. Duarte, and M. A. M. Hendrix, *IEEE Trans. Power Electron.* 23, 782 **(2008)**.
8. H. Krishnaswami and N. Mohan, A current-fed three port bi-directional dc–dc converter, *Proc. IEEE Int. Telecommun. Energy Conf. (INTELEC)* **(2007)**, pp. 523–526.
9. H. Krishnaswami and N. Mohan, Constant switching frequency series resonant three port bi-directional dc–dc converters, *Proc. IEEE Power Electron. Spec. Conf. (PESC)* **(2008)**, pp. 1640–1645.
10. H. Krishnaswami and N. Mohan, *IEEE Trans. Power Electron.* 24, 2289 **(2009)**.

Received: 21 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1625–1628, 2018

Effects of SQL Query Inclusive Computation in Application Response Time

A. M. Palanisamy*, S. Sangeetha, and R. V. Nataraj

Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamilnadu, India

In evaluating application performance with respect to end user perception, response time is one of key parameters known as Application Response Time (ART) used to evaluate a software application. ART is determined and affected by two elements such as network response time and transaction response time. Our work is concentrated in optimizing transaction response time of an application by SQL query inclusive computation. In addition to overall application response time, the effects of this approach are examined against query execution time, function process time on a web based application and the results are presented in this paper. The experimental result shows that our approach improves web applications response time considerably.

Keywords: SQL Query Inclusive Computation, Response Time, Query Execution Time, Function Process Time, Application Latency, Database Latency.

RESEARCH ARTICLE

1. INTRODUCTION

Application's transaction response time is calculated by the summation of server and client response time. The server response time is highly significant than client since it deals only with browser and navigation latencies which are negligible. On the other hand, server side response time is relying mainly on application latency, database latency, I/O latency. To optimize server response time, it is necessary to reduce application and database latencies. Many authors proposed different methods in their research work to optimize application latency by improving hardware architecture, pipeline mechanisms, cache memories, and effective resource allocation. In the case of database latency, attention has to be paid on SQL query that used to retrieve result by accessing data stored at database. The typical SQL execution unit is the query which is a collection of statements that return a result. These statements can modify the structure of a database using DDL (Data Definition Language Statements) and manipulate the content of the database using Data Manipulation Language Statements (DML). Rekha Singhal and Manoj Nambiar[3] suggested to host the application on the database server to avoid the time delays which may be introduced in the query result due to the query processing at the web server. Hossein Sheikh Attar and Tamer O. Zsu[6] proposed an alternative architecture that caches both static content, database tuples, query results at proxy servers which reduces network latency and improve the overall performance. Other optimization techniques[4] such as simplification of queries based on known precondition, memorizing the results of certain computation for later use, performing look ahead computation yields significant speedups. Toyama and Nagafuji[8] stated that a small number of query rewriting rules provide the equivalence of a query and its decompositions. Incremental query evaluation gives application system an enhanced capability for designing database, web applications to meet performance and functionality requirements while retaining simplicity of description. With several techniques and approaches discussed in literature survey, the optimization takes place at hardware level. SQL query inclusive computation approach optimizes at software level. In web based application system, SQL query inclusive computation approach can be implemented by means of SQL query rewriting with aggregated functions, Target Form Expressions (TFE) which provides a way to tune an application and its environment to perform better. TFE is a database publishing presentation extension of SQL that yields a query result presented as a document in any of several target media, for example, HTML, Java, LaTeX.

This article studies the difference in response time of a transaction executed through SQL query inclusive computation and without SQL query inclusive computation. This article is structured as follows. Section 2 describes the difference between the general and SQL query inclusive

*Author to whom correspondence should be addressed.

computation based request–response flow in web application. Section 3 illustrates the implementation of proposed work. Section 4 demonstrates the evaluation of proposed approach. Finally, conclusion in Section 5.

2. METHODOLOGY

Modern web application development follows *n*-tier architecture which separates layers of presentation, application business, and data. For ease, the application tire is even decomposed and integrated. In web based information retrieval system, user has to pass input parameters through presentation tier to application tier. The application receives input, coins an SQL query, and sends it as request to database server where the SQL query is executed and returns the fetched data as response to the application. To present the expected, visualized information to user, the response data from database server have to be parsed, computed accordingly in the application tier and information is set back to user. The basic request–response flow in *n*-tier architecture is depicted in Figure 1.

The proposed SQL query inclusive computation approach eliminates response computation process in application tier. Instead the associate response computation process in injected in SQL query and passed as request to database server. SQL query inclusive computation reduces the time taken to process response data at application tier and hence optimizes the application latency. Figure 2 depicts the request–response flow of the proposed approach.

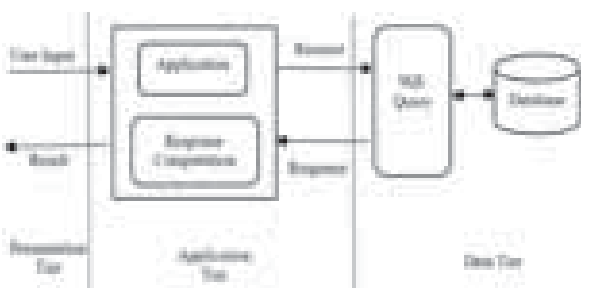


Fig. 1. Request–response flow in web applications.

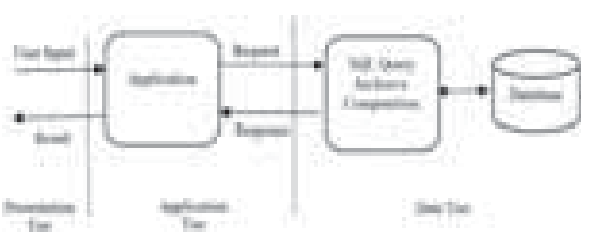


Fig. 2. SQL query inclusive computation based request–response flow in web applications.

3. IMPLEMENTATION

SQL query inclusive computation approach is implemented in a web based ERP application. The run time policy mainly depends on query actions and update actions.[4] When the query actions are optimized, it improves application's runtime. SQL query inclusive computation is devised in two different styles. To avoid parsing response data at application, one style is to construct and execute a separate SQL query with inclusive computation. The other style is altering an existing SQL query by adding required formula and textual data to it. To explain the former case, let consider a task of generating monthly payroll report of an organization. The payroll report should contain data about staff members name, identity number, department, and the amount paid to them under different salary heads such as basic, dearness allowance, provident fund, medical allowance and others. To fetch data for the report from database, developer used to write up an SQL query "SELECT field_list FROM table_name WHERE [condition]." The SQL query will be placed as a request to database server where the query will be executed and will send the result set to the application server. As a summary, the total amount value of each column should also be take place in the report. To accomplish this, generally developers used to declare separate variables each for a column that take place in the report to sum up the value by processing each iteration of the result set as mentioned in the Figure 3. This computation at application side will degrade its response time. This response result set processing at application server can be avoided by SQL inclusive computation. Instead of writing such code as in Figure 3, a separate SQL query inclusive computation can be devised as depicted in Q1 below. This additional request of this query computes only expected aggregated value and returns result set with single row.

Q1: SELECT SUM(col_1),. . .,SUM(col_*n*) FROM table_name WHERE [condition] [GROUP BY/ORDER BY].

To explain the latter case, let consider the visualization of result set data in the formats like HTML, PDF, Excel.[8]

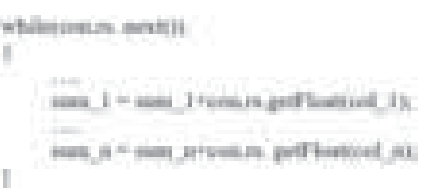


Fig. 3. Sample code to process result set data.

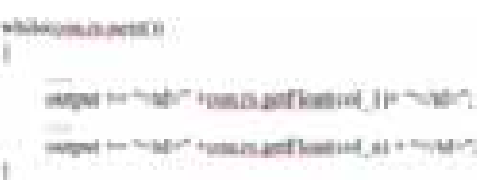


Fig. 4. Sample code to process result set data with formatting text.

The result set data has to be processed by appending certain formatting text data as shown in Figure 4. Instead of writing such formatting text as in Figure 4, the required formatting text can be included in the SQL query with function like CONCAT_WS() as depicted in Q2 below.

Q2: SELECT CONCAT_WS('</td><td>',field_list) FROM table_name WHERE [condition].

4. EVALUATION

Without query inclusive computation, data are merely sent to application server where the response data processing would take place. When more number of clients place their requests, the application server will get overloaded with processing clients' response data and load balancing has to perform in application server. This can be avoided with query inclusive computation since response data need not to be processed at application server. Even though the application response process time is considerably reduced by SQL query inclusive computation approach, the time taken to execute SQL Query Inclusive computation at data tier has also be taken into the account. This proposed approach is evaluated with respect to overall application response time, SQL query inclusive computation time at data tier.

4.1. Application Response Time Evaluation

In the aspect of application response time, SQL query without inclusive computation takes more and longer time. Sometimes in web application execution, it took longer time that led to session expire. Whereas the SQL query with inclusive computation overcomes the above said limitation and gives improved application response time. This evaluation process considers only the transaction response time as basic metric and other matrices associated with web server, application server performance are not considered since those are not basic requirements to the context of the work.

An SQL query with inclusive computation is constructed to address the former case as discussed earlier with Q1. Both queries with and without inclusive computation are executed upon 10470 tuples of a relation

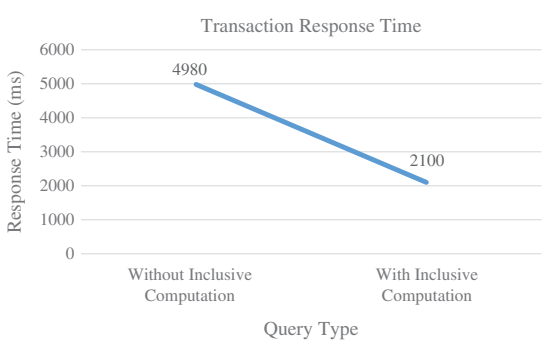


Fig. 5. Comparison of transaction response time: Case Q1.

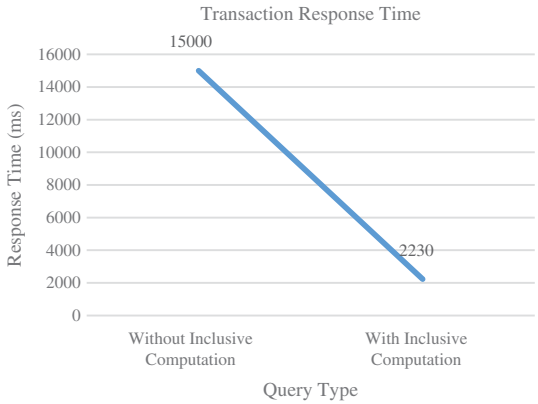


Fig. 6. Comparison of transaction response time: Case Q2.

and evaluated the application response time under both cases. The result is depicted in Figure 5, that shows query with inclusive computation optimizes the response time. To evaluate the later case as discussed earlier with Q2, queries with and without inclusive computation are exercised to generated a report holding 1100 rows. The result of later case is depicted in Figure 6, that too shows query with inclusive computation optimizes the response time considerably.

4.2. Query Execution Evaluation

With respect to SQL query execution, devising and executing a separate SQL query with inclusive computation

Table I. TFE SQL query execution.

Parameters	Duration in sec.	
	SQL query with inclusive computation	SQL query without inclusive computation
Copying to temporary table	9.71304	9.79498
Sending data	0.14391	0.08618
Statistics	0.01679	0.02000
Starting	0.01425	0.00605
Checking permissions	0.00670	0.00292
Sorting result	0.00126	0.00122
Optimizing	0.00120	0.00097
Removing temporary table	0.00094	0.00073
Opening tables	0.00065	0.00060
Updating status	0.00056	0.00059
Cleaning up	0.00045	0.00043
Freeing items	0.00022	0.00042
Preparing	0.00020	0.00016
Sorting for group	0.00006	0.00007
Closing tables	0.00005	0.00004
After opening tables	0.00004	0.00004
End	0.00002	0.00004
Init	0.00001	0.00001
Table lock	0.00001	0.00001
Executing	0.00001	0.00001
Query end	0.00001	0.00001
System lock	0.00001	0.00001
Total	9.90039	9.91549

RESEARCH ARTICLE

Table II. Parameter value-change status of TFE SQL query execution.

Variable	Value		Description
	SQL query with inclusive computation	SQL query without inclusive computation	
Created_temporary_tables	2	1	Count of temporary tables created in memory
Handler_commit	1	1	Number of internal commit statements
Handler_read_key	3007	3007	Number of requests to read a row based on a key
Handler_read_next	66242	66242	Number of index columns read with a range constraint or an index scan
Handler_read_rnd	1090	1090	Number of requests to read a row based on a fixed position. A high number indicates sorting of results or full table scans
Questions	1	1	Number of statements executed by the server
Select_range	1	1	Tables were read from the disk only in the necessary places to satisfy a limited range of conditions
Sort_rows	1090	1090	Number of rows sorted
Sort_scan	1	1	Number of sorts that were done by scanning the tables
Table_locks_immediate	16	16	The number of requests for table locks that could be granted immediately

has minimal impact since in case of Non-TFE SQL query, since it is not embedded with additional formatting data and it is not necessary for comparison evaluation. Whereas the other case has to be consider for comparison evaluation between the SQL queries with and without inclusive computation. An SQL query that considered for comparison evaluation has to perform one left outer join of two relational tables, one with 316441 records and other with 4016 records. The SQL query inclusive computation is embedded in the query considered for evaluation and executed in both fashions with SQL_NO_CACHE option. The evaluation parameters and the value against it with respect to SQL query execution time is illustrated in Table I. The data in Table I shows both SQL query with and without inclusive computation has no high variation in their execution time. The data in Table II shows the parameter value change status while executing queries with and without inclusive computation. It is found that no high variation in value changes after executing queries.

5. CONCLUSION AND FUTURE WORK

This experimental study on effects of SQL query inclusive computation in application performance proves that web application response time improves considerably. In addition to that, when optimizing response time of many transactions, the system throughput will also be optimized. This approach works well in relational databases with categorical data. Further this study has to extend experimenting this approach in columnar databases, text based IR system.

References

1. J. Armas, P. Navas, T. Mayorga, P. Rengifo, and B. Arévalo, Optimization of code lines and time of access to information through object relational mapping (ORM) using alternative tools of connection to database management systems (DBMS), *2nd International Conference on System Reliability and Safety (ICSRS)* **(2017)**, pp. 500–504.
2. A. Haris Hasanbegovic and B. Novica Nosovic, Evaluation of typical transaction based information systems response time—Lessons learned, *22nd Telecommunications*, Serbia **(2014)**, pp. 963–968.
3. Rekha Singhal and Manoj Nambiar, Extrapolation of SQL query elapsed response time at application development stage, *Annual IEEE India Conference* **(2012)**.
4. Rekha Singhal, A framework for predicting query response time, *14th IEEE International Conference on High Performance Computing and Communication* **(2012)**, pp. 1137–1141.
5. M. Akdere, U. Cetintemel, M. Riondato, E. Upfal, and S. Zdonik, Learning-based query performance modeling and prediction, *IEEE 28th International Conference on Data Engineering*, April **(2012)**, pp. 390–401.
6. M. Hossein Sheikh Attar and M. Tamer O. Zsu, *World Wide Web* 9, 21 **(2006)**.
7. D. Florescu, A. Levy, D. Suciu, and K. Yagoub, Run-time management of data intensive web-sites, *Proc. 25th International Conference on Very Large Data Bases* **(1999)**, pp. 627–638.
8. M. Toyama and T. Nagafuji, Dynamic and structured presentation of database contents on the web, *International Conference on Extending Database Technology*, Advances in Database Technology **(1998)**, pp. 451–465.

Received: 21 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1629–1633, 2018

Exploration of Service Oriented Architecture for Online Course Registration

K. Parthiban* and R. V. Nataraj

Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamilnadu, India

Each program has its own predefined curriculum. Throughout the academic system, Course registration plays a vital role in educational institutions. Managing the course registration for a semester is a tedious task as it involves several aspects such as formation of courses, allocation of course instructor and time table. Additionally, the course registration will be more tedious for the FCFS based registration as it involves to opt multiple course instructors offering multiple courses. This paper mainly focuses on exploring a Service Oriented Architecture by integrating the curriculum, Instructor management, timetable management and to deliver a highly available, reliable online course registration platform for the students.

Keywords: FCFS, Spanning Tree, CBCS.

RESEARCH ARTICLE

1. INTRODUCTION

In recent days the Choice Based Credit System (CBCS), is implemented for the students benefit which gives more flexibility to the them to select the courses for their semester as well as the course instructor for the selected course.

As the course selection process is generalized throughout the entire duration of the programme, it is difficult to maintain the registration process manually.

In this paper, we propose an architecture for managing the course registration online by integrating the curriculum, faculty and timetable allocation. The spanning tree based approach has been implemented for course registration of individual student.

2. PROBLEMS OF EXISTING COURSE REGISTRATION SYSTEM

Existing course registration system is a not fully compliant with automation. As shown in the Figure 1, the course registration is not only meant for the conduction of courses, but it also focuses the basement for further operations related to final end semester examination, including hall ticket generation, result processing and grade sheet generation.

Several problems arises in existing system is given below.

A. Curriculum management: The curriculum has been managed electronically by the application where the Board of Studies member of the department has to feed the curriculum.

B. Timetable management: The Timetable in charge of the department has to feed the time table of the individual instructor. The is no link between the timetable of the faculty and the student.

C. Student Allocation: Each individual student has to allocate for the instructor handling the course manually by the incharge. The problem behind this is that the student cannot have any provision to select the course or the faculty by self.

D. Exam Management: After the course has been completed the process of assessment for each individual has been maintained manually. The problem behind this is that there is no interlink between the curriculum management, staff management and the exam management.

In the case of CBCS based academic system the student has their own provision for selecting the course and staff on their own based certain protocols in FCFS basis. Following to the CBCS based system, it is very difficult to maintain the registration manually. The post process after the course completion is a tedious task. One cannot track the academic history of the individual student manually. The Service oriented architecture for the CBCS based course registration system has been presented and implemented in the following section.

*Author to whom correspondence should be addressed.

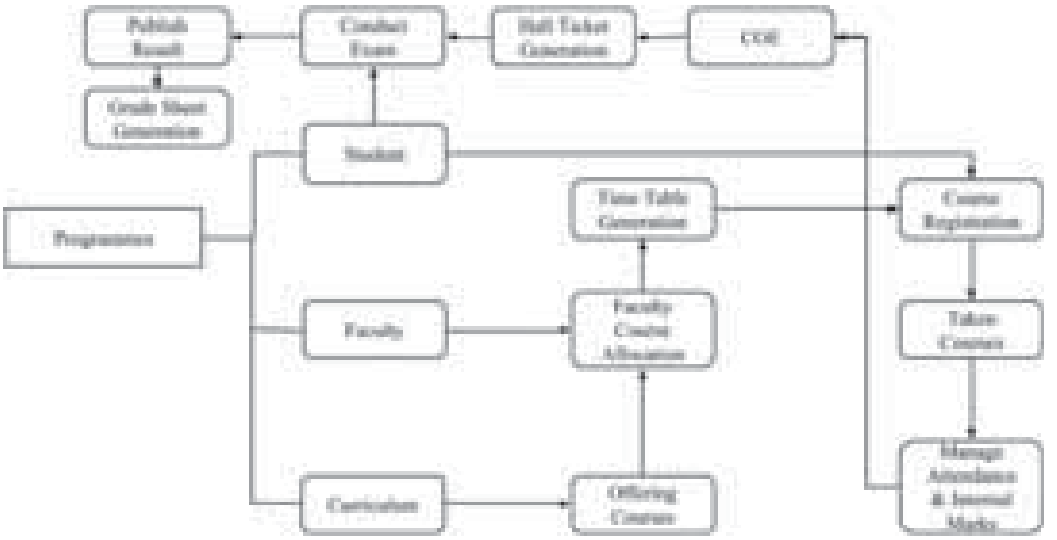


Fig. 1. Architecture of online academics system.

3. PROPOSED SOFTWARE APPROACH

The purely enterprise software solution has been proposed to overcome the difficulties of course registration process. As the architecture is purely enterprise, the core of course registration system such as curriculum, timetable, staff and student management has been integrated so that the fact has been stored electronically. As a result, the final output of the course registration such as the exam management is also being maintained as the part of the enterprise application.

The benefits of proposed approach are projected below:

- The Enterprise web based solution is easy to access in any location which allows the stakeholders to access same information and does not have to maintain a private copy of it.
- As it is enterprise, the fact has been delivered to any level of users. No duplication or wrong information will be shown.
- Integrated with curricula, so that anyone can allow to access their own curriculum details and can select on their own based on the credits.
- Enterprise system allows the student to involve on their own to select the course in their own private portal. The tree based approach allows the student to simulate on them through the dynamically generated timetable.
- Effective time management, the proposed method allows ease of use where the time taken for allocation and selection process has been reduced drastically than the manual method.

The following Figure 2 represents the flow of the proposed method.

The above mentioned flow diagram represents the workflow of the course registration. Once the issue of the course registration is assigned by the admin, the student can able to view the course and the slot slots respect to the courses. Student can able to choose the course and the slot dynamically. Once the slot has been chosen the timetable for the

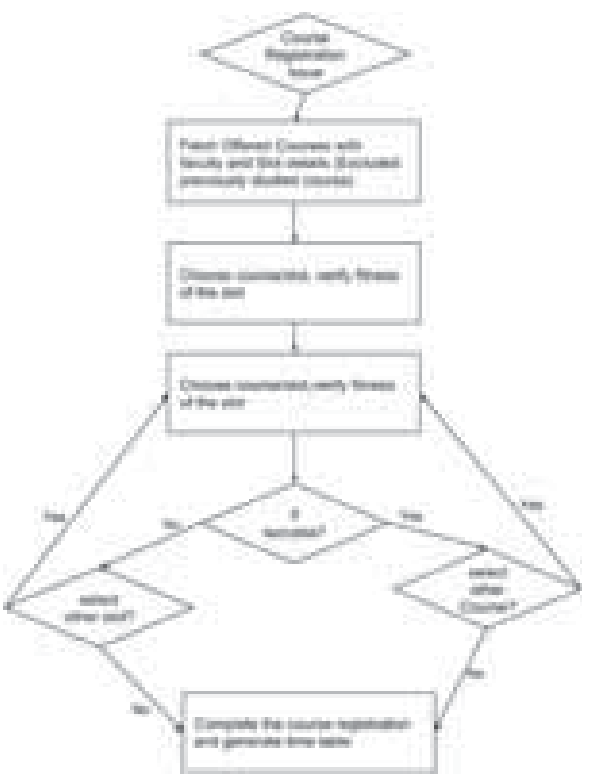


Fig. 2. Flowchart of online course registration system.

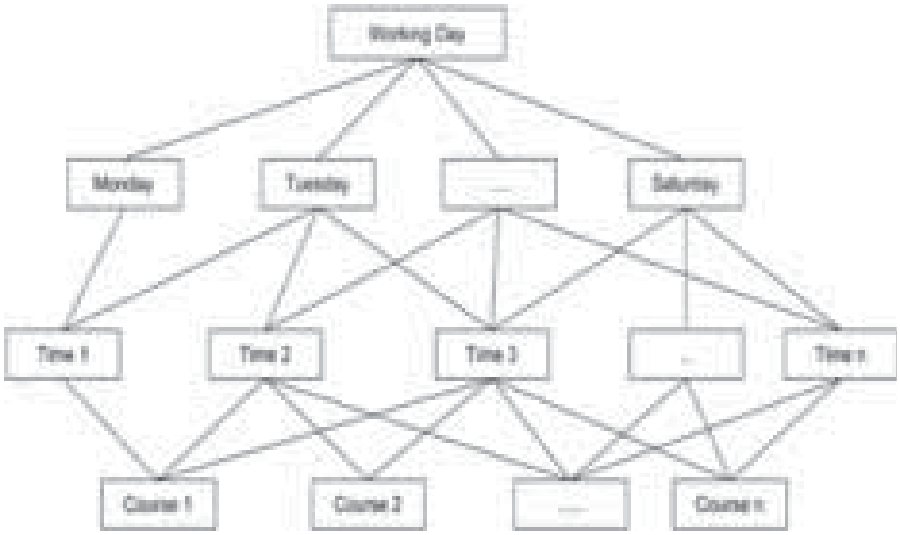


Fig. 3. Spanning tree approach of dynamic timetable generation.

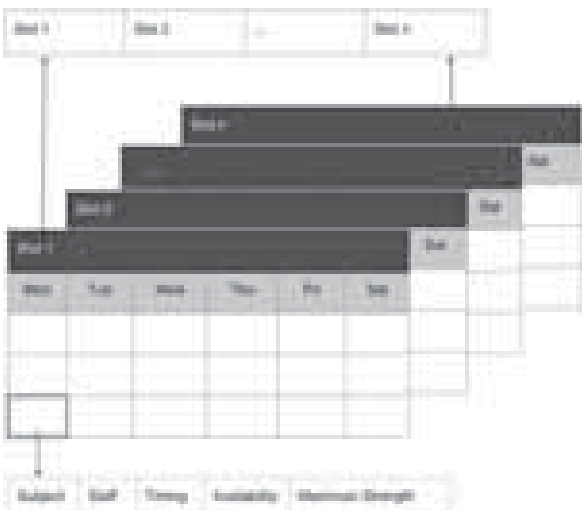


Fig. 4. Representation of allocation of timetable.

particular slot has been assigned dynamically, as illustrated in Figure 3. The student after ensuring that there is no clash between the slots, will continue to choose the other courses. Once all the courses offered to the student has been chosen the final timetable will be generated for the particular student.

4. INCORPORATING SPANNING TREE IN ONLINE COURSE REGISTRATION SYSTEM

The workflow of the proposed online course registration system is based on the Spanning tree approach. The timetable itself creates a spanning tree where it denies the cycles among the timetable creation. This indeed reduces the traversal time of the timetable of the individual course of the student.

The Figure 3 represents the Timetable of a course respect to the time and the day. Here every course may

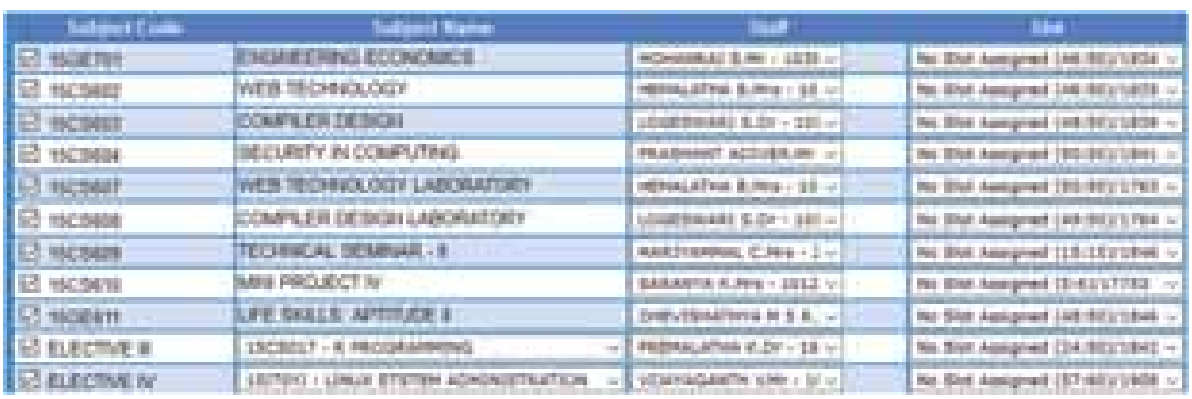


Fig. 5. Dashboard of course registration of a student.

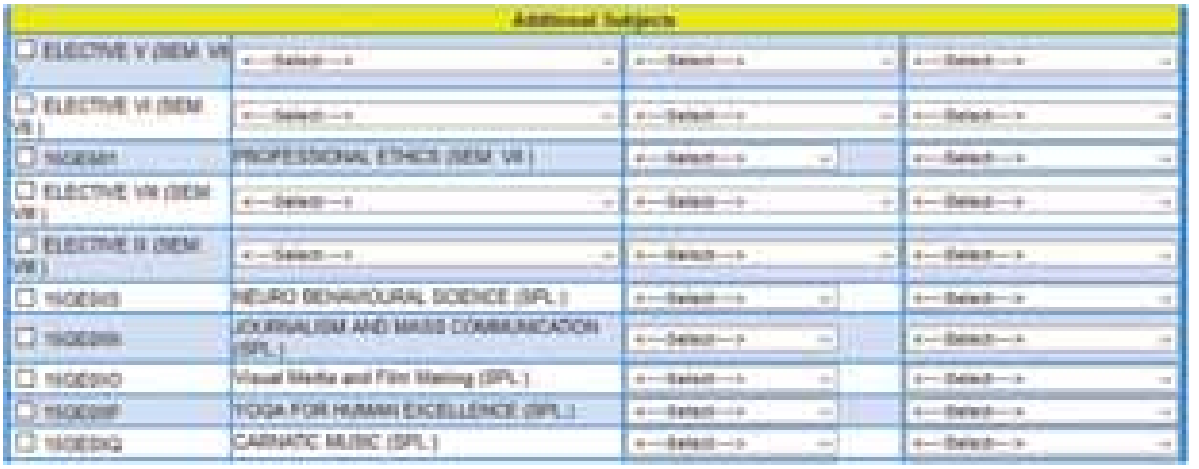


Fig. 6. Dashboard for registration of additional courses.

have related to multiple days as well as multiple timings of the same day. This indeed should not create a cyclic redundancy as it will lead to the deadlock. The spanning tree approach overcomes such occurrence by neglecting the false tuples.

5. ALGORITHMIC REPRESENTATION

The algorithmic representation for online course registration is illustrated below:

Data: (i) *Vector S contains selected slot of students*,
(ii) *C represents the chosen slot of specific subject*

Fig. 7. Dynamic timetable generation w.r.t selected slots for a student.

Output: *Acceptance or Rejection of the chosen slot*
Timetable vector T & T′
// retrieve all the timetable of each selected slot
For each $(s \in S)$
    attach $(T,$ *get timetable*$(s))$;
// retrieve the timetable for selected subject
$T' =$ *get timetable*(C);
//verify the slot chosen is already blocked by another slot
For each $(t \in T')$
        if free$(t,\ T) =$ *true*
            attach$(T,\ ti)$;
        Else
            return("*Reject*");
//accept the chosen without conflict
return("accept");

The above representation illustrates the process of CBCS based course registration provided when the curriculum allocation for the program, instructor allocation for the individual course and the timetable allocation for the instructor.

Once the above condition satisfies, the student prompted to the online course registration system the showcases the set of vector 'v' containing the pre allocated slot for the courses. Students prompted to select the courses and the instructor in FCFS basis. The temporary vector represents the timetable for the selected slot dynamically. The Online course registration prompted to success only when all the slots fit into the timetable of the individual student ensuring that there are no conflicts between the slots/course.

Figure 4 represents the allocation of timetable in the manner that the student choosing the slot dynamically.

The representation shows that there are n-number of slots and each slot contains n-number of courses respect to the timing of the course. Student can choose the any slot in FCFS manner. The slot for each course is directly controlled by the count assigned to it. One cannot choose the course if there is any conflict between the slot or between the courses.

6. USER INTERFACE

The Core subject for the specific semester which is enabled will be displayed in Figure 5. Students can choose the course instructor of the specific course with respect to the slot allocated. Student can add or drop the n credits based on the curriculum.

This dashboard shown in Figure 6, shows the additional courses of the particular semester with respect to the timetable provided no conflicts may occur. The allocation of additional courses may limit to the 'M' credits of the particular semester.

Once the course has been selected with respect to the courses opted through the portal, time table has been generated dynamically as shown in Figure 7. The timetable generated for a particular student may vary to one to another based on the add or drop of credits between them. The successful course registration of the student means that there is no conflict between the hours.

7. CONCLUSION

This paper deals with the architecture for managing the course registration online by integrating the curriculum, faculty and timetable allocation. This in result provides transparency and integrity of the curriculum management. The web access gives the added advantage by accessing securely.

This is also helpful for the students to choose their own courses independently respect to the course that the student like to study in the semester.

In future certain algorithm has to be implemented for data auditing as the data related to the online course registration is vast and need to ensure the security and reliability of the system. To ensure the scalability of the system, the top k query has to be implemented for query optimization as the dynamic timetable generation in course registration need an optimized query processing.

References

1. E. Bettinger, C. Doss, S. Loeb, A. Rogers, and E. Taylor, *Economics of Education Review* 58, 68 **(2017)**.
2. I. Ognjanovic, D. Gasevic, and S. Dawson, *The Internet and Higher Education* 29, 49 **(2016)**.
3. S. Channaruku, N. Saejiem, K. Bhumichitr, R. Jiamthapthaksin, V. Nicklamai, and K. Terdvikran, Social-aware automated course planner: An integrated recommender system for university registration system, *14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* **(2017)**, Vol. 17, pp. 544–548
4. E. S. Khorasani, Z. Zhenge, and J. Champaign, A Markov chain collaborative filtering model for course enrollment recommendations, *2016 IEEE International Conference on Big Data (Big Data)* **(2016)**.
5. J. Xu, T. Xing, and M. d. Schaar, *Personalized Course Sequence Recommendations*, IEEE **(2016)**.
6. C. Vrasidas and M. S. McIsaac, *American Journal of Distance Education* 13, 22 **(1999)**.
7. Y. AL-Bastaki, Al-Ajeeli, and Abid, *Journal of Computers and Education* 44, 327 **(2005)**.
8. L. Song, E. S. Singleton, J. R. Hill, and M. H. Koh, *The Internet and Higher Education* 7, 59 **(2004)**.

Received: 21 April 2018. Accepted: 16 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1634–1638, 2018

RESEARCH ARTICLE

Development of Multiport Converter for Hybrid Renewable Energy System

A. Angel* and A. Arunya Revathi

Department of EEE, Alagappa Chettiar Government College of Engineering and Technology, Karaikudi 630003, India

The proposed objective is to analyse the Boost derived Multiport converter (BDMC) for DC and AC loads by different kinds of energy sources using efficient power electronic converters. This Multiport converter can be used in grid application to convert the generated DC power to DC as well as AC at a time. The generation of power is taken from three renewable energy sources like solar, wind and fuel cell. The proposed Multiport converter is derived from the single switch controlled Boost converter by replacing the controlled switch with voltage source inverter (VSI). This new Multiport converter has the advantages like reduced number of switches as compared with conventional design having separate converter to supply AC and DC loads simultaneously. It provides DC and AC outputs with an increased reliability, resulting from the inherent shoot through protection in the inverter stage. Pulse width modulation (PWM) is used for controlling switches in the circuit.

Keywords: Multiport Converter, Boost Converter, PWM Technic.

1. INTRODUCTION

Energy is a fundamental aspect to people's life, and is essential not only for individuals but also for various sectors. It can be supplied from various resources which can be divided into two categories; renewable and non-renewable. Typical examples of non-renewable energy sources are petroleum, coal, and natural gas.[9]

Due to the decade of fossil fuels, to meet the load requirements renewable energy plays an important role. Renewable energy is the energy which comes from natural resources such as sunlight, wind, rain, tides and geothermal heat. These resources are renewable and can be naturally replenished. Therefore for all practical purposes, these resources can be considered to be inexhaustible, unlike dwindling conventional fossil fuels. The global energy crunch has provided a renewed impetus to the growth and development of clean and Renewable Energy Sources. Apart from the rapidly decreasing reserves of fossil fuels in the world, another major factor working against fossil fuels is the pollution associated with their combustion. Global warming is Caused by abundance of CO_2 in the atmosphere and the limitations of global resources of fossil and nuclear fuel has necessitated an urgent search for alternative sources of energy to meet the future demand. It is also important that the alternative energy sources minimize the negative environmental impact and cover the continuously increasing demand of power supply.

Renewable energy sources are known to be much cleaner and produce energy without the harmful effects of pollution unlike their conventional counterparts. Wind, solar and water energy which are non-pollution, free in their availability and renewable are considered as a promising power sources. The integration of renewable energy sources into the utility grid is an excellent option for meeting the future energy demand and environmental considerations.[3]

A proportional–integral-derivative controller calculates an error value as the difference between a desired set point and a measured process variable and applies a correction based on proportional, integral, and derivative terms, which give their name to the controller type.

A PID controller continuously calculates an error value as the difference between a desired setpoint and a measured process variable and applies a correction based on proportional, integral, and derivative terms. The controller attempts to minimize the error over time by adjustment of a control variable such as the position of a control valve, a damper, or the power supplied to a heating element, to a new value determined by a weighted sum.

1.1. Multiport Converter

Recently, there is a good attention on renewable energy sources and consequently the power electronic based systems to utilize these energy sources. The renewable energy sources depend on the environment conditions and their characteristics are variables. To provide a stable dc voltage,

*Author to whom correspondence should be addressed.

1546-1955/2018/15/1634/005 doi:10.1166/jctn.2018.7353

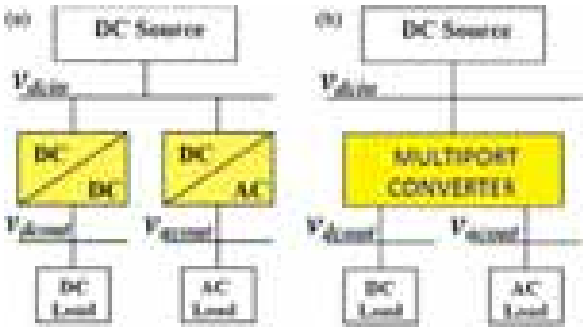


Fig. 1. Representative schematic of a multiport converter with a single dc input and simultaneous dc and ac outputs. (a) Conventional power converter-based architecture. (b) Multiport converter-based architecture.

it is necessary to use different types of dc renewable sources simultaneously. For these applications, the multiport dc/dc converters are the best solution to utilize the different types of dc voltage sources.[3] The multi-port dc/dc converters are divided into three main groups. These are single-input multi-output (SIMO), multi-input single-output (MISO), and multi-input multi-output (MIMO). Generally, these converters have simple topology, central control, high reliability and low construction cost.[5] In order to obtain a the safe and flexible electrical system with high reliability, it is necessary to use of renewable energy sources and or the combination of them with fuel cell and other energy storage elements. In this hybrid energy system (multi-input single output converters) to combine input energy sources with different electrical characteristics and generate desired output levels are used.[7] Using renewable energy sources in the multilevel inverters lead to generate converters that consist of low power switches with minimum equivalent series resistance. These converters are able to connect to the power grid without using any transformers. Moreover, it is impossible to connect the high voltage dc-link of the multilevel inverters to the low dc voltage levels of renewable energy sources; in addition, the multilevel inverters needs balance voltage level for dc-link in order to have suitable performance. However, it is important to note that the diode clamped multilevel inverter does not have the capability to balance dc-link voltage. Moreover, there are some limitations to use of high voltage elements in the high voltage applications of the conventional dc-dc converters. As a result, the single-input multi-output converter has received more attention in the applications in which photovoltaic systems or fuel cell with multilevel inverter that need different controlled voltage levels are used. In recent years, networks with renewable energy sources such as fuel cell, photovoltaic systems, wind turbines, batteries and ultra-capacitors have received more attentions because of environmental problems and increasing the reliability and flexibility. As a result, in this such systems, MIMO converters are most important role to combine and consolidate

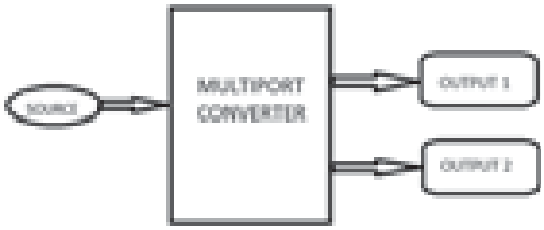


Fig. 2. Multiport converter.

the above mentioned energy sources with different current voltage characteristics, relation between these sources with output loads and generate different output voltage levels.[8]

The multi-port dc to dc converter is one of the common converters that widely used in dc renewable energy sources. The multi-port dc/dc converters are divided into single-input multi-output, multi input single-output, and multi-input multi-output converters. The multi-input multi-output dc/dc converters can combine the different types of dc voltage sources with different V–I characteristics in input side and generate different levels of dc voltages at the output side. The main aim of using these converters is reducing the cost and number of power electronic components and increasing the reliability and efficiency of the overall system. Until now, different topologies of multi-port dc/dc converters have been made. A Multi-port Power Electronic Interface (MPEI) is a self-sustainable multiple input/output static power electronic converter which is capable of interfacing with different sources, storages and loads, the integrated control system of MPEI enables both excellent system dynamic and steady state performance which renders optimal renewable energy harvesting, optimal energy management and optimal and economical utility grid interactions in a deregulated power market.

2. DESIGN OF BOOST CONVERTER

The boost converter has been designed to give an increased voltage 220 v. The formulas used for designing the converter is given in Eqs. (1), (2), (3).

$$\text{Duty cycle } (D) = 1 - (V_{\text{in}}/V_{\text{out}}) \tag{1}$$

Table I. Duty cycle V_s output voltage.

Duty cycle	DC output voltage
0.1	53.33
0.2	60
0.3	68.57
0.4	80
0.5	96
0.6	120
0.7	160
0.8	240
0.9	480
1.0	0

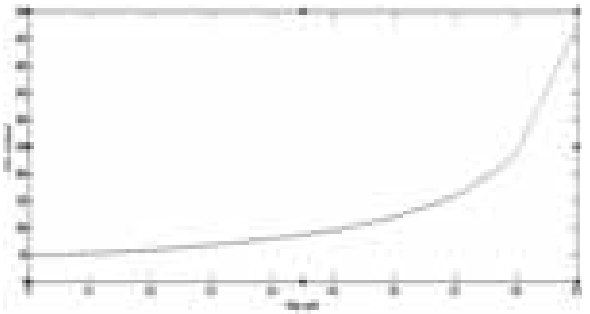


Fig. 3. DC output V_s duty cycle.

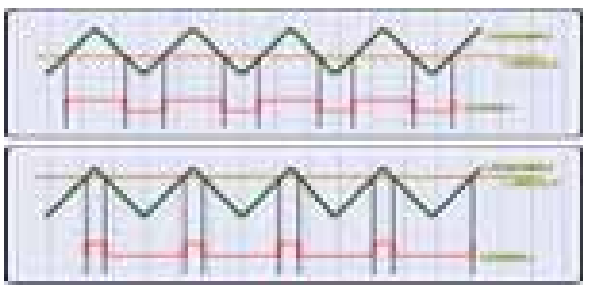


Fig. 4. PWM output with different duty cycle.

$$\text{Inductance } (L) = (D * V_{\text{in}} * (1 - D))/(2 * F_s * I_{\text{out}}) \quad (2)$$

$$\text{Capacitance } (C) = (I_{\text{out}})/(F_s * \Delta V_o) \quad (3)$$

The Table I shows the DC output voltage for the corresponding duty cycle value. It can be observed in the graph duty cycle V_s output voltage is shown in Figure 3.

3. PWM TECHNIC

Pulse width modulation (PWM) is a powerful technique for controlling analog circuits with a microprocessor's digital outputs. PWM is employed in a wide variety of applications, ranging from measurement and communications to power control and conversion by controlling analog circuits digitally, system costs and power consumption can be drastically reduced. What's more, many microcontrollers and DSPs already include on-chip PWM controllers, making implementation easy.[1]

PWM is a way of digitally encoding analog signal levels. Through the use of high-resolution counters, the duty cycle of a square wave is modulated to encode a specific analog signal level. The PWM signal is still digital because, at any given instant of time, the full DC supply is either fully on or fully off. The voltage or current source is supplied to the analog load by means of a repeating series of on and off pulses. The on-time is the time during which the DC supply is applied to the load, and the off-time is the period during which that supply is switched off. Given a sufficient bandwidth, any analog value can be encoded with PWM.

4. WORKING PRINCIPLE OF BDMC

Figure 5 shows the boost derived multiport converter in which the single switch of boost converter is replaced by voltage source inverter. The boost operation is realized by switching on both switches of a particular leg (S1–S4 or S3–S2). This is equivalent to shoot through operation as far as VSI operation is concerned. However in the operation of multiport converter is concerned this is equivalent to switching on controllable switch S_a of the conventional boost converter. The ac output is controlled using a modified version of the unipolar sine width modulation. The BDMC during inverter operation has the same circuit states as the conventional VSI. The switching scheme should ensure that the power transfer with source occurs only during V_n is positive.

The BDMC has three distinct switching intervals as described in the following.

(1) Interval I—Shoot-through interval: The equivalent circuit schematic of the BDMC during the shoot-through interval is shown in Figure 6. The shoot-through interval occurs when both the switches (either S1–S4 or S3–S2) of any particular leg are turned on at the same time. The duration of the shoot-through interval decides the boost converter duty cycle. The diode "D" is reverse biased during this period. The inverter output current circulates within the bridge network switches. Thus, BDMC allows

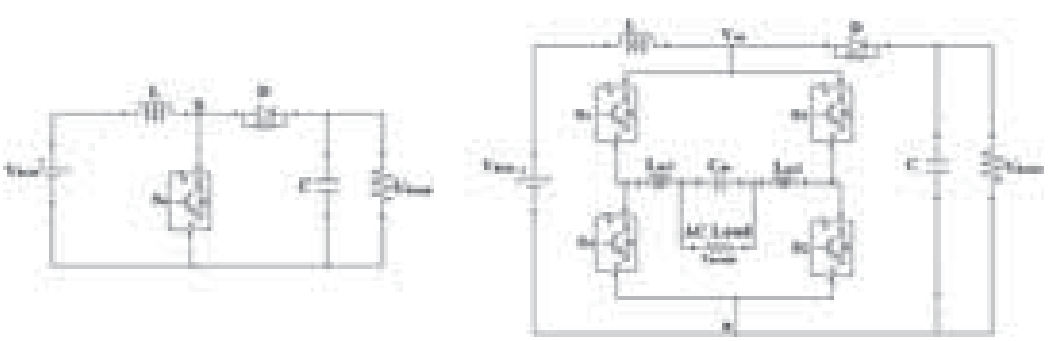


Fig. 5. Boost derived multiport converter.

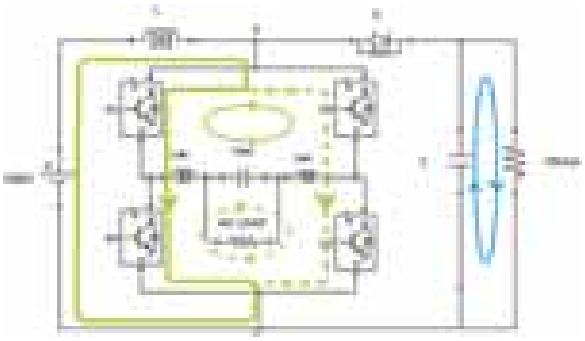


Fig. 6. Current flow direction when S1–S4 ON.

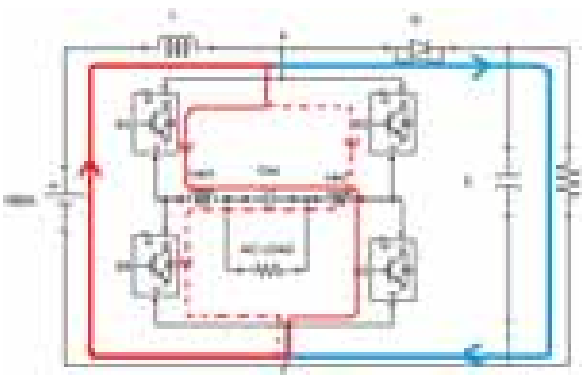


Fig. 7. Current flow direction when S1–S2 ON.

additional switching states which are strictly forbidden in a VSI.

(2) Interval II—Power interval: The power interval, shown in Figure 7, occurs when the inverter current enters or leaves the bridge network at the switch node "s." The diode "D" conducts during this period, and the voltage at the switch node (vn) is equal to the vd (neglecting the diode voltage drop). In this interval, either S1–S2 or S3–S4 is turned on.

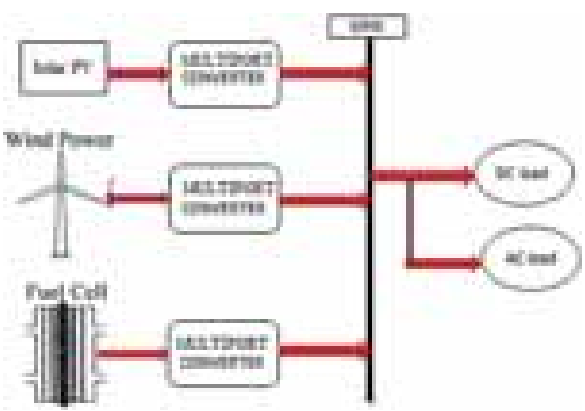


Fig. 8. Grid connected multiport converter.

5. BDMC FOR RENEWABLE ENERGY SYSTEM

Figure 8 shows the grid connected multiport converter. The power from solar, wind and fuel cell is given to the corresponding multiport converter. The boost derived multiport converter gives both DC as well as AC output.

6. CONCLUSION

New Multiport converter topologies with three renewable energy system which can supply simultaneously both DC and AC loads from a single DC supply. This multiport converter topology discussed in this paper are Boost Derived Multiport Converter (BDMC). The proposed multiport converters has the following advantages, shoot-through condition does not cause any problem on working of the circuit hence improves the reliability of the system. Independent control over AC and DC output and the converter can also be adapted to generate AC outputs at frequencies other than line frequencies by a suitable choice of the reference carrier waveform.

References

1. O. Ray, *IEEE Trans. Industry Applications* 50, 1082 **(2014)**.
2. H. Keyhani and H. A. Toliyat, A ZVS single-inductor multi-input multi-output dc–dc converter with the step up/down capability, *Proc. ECCE*, Denver, USA **(2013)**, pp. 5546–5552.
3. Moumita Das and VivekAgarwal, *IEEE Transaction on Industry Applications* 51, 4718 **(2015)**.
4. S. Daniyali, S. H. Hosseini, and G. B. Gharehpetian, *IEEE Trans. Power Electron.* 29, 775 **(2014)**.
5. M. Farhadi Kangarlu and E. Babaei, *IET Power Electron.* 6, 642 **(2013)**.
6. H. Behjati and A. Davoudi, *IEEE Trans. Ind. Electron.* 49, 1464 **(2013)**.
7. J. Hui, A. Bakhshai, and P. K. Jain, A hybrid wind-solar energy system: A new rectifier stage topology, *Proc. APEC*, Palm Springs, USA **(2010)**, pp. 1048–2334.
8. V. Mummadi, Multi input integrated buck-boost converter for photovoltaic applications, *Proc. ICSET*, Singapore **(2008)**, pp. 546–551.
9. B. Mangu, S. Akshatha, D. Suryanarayana, and B. G. Femandes, *IEEE Journal of Emerging and Selected Topics in Power Electronics* 4, 1086 **(2016)**.
10. L. Solero, A. Lidozzi, and J. A. Pomilio, *IEEE Trans. Power Electron.* 20, 1007 **(2005)**.
11. A. Khaligh, J. Cao, and Y. J. Lee, *IEEE Trans. Power Electron.* 24, 862 **(2009)**.
12. H. L. Do, *IEEE Trans. Power Electron.* 25, 1193 **(2010)**.
13. A. Nami, F. Zare, A. Ghosh, and F. Blaabjerg, *IET Power Electron.* 3, 197 **(2009)**.
14. M. Patterson, N. F. Macia, and Arunachala M. Kannan, *IEEE Transactions on Energy Conversion* 30, 359 **(2015)**.
15. F. Nejabatkhah, S. Danyali, S. H. Hosseini, M. Sabahi, and S. M. Niapour, *IEEE Trans. Power Electron.* 27, 2309 **(2012)**.
16. H. Behjati and A. Davoudi, A MIMO topology with series outputs: An interface between diversified energy sources and diode clamped multilevel inverter, *Proc. APEC*, Orlando, USA **(2012)**, pp. 546–551.
17. A. Werth, N. Kitamura, and K. Tanaka, *IEEE Transaction on Smart Grid* 6, 1621 **(2015)**.
18. M. Jafari, G. Hunter, and J. G. Zhu, A new topology of multi input multi output buck-boost dc–dc converter for

RESEARCH ARTICLE

microgrid applications, *Proc. PECon.*, Kota Kinabalu **(2012)**, pp. 286–291.
19. K. Eguchi, Design of a capacitor based MIMO dc–dc converter and its optimal control method, *Proc. ICINIS*, Kunming, Yunnan **(2011)**, pp. 141–144.
20. S. Saggini and P. Mattavelli, Power management in multi source multi load energy harvesting systems, *Proc. EPE*, Barcelona, Spain **(2009)**, pp. 1–10.
21. E. Babaei and Okh. Abbasi, *IET Power Electronics* 9, 9 **(2016)**.
22. A. Kwasinski and P. T. Krein, Multiple-input dc–dc converters to enhance local availability in grids using distributed generation resources, *Applied Power Electronics Conference, APEC 2007—Twenty Second Annual IEEE*, February–March **(2007)**, pp. 1657–1663.
23. A. S. W. Leung, H. S. H. Chung, and T. Chan, A ZCS isolated full-bridge boost converter with multiple inputs, *Power Electronics Specialists Conference 2007*, PESC 2007, IEEE, June **(2007)**, pp. 2542–2548.
24. B. Ozpineci, L. M. Tolbert, and D. Zhong, Multiple input converters for fuel cells, *Industry Applications Conference, 2004, 39th IAS Annual Meeting, Conference Record of the 2004 IEEE*, October **(2004)**, Vol. 2, pp. 791–797.
25. Salem Alshibani, Vassilios G. Agelids, and RukmiDutta, *IEEE Transaction on Sustainable Energy* 5, 10 **(2014)**.

Received: 23 April 2018. Accepted: 16 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1639–1643, 2018

Internet of Things (IoT) Elements, Trends and Applications

S. Dhiviya[1,*], S. Malathy[1], and D. Rajesh Kumar[2]

[1] *Department of CSE, Kongu Engineering College, Perundurai, Tamilnadu, India*
[2] *Department of IT, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India*

Internet of Things (IoT) technology is alive in a wide continuum of systems, sensors and networked products, which provides the benefits of network interconnections with new-fangled abilities. This Internet of Objects change everything including the human ourselves which brings enlighten for disabled people. The IoT is an interrelated computing device, mechanical and digital machines, objects, animals or people with unique identifiers and the ability to transfer data. This provides the methodology to change the agriculture, industry and many useful things in all fields. This paper depicts about the IoT elements, new trends and future applications of IoT.

Keywords: Sensors, Radio Frequency Identification, Network, Communication.

RESEARCH ARTICLE

1. INTRODUCTION

The IoT is considered to be an economic wave in the industrial era. This intelligent methodology uses different protocols to exchange information through the sensing devices. It is a broadening of Internet which further leads to achievement of identification of things, location and tracking of things. This comprehensive technology has the ability to analyze both living and non-living things in the real world.

This methodology enables the development of small-scale devices with unique identification and computing capability with sensors and actuators embedded and are connected through both wireless and wired sensor networks. The technology development is done at low cost and size. The main characteristics of IoT are:

Perception: The sensors, Radio Frequency Identification (RFID), barcode are used to get the data about any object anywhere. This also includes places and object identification and recognition of things.

Transmission: With the available of networking and communicating technologies the data can be accessed at any time. The IoT provides reliable transmission between machine to machine and mobile to machine.

Processing: Cloud computing helps in intelligent processing of IoT data. The service provider process millions or billions of data using cloud computing.

The rest of the paper is organized as follows. Section 2 presents the previous work related to IoT. Section 3 discusses the Position and Circumstances of IoT in detail. Section 4 reports the main applications and its trends. Section 5 provides the conclusion of the paper and possible direction for future work.

2. RELATED WORK

Payam Barnaghi et al.[1] surveyed development of IoT technology in both industry and academic. The interoperability among the "Things" which avoids generic solutions to a problem. The semantics has been added to combine both knowledge engineering and artificial intelligence to present the information. The volume, velocity and volatility of the data provide various challenges to the existing systems.

Andrea et al.[2] presented how IoT is applied for smart cities. This paper describes about the advanced used of technologies for the administration of large city. The paradigm includes monitoring of noise, congestion control, energy consumption, parking, and automation of buildings. The IoT architecture provides practical solution to cities.

Somayya et al.[3] formulated a review on Internet of Things characteristics, basic definition, requirements for usage. The technology is applicable to all fields. This provides enormous use of IoT applications in a simpler way.

Jayavardhana et al.[4] illustrated a survey on cloud implementation network. The communication between the devices and data storage in private and public cloud has been explained.

Zhang et al.[5] proposed a method based on improvement strategy of medical wireless application and internet development. This paper depicts about the unified platform of

*Author to whom correspondence should be addressed.

medical IoT in relation with features of medical data, and provides fundamental framework, design, constitution and coverage model in detail to afford the technological foundation for insisting the medical process management based on current technologies of internet of things.

Jaladi et al.[6] presented the new stage of internet of things like Bluetooth and Zigbee technologies. The web server collects the raw data from the environment and accessed data are displayed to the end user. The collected data are sent to the cloud environment and utilized whenever needed. The different kinds of sensors such as to determine the temperature, pressure, light and pollution.

Vasisht et al.[7] proposed different methodology in increasing the productivity of agriculture. The new techniques are depicted to reduce the loss and input costs in agriculture. This paper presents a gateway and duty cycling approach for increasing the agriculture production in moisture land. This also provides a survey on soil, water content of the land using precision maps.

Lokesh et al.[8] designed a wireless robot for improving the activities of agricultural systems. The mobile robot is used in an agriculture field and analyzed by taking reports from the field. The robot worked well and showed up its efficiency in the smart agriculture system. This assists in improving the moisture of the soil, roots and necessary water content for the crop.

Akyildiz et al.[9] explained all the layers based on sensor networks and nodes. The driving factors for the sensor networks like fault tolerance, addressing schemes has been surveyed. Each network layer is explained with the keen details to route the data and how it is getting transferred. This survey provides methodology to develop a protocol for network.

Chang et al.[10] illuminated that how networks is applicable in the cultivation process. The networks provides valid description of how to yield a continuous growth in the crop. The rainy season is utilized for which crops and how it provides sustainable moisture in the soil is given the sensor networks for a particular type of field. This information improves the farmer's growth and production of cultivation.

Sathish et al.[11] described that wireless networks are liable to cloning attacks which leads to enormous destructive factors. This at beginning stage attacks a node and make changes in it and further it takes control over the entire network. To avoid this factor a new methodology has been developed which provides solution to these kinds of attacks.

Yugang et al.[12] reviewed the prediction method on mining. This is mainly based on heterogeneous network on combining structural and textual data. A framework has been developed to integrate textual and structural data. This showed that the proposed methodology predicts and performs at better level than other methods of prediction.

Rajesh et al.[13] projected to identify these clone nodes. The Residual Energy and Gravitational Search Algorithm based Simulated Annealing (RE-GSASA) method has been proposed to investigate the clone node in network. The performance of RE-GSASA method in terms of optimized energy consumption, improves clone node detection probability rate by reducing the node detection time.

The above discussed techniques have various advantages and issues in bringing up the Internet of Things. These problems can be resolved by using different sensors and high network capability. The elements of IoT has different components and provides reliable data collection. The applications of IoT shows the recent development and usage of internet of things in every place.

3. ELEMENTS AND POSITION

The main components of IoT are are (i) Hardware (ii) Middleware (iii) Presentation. The Figure 1 represents the elements of IoT. The IoT elements are:

A. *Radio Frequency Identification* (*RFID*):

RFID[6] is useful in design of microchips for wireless data communication. This design is used to identify any device automatically using its barcode. Passive RFID uses interrogation signal to communicate with the RFID reader which is mainly useful in retail and supply management. Active RFID readers have inbuilt battery supply and they initiate communication. The active RFID tags are mainly used in cargo monitoring.

B. *Sensor Networks*:

Wireless Sensor network[1] is a software and hardware package. The recent trends of wireless communications are at available efficient, low cost, low power miniature devices which improves the utilization of a sensor network in various aspects. Active RFID which is similar to low end WSN works with the slender processing capability. The components of WSN network include:

(a) Hardware-WSN[13] core hardware consists of sensor interfaces, power supply, transceiver units and processing units. This comprises of modern sensor nodes which have the capacity to commune using one frequency band.

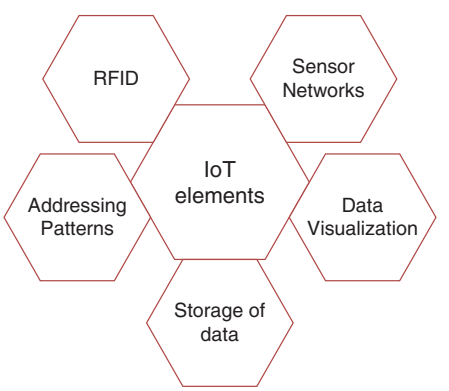


Fig. 1. Elements of IoT.

(b) Communication Stack—WSN nodes interact among themselves to transfer data in single or multi-hop[3] to a base station. Node drop outs, and consequent degraded network lifetimes, are frequent. Communication stack act as a router to WSN subnet and Internet.

(c) Middleware—The importance of middleware for sensor networks is to fortify the maintenance, deployment, development and execution of sensing-based applications. This helps in formulating complex high-level sensing tasks, communicating to WSN,[1] coordination of sensor nodes to split the task and distribute it to the individual sensor nodes, data fusion and reporting the result. The unique property of WSN middleware are application knowledge in nodes and adaptive fidelity method.

(d) Secure Data aggregation—The secure data aggregation method is essential for expanding the lifetime of the network. This methodology assures reliable data collection from sensors. The network topology[9] should have the capability to heal node failure. Ensuring security is critical as the system is automatically linked to actuators and protecting the systems from intruders becomes very important.

C. *Addressing Patterns*:

The essential factors for constructing a unique address are: reliability, uniqueness, persistence and scalability. The element that is to be connected and which is already connected need to be determined by unique identification, location and functionality. The IPV4[2] helps in identifying devices geographically but these devices cannot be determined as individuals. The inclusion of networks and devices should not crate the performance of the network, the working of the devices, the reliability of the data or the efficient use of the devices.

Uniform Resource Name develops copy of the resources which can be fetched through the URL. Using the advantage of metadata for sending data form a database to user.

D. *Storage of Data*:

The data storage is an important issue since the data has to be maintained in a secured and reliable way. The data is mainly used for monitoring and inclination. The artificial intelligence[8] methodology should be created for maintaining these stored data in an efficient and for the purpose of decision making. The schematic system must be constructed in terms of hardware and software basis which is applicable to IoT.

E. *Data Visualization*:

In the era of recent trends visualization of data is quite demanding factor. Data has to be viewed and easily understood by the user so that visualization is easier for everyone. This leads in better decision making progress which converts a raw data into knowledge. Visualization should contain both raw data and transformed data which is especially useful for a final user.

4. APPLICATIONS AND TRENDS

4.1. Applications

There are many applications[4] concerned with Internet of Things which can be differentiated based on their network, scalability, availability and its impact. The Figure 2 shows the picture of real time applications which are in use now. Some of the trending applications are as follows:

A. *Smart Home*:

Smart Home has become the radical extent of achievement in the residential areas and survey says. Smart homes will become as familiar as smartphones. Smart Home products are prosperous enough to save time, energy and money. These smart homes related to IoT provides a new experience to the user.

B. *Digital Health*:

Digital health makes prominent development in the medical treatment. This helps people to get report on their health and gives prescription. Even there are many health apps which provides what food should be taken on a daily basis to keep a balanced diet. Tele health is mainly concerned empowering people to live healthier life by wearing connected devices. The analyzed data will support in personalized analysis of an individual's health.

C. *Power grids*:

Power grids are going to be popular in all industrial and household places which succor in energy saving. The concept behind the smart grids is to gather data in a computerized manner and figure out the behavior of electricity consumers and suppliers for improving the power usage. Smart Grids will also be able to identify sources of power usages quickly and making possible distributed energy system.

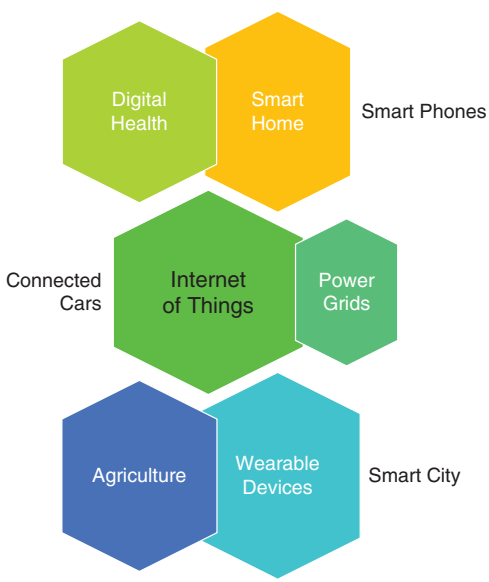


Fig. 2. Internet of things applications.

RESEARCH ARTICLE

D. *Agriculture*:

Need for food has been increased in a higher range because of world's population. Advanced techniques[7] are now utilized in farming to improve the agriculture level. The moisture of soil, nutrients available in the soil, water level are sensed using IoT.

E. *Connected Cars*:

Connected cars is mainly concerned with automotive internal functions of vehicles. The cars will reach the destination of user just by recognizing the voice. The vehicle searches the location using google map. The major companies like Tesla, BMW, Apple, and Google are working on bringing the solution for connected cars in automobiles.

F. *Wearable Devices*:

Smart watches are the most common form of wearable technology. Apple Watch, smart watches are some of the most popular wearables. Fitness tracker is used to track of your activity throughout the day. Google Glass is a headset that displays information in a smartphone like hands free format and lets the wearer give voice commands to the device.

G. *IoT Connectiviy and Business Model*:

The wireless connections and protocols has been developed for business models with lower costs of ownership. Different technologies such as Sigfox, LoRaWAN and 3GPP's narrowband (NB) are analyzed to connect to the sensors. New Business owners can use these different methodologies to increase the benefit of their company. Speed, accessibility, remote work are the things in which IoT plays an essential role in developing the business.

H. *Smart Parking*:

Cloud based integrated parking system has been designed now in smart cities. Since smart cities are in recent trends now to improve the reliability the smart parking lots are used. Parking lot application helps the user to identify the parking area and location to place their vehicle. This information is gained from the map to locate the vehicle in the free space.

I. *Swimming Pool Automation*:

The main objective of swimming pool automation process is to maintain the water hassle-free of germs and bacteria for the bathers. This automation process helps in maintaining the sanitizer level, chlorine level, pH control, TDS control and in filtration process. The Integrated Controllers monitor the pressure of the filter with sensors and cleans it automatically whenever necessary.

J. *Intelligent Shopping*:

Intelligent Shopping using RFID tag and reader in shopping cart. The secure system is used for smart shopping and each product is tagged using RFID. The billing process can be done using shopping cart. Smart placing of things can also be done using the smart shopping. This provides feasible system for shopping and inventory management system can also be done easier. Smart Shopping has intelligent and secure communication protocol to provide security.

K. *Offspring Care*:

Offspring care controls the growing conditions of animal in farms to take care of its health. The diagnosis of harmful gases and air quality in the environment is done using sensors. The animal tracking is also done in open pastures. Daily food and water supply is also maintained for animals.

L. *Ultraviolet Radiation*:

Ultraviolet is an electromagnetic radiation lesser than that of visible light but higher than X-rays. The long-wavelength radiation causes chemical reactions, and many substances to fluoresce. Adequately, biological effects of UV are more than its heating effects. An UV sensor is fixed with the controller to provide information about UV rays. Using this information can decide to travel in that particular area. The temperature sensor is connected to the controller through ADC to track the temperature. This information is published on to the web server using IoT module.

M. *Forest Fire Detection*:

Forest fire is also known as hill fire is an uncontrolled fire which occurs in wild areas. The sensor system detects these kinds of fires as early as possible so as to prevent the harmfulness. The sensed data is collected using Arduino board and transferred to the base station wirelessly. The information is continuously sent via Bluetooth. The alarm is attached along with the sensor to alert the people in forest area.

4.2. IoT Trends

The Internet of things (IoT) is developing at a faster rate which will fascinate the IoT industry. IoT trends continues to progress at an incredibly faster pace. People are at keen interest to know the innovation of IoT. The groundbreaking impact of the Internet of Things on our lives like ATMs that report criminal factors around the area, forks that insist you to eat slowly and IP address for each organ of your body for doctors to diagnose. Connected devices are becoming an important part of our daily lives. But still the lack of standardization remains. The difficulties of IoT standardization are basicaly classified into three categories: platform, applications and connectivity. These three are inter-related process and missing one will halt the standardization process. As IoT is getting elaborated there will be an increase in number of devices connected to the network in different areas for business models. The pros of using smart devices boost customer engagement, increases visibility, and ease of communication that comprise new human-machine interfaces such as voice user interface or Chatbot. Security is the next factor which will confound the challenges of IoT.

Blockchain is a new factor for IoT Security. The astounding conquest of Cryptocurrency, which is built on Blockchain technology, helps in numerous transactions, which further reduces cost. Blockchain enhances security,

privacy of all people engaged in transcations and transparent pattern of transactions. Blockchain is protected by a private key eventhough it is public. IoT investments will continue at a higher rate since people get adapted to the usage of IoT devices. Both software and hardware investments will rapidly increase at a greater pace. This is essential for the applications and sensors to be used in IoT.

Decision-making and implementation of action happens via IoT devices and only transfers appropriate information to the cloud which is coined as "Fog computing." The fog nodes, can be placed anywhere with a network connection: on oil rig, on a factory floor, on top of a power pole, on a railway track, in a vehicle. Any thing which computes, stores and connects can be a fog node. The transformational business models will grow up in many IoT verticals reinforced by Big Data and Aritifical Intelligence tools. The value is mainly based on the convenience of the service for end customers and the usage of information that is collected, analyzed, and fed back into suppliers' business models. The main capacity for IoT business model transformation expands beyond this, to enclose a huge variety of more complex, as-a-service business models that distracts available companies, particularly for areas such as heavy industry, transport and logistics.

5. CONCLUSION AND FUTURE WORK

The communcation capability gives the clearer vision of Internet of Things, how sensing is done and how data is stored and maintained. The growth of the next generation mobile system is based on the vision of the users in developing new applications. The elements and position signifies the sensing network, data security, aggregation of data, representation of data. The applications of IoT depicts the recent trends and devices which are currently available for the end user. Finally IoT saves the energy and time of man and which provides the smart environment.

The work can be extended as collaborative Platform-as-a-Service (PaaS) model works like a data highway, where customers and suppliers maintain and use tool-related data, based on instruction and access rights. The mobile application and the cloud-based platform work together to collect input from the app, sensor data to track stock movements from warehouses, and customer data. This will enable machine learning and predictive forecasting to further reduce downtime and enable a faster productive workforce. The automotive workforce can be created and new machine parts can be ordered without human interaction.

References

1. P. Barnaghi, W. Wang, C. Henson, and K. Taylor, *ACM* 8 **(2012)**.
2. A. Zanella and L. Vangelista, *IEEE Internet of Things Journal* 1, 22 **(2014)**.
3. Somayya Madakam, R. Ramaswamy, and Siddharth Tripathi, Internet of things (IoT): A literature review 164 **(2015)**.
4. J. Jin, J. Gubbi, and S. Marusic, *IEEE Internet of Things Journal* 1, 112 **(2014)**.
5. H. Zhang and W. Ni, *Bioinformatics and Biomedicine (BIBM)* **(2017)**.
6. A. R. Jaladi, K. Khithani, P. Pawar, K. Malvi, and G. Sahoo, *Knowl. Inform. Syst.* 4, 1371 **(2017)**.
7. Deepak Vasisht, Zerina Kapetanovic, and J. Won, FarmBeats: An IoT platform for data-driven agriculture, *14th USENIX Symposium on Networked Systems Design and Implementation* **(2017)**, pp. 515–529.
8. Lokesh Krishna, O. Silver, and Wasswa Fahad Malende, Internet of Things Application for Implementation of Smart Agriculture System, IEEE **(2017)**.
9. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, *IEEE Communications Magazine* 40, 104 **(2002)**.
10. H. Chang, N. Zhou, X. Zhao, Q. Cao, M. Tan, and Y. Zhang, A new agriculture monitoring system based on WSNs, *2014 12th International Conference on Signal Processing (ICSP)* **(2014)**, pp. 1755–1760.
11. R. Sathish and D. R. Kumar, Dynamic detection of clone attack in wireless sensor networks, *International Conference on Communication Systems and Network Technologies* **(2013)**, pp. 501–505.
12. J. I. Yugang, L. I. Yitong, and S. Chuan, *Journal of Computer Applications* 3201 **(2017)**.
13. D. Rajesh Kumar, A. Shanmugam, C. Palanisamy, and A. M. Natarajan, *Asian Journal of Research in Social Sciences and Humanities* 6, 201 **(2016)**.
14. L. Atzori, A. Iera, and G. Morabito, *Comput. Netw.* 54, 2787 **(2010)**.
15. A. Sangaiah, A. Thangavelu, and V. Meenakshi Sundaram (eds.), Cognitive computing for big data systems over IoT, Lecture Notes on Data Engineering and Communications Technologies, Springer, Cham, Vol. 14.
16. C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, *IEEE Commun. Surveys and Tutorials* 16, 414 **(2014)**.
17. D. Rajesh Kumar and Shanmugam, A hyper heuristic localization based cloned node detection technique using GSA based simulated annealing in sensor networks **(2017)**.
18. S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, *Comput. Netw.* 76, 146 **(2015)**.
19. L. Liu, *IEEE Transactions on Signal Processing* 62, 1850 **(2013)**.
20. M. Zhao, X. Wang, and S. Feng, *IEEE Wireless Communications Letters* 19, 1540 **(2015)**.
21. http://www.wired.co.uk/article/internet-of-things-what-is-explained-iot.
22. http://www.engpaper.com/iot-2017.htm.
23. http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6488907.

Received: 23 April 2018. Accepted: 16 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1644–1648, 2018

RESEARCH ARTICLE

Sentiment Analysis of Twitter Data Using Machine Learning Algorithm

K. Ganagavalli*, A. Mangayarkarasi, T. Nandhinisri, and E. Nandhini

Department of CSE, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamilnadu, India

Online social media is a powerful platform for dissemination of information during important real-world events. Over last few years social media has also been used to disseminate misinformation in the form of rumours, hoaxes, fake images, and videos. The spread of such untrustworthy content online has caused the loss of money, infrastructure and threat to human lives in the offline world. This paper is mainly focussed on Twitter, which is one of the most popular microblogging web services today. As a social media website, Twitter has become the third most popular behind only Facebook and YouTube. Its user base statistics ensure a wide audience for business to engage with. This paper focuses on a method to obtain sentiments of users from twitter and using it to analyse whether they can become threat to particular person or society. Machine learning is one of the technologies which is used for decision making and also make predictions by learning more from dataset. Since huge amount of data is streaming very fast, it needs to be classified in short duration.

Keywords: Social Media, Analytics, Machine Learning, Data Cleaning, Pre-Processing.

1. INTRODUCTION

Big data analytics is the process of analyzing large volume and variety of data sets which can be further used to identify hidden patterns, unknown similarities, upcoming trends, customer priorities and other useful information that can help organizations for improvising their business decisions. Big data analytics allows data scientists and business users to make better and faster decisions using the previously available historical data. Using advanced analytics techniques such as text analytics, machine learning, predictive analytics, data mining, statistics, and natural language processing, businesses can analyze previously untapped data sources independent or together with their existing enterprise data to gain new insights resulting in significantly better and faster decisions. Data coming through.

Twitter has been increasing in popularity from last few years. This data contains sentiments of people who express themselves by giving their views, sharing pictures, also tweeting about matters which are going on. Twitter data not only contains sentiments of this people but also of some criminals who communicate with each other. There are many challenges to using Twitter as an information source for crime prediction. Tweets may contained more number of misspelled words and symbol usage for expressing different types of emotions diagrams and symbols. To make matters worse, Twitter imposes a 140-character limit on the length of each tweet, encouraging the use of these and other message shortening devices. Finding this information is very crucial for respective authorities to stop any criminal activity from happening.

The main challenge in front of IT world is to store and analyze huge quantities of data. Every single day data is generated in huge amount from various fields like Geography, Engineering, and Economics and Science etc. To analyze such huge amounts of data for better understanding of users there is a need to develop data intensive applications which are highly available, highly scalable and based on reliable storage system. Since twitter data is real time data, the salient features of same are given.

2. DOMAIN

Twitter users post their opinions or thoughts about different topics and domains using simple and short messages. The topics for discussion need tot to be specific and it can be any topic. This is different from a large percentage of past research, which focused on specific domains such as movie reviews. To analyze this fast streaming data Map Reduce framework can be used since it gives support for parallel map reduce jobs. To obtain valuable information from such streaming data is by extracting words from sentiments and views given. These extracted words from data can be used to find the person who might become threat to

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1644/005 doi:10.1166/jctn.2018.7355

society or any individual and risk modelling can be done. By this information respective authorities can keep track of all those users who are person of interest. This may reduce the crime from happening or at least minimize or mitigate the intensity of damage. Machine learning technology such as Naive Bayes Classifier is used for classification of sentiments based on extracted words. Naive Bayes Classifier works very easily and faster in predicting dataset and requires less training dataset. Big data has made recommendation system more important for the users as it predicts correct piece of information out of huge amount of information. However it also leads to problem of scalability of algorithm and systems. For analyzing the tweets, user have to take polarity into consideration using different types of dictionaries.

2.1. Lexical Dictionary

It mainly consists of most of the English words which will help to analyze and segregate the tweets by matching the word in the tweet with the words in the lexical dictionary. It also consists of phrases, idioms, multi words and headwords.

2.2. Acronym Dictionary

It is used to expand all the acronyms and abbreviations which will generate words which can be analyzed using lexical dictionary.

2.3. Emoticon Dictionary

A tweet may contain different types of emotions based on the users' thoughts and it can be analyzed using a sentimental dictionary. Emoticons are essentially the textual portrayal of the tweeter's mode and convey some meaning.

2.4. Stop Words Dictionary

These are the words in the tweet which are of no consequence as they do not have any polarity and need not be analyzed. So they are eliminated and tagged as stop words. User keep a dictionary with the complete list of all stop words for example able, are, both, etc.

3. ALGORITHM

3.1. Word Matching Algorithm

The word matching algorithm is used to analyze the entire tweet which is stored in database. The algorithm calculates the sentiment based on the words which are present and number of times it has occurred. The words which make a user person of interest are already predefined based on which overall sentiment value is assigned to the tweet of user. These sentiments help to recognize whether the sentence has a positive impact or negative impact based on the final calculation done for the sentence. Tweets which are coming from database are sent through word matching algorithm which looks up for the words and matches with the predefined ones based on which word matching algorithm calculates the sentiment value for every tweet and generates final result.

3.2. Naive Bayesian Classifier

Naive Bayes has been one of the easiest ways of using machine learning technologies for classifying tweets. The proposed methodology for applying Naïve Bayes algorithm splits into 2 phases, Map and Reduce Phase.

3.3. Map Phase

The working of Map phase consists of two major tasks. First, creating a hash map for retrieval of polarity of each word. Secondly, processing the overall polarity of the tweets by applying Naive Bayes algorithm. The map( ) method in Map Reduce phase reads the content of the SentiwordNet dictionary from a file and transform into the Hash map for key-value based polarity retrieval of words. From here, the polarity of each word is stored in the hash map for faster processing. Now, the map( ) method read tweets line by line from the file. Map method parses each and every word and generates tokens. Each token has polarity available in the hash map. The polarity is fetched for each word and calculates the overall polarity of a single tweets using probabilistic model.

3.4. Reduce Phase

The reduce( ) method collects the overall polarity of each tweets and transform into 5 different categories as extreme positive, positive, extreme negative, negative and neural. The reduce( ) method iteratively work to collect various sentiments and based on polarities it classify and write the output on HDFS.

4. SENTIMENTAL ANALYSIS— SYSTEM DESIGN

Sentiment analysis is defined as a process which automates analyzing of user attitudes, customer opinions, and their emotions from feedback data, customer reviews in social media, speech, tweets and database sources through Natural Language Processing. In Sentiment analysis involves opinions in tweets will be classified into different categories like "positive" or "negative" or "neutral," "angry," "anticipation." Textual information present on web is majorly classified into either of the two categories: fact data and sentiment data. Fact data are the objective terminologies concerning different entities, issues or events. Whereas sentiment data are the subjective terms, that define individual's opinions or beliefs for a particular entity, product or event.

4.1. Data Sources

Selection of data source to conduct the sentiment analysis plays a significant role. Social media platforms as

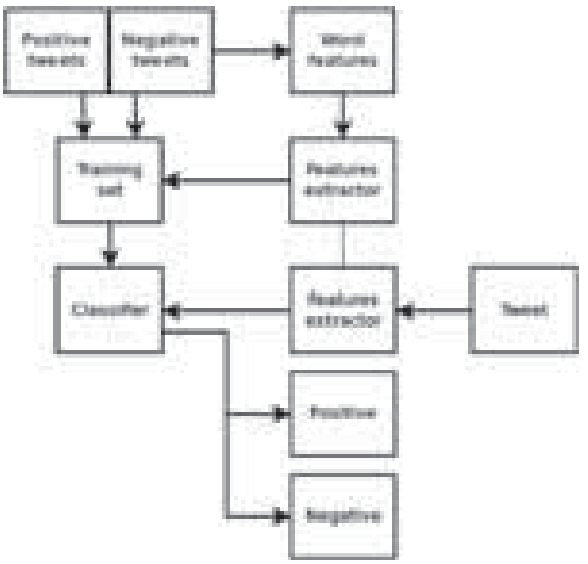


Fig. 1. Sentimental analysis architecture.

the data sources are broadly categorized into three general categories: blogs, micro-blogging sites, and review site. Among all categories, a micro-blogging site such as Twitter has gained higher popularity due to its limited strength of the content and publically availability of data.

4.2. Data Pre-Processing

Mining of Twitter data is a challenging task. The collected data is raw data. In order to apply classifier, it is essential to pre-process or clean the raw data. In the pre-processing phase the data will be cleaned. From the extracted data unwanted data parts like comma, special characters, white spaces emotions, different types of notations URLs, different slang words used in specific parts will be removed and the data will be converted into uniform case. So that it could be processed efficiently. The following steps show the pre-processing procedure.

- Remove the Twitter notations such as hash tags (#), retweets (RT), and account Id (@).
- Remove the URLs, hyperlinks and emoticon. It is necessary to remove non letter data and symbols as are dealing with only text data.
- Remove the stop words such as are, is, am etc. The stop words do not emphasize on any emotions, it is intended to remove them to compress the dataset.
- Compress the elongated words such as happy into happy.
- Decompress the slag words such as g8, f9. Generally slang words are adjectives or nouns and they contain the extreme level of sentiments. So it is necessary to decompress them.

Fig. 2. Sentimental analysis.

4.3. Feature Extraction

The pre-processed dataset has various discrete properties. In feature extraction methods, extract different aspects such as adjectives, verbs and nouns and later these aspects are identified as positive or negative to detect the polarity of the whole sentence. Followings are the widely used Feature Extraction methods.

- Terms Frequency and Term Presence: These features denote individual and distinct words and their occurrence counts.
- Negative Phrases: The presence of negative words can change the meaning or orientation of the opinion. So it is evident to take negative word orientation in account.
- Parts of Speech (POS): Finding nouns, verbs, adjectives etc. as they are significant gauges of opinions.

5. IMPLEMENTATION—TWITTER DATA ANALYSIS

Twitter shares its data in document store format (JSON–JavaScript Object Notation) and allows developers to access it using APIs. Twitter APIs can be accessed only via authenticated requests. Twitter uses Open Authentication (OAuth) and each request must be signed with valid twitter credentials. OAuth provide a safer alternative because the user generally logs into twitter and approves sharing of his information to each application. The user can anytime de-authenticate certain applications from accessing his tweets. API users can use the REST API (for pull access—where must specify the user credential whose data is sought) or Streaming API (for push access—once a request is made they provide continuous stream of updates with no further input required from user). Twitter also places something called a Rate Limit (restriction on amount of data calls that a particular application can make) based on which data gets shared. Once have received data from twitter and placed it in our NOSQL database can apply Map-Reduce/Graph theory approaches to analyses the same.

5.1. Twitter 4J

The tweet retrieval process needs access tokens from the twitter developer site and a piece of code which perform the operation of retrieving those tweets. As the base language used will be java, choose to implement the Java library called Twitter 4J. This library is developed for the twitter API. With Twitter 4J library, A java application can be easily integrated with the twitter service for any analysis purposes. 100% Pure Java—works on any Java Platform v5 or later, Android platform and Google App Engine ready, zero dependency: Additional jars are not required, already built gzip support. There are some system requirements that need to be followed for the Twitter 4J java library to successfully operate. The library supports Windows and Unix Operating systems with Java 1.5 or higher versions installed on it. A java document is also provided in case a user needs to find the method name, syntax, or the root package while implementing the code. To use the java library, the user just needs to add the .jar file to the java application class path.

5.2. Twitter Data Processing

In this paper, the overall process of data collection from twitter social media, pre-processing/cleaning of collected data, performing geographic analysis and sentiment analysis over the collected data including the implementation details of the proposed solution is provided. The tweet analysis used to detect crime consisted of the following steps:

1. Collect tweets
2. Clean and parse the data
3. Conduct geographic analysis on the extracted tweets
4. Conduct sentiment analysis on the extracted tweets.

5.3. Functional Requirements

1. Connected to twitter and fetched tweets based on keywords related to terrorism within the Kenyan context.
2. No duplication of the fetched tweets.
3. Retrieved the metadata of each tweet along with text content and coordinates (the longitude and the latitude). The tweets were stored in the database with the following headers;
   - Tweet Id.
   - Date and time of tweet creation.
   - Actual text of the tweet.
   - Sender's place.
   - Geo coordinates (latitude and longitude).
   - Sent http request to twitter for every 30 seconds interval to fetch the tweets.
4. Created a database table to store tweet text along with extracted information of geo-coordinates.
5. Performed cleaning of tweets and stored in a database table.
6. Performed sentiment analysis on each tweet text and calculated polarity score.

6. EXPERIMENTAL RESULTS

The experiment is done a set of tweets extracted from twitter API. IT is pre-processed and cleaned. Then it is analyzed using sentimental dictionary to categorize words into different types of emotions. Based on the count of words shared in different emotions a plot will be plotted for providing a graphical representation on the analyzed report.

The word count shared on different emotions for tweet is shared below.

A graph is plotted on the given emotion analysis report.

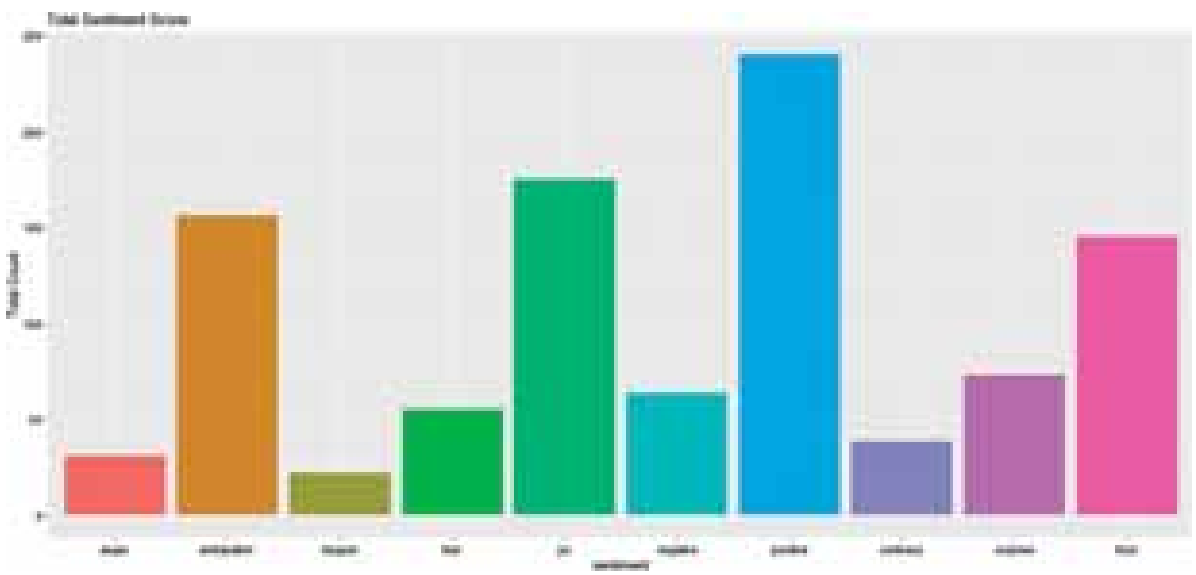


Fig. 3. Sentimental analysis report.

7. CONCLUSION

Twitter Data in the form of opinion, feedback, reviews, remarks and complaint are treated as big data and it cannot be used directly. These data first convert as per requirement. It is proposed to analyze the real time live tweets from twitter using Twitter API, and the large volume of data makes the application suitable for Big Data Analytics. This paper can be extended to predict location of a tweet based on the tweet's information and the user's information in the future. It has presented a simple and complete system for sentiment analysis on large dataset generated from twitter using with word matching algorithm and Naive Bayes. It has implemented word matching algorithm on top of to improve the efficiency of system generating results.

References

1. B. Liu, E. Blasch, Y. Chen, D. Shen, and G. Chen, Scalable sentimentclassification for big data analysis using naïve bayes classifier, *Scalable IEEE Intl. Conf. on Big Data*, October **(2013)**.
2. Hadoop The Definitive Guide, 4th edn., By Tom White, Oreilly **(2015)**.
3. Aditya Bhardwaj, Vanraj, and Ankit Kumar, Yogendra Narayan, and PawanKumar, Big data emerging technologies: A case study with analyzing twitter data using apache hive, *Proceedings of 2015RAECS UIET Panjab University Chandigarh*, December **(2015)**.
4. Data Mining Book Concepts and Techniques, Jiawei Han and Michelinekamber **(2011)**.
5. Ramesh, R. Divya, G. Divya, D. Merin, and K. Kurian, *IJIRST—International Journal for Innovative Research in Science and Technology* I **(2015)**.
6. CloudAcademy-http://cloudacademy.com/blog/naive-bayes-classifier/.
7. Ajinkya Ingle, Anjali Kante, Shriya Samak, and Anita Kumari, *International Journal of Engineering Research and General Science* 3, 144 **(2015)**.
8. Garvit Bansal, Anshul Gupta, Utkarsh Pyne, Manish Singhal, and Subhasis Banerjee, A Framework for Performance Analysis and Tuning in Hadoop Based Clusters **(2015)**.
9. Swati Sharma and Manoj Sethi, *International Research Journal of Engineering and Technology (IRJET)* **(2015)**, e-ISSN: 2395-0056, pISSN: 2395-0072.

Received: 24 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1649–1655, 2018

Big Health Data Processing with Document-Based NoSQL Database

N. R. Gayathiri[1], D. David Jaspher[2,*], and A. M. Natarajan[3]

[1]Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam 638401, India
[2]PG Research Scholar, Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam 638401, India
[3]Department of Computer Science, KPR Institute of Engg and Technology, Coimbatore 641407, India

RESEARCH ARTICLE

Health care is an essential thing in maintenance or improvement of health. The volume of Healthcare data includes variable content sorts, sounds, and pictures and is expanding every day. In this way, the capacity and preparing of these information is an important and testing issue. By and large, Relational databases are utilized for putting away Healthcare information which are not ready to deal with the monstrous and different nature of them. Techniques: This examination introduces the model in light of NoSQL databases for the capacity of Healthcare data. In spite of various kinds of NoSQL databases, Document Oriented DBs were chosen by a study on the idea of Healthcare Data. The information were appropriated on the database by applying the distributed property. With the results the comparison of the efficiency in fetching data with respect to singleton and distributed database using NoSQL were analyzed. Based on the observation the model has given the results with respect to the time required to fetch records and the no of records that were taken under singleton and distributed databases and this has also been implemented with java and their efficiencies of retrieval was checked along with Hashing and B-tree.

Keywords: Document Oriented DBs, NoSQL DB's, Hashing and B-Tree.

1. INTRODUCTION TO BIG DATA

The term Big data is used to describe the data in large volume either structured or unstructured and is in an increasing state with respect to a business or historical records on a day-to-day basis. Big data can be analyzed for insights that lead to better decisions and strategic business moves, day to day social activities, improvements in future prediction and analysis based on historical records and so on.[4] This helps in Organizations to achieve consistency and higher accuracy on future works and prediction of data. The Big data characterized by the following 3v's.

1.1. Volume

Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data. In the past, storing it would've been a problem—but new technologies (such as Hadoop) have eased the burden. It is estimated that data volume is increasing 40% per year, and will grow 44 times between 2009 and 2020,[1] Big companies are no strangers to Big Data. In early 1980's UPS began to capture and track data on package movements that now number 16.3 million packages per day while responding to 39.5 million tracking requests per day. But as of now it is estimated an amount over 16 petabytes of data.[2] Wal-Mart records more than 1 million customer transactions per hour, generating more than 2.5 petabytes of data.[6] And in a survey report 17% of companies currently managing more than a petabyte of data with an additional 22% reporting hundreds of terabytes.[7]

1.2. Velocity

Data streams are growing in an unprecedented speed and are dealt with in a timely manner. The ability to represent data accessibility and storing them are represented as velocity of data. The data velocity is represented as follows.

A data whose value is independent of time or its value remains unchanged for any test attribute can be said as data in rest. Whereas data-in-Motion is described as the data being captured at a rate or with a lag time that makes it useful. This incorporates the characteristics of timeliness or latency Consistency and completeness of fast moving streams of data as one concern, thereby matching them to specific outcome events, a challenge raised under variety is another. This can be thought of as a fire hose of incoming data that needs to be captured, stored, and analyzed.

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1649/007 doi:10.1166/jctn.2018.7356

It describes that how long the data will be valuable, is the data is valuable only over a certain period or is it permanently important. Understanding this dimension of velocity in the data can be stored permanently or discarded over a time frequency. This is to determine the speed of the data that to be stored and retrieved. This is one of the major determinants of NoSQL storage, retrieval, analysis, and deployment architecture that companies must work through today. The analytical tools like Hadoop and NoSQL are used prior to the site with real time analytics of big data.

1.3. Variety

Data are available in different formats such as structured, numeric data in traditional databases and unstructured data like text documents, email, video, audio, stock ticker data and financial transactions. Variety describes different formats which could not be stored in a structured database, it's estimated by some studies to account for 90% or more of the data in organizations. Unstructured data includes documents, emails, social media text messages, video, still images, audio, graphs, devices, RFID tags, machine logs, cell phone GPS signals, and the output from all types of machine-generated data from sensors, DNA analysis devices, and more. This type of data is characterized as unstructured or semi-structured and has existed all along.

Variety is also used to refer data from many different sources, what's changed is that through analysis it can yield new realization and valuable insights that are not previously available. Since data are of different formats and from different sources these has challenges in the processing which includes storing and retrieving these data types quickly and effectively. Unstructured data is growing much more rapidly than structured data. Gartner estimates that unstructured data doubles every three months and offers the example that there are seven million web pages added each day.[8]

2. CHALLENGES IN BIG DATA

2.1. Storing of Big Data

Since the volume of the data is very high and growing, Storing of such large volume of data matters, more number of data that an organization has, more the complexity arises in storing the data.

2.2. Security of Big Data

Prevention of data from data theft and other criminal activities are important, This includes the activities like transaction logs and data Validation and filtration of endpoint inputs, Securing distributed framework calculations and other processes, Securing and protecting data in real time, Protecting access control method communication and encryption, Protecting access control method communication and encryption, Granular auditing, Granular access control, Privacy protection for non-rational data stores[10] etc.

2.3. False Positives

According to Jack Levis, It is difficult to draw the useful insights without solid analytical models from Big Data in place, sometimes 'thinking fast' can lead to false positives.

2.4. Incorrect Findings

Big Data can guide you to a more accurate prediction of the future, but it should not be taken as face value, there needs to be a human element involve in a process, analyze and find conclusions. Big data can be used to make future prediction but it should not be consider as the face value there need to be human involvements in processing analyzing and finding conclusion in order to avoid false results.

2.5. Big Data Applications

Organizational Goals are improved by analyzing the previous contacts, customer reports, purchase history, and other sources to structure special offers and purchase recommendations well suited to their tastes. Henceforth big data analysis plays an important role in business analysis. It also plays a major role in fields like Education, Financial Accounting, Health care, Weather forecasting, Economic welfare Accountings etc.

2.6. Big Data Handling

Cloud computing provides flexible infrastructure and high storage capacity for Big Data applications.[2] The MapReduce framework is most preferable for processing huge volume of unstructured data set in Big Data. Sedayaoet et al.,[4] suggested to use Hadoop to analyze the anonymized data and obtain useful results for the Human Factors analysts. At the same time, the requirements of anonymization were learned and anonymized data sets need to be carefully analyzed to determine whether they are vulnerable to attack. Anonymization tools were found intended for the enterprise generally did not seem to consider the quality of anonymization and does not clearly state whether an anonymized data set was vulnerable to correlation attacks.

3. ROLE OF BIG DATA IN HEALTH CARE

HealthCare Organizations have data in large volumes, this data may be either structured or unstructured and are clinically relevant. This are collected from various relevant places like labs government surveys, Local healthcare centers, imaging systems, Physical Notes so on. But where this data are collected and how they are stored and processed. The goal of the Big data healthcare is to gather the overall data from the patient Information and analyze it to achieve the critical objectives as given below.

Building Sustainable Healthcare Systems are focused on understanding the patients and improving the patient care thereby also promotes to promote the effective utilization of the resources, thereby reduce the cost for the care.

Collaborate to improve the care and outcomes a tailored and effective healthcare program is an important factor. This would be achieved by means of determining the ways to understand the patient individually and improve patient engagement, personalize Healthcare Initiatives, Research and development with the quality of service rendered and so on.

Increase Access to HealthCare Makes the resource easy to access, in order for the population to thrive, Educating consumers on preventive care can improve health and reduce the demand and waste of healthcare resources.

4. NoSQL

NoSQL is abbreviated as Not-only-SQL, it is a way to deal with database outline that can accommodate a wide assortment of information models, including key-esteem, report, columnar and chart groups. NoSQL, which remain for "not just SQL," is another option to customary Relational databases in which data is put in tables and data pattern is deliberately outlined before the database is manufactured. NoSQL databases are particularly helpful for working with expansive arrangements of conveyed information. This was motivated to include the simplicity of design and controlled availability of data, this also works well with horizontal scalability to clustering of machines, the databases used are unstructured and are different from relational databases this makes them faster to operate (i.e., store and retrieve). The data structures used in the no-sql are more flexiable than the normal relational databases.

4.1. Key Features of NoSQL

NoSQL systems features "Share nothing" as its key, Making this database to become horizontal scaling-replicating and partitioning data over many servers.[19] With this feature, NoSQL systems can support simple read/write operations for a huge volume of data in a lesser time. NoSQL systems don't provide ACID (Atomicity, Consistency, Isolation, Durability) properties, but follow BASE property. BASE is acronym for Basically Available, Soft state and eventually consistent. Basically available means that data are available of the time.[20] Soft state means data is not consistent all the time but will be in eventually consistent state. NoSQL features the following four different classes.

Key-value store: The data are stored as key-value pairs.[21] This data structure is also known as "hash table" where the data are retrieved by keys.[22] The examples of key-value stores are Redis, Memcached.

Document store: The data are stored in collections, these collections contain key-value pairs which encapsulate key value pairs in JSON (Javascript Object Notation) or JSON like documents.[22] MongoDB, CouchDB are some of the examples of Document store databases. Since values are not opaque to the system, data can be queried by values as well as keys. It is used for applications in which data is changed occasionally like Customer Relationship Management System.[21]

Column family: The data are stored as a set of rows and columns.[21] Here the columns are grouped accordingly relationship of data. Cassandra and HBase are the well-known examples of the column family databases.

Graph database: These are databases that represents the data in the form of graph. Neo4j is a well-known example of the Graph Database.[22]

4.2. Modeling of NoSQL

Over the most recent couple of years, utilization of NoSQL database has been spread in various areas in view of their capacity to manage new necessities of uses. Additionally, NoSQL databases are getting to be the favored choice for putting away noteworthy sums of excess information in the Cloud. NoSQL DBs introduce new capacity structures that give high versatility, accessibility and quick recovery prerequisites for overseeing unstructured and incompletely organized information. Numerous NoSQL DBs are open source and they are less expensive per terabyte than conventional DBs. These DBs are more proper for the electronic information.[13] In light of the information show, there are four general composes for these databases: Key-esteem DB, Columnar DB, Document-Based DB and Graph-Based DB.[12] Key-esteem DBs have the minimum complex structure. In this model, information are put away and recovered by predefined key with an autonomous esteem. The key of a key/esteem match is a special incentive in a set and it is utilized for getting to the information.[13] In Columnar DBs, the section is the significant part. It comprises of related information gathered closely. The information are put away in a section family premise that is normally characterized at the design or startup time. Segments can store any information composes viably.[12]

Report based DBs conceivably sort out immense and complex archives. This model backings different kinds of monstrous records with adaptability to include any quantities of fields in any length. The record is considered overall protest; and it isn't part into name/esteem sets. Report databases permit ordering of archives on both essential identifier and its properties.[11–13] Diagram based databases utilize diagram structures with hubs, edges and properties for putting away information.[14]

Thus, there is a long assortment of NoSQL databases in various information models. Clients must choose one of them in view of their application necessities. In the following section, we are going to contend around a suitable NoSQL database for wellbeing information, contingent upon information attributes.

4.3. Making Distributed Database in NoSQL

Usually Relational Databases scale vertically, where a single server hosts the entire database to ensure acceptance performance for tables using cross joins and transactions.

But for data which grows rapidly the data have to be scaled horizontally, which includes adding of servers instead of handling the data with a single server. Using SQL 'Sharding' a database across many servers can be accomplished by means of SANs and other complex activities to make the hardware act as a single server.

NoSQL databases, then again, for the most part bolster auto-sharding, implying that they locally and consequently spread information over a discretionary number of servers, without requiring the application to try and know about the sythesis of the server pool. Information and question stack are naturally adjusted crosswise over servers, and when a server goes down, it can be rapidly and straightforwardly supplanted with no application disturbance.

Distributed computing makes this essentially less demanding, with suppliers, for example, Amazon Web Services giving for all intents and purposes boundless limit on request, and dealing with all the vital framework organization undertakings. Engineers never again need to develop mind boggling, costly stages to help their applications, and can focus on composing application code.[15] Ware servers can give an indistinguishable preparing and capacity abilities from a solitary top of the line server at a small amount of the cost.

Most NoSQL databases additionally bolster programmed database replication to keep up accessibility in case of blackouts or arranged support occasions. More modern NoSQL databases are completely self-mending, offering robotized failover and recuperation, and in addition the capacity to appropriate the database over numerous geographic areas to withstand provincial disappointments and empower information limitation. Not at all like social databases, NoSQL databases for the most part have no necessity for discrete applications or costly additional items to execute replication.

4.4. Appropriate Model for Health Care

On account of NoSQL models, column based database is usable for a ton of little constant peruses and writes. The data are recovered by determined example, and a couple of Attributes are included in queries. For getting to superior written work, key-esteem models are recognizable. This model likewise worked with a level information model, and question is performed in characterized keys. Document Oriented database works over a wide assortment of access examples and information composed.

4.5. Analysis of Document Oriented Database

A Document Oriented database is intended for storing, recovering, and overseeing record document-oriented, or semi organized information. Document-Oriented databases are one of the primary classes of NoSQL databases. The focal idea of a Document-Oriented databases is the idea of a Document. While each Document-Oriented databases execution varies on the points of interest of this definition, all in all, they all accept records typify and encode information (or data) in some standard format(s) (or encodings) being used incorporate XML, YAML, JSON and BSON, as well as double structures like PDF and Microsoft Office records (MS Word, Excel, etcetera).

4.6. Data Model Evaluation Criteria

Relational Database Management framework won't leave for huge databases application on the grounds that have a downside to the analysis like tedious and execution speed and scaling of questions, they are characterize necessary. In any case, the capacity prerequisite for the new age of uses are tremendous extraordinary from legacy applications. We can pick NoSQL (MongoDB) rather than MySQL due to two factors, usability and timing execution. We reason that if your web application is information serious what's more, stores bunches of huge information, questions loads of information, and by and large lives database, at that point you better do that effectively or have assets (i.e., cash) to consume. It's a value improving the situation all associations for creating applications. Lastly, the report finishes up by technique characterize a database coordination strategy by utilizing a middleware between the two layers. In this strategy, application does not need to consider about the multifaceted nature of hidden database layer with MongoDB Databases. There information appropriation and capacity. They need to utilize the essential SQL question dialect to get result from the database and all the arrangement transformation tenets will be finished by the Metadata since information bring from database as JSON design. The framework was proposed in light of the fact that MongoDB has recently appeared, though the standard SQL dialect has been over years and, thusly on the off chance that we blend the two we can utilize the highlights of both the database. In spite of the fact that, NoSQL (MongoDB) has the upside of level development, yet for complex SQL asks for, it can't bolster them extremely well. For the Query in light of KEY/VALUE and gigantic information stockpiling necessities, NoSQL is an exceptionally worth doing decision for me and every single other engineer also, associations who's created enormous applications.

5. RESULTS AND DISCUSSION

The performance of the model is discussed from the perspectives of query time in several operations, data preparation, flexibility and extensibility are observed with the test dataset taken from the health care survey report.

Table I. Parameters of the system for implementation.

System parameter	Values
Operating system	Windows 10 pro-64 bit
Processor	Intel core i5-2.2 GHz
NoSQL database	Mongo DB 3.4.10 2008R2 plus
HDD	60 GB
RAM	6 GB
Data	Annual health survey across various states

Table II. Performance evaluation for 100000 records.

No. of records taken	Query	No. of records fetched	Execution time in MongoDB without shard	Execution time in MongoDB with shard
100000	Query 1	100000	0.091	1.048
	Query 2	4330	0.030	1.053
	Query 3	629	0.054	1.046
	Query 4	1	0.163	0.122

5.1. Test Data and Environment

The data is collected from the annual health survey report taken by the government in various locations and merged to present a single health care data report. This data are structured and have higher volume, The processing of this data was made using both Singleton and sharded databases in NoSQL databases. The non-relational database was implemented in shard technique by three nodes, and the test was repeated in same conditions as what done with the singleton NoSQL database.

Document-Based model stores data in collections with BSON format (a kind of text format). So, data must be converted to text for importing to Document-Based database. Converting tables to text is performed in different ways including the use of SQL to Text software, establishing of "result to file property" programming methods in NoSQL databases.[3]

5.2. Query Time

The performance evaluation of the NoSQL is compared among the collections having various size of records which includes as 100000, 500000, 1000000 data in the collection.

Queries were generated to fetch the no of records with respect to each individual collection. The Queries are distinguished on the base of number of predicates which is the number of conditions that are needed to fetch the required amount of data. And the execution time describes the time in milli seconds Query 1 represents the query to fetch all the records, Query 2 Represents the Query with a single predicate and to fetch the values that match the given predicate. Similarly Query 3 and Query 4 are also to fetch records based on conditions with two and three Predicates. The execution time for fetching the records in MongoDB with Sharded and Unsharded Records are given as in the table above.

Table III. Performance evaluation for 500000 records.

No. of records taken	Query	No. of records fetched	Execution time in MongoDB without shard	Execution time in MongoDB with shard
500000	Query 1	500000	0.047	1.085
	Query 2	188	0.225	2.053
	Query 3	185	0.067	0.297
	Query 4	1	0.705	0.885

Table IV. Performance evaluation for 1000000 records.

No. of records taken	Query	No. of records fetched	Execution time in MongoDB without shard	Execution time in MongoDB with shard
1000000	Query 1	1000000	0.018	0.057
	Query 2	192892	0.023	1.021
	Query 3	14044	0.020	1.122
	Query 4	1	0.494	1.81

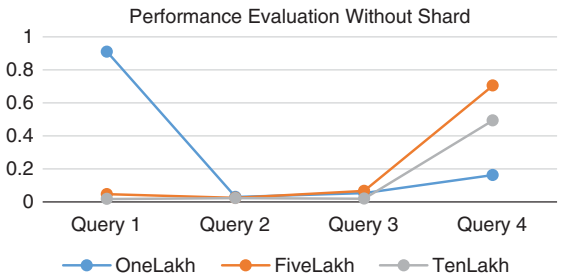


Fig. 1. Performance of MongoDB with singleton database.

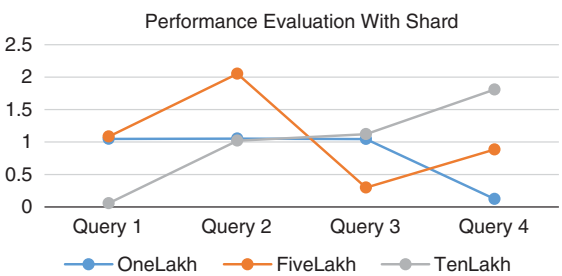


Fig. 2. Performance of MongoDB with distributed database.

6. MongoDB INDEXING

MongoDB's native spatial index also called 2dsphere, it is a flat coordinate system that first partitions the data at multiple resolution levels and then indexes the resulting cells with B-tree.[23] That is, 2dsphere only supports the access of spatial data with the coordinates of latitude and longitude. However, in geographical applications which focus on city/county scale, planar Cartesian coordinate is widely used.[16] Thus, constant and time-consuming data transformation between two coordinate systems is a necessity when applying 2dsphere index to these applications,

RESEARCH ARTICLE

not to mention that sophisticated spherical computation is much more costly than Cartesian computation in spatial operations.[16] In computer science, a *B tree* is an N-array tree with a variable but often have large number of children per node. A B tree consists of a root, internal nodes and leaves,[17] a *B-tree* is a tree data structure that helps in keeping the data sorted and allows performing searches, sequential access, insertions, and deletions in logarithmic time.

6.1. Performance Evaluation with Hashing

Hashing is a strategy to change over a scope of key esteems into a scope of files of a cluster. It might happen that the hashing procedure is utilized to make an effectively utilized file of the cluster. In such a case, we can look through the following unfilled area in the exhibit by investigating the following cell until the point that we locate a vacant cell. This strategy is called linear probing. In our case the NoSQL MongoDB was implemented in java and their results were tested for fetching and retrieval time with Hashing and B-tree and their effects of retrieval in fetch time with respect to the hashing and B-tree were taken. It is noted that the performance of the retrieval in comparatively higher than that of with the B-tree.

6.2. Hash Algorithm

Step 1: Establish the connection between Java and MongoDB.
Step 2: Set up the Key value in java (e.g., $i = 0$, to refer as an index from 0 in start of the record).
Step 3: Fetch the records from the mongo connection using Query.
Step 4: For each record fetched from the query implement the hash table and increase the key value for each record.
Step 5: For each record in the hash table generate key with the hash value.
Step 6: Execute the results.

The time complexity of Hashing Algorithm is O(1), Whereas the time complexity for B-Tree is O(log(n)), that is when the number of records increases there is a gradual increase in the execution time for the B-tree. For the calculation of fetch time with a query in MongoDB with Java implemented with B-tree and Simple Hashing. The execution time are 74.814 sec in Hashing and 109.628 secs in B-tree. The evaluation status are as follows.

The Queries describe the query to fetch the no of records which are 1,00,000 records for Query-1, 7741 records for Query-2, 2064 records for Query-3, 147 records for Query-4.

The figure shows the comparison of performance in java with respect to the use of Hash Function and B-tree operation.

Table V. Performance evaluation of MongoDB java with hashing and B-tree.

Query	Execution time with hash	Execution time with B-tree
Query 1	5 min 43.342 sec	7 min 42.213 sec
Query 2	1 min 14.361 secs	1 min 49.321 sec
Query 3	24.121 sec	23.567 sec
Query 4	0.916 sec	0.623 sec

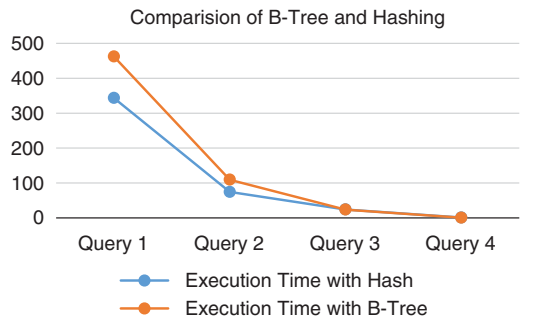


Fig. 3. Comparison of B-tree and hash performance.

The Evaluation status thus shows that the execution time in B-Tree is increasing with respect to the decrease in the no of records. This is due to its time complexity log(n). In which there is a linear increase in time with increase in the no of records.

7. CONCLUSION

In the present paper, considering the present difficulties with regard to Health care information administration, a model in view of the Document Oriented databases was displayed. Keeping in mind the end goal to assess the execution of this model contrasted with past one, these two models were actualized by utilizing genuine Healthcare data. Accepting the inquiry time, by adding Shard highlight to Document based DB, the speed of information recuperation activities exceedingly expanded also, achieved nearly to SQL Server. Not with standing, Sharding influenced the written work speed in this database and backed off the "express" activities more than the past state. Thinking about information arrangement, the new model has the limit of putting away semi-organized and unstructured information with no arrangement require. This model has higher adaptability than that is orient to any structured dbs. The extensibility of document oriented DB is high, while the structured relational model dB's have few restrictions in this respect.

The present investigation gave an answer for putting away and preparing of Health data. The appraisals performed in various settings demonstrate the adequacy of the arrangement. Later on, this model might be created through applying current stockpiling what's more, handling devices, for example, delineate programming models. In this model, a significant issue is confide in

information security and protection by clients. To acquire this issue, different security approaches in Cloud condition must be explored and perfect strategy ought to be presented for these conditions. And the results were been tested with the efficiency of retrieval using Hashing and B-Tree.

References

1. Mckinsey Global Institute, Big Data, [Report, 2011], McKinsey Global Institute, New York.
2. S. Vennila and J. Priyadarshini, Department of Computer Science, Scalable Privacy Preservation in Big Data a Survey, Vellore Institute of Technology.
3. Z. Goli Malekabadi, M. Sargolzaei Avan, and M. K. Akbari, An effective model for store and retrieve big health data in cloud computing.
4. Q. Zhou, B. Xia, W. Xue, C. Zeng, R. Han, and T. Li, An Advanced Inventory Data Mining System for Business Intelligence, Automation Department, Xiamen University, Xiamen, Fujian, China.
5. K. B. Sundhara Kumar, Srividya, and S. Mohanavalli, A Performance Comparison of Document Oriented NoSQL Databases.
6. Open Data Center Alliance, Big Data Consumer Guide.
7. J. McKendrick, Research Analyst, Big Data Opportunities Survey, Produced by Unisphere Research, A Division of Information Today, Inc., May **(2013)**, Sponsored by SAP.
8. T. H. Davenport and J. Dyche, Big Data in Big Companies, SAS Institute Inc., May **(2013)**.
9. Big Data, G. O. Strawn, NITRD.gov, NITRO Big Data.pdf April **(2013)**.
10. http://dataconomy.com/2017/07/10-challenges-big-data-security-privacy/.
11. N. Leavitt, Will NoSQL Databases Live Up to Their Promise, IEEE Computer Society **(2010)**.
12. S. Tiwari, Professional NoSql, JohnWiley & Sons, Inc., Indiana **(2011)**.
13. N. C. Brad, Data De-Duplication in No SQL DB **(2012)**.
14. I. Robinson, J. W. E. Eifrem, and N. J. M. Loukides (eds.), Graph Databases, O Reilly Media, United States of America **(2013)**.
15. W. Vogels, (2012-01-18), Amazon DynamoDB—A fast and scalable NoSQL database service designed for internet scale applications, All Things Distributed, Retrieved 2017-03-06.
16. L. Xiang, X. Shao, and D. Wang, Providing R-Tree Support for MongoDB, State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan, China, (geoxlg, shaoxiaotian, wangdehao)@whu.edu.cn.
17. Ramez Elmasri Navathe and B. Shamkant, Fundamentals of Database Systems, 6th edn., Pearson Education, Upper Saddle River, N.J. **(2010)**.
18. MongoDB Btree Index and Fatel Index, Ref https: www.slideshare .netmongodb20121024-mongodbboston.
19. R. Cattell, *SIGMOD Rec.* 39, 12 **(2011)**.
20. O. Sutinen, NoSQL—Factors Supporting the Adoption of Non-Relational Databases, M.Sc Thesis, Dept. Comput. Sci., Tampere Univ., Finland **(2010)**.
21. Rupali Arora and Rinkle Rani Aggarwal, *International Journal of Scientific and Engineering Research* 1 **(2013)**.
22. Kamaljeet Kaur and Harpreet Kaur, *International Journal of Advanced Research in Computer Science*.
23. https://docs.mongodb.com/manual/indexes/MongoDb Documentation.

Received: 24 April 2018. Accepted: 17 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1656–1661, 2018

RESEARCH ARTICLE

Microarray Analysis Using Advanced Particle Swarm Optimization

R. M. Suruthi* and K. Umamaheswari

Department of Information Technology, PSG College of Technology, Coimbatore 641004, India

Microarray technology concurrently monitors the expression levels of thousands of genes across collection of related samples. Microarray basically consists of large number of gene sequences under multiple conditions. Recent advances in DNA microarray technology allow scientists to measure the expression levels of thousands of genes simultaneously in a biological organism. This made it possible to generate databases consisting of cancerous and normal tissues. Microarray data are extremely asymmetric in dimensionality, highly redundant and noisy. The approach focuses on the feature gene selection and classification, which employs an optimization algorithm to select a subset of genes. The proposed method consist of feature selection approach using filtering technique for reducing the high dimensional microarray data. Optimized feature subset has to be selected using advanced PSO (Particle Swarm Optimization) approach which could reduce computation speed and improve performance accuracy. Classification of the features can be done, using a SVM (Support Vector Machine) classifier to evaluate the performance.

Keywords: Microarray, PSO, SVM, Feature Selection.

1. INTRODUCTION

A Microarray database is a repository containing gene expression data. The key uses of a microarray database are to store the measurement data, manage a searchable index, and make the data available to other applications for analysis and interpretation. Microarrays, known as DNA chips or sometime called gene chips. This makes it possible to measure simultaneously the expression level in a cell or tissue sample for each gene represented on the chip. DNA microarrays can be used to determine which genes are being expressed in a given cell type at a particular time and under particular conditions. This allows us to compare the gene expression in two different cell types or tissue samples, where we can determine the more informative genes that are responsible for causing a specific disease or cancer.

Interest in the analysis of microarray data has considerably increased. A typical micro-array will investigate thousands of genes, recording their expression levels over hundreds of samples. A major use of microarray data is to classify genes with similar expression profiles into groups in order to investigate their biological significance. The relation between microarray analysis and their biological significance is often difficult.

Microarray data sets are commonly very large, and analytical precision is influenced by a number of variables. So it is extremely useful to reduce the dataset to those genes that are best distinguished between the two cases or classes (e.g., normal vs. diseased). Such analyses produce a list of genes whose expression is considered to change and known as differentially expressed genes.

Recently, microarray technologies have opened up many windows of opportunity to investigate cancer diseases using gene expression. The primary task of a microarray data analysis is to determine a computational model from the given microarray data that can predict the class of the given unknown samples. The accuracy, quality, and robustness are important elements of microarray analysis. The accuracy of microarray dataset analysis depends on both the quality of the provided microarray data and the utilized analysis approach or objective. However the curse of dimensionality, the small number of samples and the level of irrelevant and noise genes makes classification task of a test sample more challenging. Those irrelevant genes not only introduce some unnecessary noise to gene expression data analysis, but also increase the dimensionality of the gene expression matrix. This results in the increase of the computational complexity in various consequent research objectives such as classification and clustering.

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1656/006 doi:10.1166/jctn.2018.7357

2. LITERATURE SURVEY

Feature selection has become the core areas of research for applications where dataset contains tens or hundreds or thousands of variables are available. These areas include text processing of internet documents, gene expression array analysis and combinatorial chemistry.

A wide range of such problems that arise are,

- providing better definition of the fitness function
- feature construction.
- feature ranking.
- efficient search methods.

In gene selection problem, the variables are gene expression coefficients corresponding to the abundance of mRNA in a sample, for a number of patients. The task of classification is to separate healthy patients from cancer patients, based on their gene expression "profile." Usually fewer than 100 examples (patients) are available altogether for training and testing. But the number of variables in the raw data changes from 6000 to 60,000. Some initial filtering usually brings the number of variables to thousand. Because the abundance of mRNA varies by several orders of magnitude depending on the gene, the variables are usually standardized. A DNA microarray is a collection of DNA spots attached to a solid surface. DNA microarrays are used to measure the expression levels of thousands of genes simultaneously. In DNA microarray experiments, the sample from which the expression levels of genes are to be measured is applied on the DNA microarray, the expression level of genes are measured. The resultant dataset is known as DNA microarray dataset, which is used for cancer classification. Various classifiers such as Support Vector Machine (SVM), Naïve Bayes, *K* Nearest Neighbour are employed for cancer classification. The feature selection is optimized currently by using swarm intelligence. In the existing system, the various types of particle swarm optimization is used for feature selection which is discussed in this chapter.

Xi et al.[1] proposed a binary quantum-behaved particle swarm optimization (BQPSO) for cancer feature gene selection, using support vector machine (SVM) for classification of gene samples. The proposed BQPSO algorithm is a discretized version of original QPSO for binary optimization problems. The procedure for cancer feature gene selection and classification is based on BQPSO and SVM with leave-one-out cross validation (LOOCV). The results show that BQPSO/SVM has advantages in accuracy, robustness, and the number of feature genes selected compared with the other two algorithms. This method produced less classification accuracy and more time consuming.

Moradi and Gholampour[2] proposed HPSO (Hybrid Particle Swarm Optimization) which uses a local search strategy that is combined with particle swarm optimization to select the less correlated and salient feature subset. The search process of the particle swarm optimization is to select distinct features by using the local search stratergy. Local search strategy shall be integrated with the multi-objective feature selection optimization methods to enhance the classification accuracy and reduce the number of features. The drawback of this method is that it increases the local search operations by clustering the features into several groups and also produces less classification accuracy.

Dhivya and Umamaheswari[3] proposed an unsupervised filter based method known as dynamic MBPSO (D-MBPSO) which integrates MBPSO into filter approach by defining new fitness function and it is independent of any learning model. Filter approach is used to quantify the relevance based on the intrinsic properties of the data. This produced reduced feature subset with good classification accuracy. This D-MBPSO method is time consuming because the filter approach has to evaluate the intrinsic properties of the data.

Ahmed F. Ali and Mohamed A. Tawhid[4] proposed HPSOGA that is based on three mechanisms. The first mechanism applies particle swarm optimization The second mechanism is to reduce the dimensions and the partition the population by dividing the population into sub-populations and applying the crossover operation in each sub-population in order to increase the diversity of the search in the algorithm. The last mechanism is applied to avoid the premature convergence and trapping in local minima by using the genetic mutation operator in the whole population. This method lead to premature convergence and lead to less classification accuracy.

Suresh Dara et al.[6] proposed a rough based hybrid binary PSO method in which redundant features are eliminated using rough set theory which is called as distinction table. This method uses a heuristic based fast processing stratergy. Objective function is used for optimization. The fitness function are used to reduce the features and increase the classification accuracy.

The PSO algorithm is an optimization algorithm and reduces the feature subset which consist of thousands of genes.

3. PSO ALGORITHM

PSO algorithm is a optimization and stochastic method developed by Eberhart and kennedy in 1995 based on social behaviors of birds flocking. PSO is similar to Genetic algorithm and other evolutionary algorithms. There are many variants of PSO. PSO is based on swarm intelligence approach, where the population has been searched for the global optimum solutions by updating the generations. Particles move through the search space to find the best position.

PSO is initialized with a group of random particles or solutions and then searches for optima by updating generations. In every iteration, each particle is updated by

following two "best" values. The first one is the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called pbest. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the population. This best value is a global best and called gbest. After finding the two best values, the particle updates its velocity and positions using the Eqs. (1) and (2) respectively.

$$v[\,] = w \times v[\,] + c1 \times \text{rand}[\,] \times \text{pbest}[\,] - \text{present}[\,] + c2 \times \text{rand}[\,] \times (\text{gbest}[\,] - \text{present}[\,]) \quad (1)$$

$$[\,] = \text{present}[\,] + v[\,] \quad (2)$$

$v[\,]$ is the particle velocity, persent[ ] is the current particle (solution). pbest[ ] and gbest[ ] are defined as stated before. rand( ) is a random number between (0, 1). $c1$, $c2$ are learning factors which is generated using random function. Eqs. (1) and (2) define the classical version of PSO algorithm. A constant, $V_{\max}$, was introduced to arbitrarily limit the velocities of the particles and improve the resolution of the search. The maximum velocity $V_{\max}$, serves as a constraint to control the global exploration ability of particle swarm. Further, the concept of an Inertia Weight was developed by Shi and Eberhart in 1998 to better control exploration and exploitation. The motivation was to eliminate the need for $V_{\max}$. The resulting velocity update is changed in Eq. (3),

$$[\,] = w \times [\,] + c1 \times [\,] \times \text{pbe}[\,] - \text{present}[\,] + c2 \times \text{rand}[\,] \times (\text{gbest}[\,] - \text{present}[\,]) \quad (3)$$

As there is a large effect of initial velocity in the balancing of exploration and exploitation process of swarm, Inertia Weight (w) is used to control the velocity. In the existing system of traffic light scheduling using PSO, the authors have used Linearly Decreasing Inertia weight for velocity updation. In a Linearly Decreasing Weight Particle Swarm Optimization (LDWPSO), a linearly decreasing inertia factor was introduced into the velocity of the updated equation from the original PSO. The performance of LDWPSO is significantly improved over the original PSO because LDWPSO balances out the global and local search abilities of the swarm effectively. In LDWPSO, W_{ldw} is the inertia weight which linearly decreases from 0.9 to 0.4 through the search process. The equation for the linearly decreased weight is written in Eq. (4),

$$W_{\text{ldw}} = (W_{\max} - W_{\min}) \times \frac{\text{Iteration\_max} - \text{Iteration}}{\text{Iteration\_max}} + W_{\min} \quad (4)$$

$W_{\max}$ is 0.9, $W_{\min}$ is 0.4 and Iteration_max is the maximum number of the allowed iterations and Iteration is the current iteration.

Linear Decreasing Inertia Weight PSO (LDIW-PSO) algorithm is known to have the shortcoming of premature convergence in solving complex (multipeak) optimization problems due to lack of enough momentum for particles to do exploitation as the algorithm approaches its terminal point and considering the whole dataset consumes more amount of time. In Global-Local Best Inertia Weight, the Inertia Weight is based on the function of local best and global best of the particles in each generation. It neither takes a constant value nor a linearly decreasing time-varying value.

4. PROPOSED SYSTEM

In Advanced PSO approach, the data is to be preprocessed and then the data is sent to advanced PSO, where the relevant features are selected and given for classification. Figure 1 represents the flow of the proposed system.

PSO approach consist of different phases namely,

1. Data preprocessing
2. Feature Selection
3. Classification
4. Performance Evaluation.

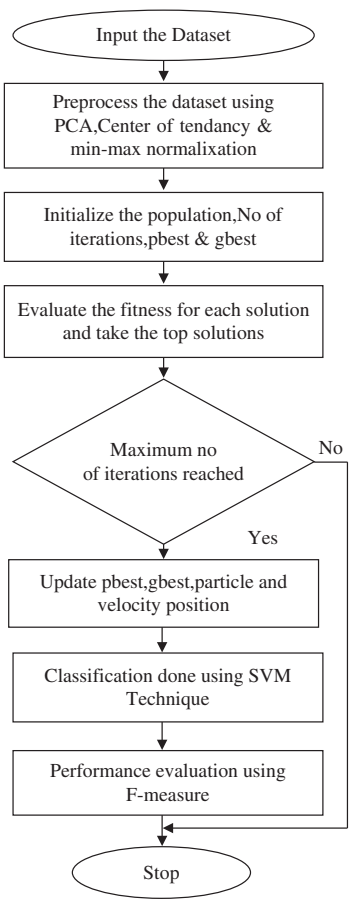


Fig. 1. Flow chart of the proposed system.

4.1. Data Preprocessing

Pre-processing of the data is carried out in three major steps namely dimensionality reduction, filling the missing values and data normalization.

- Dimensionality Reduction

Dimensionality reduction is the process of reducing the dimensions of the original dataset by obtaining a set of principal variables. The main linear technique for dimensionality reduction is principal component analysis, performs a linear mapping of the data to a lower-dimensional space in such a way that the variance of the data in the low-dimensional representation is maximized. The original space (with dimension of the number of points) has been reduced (with data loss, but hopefully retaining the most important variance) to the space spanned by a few eigenvectors.

Transforming the p original variables $X = [x1, x2, \ldots, xp]$ to a new set of K predictor variables, $T = [t1, t2, \ldots, tK]$, which are linear combinations of the original variables is given in Eq. (5)

$$U_k = \arg\max \mathrm{Var}(X_u)u'u = 1 \tag{5}$$

Where U_k is the predictor variable and X_u is the original variables.

- Filling the missing values

Missing data is one of the major issues, which leads to predisposition result. Missing Data will not make any impact on the result if its percentage is less, but if it exceeds more than 15% then it will surely hinder the result achieved after applying data mining techniques.

Centre of tendancy is a most commonly used technique and replaces missing values on a variable with the mean estimation of the observed values is given in Eq. (6),

$$a_j(x_i) = \sum a_j(x_i)/n \tag{6}$$

where n is the total number of instances and a_j denotes the jth attribute.

- Data normalization

Min–max normalization is often known as feature scaling where the values of a numeric range of a feature of data are reduced to a scale between between 0 and 1 represented in Eq. (7),

$$a_j(x_i) = (a_j(x_i) - \min_j)/(\max_j - \min_j) \tag{7}$$

where $a_j(x_i)$ is the value of the gene data for attribute a_j, $\min_j$—minimum value in the jth attribute and $\max_j$—maximum value in the jth attribute.

4.2. Feature Selection

Advanced Particle Swarm Optimization, is a population-based metaheuristic inspired by the social behavior of birds within a flock, and was initially designed for continuous optimization problems. In PSO, each potential solution to the problem is called a particle position and the population of particles is called the swarm. In this algorithm, the velocity of each particle are updated using the Eq. (8),

$$[\,] = w \times [\,] + c1 \times [\,] \times \mathrm{pbe}[\,] - \mathrm{presnt}[\,] + c2 \times \mathrm{rand}[\,] \times (\mathrm{gbest}[\,] - \mathrm{present}[\,]) \tag{8}$$

Where W is the inertia weight calculated for the particle. $c1$ and $c2$ are constants which is equal to 0.5 rand_1 and rand_2 are random numbers between 0 and 1 that is generated using the rand( ) function. Pbest[ ] is the particle best. Gbest[ ] is the global best for the whole population.

Inertia factor W according to distance of the particles of a particular generation from the global best. The value of W for each particle is given by the equation,

$$W = W_0 \times \left(1 - \frac{\mathrm{dist\_}i}{\mathrm{max\_dist}}\right) \tag{9}$$

where W_0 is the random number between 0.5 and 1, dist_i is the hamming distance of i-th particle from the global best particle and max_dist is the maximum possible hamming distance.

The position update is done using the Eq. (2). The particle with maximum fitness value is considered as the best particle as this is maximization problem where the accuracy is to be improved.

The pseudocode of the advanced PSO is stated in Figure 2.

4.3. Population Initialization

Population is initialized with random values of 0 and 1 using the random function. The size of the population is equal to the total size of the dataset and the number of columns is equal to the number of attributes excluding the class attribute.

```
For each particle
  Initialize particle
  Calculate fitness value according to the equation
    fitness = 1/(∑dataset(i, j) + 1)
  Arrange calculated fitness for each solution in descending order
  Take the first half of the fitness
End
Do
  For each particle
    Calculate fitness value according to the below equation
    fitness = 1/(∑dataset(i, j) + 1)
    If the fitness value is better than the best
fitness value (pBest) in history
    set current value as the new pBest
  End
  Choose the particle with the best fitness value of
all the particles as the gBest
    For each particle
    Calculate particle velocity according to Eq. (8)
    Update particle position according to Eq. (2)
  End
While maximum iterations or minimum error criteria is not attained
```

Fig. 2. Pseudocode for advanced PSO.

4.4. Fitness Evaluation

The fitness value of each solution is evaluated by adding up the "values" of all the genes. This is designated in Eq. (10),

$$S = \sum(\text{value of each gene in the dataset}) \quad (10)$$

Our aim is to maximize the fitness value. Fitness is evaluated using the Eq. (11),

$$\text{Fitness} = 1/(S+1) \quad (11)$$

Fitness value calculated for each solution is called as the pBest (Partcile Best). gBest (Global Best) is the maximum of pBest. Arrange all the pBest values in descending order to segregate the total population size into half for further processing. when this is done to reduce the time and to improve the overall performance of the system.

4.5. Velocity and Position Updation

Velocity and positions are updated in order to evolve the particles to give the best solutions using the Eqs. (8) and (2).

4.6. Classification

Support vector machines (SVMs) are a set of related supervised learning methods that analyze data and recognize patterns, used for classification (machine learning) and regression analysis. The standard SVM is a nonprobabilistic binary classifier or binary linear classifier that predicts, for each given input, which of two possible classes. A support vector machine constructs a hyperplane or set of hyperplanes in a high or infinite dimensional space, which can be used for classification, regression or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

4.7. Performance Evaluation

In statistical analysis of binary classification, the F1 score (also F-score or F-measure) is a measure of a test's accuracy. It considers both the precision p and the recall r of the test to compute the score: p is the number of correct positive results divided by the number of all positive results, and r is the number of correct positive results divided by the number of positive results that should have been returned. The F1 score is the harmonic average of the precision and recall, where an F1 score reaches its best value at 1 (perfect precision and recall) and worst at 0. In the breast cancer dataset, the whole dataset is first divided into testing and training and then classification of the data is using the Eqs. (11)–(13).

$$\text{Precision} = \text{TP}/(\text{TP}+\text{FP}) \quad (12)$$

$$\text{Recall} = \text{TP}/(\text{TP}+\text{FN}) \quad (13)$$

$$\text{F1 score} = 2*(\text{Precision}*\text{Recall})/(\text{Precision}+\text{Recall}) \quad (14)$$

Where TP is the True Positive, FP is the False Positive, FN is the False Negative.

5. EXPERIMENTAL RESULTS

Any dataset which has integer or floating point values as its attributes can be given as input. In this project, breast cancer dataset is used which is taken from UCI Repository database. The implementation of the algorithm for the given breast cancer dataset is done using MATLAB The attribute of the dataset contains the values ranging from 1 to 10 and details of the dataset is given in Table I.

The dataset is preprocessed by using the preprocessing techniques. The techniques that are used on the data are Principal Component Analysis (PCA), Center of tendancy (Mean) and Min-Max normalization. After preprocessing the values in the dataset are in the range 0 to 1. This preprocessed dataset is given for feature selection which contains advanced PSO method. Using advanced PSO method, the population is randomly initialized with values 0 and 1. The size of the population in PSO algorithm is equal to the size of the dataset. Initially the fitness is calculated and then the fitness is sorted based on descending order and half of the dataset which contains the maximum fitness is selected for further evaluation. Then velocity and position is updated for each and every iteration. The solution with maximum fitness is selected for further classification. Advanced PSO algorithm runs for about 0.2000 ms. For classification, initially the dataset is divided into testing and training data. 70% of the original data is used for training and the remaining 30% of the data is used for testing. SVM Classifier is used for classification and classification error is generated. The Table II that is given below consists of the best solutions, which is taken as an output of

Table I. Description of the dataset.

Dataset	Breast cancer
Source	https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(original)
No. of instances	699
No. of attributes	10 + 1 class attribute
Class	2 for benign, 4 for malignant

Table II. Solutions with classification error.

Sol. no.	Solutions									Loss
1	1	0	1	1	0	1	1	0	0	0.0141
2	1	1	0	0	1	1	0	1	0	0.0181
3	1	0	1	0	0	1	1	0	0	0.0118
4	0	0	0	0	1	0	0	0	0	0.0471
5	0	1	1	1	1	1	1	1	0	0.0208
6	1	0	1	0	0	1	1	0	0	0.0118

Table III. Testing and training accuracy of solutions.

Sol. no.	Training accuracy	Testing accuracy
1	96.8815	99.061
2	96.8815	99.0654
3	96.4803	99.2974
4	88.0503	96.1353
5	96.4509	98.829
6	96.281	99.2974

Table IV. Comparison of accuracy for PSO and advanced PSO.

PSO	Advanced PSO
98.2974	99.2974

advanced PSO algorithm and classification error for each solution.

The performance of the developed algorithm is measured using F1-Score or *F* Measure. Performance accuracy is calculated for both testing and training data for advanced PSO algorithm. About 95% of accuracy is obtained using SVM classifier and it is tested for different solutions. The solution which is nothing but the population, which provides less classification error and more accuracy in both testing and training datasets is said to be efficient solution. The testing and training accuracy for classification is given along with the solution number is given in Table III.

The solution number 3 and 6 provides better training and testing accuracy. The advanced PSO performs better when compared to PSO algorithm. Performance evaluation for both PSO and advanced PSO algorithm is given in Table IV.

6. OBSERVATION AND FUTURE ENHANCEMENTS

The data for cancer diagnosis or classification continues to grow and each study have limited power and validity, there is a need for more accurate analysis. The use of advanced PSO algorithm produced better accuracy. The traditional PSO algorithm is more time consuming and produces less fitness. This eventually lead to the development of advanced PSO algorithm which selects the features with high fitness for feature selection. This method in turn reduces the dataset into half by selecting the top ranked features. Use of SVM technique for classification has produced greater accuracy levels with reduced feature sets. Further enhancements to advanced PSO algorithm can be done which combines some other evolutionary algorithms.

References

1. M. Xi, J. Sun, L. Liu, F. Fan, and X. Wu, Cancer Feature Selection and Classification Using a Binary Quantum-Behaved Particle Swarm Optimization and Support Vector Machine, Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine **(2016)**, Vol. 2016, pp. 1–9.
2. P. Moradi and M. Gholampour, *Applied Soft Computing* 43, 117 **(2016)**.
3. M. Dhivya and K. Umamaheswari, D-MBPSO: An unsupervised feature selection algorithm based on PSO, *Proceedings of the 6th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2015), Held in Kochi, India during December 16–18, 2015, Innovations in BioInspired Computing and Applications, Advances in Intelligent Systems and Computing*, © Springer, International Publishing Switzerland **(2016)**, Vol. 424, pp. 359–369.
4. Ahmed F. Ali and Mohamed A. Tawhid, *Ain Shams Engineering Journal* 8, 191 **(2016)**.
5. Barnali Sahu and Debahuti Mishra, A novel feature selection algorithm using particle swarm optimization for cancer microarray data, *International Conference on Modeling Optimization and Computing* **(2012)**, pp. 27–31.
6. Suresh Dara, Haider Banka, and Chandra Sekhara Rao Annavarapu, A Rough Based Hybrid Binary PSO Algorithm for Flat Feature Selection and Classification in Gene Expression Data, Springer Publications, Berlin, Heidelberg **(2017)**.
7. Y. Shi and R. Eberhart, A modified particle swarm optimizer, *In the Proceedings of IEEE International Conference on Evolutionary Computing* **(1998)**, pp. 69–73.
8. L. Y. Chuang, H. W. Chang, C. J. Tu, and C. H. Yang, *Comput. Biol. Chem* 32, 29 **(2008)**.
9. J. Tang, S. Alelyani, and H. Liu, Feature selection for classification: A review, Data Classification: Algorithms and Applications, CRC Press, Cambridge **(2014)**.
10. J. Sun, W. Xu, and B. Feng, A global search stratergy of quantum-behaved particle swarm optimization, *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems*, Singapore **(2004)**, Vol. 1, pp. 111–116.
11. Chandra Sekhara Rao Annavarapu, Suresh Dara, and Haider Banka, *EXCLI Journal* 460 **(2016)**.
12. M. Banerjee, S. Mitra, and H. Banka, *IEEE Trans. Syst. Man. Cybern. Part C (Applications and Reviews)* 6, 22 **(2007)**.
13. G. Chandrashekar and F. Sahin, *Comput. Elec. Engin.* 40, 16 **(2014)**.

Received: 24 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1662–1665, 2018

RESEARCH ARTICLE

A Survey Based on Human Emotion Identification Using Machine Learning and Deep Learning

C. Akalya Devi[1], D. KarthikaRenuka[1], and S. Soundarya[2,*]

[1]*Dept. of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India*
[2]*PG Scholar, Dept. of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India*

An Emotion can be identified by various activities of a person. A human can express their feelings or Emotions in different ways. Emotions are classified into different types like happy, disgust, anger and so on. Emotions can be detected or identified in variety of ways. Emotion can be recognized by body gestures, EEG signals, face images, text and so on. A survey or review is made on identifying the human emotions using EEG signals with help of Machine Learning and Deep Learning Techniques. In this paper, a review is done to find the best algorithms for finding an Emotion of a person by analyzing various Machine Learning and Deep Learning algorithms with the help of previous papers.

Keywords: Emotions, Machine Learning, Deep Learning, EEG.

1. INTRODUCTION

Emotion plays major role in human daily life, not only for interaction but also in many part of our life. Emotion detection technology is making a huge difference in how we leverage text analysis and EEG signals. Especially, in the field of marketing and health care. Emotion Detection and Recognition from EEG signals is a recent field of research that is closely related to Sentiment Analysis. Sentiment Analysis aims to detect positive, neutral, or negative feelings from EEG signals, whereas Emotion Analysis aims to detect and recognize types of feelings through the expression of texts, such as *anger*, *disgust*, *fear*, *happiness*, *sadness*, and *surprise*. A survey on various techniques in machine learning and deep learning for detecting or identifying human emotions has been discussed in this paper.

2. RELATED WORKS

Emotion detection using EEG signals is a universal issue and causes difficulties due to uncertain physical and psychological characteristics of emotions.

In Paper [1], they analysed the previous works from 2009 to 2016, that propose novel methods for the recognition of emotions through EEG signals. Their analysis draws on two perspectives:

1. A set of recommendations to avoid the common pitfalls of this type of research area.
2. Another more specific that considers the aspects involved in the process of recognizing emotions from EEG signals (e.g., subjects, features extracted, classifiers, etc.).

For this they analysed all papers from 2009 to 2016 and they go through each and every algorithm they used for EEG signals pre-processing, feature extraction and classification. In this survey paper, they derive a set of best practice recommendations to help researchers produce well validated and high-quality works, and also able to be reproducible and replicable.

In Paper [2], they reviewed emotion classifications, emotion evoking experiments and emotion recognition algorithms. They proposed and implemented a novel fractal dimension based algorithm for recognition of emotions from EEG in real time. They implemented the algorithm along with Haptek system.

In Paper [3], a DREAMER database had taken for analysis of emotion from EEG signals. The database includes EEG and ECG recordings from 23 participants, where each participant rated his emotional response along the scales of valence, arousal, and dominance, after watching each of the 18 film clips selected to elicit specific emotions. The DREAMER database made publically available after publication of this paper. In order to give the opportunity to researchers to evaluate their algorithm using this database.

In this Paper [4], they classified emotions using EEG signals using various Deep Learning and Machine Learning algorithms. They compared the results each and every algorithm they used for emotion classification. In their

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1662/004 doi:10.1166/jctn.2018.7358

work they used Deep models like DBN (deep belif networks) and DBN-HMM (A hidden markov model (HMM) is integrated along with DBN). They also compare the performance of the deep models to KNN, SVM and Graph regularized Extreme Learning Machine (GELM). The average accuracies they got for DBN-HMM, DBN, GELM, SVM, and KNN in their experiments are 87.62%, 86.91%, 85.67%, 84.08%, and 69.66%, respectively. From their experimental results they show that DBN and DBN-HMM models improve the accuracy of EEG based emotion classification in comparison with other methods.

In Paper [5], they used HCNN (hierarchical convolutional neural network) for classification of different states of emotion. They also used stacked autoencoder (SAE), SVM, and KNN as competing methods. As a result deep models (HCNN and SAE) show absolute advantage over traditional shallow models (SVM and KNN). They confirmed that the high-frequency wave bands Beta and Gamma were the most suitable bands for emotion reading. HCNN gives them the higher accuracies when compared with other methods, and SAE is slightly inferior than HCNN.

In Article [6], the classification of emotion have implemented using the DLN (deep learning network) and machine learning algorithms like SVM and Naive Bayes classifier. Then they enhanced with covariate shift adaptation of the principal components. The deep learning network they used constitute of three stack autoencoders and two softmax classifiers for valence and arousal state classifications. As a result of their work, they got classification accuracy of the DLN with PCA + CSA is 53.42% and 52.05% to classify three levels of valence states and three levels of arousal states. Thus they conclude DLN provides better accuracy performance compared to SVM and Naive Bayes classifier.

In Paper [7], the different feature extraction and classification algorithms for EEG signals have discussed. In this paper, they focused on several feature extraction methods like,

- Discrete wavelet Transformation
- Higher Order Crossing
- Principal Component Analysis
- Nonlinear Dynamic Analysis
- Independent Component Analysis
- Short Time Fourier Transform and Mutual
- Information
- Statistical Based Feature.

The classification algorithms they discussed in this paper is,

- Neural Network
- K-Nearest Neighbor
- Support vector Machine
- Linear Dynamic Analysis.

They discussed the advantages and disadvantages of all feature extraction and classification algorithms mentioned above. They conclude from that neural network and SVM provides higher accuracy when compared with other algorithm mentioned above.

In Paper [8], they made a survey on classification and feature extraction for EEG signals. The feature extraction techniques they used for their study, were PCA and ICA techniques. The classification algorithms they used for their study were Support Vector Machine and K-means Algorithm. They just discussed about the methods in their survey paper.

In Paper [9], they done a survey based on feature extraction methods and feature classification techniques. They discussed the advantages and disadvantages of each and every method they mentioned in their paper for feature extraction and feature classification techniques for classification of different types of emotions using EEG signals.

In Paper [10], the discussion done to tell the effectiveness of Neural Networks to classify user emotions using EEG signals from the DEAP database. They used two deep learning models to classify the user emotion states. The deep learning models they used were DNN and CNN. They classified the emotion for valance and arousal classes using EEG signals from DEAP database. From their comparison they found CNN provides better results when compared to DNN. They also compare our results with that of, for classification in 3 classes and for classification in 2 classes, which represent the previous state of the art classification accuracy for both Valence and Arousal.

3. MACHINE LEARNING TECHNIQUES

Machine learning is a scientific discipline that deals with the construction and study of algorithms that can learn from data (R. Kohavi and F. Provost 1998). Such algorithms operate by building a model based on inputs and using these inputs to make predictions or decisions, rather than following only explicitly programmed instructions (C. M. Bishop, 2006). Specifically in emotion detection, Machine learning algorithms are used to learn how to detect emotions. These approaches can be divided into supervised and unsupervised learning. A. Supervised Learning approaches rely on a labelled training data, a set of training examples. The supervised learning algorithm analyses the training data and infers a function, which we use for mapping new examples.

In the Machine Learning (ML) approach to sentiment analysis, the classifier automatically learns the properties and categories from the pre-classified training documents. ML based classification is called as supervised learning because this process is guided by the labeled training set. There exists various machine learning approaches, some of them are explained as below.

3.1. Support Vector Machines

The fastest and most widely used algorithm among all the ML algorithms is the support vector machine (SVM).

A binary SVM is a hyperplane separating the feature space of positive instances from the feature space of negative instances. During the training phase, the hyperplane that can separate the positive feature space from the negative feature space with a maximal margin is chosen. The margin is distance of the nearest point from the positive and negative sets to the hyperplane. Support vectors, the subset of the training instances, determine the hyperplane for a SVM. SVM classifiers perform extremely well irrespective of the dimensionality of the feature space.

3.2. Decision Tree Classifiers

A Decision Tree (DT) classifier resembles a tree in which the features are represented by nodes and the edges leaving a node are labeled by the feature weight and leaves represent the categories. The tree is constructed based on a recursive procedure. At each step a feature 'F' is picked and the training collection is divided into two groups, one containing 'F' and another not containing 'F'. This procedure is carried out until only documents of a single category remain. A leaf is generated at the end of this procedure. Information gain or entropy is used to choose a feature at each step. Generally the DT classifiers are used as baseline classifiers.

3.3. Fuzzy Logic

Fuzzy logic based systems are handy in real life situations where the decision to be taken are based on multiple criteria with complex interlink among them. It is very true for a sentiment analysis process in which the system must be able to understand the sentiment expressed by a customer in a review based on the statements about various features of the product or service. For example, in a movie review, the reviewer may praise the director for the usage of technology and blame on the acting and story. Deciding on the overall sentiment as positive or negative depends on the opinion words or phrases used by the reviewer for each of the features. When the number of features is more, the complexity in the decision making gets added and hence the decision making becomes tough. In such situations, fuzzy logic can be effectively used.

3.4. Probabilistic Classifier

Naïve Bayes Algorithm Probabilistic classifiers use the Bayes' theorem to calculate the probability $P\ (c \mid d)$, that a document belongs to a category c.

$$P(c \mid d) = P(d \mid c)P(c)/P(d)$$

In order to determine the probability $P(d \mid c)$, it is assumed that the coordinates representing the document as a feature vector are independent. These classifiers are called as Naïve Bayes (NB) classifiers.

3.5. Neural Networks

Neural network (NN) can be designed to carry out the task of opinion mining. The features of a document are the input nodes, the output nodes deliver the category. The dependence relations are taken care by the link weights (M. Lam, 2004). Generally the NN are trained by back propagation, i.e., the training documents are fed into the input nodes and if a wrong classification occurs, the error is propagated back in the network to minimize the error by adjusting the link weights. Perceptron is the simplest kind of a NN, which has only two layers viz. input layer and output layer. A multilayer perceptron contains one or more hidden layers between the input and output layers.

4. DEEP LEARNING TECHNIQUES

Emotion Detection can be performed efficiently by implementing different deep learning models, which have been extended recently. These models include CNN (convolutional neural networks), DNN (deep neural networks) and DBN (deep belief networks). This describes the efforts of different researchers toward implementing deep learning models for performing the emotion detection.

4.1. Convolutional Neural Networks (CNN)

The CNN (convolutional neural network) includes pooling layers and sophistication as it gives a standard architecture to map the signals of variable length into signals of fixed size scattered vectors. This study has proposed a novel convolutional neural network (CNN) framework for classification of emotions using EEG signals. CNN has been implemented using Tensorflow on a windows machine. As CNN enhance its performance by increasing its size and depth, so a very deep CNN model, inspired by Tensorflow is proposed with 22 layers for Emotion analysis. CNN use four different layers in it for classification. The layers in CNN were convolutional layer, pooling layer, flattening layer and the fully connected layer. Each layer performs some operation and output from each layer will be given as a input to the next layer. Pooling layer will perform down-sampling in it. Fully connected layer is the layer where we get the classification result.

4.2. Deep Belief Networks (DBN)

Deep belief networks (DBNs) includes several hidden layers, composed by RBM (restricted Boltzmann machines). DBN has been proved efficient for feature representation. It utilizes the unlabeled data and fulfills the deficiencies of labelled analysis issues. The features of language specific and inter language have been presented through building multiple weakly shared layers of features. In comparison with existing studies the proposed work address the challenge of shortening overlap among feature spaces of both source and target language data through cross lingual information transfer process using back propagation.

DNNs used for transformation of information from source to target language. The experiments have been conducted for emotion classification tasks of EEG signals taken from brain waves. When we add more hidden layers it gives better results for us.

5. CONCLUSION

As for Machine Learning approaches, the supervised learning approach is more used in emotion detection because it usually leads to better results than unsupervised learning. The unsupervised learning approaches need labelling training examples and annotating of examples, which is a time-consuming task. For this reason, several researches have analyzed and realized. Although, it is important to use good feature extraction techniques in Machine Learning algorithms to obtain good results. Moreover, in emotional detection systems based on machine learning approach, we have detected that most of these systems use features based on a pre-processed EEG signals. Hence, we propose a new direction focuses on deep model analysis, since we consider that if we use features based on a deep analysis on EEG signals we could improve the emotional detection systems. Deep learning consists of numerous effective and popular models, these models are used to solve the variety of problems effectively. Different studies have been discussed in this review to provide a deep knowledge of the successful growing of deep learning applications in the field of Emotion analysis.

References

1. S. M. Alarcao and M. J. Fonseca, *IEEE Transactions on Affective Computing* **(2017)**.
2. Y.-J. Liu, M. Yu, G. Zhao, J. Song, Y. Ge, and Y. Shi, *IEEE Transactions on Affective Computing* **(2017)**.
3. S. Katsigiannis and Naeem Ramzan, *IEEE Journal of Biomedical and Health Informatics* **(2017)**.
4. W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, EEG-based emotion classification using deep belief networks, *2014 IEEE International Conference on Multimedia and Expo (ICME)*, IEEE **(2014)**, pp. 1–6.
5. J. Li, Z. Zhang, and H. He, *Cognitive Computation* 1 **(2017)**.
6. S. Jirayucharoensak, Setha Pan-Ngum, and Pasin Israsena, *The Scientific World Journal* 2014 **(2014)**.
7. M. Sreeshakthy, J. Preethi, and A. Dhilipan, *International Journal of Information Technology and Computer Science (IJITCS)* 8, 19 **(2016)**.
8. Shamla Mantri, Vipul Patil, and Rachana Mitkar, *International Journal of Engineering Research and Development* 4, 24 **(2012)**.
9. F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, *Journal of Neural Engineering* 4, R1 **(2007)**.
10. Samarth Tripathi, Shrinivas Acharya, Ranti Dev Sharma, Sudhanshu Mittal, and Samit Bhattacharya, Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset, AAAI **(2017)**, pp. 4746–4752.

Received: 24 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1666–1674, 2018

RESEARCH ARTICLE

Design of Coarse Grained Architecture with Gated Clock Technique for Low Power Applications

Yazhinian Sougoumar[1,*] and Tamilselvan Sadasivam[2]

[1] *Research Scholar, Department of Electronics and Communication Engineering, Pondicherry Engineering College, Pondicherry 605014, India*
[2] *Department of Electronics and Communication Engineering, Pondicherry Engineering College, Pondicherry 605014, India*

In recent times, reconfiguration system is widely used due to their combination of efficiency and flexibility. Hardware based reconfiguration offers to use the same hardware for the different purpose of the optimized hardware utilization. Coarse grained (CGRA) and fine grained (FGRA) are the two types of reconfigurable architectures used in the present time. In the CGRA, the hardware functionality is specified at the word level. Likewise, the FGRA, hardware functionality is specified at a bit level. Coarse grained reconfigurable architecture (CGRA) needs many processing elements (PEs) and configuration memory for reconfiguration operation. CGRA reconfigure the processing element from one process to another process dynamically without any performance degradation. The CGRA system is designed with 16 processing elements; it's arranged in a 4×4 matrix structure. Each processing elements are configured by more than one operation. The operation can be selected and changed by using the selection bits. According to the selection bit, the operation of the processing element is changed; likewise, reconfiguration bit is used to reconfigure the processing elements. Although the reconfiguration architecture provides higher performance and flexibility, it consumes a significant amount of power. Reducing power in the CGRA and FGRA is the very critical issue; avoid this kind of situation by proposed the clock gating and pipelining reuse techniques. The proposed techniques avoid the unwanted usage of the clock signal during the operation of processing elements. Interconnections between the processing elements are essential in the CGRA structure. The proposed architecture was simulated by Modelsim 6.3c, synthesized and analyzed by Xilinx ISE simulator and Quartus with cyclone II device.

Keywords: Coarse Grained Reconfigurable Architecture (CGRA), Fine Grained Reconfigurable Architecture (FGRA), Clock Gating, Pipeline Reuse, Quartus II, Xilinx ISE.

1. INTRODUCTION

Reconfiguration provides high performance at low power. It is mostly used for mobile and embedded systems applications. Reconfigurable computing architecture is a process of joining the flexibility of software with the high performance of hardware by processing the high-speed computing like Field programmable gate array (FPGAs). FPGAs are also support the partial reconfiguration. It is the process of changing half of the portion of reconfigurable hardware, while the other part is still running. Advances in the silicon technologies, increase the number of transistors that can be integrated into a single device. Chip manufacturer uses this integration technique to fabricate the system on a single silicon device. FPGA manufacturers also using this technique to implement the DSP blocks, multipliers, a memory block and many logic elements along with the reconfigurable fabric system. The reconfigurations in the devices are reduced power consumption, flexibility and hardware reuse. Reconfiguration architecture contains fine grained and coarse grained reconfigurable architecture. Fine grained reconfigurable architecture programmable block and routing logic operated at a bit level. It is based on an array of homogeneous reconfigurable processing elements, which can be configured as logic interconnects. This design provides flexibility for distributing the hardware resources between the logic interconnects. It could obtain high performance for a huge number of algorithms, because of its bit level reconfiguration. It provides an efficient communication medium for exchanging operations including memories. Likewise, the coarse grained reconfigurable architecture (CGRA) operated at the word level. CGRA is heterogeneous. CGRA is building an array of functional units which communicate using local and global interconnects. Functional units divided

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1666/009 doi:10.1166/jctn.2018.7359

into logic and memory elements. All the functional units are dynamically reconfigurable to support the run time application. It is easier to program and fast reconfiguration when compared to fine grained reconfigurable architecture (FGRA). CGRA provides mapping flexibility to reduce the reconfiguration run time and achieve higher performance using the word level operation. Both the reconfiguration is constructed in the FPGAs for achieving high speed and low power applications. Modern FPGAs are used for much application like medical, industrial, video and image processing etc.

2. RELATED WORKS

Coarse grained reconfigurable architectures (CGRAs) need many logic elements (LEs) and a configuration memory block for reconfiguration of its LE array. It consumes more power; reducing power is very difficult for CGRA. Kim et al.[1] proposed the reusable context pipelining (RCP) architecture to reduce power consumption caused by reconfiguration. The characteristic of loop pipelining provides the power reduction. It is multiple instruction streams, and multiple data stream style execution model RCP efficiently reduces the power consumption in configuration memory without any performance changes. The proposed architecture is supported for low power reconfiguration and hybrid configuration. The architecture used to achieve power saving in the reconfigurable architecture; the performance is same as traditional CGRA.

The coarse grained architecture contains context memory and the data memory organizations. Performance, area and power are the parameters to consider, achieve a greater tradeoffs. Wang et al.[2] proposed the two techniques named as hierarchical configuration context (HCC) and the lifetime based data memory organization (LDO) concentrating on the context memory and the data memory. The proposed techniques are to compress the on-chip memory space and reduce the reconfiguration time. Hiware and Padole[3] presented a configuration memory based dynamic coarse grained reconfigurable multicore architecture. CGRA is a reconfigurable system, using the byte computing. Coarse grained architecture changes the hardware configuration dynamically depending on the application. Proposed dynamic reconfigurable architecture changes its hardware based on configuration codes stored in the configuration memory.

NATURE architecture uses the concept of temporal logic folding and fine grained dynamic reconfiguration to increases the logic density. So Lin et al.[4] have been proposed the fine-grain dynamically reconfigurable (FDR) that consists of an array of homogeneous reconfigurable processing elements (PEs). Each PEs can be configured into the LUTs. The proposed architecture eliminates most of the long distance and global wires, which occupy a huge area in the conventional FPGAs. Alnajjar et al.[5] proposed the novel concept named as flexible reliability in the coarse-grained reconfigurable architecture. It deals with soft errors and aging. TMR, DMR and SMS mode provides the error correction in the configuration. TMR mode provides error correction; likewise, DMR provides error detection. The aging reduces the circuit delay and power consumption by the system.

Conventional FPGAs have long reconfiguration latency. To overcome these limitation novel dynamically, reconfigurable architecture is proposed by Warrier et al.[6] It enables run time reconfiguration and hardware reuse. Modern FPGAs are used in computation intensive applications, coarse grain DSP blocks are needed to boost the performance. Reconfigurable DSP block design for dynamically reconfigurable architecture used for different arithmetic functions in different clock cycles. Yin et al.[7] have been proposed the Nested loop is pipelining on CGRA. Loops are implemented on CGRA for acceleration. The nested loop pipelining used to exploit the parallelism of loops. To use of affine transformation and polyhedral model reduces the loop pipelining problem. Parallelism improves the performance of the CGRA.

Kim et al.[8] presented a coarse grained reconfigurable architecture (CGRA) based multi-core architecture provides high performance by kernel-level parallelism. Existing architecture suffer from power consumption and performance bottleneck. The proposed ring based sharing fabric (RSF) to boost their flexibility. Filho et al.[9] introduced an architecture description language targeted to describe the coarse grain reconfigurable architecture template. This language allows fast modeling and analysis of the architecture. The proposed language enables a formal validation, analysis and estimation of hardware cost. The concept of reconfigurable computing and reconfigurable hardware was presented by konstantinos et al.[10] different types and classification of reconfigurable architecture is helpful for developing the architecture in the future.

3. GRANULARITY OF RECONFIGURATION

Reconfigurable systems are classified by the granularity. It shows which architecture is best for reconfiguration. Reconfiguration architecture can be classified into fine-grained, coarse-grained, medium-grained and mixed-grained. Here discuss the fine-grained and coarse grained reconfigurable architecture. Reconfigurable hardware with bit level is named as fine-grained reconfigurable hardware architecture; hardware operated with word level is called as coarse-grained reconfigurable hardware. Coarse grained architecture is easy to design and reconfigure but less flexible than fine-grained architecture. But the coarse-grained architectures are faster and more efficient because it runs in word level operation. CGRA system is trying to improve the performance and reduce the power consumption of the reconfigurable system.

In fine-grained architectures, the basic programmed building block consists of combinational and sequential

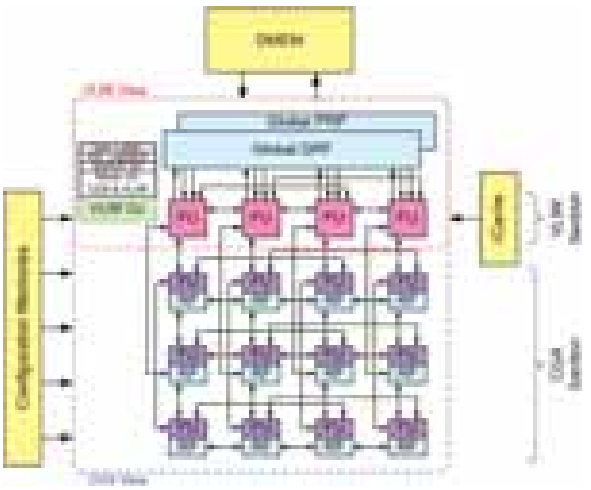


Fig. 1. General CGRA architecture.

circuits. The logic element in the architecture decides the operation of the reconfigurable system. Here the logic element can be programmed into a simple logic function, such as two-bit adder. These elements are connected to the reconfigurable interconnection. The task of the logic element is simple, so it provides more flexible. But comes to more critical task the performance and efficiency are very low. In case logic blocks are replaced by an 8-bit adder, the architecture provides low efficiency and occupies more space in the fine-grained implementation. Medium-grained reconfigurable systems use the logic block of large granularity. Operations are optimized in the medium-grained architecture. Because of these, the chip area is reduced when compared to fine-grained architecture.

CGRA architecture is shown in Figure 1; Coarse-grained architecture logic blocks are optimizing the large computations during the operation. It will perform these operations more quickly and also consume less area. Coarse-grained architecture is more efficient than the fine-grained architecture for implementing the functions. Because it is operated at word level operation. It requires less amount of reconfiguration information for reconfigurable hardware with fine granularity. Less amount of reconfiguration information will accelerate the speed of reconfiguration. Reconfiguration time is short when compared to fine-grained architecture. Interconnection overhead is small in the coarse-grained architecture.

4. TYPES OF RECONFIGURATION

The Reconfiguration scheme of the reconfigurable system is divided into three categories: (1) Logic Reconfiguration, (2) Static Reconfiguration, (3) Dynamic Reconfiguration. The reconfiguration of both the logic elements and the interconnection is achieved by using SRAM memory bit to control the configuration of transistors. The functionality

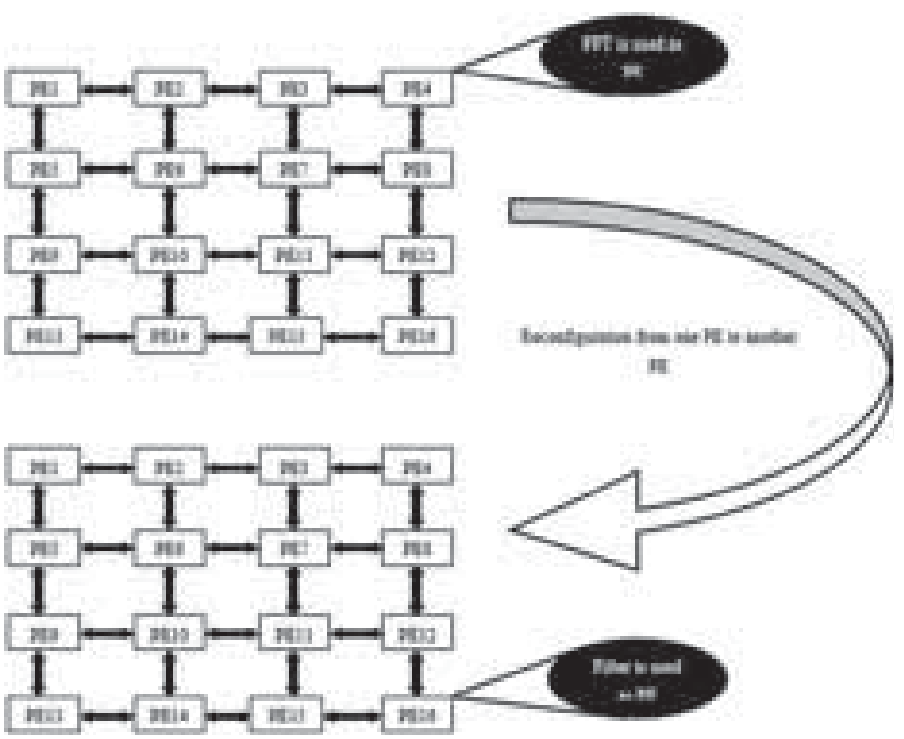


Fig. 2. Reconfiguration from one PE to another PE.

Table I. Comparison between fine grained and coarse grained architecture.

Characteristics	Fine grained	Coarse grained
Operation	Bit level	Word level
Flexibility	High	Medium
Performance	Medium	High
Reconfiguration time	Large	Small
Interconnection	Long	Short

of the logic elements, I/O units, and the interconnection network is changed by getting bit stream of reconfiguration data onto the hardware.

Static reconfiguration changes the hardware configuration when the system stops entire processing. But the dynamic reconfiguration allowed changing the hardware configuration when the system is processing the function. Static reconfiguration also referred as compile time reconfiguration. It is simplest and most common approach for implementing the application. Likewise, Dynamic reconfiguration also referred as run time reconfiguration. It uses a dynamic allocation steps that reallocate the hardware at run time. This type of reconfiguration increases the performance, flexibility and functional density. It is based on the concept of virtual hardware. In Figure 2 shows the reconfiguration of the processing element (PE) from one operation to another operation. For example, the fourth processing element performs both the FFT and filtering operations, and it can be done by using dynamic reconfiguration without any performance degradation.

5. DESIGN OF PROCESSING ELEMENTS (PE)

Here designed the 16 processing elements into a single CGR architecture. The CGRA system has been designed with the array of 4×4 PEs. The arrays of processing elements perform operations according to the opcode given. The lists of Processing Elements designed in the system are:

1. Data Encryption 2. Data Decryption 3. Advanced Russian Peasant Multiplier 4. FIR filters 5. Comparator 6. Fast Fourier Transform (FFT) processor 7. Hamming encoder 8. Hamming decoder 9. Convolution 10. Factorial 11. Cyclic redundancy check (CRC) 12. ALU 13. QAM Modulation 14. QAM Demodulation. These are processing elements used in the CGRA architecture.

5.1. Data Security

In Real time all system needs security to protect the data from the attackers. Various data security techniques are available on the market; one of the efficient security techniques is Advanced Encryption Standard (AES). It is one of the commercial security algorithms used for satellite communication, net-banking, etc. It completes the operation by two processes, named as encryption and decryption. AES uses the symmetric block cipher technique. The Same key is used for both encryption and decryption operations.

Table II. List of operations according to the selection line.

PE no.	Selection line	Operation
1	0000	Data encryption
2	0001	Data decryption
3	0010	Advanced Russian peasant multiplier
4	0011	FIR filter
5	0100	Comparator
6	0101	FFT
7	0110	Hamming encoder
8	0111	Hamming decoder
9	1000	Convolution
10	1001	Factorial
11	1010	Cyclic redundancy check
12	1011	ALU
13	1100	QAM modulation
14	1101	QAM demodulation

5.2. FIR Filter

Digital Signal Processing (DSP) operations are widely used in wireless communication Technologies to control and guide the signal flows. Convolution, Correlation, Frequency Transformation and filtering are the important operations of DSP applications.[14] Finite Impulse Response (FIR) filter is used to filter the noise/unwanted signals at finite impulse durations. Multiplication and Accumulation (MAC) unit estimates the duration of periodic impulses. Designed FIR filter compromise the parameters like low power, less area and less delay while processing the signal.

5.3. Russian Peasant Multiplication

Modified Russian Peasant Multiplication (RPM) is the best multiplication algorithms which based on "Multiply Divide" principle. In the perspective of digital implementation, left-shifters and right-shifters are used to multiply and divide '*n*' bit binary data. Hence, shifters based digital circuit is designed to perform the multiplication operation by using Russian Peasant Multiplication algorithm. Developed Digital Multiplication called as "Digital Russian Peasant Multiplier (DRPM)."

5.4. 128-Point FFT Architecture with Combined SDF-MDC Structure

To compute a 128-point FFT, Fast Fourier Transformation (FFT) is one of the digital signal processing operations used for frequency transformation. In nature, IFFT is used to analyze the frequency characteristics of discrete time domain signals. On the other hand, FFT is used to analyze the timing characteristics of discrete frequency response. The 128-point FFT architecture is designed as one of the processing element in the coarse grained architecture. This processing element operates like time to frequency signal conversion.[11] It includes Mixed Radix

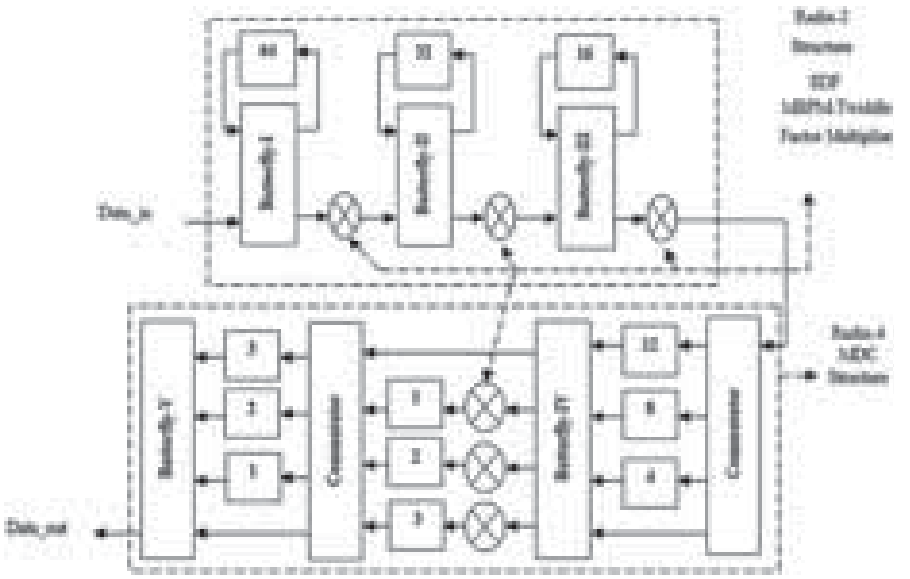


Fig. 3. Structure of 128-point new mixed radix-2 and radix-4 pipeline combined SDF-MDC FFT unit.

pipelined FFT with combined Single path delay feedback (SDF)-multipath delay commutator (MDC) structure and modified bit parallel multiplier for twiddle factor multiplication. SDF structure is used to reduce the delay in the FFT architecture.[12] Likewise, the MDC structure is used to reduce the hardware utilization and power consumption of the FFT architecture.

Figure 3 shows the combined structure of SDF and MDC. Commutator unit is used to change the position of real and imaginary input data points for optimizing or reducing the hardware complexity of the circuit. The proposed mixed radix-2 and radix-4 pipeline FFT architecture can process a continuous sequential flow of data. The proposed butterfly structures can read the both real and imaginary inputs and store both outputs in a single clock cycle. The mixed SDF-MDC structure has the less computational path and also enhances the performances of FFT processor. Developed 128-point Mixed SDF-MDC architecture has utilized to implement a single stage of Single-Path Delay Feedback, and all other stages are utilized in Multi-Path Delay Commutator.

6. CLOCK GATING TECHNIQUE

The clock signal has major sources of power dissipation because of high frequency. The clock signal is widely used for synchronization purpose; it does not compute any operations, and it does not carry any information in the logic circuits. Avoid the unnecessary clock activities in the logic circuits can be achieved by technique named as clock gating. Clock gating is used to control the power dissipation occurred by the clock net, reducing the unwanted switching on the part of the clock net by disabling the clock. It also reduces the unnecessary activities inside the gated module for saving power. Gated clock technique is easily accepted, to optimize the power and it can be applied at a system level, gate level and register transfer level (RTL). Power consumption in the circuit can be classified into dynamic power and leakage power. Leakage power is constant power according to the device used. Dynamic power is not a constant power; it will be varied according to the process of the circuit. The clock signal is the major sources of dynamic power dissipation. In synchronous circuits, the clock of some sequential elements can be disabled without affecting the logic functionality. This technique is known as clock gating. Clock gating is the most effective and widely used technique for power saving. Clock gating has been performed at different levels of abstraction and granularities. Likewise, power gating is a technique effectively reduces the active leakage power.

6.1. Architecture Level Clock Gating

This type of clock gating method is used in the system on chip design; because SoC design contains multiple processors. In this design, the clock signal of the entire processor or a particular module is disabled until it receives the request from the hardware power management unit. These are normally used while the processor is in sleep or wait for interrupt state.

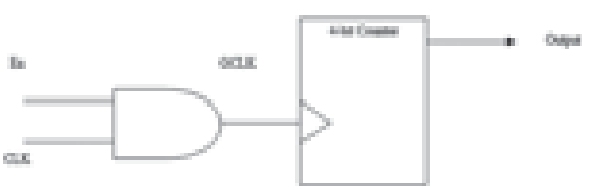


Fig. 4. AND based clock gating structure.

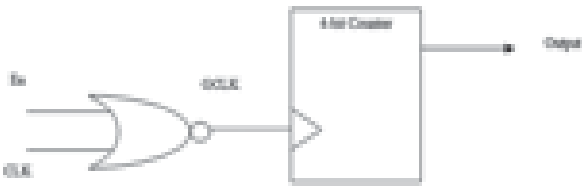


Fig. 5. NOR based clock gating structure.

6.2. Micro Architecture Level Clock Gating

This type of architecture is used for large architecture systems are composed of different logic blocks. Every block is mutually exclusive and it also not exploited at the same time. In the architecture many blocks are present in an idle condition, so the designers locate the idle blocks and include the clock gating technique for enabling or disabling the clock of these blocks.

6.3. Types of Clock Gating

There are different clock gating techniques are available in the system on chip design. These are: AND gate based technique, NOR gate based technique, latch based AND gate, latch based NOR gate, MUX based techniques are the different level of gating methods.

6.3.1. AND Gate Based Technique

AND gate based clock gating technique is widely used in the system on chip processor, because of its simple logic. Most of the case two input AND gate is used in the design. One input of the AND gate is clock signal while the second input of the AND gate in control signal. The second input fully controls the output of the AND gate. Output clock from the AND gate is used to control the sequential system circuits.

6.3.2. NOR Gate Based Technique

NOR gate based technique is suitable for clock gating, but it performs the actions only on the positive edge of the global clock. Two input NOR gate is used to generate the output signal. It generates outputs when one of the inputs is positive.

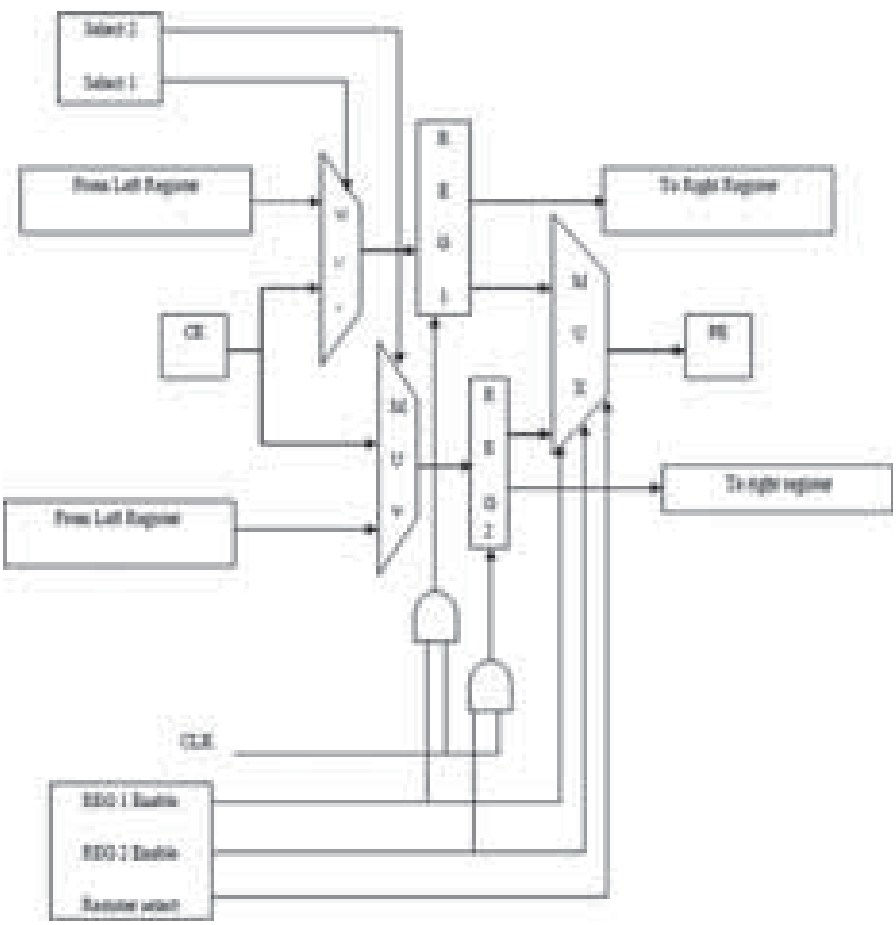


Fig. 6. Connection between CE and PE.

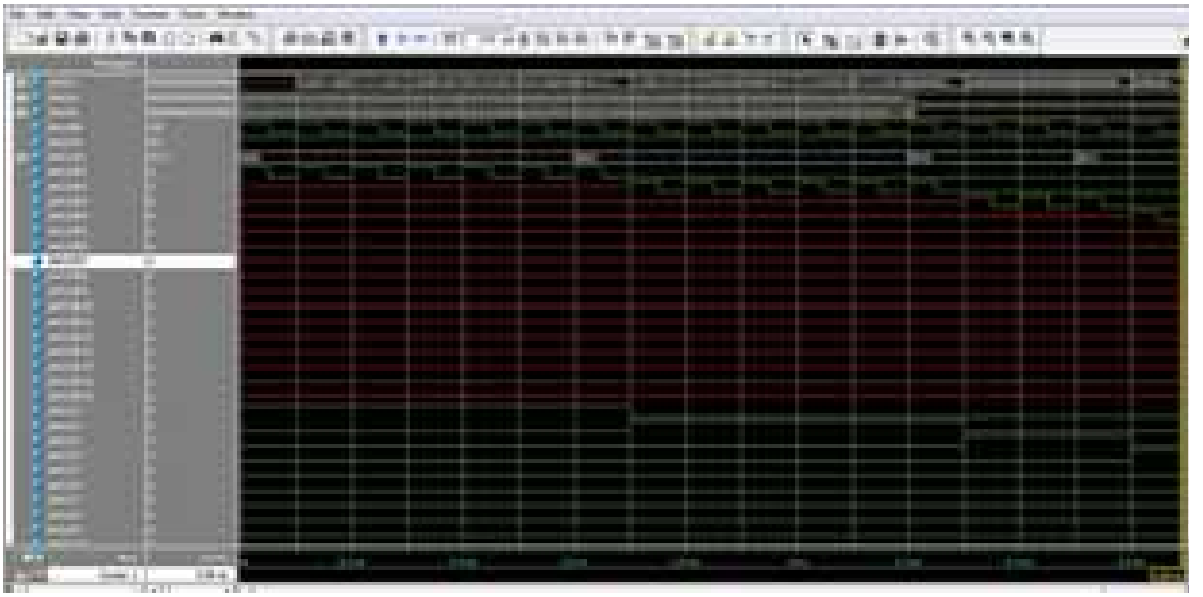


Fig. 7. Simulation result of CGRA with clock gating architecture.

7. PIPELINING REUSABLE TECHNIQUE FOR POWER REDUCTION

Coarse-grained reconfigurable architecture (CGRA) needs a lot of processing elements and a configuration memory for reconfiguration. The CGRA architecture takes more amount of power during the reconfiguration operations. Power saving is an important process in the reconfiguration approach.

Figure 6 shows the connection between the processing element and cache element. The above architecture contains multiplexer for selecting input according to the selection line. Clock pipelining reusable architecture is the technique to save the power, also avoid the unwanted power usage without any performance degradation. When mapping onto the reconfigurable architecture with loop pipelining, consider two mapping techniques: spatial mapping and temporal mapping. In the spatial mapping, each PE executes a fixed operation with a static configuration. This means that loop body is spatially mapped onto the reconfigurable array. The advantage of spatial mapping is no need of reconfiguration during execution of a loop because of the fixed functionality of each PE. However, one disadvantage is that spreading all the operations of the loop body over the limited reconfigurable array may require too many resources. In the temporal mapping, a PE executes multiple operations by changing the configuration dynamically within a loop.

8. RESULTS AND DISCUSSION

The structure of coarse grained reconfigurable architecture and fine grained reconfigurable architecture was designed by using Verilog HDL. Also, the number of processing of processing elements was done the same Verilog code. The simulation results of the CGRA system has been processed by using Modelsim XE 6.3C and those the results has been synthesized and analyzed using Xilinx ISE 10.1 and Quartus with Cyclone II device. A simulation result of the CGRA with clock gating is shown in Figure 8. Every processing element parameters are analyzed individually and shown in the table below. Table III shows the comparison results of conventional FFT and the proposed FFT. It explains the proposed FFT processing element is better in delay, area and power. Another processing element, Russian peasant multiplication and the FIR filter was compared with existing architecture, % reduction of delay, power; LUT and slices are shown in Tables IV and V.

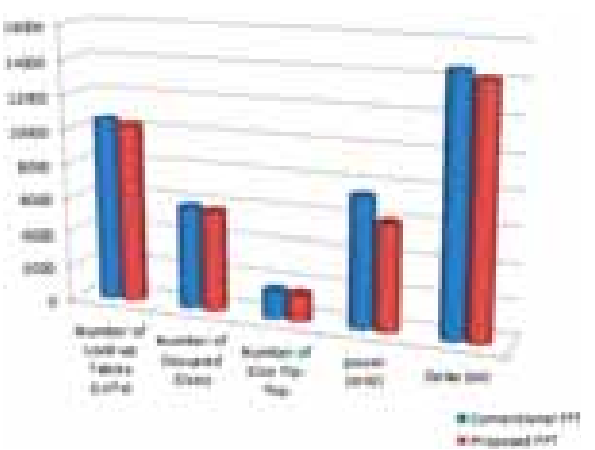


Fig. 8. Comparison graph of conventional and proposed FFT.

Table III. Comparison between conventional FFT and proposed SDF-MDC FFT.

Parameters noticed	Conventional FFT	Proposed FFT	% reduction (%)
Look up tables (LUTs)	10,542	10,317	2.1
Number of occupied slices	5,863	5,732	2.2
Number of slice flip-flops	1,590	1,554	2.3
Delay (ns)	14.625 ns	14.18 ns	3.4
Power (W)	7.438 w	6.093 w	18

Table IV. Comparison of existing multiplier with proposed multiplier.

Parameters noticed	Existing Russian peasant multiplier design	Proposed Russian peasant multiplier design	% reduction (%)
Number of LUTs	725	687	5.2
Number of occupied slices	394	376	4.5
Delay (ns)	47.848 ns	46.765 ns	2.2
Power (W)	0.321 w	0.303 w	6.5

The parameters of the architecture like logic utilization, time period and power consumption was analyzed, and the results are taken from Altera Quartus II. Comparison results of coarse grained architecture by with clock gating and without clock gating is shown in Table VI, and the graphical representation shown in Figure 10. Likewise, comparison results of conventional routing and the bidirectional routing is shown in Table VII, graphical representation of parameters difference shown in Figure 11. Different parameters are estimated for fine grained and coarse grained architecture, and the estimated results are shown in Table VIII.

Table V. Comparison between existing FIR filter and proposed FIR filter.

Parameters noticed	Existing filter	Proposed filter	% reduction (%)
Number of LUTs	89	67	24.7
Number of occupied slices	87	46	47
Number of slice flip-flops	94	35	62
Delay (ns)	4.6 ns	4.3 ns	6.5
Power (W)	0.252 W	0.219 W	13

Table VI. Comparison of coarse grained architecture by with clock gating and without clock gating.

Parameters noticed	With clock gating	Without clock gating	% reduction (%)
Total logic elements	15,056	15,568	3.2
Total registers	291	327	11
Total pins	387	389	0.5
Total thermal power dissipation (W)	33.013 W	38.293 W	13.78
Frequency (MHz)	420.17 MHz	358.55 MHz	14.6
Time period (ns)	2.3 ns	2.78 ns	17.2

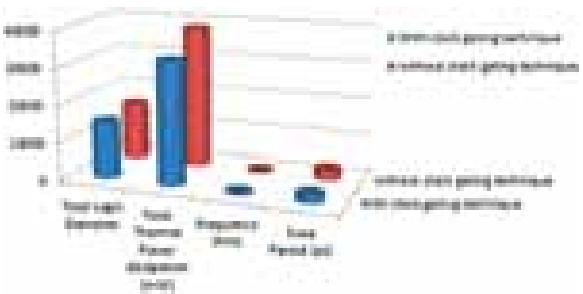


Fig. 9. Comparison graph of CGRA with and without clock gating technique.

Table VII. Parameter estimation of FGRA and CGRA.

Estimated parameters	Coarse grained reconfigurable architecture (CGRA)	Fine grained reconfigurable architecture (FGRA)
Total logic elements	14,199	14,076
Total registers	390	256
Total pins	387	387
Total thermal power dissipation (mW)	238.35 mW	238.49 Mw
Frequency (MHz)	420.17 MHz	420.17 MHz
Time period (ns)	2.380 ns	2.380 ns
Delay (ns)	48.270 ns	50.823 ns

9. CONCLUSION

In this paper, the Coarse grained and the Fine grained reconfigurable architecture were designed. The flexible, reconfigurable architecture for efficient realization is suitable for applications. In the architecture, different processing elements were designed for different applications. Power consumption is very crucial for the coarse grained and fine grained reconfigurable architecture. Unwanted power consumption by the circuit can be avoided by clock gating and pipeline reuse approach. It offers 13.78% reduction of power consumption, 3.2% reduction of logic utilization and 14.6% improvement of operating frequency compared to the existing architecture. The synthesized report of the architecture is taken from Xilinx ISE and Quartus II tools.

References

1. Y. Kim, R. N. Mahapatra, I. Park, and K. Choi, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 17, 593 **(2009)**.
2. Y. Wang, L. Liu, S. Yin, M. Zhu, P. Cao, J. Yang, and S. Wei, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 22, 983 **(2014)**.
3. D. Padole and R. Hiware, Configuration memory based dynamic coarse grained reconfigurable multicore architecture, *IEEE Region 10 Conference* **(2013)**, pp. 1–5.
4. T. J. Lin, W. Zhang, and N. K. Jha, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 22, 2607 **(2014)**.
5. D. Alnajjar, H. Konoura, Y. Ko, Y. Mitsuyama, M. Hashimoto, and T. Onoye, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 21, 2165 **(2013)**.

RESEARCH ARTICLE

6. R. Warrier, L. Hao, and W. Zhang, Reconfigurable DSP block design for dynamically reconfigurable architecture, *IEEE International Symposium on Circuits and Systems (ISCAS)* **(2014)**, pp. 2551–2554.
7. S. Yin, D. Liu, Y. Peng, L. Liu, and S. Wei, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 24, 507 **(2016)**.
8. Y. Kim, H. Joo, and S. Yoon, *IET Circuits, Devices and Systems* 10, 251 **(2016)**.
9. J. O. Filho, S. Masekowsky, T. Schweizer, and W. Rosenstiel, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 17, 1247 **(2009)**.
10. Z. Qian and M. Margala, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 24, 3008 **(2016)**.
11. S. Qiao, Y. Hei, B. Wu, and Y. Zhou, An Area and Power Efficient FFT processor for UWB Systems, IEEE **(2007)**, pp. 582–585.
12. P. S. V. Srivatsava and V. Sarada, *International Journal of Engineering Research and Technology (IJERT)* 3, 1600 **(2014)**.
13. S. Kakde, S. Khan, P. Dakhole, and S. Badwaik, Design of area and power aware reduced complexity wallace tree multiplier, *IEEE International Conference on Pervasive Computing (ICPC)* **(2015)**, pp. 1–6.
14. X. Lai and Z. Lin, *IEEE Transactions on Signal Processing* 62, 4532 **(2014)**.
15. S. Adham Hadi, *Journal of Scientific and Engineering Research* 2 **(2015)**.

Received: 24 April 2018. Accepted: 21 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1675–1682, 2018

A Secured Mechanism for Protecting the Fingerprint Template in Biometric Authentication

P. Jayapriya[1,*] and R. Manimegalai[2]

[1] *Research Scholar, Department of Information Technology, PSG College of Technology, Coimbatore 641004, Tamil Nadu, India*
[2] *Department of Information Technology, PSG College of Technology, Coimbatore 641004, Tamil Nadu, India*

RESEARCH ARTICLE

Cryptosystems that combine biometric traits for providing security are called as bio-cryptosystems. The objective of this work is to integrate biometric traits and unique features with conventional cryptography to enhance security during authentication. Unique features such as ridges, valleys, and bifurcation are extracted from the individual's fingerprint for authentication and identification in securing systems. As the features extracted from the fingerprint template are very unique and user friendly, most of the applications use fingerprint for identification. Biometric template protection and matching algorithms should be faster in identifying and authenticating users. In biometrics applications, when a fingerprint is entered into the system, only a "*template*" of the fingerprint, not the image of the fingerprint is stored. The fingerprint template can be generated using both symmetric and asymmetric algorithms and stored in the database. The template should be protected from unauthorized users and invaders. Binary values representing salient features of biometrics can also be used along with an encryption algorithm to generate the key into applications. This work proposes a secured mechanism which combines the fingerprint biometric with encryption to generate the key. The proposed technique reduces the time to retrieve the template from database when compared to existing symmetric algorithms.

Keywords: Fingerprint Template, Encryption, Decryption, RSA Algorithm.

1. INTRODUCTION

Biometric is a unique measure of every individual. Biometric characteristics are used for recognizing and verifying the identity of a human being. Biometric measures are integrated with cryptography to protect information from attackers. The widely used biometrics for security are fingerprint, hand, eye, face, and voice.[1] Biometrics can be combined with other technologies such as smart cards, encryption keys, and digital signatures in order to provide better security features during authentication.[2] Biometrics combined with cryptography techniques are increasingly used for secured transactions and personal data privacy and are called as Bio-crypto systems. They are mainly required at federal and state governments, and e-commerce applications. Biometric-based authentication is useful in workstation, network and domain access, single sign-on, application logon, data protection, remote access to resources, transaction security and web security. There are two strategies followed in bio-cryptosystems: (i) Protecting the cryptographic key using biometric template, and, (ii) generating cryptographic key using biometric traits. Bio-cryptosystems can be classified into three different models are (i) Key release bio-cryptosystems: In this model, the key is released when the biometric template is successfully matched with query template. (ii) Key binding bio-cryptosystems: In this model, the biometric template is protected by binding it within a cryptographic frame work. (iii) Key generation bio-cryptosystems: In this model, a key is directly derived from the template and stored in the database. Here the biometric template is not stored.[3]

1.1. Fingerprint Characteristics

A fingerprint is a distinct pattern of ridges and valleys on the finger surface of an individual. A ridge is defined to be a single curved segment, whereas, a valley is an area between two adjacent ridges. So the dark areas of the fingerprint are called ridges and white area that exists between them is called as a valley. Fingerprint recognition looks for unique patterns of ridges and valleys that are present in an individual's fingerprint. These patterns are unique to every individual and thus help to identify individuals from an entire population. Fingerprints

*Author to whom correspondence should be addressed.

are inherent to individuals and can neither be lost nor stolen which makes it highly accurate and reliable. Moreover, the availability of low-cost fingerprint readers coupled with easy integration capabilities has led to the widespread deployment of fingerprint biometrics in all types of organizations.[6]

2. LITERATURE SURVEY

Generating cryptography key using biometric is more secured, it can't be stolen by an unauthorized person and there is no need to remember the password. Three factors that need to be considered for generating the key are key entropy, key uniqueness and, key stability.

Key entropy ensures that the strength of the biometric key is stronger than the key generated by cryptography. Biometric features will make the key unique. The template generated during enrollment phase may vary with query template. The variation occurs due to illumination, orientation and some physical changes in biometric trait used. Various techniques such as a fuzzy vault, fuzzy commitments are used for generating stable keys in Ref. [3].

Various methods for key generation and key release are available in the literature. Soutar et al. have proposed a key-release algorithm in which cryptography key is stored as part of the user's database in Ref. [4]. Later Roginsky has proposed that user authentication and key release can be completely decoupled in Ref. [6]. The problems in the above mentioned methods are, there is no way to ensure who produced the key and also the biometric template can be easily spoofed by the attackers. There is a possibility that the key may be stolen by the spoofers during enrollment process before the key is stored in the database. Key generation technique avoids some problems that occur in the key release techniques. In key generation binding the secret key to the biometric information and there is no need to store the biometric template. Clancy has proposed a key generation algorithm for accessing the biometric template in Ref. [5]. Russell et al. have proposed a method of transformation and matching algorithm for fingerprint using cancelable biometric which give better performance in Ref. [7]. Rathgeb and Uhl have created biometric keys from the binary iris code. Fuzzy biometric data is used in Ref. [8]. Hao et al. have proposed the 140 bits of key generation from the iris. They have developed two-way error correction technique that merges Hadamard and Reed-Solomon codes.[10]

3. PREPROCESSING STEPS INVOLVED IN FINGERPRINT RECOGNITION

3.1. Preprocessing

The images with a low level of abstraction are improved through preprocessing technique. The aim of the preprocessing is to enhance features of the image for further processing. The image enhancement is the process of adjusting digital images so that the results are suitable for further image analysis. The removal of noise, sharpen or brighten the image, making it easier to identify key features. Various stages of Image pre-processing techniques are shown in Figure 1.

After image acquisition, the image is normalized so that it has a pre-specified mean and variance. The orientation of the image is estimated from the normalized input image. The frequency image estimation is computed from the normalized input image and estimated orientation image. The histogram equalization is a technique for adjusting image intensities to enhance the contrast of an image. Filtering enhances the image by using smoothing, sharpening, and edge enhancement. Figure 2 shows the minutiae points of a fingerprint image, crossovers are formed when two ridges cross each other. The core is center of the fingerprint template. Bifurcation is the point where a single ridge branches out into two or more ridges. Ridge ending is the point where the ridge ends suddenly. Ridge dots are very small ridges. Islands are slightly longer than dots and occupy a middle space between two diverging ridges. Ponds or Lakes are the empty space between two diverging ridges. Spurs are a notch protruding from a ridge. Bridges are the small ridges that join two longer adjacent ridges. Ridge endings and ridge bifurcations are the most commonly used minutiae type since all other types of minutiae are based on a combination of these two types.[10]

3.2. Minutiae Extraction

Minutiae points are the major features of a fingerprint image and are used in the matching of fingerprints. These minutiae points are used to determine the uniqueness of

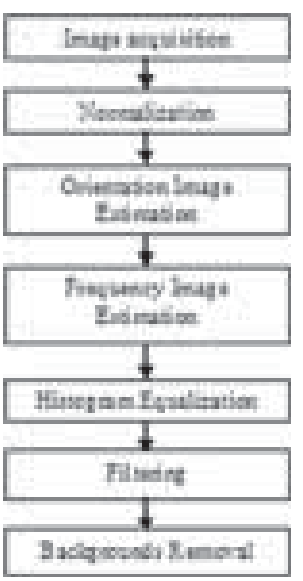


Fig. 1. Steps involved in image pre-processing.

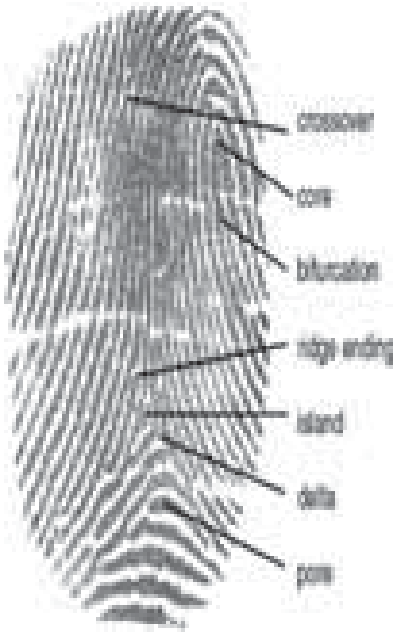


Fig. 2. Minutiae points.

a fingerprint image. A good quality fingerprint image can have 25 to 80 minutiae depending on the fingerprint scanner resolution and the placement of the finger on the sensor. Minutiae can be defined as the points where the ridge lines end or fork. So the minutiae points are the local ridge discontinuities and can be of many types.

3.3. Minutiae Extraction Techniques

A fingerprint is produced when the finger is pressed against the smooth surface of the sensor device. The process of extracting various evident characteristics in a fingerprint is called feature extraction. Various methodologies for minutiae extraction are shown in Figure 2.

3.3.1. Binary Fingerprints

Fingerprint binarization is the process of converting an 8-bit gray-scale fingerprint image into a 1-bit ridge image. This is virtually equivalent to thresholding. Post-processing of binarized images, such as smoothing, is also important because there may loss of information in the image. Fingerprint image binarization, as a critical part of fingerprint image preprocessing, is the precondition of fingerprint image. In order to utilize the texture characteristics of the fingerprint image effectively, the orientation information of fingerprint image is considered. Then, adaptive local threshold binarization and post-processing algorithms are combined with the orientation information. The contrastive experimental results indicate that the presented algorithm can achieve better binarization effects than the adaptive binarization algorithm. Moreover, the algorithm has a good capability of repairing the disconnected line, eliminating the holes and removing crossed lines in the fingerprint image to a certain extent, which has important meaning for improving the minutiae extraction precision.

3.3.1.1. Unthinning Binary Images. There are three methods of minutiae extraction from unthinned binarized images. (i) Chain code processing, (ii) run-based methods, (iii) ridge flow and Local pixel analysis based methods. The chain code processing technique is based on the chain-code representation of pixel image and it can be recovered fully from the chaincode of its contour. Run based method is on the horizontal and vertical run-length encoding from binary images. This process results in fast minutiae extraction without requiring a computationally expensive thinning process. Ridge flow and local pixel analysis is a square based method to extract minutiae from unthinned binary images in a 3×3 square mask created around each pixel in the fingerprint image and the average of pixels is computed.

3.3.1.2. Thinning Binary Images. Thinning of binary images is commonly used to tidy up the output of edge detectors by reducing all lines to single pixel thickness. Thinning is normally applied to binary images, and produces another binary image as output. Thinning algorithm is a morphological operation that is used to remove selected foreground pixels from binary images. It preserves the topology of the original region while throwing away most of the original foreground pixels. Figure 3 shows the result of a thinning operation on a simple binary image.

Crossing number based

Crossing number based is the most widely used method of minutiae extraction for thinned binarized images category. It is preferred over other methods because of its computational efficiency and intrinsic simplicity.

Mathematical Morphology

Mathematical Morphology is the analysis of signals in terms of shape. This simply means that morphology works

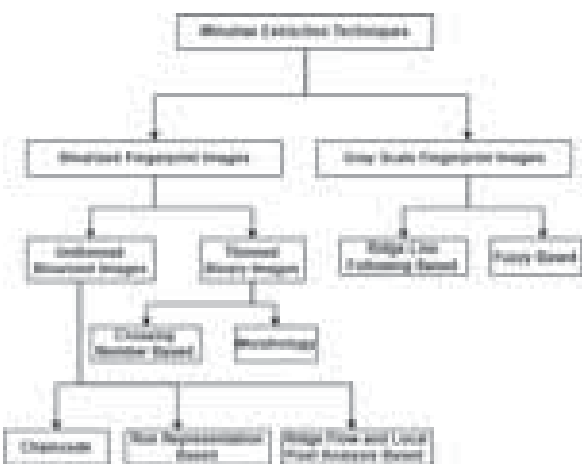


Fig. 3. Minutiae extraction techniques.

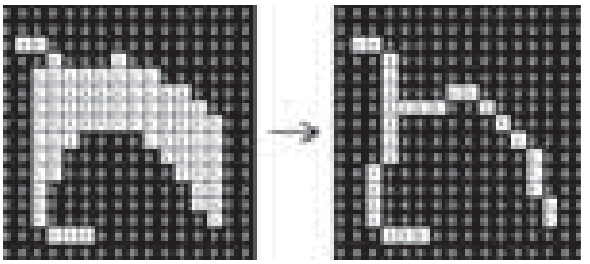


Fig. 4. Image before and after thinning.

by changing the shape of objects contained within the signal. Mathematical morphology was developed in 1970's by G. Matheron and J. Serra. Morphology has several advantages over other techniques, especially when applied to image processing. It preserves edge information and works by using shape-based processing. It can be designed to be idempotent and computationally efficient. Thinning image on a simple binary image is shown in Figure 4. It is commonly used to tidy up the output of edge detectors by reducing all lines to single pixel thickness. Thinning is normally applied to binary images to improve the recognition performance.

3.3.2. Grayscale Fingerprint Images

Some of the disadvantages for binarization and thinning are: (i) it is time consuming, (ii) information may be lost during binarization (iii) not suitable for low-quality images. Many methods to extract minutiae from grayscale fingerprint images without binarization and thinning. The methods are as follows: In a gray-level fingerprint image, ridges and valleys in a local neighborhood form a sinusoidal-shaped plane wave which has a well-defined frequency and orientation. Ridgeline following based techniques directly extracts the minutiae from the grayscale image by following the ridge flow lines with the help of local orientation field. A ridgeline is defined as a set of points that are local maxima along one direction. The ridgeline extraction algorithm tries to locate the local maximum relative to a section orthogonal to the ridge direction. In Fuzzy based technique extraction techniques a grayscale image is observed to have two distinct levels of gray pixels. One such level is formed by darker pixels and constitutes the ridges. The other level is formed by the lighter pixels and constitutes the valleys and furrows.

4. MODIFIED RSA FOR FINGERPRINT KEY GENERATION

4.1. Modified RSA

RSA is a cryptosystem which is used to protect by encrypting sensitive data. RSA is well known public-key cryptography. It is an asymmetric cryptography, uses two different mathematically linked keys, public and private keys. The private key is kept secret and public key is shared with everyone. One key is used to encrypt and other is used to decrypt data. RSA is a powerful asymmetric algorithm. It assures authenticity and confidentiality of data storage. RSA is based on the fact that larger multiplication is difficult to factorize. Biometric binary data is integrated with Modified RSA for key generation.

4.2. Algorithm Modified RSA-Key Generation ( )

1. Generate a pair of large distinct prime numbers p and q
2. Compute the modulus n as $n = p * q$.
3. Select a continuous public exponent e between 3 and $n-1$ that is relatively prime top-1 to $q-1$.
4. Compute the private exponent d from e, p, and q.
5. Output (n, d) as the private key and (n, e) as the public key.
6. Encryption
7. Encrypt $(m) = c = m^e \pmod{n}$
8. The message m is input, the resulting message is ciphertext c. Encryption operation in RSA is exponentiation to the eth power.
9. Decryption
10. Decrypt $(c) = m = c^d \pmod{n}$.

The ciphertext c is input, the resulting original message is revoked. Decryption operation in RSA is exponentiation to the dth power. The e and d exponent relationship ensures that decryption and encryption are inverses.

4.3. Work Flow

The proposed methodology improves Enhanced Encryption for Fingerprint Protection. The input Fingerprint image is taken in .tif format. The given image is preprocessed, it includes binarization and thinning. Thinning

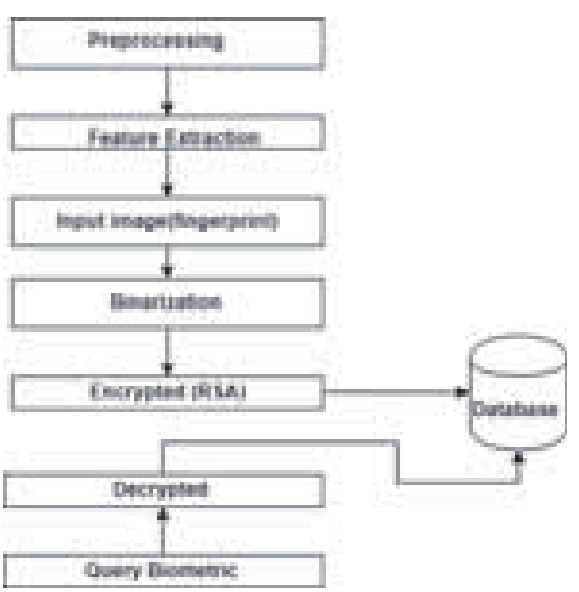


Fig. 5. Block diagram of the proposed methodology.

refers to the process of removing unnecessary pixels in the input image. Minutiae are extracted from the thinned image where Minutiae points are the major features of a fingerprint image and are used in the matching of fingerprints. Ridge endings and ridge bifurcations are the most commonly used minutiae types which are extracted from the given fingerprint image. The resultant image is converted into the decimal using ID conversion after the transpose of the minutiae extracted image. The obtained decimal string is converted into a binary string using MATLAB function such as dec2bin. This string is written into a text file. The binary string is read from the file and given to the RSA algorithm for encryption. In the implementation of RSA Algorithm two inputs, such as p and q are given manually, where p and q are prime numbers. The value of e and d are calculated. Then the binary string is converted into Ciphertext.

This Cipher text is considered as the key for security improvement.

5. IMPLEMENTATION

The fingerprint acquired from the sensor and is given as input the mat lab software which is support for windows. Other than mat lab, numpy Scilab and the R language are some other software languages for implementing the image processing coding.

5.1. Experimental Results

The input fingerprint template is stored in .tif format. The digital image is binarized which is referred to as the process of converting the image in the black and white image. The input binarized image is processed to remove the unnecessary pixels. This process is known as thinning. Hence the preprocessing of the given fingerprint image is done (Fig. 7).

The fingerprint acquired from the sensor and is given as input the mat lab software which is support by windows. Other than mat lab, numpy Scilab and the R language are some other feature extraction techniques.

Minutiae points refer to the unique features in the fingerprint of human's. This helps in identifying the individuality of the person (Fig. 8 Minutiae Extraction). The minutiae are converted into a binary string. The binary string is written into a text file. The RSA algorithm reads the binary string from the text file and encrypts the same (Fig. 9 Encryption).

5.2. Results

The symmetric cryptography a same secret key is shared by both sender and receiver. (i) The DES is a symmetric cryptographic algorithm used for encryption and decryption of the message. In DES only one secret key is used for both encryption and decryption of the message. The key size of DES is 56-bit. The key size of 3DES is 168 bit. It is slower than DES. It performs operation three times on each block of data. (ii) AES is different from DES and 3DES due to it has variable key sizes such as 128,192,256 bits. In Asymmetric cryptography, in RSA two keys are used one is public key and second is a private key. The public key is openly accessible to everyone in the

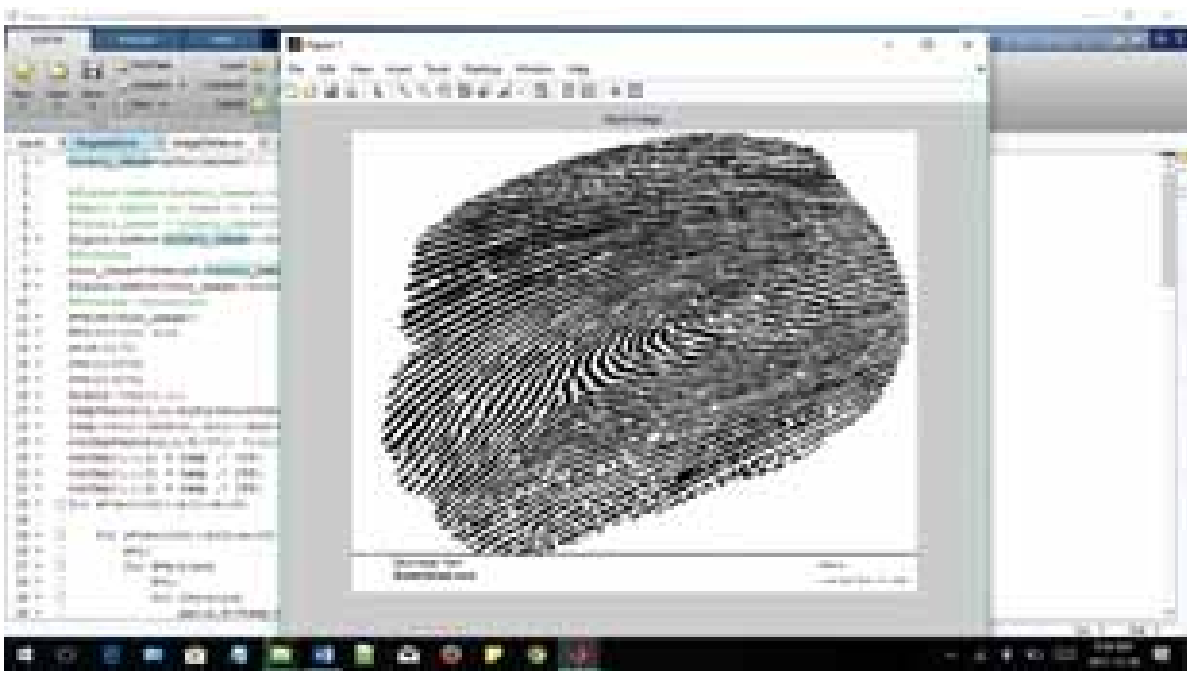


Fig. 6. Input image of a fingerprint.

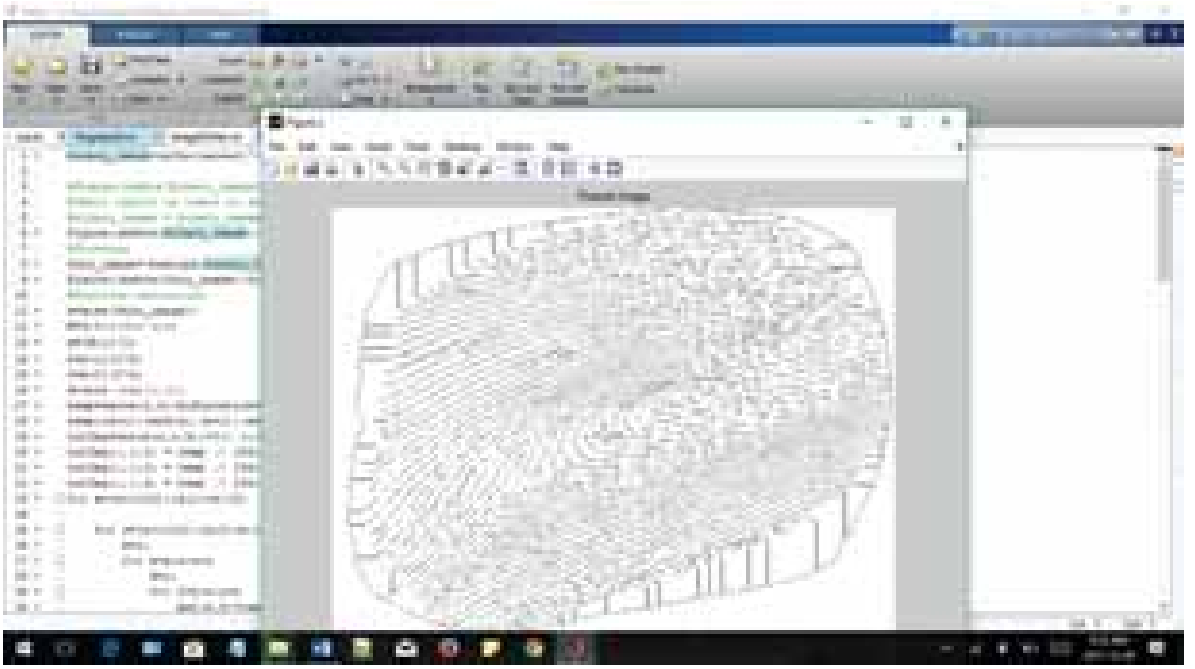


Fig. 7. Thinned image of a fingerprint.

cryptosystems and the private key is kept confidentially. ElGamal is an alternative for RSA asymmetric key encryption algorithm that is based on the Diffie-Helman key exchange.[21]

The key size with their generation time is given in Table I. For a key generation the binary bits are generated by the integer input in cryptography algorithms. Here the binary value of biometric fingerprint template is used for a

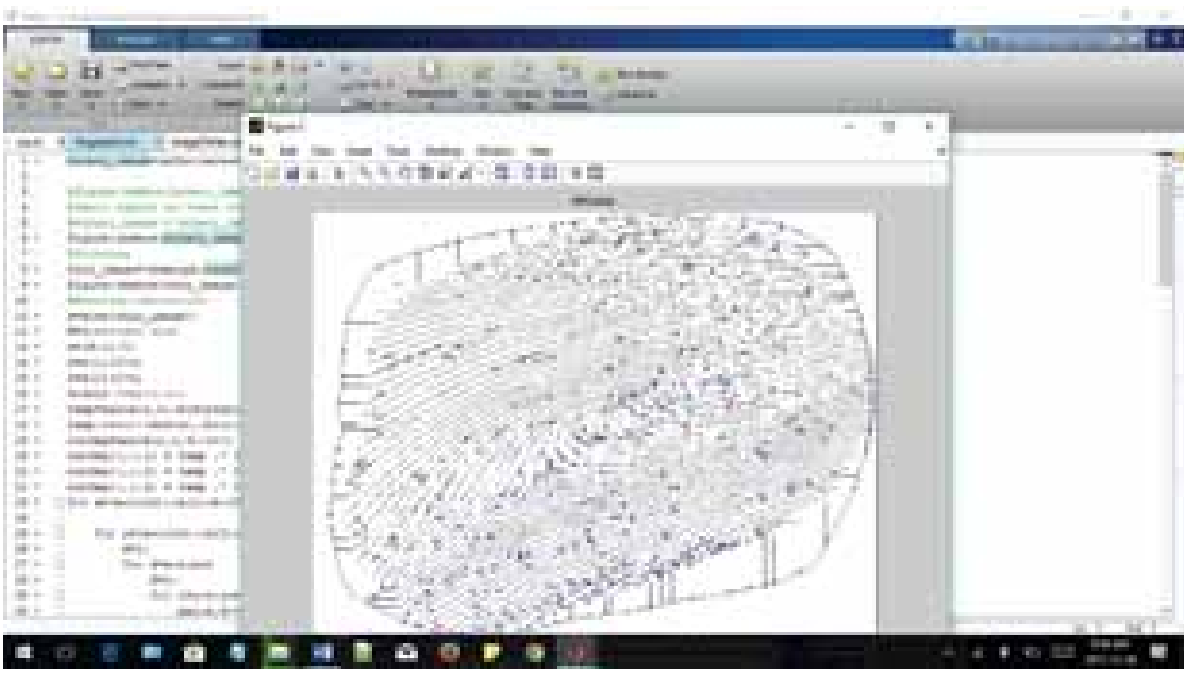


Fig. 8. Minutiae extraction of a fingerprint.

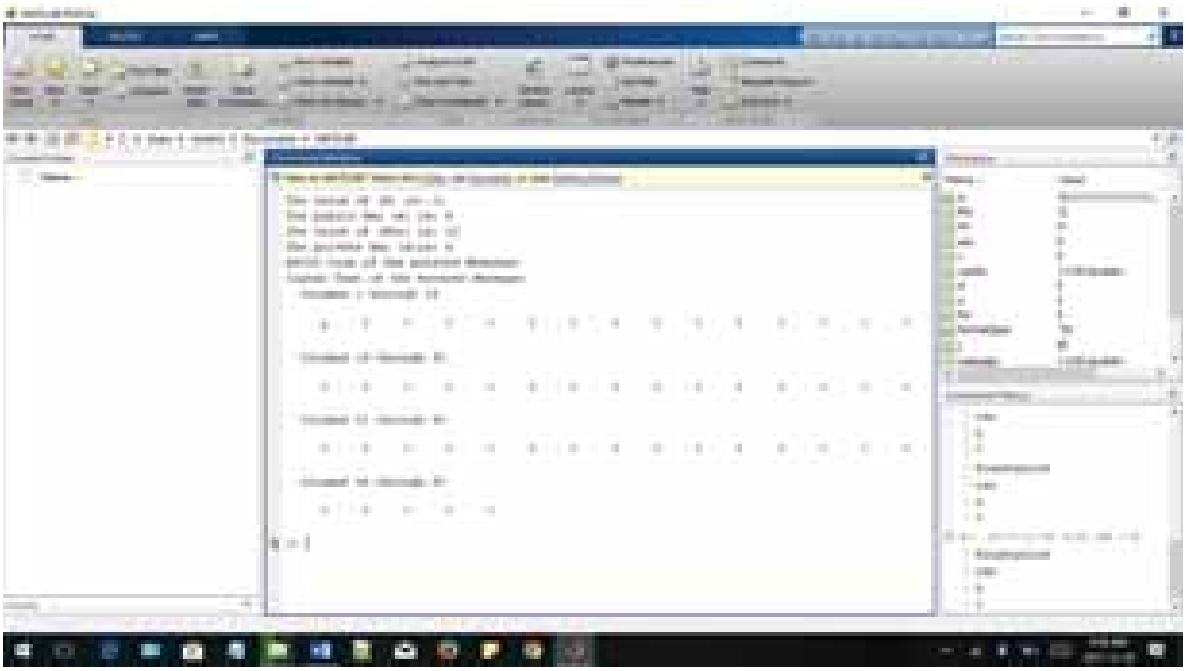


Fig. 9. Encrypted fingerprints.

key generation which is strongly authenticated. The security level of the proposed bio-cryptosystems was designed using MATLAB 2014. RSA algorithm is safe and secure for its users through the use of complex mathematics. RSA algorithm is hard to crack since it involves factorization of prime numbers which are difficult to factorize. Moreover, RSA algorithm uses the public key to encrypt data and the key is known to everyone, therefore, it is easy to share the public key. The proposed algorithm is compared with the performance of other various symmetric algorithms and the results are listed in Table II.

The performance analysis shows that the performance of the modified RSA algorithm is very high when compared with 3-DES and AES algorithms.

Table I. Key size with their generation time.

Cryptographic algorithms	Key size (bits)	Generation time (milliseconds)
Symmetric		
DES	56	29
AES	128	75
Asymmetric		
RSA	1024	287
ElGamal	160	86

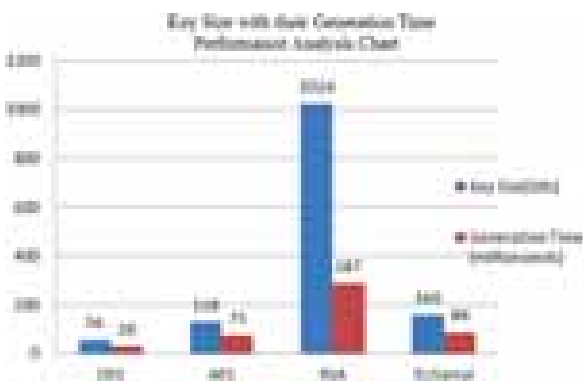
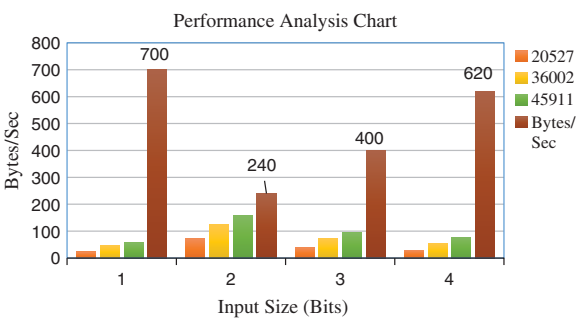


Table II. Performance analysis.

Input size	DES	3DES	AES	MRSA
20527	24	72	39	30
36002	48	123	74	56
45911	57	158	94	78
Bytes/Sec	700	240	400	620



6. CONCLUSION

The proposed method designing using fingerprint template which protects the template using RSA has been implemented using MATLAB. The performance of the RSA algorithm is comparatively improved the time required for key generation speed and highly secured the protection of the template. Biometrics plays an important role in encryption of messages in our day to day life. The usage of fingerprints as a biometric measure is a predominant one. The above system does the encryption of fingerprint image which is least prone to attack. As RSA algorithm is used for encryption the encrypted fingerprint image can't be found easily. This encryption can be applied at the login module of a web application where the password is the fingerprint of the user. Thus the password can't be hacked by the attackers. While the ECC-free key binding scheme is still a new direction to study, it is believed that the proposed scheme has wide room to improve security, privacy, and accuracy performance aspects in the future.

References

1. M. Figarella, *Biometric Technology Today* 2017, 5 **(2017)**.
2. M. Fairhurst, C. Li, and M. Da Costa-Abreu, *IET Biometrics* 6, 11 **(2017)**.
3. A. K. Jain, K. Nandhakumar, and A. Nagar, *Journal on Advances in Signal Processing* 2008, 17 **(2007)**.
4. C. Soutar, D. Roberge, S. A. Stojanov, R. Gilroy, and B. V. K. Vijay Kumar, *Proceedings of the SPIE-Optical Security and Counterfeit Deterrence Techniques II* 3314, 178 **(1998)**.
5. T. C. Clancy, N. Kiyavash, and D. J. Lin, Secure smartcard-based fingerprint authentication, *Proceedings ACM SIGMM Multimedia, Biometrics Methods, and Workshop* **(2003)**, pp. 45–52.
6. A. Roginsky, A New Method for Generating RSA Keys, International Business Machines Consulting Group, Springer **(2004)**.
7. R. Ang, R. Safavi-Naini, and L. McAven, Cancellable key-based fingerprint templates, *Australasian Conference on Information Security and Privacy* **(2005)**, pp. 242–252.
8. C. Rathgeb and A. Uhl, *IET Computer Vision* 5, 389 **(2011)**.
9. G. I. Davida, Y. Frankel, and B. J. Matt, On enabling secure applications through offline biometric identification, *Proceedings of the IEEE Privacy and Security* **(1998)**, pp. 148–157.
10. F. Hao, R. Anderson, and J. Daugman, *IEEE Transactions on Computers* 55, 1081 **(2006)**.
11. Priyanka Rani and Pinki Sharma, *Proceedings of International Journal of Advanced Research in Computer Science and Technology* 2 **(2014)**.
12. Manisha Redhu and S. Balkishan, *Proceedings of International Journal of Engineering Research and Applications* 3, 2488 **(2013)**.
13. W. Stallings, Cryptography and Network Security, Principles and Practices, Fourth edn., Pearson, Prentice Hall **(2006)**.
14. D. Gollmann, Computer Security, Third edition, Wiley **(2013)**.
15. G. Zhen, W. Li, and C. Zhan, Cryptographic key generation from biometric data using lattice mapping, *Proceedings of the 18th International Conference on Pattern Recognition* **(2006)**, Vol. 4, pp. 513–516.
16. C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B. V. K. Vijaya Kumar, *Proceedings of Optical Pattern Recognition IX* 3386 **(1998)**.
17. H. C. Lee and R. E. Gaensslen (eds.), Advances in Fingerprint Technology, CRC Press, Elsevier, New York **(2001)**.
18. Ravi Subban and Dattatreya P. Mankame, A Study of Biometric Approach Using Fingerprint Recognition, Lecture Notes on Software Engineering, May **(2013)**, Vol. 1.
19. U. U. Manikdurge and A. M. Shah, *International Journal of Science, Engineering and Technology Research* 5 **(2016)**.
20. J.-P. Aumasson, *Computer Fraud and Security* 6 **(2017)**.
21. Om Preeti Chaurasia, *I.J. Image, Graphics and Signal Processing* 6, 29 **(2012)**.
22. Faiqa Maqsood, Muhammad Ahmed, Muhammad Mumtaz Ali, and Munam Ali Shah, *International Journal of Advanced Computer Science and Applications* 8 **(2017)**.

Received: 24 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1683–1687, 2018

Missing Data Prediction in Micro-Array Using Enhanced Regression Method

M. Karthiga[1,*], S. Sankarananth[2], S. Nevetha[1], and R. R. Asmidha[1]

[1] *Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Erode 638401, India*
[2] *Department of Electrical and Electronics Engineering, Excel College of Engg and Technology, Erode 637303, India*

The most intricate problem in data mining is data missing. Health care industries role out with huge number of data and preserving those data without any loss is a big problem in present scenario. Transferring large amount of health care data from one location to other location might cause data loss. These loss in data result in missing one or more attributes from the database table. These missed out data are predicted to get the approximate results. But accuracy is not guaranteed in case of prediction. So a multidimensional array model with modified advanced regression is implemented to predict the missing data. From the available dataset, two databases are created for normal patients and cancer patients. The reformed progressive regression will equate the existing spatial database with the normal database from the input database. So the result will be obtained from the dataset. Whether the patient will affect from cancer or not, also their infection ration percentage can be find out, along with the missing values in the database during the time of data migration. An enhanced progressive regression was introduced, to determine the missing values in the dataset. Bayesian Principal Component Analysis (BPCA) obtains the lowest normalized root-mean-square error on 82.14% of all missing rates. The accuracy rate of our proposed algorithm is 95.79%.

Keywords: Database, Missing Data, Micro Array, Advanced Regression, BPCA, Spatial Database.

RESEARCH ARTICLE

1. INTRODUCTION

Now-a-days, the volume of data is growing larger day by day. Query based data retrieval is used to extract the useful information from the database.[1] But in case of huge number of data, query tools are not as much efficient in retrieving information from the database. Several data mining techniques are available for extracting the relevant data from the database in an efficient manner.[2] These techniques concentrate on specific relevant records to retrieve knowledge from the database. In our proposed method, database for cancer patients is collected. The readings are noted down from the blood test samples of the infected patients. The data set for about 500 cancer patients is collected from Sri Ramakrishna Hospital, Coimbatore.

2. RELATED WORKS

The data set of Microarray gene expression normally contains missing values in nature. These missing data would result in negative result analysis. In-order to overcome these missing data, lot of algorithms has been proposed to ascribe the missed out values.[6] Ignoring or deleting directly the genes with missing values results in huge loss of valuable information. Consequent data analysis may also provide negative results due to this missing data. The pure microarray data without any missing data will definitely help the researchers in predicting the disease more accurately. To overcome such issues lot of imputation algorithms like Bayesian principle, kNN, SVD and local lease squares (LSS) approach have been proposed. Though these algorithms performs good on handling missing data, they also have certain limitations. In our proposed methodology, micro gene expression for cancer disease is taken and the missing values are predicted using hybrid imputation algorithm.

3. EXISTING METHOD

Decision tree classification algorithm has been used in the existing methodology. This classifier divides the instance space using hyper planes and the classifier is built from a root. The root represents an attribute and the root is divided based on the values of the attribute. Each new iteration divides the tree into new division until an end criterion is achieved. The terminal leaf nodes are assigned with class labels representing the classification outcome.

*Author to whom correspondence should be addressed.

End criterion should be chosen with great care in order to avoid over fitting of huge trees and under fitting of small trees. Pruning method helps to avoid over fitting of trees. A new instance is directed from the root of the decision tree down to a leaf that best fits the classification outcome. Decision tree methodology is efficient in producing results, but they are unstable as the increase in training datasets result in abnormal values.

The below formula shows the Entropy using the frequency table of one attribute

$$E(S) = \sum_{i=1}^{c} -p_i \log_2 p_i \quad (1)$$

This existing system deals with decision trees in data mining based missing data prediction. This is the most common method. Decision tree learning with prediction is also commonly used in data mining. There is no goal to create a model that predicts the value of a target variable based on several input variables.

4. PROPOSED WORK

The proposed work is that we have modified the id3 algorithm using decision tree classification method and included the pre processing steps for the cancer data set to improve the accuracy of the classifier. In Figure 2, the data set has *x* values missing in it. In Figure 3, using the pre processing steps of the data set, we have resolved it. Also the data set has data conflicts in it. And we have proposed an approach to resolve it. The concept of conditional entropy measure is used in the id3 algorithm and modified it. Then after pre processing the data set, it is supplied to the modified algorithm which constructs the decision tree and thus it proves to increase the accuracy of the classifier.

In Figure 1, a portal has been created where any patient can check their own reports through online and predict their occurrence of cancer themselves.

Attribute Selection in multi array model

Given a collection S of c outcomes in multi array

$$\text{Entropy(S)} = \text{S} - \text{p(I)log2p(I)}$$

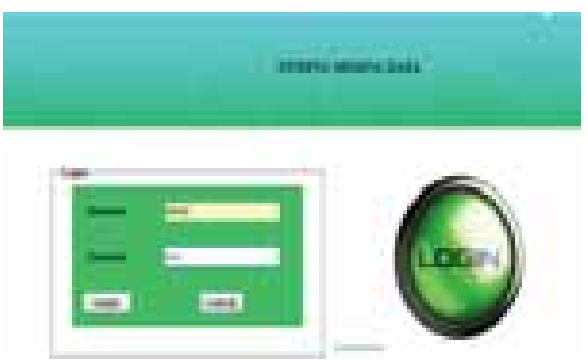


Fig. 1. A portal for uploading the reports and predicting the occurrence of cancer.

where p(I) is the proportion of S belonging to class I. S is over c. Log2 is log base 2.

Note that S is not an attribute but the entire sample set.

If S is a collection of 14 examples with 9 YES and 5 NO examples, then

$$\text{Entropy(S)} = -(9/14)\text{Log2}(9/14) - (5/14) \times \text{Log2}(5/14) = 0.940$$

Notice entropy is 0 if all members of S belong to the same class (the data is perfectly classified). The range of entropy is 0 ("perfectly classified") to 1 ("totally random").

Gain(S, A) is information gain of example set S on attribute A is defined as

$$\text{Gain(S, A)} = \text{Entropy(S)} - \text{S}((|\text{S}_\text{v}|/|\text{S}|) * \text{Entropy}(\text{S}_\text{v}))$$

Where: S is each value v of all possible values of attribute A.

S_v = subset of S for which attribute A has value v.

$|S_v|$ = number of elements in S_v.

Algorithm Steps

Step 1: Begin the process;

Step 2: Let X be the input dataset (raw data set with missing data);

Step 3: Covert excel into SQL format;

Step 4: String Excel = Server.MapPath("missing data/" +FileUpload1.FileName.ToString()); (Conversion Process)

Step 5: String connect = "Provider = Microsoft.Jet .OLEDB.4.0; Data Source =" + Excel;

Step 6: Consider all the data as a data Array;

$$= \begin{pmatrix} B_{1,1} & B_{1,2} & \dots & B_{1,p} & A_{1,1} & A_{1,2} & \dots & A_{1,n-p} \\ B_{2,1} & B_{2,2} & \dots & B_{2,p} & A_{2,1} & A_{2,2} & \dots & A_{2,n-p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ B_{k,1} & B_{k,2} & \dots & B_{k,p} & A_{k,1} & A_{k,2} & \dots & A_{k,n-p} \end{pmatrix}$$

Step 7: $B_{1,1}$ will be the initial node and $A_{k,n-p}$ will the end of the row;

Step 8: Array will allow search for row and Column;

Step 9: Finding Null value in the array;

Step 10: {arr.Add(rd[0].ToString());}: Array execution will done;

Step 11: for (int $i = 0$; $i < n$; $i++$): Loop will find all missing value for column and row;

Step 12: arr[i].ToString() +: String will verify the value;

Step 13: Let nIG = missing.Count;

Step 14: if (missing[j] == "") {numofnullIG = numofnullIG + 1;}: finding index for null value

Step 15: for (int $i = 0$; $i <$ nIG; $i++$): if (missing[i] == "") and inde = i; Fixing X value

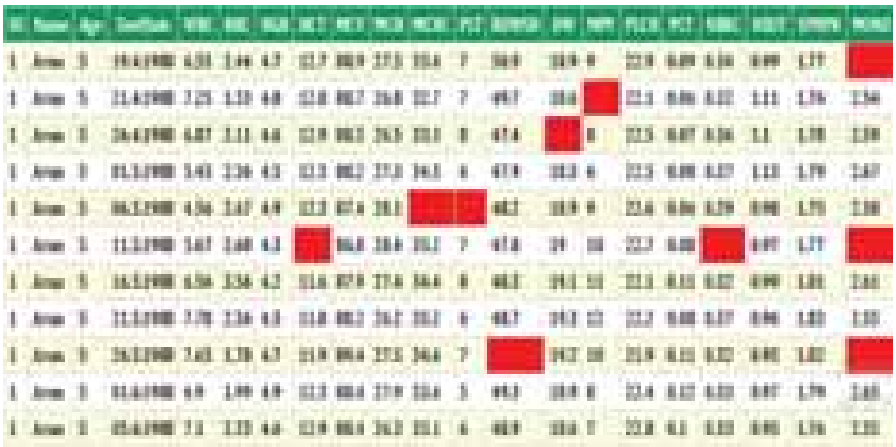


Fig. 2. Prediction of missing values in the dataset.

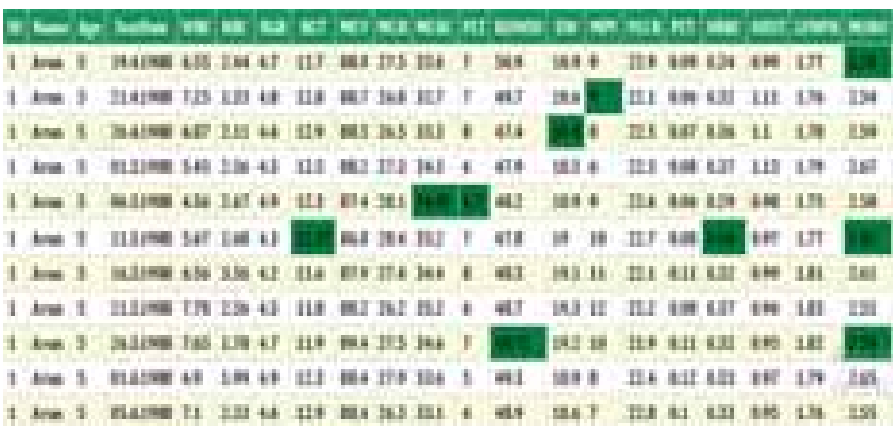


Fig. 3. Fixing of missing values in the data set.

Step 16: val = val + Convert.ToDouble(missing[i].ToString());
Fine tune the result
Step 17: finalvalue = nIG − numofnullIG; finding min and max value
Step 18: finalvalue = val/finalvalue;
Step 19: finalvalue = Math.Round(finalvalue, 2);
Step 20: Fix value from $B_{1,1}$ to $A_{k,n-p}$ will the end of fixing
Step 21: Repeat the process
Step 22: Stop.

5. RESULTS AND DISCUSSION

Figure 4 is the initial result from the data set. This result shows the occurrence of cancer with day to day percentage calculation of the patients.

From the above given results, the accuracy of the prediction has been improved and the result has been generated in a very short period of time while comparing with the existing system as depicted in the Figures 5–8.

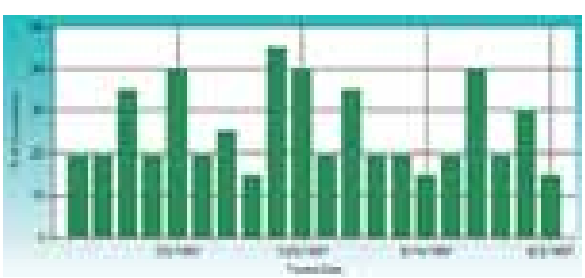


Fig. 4. Percentage of occurrence of cancer.

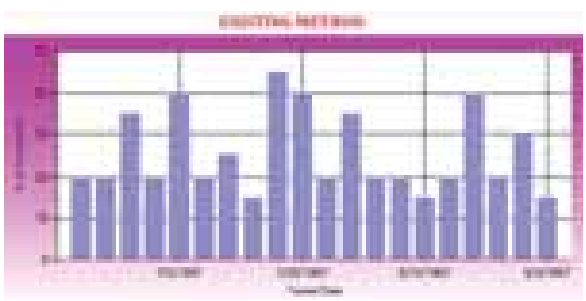


Fig. 5. Percentage of cancer occurrence using existing method.

Tested Date	Percentage of Occurrence	CPU Time
6/19/1997	17	8
6/23/1997	17	8
6/27/1997	12	7
7/1/1997	17	8
7/5/1997	17	8
7/9/1997	17	8
7/13/1997	12	7
7/17/1997	12	8
7/21/1997	42	8
7/25/1997	17	8
7/29/1997	17	8

Fig. 6. Efficiency shown in existing method.

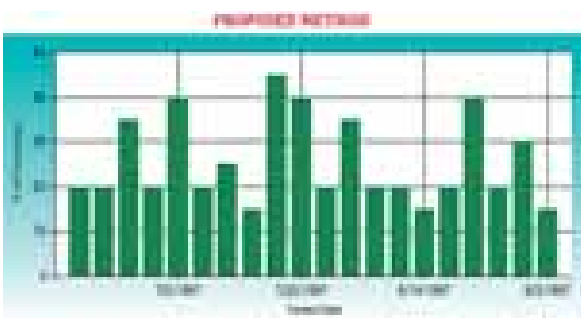


Fig. 7. Percentage of cancer occurrence using proposed method.

Tested Date	Percentage of Occurrence	CPU Time
6/19/1997	28	6
6/23/1997	28	7
6/27/1997	28	7
7/1/1997	28	7
7/5/1997	42	7
7/9/1997	28	6
7/13/1997	28	7
7/17/1997	28	7
7/21/1997	42	6
7/25/1997	42	6
7/29/1997	28	7

Fig. 8. Efficiency obtained from proposed method.

6. FUTURE WORK

Even our current system works better, we need some enhancement to make the system more better. Our future works related to inductive learning. Inductive learning algorithms have been suggested as alternatives to knowledge acquisition for expert systems. However, the application of machine learning algorithms often involves a number of subsidiary tasks to be performed as well as algorithm execution itself. It is important to help the domain expert manipulate for a specific algorithm, and subsequently to assess the algorithm results. These activities are often called as post processing and knowledge distribution. The future enhancement discusses issues related to the application of the supervised learning algorithm, an important representative of the inductive learning family. A prototype workbench which has been developed to provide an integrated approach to the application of supervised learning is presented. The design rationale and the potential use of the system are justified. Finally, future directions and further enhancements of the workbench are discussed to make the system better.

7. CONCLUSION

Regression is very essential to organize data, retrieve information correctly and swiftly. Implementing Machine learning to classify data is not easy given the huge amount of heterogeneous data from a raw data set. The proposed algorithm depends entirely on the accuracy of the training data set for building its decision trees. The AR algorithm learns by supervision. It has to be shown what instances have what results. The data in the data set are unpredictable, volatile and most of it contains missing data. The way forward for Information retrieval in the dataset, is considered as a micro array methods according to this algorithms which are unsupervised and reinforcement learners are used to classify and retrieve the missing data. Thus the research explains the trends, threads and process of the proposed algorithm which was implemented for finding the missing values and predicting blood cancer disease in a successfully manner.

References

1. S. A. Armstrong, J. E. Staunton, L. B. Silverman, R. Pieters, M. L. den Boer, M. D. Minden, S. E. Sallan, E. S. Lander, T. R. Golub, and S. J. Korsmeyer, *Nat. Genet.* 30, 41 **(2002)**.
2. T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, and M. A. Caligiuri, *Science* 286, 531 **(1999)**.
3. S. Kim, E. R. Dougherty, Y. Chen, K. Sivakumar, P. Meltzer, J. M. Trent, and M. Bittner, *Genomics* 67, 201 **(2000)**.
4. R. Jornsten, H. Wang, W. J. Welsh, and M. Ouyang, *Bioinformatics* 21, 4155 **(2005)**.
5. L. P. Bras and J. C. Menezes, *Syst. Biol. (Stevenage)* 153, 105 **(2006)**.
6. A. W. Liew, N. F. Law, and H. Yan, *Brief Bioinform.* 12, 498 **(2011)**.
7. O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman, *Bioinformatics* 17, 520 **(2001)**.
8. S. Oba, M. A. Sato, I. Takemasa, M. Monden, K. Matsubara, and S. Ishii, *Bioinformatics* 19, 2088 **(2003)**.
9. H. Kim, G. H. Golub, and H. Park, *Bioinformatics* 21, 187 **(2005)**.
10. Z. Cai, M. Heydari, and G. Lin, *J. Bioinform. Comput. Biol.* 4, 935 **(2006)**.

11. X. Zhang, X. Song, H. Wang, and H. Zhang, *Comput. Biol. Med.* 38, 1112 **(2008)**.
12. W. K. Ching, L. Li, N. K. Tsing, C. W. Tai, T. W. Ng, A. Wong, and K. W. Cheng, *Int. J. Data Mining Bioinform.* 4, 331 **(2010)**.
13. K. O. Cheng, N. F. Law, and W. C. Siu, *Pattern Recog.* 45, 1281 **(2012)**.
14. T. H. Bø, B. Dysvik, and I. Jonassen, *Nucl. Acids Res.* 32, e34.1 **(2004)**.
15. M. Ouyang, W. J. Welsh, and P. Georgopoulos, *Bioinformatics* 20, 917 **(2004)**.
16. M. K. Choong, M. Charbit, and H. Yan, *IEEE Trans. Inf. Technol. Biomed.* 13, 131 **(2009)**.
17. R. Jornsten, H. Y. Wang, W. J. Welsh, and M. Ouyang, *Bioinformatics* 21, 4155 **(2005)**.
18. X. Pan, Y. Tian, Y. Huang, and H. Shen, *Genomics* 97, 257 **(2011)**.
19. X. Gan, A. W. C. Liew, and H. Yan, *Nucl. Acids Res.* 34, 1608 **(2006)**.

Received: 24 April 2018. Accepted: 24 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1688–1694, 2018

RESEARCH ARTICLE

E-Mail Spam Classification Using Machine Learning in Distributed Environment

V. Sri Vinitha[1,*], D. Karthika Renuka[2], and A. Bharathi[1]

[1] *Department of IT, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamil Nadu, India*
[2] *Department of IT, PSG College of Technology, Coimbatore 641004, Tamil Nadu, India*

Electronic mail is a digital mechanism for exchanging messages through Internet or intranet communication platforms. Email spam, also known as junk email or unsolicited bulk email (UBE), is a subset of spam that involves nearly identical messages sent to numerous recipients by email. Spam's are usually sent to try to get the person to buy something or do something else that will cause gain for the sender. Spammers collect e-mail addresses, contact numbers from chat rooms, websites, customer lists and newsgroups etc. They are sent for the purposes of advertising, phishing, spreading malware, etc. Fraudulent spam also comes in the form of phishing emails, which are emails disguised as official communication from banks, online payment processors or any other organizations a user may trust. Phishing emails typically direct recipients to a fake version of the organization's website, where the user is prompted to enter personal information, such as login and credit card details. Spam email may also deliver other types of malware through file attachments or scripts, or contain links to websites hosting malware. In order to avoid these problems, spam classifications is done which filter the spam e-mails from inbox, moves it to the junk e-mail folder. The proposed e-mail spam classification system consists of two phases, such as training phase and testing phase. At first, input e-mail data is pre-processed in order to remove any errors or missing data in the dataset. After pre-processing, spam emails are classified using ensemble classification algorithm. In order to improve the performance, email spam classification is done using Random Forest in the Hadoop single node cluster. Both the results are compared through evaluation metrics namely, sensitivity, specificity and accuracy.

Keywords: Spambase Dataset, Logistic Regression, Support Vector Machine, Random Forests, Hadoop, Map Reduce.

1. INTRODUCTION

Email is the most widely used features of the internet. It is used for sending and receiving messages from anyone with an email address to anywhere in the world. Email spam, also known as junk email, is a type of electronic spam where unsolicited messages are sent by email. Many email spam messages are commercial in nature but may also contain disguised links that appear to be for familiar websites but in fact lead to phishing web sites or sites that are hosting malware. Spam email may also include malware as scripts or other executable file attachments (trojans). Spammers collect email addresses from chatrooms, websites, customer lists, newsgroups, and viruses that harvest users' address books. These collected email addresses are sometimes also sold to other spammers. A spammer typically sends an email to millions of email addresses, with the expectation that only a small number will respond or interact with the message. Spam classification is the process of filtering spam e-mail from inbox and moved to the junk e-mail folder.

2. RELATED WORKS

Spam is a very common issue faced by users on the internet, as it not only hinders the performance of a network, but it also wastes space and time, and also spread viruses, malware, spyware and cause consequent system failure. Mukund[1] proposes cognition to improve the performance of the distributed system. In this paper, spam e-mails are labelled using Machine Learning and the Naive Bayesian approach in MapReduce framework (hadoop) and multicast update. The comparative analysis of these two methods were made. The simulation results show that, the new reduce ( ) method reduces the time taken by MapReduce jobs by a significant amount. The reduction is maximum

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1688/007 doi:10.1166/jctn.2018.7362

for a mediumly large number of systems and decreases as the number of systems increase.

Nowadays E-mail is used for exchanging information. Unsolicited Bulk E-mail (UBE) known as spam mail is considered to be a major threat. Spam causes several problems, some of them resulting in direct financial losses. In real time,[2] processing is done with a large and high dimensional data. Because of that curse of dimensionality problem occurs, to overcome this firefly feature selection method is used to select the best features. With that selected features, spam classification is done using Naïve Bayes classifier. These two processes are effectively implemented in the distributed map-reduce framework and result is compared with the existing methods such as particle swarm optimization and neural network for spam base and CSDMC2010 SPAM corpus. The proposed system provides more accuracy.

Electronic mail, is a most popular, fastest and cheapest way of exchanging digital messages using Internet and becoming an integral part of everyday life for millions of people. However, unwanted or junk e-mail also known as spam became a foremost problem on the today's Internet and it not only consume users' time and energy to recognize and eliminate the undesired messages, but also become origin of many frustrating problems for instance taking up restricted mailbox space, overwhelming important individual emails, and wasting network bandwidth. To solve these problems, the users of e-mail should have automated tool that can filter the spam e-mails automatically. In this study,[3] the experiments were conducted for spam emails filtering task on the dataset obtained from UCI Machine Learning Repository separately using ten machine learning algorithms with ten-fold cross validation. The result obtained shows that classifier Random Forest is out performing with AUC, accuracy and MCC value up to 0.987, 0.955 and 0.906 respectively.

In this paper [4], author proposed a novel technique based on deep learning techniques to address the problem of spam drift and information fabrication. The syntax of each tweet will be learned through WordVector Training Mode. Based on the preceding dataset representation, binary classifier is constructed. In experiments, 10-day real Tweet datasets is collected and implemented in order to evaluate the deep learning techniques. The performance of different classifier is studied and compared deep learning method with other existing text-based methods. Further, it is also compared with non-text-based detection techniques. According to the experiment results, proposed method was more accurate.

With the growth of the internet and the increasing importance of emails in our daily lives, spams have become a common phenomenon posing serious threats, as it gives rise to undesired emails. Image spams are prevalent with the increased use of emails nowadays which cannot be efficiently detected even with latest techniques. This paper [5] proposes a solution for image spam detection using SVM with Gaussian kernel based classifier. Extensive experiments were conducted on ISH, TREC07, Dredze and personal images datasets considering the four performance parameters: accuracy, recall, precision, and F-measure. Evaluation of results indicates that Support Vector Machine (SVM) with Gaussian kernel based classifier outperforms comparatively than the other consider spam detection methods based on performance parameters.

3. PROPOSED METHODOLOGY

Figure 1 represents the overall process of proposed Email spam classification using Random Forest algorithm in Distributed environment.

3.1. Dataset

UCI repository Spambase dataset is used for classifying spam and ham mails. It contains 4601 instances and 57 attributes. The last column of the dataset denotes whether the mail is spam (1) or not (0). Most of the attributes indicate whether a particular word or character was frequently occuring in the e-mail. The description of the attributes is given below,

- 48 continuous real [0,100] attributes of type word_freq_WORD
- 6 continuous real [0,100] attributes of type char_freq_CHAR]
- 1 continuous real [1,...] attribute of type capital_run_length_average
- 1 continuous integer [1,...] attribute of type capital_run_length_longest
- 1 continuous integer [1,...] attribute of type capital_run_length_total
- 1 nominal {0,1} class attribute of type spam.

3.2. Data Preprocessing

Data pre-processing is a technique in data mining that is used to transform the raw data into clean dataset. In real

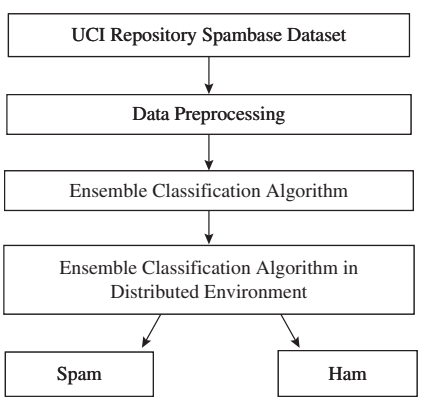


Fig. 1. Overall process of email spam classification.

world, some data has missing attribute values, inconsistent and is likely to contain errors or outliers, which leads to misclassification. So, pre-processing is done before feeding data into the algorithm.

Data goes through a series of steps during pre-processing:

Data Cleaning: Data is cleansed through processes to fill missing values, smooth noisy data, identify or remove errors and resolve inconsistencies.

Data Integration: Integrates data from different sources and conflicts within the data are resolved.

Data Transformation: Data is transformed into the form appropriate for mining using Normalization and Aggregation.

Data Reduction: It is helpful is analyzing reduced representation of the dataset without compromising the integrity of original data and producing the similar analytical results.

Data Discretization: It reduce the number of values for a given continuous attribute by dividing the range of a continuous attribute into intervals.

3.3. Ensemble Classification Algorithm

Ensemble methods (Fig. 2) are techniques that create multiple models and then combine them to produce improved results. Ensemble methods usually produces more accurate solutions than a single model would. In this paper, Logistics Regression, Support Vector Machine and Random Forest algorithm are used to classify the spam and ham emails. Logistics Regression is suitable for classification tasks. This algorithm applies a logistic function to a linear combination of features to predict the outcome of a categorical dependent variable based on predictor variables. Support Vector Machine is a supervised machine learning algorithm which can be used for both classification and regression analysis. Here, data is plotted in n-dimensional space with the value of each feature being the value of a particular coordinate. Then classification is performed by finding the hyper-plane that differentiate the two classes very well. Random forest is a tree-based algorithm which involves building several trees (decision trees), then combining their output to improve generalization ability of the model. The method of combining trees is known as an ensemble method. Random forests algorithm can be used for both classification and regression tasks. Ensembling is nothing but a combination of weak learners (individual trees) to produce a strong learner.

3.4. Ensemble Classification Algorithm in Distributed Environment

Hadoop cluster is a special type of computational cluster designed for storing and analyzing vast amount of unstructured data in a distributed computing environment. These clusters run on low cost commodity computers. Distributed environment is chosen to generate data at various geographical locations and needs to be available locally most of the time. The data and software processing is distributed to reduce the impact of any particular site or hardware failure. In this paper, Random Forest classification is done in the Hadoop single node cluster. In single node cluster, only one DataNode running and setting up all the NameNode, DataNode, ResourceManager and NodeManager on a single machine. A DataNode stores data in the Hadoop File System. A functional file system has more than one DataNode, with the data replicated across them, so that if one data is lost it will process with the other data. NameNode keeps the directory of all files in the file system, and tracks where across the cluster the file data is kept. JobTracker coordinates the parallel processing of data using

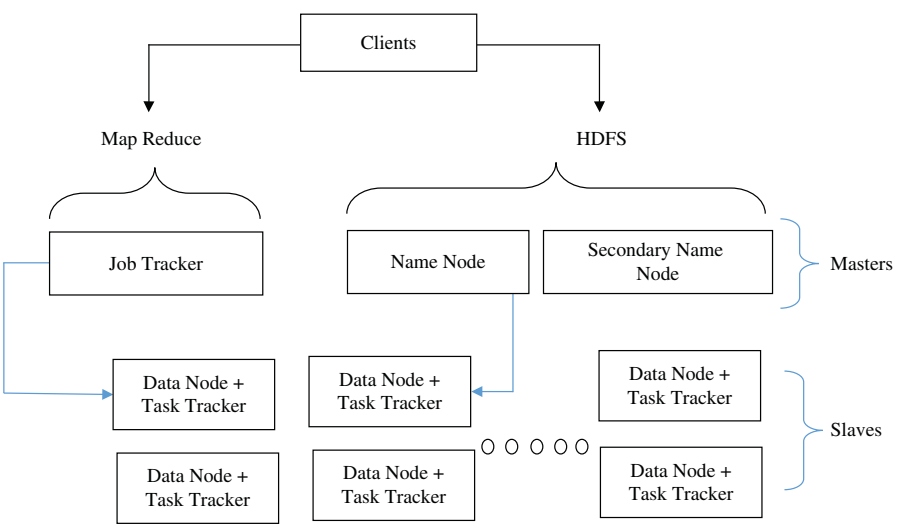


Fig. 2. Ensemble classification algorithm in distributed environment.

MapReduce. A TaskTracker is a node in the cluster that accepts tasks-Map, Reduce and Shuffle operations-from a Job Tracker. Secondary Namenode whole purpose is to have a checkpoint in HDFS. It is just a helper node for namenode.

4. IMPLEMENTATION

4.1. Hadoop Installation Steps

Apache hadoop is supported by all flavors of GNU/LINUX platforms or hadoop installation can also be done through Virtual Machine which has Linux in it. Hadoop framework is written in java. Before starting the hadoop installation, check java is installed and latest package is updated from all the repositories and PPA's. Dedicated Hadoop user account is used for running Hadoop. It is recommended to separate the Hadoop installation from other software applications and user accounts running on the same machine. Add the user hduser and the group hadoop to the local machine. Make this hduser as sudoer (admin). The hadoop control scripts rely on SSH to perform cluster-wide operations like stopping and starting all the daemons in the clusters. Hadoop requires SSH access to manage its nodes, i.e., remote machines plus local machine.

For a single-node setup of Hadoop,[15] configure SSH access to localhost for the hduser user. So, we need to have SSH up and running on our machine and configured it to allow SSH public key authentication. To work seamlessly, SSH needs to be setup to allow password-less login for the hadoop user from machines in the cluster. The simplest way to achieve this is to generate a public/private key pair using RSA algorithm and share it with different users. Then add the newly created key to the list of authorized keys so that Hadoop can use ssh without prompting for a password. Enable SSH access to local machine with this newly created key. The final step is to test the SSH setup by connecting to the local machine with the hduser user. The step is also needed to save the local machine's host key fingerprint to the hduser user's known hosts file. If the SSH connection fails, then

- Enable debugging with ssh-vvv localhost and investigate the error
- Check the SSH server configuration in /etc/ssh/sshd_config. If any changes made to the SSH server configuration file, then force a configuration reload with sudo /etc/init.d/ssh reload.

Install hadoop and move the contents of hadoop to a directory /usr/local/hadoop. Sometimes will get an error. This error can be resolved by logging in as a root user, and then add hduser to sudo. To complete hadoop setup, the following files have to be modified,

- ~/.bashrc
- /usr/local/hadoop/etc/hadoop/hadoop-env.sh
- /usr/local/hadoop/etc/hadoop/core-site.xml
- /usr/local/hadoop/etc/hadoop/mapred-site.xml.template
- /usr/local/hadoop/etc/hadoop/hdfs-site.xml

The path where the java has been installed should be identified to set the JAVA_HOME environment variable. Now edit the .bashrc file in home directory and move the end of file to hadoop. Modify the hadoop-env.sh file to set JAVA_HOME so that the value of JAVA_HOME variable will be available to Hadoop whenever it is started up. The /usr/local/hadoop/etc/hadoop/ core-site.xml file contains configuration properties that Hadoop uses when starting up. This file can be used to override the default settings that Hadoop starts with. By default, the /usr/local/hadoop/etc/hadoop/ folder contains /usr/local/hadoop/etc/hadoop/mapred-site. xml.template file which has to be renamed/copied with the name mapred-site.xml. The mapred-site.xml file is used to specify which framework is being used for MapReduce. The /usr/local/hadoop/etc/hadoop/hdfs-site.xml file needs to be configured for each host in the cluster that is being used. It is used to specify the directories which will be used as the *namenode* and the *datanode* on that host.

To start using, format the hadoop file system. The format command should be issued with write permission since it creates *current* directory under /usr/local/hadoop_store/hdfs/namenode folder. The *hadoop namenode-format* command should be executed once before we start using Hadoop, otherwise it will destroy all the data on the Hadoop file system. Now, the newly installed single node cluster can be started using *start-all.sh* or (*start-dfs.sh* and *start-yarn.sh*). To stop all the daemons running on the machine, run *stop-all.sh* or (*stop-dfs.sh* and *stop-yarn.sh*).

4.2. Logistics Regression

Logistic Regression[7] is a binary classification algorithm. It is also known as logit regression because it predicts the probability of occurrence of an event by fitting data to a logit function. Logistic Regression[8] is used to estimate discrete values based on the given set of independent variables. The output it predicts lies between 0 and 1 as in Figure 3.

The log odds of the outcome are modeled as a linear combination of the predictor variables.

$$\text{odds} = p/(1-p) = \text{probability of event occurrence} / \text{probability of not event occurrence} \quad (1)$$

$$\ln(\text{odds}) = \ln(p/(1-p)) \quad (2)$$

$$\text{logit}(p) = \ln(p/(1-p)) = b0 + b1X1 + b2X2 + b3X3 + \cdots + bkXk \quad (3)$$

In the above odds, p is the probability of presence of the characteristic of interest. It chooses parameters that maximize the likelihood of observing the sample values rather than that minimize the sum of squared errors.

RESEARCH ARTICLE

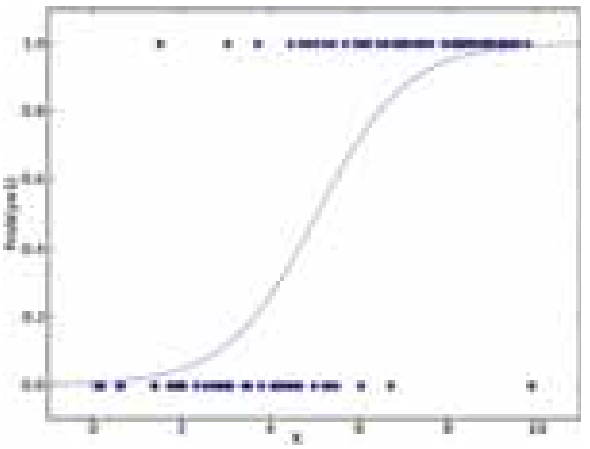


Fig. 3. Logistic regression.

4.3. Support Vector Machine

SVM, as a supervised learning technique, is a conjunction of linear learning machine and kernel function. It creates a hyperplane which divides two classes of variables by maximizing the margin between the hyperplane and the closest data points by assigning a set of weights w to the feature vector as in Figure 4. In the SVM only points that lie close to the decision boundary influence the optimal solution. Thus, the decision function is specified only by a subset of training samples, called the support vectors. In soft-margin SVM approach some observations are allowed to overlap to wrong classes. Formally, it can be presented in the form of finding an optimal solution to the quadratic problem.

The primal problem can be transformed into easier in computation dual form:

$$\max_{\alpha \in R: 0 \le \alpha \le C} \sum \alpha_i - 1/2(\alpha \cdot y)^T X^T X(\alpha \cdot y) \qquad (4)$$

where expression $X^T X$ denotes kernel function.

Using stack variables, one is able to control if observations can be predicted on the wrong side of the margin. It allows us to take under consideration unobserved features which may still influence observations and thus cause noise in data. Large volume of parameter C will discourage overlapping of classifications and in extreme cases may lead to over fitting. Small value of C will lead to loosening of our requirements and to the smoother decision boundary.

What makes SVM attractive in e-mail classification problems is its robustness and ability to handle large feature spaces. Since the algorithm is not trying to minimize the error rate, but rather separate the patterns in high dimensional space, the result is that SVMs are quite insensitive to the relative size of each class.

4.4. Random Forests

Random Forest is a supervised ensemble-learning model used for classification and regression. This algorithm creates the forest with a number of trees and make it random as in Figure 5. A Random Forest consists of a collection or ensemble of simple tree predictors, each capable of producing a response when presented with a set of predictor values. For classification problems, this response takes the form of a class membership, which associates, or classifies, a set of independent predictor values with one of the categories present in the dependent variable. Alternatively, for regression problems, the tree response is an estimate of the dependent variable given the predictors. A Random Forest consists of an arbitrary number of simple trees, which are used to determine the final outcome. For classification problems, the ensemble of simple trees vote for the most popular class. In the regression problem, their responses are averaged to obtain an estimate of the dependent variable. Using tree ensembles can lead to significant improvement in prediction accuracy because of the higher number of trees in the forest.

There are two stages in Random Forest algorithm, one is random forest creation, the other one is to make a prediction from the random forest classifier created in the first stage.

Pseudocode for Random Forest creation:

1. Randomly select "K" features from total "m" features where $k \ll m$
2. Among the "K" features, calculate the node "d" using the best split point

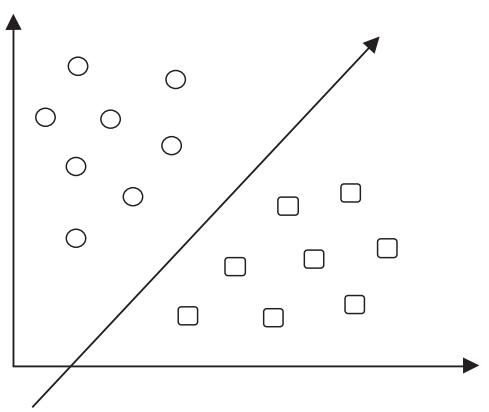


Fig. 4. Support vector machine.

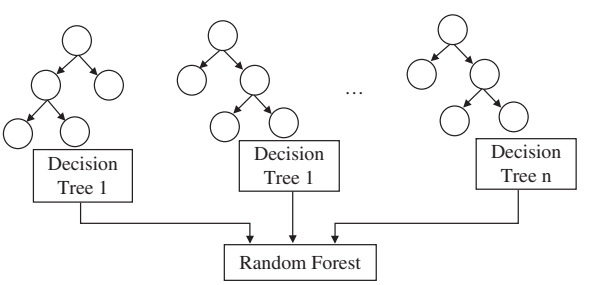


Fig. 5. Random forest algorithm.

3. Split the node into *child nodes* using the *best split*
4. Repeat the 1 to 3 steps until "*l*" number of nodes has been reached
5. Build forest by repeating steps 1 to 4 for "*n*" number times to create "*n*" *number of trees*.

After creating random forests, the next stage is to make predictions. The pseudocode for random forest prediction is,

1. Takes the *test features* and use the rules of each randomly created decision tree to predict the outcome and stores the predicted outcome (target)
2. Calculate the *votes* for each predicted target
3. Consider the *high voted* predicted target as the *final prediction* from the random forest algorithm.

5. RESULT ANALYSIS

Figure 4 shows the performance comparison for spambase dataset. The performance analysis has been made by using the evaluation metrics such as precision, recall and accuracy.

5.1. Evaluation Metrics

5.1.1. Precision

It is the ratio of true positive to true and false positives as in Figure 6. This determines how many identified objects in a class were correct.

$$\text{Precision (P)} = \text{TP}/(\text{TP}+\text{FP}) \quad (5)$$

5.1.2. Recall

It is the ratio of true positives to the number of true positives and false negatives as in Figure 7. This determines how many objects in a class are misclassified as something else.

$$\text{Recall (R)} = \text{TP}/(\text{TP}+\text{FN}) \quad (6)$$

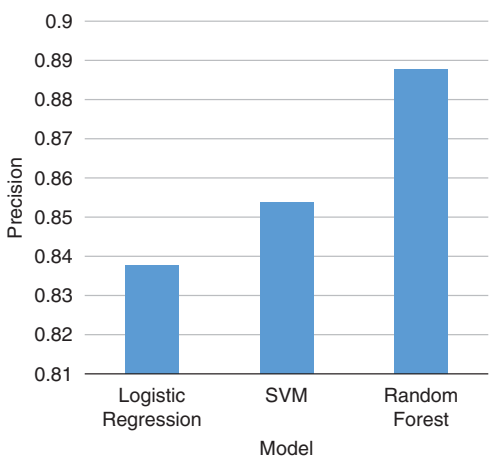


Fig. 6. Performance analysis using precision.

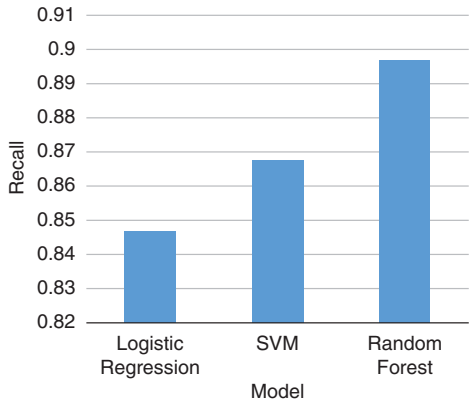


Fig. 7. Performance analysis using recall.

5.1.3. Accuracy

It is defined as the sum of all true positives and true negatives to the total number of test instances as in Figure 8. This measures the overall accuracy of the classifier.

$$\text{Accuracy (A)} = (\text{TP}+\text{TN})/(\text{TP}+\text{TN}+\text{FP}+\text{FN}) \quad (7)$$

Where,

- *True Positive* (*TP*) is the number of spam documents correctly classified as spam.
- *True Negative* (*TN*) is the number of non-spam documents correctly classified as non-spam.
- *False Positive* (*FP*) is the number of spam documents classified as non-spam.
- *False Negative* (*FN*) is the number of non-spam documents classified as spam.

By analyzing the plotted graph, the performance of proposed e-mail spam classification is significantly improved by using the Random Forest algorithm compared to Logistic Regression and Support Vector Machine. Ensemble methods use multiple learning algorithms to obtain better predictive performance.

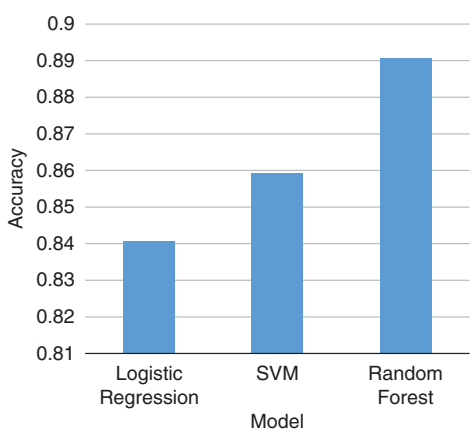


Fig. 8. Performance analysis using accuracy.

Table I. Comparison of precision, recall and accuracy.

	Model	Precision	Recall	Accuracy
0	Logistic regression	0.837624	0.846754	0.840721
1	SVM	0.853807	0.867646	0.859390
2	Random forest	0.887624	0.896754	0.890721

6. CONCLUSION

Nowadays, E-mail plays a vital role in exchanging information. Spam is unwanted or irrelevant email sent to the user, as it wastes user time and also wastage of bandwidth. In order to avoid these spam email, spam classification is done. Here, Ensemble classification algorithm is used to classify the spam and ham mails. Initially, input email is pre-processed to remove missing values and errors. After pre-processing, spam email is classified using Random forests algorithm. The same pre-processed dataset is implemented in the distributed environment using Hadoop single node environment in order to improve the performance of the classifier. The results for the spam classification is validated through evaluation metrics namely, accuracy, sensitivity and specificity. Proposed system in Table I shows that the Random Forests classification algorithm in the Distributed environment using Hadoop single node environment gives better accuracy compared to Logistic Regression and Support Vector Machine.

7. FUTURE WORK

Further, the performance can be improved by selecting the best features using feature selection method. After selecting the best features, spam mails can be classified by using deep learning statistics to improve the accuracy.

Acknowledgment: Our sincere thanks to University Grants Commission (UGC), Hyderabad for granting the funds to carry out this work.

References

1. Y. R. Mukund, Sunil Sandeep Nayak, and K. Chandrasekaran, *Open Journal of Big Data (OJBD)* 1, 16 **(2015)**.
2. Karthika Renuka Dhanaraj and Visalakshi Palaniswami, *Australian Journal of Basic and Applied Sciences* 8, 118 **(2014)**.
3. Manish Kumar, *International Journal of Innovative Research in Computer and Communication Engineering* 4, 3200 **(2016)**.
4. T. Wu, S. Liu, J. Zhang, and Y. Xiang, Twitter Spam Detection based on Deep Learning, ACM **(2017)**.
5. Prashant Kumar and Mantosh Biswas, SVM with gaussian kernel-based image spam detection on textual features, *International Conference on Computational Intelligence and Communication Technology* **(2017)**.
6. Mohammed Nazim Feroz and Susan Mengel, Examination of data, rule generation and detection of phishing URLs using online logistic regression, *IEEE International Conference on Big Data* **(2014)**, pp. 241–250.
7. Adi Wijaya and Achmad Bisri, Hybrid decision tree and logistic regression classifier for email spam detection, *International Conference on Information Technology and Electrical Engineering* **(2016)**.
8. Ahmed Hamza Osman, Hani Moaiteq Aljahdali, *International Journal of Advanced Computer Science and Applications* 8, 420 **(2017)**.
9. Shashank Mishra and D. Malathi, Behaviour analysis of SVM based spam filtering using various parameter values and accuracy comparison, *IEEE International Conference on Computing Methodologies and Communication* **(2017)**.
10. T. Kumaresan, S. Saravanakumar, and R. Balamurugan, Visual and Textual Features Based Email Spam Classification Using S-Cuckoo Search and Hybrid Kernel Support Vector Machine, Springer **(2017)**.
11. Shradhanjali and Toran Verma, *International Journal of Advance Research, Ideas and Innovations in Technology* 3, 1491 **(2017)**.
12. Mohammed Zakariah, *International Journal of Engineering and Innovative Technology (IJEIT)* 4, 189 **(2014)**.
13. Bhagyashri U. Gaikwad and P. P. Halkarnikar, *International Journal of Advanced Computer Engineering and Communication Technology (IJACECT)* 2, 1 **(2013)**.
14. A. Akinyelu and A. Adewumi, *Journal of Applied Mathematics* **(2014)**.
15. Abdelrahman AlMahmoud, Ernesto Damiani, Hadi Otrok, and Yousof Al-Hammadi, *IEEE Transactions on Big Data* **(2017)**.

Received: 24 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1695–1699, 2018

Fabrication Design of Axially Rotated Square Resonator Based Compact Ultra Wide Band-Band Pass Filter Using Tight Coupling

S. Oudaya Coumar* and S. Tamilselvan

Department of Electronics and Communication Engineering, Pondicherry Engineering College, Puducherry 605014, India

RESEARCH ARTICLE

This paper is about a axially rotated Ultra wideband BPF with tight coupling line. A novel axially rotated square resonator integrated with inter-digital coupling otherwise called as tight coupling on both sides plays a key role in this filter design. Design and EM Simulation of the UWB filter's characteristics are discussed in this work. The proposed UWB filter produces tremendous bandwidth ranges from 2.5 GHz to 8 GHz. It also exhibit out band characteristics which meets the FCC's requirement. The simulation study of the projected filter is executed by electromagnetic tool. The filter evaluation parameters like return loss, insertion loss, phase and group delay are obtained and their responses are analyzed. The complete size measurement of the filter is achieved to be 39 mm × 3.2 mm × 1.6 mm on accounting the above features. The calculated bandwidth ratio with respect to center frequency of the filter is obtained about 115%. The performance parameter results shows the filter's return loss (S_{11}) and insertion loss (S_{21}) are −40 dB and −1 dB respectively.

Keywords: Ultra Wideband Filter (UWB), Square Resonator, Insertion Loss, Return Loss, Group Delay Analysis.

1. INTRODUCTION

The growth of Wireless radio communication technology in the engineering field over few years is emerging drastically. It is believed that wireless communication technology will continue to grow in the near future since the necessity of wireless communication receivers are increasing day by day because of rapidly developing wireless systems. The major development in wireless communication is mobile radio communication system which has wide range of applications like data, video, ranging in mobile phones. The ultimate goal of wireless communication system is to provide the best services in the limited power and within the allocated spectrum. Nowadays mobile users are tremendously increasing in large number, whereas there is a tough challenge in managing crowding of spectrum and coexistence of wireless devices whose operating frequency is close to one another will become a major issue. The mobile service providers are facing lot of challenges in allocating bandwidth and best data rate to all the users from limited availability of resources. In this regard, an innovative and efficient technology is needed that can able to satisfy all the requirements of service providers and consumes as well. Ultra wideband is one such technology which can able to meet out all the expectations of mobile service providers and it can coexists with licensed and unlicensed narrowband very well.[1]

As per Federal Communication Commission-USA declaration, the frequency band range starts from (3 to 10) GHz is assumed to be license free band. This wide band is mainly suitable for high speed indoor data communication applications and this wide range band is termed to be Ultra wideband. The Ultra wideband technology has created excellent opportunities and increasing research interests towards wireless industries. More researches has been recorded in the literature towards exploring various segment of the UWB system.[2]

UWB technology is extremely the best support for ultra-high speed radio applications because it has huge bandwidth which provides very high data rate. UWB system has certain special features namely, it operates in low-power which is near to −41 dBm, requires minimum cost to design the communication system and can able to achieve positioning. These features can give way for plenty of applications such as logistics, security applications, medical applications, high speed data applications and military/navigation applications.[3] UWB system also have following some unusual characters, it occupies a bandwidth very much greater than 0.5 GHz and fractional bandwidth larger than 20% of the actual value. It is obvious that, the larger bandwidth provides a very high

*Author to whom correspondence should be addressed.

data handling capacity as per Shannon's capacity theorem. This results to increase the processing gain thereby it allows large number of users can access the system with equal best services.

In this rapidly developing wireless system, the challenging deal is to design the various system elements that should be compatible and match the UWB system's requirements. In this aspects, design of band pass and band stop filters received lot of attention among many researchers as they are the unavoidable segment of UWB system.[4,5] In the recent days, the wireless communication industry focuses much on miniaturization of receiver system as much as possible. In this regard, many researchers are working towards the advancements in miniaturization with least cost on filters as they are the essential block of any receiver system and crack the attraction of end users as well.

The proposed axially rotated UWB-BPF filter consists of simple compacted square shaped resonator. The structural design of the BPF is unique and simple. The BPF is designed in such a way that it has limited circuit elements when compared with the existing complex filter designs in the literature. The organization of the paper is given by: Section 2 details the operation of square shaped resonator design model. In Section 3, the proposed compact UWB-BPF using square resonator is described. Simulation results and analysis are discussed in Section 4 and finally Section 5 concludes the paper.

2. SQUARE RESONATOR DESIGN MODEL

The square resonator principle is similar to logic of distributed transmission line confined in a closed area. The logical representation of the resonator comprises of the feed lines for input signal, inter-digital coupling lines for strong matching, and the resonator structure to produce sustained oscillation.[6] Figure 1 shows a type of square shaped resonator based UWB-BPF arrangement. The power to be transmitted is fed into any of the ports which pass through the feed lines, couple lines and resonating unit thereby the power reached the output port. The filter in the structure thereby the performance is reciprocal. The resonant frequencies of the square resonator remains unaffected only when the feed lines and gap between the resonator is constant.

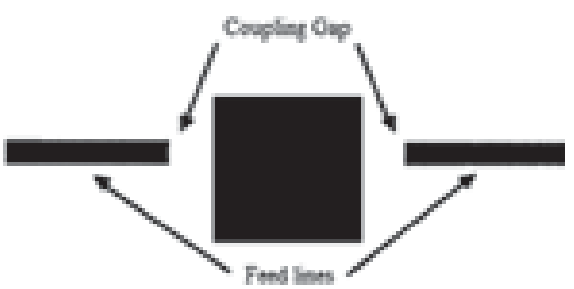


Fig. 1. Microstrip square resonator.

In the literatures, this fashion of coupling is denoted as loose or weak coupling. This types of weak coupling exhibits trivially lesser capacitance in between the coupling lines and the resonator. If the gap between the feed lines and the resonator is small, then there will be an increase in capacitances which in turn produces strong coupling. This effects may results in the deviation of resonant frequencies from the desired resonant frequencies of the square resonator. In order to specifically design the sustained oscillation of square resonator, the capacitances values of the coupling should be taken into consideration. When the average value of the sides of the square resonator is same as that of the integral multiple of a transmission line wavelength, then resonance will takes place. This resonant frequency can be mathematically represented in Eqs. (1) and (2) as

$$2\pi r = n\lambda_g \tag{1}$$

$$f_0 = \frac{nC}{2\pi r\sqrt{\varepsilon_{\text{reff}}}} \tag{2}$$

Where, λ_g guided wavelength, 'r' mean side of the square, 'n' mode value, f_0 resonant frequency, 'c' light speed in free space, $\varepsilon_{\text{reff}}$ effective dielectric constant. This mathematical representation is applicable for the case of loose coupling because coupling gap effects are not considered. The resonant frequencies for various modes can also be manipulated from the above mentioned expression because λ_g is dependent on frequency.[7–9] In the starting mode, the field will be maximum at the position of coupling gap and the there is no field takes place at normal from the coupling gap locations.

3. UWB BANDPASS FILTER USING SQUARE RESONATOR

The geometric design topology of the proposed axially rotated UWB-BPF can be viewed in Figure 2. The axially rotated UWB-BPF considered in the work is mainly comprised of a simple compact square resonator structure. The BPF is designed on assuming (FR4) as substrate with dielectric height of 1.6 mm and a relative permittivity of 4.4.

This substrate is compatible with all the elements of communication system. The cross section of such microstrip substrate used for the UWB-BPF design is shown in Figure 3. The square resonator based axially rotated UWB band pass filter is a simple circuit and easy to design whose optimal geometric values are listed in Table I.

The logical structure of the square resonator will allow only the waves which is an integral multiple of guide's wavelength and equals to the mean or average side. There are many more complicated circuits with complex topology are available in the literature.[10] However, this proposed simple filter topology produces better results on

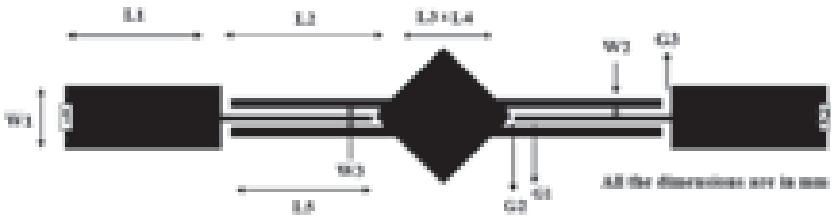


Fig. 2. Design topology of the UWB-BPF.

par with those complex structures. This types of resonator based circuits are very much useful for various applications.

The square resonator circuits can also be implemented over distributed transmission lines. Thickness determination of feed lines, coupling mode and type of resonator are chosen depending upon the applications and operating frequency ranges. Most resonator topologies developed in the past years are microstrip line, coplanar waveguide, slotted line and defected-ground structure.[11, 12]

The signal is feed through feed lines whose width is designed in such a way to match the port impedance to 50 ohms. The inter connected multiple finger like coupling line arrangement is defined as inter digital which is associated with square shaped MMR (Multi Mode Resonator). It offers wide operating bandwidth and enrich out band performance.[13, 14] The inter digital coupling line and MMR are united together and produces sustained oscillation for the wide range of frequencies (3.1–10.6) GHz. This range of frequency is accurately corresponding the operating frequency of the UWB system building. Hence it is best candidate for UWB band pass filter. The dimensions and the size specifications of the BPF are mentioned in the Figure 2.

Fig. 3. Cross section of microstrip substrate.

Table I. Optimal geometric values of the projected UWB-BPF.

Filter elements	Descriptions	All dimensions are in mm
L	Length of the substrate	39
W	Width of the substrate	20
H	Thickness of the substrate	1.6
L1	Length of the feed line	8
W1	Width of the feed line	3.2
L2	Length of inter digital coupling line	8.3
W2	Width of inter digital coupling line	0.5
L3	Side of the square	6
L4	Side of the square	6
L5	Length of the middle inter digital coupling line	8.2
W3	Width of the middle inter digital coupling line	0.3
G1	Gap between inter digital coupling line	0.3
G2	Gap between coupling line and resonator	0.45
G3	Gap between feed line stub and coupling line	0.35

4. SIMULATION RESULTS AND ANALYSIS

This section deals with the discussion on the result analysis of the projected UWB-BPF. The simulation is carried out through full wave 3D EM simulation tool and its corresponding performances are evaluated. To investigate the characteristics over wide range of pass band, the effects of changing the various dimension of filter element with respect to the other elements are investigated in this work. This filter is analyzed for various parameters like strip width, length and thickness by assuming variable values for one parameter and maintaining constant with respect to other parameters.

The computer-generated Scattering characteristics S_{11}, S_{21}, phase, group delay and Q factor of the proposed UWB-BPF using square shaped resonator are observed in Figure 4. The response clearly infers that the band pass filter possess tremendous insertion loss value of −1 dB and the extreme small profile of return loss is obtained around −40 dB. These values are obtained by assuming −25 dB bandwidth line. The fractional bandwidth is manipulated from the response using bandwidth and resonant frequency ratio which is approximately 115%.

The phase variation response simulation of the filter with respect to S_{21} parameter values are clearly denoted in Figure 5. It is evident from the result that the projected BPF is having excellent in phase character in the pass band region. The significant and most important quality of a filter is that the filter's phase should be linear over the frequency range of the pass band, where the phase response with respect to return loss. Linear phase of the filter should

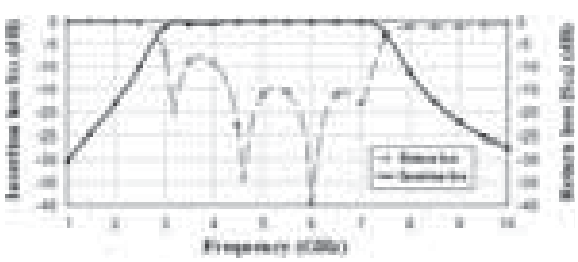


Fig. 4. Simulated S_{11} and S_{21} values of the UWB-BPF.

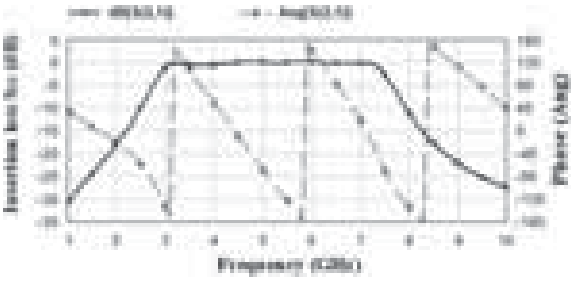


Fig. 5. Simulated phase of S_{21} of the projected UWB-BPF.

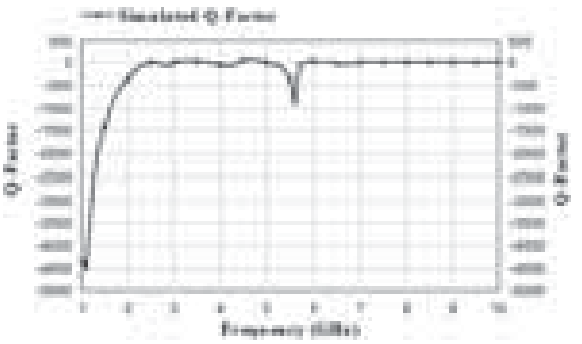


Fig. 6. Q factor response of the projected UWB-BPF.

have minimum group delay over the desired band. The phase of the concerned filter is appreciably linear so the best suited for UWB services throughout the pass band.

As far as filter design considered, the quality factor is a dimensionless factor that explains the extent to which how a resonator produced sustained oscillations for long period. It analyze the behavior a resonator's bandwidth by accounting its center frequency of the intended pass band

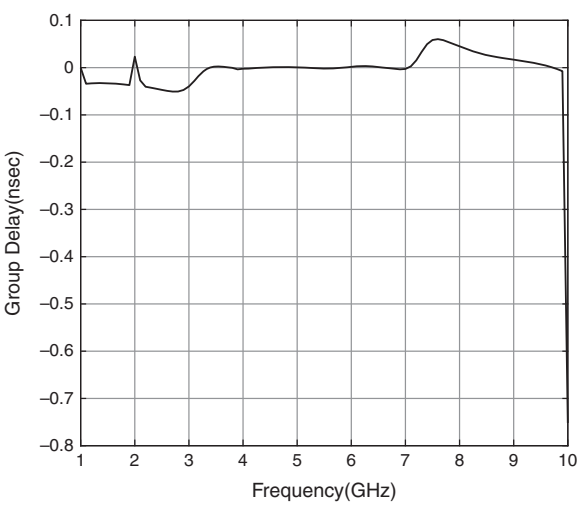


Fig. 7. Group delay response of the projected UWB-BPF.

Fig. 8. Prototype model of the UWB-BPF.

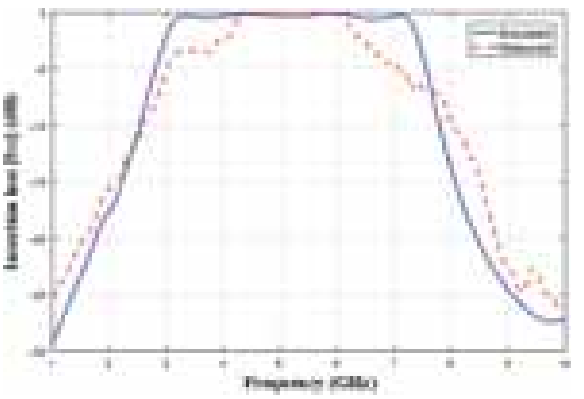


Fig. 9. Simulated and measured S_{21} of the proposed UWB-BPF.

range. The simulated Q factor response with respect to frequency for the concerned BPF is displayed in Figure 6.

On considering the UWB applications, investigation of uniform group delay is highly inevitable and needed. Figure 7 depicts the group delay simulation response of the UWB-BPF. The response implies the projected filter is having well excellent linear signal transfer. It is also measured from the Figure 7 that the group delay value lies

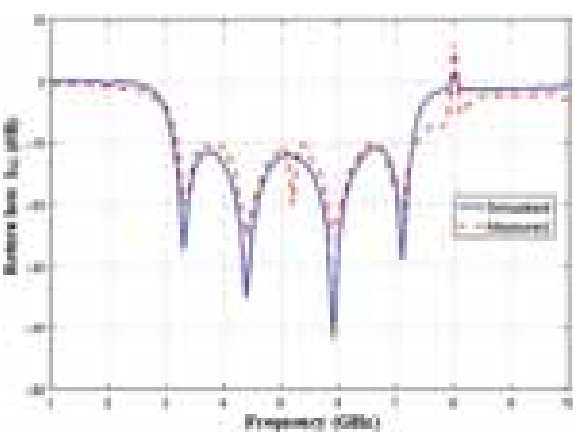


Fig. 10. Simulated and measured S_{11} of the proposed UWB-BPF.

Table II. Values observed from simulation.

Parameters	Simulated values of UWB-BPF	Measured values of UWB-BPF
Insertion loss $\|S_{11}\|$ in dB	−40 dB	−25 dB
Return loss $\|S_{21}\|$ in dB	−1 dB	−3 dB
Pass band	2.5 GHz to 8 GHz	2 GHz to 8 Ghz
Bandwidth	5.5 GHz	6 GHz
Phase at pass band	Linear	**Linear**
Group delay	< 0.1 nsec	**<0.1 nsec**
Size	39 mm × 3.2 mm × 1.6 mm	39 mm × 3.2 mm × 1.6 mm

below 0.1 ns in the UWB's pass band region. This feature confirms the UWB system has minimum distortion when implemented practical scenario.

The prototype realization of UWB-BPF is displayed in Figure 8. The insertion loss and return loss of the fabricated filter is measured using Network analyzer Model HP8757D as shown in Figures 9 and 10 respectively. A small disagreement between simulation and measured results are due to effects from fabrication accuracy and may be of calibration error of the SMA connectors.

Table II listed the essential and required filter performance parameters of the projected UWB-BPF.

5. CONCLUSION

Simple and compact topology of square resonator based Ultra-wideband (UWB) is detailed in this work. The band pass filter is realized by square shaped resonator associated with inter-digital conductors. The square shaped resonator produces wide pass band in the range of 2.5 (GHz) to 10 (GHz) which is UWB. An UWB bandpass filter having higher wide band is very much needed and highly required in real time UWB radio environment. The size of the square shaped resonator and inter digital coupling line designs can have flexibility to vary easily based on the wide band requirements. The simulation design and performance behavior of the UWB-BPF is explicitly deliberated. The filter considered in this work demonstrates smooth and uniform wide bandwidth ranges from 2.5 GHz to 8 GHz with the size of 39 mm × 3.2 mm × 1.6 mm. This circuit will also have excellent out of band performance in terms of roll on and roll off. On accounting all the discussed structural and performance characteristics, the fractional bandwidth is calculated about 115% with return loss and insertion loss of about −40 dB and −1 dB respectively. Thus the S parameter, phase, group delay and wide operating band behavior of simulated and practically measured values of UWB-BPF conveys better performance, so it is optimally fit for wide band wireless radio system applications.

References

1. Federal Communications Commission, Finf Repom and Order, Revisian of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems, adapted February **(2002)**.
2. Y.-Q. Liu and X.-Z. Ke, *Progress in Electromagnetics Research Letters* 59, 21 **(2016)**.
3. J. S. Hong and M. J. Lancaster, Microstrip Filters for RF/Microwave Applications, John Wiley & Sons, Inc., New York **(2001)**.
4. D. M. Pozar, Microwave Engineering, Wiley, New York **(1998)**.
5. R. Ludwig and P. Bretchko, RF Circuit Design Theory and Applications, Prentice-Hall Publications, London **(2000)**.
6. D. A. Letavin, Compact bandpass filter with microstrip resonators, *18th IEEE International Conference on Micro/Nanotechnologies and Electron Devices* **(2017)**.
7. S. Sun and L. Zhu, *IEEE Microw. Wireless Component. Letter* 16, 440 **(2006)**.
8. J.-T. Kuo and E. Shih, *IEEE Transactions on Microwave Theory and Techniques* 51, 1554 **(2003)**.
9. L. Lin, S. Yang, S.-J. Sun, B. Wu, and C.-H. Liang, *IET Electronics Letters* 50, 1218 **(2014)**.
10. Bahman Mohammadi, Arash Valizade, Pejman Rezaei, and Javad Nourinia, *IET Microwaves, Antennas and Propagation* 9, 64 **(2015)**.
11. T. Zhang, F. Xiao, J. Bao, and X. Tang, *IET Journals and Magazines, Electronics Letters* 52, 210 **(2016)**.
12. P. Cheong, T.-S. Lv, W.-W. Choi, and K.-W. Tam, *IEEE Microw. Wireless Compon. Lett.* 21, 77 **(2011)**.
13. Azzedin Naghar, Otman Aghzout, Ana Vazquez Alejos, Manuel Garcia Sanchez, and Mohamed Essaaidi, *IET Microwaves, Antennas and Propagation* 9, 1786 **(2015)**.
14. N. Janković, G. Niarchos, and V. Crnojević-Bengin, *Electronics Letters on IET Journals and Magazines* 52, 372 **(2016)**.

Received: 24 April 2018. Accepted: 9 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1700–1705, 2018

RESEARCH ARTICLE

Middleware for Service Composition in Cyber Physical System

Swati Nikam[1,*] and Rajesh Ingle[2]

[1] *Department of Computer Engineering, Dr.D.Y.Patil Institute of Technology (DIT), Affiliated to Savitribai Phule Pune University, Pimpri, Pune 411018, Maharashtra, India*
[2] *Department of Computer Engineering, Pune Institute of Technology (PICT), Affiliated to Savitribai Phule Pune University, Pune 411043, Maharashtra, India*

Cyber Physical Systems, faces many inherent challenges due to the tight coupling between cyber and physical world, heterogeneity at various scales, unreliable networking, close integration of computation and physical process, complexity at multiple temporal as well as spatial scale and highly dynamic operating conditions. Various core research challenges present in CPS are namely resource management, service composition, protocol standardization, achieving quality of service. Service composition means to combine more than one services when individual service is not sufficient as per the demand. While doing service composition, we may require the correct and timely provisioning of the resources because in the context of CPS, resources and services are equally important. Our work focuses on service composition and resource provisioning simultaneously. In this work, we design the middleware for Service Composition and Resource provisioning.

Keywords: Cyber Physical Systems, Resource Provisioning, Service Composition, Middleware.

1. INTRODUCTION

The term Cyber Physical Systems was first coined in by Gill in 2006 and gained much attention from academia as well as industry as it is very potential due to capability to depict the interaction between cyber and physical space. The application domain of CPS are mission critical areas such as transportation, health care, energy, defence, homeland security, smart grid, water distribution and many more. Each domain is becoming smarter day by day, as smartness in adopted in every service and every device. This gives rise to need of considering resources along with services. To coordinate the variety of services and numerous resources present, middleware seems to be the best approach. Sufficient discussion is done on core challenges of Service composition and Resource Provisioning in the context of CPS.[1,2] Middleware provides aims to hide the technological details. Many researchers have proposed variety of middleware in CPS with different perspective which is sometimes limited to theoretical framework discussions or sometimes limited to a particular domain or sometimes incapable to handle the primary challenges of CPS itself. The paper is organised as follows. Section 2 gives overview of related literature in the CPS domain. Section 3 gives on overview of challenges in middleware. Section 4 talks about our middleware design followed by conclusion in Section 5.

*Author to whom correspondence should be addressed.

2. LITERATURE REVIEW

Middleware[3] is the software that acts like a {glue, intermediary, broker, middle man, interpreter, abstraction provider, consolidator, integrator, facilitator, or connector} between one or more applications and other application(s) or between one or more applications and the underlying infrastructure including the platforms and operating systems used. Main functionality/features of middleware can be summarised as follows.

In the domain of CPS, many researchers have adopted middleware approach focusing on any particular aspects or a particular application or a particular QoS attributes. The prominent work is discussed here.

Bourcier[4] have implemented a middleware to enable the management of software deployment and the dynamic reconfiguration of these CPS systems. It has considered Component Based Systems and the model at runtime paradigm as the basis for the design of CPS.

Kiril[5] have proposed a middleware approach and demonstrated that service-oriented computing approach which improves portability without loss of performance. An extremely light-weight and flexible method for local and remote service interaction is proposed.

Dabholkar[6] presented an approach to systematically specialize general-purpose middleware used to host CPS. Their approach is based on the principles of Feature-Oriented Software Development (FOSD).

 1546-1955/2018/15/1700/006 doi:10.1166/jctn.2018.7364

Calyo[7] have talked about CPS requirement of new computing and networking technologies and they have analyzed the applicability of different middleware technologies as data distribution means for CPS.

Ismael[8] have focused on applications requirement such as autonomy, fault tolerance, energy efficiency and configurability issues. Their work described the design principles of CPS which needs cooperation.

Franke[9] have presented a flexible and semantic middleware (SeMiWa) for cyber-physical systems. The work focused on the challenge of controllability in a heterogeneous wireless infrastructure of a smart home environment. To solve the problem of interoperability, they have used semantic annotated XMPP messages as transmission format.

Kim[10] have worked on the issue of the design and implementation of real-time mechanisms in Etherware (a middleware developed at the University of Illinois). Looking at the safety requirement of CPS author has worked on temporal predictability and flexibility of Etherware.

Mundra[11] worked on Distributed CPS. Understanding the need to design reconfigurable, robust, validated and consistent real time middle ware systems with end-to-end timing, they have designed Distributed CPS-Hidden Markov Model considering processor efficiency and data validation.

Zhang[12] explained about need to handle aperiodic and periodic events. They have worked on design, implementation and performance evaluation of the first configurable component middleware services for admission control and load balancing of aperiodic and periodic event handling in distributed cyber-physical systems with end to end timing constraint.

Park[13] have discussed about a global CPS network that integrates different CPS networks appears in the near future. They have proposed a CPS middleware framework that ensures interoperability and communication between heterogeneous components in a global CPS network through IPv6.

Authors[14] have discussed challenges from application development point of view. They have put forward middleware challenges like Efficient integration, resource management, secure service, advanced communication services, load balancing, scalability, global reference time, supporting autonomous operations.

Mohamed[15] have proposed the middleware for processing distributed query along with data management algorithms.

Cuzzcrea[16] have proposed service oriented middleware to build CPS applications. They have also categorised CPS applications based on their scope and coverage as Nanoscale, body area CPS, local area CPS, Mobile ad hoc CPS, Wide area CPS. Also a two level broker architecture is presented.

Service oriented middleware can play an important role in facilitating the design, development and implementation of service oriented systems. Mohamed[17] have discussed about few SOM approaches in distributed environment along with its challenges.

Drenker[18] have focused on the need for dependability in the future generation of societal scale cyber-physical systems.

Garcia-Valls[19] have compared performance of various middle wares from scalability point of view in the context of distributed cyber physical systems.

Noguero[20] have presented time-triggered middleware architecture that provides fault tolerance and dynamic reconfiguration at run-time taking into account the available resources of the underlying infrastructure, i.e., devices and communication networks.

Some middlewares are designed focusing on a particular application. Iqubal[21] have designed a middleware framework for continuous monitoring of water distribution network to handle real-time data management, in-network collaborative event detection and hydraulic modelling.

The types of Middleware found in CPS domain are Context Aware, Service Oriented, QoS aware, Domian Specific middleware etc.

There is a tremendous literature on middleware not only in the domain of CPS, but also Distributed CPS, Networked CPS, Large scale CPS, Cyber Physical Cloud Computing. We have studied the prominent work in selected domains.

3. CHALLENGES IN THE MIDDLEWARE

The literature study points to some of the challenges in middleware which are as follows.

(1) *Need of generic middleware*: Looking at the variety of CPS application domains offering basic CPS services, there is a need to design and develop a generic middleware which can reuse already existing CPS components,

Layer	Functions/Algorithms	Components
Application layer	Seamless service to user	User interface through web or application
Service layer	Task dependency, Service Discovery, Service ranking, Service selection, Service composition	CPS middleware
Control layer	CPS model generation	
Context processing layer	Resource system comparator	CPS system
Resource layer	Resource provisioning, Resource node comparator, Service dependency resolution, Service deployment, Service execution	Resource system
	Task execution	Resource node

Fig. 1. Layered architecture.

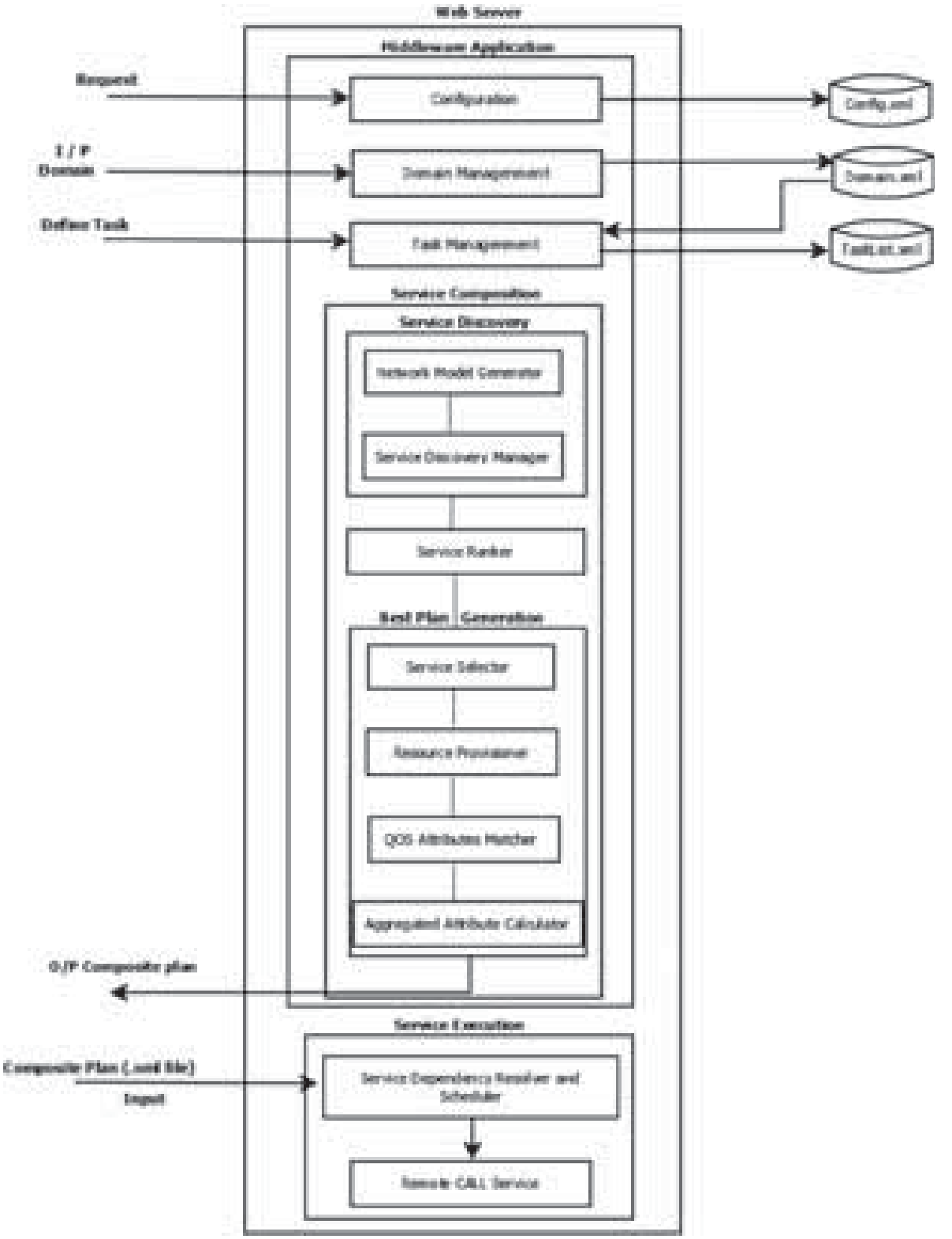


Fig. 2. Architectural block diagram of middleware for service composition and resource provisioning.

services. It should provide a base for enhancing inherant CPS challenges as well as to provide a common platform where services can be added, updated, enhanced or redesigned to cater the demands of CPS applications.
(2) *Reuse of existing resource*: Irrespective of the heterogeneity present in the various resources in terms of types, use, vendors, communication capability, operating standards, interfaces. The already deployed but underutilized resources should be utilized and thus can improve overall resource utilization.
(3) *Enabling smooth, efficient and Seamless integration*: The middleware should be able to integrate not only between services provided by different service providers but also between heterogeneous components present in CPS which differ in communication capabilities, synchronization capabilities, security capabilities.
(4) *Less complex*: The complexity will restrict the usage by application users, so it should be less complex.
(5) *Domain independent*: The middleware should consider the specific requirements of domain, but its working should not restrict to a particular domain only.
(6) *Understanding of joint dynamics*: Designing middleware for a particular domain has limited issues, but in case of CPS, it should consider the joint dynamics efficiently and fully.
(7) *Interoperability Support*: The middleware should support interoperability so that it can work seamlessly amongst variety of CPS applications.
(8) *Security Support*: As the scope of CPS becomes wider the possibility of threat also increases, so it should consider the necessary security requirements.
(9) *Safety Support*: A large scale CPS should be safe.
(10) *Sustainability Support*: The CPS application development supported by middleware should be sustainable.

4. MIDDLEWARE ARCHITECTURAL DESIGN

4.1. General Layered Architecture

We have studied various architecture present in the literature and proposed the generic layered architecture along with the functions as shown in Figure 1.

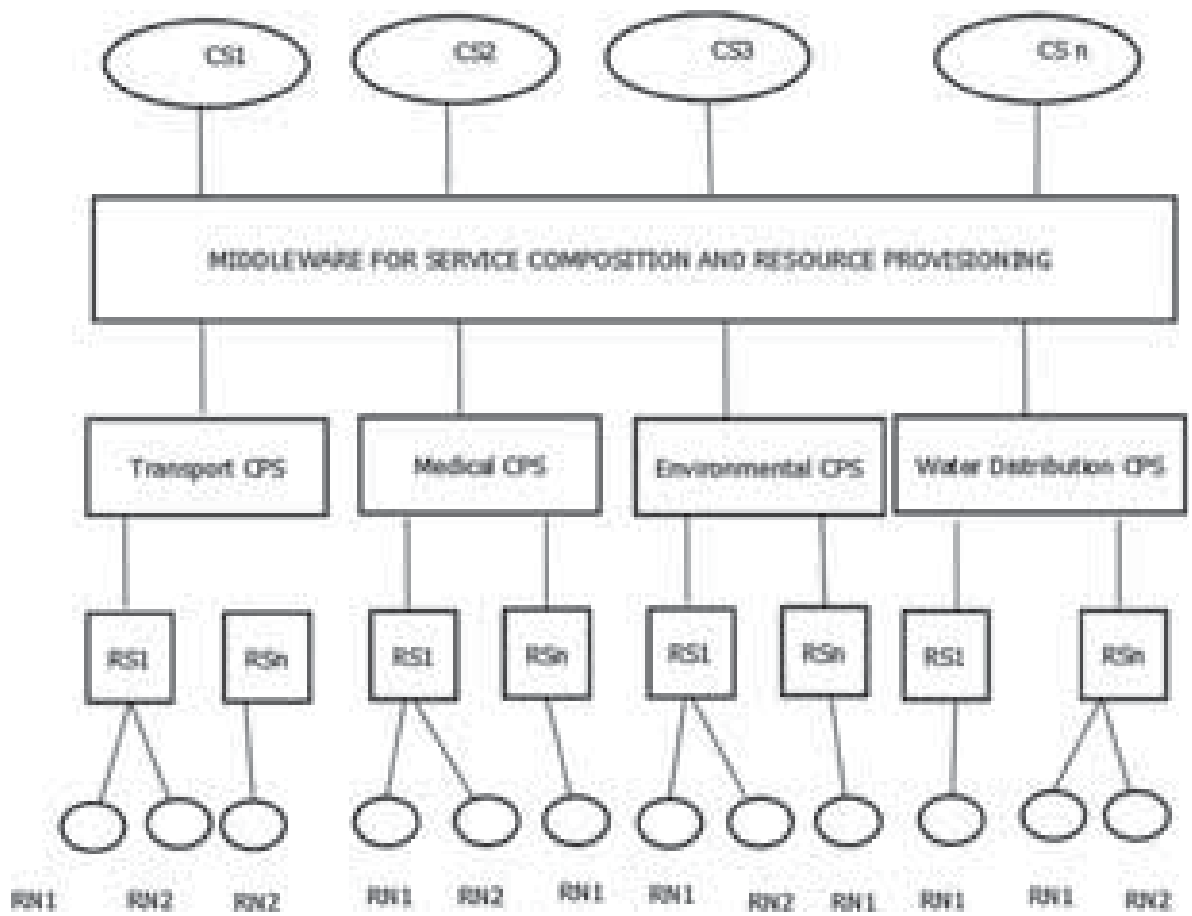


Fig. 3. Middleware for service composition and resource provisioning.
Notes: CS: Composite Service, RS: Resource System, RN: Resource Node.

RESEARCH ARTICLE

4.2. Architectural Block Diagram

Our work thus focuses on designing and implementing a middleware which will perform service composition on already existing CPSs. Also it considers challenges in resource provisioning along with their individual issues as well as integration issues. Service composition[1, 2] has different phases like service discovery, service ranking, best service selection, service deployment and service execution. The different components of middleware are shown in Figure 1 along with their interaction with other components is shown in Figure 2.

4.3. Working

The working of the middleware is as follows.
(1) The proposed middleware communicates with different CPS system under which various Resource systems are registered and in turn Resource system is responsible for managing all the resources. Thus CPS model is generated.
(2) As per the requirement for the given request of service composition, it is decomposed into task list.
(3) Task dependency is resolved.
(4) For each task, the services are discovered.
(5) The services are ranked as per our ranking algorithm.
(6) Then the best service is selected. While doing this QoS attributes of service as well as resources are taken into consideration.
(7) To deploy the service, a best Resource Node is also found with the help of Resource Provisioner, which uses Resource Virtualization technique. As resources are virtualized so the multiple virtual copies are also available in the middleware for resource provisioning.
(8) Then a best composition plan is selected.
(9) Service dependencies are also resolved.
(10) Services are scheduled and then executed.

4.4. Case Study

Here our Middleware as shown in Figure 3 communicates with different domain specific CPS systems like Transport CPS (T-CPS), Medical CPS (M-CPS), Environment CPS (E-CPS), Water Distribution CPS (W-CPS). Under such CPS systems, various Resource systems are registered and in turn Resource system is responsible for managing all the Resource Nodes. Sufficient literature is found on Resource Virtualization. It has many advantages which are as follows.[22, 23]
(1) Precise information is available through limited number of sensors.
(2) Reduced energy consumption and hence reduced cost of network.
(3) Increased availability of data in turn affects reliability.
(4) Multi tenancy of service so increase in number of applications.
(5) Increased flexibility in network administration, reconfiguration and scalability.

5. CONCLUSION

We have studied various middlewares along with their challenges. We elaborated expectations from middleware. Then we have presented layered architecture along with our middleware design. Our ongoing work involves design and implementation of above discussed middleware for service composition and resource provisioning. Under that, a set of algorithms are developed and implemented in simulated environment.

Acknowledgment: I am thankful to Dr. D. Y. Patil Institute of Technology for facilitating in all aspect to carry out this research.

References

1. Swati Nikam and Rajesh Ingle, Resource provisioning algorithms for service composition in cyber physical systems, *IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI)* **(2014)**, pp. 2797–2802.
2. Swati Nikam and Rajesh Ingle, *International Journal of Computer Science and Information Security* 15, 192 **(2017)**.
3. Jameela Al-Jaroodi and Nader Mohamed, *Concurrency and Computation: Practice and Experience* 24, 1919 **(2012)**.
4. A. Padilla, F. Javier, F. Weis, and J. Bourcier, Towards a model @ runtime middleware for cyber physical systems, *Proceedings of the 9th Workshop on Middleware for Next Generation Internet Computing* **(2014)**.
5. M. Kirill and A. Gul, Building portable middleware services for heterogeneous cyber-physical systems, *3rd International Workshop on Software Engineering for Sensor Network Applications, SESENA Proceedings* **(2012)**, pp. 31–36.
6. Dabholkar Akshay and Gokhale Aniruddha, An approach to middleware specialization for cyber physical systems, *29th IEEE International Conference on Distributed Computing Systems Workshops* **(2009)**, pp. 73–79.
7. I. Calvo, I. Etxeberria-Agiriano, and A. Noguero, Distribution middleware technologies for cyber physical systems, *9th IEEE International Conference on Remote Engineering and Virtual Instrumentation (REV)* **(2012)**, pp. 1–4.
8. I. Etxeberria-Agiriano, I. Calvo, A. Noguero, and E. Zulueta, Configurable cooperative middleware for the next generation of CPS, *9th IEEE International Conference on Remote Engineering and Virtual Instrumentation (REV)* **(2012)**, pp. 37–42.
9. M. Franke, C. Seidl, and T. Schlegel, A seamless integration, semantic middleware for cyber-physical systems, *10th IEEE International Conference on Networking, Sensing and Control (ICNSC)* **(2013)**, pp. 627–632.
10. K. Kyoung-Dae, Middleware and Control of Cyber-Physical Systems: Temporal Guarantees and Hybrid System Analysis, Ph.D. Thesis, University of Illinois at Urbana-Champaign **(2011)**.
11. Ankit Mundra, Bhagvan K. Gupta, Geetanjali Rathee, Meenu Chawla, Nitin Rakesh, and Vipin Tyagi, Validated Real Time Middle Ware for Distributed Cyber Physical Systems Using HMM, arXiv preprint arXiv:1304.3396 **(2013)**.
12. Y. Zhang, C. Gill, and C. Lu, Reconfigurable real-time middleware for distributed cyber-physical systems with aperiodic events, *28th International Conference on Distributed Computing Systems (ICDCS)* **(2008)**, pp. 581–588.
13. S. O. Park, J. H. Park, and Y.-S. Jeong, *Mobile Networks and Applications* 18, 110 **(2013)**.
14. Nader Mohamed, Jameela Al-Jaroodi, Sanja Lazarova-Molnar, and Imad Jawhar, *Scalable Computing: Practice and Experience* 18, 331 **(2017)**.

15. A. Cuzzocrea, J. Cecilio, and P. Furtado, An effective and efficient middleware for supporting distributed query processing in large-scale cyber-physical systems, *International Conference on Internet and Distributed Computing Systems*, Springer, Cham **(2014)**, pp. 124–135.
16. Nader Mohamed, Sanja Lazarova-Molnar, Imad Jawhar, and Jameela Al-Jaroodi, Towards service-oriented middleware for fog and cloud integrated cyber physical systems, *IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, IEEE **(2017)**, pp. 67–74.
17. Jameela Al-Jaroodi and Nader Mohamed, *Journal of Network and Computer Applications* 35, 211 **(2012)**.
18. G. Denker, Nikil Dutt, Sharad Mehrotra, M.-O. Stehr, C. Talcott, and Nalini Venkatasubramanian, *Journal of Internet Services and Applications* 3, 41 **(2012)**.
19. M. García-Valls, C. Calva-Urrego, A. Juan, and A. Alonso, *Computer Standards and Interfaces* 51, 95 **(2017)**.
20. A. Noguero, I. Calvo, and L. Almeida, A time-triggered middleware architecture for ubiquitous cyber physical system applications, *International Conference on Ubiquitous Computing and Ambient Intelligence*, Springer **(2012)**, pp. 73–80.
21. Mudasser Iqbal and H. B. Lim, A cyber-physical middleware framework for continuous monitoring of water distribution systems, *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, ACM **(2009)**, pp. 401–402.
22. D. Himansu, BighnarajNaik, BibudenduPati, and C. R. Panigrahi, *International Journal of Grid Distribution Computing* 7, 121 **(2014)**.
23. Imran Khan, FatnaBelqasmi, RochGlitho, Noel Crespi, Monique Morrow, and Paul Polakos, *IEEE Communications Surveys and Tutorials* 18, 553 **(2016)**.

Received: 24 April 2018. Accepted: 9 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1706–1711, 2018

RESEARCH ARTICLE

Service Composition Issues in Cyber Physical Systems and Middleware Design

Swati Nikam[1,*] and Rajesh Ingle[2]

[1] *Department of Computer Engineering, Dr.D.Y.Patil Institute of Technology (DIT), Affiliated to Savitribai Phule Pune University, Pimpri, Pune 411018, Maharashtra, India*
[2] *Department of Computer Engineering, Pune Institute of Technology (PICT), Affiliated to Savitribai Phule Pune University, Pune 411043, Maharashtra, India*

Today Cyber Physical Systems (CPS) has spanned through all Smart Applications of day to day life. These applications range from Smart Transportation, Smart Water Management, Smart Grid, Smart Office, Smart Home Automation and many more. The services provided by CPS applications is making our life easy but in comparisons with software services provided by Service Oriented Architecture (SOA), the services provided by CPS are different in many aspects. As there is a tight coupling of cyber and physical space so while defining and describing the CPS Service, we need to take into consideration the involvement and impact of physical space too. Besides that Service Composition is one of the major challenge present in Cyber Physical systems which by definition means to combine more than one services to achieve a specific task where individual services are not sufficient. So we need to consider provisioning of appropriate resource also in order to achieve the job of service composition. And while solving this Service Composition and Resource Provisioning challenges, we need to take care of the inherent challenges of CPS as well. In this paper we present issues in SC and then a middleware approach to resolve issues of Service Composition and Resource Provisioning challenge which considers Resource as well as Service simultaneously.

Keywords: Cyber Physical Systems, Resource Provisioning, Service Composition, Middleware.

1. INTRODUCTION

Cyber Physical system is the term coined by Gill in 2006 which is defined as Recently, Cyber Physical Systems (CPS) has attracted much research attention from both academia and industry. CPS applications are spanned through various domains like intelligent transportation, precision agriculture, Health CPS, water and mine monitoring, aerospace, smart buildings, structural health monitoring so on. In cyber physical systems, control, communication and computing in involved. It is highly heterogeneous in terms of devices, computing platforms and interconnection technology. There are some inherent characteristics of CPS in itself like unreliable networking (as it can be wired/wireless network, WLAN, Bluetooth, GSM, etc. are distributed systems), tight environmental coupling, close integration of computation and physical process, complexity at multiple temporal and spatial scale, dynamic operating condition etc. To understand CPS one should be familiar with basics of Wireless sensor network, Communication technologies, embedded systems and artificial intelligence. Currently many researchers are working on different research challenges present in CPS, out of which Service Composition is a significant one.

2. LITERATURE REVIEW

Service composition means to combine more than one services to perform a specific task which can not be performed by individual service. Many researchers have identified this challenge and worked towards it. Service computing technologies have been widely applied to many application domains to facilitate rapid system integration for desired goals.[23] Tremendous Literature is present on Service Oriented Computing (SOC) and Service oriented Architecture (SOA). As we know SOA promotes to provide software functionalities as interoperable service with their interface defined as protocols. Already Service Composition is very well studied in Web Service, Cloud computing, Wireless sensor network domain by

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1706/006 doi:10.1166/jctn.2018.7365

many. The existing techniques like UDDI, OWL-S, SOAP, WSDL etc. make it convenient to understand the process of service composition. But the characteristics, requirement, context of each above domains are different, so it may not be directly applied in the context of CPS. Based on the scope of service composition it can be categorized as

(1) Service Composition within CPS and
(2) Service Composition in Networked CPS.[1–3]

Before understanding the issues of service composition algorithms first we focus on the difference between Software service and CPS service as shown in Table I.

Because of requirement of different treatment needed by CPS Service and because of inherent characteristics of CPS domain itself, SOA model may not be readily applicable in CPS modeling, composition and management of CPS service.

Also while combining these CPS services it needs to consider resource as well, so along with SC we need to take care that, the required resources are also provisioned for successful execution of services. So we have considered service composition and resource provisioning as the focus of our work. CPS is inherently dynamic, nondeterministic in nature, so service composition needs more exploration.[4] Different phases of Service composition are as follows. Many of them have worked in one or more than one aspect of service composition which can be summarized as follows.

Service composition phases	Literature
Service description	[17]
Service naming	[22]
Service discovery	[19][20][21]
Service selection	[19][20][21]
Service deployment	[20]
Service execution	[21]

Apart from this Kaiyu[5] talk about Component model based composition mechanisms like Monolithic, Object Oriented, Application Oriented, Component Based, Service Oriented. Huges[6] have presented Composition Challenges as Compensability and Composition ability.

Huanag[7] have presented an ontology model for physical entity representation and used Artificial Intelligence planning for service composition. Bastani[8] have presented an efficient framework for service composition which uses 2 level composition first at abstract level and second at context level.

Son et al.[9] have considered context in the SC. Kim[10] have presented a semantic framework for reconfiguration of instrumented Cyber Physical Spaces in which they have considered an event model. Meng[11] has focused on compositional modeling and composition reasoning of QoS properties.

Table I. Difference between software service and CPS service.

Aspects	Software service	CPS service
Dependency on physical environment	Are not dependent on physical environment so does not impact on availability of service and in turn QoS	Highly dependent on physical environment so impacts on availability and QoS
Multiple instances	Can have multiple instances	Can have multiple instances but only one will be served at a time.
Contention problem	Contention problem may not occur	Contention problem occurs because only one physical service can be served at a time
Context consideration	Need not take into consideration the context while describing service.	Contexts needs to be considered while describing the service.
Predictability of service execution	More predictable in terms of successful service execution	Less predictable in terms of successful service execution because of joint dynamics involved in cyber and physical space
Real time	Need not require real time consideration	Requires real time consideration
Resource consideration	No physical resources are considered as a part of definition	Physical resources needs to be considered
Current approaches	SOA	PE-SOA,[21] Extended OWLS[8]

Pascal[12] have presented a group based programming abstraction for CPS which helps to group various sensors and actuators. It facilitates feedback control mechanisms for dynamic group membership update and requirements based on feedback from the current mechanisms. It has also compared other grouping abstractions like Hood, Abstract region, Logical Neighbours, Scope etc. Swaroop[13] have presented dynamic service composition based on graph theory and service repository so dynamically a complex service can be delivered.

Liu[14] have presented resource aggregation and dynamic service composition framework in manufacturing domain based on runtime adjustment strategies. Franke[15] have presented an approach where seamless integration of devices is possible.

Peng[16] have carried out composition analysis of components and also formal verification based on service oriented architecture is discussed. Service composition also has to consider the *scalability issues* as there is a need to check that, considering the context of CPS whether service can be extended or not.[17]

Yajing[18] have discussed about Collaboration Problem amongst Multiple CPS where they have extended OWL-S

RESEARCH ARTICLE

framework to address the functionalities. They have used the *Abstract Service and Concrete service Representation*. The idea behind Abstract service is, as there are normally more than one physical systems of the same model, which typically provide the same set of service so represent them using Abstract service and at the time of execution.

Here authors[19] have presented efficient *context-aware service composition framework* along with algorithm of atomic service filtering algorithm.

Service composition problem is discussed as two phase context sensitive service composition optimization problem[20] which is solved using Particle Swarm Optimization Method. They have also proposed PE-SOA ontology model for the same along with case study discussion of traffic accident rescuing task.

Huang[21] have extended the conventional SOA model, where PE-ontology is used to connect different physical entities based on their capabilities (services they can provide) and PE-SOA model helps in service specifications that are suitable for physical entities. Thus it simplifies PE service specifications and reduces the complexity of the reasoning procedure.

Authors[22] have given details about name centric service architecture for CPS. Services register to the service bus with their local daemon and with URN names. This can simplify an orchestration of higher-level services that can be combined from basic services distributed in the CPS. But the challenge is, using named services requires a consistent naming system within the CPS.

Authors[23] have proposed improvement in resolution process which means binding of abstract service with concrete service. They have proposed profile based execution and consistency checker using the detector and matchmaking mechanism.

Authors[24] have proposed event driven SOA where devices are exposed to application developer as a web service. Also they have developed special purpose event handling services. They have reviewed SSCL (specialized service composition language) and also discussed Python and Spreadsheet approach for event driven SOA.

3. ISSUES IN SERVICE COMPOSITION

From the above literature study, the issues of service composition can be listed which are as follows.

(1) Most of the papers have focused on Service Oriented Architecture which works well for web service but in the context of CPS it needs to consider resources simultaneously.

(2) Also the methods provides a particular domain specific service composition but there is lack of generalized solution which can work ubiquitously.

(3) One very important point is service alone should not be considered, but has to be considered together with service, resource, context of the cyber and physical environment and dependency between resources and service.

(4) A centralized approach for service composition which faces problem from scalability point of view.

(5) The work is limited to CPS, but Networked CPSs are not considered.

(6) Researchers have discussed few phases of service composition but nobody has discussed all the phases of service composition in CPS context, it is revolving around Wen service, Wireless Sensor Network only.

(7) There is a limited work found on methodology details in the context of service composition, most of them have discussed conceptual work or partial details.

4. MIDDLEWARE DESIGN FOR CPS

4.1. General Layered Architecture

We have studied various layered architecture present in the literature and proposed the generic layered architecture along with the functions as shown in Table II.

4.2. Architectural Block Diagram

Our work thus focuses on designing and implementing a middleware which will perform service composition on already existing CPSs and for that the required resource provisioning needs to be considered along with their individual issues as well as integration issues. Service composition[25, 26] has different phases like service discovery, service ranking, best service selection, service deployment and service execution. The different components of middleware are shown in Figure 1 along with their interaction with other components is shown in Figure 2.

4.3. Working

The working of the middleware is as follows.

(1) The proposed middleware communicates with different CPS system under which various Resource systems are

Table II. Layered architecture.

Layer	Functions/Algorithms	Components
Application layer	Seamless service to user	User interface through web or application
Service layer	Task dependency, Service discovery, Service ranking, Service selection, Service composition	CPS middleware
Control layer	CPS model generation	
Context processing layer	Resource system comparator	CPS system
Resource layer	Resource provisioning, Resource node comparator, Service dependency resolution, Service deployment, Service execution	Resource system
	Task execution	Resource node

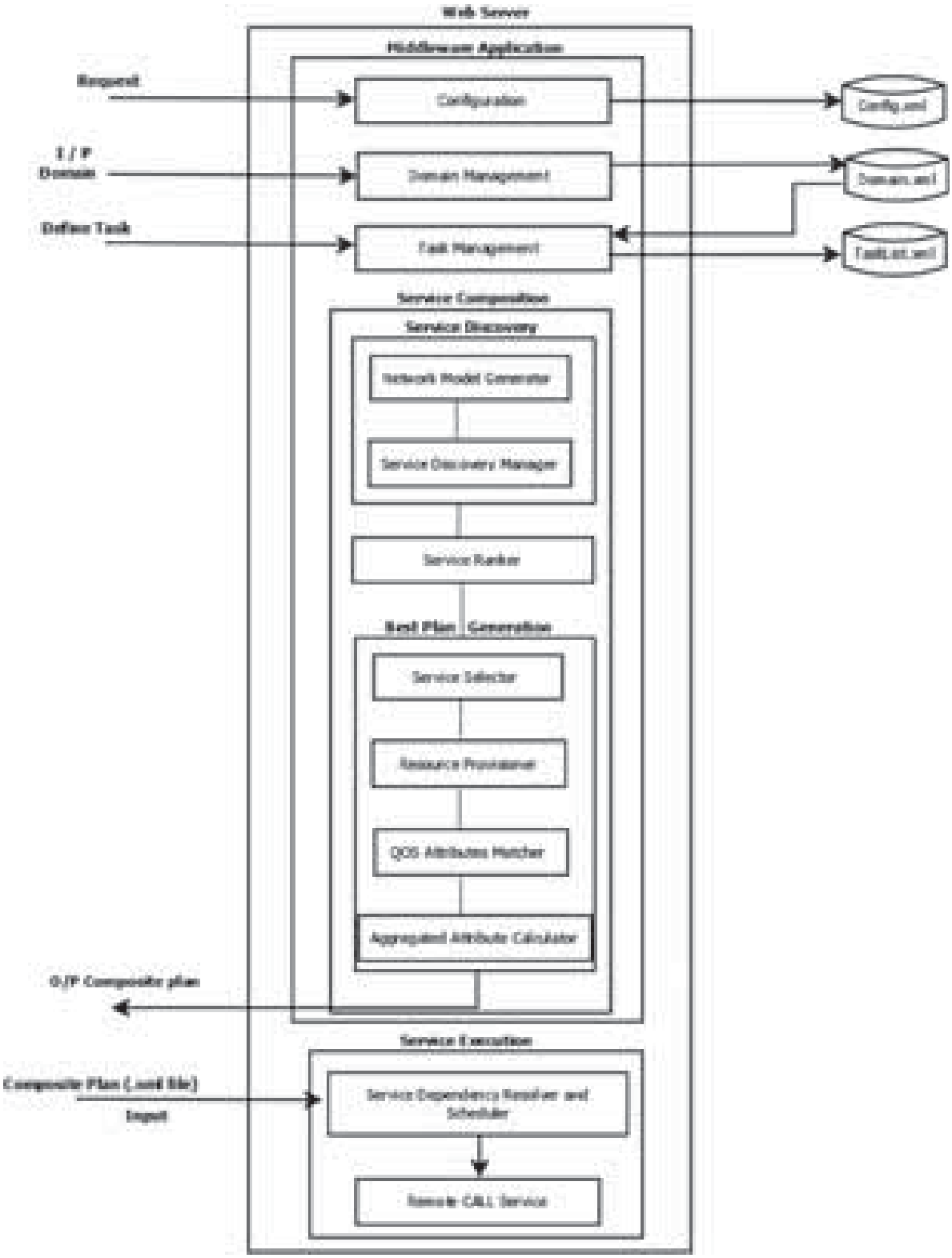


Fig. 1. Architectural block diagram of middleware for service composition and resource provisioning.

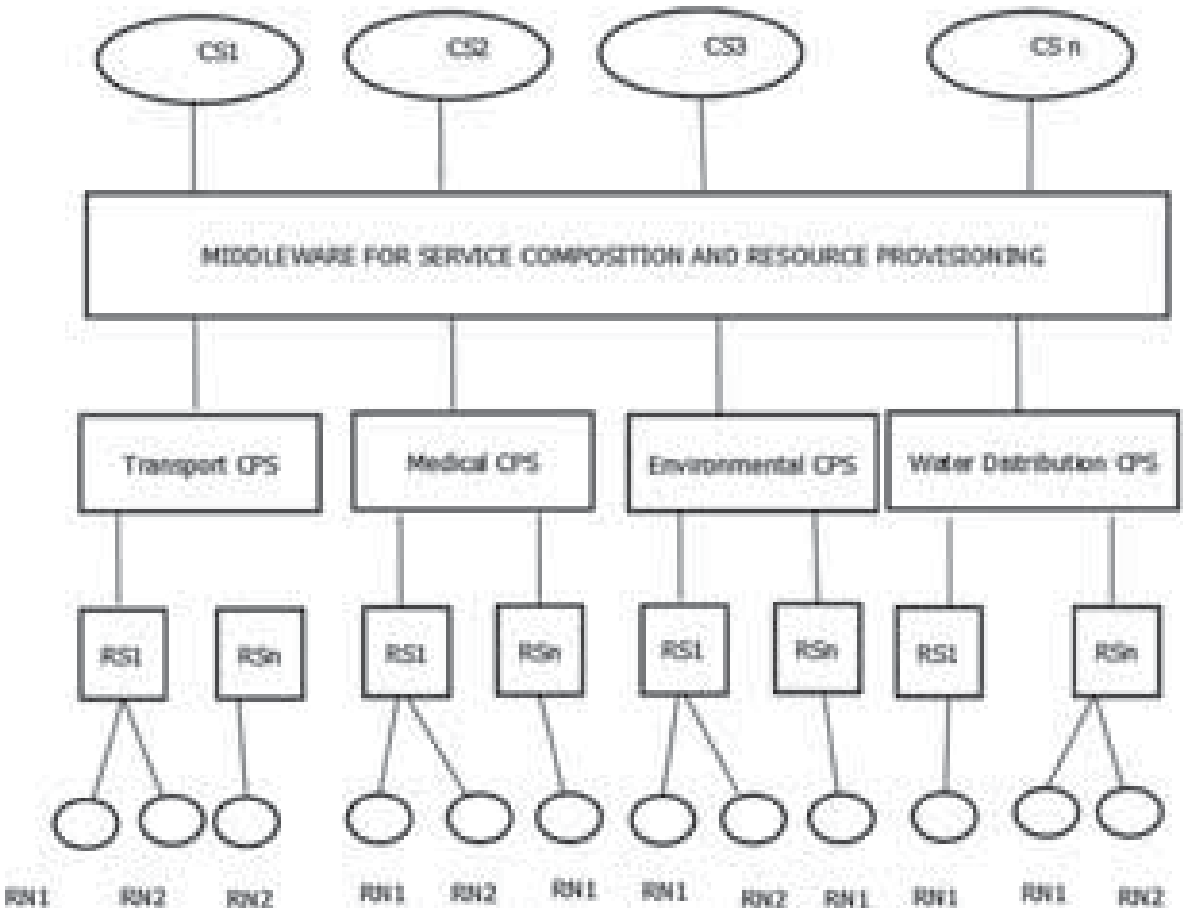


Fig. 2. Middleware for service composition and resource provisioning. PS = Physical sensors, VS = Virtual sensors, T-CPS = Transport CPS, M-CPS = Medical CPS, E-CPS = Environmental CPS, W-CPS = Water Distribution CPS, RS = Resource System, RN = Resource Node.

registered and in turn resource system is responsible for managing all the resources. Thus CPS model is generated.
(2) As per the requirement for the given request of service composition, it is decomposed into task list.
(3) Task dependency is resolved.
(4) For each task, the services are discovered.
(5) The services are ranked as per our ranking algorithm.
(6) Then the best service is selected. While doing this QoS attributes of service as well as resources are taken into consideration.
(7) To deploy the service, a best Resource Node is also found with the help of Resource Provisioner, which uses Resource Virtualization technique. As resources are virtualized so the multiple virtual copies are also available in the middleware for resource provisioning.
(8) Then a best composition plan is selected.
(9) Before executing the services as per the Best Composition Plan, service dependencies are also resolved and services are scheduled.

4.4. Case Study

Here our Middleware communicates with different domain specific CPS systems like Transport CPS (T-CPS), Medical CPS (M-CPS), Environment CPS (E-CPS), Water Distribution CPS (W-CPS). Under such CPS systems, various Resource systems are registered and in turn Resource system is responsible for managing all the Resource Nodes. We assume can be utilised fully for resource provisioning. All the resources are made available with the help of virtualization techniques. Sufficient literature is found on Resource Virtualization. It has many advantages which are as follows.[27, 28]
(1) Precise information is available through limited number of sensors.
(2) Reduced energy consumption and hence reduced cost of network.
(3) Increased availability of data in turn affects reliability.
(4) Multi tenancy of service so increase in number of applications.
(5) Increased flexibility in network administration, reconfiguration and scalability.

5. CONCLUSION

We have discussed various issues in Service Composition. Our ongoing work involves design and implementation of above discussed middleware for service composition and resource provisioning. Under that, a set of algorithms are developed and implemented and also validated in actual and simulator environment.

We have discussed the various issues in Service Composition in the Context of Cyber Physical Systems. Further we have also discussed that how middleware approach can help to solve these issues. Our future work includes the validation and implementation of this middleware in real time.

Acknowledgment: I am thankful to Dr. D. Y. Patil Institute of Technology for being facilitator to carry out this research.

References

1. M. Kim, M.-O. Stehr, and C. Talcott, *Science of Computer Programming* 78, 2453 **(2013)**.
2. Moreno, J. Damm, M. Haase, J. Grimm, C. Holleis, and Edgar, *Forum on Specification and Design Languages (FDL)* 8, 68 **(2012)**.
3. M. Kim, M.-O. Stehr, J. Kim, and S. Ha, An application framework for loosely coupled networked cyber-physical systems, *8th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC)* **(2010)**, pp. 144–153.
4. L. E. Ashford and S. A. Seshia, *Introduction to Embedded Systems: A Cyber-Physical Systems Approach*, MIT Press **(2016)**.
5. K. Wan, D. Hughes, K. Man, T. Krilavicius, and S. Zou, *International Journal of Design, Analysis and Tools for Circuits and Systems* 2, 30 **(2010)**.
6. K. Wan, D. Hughes, Ka Lok Man, and Tomas Krilaviçius, Composition challenges and approaches for cyber physical systems, *IEEE International Conference on Networked Embedded Systems for Enterprise Applications (NESEA)* **(2010)**, pp. 1–7.
7. J. Huang, F. Bastani, I. L. Yen, and W. Zhang, A framework for efficient service composition in cyber-physical systems, *Fifth IEEE International Conference on Service Oriented System Engineering (SOSE)* **(2010)**, pp. 291–298.
8. J. Huang, F. Bastani, and I.-L. Yen, An efficient framework for service composition for CPS, *International Conference on Information Technology, Computer Engineering and Management Sciences* **(2011)**.
9. S. N. Han, G. M. Lee, and N. Crespi, Context-aware service composition framework in web-enabled building automation system, *6th International Conference on Intelligence in Next Generation Networks (ICIN)* **(2012)**, pp. 128–133.
10. K. Minyoung, D. Massaguer, N. Dutt, S. Mehrotra, S. Ren, M.-O. Stehr, C. Talcott, and N. Venkatasubramanian, A semantic framework for reconfiguration of instrumented cyber physical spaces, *Workshop on Event-Based Semantics, CPS Week* **(2008)**.
11. S. Meng, Challenges on co-ordination for CPS, *Proceedings of second International Symposium on Computer, Communication, Control and Automation (ISCCCA)* **(2013)**.
12. P. A. Vicaire, E. Hoque, Z. Xie, and J. A. Stankovic, Bundle: A group-based programming abstraction for cyber-physical systems, *IEEE Transactions on Industrial Informatics* 8, 379 **(2012)**.
13. K. Swaroop, M. Kumar, and B. A. Shirazi, Dynamic service composition in pervasive computing, *IEEE Transactions on Parallel and Distributed Systems* 18, 907 **(2007)**.
14. L. Weining and J. Su, A solution of dynamic manufacturing resource aggregation in cps, *6th IEEE Joint International on Information Technology and Artificial Intelligence Conference (ITAIC)* **(2011)**, pp. 65–71.
15. M. Franke, C. Seidl, and T. Schlegel, A seamless integration, semantic middleware for cyber-physical systems, *10th IEEE International Conference on Networking, Sensing and Control (ICNSC)* **(2013)**, pp. 627–632.
16. P. Wang, Y. Xiang, and S. H. Zhang, Cyber-physical system components composition analysis and formal verification based on service-oriented architecture, *Ninth IEEE International Conference on e-Business Engineering (ICEBE)* **(2012)**, pp. 327–332.
17. Kumar Padmanabh, *Journal of Indian Institute of Science* 93, 499 **(2013)**.
18. Z. Yajing, Abstract Cyber Physical Systems Service Composition, Service Life Cycle Tools and Technologies, Methods, Tends and Advances, Jonathan Lee, Shan Ma, Alen Liu, IGI Global **(2012)**, Chap. 14, pp. 2012–2014.
19. T. Wang, L. Cheng, and K. Zhang, *Advances in Information Sciences and Service Sciences (AISS)* 4, 11 **(2012)**.
20. W. Tao, A two-phase context-sensitive service composition method with a workflow model in cyber-physical systems, *Proceedings-17th IEEE International Conference on Computational Science and Engineering, CSE 2014, Jointly with 13th IEEE International Conference on Ubiquitous Computing and Communications, IUCC 2014, 13th International Symposium on Pervasive Systems* **(2015)**, pp. 1475–1482.
21. J. Huang, F. Bastani, I.-L. Yen, J. Dong, W. Zhang, F.-J. Wang, and H.-J. Hsu. Extending service model to build an effective service composition framework for cyber-physical systems, *IEEE International Conference on Service-Oriented Computing and Applications (SOCA)* **(2009)**, pp. 1–8.
22. H. Horst, T. Teubler, and S. Fischer, Name-centric service architecture for cyber-physical systems (short paper), *6th IEEE International Conference on Service-Oriented Computing and Applications (SOCA)* **(2013)**, pp. 77–82.
23. R. Alcarria, B. Bordel, and A. Jara, Flexible service provision in context-aware cyber-physical systems, *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, Springer, Cham **(2017)**, pp. 873–883.
24. S. Siniša, D. Škvor, and M. *Automatica* 53, 294 **(2015)**.
25. Swati Nikam and Rajesh Ingle, *International Journal of Computer Science and Information Security* 15, 192 **(2017)**.
26. Swati Nikam and Rajesh Ingle, *International Journal of Computer Science and Information Security* 15, 192 **(2017)**.
27. H. Das, B. Naik, B. Pati, and C. Rani Panigrahi, *International Journal of Grid Distribution Computing* 7, 121 **(2014)**.
28. I. Khan, F. Belqasmi, R. Glitho, N. Crespi, M. Morrow, and P. Polakos, *IEEE Communications Surveys and Tutorials* 18, 553 **(2016)**.

Received: 24 April 2018. Accepted: 9 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1712–1718, 2018

RESEARCH ARTICLE

Design and Analysis of Multiply and Accumulation Units Using Low Power Adders

R. Nirmal Kumar, S. Karthick, R. S. Valarmathi, and D. Rajesh Kumar*

Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam, Erode 638401, India

The prevalent blocks used in digital signal processing hardware are the adder, multiplier and delay elements. Better the performance of adder structure better will be the performance of multipliers in total aspect. Reducing power dissipation, delay and area at the circuit level is considered as one of the major factors in developing low power systems. In this paper we have introduced a new (i) 8 transistor (8T) full adder (ii) Proposed Shannon based (8T) adder using pass transistor logic which has better power, delay performance than the existing adders. Performance comparison of the proposed 8T adder has been made by comparing its performance with 10T SERF, 10T CLRCL, and the existing 14T full adders. The proposed 8T full adder structure has improved performance characteristics and suitable for Array, Carry Save and Dadda multipliers. Also three versions of 3 tap FIR filter namely Broadcast, Unfolded Broadcast, Unfolded and Retimed Broadcast structures have been implemented using three different multipliers. Each of these multipliers used for the Filters are implemented using all the existing full adders and the proposed 8T full adder. Results show that circuits implemented using proposed 8T full adder has better power, delay and cascaded performance when compared with the peer ones. All the simulations were carried out using TSMC Complementary Metal Oxide Semiconductor (CMOS) 120 nm technology file with a supply voltage of 1.8 V. Tools used are Tanner EDA tools.

Keywords: Array Multiplier, Carry Save Multiplier, FIR Filter Dadda Multiplier, Unfolding and Retiming, Tanner EDA Tools.

1. INTRODUCTION

This Full Adder is a basic block in all digital circuits. A small change in transistor count, power and delay will cause a drastic change in the performance of a large VLSI circuit. The performance of multipliers depends on the full adder used. The important parameters to be considered while designing a full adder are power consumption, delay, area, full swing operation and performance while cascading adders in a multiplier structure. Three existing low power full adder cells are 10 transistor SERF adder, 10 transistor CLRCL adder and 14 transistor full adder. The circuit diagram of 10T SERF adder[1] is shown in Figure 1. It can be seen that it has two four transistor xnor structures to perform the sum operation. Both these xnor does not have ground so it has very low power consumption since there is no direct path from V_{dd} to ground. The carry logic is generated by pass transistor logic. Though SERF adder[1] consumes less power it suffers from threshold loss problem as both sum and carries regenerated from pass transistor logic.

In order to rectify the defects in the SERF adder,[1] the 10T CLRCL[1] adder was introduced. The main aim of the design is that the carry signal must not suffer from distortion as it propagated. The circuit diagram of 10T CLRCL[1] is shown in Figure 2. It consists of two transistor xor structures. The inverters used prevent threshold loss problem and propagate the full swing signals to generate the carry. Though the carry signal has full swing operation, this circuit consumes more power. The circuit diagram of Existing 14T full adder[2] is shown in Figure 3. It has a four transistor xor structure and an inverter. The carry is generated using transmission gate logic and the sum is generated from pass transistor logic. The power consumed by this circuit is less when compared with that of 10T CLRCL full adder[2] and more when compared with 10T SERF full adder.[1] DSP Journal, Volume 9, Issue 1, June, 2009.

With the aim of further minimizing the number of transistors, pass transistor logic based XOR and XNOR circuits[2] were used and as a result the 14T full adder circuit of Figure 2 was designed. This circuit among all 14T full adder circuits[3, 4] shows the better results for delay

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1712/007 doi:10.1166/jctn.2018.7366

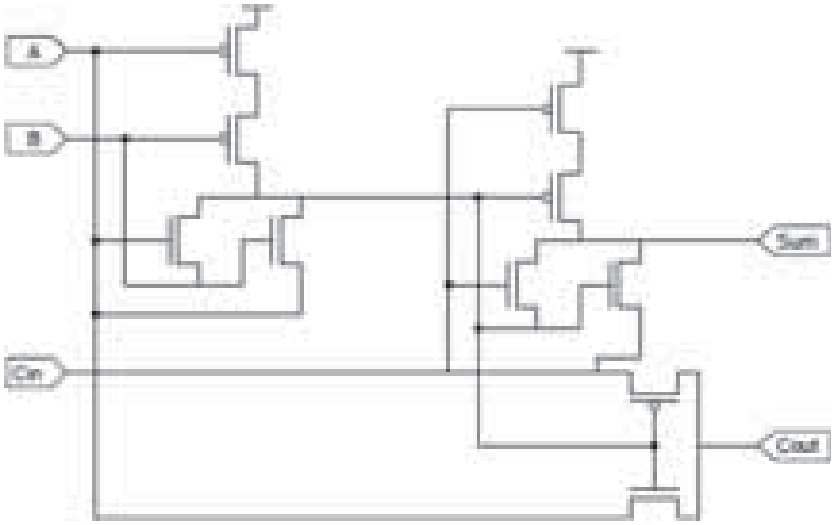


Fig. 1. Circuit diagram of 10T SERF full adder.

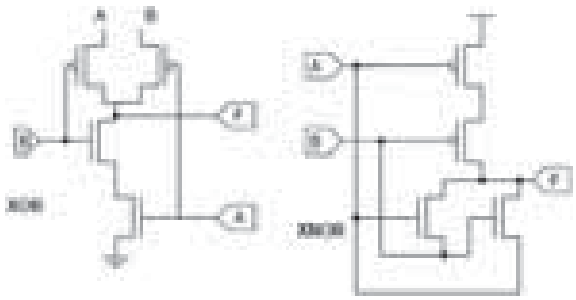


Fig. 2. Circuit diagram of XOR and XNOR structures.

and power as compared to 16T full adder but it suffers from the threshold loss problem of approximately 0.4 v. It works well in high performance multipliers with low power consumption. The further designed 10T full adder[5] of Figure 3 uses inverter-based 4T XOR gates in their design and shows remarkable improvements in power and delay. It also reduces the silicon area. This reveals better performance than the SERF 10T adder cell.[5] The drawback of this circuit is that it also suffers from threshold loss problem of 0.35 v approximately equal to 14T adder circuit.

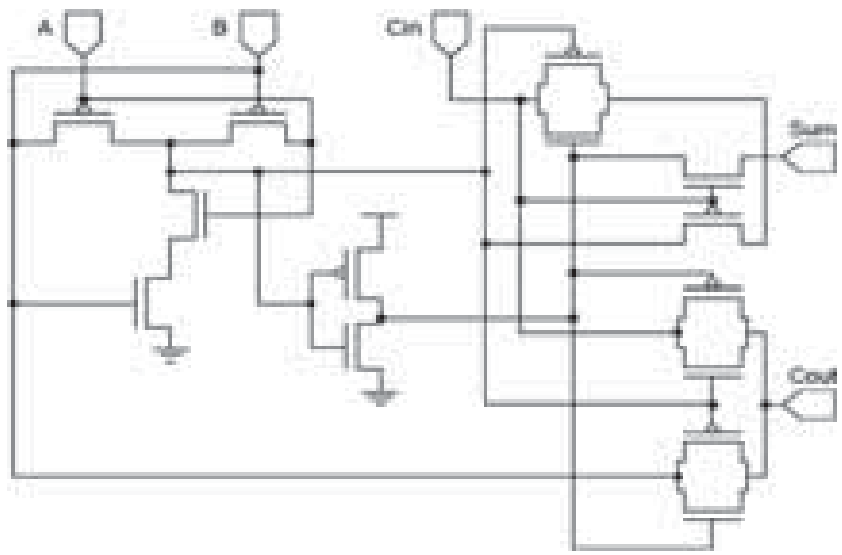


Fig. 3. Circuit diagram of full adder using T.G.

2. PROPOSED 8T FULL ADDER

This design of proposed full adder is based on three transistor XOR gates. It acquires least silicon area. The design of 3T XOR gate is shown in Figure 4. The heart of the design is based on a modified version of a CMOS inverter and a PMOS pass transistor.

The design of 8T full adder cell having the least number of transistors using 3T XOR gates is shown in Figure 5. The Boolean equations for the design of the 8T full adder are as follows.

This circuit shows approximately 45% improvement in threshold loss (0.2 v) as compared to other adders stated above. It is the fastest and consumes least power. Therefore we report it to be the best on account of power consumption, delay and threshold loss.

3. SIMULATIONS AND PERFORMANCE ANALYSIS

We have performed simulations using Tanner EDA tool in 90 nm and 130 nm technologies; with supply voltage ranging from 1 v to 2.6 v in steps of 0.2 v. To establish an impartial testing environment each circuit have been tested on the same input patterns. Comparative studies on the different adders found in literature have been done using 90 nm and 130 nm technologies. Studies have been done with 16T, 14T, 10T and the proposed 8T full adder cells. In our proposed adder circuit we have shown that reducing the aspect ratio of the PMOS transistors results in better performance while keeping the threshold loss constant. All paragraphs must be indented. All paragraphs must be justified, i.e., both left-justified and right-justified.

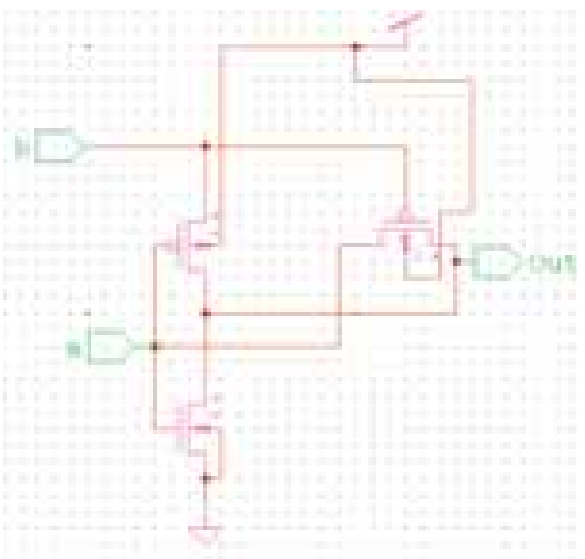


Fig. 4. Proposed XOR gate (3T).

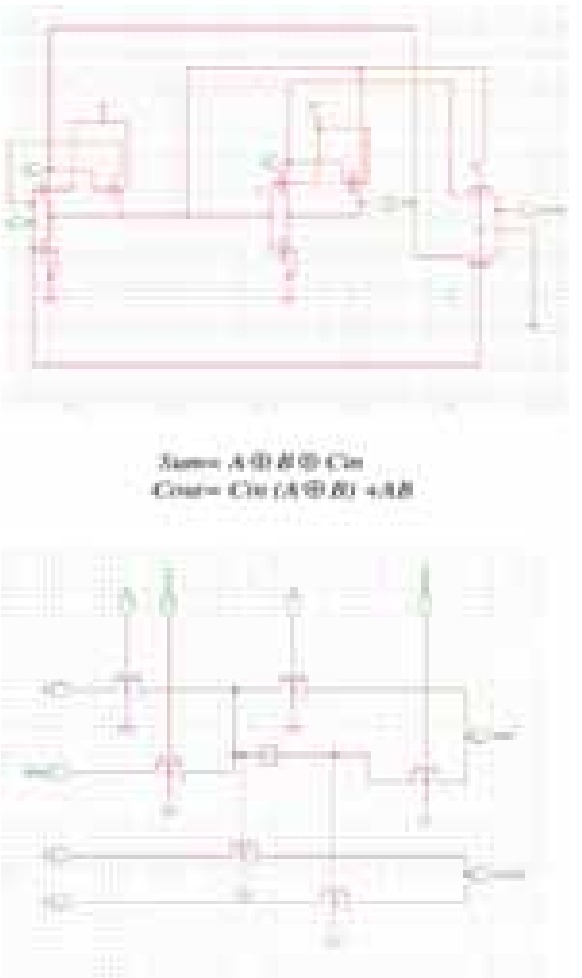


Fig. 5. Proposed (8T) adder.

The result of the comparative study shows that the performance of the proposed 8T full adder cell is the best among all. The 8T full adder cell also occupies the minimum silicon area on chip amongst all the full adders reported so far in the literature. The small silicon area of the proposed full adder cell makes it potentially useful for building compact VLSI circuits on a small area of chips. The threshold loss, delay and power dissipation of the proposed full adder cell is much less compared to any other adder. The net effect is that our proposed 8T full adder cell shows a much better performance compared to any other adders available in the literature. Figures 6–11 shows the comparative analysis of the circuits stated above at 90 nm and 130 nm technology. The simulation results reveal that the proposed 8T full adder is proven to be the best if the main design aspects of area covered on chip, threshold loss, and delay and power consumption are the ultimate goals.

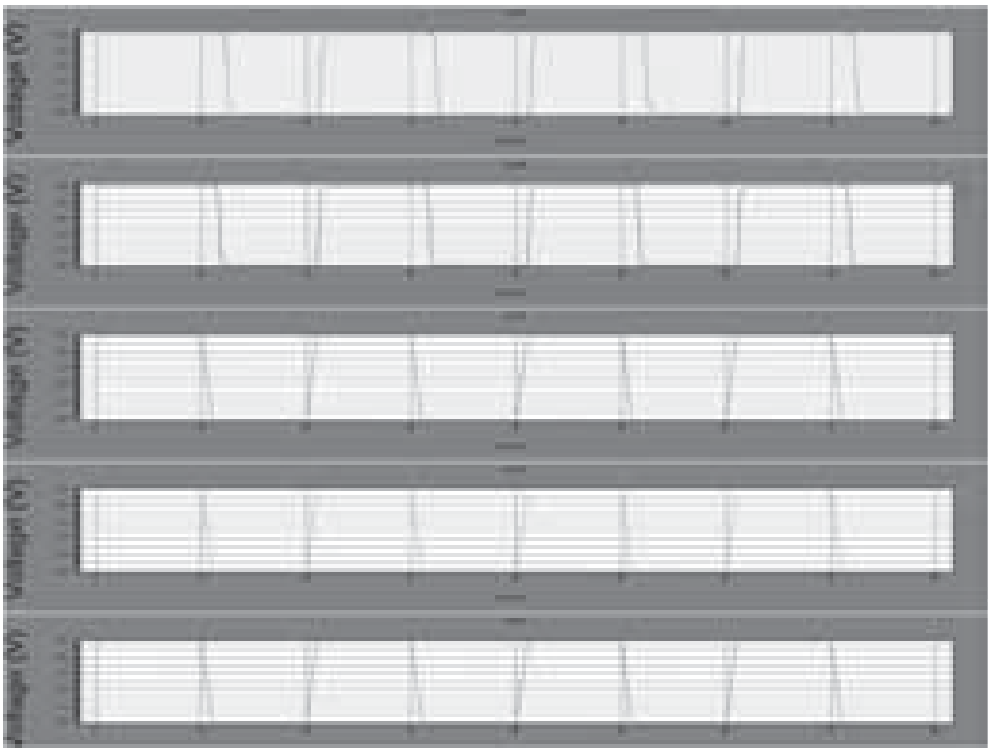


Fig. 6. Full adder output waveform.

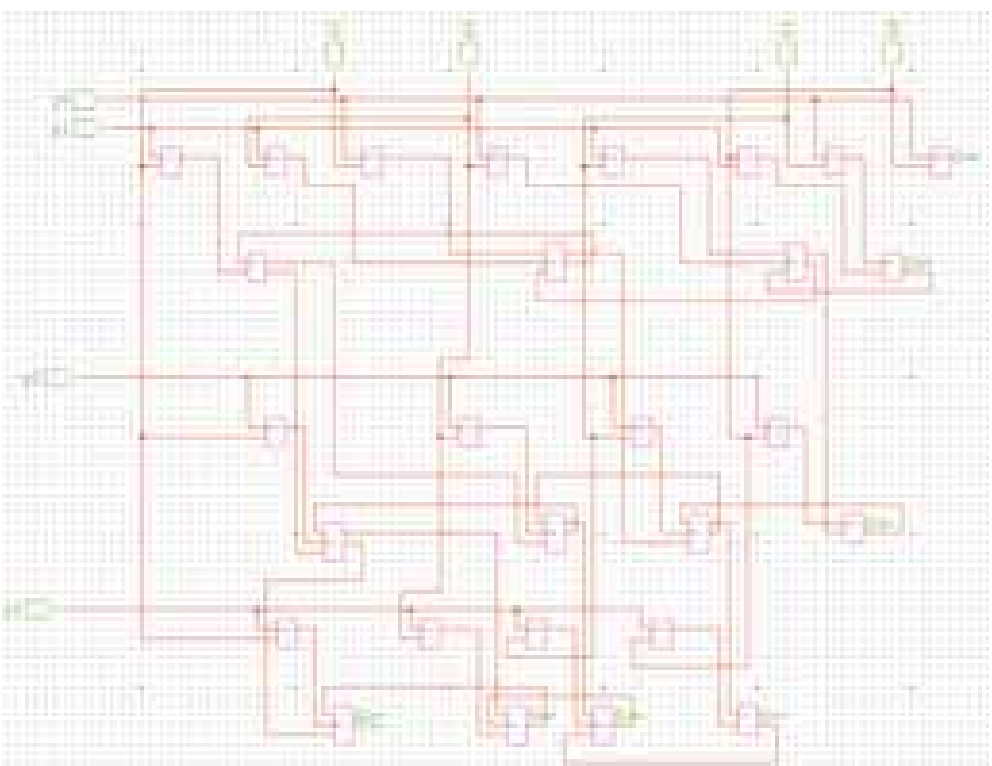


Fig. 7. Multiplier diagram.

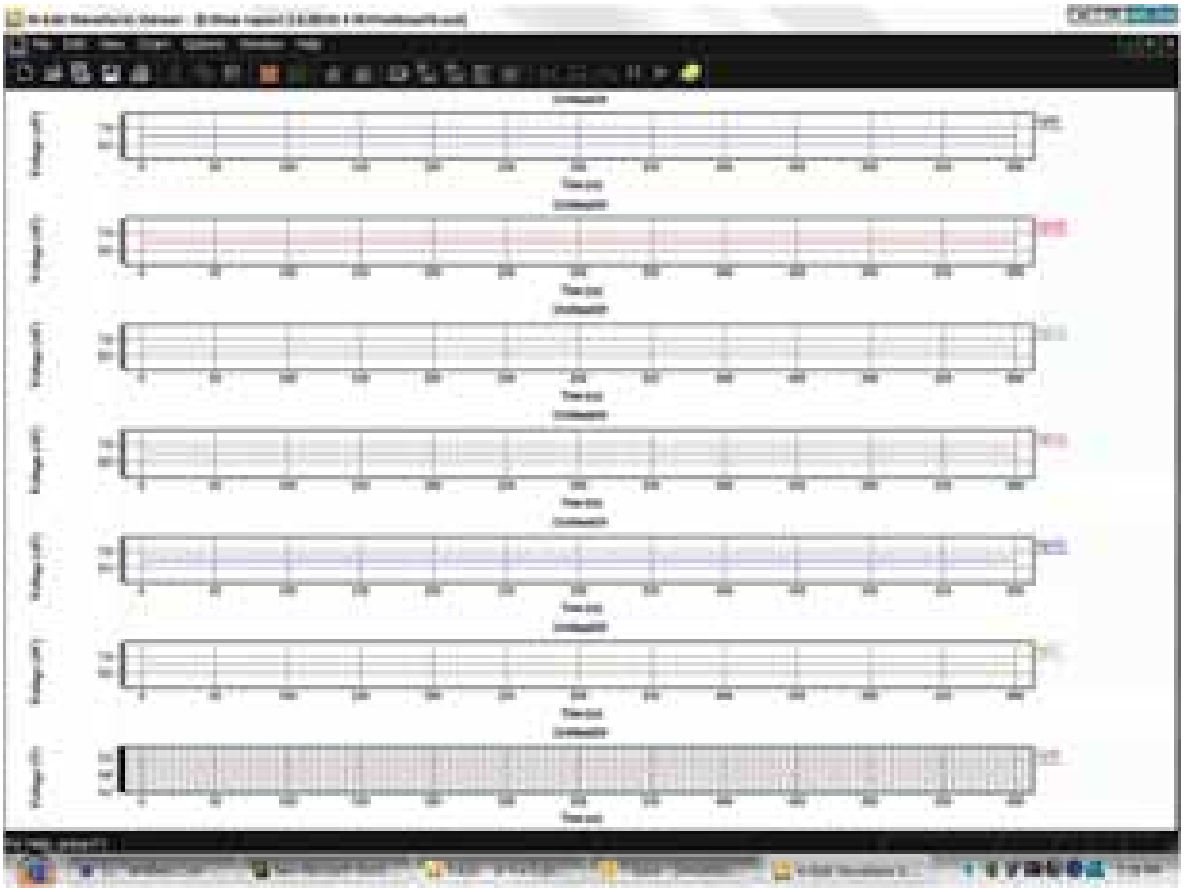


Fig. 8. Multiplier output waveform.

The Table I shows the power, delay and power-delay product comparison of three existing full adders and the proposed 14T full adder. It can be seen that 10T SERF has the least power and power-delay product but it suffers from severe threshold loss problem which leads to circuit malfunction when cascaded in larger circuits. The proposed 14T full adder has less power consumption and delay when compared with existing 14T and 10T CLRCL full adders. Figure 6 shows the output waveforms containing the sum and carry signals of all the full adders. The input pattern is varied from 000 to 111 with input changing every nano-second so as to calculate the delay from the wave form. It can be seen from the waveform that carry (ca) of 10T SERF suffers from threshold loss problem (3rd and 5th bit). This is the main drawback of 10T SERF full adder. In 10T CLRCL waveform the carry (ca) signal suffers from threshold loss problem also the width of the glitches show that it has slow operation. It can be seen that the proposed 14T and existing 14T full adders have similar waveform but proposed 14T adder is faster than existing 14T adder because there exists a glitch at 2ns in the sum signal of existing 14T adder. Also proposed 14T adder consumes less power than the existing 14T adder.

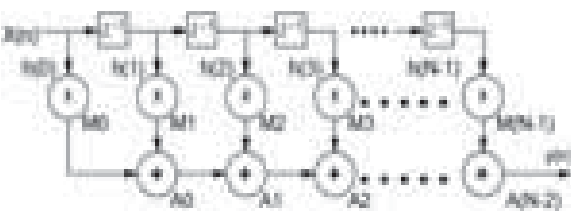


Fig. 9. FIR filter in direct form.

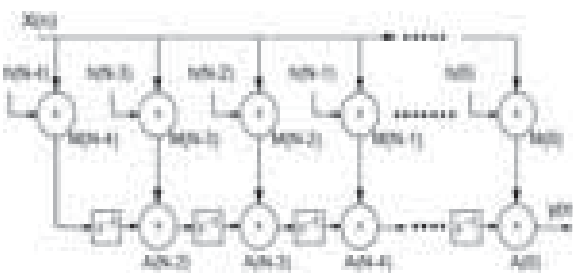


Fig. 10. FIR filter in direct form.

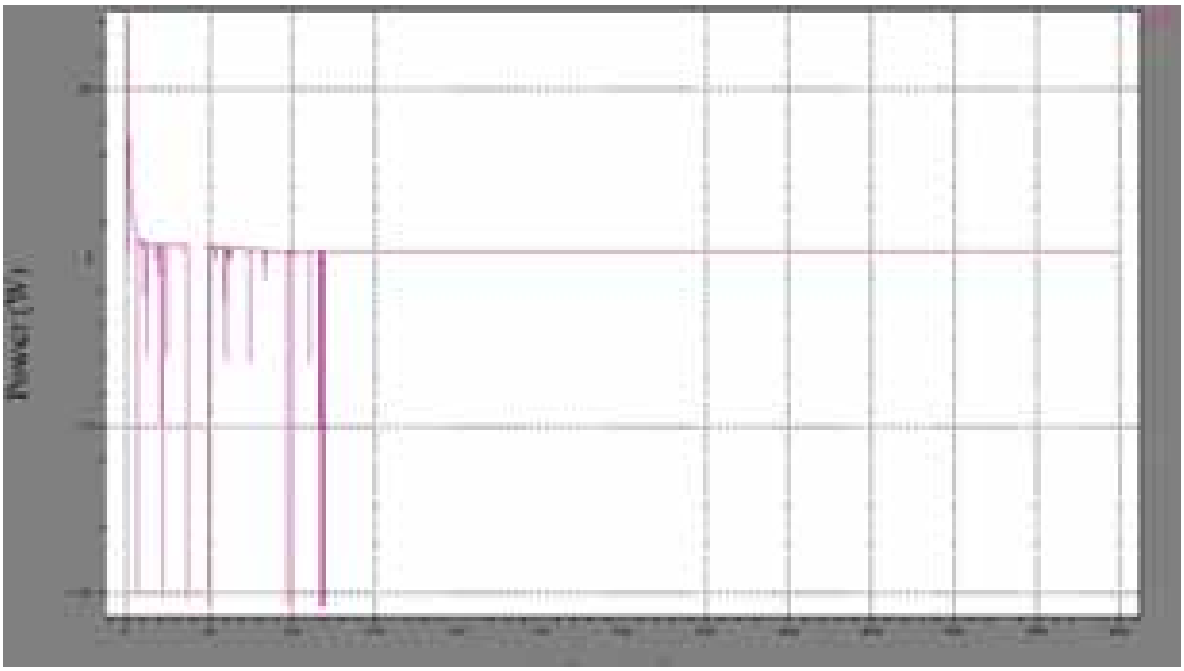


Fig. 11. Filter output waveform.

Table I. Comparison of power and delay of different full adders.

Full adder	Power (in μW)	Delay (in ns)	Power delay product (10^{-15} Ws)
10T CLRCL	21.074	0.1680	3.540432
10T SERF	7.990	0.1078	0.861322
EXISTING 14T	16.960	0.1399	2.372704
PROPOSED 8T	13.074	0.1013	1.3243962
PROPOSED 8T	8.12	0.1013	1.315296

Table II. Power, delay comparison of 4×4 array multipliers using different adders.

Full adder	Power (in μW)	Delay (in ns)	Power delay product (10^{-15} Ws)
10T CLRCL	72.562	0.459	33.305958
10T SERF	55.493	0.531	29.46678
EXISTING 14T	69.436	0.395	29.46678
PROPOSED 8T	65.665	0.350	27.42722
PROPOSED 8T	58.493	0.340	26.42722

Table III. Power, delay comparison of 4×4 carry save multiplier using different adders.

Full adder	Power (in μW)	Delay (in ns)	Power delay product (10^{-15} Ws)
10T CLRCL	388.42	1.611	625.74462
10T SERF	305.90	2.882	881.6038
EXISTING 14T	335.16	2.39	801.0324
PROPOSED 8T	333.02	1.328	442.25056
PROPOSED 8T	323.02	1.318	432.25056

4. MULTIPLIERS

In this paper three different multipliers are considered for analysis of the adders. The multipliers are the structures where there will be many cascading stages of the full adder, so the performance of the full adders while cascading to many stages can be easily studied by analyzing the power, delay, power-delay product of the different multipliers made from different adders. The multiplier structures that have been considered for analysis are Array, Carry-Save and Dadda. Both 4×4 and 8×8 versions of the above three multipliers have been implemented and their performance comparison has been made.

The output results of 4×4 and 8×8 array multiplier show that the multiplier with SERF adder consumes low power but the output is highly distorted. Though the Proposed 14T consumes more power than SERF adder it is faster than it. The Proposed 14T full adder has the least value of power-delay product. The output results of 4×4 and 8×8 carry save multiplier show that the multiplier with PROPOSED 14T adder the least value of

Table IV. Power, delay comparison of 4×4 carry save multiplier using different adders.

Full adder	Power (in μW)	Delay (in ns)	Power delay product (10^{-15} Ws)
10T CLRCL	94.478	0.561	53.00216
10T SERF	51.927	0.539	27.98865
EXISTING 14T	58.818	0.463	27.23273
PROPOSED 8T	62.466	0.373	23.29982
PROPOSED 8T	51.427	0.353	22.29982

Table V. Power, delay comparison of 8×8 carry save multiplier using different adders.

Full adder	Power (in μW)	Delay (in ns)	Power delay product (10^{-15} Ws)
10T CLRCL	390.15	2.22	866.133
10T SERF	295.18	1.9	560.842
EXISTING 14T	335.73	2.068	694.28964
PROPOSED 8T	316.88	0.9	285.192
PROPOSED 8T	301.88	0.89	275.192

power-delay product. The Tables IV and V shows the 4×4 and 8×8 Dadda multiplier results using different full adders.

5. FIR FILTER

Three versions of 3-tap FIR filter[5] have been implemented in this section. They are the Broadcast 3 tap FIR filter, Unfolded Broadcast FIR filter and Unfolded and Retimed Broadcast FIR filter. All these three structures have been implemented using three types of multipliers (8×8 bit) with four types of the above described adders.

Figure 7 shows the FIR structure in direct form, here the critical path consists of a multiplier and *n* full adders. Figure 8 shows the broadcast structure of FIR filter where the critical path consists of one full adder and one multiplier.

6. RESULTS OF FIR FILTER

Product comparison graphs of 3 tap FIR Broad cast structure implemented by Array, Carry Save and Dadda multipliers. Each of these multipliers is designed using four different full adders (10T CLRCL, 10T SERF, proposed 8T and Existing 14T). It can be seen that the FIR filter using proposed 8T has least power delay product in all the cases.

7. CONCLUSIONS

The Proposed 8T full adder has 33% power savings when compared with Existing 14T Full adder and 38% power savings when compared to 10T CLRCL adder. The Proposed 8T full adder has the least delay when compared with its peers. Three multipliers Array, Carry-Save and Dadda has been designed using four different full adders and their performance has been compared. The multipliers implemented using proposed 8T full adder have least power delay product in all the cases. Three forms of 3 tap FIR filter namely broadcast, unfolded broadcast and unfolded retimed broadcast have also the results reveal that Filter structures implemented using Proposed 8T full adders posses the least power delay product. This reveals the compatibility of the Proposed 8T full adder for DSP applications.

References

1. J.-F. Lin and Y.-T. Hwang, A novel high-speed and energy efficient 10-transistor full adder design 54 **(2012)**.
2. E. Abu-Shama and M. Bayoumi, A new cell for low power adders, *Proc. Int. Midwest Symp. Circuits Syst.* **(1995)**, pp. 1014–1017; *DSP Journal* 9 **(2009)**.
3. T. Kowsalya, *ICGSTPDCS* 8 **(2008)**.
4. Deepak, G. Meher, and P. K. Sluzek, Performance characteristics of parallel and pipelined implementation of FIR filters in FPGA platform, *International Symposium on Signals, Circuits and Systems, 2007, ISSCS 2007*, July **(2007)**.
5. N. Zhuang and H. Wu, *IEEE J. Solid-State Circuits* 27, 840 **(1992)**.
6. J. Wang, S. Fang, and W. Feng, *IEEE J. Solid-State Circuits* 29, 780 **(1994)**.
7. A. M. Shams and M. Bayoumi, *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.* 47, 478 **(2000)**.
8. C. S. Wallace, *IEEE Trans. Electronic Computers* EC-13, 14 **(1964)**.
9. D. Rajesh Kumar, A. Shanmugam, C. Palanisamy, and A. M. Natarajan, *Asian Journal of Research in Social Sciences and Humanities* 6, 201 **(2016)**.
10. L.-F. Chao and E. H.-M. Sha, Efficient retiming and unfolding, *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993, ICASSP-93*, NJ, April **(1993)**.
11. D. Rajesh Kumar and R. Sathish, Dynamic detection of clone attack in wireless sensor networks, *The Proceedings of International Conference on Communication Systems and Network Technologies, (CSNT 2013) and IEEE Digital Xplore Held at MSIR Labs*, Gwalior **(2013)**, Vol. 1, pp. 501–505, ISBN 978-0-7695-3.
12. F. N. L. Op't Eynde, P. F. M. Ampe, L. Verdeyen, and W. M. C. Sansen, *IEEE Journal of Solid-State Circuits* 25, 265 **(1990)**.
13. P. Merolla and K. Boahen, A recurrent model of orientation maps with simple and complex cells, *Advances in Neural Information Processing Systems*, MIT Press, MA, Cambridge, December **(2004)**, Vol. 16, pp. 995–1002.
14. D. Rajesh Kumar and A. Shanmugam, A hyper heuristic localization based cloned node detection technique using GSA based simulated annealing in sensor networks, Cognitive Computing for Big Data Systems Over IoT, Springer International Publishing, Springer International Publishing, USA **(2017)**, pp. 307–335.
15. G. Indiveri, E. Chicca, and R. Douglas, *IEEE Trans. Neural Netw.* 17, 211 **(2006)**.

Received: 25 April 2018. Accepted: 14 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1719–1727, 2018

Bitstream Compression for High Speed Embedded Systems Using Separated Split Look Up Tables (LUTs)

R. Nirmal Kumar, V. Chandran, R. S. Valarmathi, and D. Rajesh Kumar*

Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam, Erode 638401, India

Bitstream compression is widely used to improve both communication bandwidth and memory requirement in high speed embedded systems. One of the popular bitstream compression techniques is Dictionary based bitstream compression (DBC) since it has better compression efficiency and fast decompression mechanism. The Bitmasking based compression techniques are increases the effectiveness of the DBC by recording the changed bits using bitmasks. Two efficient bitstream compression techniques with and without bitmasking are proposed for improving compression ratio (CR) and bit saving rate (BSR), named as Separated split LUT + Bit masking and Separated split LUT Algorithm. The proposed techniques are implemented in a 28 nm Qualcomm DSP core design project. Our proposal has achieved a compression rate of 34.5% and 7% improvement in code compression when compared to the Huffman traditional method.

Keywords: Bitstream Compression, Dictionary Based Bitstream Compression, Bitmasking Based Bitstream Compression, CR, BSR.

RESEARCH ARTICLE

1. INTRODUCTION

In general (In real time applications), Embedded System is nothing but the combination of system hardware and software. In addition to that High Speed embedded systems need to execute numerous applications simultaneously, which indirectly creates an impact on the on-chip memories. On-chip memories or Storage space are most essential one in all high speed embedded system applications such as Industrial machines, chips in telecom switching equipment, agricultural devices, automobiles, ambient devices, medical equipment, household appliances, vending machines, cameras, airplanes and toys as well as mobile devices.

In real time embedded applications, the required memory constraints are depends on the system complexity and amount of data to be used to process the particular application.[2] If the above mentioned two criteria are increased by some factors, the needed memory capacity will also simultaneously increase. For example, minimum 60 MB of bitstream are installed in current state-of-the-art cars.[1] The cost of an IC (integrated circuit) is strongly related to the bitstream size, and most of the chip area is occupied by the memory chip, a considerable reduction of the cost can be achieved by the reduction of embedded memory size.

One of the techniques used to reduce the memory size consumption of embedded applications is bitstream compression. The main aim of bitstream compression is to reduce the input bit streams (e.g., a data file, a speech signal, an image, or a video signal) as possible compare than the original input source. On the other hand, we need to recover the original bitstream from the compressed bitstream by decompression algorithm. But in our proposed work, we just focus on the bitstream compression algorithms. The efficiency of this algorithm is discussed in terms of compression ratio (CR) and Bit Saving Rate (BSR). The term compression ratio is defined as the ratio between compressed bitstream and original bitstream. For good compression, both the CR and BSR should be low and high respectively.[6, 14]

$$\text{Compression Ratio (CR)} = \frac{\text{Compressed bitstream}}{\text{Original bitstream}} \quad (1)$$

$$\text{Bit Saving Rate (BSR)} = 1 - \text{CR} \quad (2)$$

$$\text{Compression} = \frac{\text{Original bitstream} - \text{Compressed bitstream}}{\text{Original bitstream}} \quad (3)$$

The general frame work for bitstream compression and decompression is shown in Figure 1. In our frame work, we are discussed about the overall general structure and process of compression and decompression engine. The diagrammatic representation or flow of compression and decompression engine is shown in Figure 1. This frame work is discussed under two main terms i.e., behavior of

*Author to whom correspondence should be addressed.

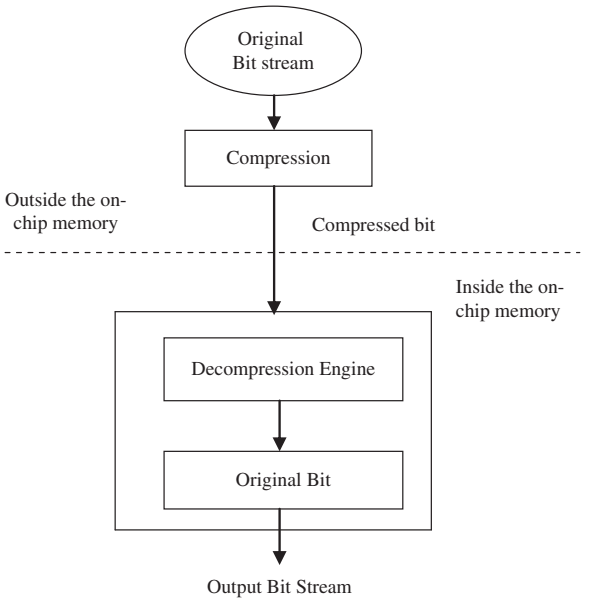


Fig. 1. General frame work of bitstream compression and decompression.

bitstreams by outside and inside the on-chip memories in embedded system. From the figure, both the memory parts are separated by the dotted line.

At first the input bit stream or the data (instructions) to be compressed is given to the compression block, which is placed outside the on-chip memories. One of the lossless compression techniques is used to compress the particular word length bitstream in compression block. Output of the compression block i.e., the compressed bit should be less than the original input bitstream. The size of the compressed bit is directly depends on which type of compression technique used. The compressed bit is entered into the decompression engine to recover the original data without make any loss on that. These extractions of original bitstream from the compressed bitstreams are handled inside the on-chip memories. The final output of this frame work is equal to our original bitstream (instructions) which is to be used as the input of this frame work.

The rest of this paper is structured as follows. In Section 2, the works related to proposed algorithms are discussed briefly. The proposed Separated split LUT + Bit masking Algorithm and Separated split LUT Algorithm are presented with the proper examples in Section 3. Section 4 presents the Experimental Results and Analysis of presented algorithms and Section 5 concludes this paper.

2. RELATED WORKS

Most of the computer systems needed large amount of memory to process the coding. Embedded systems not have this much of possibility to process the code or bitstream, because it does not have a large memory. For that it needs to compress the processing bitstream as much as possible before store in to the memories. In addition to that this should be falling under lossless bitstream compression for better recovering efficiency. Some of the considerable lossless bitstream compressions are run length encoding, Golomb code, Huffman code, Lempel Ziv, Dictionary Based Compression, Bit masking, which are reviewed below.

Very simplest lossless compression encoding method is run length encoding (RLE).[18] The most useful data like simple graphic images such as line drawings, icons and animations that contain many consecutive ones or zeroes. Jagadeesh et al.[13] proposed that the test data compression technique to reduce the test data size without affecting the overall system performance and behavior. This technique mainly contributions the follows: 1. The maximum bit matching pattern is generated from an efficient bitmask selection technique for test data compression; 2. an sufficient dictionary selection method is developed here based on the bitmask based compression; and 3. The test data compression technique is proposed an efficient dictionary and bitmask selection to minimize the testing time and memory storage space. Combined bitmask based compression and run length coding[3] provides better compression ratio and decompression efficiency for frequent data. But it is not suitable for the bitstream which do not have any consecutive runs (ones or zeroes) as it increases the file size largely.

To avoid continuous variation in compression, the Golomb Coding is used instead of Run Length encoding. Saranya et al.[4] proposed the modified Decode Aware Compression method, which replaces the Runlength encoding compression technique by Golomb Encoding technique. The large size data is compressed into smaller one by applying the Golomb Encoding technique. The achieved Compression Ratio of the proposed method is independent of the decompression hardware and depends on the entropy of the configuration bitstream. Configuration time of FPGA is depends only on the data rate, the size of the configuration bit-stream, and the speed of a memory stores the configuration data. Anshuman Chandra et al. proposed a new variable-to-variable-length Golomb code[5] for test-data compression and decompression architecture. Very high compression and a low-cost on-chip decoder are the advantage of Golomb coding.

Bonny and Henkel[15] used Lempel Ziv Storer Szymanski compression algorithm with a filled buffer technique and extended blocks used to compress VLIW instructions. For further compression, the Huffman coding is used as the next method in the proposed extended blocks.

Compression is an efficient way to store the bitstreams into the limited size of embedded memory. Huffman coding is one of the lossless compression techniques. The main aim is to show the design and implementation

of a circuit which performs a coding of text files without considering the main memory of the system. Wolfe and Chanin[7] are used Huffman coding on Microprocessor not including Interlocked Pipeline Stages processors. And which is implemented by a precache structure with modified architecture of cache. Recent research in bitstream compression made a continuation to the use of memory addresses for compression. Larin and Conte[8] applied Huffman coding to VLIW (Very Long Instruction Word) processors. The bounded Huffman coding and bitstream placement algorithm is proposed by Qin and Mishra[9] to compress and replace compressed instructions. The Huffman-based multilevel dictionary[10, 11] is proposed to compress two adjacent instruction sequences as well as for single instruction compression.

Bonny et al. proposed a novel hardware supported approach to optimize the number and size of generated LUTs and to improve the compression ratio in embedded system and with statistical and dictionary-based compression schemes. In *statistical schemes*, the codeword is choosing based on the frequency of instruction sequences which replace the original data instructions. The most frequent data are replaced by the shorter codeword and the less frequent data are replaced by the shorter codeword. In *dictionary-based schemes*,[11] the sequences of instructions are replaced by a single new codeword and which is also used as an index value to a dictionary that has the original instruction sequence.

BitMask-based code compression is an improved version of Dictionary-Based Code Compression (DCC). Bitmask is involved to compress the larger bitstream into small compare than DCC. The purpose of BitMask is to mask mismatched values by its positions. Wang et al.[12] proposed a small separated dictionary and variable mask numbers were used with the BitMask algorithm to diminish the bitstream length of highly frequent instructions. And also a novel dictionary selection algorithm is proposed to improve the instruction matching rates. The fully separated dictionary method is used to reduce compression ratio (CR) by improving the performance of the decompression engine. In this research work, the bitstream compression is focused on two directions: (1) applying existing compression methods to various architectures for optimization and (2) combining several approaches to improve the performance, including CR.

2.1. Dictionary Based Code Compression (DCC)

Dictionary-based code compression techniques provide compression efficiency as well as fast decompression mechanism. DCC is the method in which the LUT are constructed based on their frequent data occurrence. The most frequently occurred data are stored in small LUT with index number. Similarly the data are inserted in large LUT with frequent data repetition until the LUT is full. The tag bits are used to identify whether the data are compressed with LUT access or not. The dictionary based code

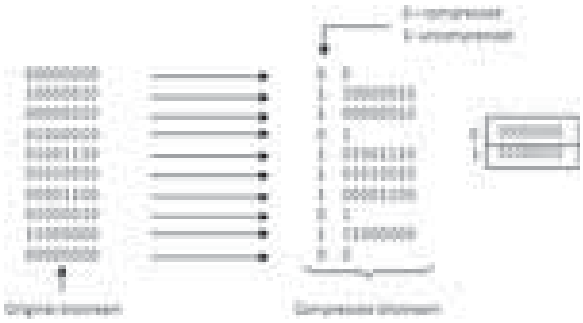


Fig. 2. Dictionary code compression.

compression provides efficient code compression with less hardware overhead. The following examples in Figure 3 shows the dictionary based code compression which provide clear knowledge how it works. The following example has a compression ratio of 2.5%. The original program has 8-bit code size and they are compressed with the help of value stored in LUT and access by index number. The next section shows a bit mask code compression.

2.2. Improved Dictionary Based Approach

The basic idea is to determine the instruction sequences that are different in few bit positions (hamming distance) and store that information in the compressed program and update the dictionary (if necessary). The compression ratio will depend on how many bit changes are considered during compression. If more bit changes are allowed, more matching sequences will be generated. However, the size of the compressed program will increase depending on the number of bit positions. Figure 3 shows the improved dictionary based scheme considering 1-bit change. An extra field is necessary to indicate whether mismatches are considered or not. In case a mismatch is considered, another field is necessary to indicate the bit position that is different from an entry in the dictionary. For example, the third pattern (from top) in the original program is different from the first dictionary entry (index 0) on the sixth bit position

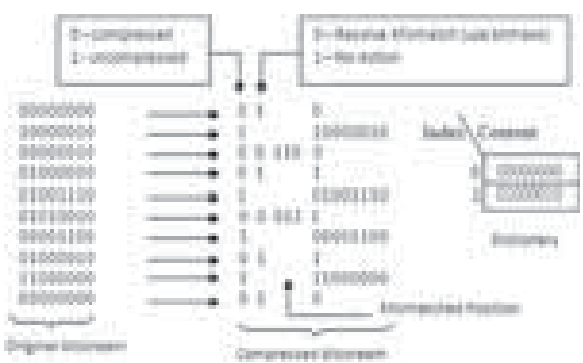


Fig. 3. Dictionary code compression.

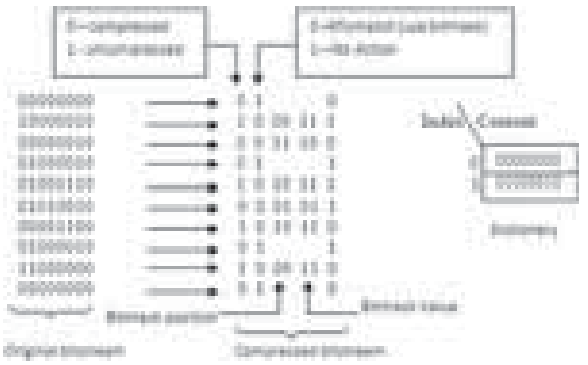


Fig. 4. Bit mask code compression.

(from left). The compression ratio for this example is 95%.

2.3. Bit Mask Code Compression (BCC)

Bitmask based compression an enhancement on the dictionary based compression scheme, which helps us to get more matching patterns. In dictionary based compression, each data is compressed only if it completely matches with a dictionary entry. The data that are a directly coordinated are compressed with 3 bits. The first bit represents whether it is uncompressed (using 1) or compressed (using 0). The second bit shows whether it is compressed using bitmask (using 0) or not (using 1). The last bit indicates the dictionary index. In Figure 4 the last bit indicates the dictionary index. Data that are compressed using bitmask requires 7 bits. The data compressed by using bitmask approach, are indicated by initial two bits. The next two bits indicate position of the mask and followed by two bits that indicate the value of bit in that position.

The data which is different for more than 1 bit is left uncompressed. Dictionary-based code compression[4] is commonly used in FPGA systems, because it can reach a proficient CR, possess a fairly straightforward decoding. A new algorithm is to develop the compression concert (including the CR) with a lesser hardware operating cost. Based on the Bit Mask code compression[2, 3] algorithm, a tiny divided dictionary is projected to limit the codeword length of high-frequency instructions, and a original dictionary selection algorithm is proposed to achieve more satisfactory instruction selection, which in turn may reduce the average CR. The corresponding examples show how it works with Bitmask approach.

3. PROPOSED SEPARATE SPLIT LUT ALGORITHM

In this section, the proposed algorithms are described. Separated LUTs are used to reduce the bitstream length of high-frequency instructions.

Fig. 5. Proposed flow diagram.

3.1. General Flow of Proposed Algorithm

The general flow diagram of bitstream compression of our proposed separated split LUT algorithm is shown in Figure 5. Initially, the flow chart gets the original input bitstream, LUT size and number of rows in LUT as the inputs. Finally, we derive the output as compressed bitstream, compression ratio (CR) and saving rate.

The bitstream, which we got as the input to this algorithm is segmented based on the size of LUT. And further, the segmented input bitstream are involved to pick up the frequently occur bitstream, which is stored into concern LUT (Small/Big LUT). If R number of rows needed for designing both small and big LUT, then select R frequent bitstreams among the input segmented bitstream. The most frequent bitstream is stored into the small LUT and the Big LUT is used to store the R-1 frequent bitstreams. Here we are discussed two proposed compression algorithm for bitstream compression. One is separated split LUT and another one is separated split LUT + Bit masking. In the next flow step, the actual compression is takes place between the input bitstream segments and the separated LUTs based on any one of the proposed compression algorithm (set of procedures).

Table I. Small LUT of Figure 2.

01		Small LUT
0	Direct match	0000 0000 0000 0000 0000 0000 0000 0000
		Separate split small LUT
10	Left match	0000 0000 0000 0000 XXXX XXXX XXXX XXXX
11	Right match	XXXX XXXX XXXX XXXX 0000 0000 0000 0000

	Original bitstream								Compressed bitstream			
D1	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D2	0000	1111	0000	1111	0000	1111	0000	1111	10	00		
D3	0101	1111	1010	1111	0000	1111	0000	1111	10	11	0	0101 1111 1010 1111
D4	0000	1000	0000	0000	0000	0000	0000	0000	11	0		00100 0
D5	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D6	0000	1111	0000	1111	0000	1111	0000	1111	10	00		
D7	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D8	1010	1010	1010	1010	1010	1010	1010	1010	10	01		
D9	0101	1111	0101	0101	0110	1001	0000	1111	00	01011111010101010110100100001111		
D10	0000	1111	0000	1111	0000	1111	0000	1111	10	00		
D11	1010	1010	1010	1010	1110	1111	1110	1111	10	10	1	1110 1111 1110 1111
D12	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D13	1010	1010	1010	1010	1010	1010	1010	1010	10	01		

Fig. 6. Sample separated split LUT + bit masking compression for 416 bits (input) entries.

3.2. Separated Split LUT + Bit Masking Algorithm

Separated split LUT + Bit masking Algorithm is one of the proposed lossless compression algorithm and which is derived from the traditional Dictionary based algorithm.[11] In this proposed algorithm, two main concepts are discussed i.e., Right and Left splits LUT and bit making for better compression ratio compare than the existing algorithms. For compression, the traditional methods are comparing the whole segmented input bitstream to the bitstream (frequent data) stored in the LUT. But in our research work, the whole and partial (splitting) segmented input bitstream are compared with the whole and split LUTs for efficient compression. For example, 2*n* bit stored bitstream is splits as *n* bits.

Algorithm Steps

1. Let N bit as the input original bitstream
2. Number of rows in small LUT should be less as compared to big LUT rows. Let R be the total number of rows in both LUTs.
3. Set the size of small and big LUTs as 'L' bits and the index values fixed based on number of big LUTs rows.
4. In addition to that both the LUTs are further splitted (say Left and Right splitting) for best compression (split size of bits are half of LUTs size i.e., L/2)
5. Bitmask size
6. For compression
   (i) 00—Uncompressed
   (ii) 01—Compressed with small LUT

   0—direct match
   10—Left splitting is matched with input segment
   11—Right splitting is matched with input segment

   11 10—Compressed with Big LUT (with 2 index values)
   00—Direct match with index 0
   01—Direct match with index 1
   10x—Left splitting is matched with input segment
   11x—Right splitting is matched with input segment

   12 11—Involved Bitmask
   0—Bit mask with small LUT
   11—Bit mask with big LUT with index '0'
   11—Bit mask with big LUT with index '1'
7. Based on the LUTs size, the input bitstream is segmented and ready for compression.

Table II. Big LUT of Figure 2.

Index	10		Big LUT
0	00	Direct match	0000 1111 0000 1111 0000 1111 0000 1111
			Separate split big LUT
	10 0	Left match	0000 1111 0000 1111 XXXX XXXX XXXX XXXX
	11 0	Right match	XXXX XXXX XXXX XXXX 0000 1111 0000 1111
1	01	Direct match	1010 1010 1010 1010 1010 1010 1010 1010
			Separate split big LUT
	10 1	Left match	1010 1010 1010 1010 XXXX XXXX XXXX XXXX
	11 1	Right match	XXXX XXXX XXXX XXXX 1010 1010 1010 1010

Table III. Bit mask of Figure 2.

11	Bit mask
0	Bit mask with small LUT
10	Bit mask with big LUT with index '0'
11	Bit mask with big LUT with index '1'

Table IV. Compression using bit mask for 32 bits.

Bit mask		
Original data—XXXX YYYY XXXX YYYY XXXX YYYY XXXX YYYY		
	Bit changed position	Masking
ZXXX YYYY XXXX YYYY XXXX YYYY XXXX YYYY	0	00000
X**Z**XX YYYY XXXX YYYY XXXX YYYY XXXX YYYY	1	00001
XX**Z**X YYYY XXXX YYYY XXXX YYYY XXXX YYYY	2	00010
XXX**Z** YYYY XXXX YYYY XXXX YYYY XXXX YYYY	3	00011
⋮	⋮	⋮
XXXX YYYY **Z**XXX YYYY XXXX YYYY XXXX YYYY	8	01000
XXXX YYYY X**Z**XX YYYY XXXX YYYY XXXX YYYY	9	01001
XXXX YYYY XX**Z**X YYYY XXXX YYYY XXXX YYYY	10	01010
⋮	⋮	⋮
XXXX YYYY XXXX YYYY XXXX YYYY XXXX Y**Z**YY	29	11101
XXXX YYYY XXXX YYYY XXXX YYYY XXXX YY**Z**Y	30	11110
XXXX YYYY XXXX YYYY XXXX YYYY XXXX YYY**Z**	31	11111

	Original bitstream								Compressed bitstream			
D1	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D2	0000	1111	0000	1111	0000	1111	0000	1111	10	00		
D3	0101	1111	1010	1111	0000	1111	0000	1111	10	11	0	0101 1111 1010 1111
D4	0000	1000	0000	0000	0000	0000	0000	0000	01	11		0000 1000 0000 0000
D5	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D6	0000	1111	0000	1111	0000	1111	0000	1111	10	00		
D7	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D8	1010	1010	1010	1010	1010	1010	1010	1010	10	01		
D9	0101	1111	0101	0101	0110	1001	0000	1111	00	01011111010101010110100100001111		
D10	0000	1111	0000	1111	0000	1111	0000	1111	10	00		
D11	1010	1010	1010	1010	1110	1111	1110	1111	10	10	1	1110 1111 1110 1111
D12	0000	0000	0000	0000	0000	0000	0000	0000	01	0		
D13	1010	1010	1010	1010	1010	1010	1010	1010	10	01		

Fig. 7. Example for 416 input bits (method—separated right and left split LUT).

8. N bit original inputs are segmented as N/L segments with L bits named as D1, D2, D3, D4,…, D(N/L).
9. From the segmented bitstreams, find F frequent bitstream segment.
10. The most frequent segment among other segment is placed in small LUT (s)
11. The next frequent one is stored in big LUT with index 'x' (bx)
12. Continue the step 11 up to fill all the rows (R-1) in the big LUT.
13. If the original bitstream segment D1 is matched to small LUT, then which is falling under compression condition step 6(i).
14. Likewise, we do this procedure for all original input bitstream segments by compression procedure in step 6.
15. **Output**: compression ratio, bit saving rate, number of bits in compressed bitstream.

Example for 416 input bits (Method—Right and Left split + Bitmask).

The Table I shows the design of Small LUT for Figure 6. The most frequent bitstream among all segmented bitstream is stored in the small LUT. The resultant compression bitstream show that the first two bits explain the compression happened belongs to small LUT and the follow up bits indicates what type of matching is occur.

3.3. Separated Right and Left Split LUT

In Figure 7 Separated Right and Left split LUT is another one proposed lossless compression algorithm in our research work. In this novel algorithm, the segmented input bitstream is compressed directly with the concern LUTs or else which is further virtually splits and matched with particular LUTs for better compression. The virtual or partial splitting is done as n bits for $2n$ bit stored or

Table V. Power comparison.

	Power (mW)							
	Compression techniques							
Bench mark circuits	Huffman	LZ	Run length	Golomb	Dictionary based code compression	Bitmask code compression	Separated split LUT	Separated split LUT + Bit masking
Sha	528	483	491	611	445	464	592	539
FFT	567	519	549	650	512	488	516	517
CRC	556	555	597	677	525	562	667	670
Dijkstra	565	593	587	704	597	726	659	717
Susan	568	670	622	730	625	734	733	739

Table VI. Area (device utilization) comparison.

	Area							
	Compression techniques							
Bench mark circuits	Huffman	LZ	Run length	Golomb	Dictionary based code compression	Bitmask code compression	Separated split LUT	Separated split LUT + Bit masking
Sha	9732	9705	12072	20644	5198	6459	7306	6904
FFT	10734	11130	13718	21223	6096	7139	7197	7289
CRC	9757	11499	15944	22458	6884	8472	9050	9064
Dijkstra	11051	12836	16569	24182	7740	9843	9432	10142
Susan	11764	14245	17212	29993	8018	11487	10592	10422

Table VII. Delay comparison.

Delay (ns)								
	Compression techniques							
Bench mark circuits	Huffman	LZ	Run length	Golomb	Dictionary based code compression	Bitmask code compression	Separated split LUT	Separated split LUT + Bit masking
Sha	25.833	16.490	17.645	24.976	12.888	12.610	13.743	13.943
FFT	26.531	16.685	18.145	24.348	13.138	12.613	14.533	12.963
CRC	26.102	16.945	18.995	24.406	13.268	12.813	14.218	14.218
Dijkstra	27.780	16.070	18.420	24.906	13.263	13.758	13.088	14.368
Susan	27.283	17.268	18.545	25.767	13.363	13.263	13.405	13.065

segmented input bitstream. This algorithm is not discussed about any bitmasking concepts (for single bit change from $2n$ bits).

Algorithm Steps

1. Let N bit as the input original bitstream
2. Number of rows in small LUT should be less as compared to big LUT rows. Let R be the total number of rows in both LUTs.
3. Set the size of small and big LUTs as 'L' bits and the index values fixed based on number of big LUTs rows.
4. In addition to that both the LUTs are further splitted (say Left and Right splitting) for best compression (split size of bits are half of LUTs size i.e., L/2)
5. For compression
   00—Uncompressed
   01—Compressed with small LUT
   0—direct match
   10—Left splitting is matched with input segment
   11—Right splitting is matched with input segment
   10—Compressed with Big LUT (with 2 index values)
   00—Direct match with index 0
   01—Direct match with index 1
   10x—Left splitting is matched with input segment
   11x—Right splitting is matched with input segment
6. Based on the LUTs size, the input bitstream is segmented and ready for compression.
7. N bit original inputs are segmented as N/L segments with L bits named as D1, D2, D3, D4,…, D(N/L).
8. From the segmented bitstreams, find F frequent bitstream segment.
9. The most frequent segment among other segment is placed in small LUT (s)
10. The next frequent one is stored in big LUT with index 'x' (bx)
11. Continue the step 11 up to fill all the rows (R-1) in the big LUT.
12. If the original bitstream segment D1 is matched to small LUT, then which is falling under compression condition step 5(i).
13. Likewise, we do this procedure for all original input bitstream segments by compression procedure in step 5.

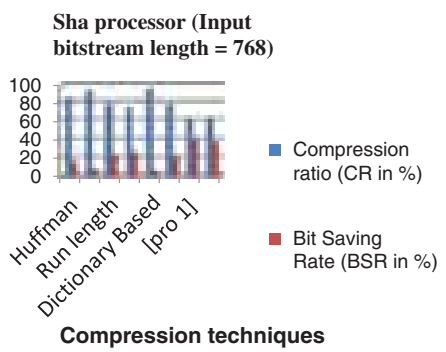


Fig. 8. CR and SBR comparison on sha processor with 768 entries.

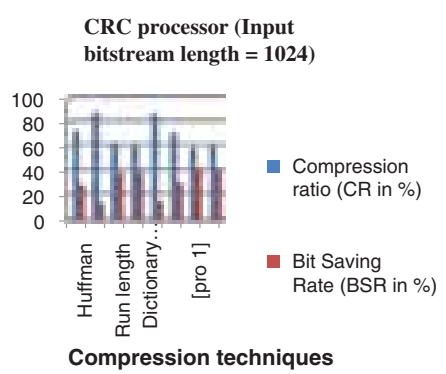


Fig. 9. CR and SBR comparison on CRC processor with 1024 entry.

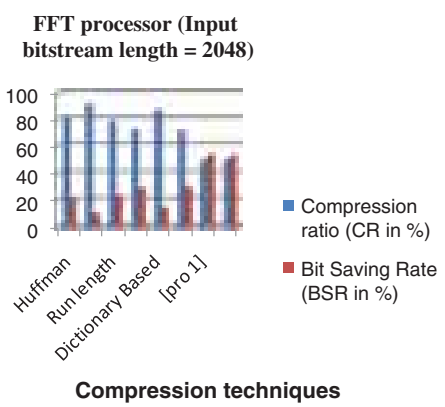


Fig. 10. CR and SBR comparison on FFT processor with 2048 entry.

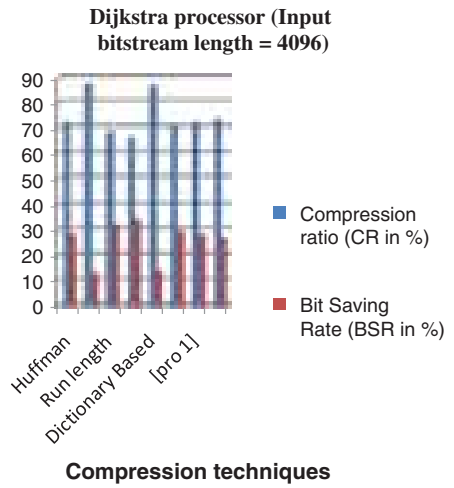


Fig. 11. CR and SBR comparison on Dijkstra processor with 4096 entry.

14. **Output**: compression ratio, bit saving rate, number of bits in compressed bitstream.

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, experimental results and comparison analysis made between different benchmarking circuits like sha, FFT, CRC, Dijkstra, Susan and the various compression techniques including our proposed techniques are presented.

The achieved compression and bit saving rate for sha processor with 768 entries is shown in Figure 8. In general, the compression ratio and saving rate are indirectly proportional i.e., the decreasing CR and increasing BSR projects the better compression performance. From this figure, our proposed algorithms give the best compression parameters.

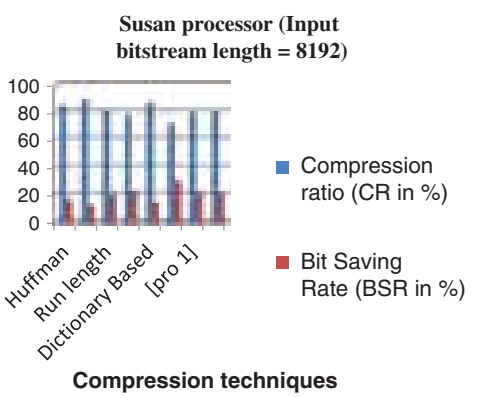


Fig. 12. CR and SBR comparison on susan processor with 8192 entry.

The CR and BSR achieved by CRC processor with 1024 entries, FFT processor with 2048 entries, dijkstra with 4096 entries and susan with 8192 entries are shown in Figures 9–12 respectively.

5. CONCLUSION

The compression algorithms are very simple and should not be power, time and area consuming in order to improve the overall process of the system. The lossless code compression algorithms presented herein is Separated split LUT + Bit masking and separated split LUT Algorithms that aims to improve the compression and bit saving rate. Our technique performs well for different input bitstream entries and dictionary sizes. The comparison results presented here shows that the efficiency of our proposed compression algorithms was best among all other bench mark circuits. Important performance metric of these days is power, delay and area utilization, which are also taking into account for result analysis.

Our experimental results show that our approach reduces the original program size by up to 45%. Our technique outperforms all the existing dictionary-based techniques by an average of 15%, giving compression ratios of 55%–65%. We also proposed the design of a simple and fast decompression unit that is capable of decoding an instruction per cycle as well as performing parallel decompression.

References

1. M. Broy, Automotive software and systems engineering, *Proc. ACM/IEEE Int. Conf. Formal Methods Models Co-Design* **(2011)**, pp. 143–149.
2. T. Bonny and J. Henkel, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 16, 1696 **(2008)**.
3. X. Qin, Chetan Muthry, and Prabhat Mishra, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 19, 411 **(2011)**.
4. R. Saranya, S. Kousalya Devi, and V. Lakshmi Prabha, *International Journal of Computer Applications* 21.
5. Anshuman Chandra and Krishnendu Chakrabarty, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 20, 355 **(2001)**.
6. S.-W. Seong and Prabhat Mishra, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 27, 673 **(2008)**.
7. A. Wolfe and A. Chanin, Executing compressed programs on an embedded RISC architecture, *Proc. 25th Annu. Int. Symp. Microarchitecture*, December **(1992)**, pp. 81–91.
8. D. Rajesh Kumar and A. Shanmugam, A hyper heuristic localization based cloned node detection technique using GSA based simulated annealing in sensor networks, Cognitive Computing for Big Data Systems Over IoT, Springer International Publishing **(2017)**, pp. 307–335.
9. S. Y. Larin and T. M. Conte, Compiler-driven cached code compression schemes for embedded ILP processors, *Proc. 32nd Annu. Int. Symp. Microarchitecture*, IEEE Computer Society Washington, DC, USA, November **(1999)**, pp. 82–91.
10. X. Qin and P. Mishra, Efficient placement of compressed code for parallel decompression, *Proc. 22nd Int. Conf. VLSI Design*, January **(2009)**, pp. 335–340.
11. W. R. A. Dias, E. D. Moreno, and I. N. Palmeira, A new code compression algorithm and its decompressor in FPGA-based hardware,

Proc. 26th Symp. Integr. Circuits Syst. Design (SBCCI), September **(2013)**, pp. 1–6.

12. X. Kavousianos, E. Kalligeros, and D. Nikolos, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 26, 1070 **(2007)**.

13. W. J. Wang and C. H. Lin, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 24, 266 **(2016)**.

14. D. Rajesh Kumar, A. Shanmugam, C. Palanisamy, and A. M. Natarajan, *Asian Journal of Research in Social Sciences and Humanities* 6, 201 **(2016)**.

15. E. Chicca, D. Badoni, V. Dante, M. D'Andreagiovanni, G. Salina, S. Fusi, and P. Del Giudice, *IEEE Trans. Neural Netw.* 14, 1297 **(2003)**.

16. S. Jagadeesh, T. Venkateswarlu, and M. Ashok, *International Journal of Engineering Research and Technology (IJERT)* 1, 1 **(2012)**.

17. Satyendra N. Biswas, Sunil R. Das, Mansour H. Assaf, and Altaf Hossain, Test vector compression technique in system-on-chip, *Instrumentation and Measurement Technology Conference*, May **(2009)**, pp. 1126–1131.

18. D. Rajesh Kumar and T. Sathish, Dynamic detection of clone attack in wireless sensor networks, *Proceedings of International Conference on Communication Systems and Network Technologies, (CSNT 2013) and IEEE digital Xplore held at MSIR labs*, Gwalior, Vol. 1, pp. 501–505, ISBN 978-0-7695-3.

19. M. Oster, A. M. Whatley, S.-C. Liu, and R. J. Douglas, A hardware/software framework for real-time spiking systems, *Proc. 15th International Conf. Artificial Neural Netw.: Biological Inspirations (ICANN'05)*, September **(2005)**, Vol. 3696, pp. 161–166.

20. T. Bonny and J. Henkel, FBT: Filled buffer technique to reduce code size for VLIW processors, *Proc. IEEE/ACM Int. Conf. CAD (ICCAD)*, November **(2008)**, pp. 549–554.

Received: 25 April 2018. Accepted: 14 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1728–1733, 2018

RESEARCH ARTICLE

Economics Analysis of Amazon EC2 Instances

K. P. Sampoornam[1,*], Hyder Ali Segu Mohamed[2], and R. Priyadarshini[1]

[1] *Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Erode 638401, Tamilnadu, India*
[2] *Department of Computer Engineering, Buraydah Colleges, 51418, Kingdom of Saudi Arabia*

The monitoring and the analysis of public clouds gains momentum, due to their widespread exploitation by individual users, researchers and companies for their daily tasks. This paper proposed an algorithm for optimizing the cost and the utilization of a set of running Amazon EC2 instances by resizing them appropriately. The algorithm, namely Cost and Utilization Optimization algorithm, receives information regarding the current set of instances used (their number, type, utilization) and proposes a new set of instances for serving the same load, so as to minimize cost and maximize utilization, or increase performance efficiency. Cost and Utilization Optimization is integrated in Smart cloud Monitoring (SuMo), an open-source tool. It is develop for collecting monitoring data from Amazon Web Services and analyzing them. A number of experiments are performed, using input data that correspond to realist Amazon Web Services configuration scenarios, which exhibit the benefits of the Cost and Utilization Optimization algorithm.

Keywords: Public Clouds, Monitoring, Amazon Web Services, Cost, Utilization, Optimization.

1. INTRODUCTION

Cloud computing provides resources as a service over the network, enabling their efficient and flexible management. Cloud computing also replaces Information Technology (IT) costs of ownership and operation with variable costs that depend on the use of the cloud resources. As a result, today, an increasing number of individual users, researchers and companies, trust their computing and storage tasks on public and private clouds. The former provides resizable compute capacity as a public service (e.g., Amazon Web Services—AWS, Rackspace), while the latter are built by organizations' own infrastructure, using cloud computing toolkits (e.g., Open Nebula, Eucalyptus) and providing computing services to their employees or customers.

IT monitoring is an important operation for cloud resource management, as for all kinds of IT systems. The sheer number of cloud resources makes it difficult for a simple user or administrator to effectively monitor them or control the parameters that determine their proper use. Estimates place the number of servers in commercial data centers, used for providing public or private cloud services, to 450.000 servers, while even larger ones (in the order of 106 to 107 machines) have been envisioned for the future. Also, organizations may utilize hundreds of virtual instances of a public cloud provider for their operations. A good monitoring entity ensures that the monitored resources run uninterrupted, with acceptable performance and utilization, keeping the associated (e.g., energy) costs low.

There are a number of important differences between monitoring private and public clouds, both regarding the way monitoring is performed and regarding the parameters of interest and the requirements. In private clouds the administrator has full access to the resources, being able to monitor any kind of parameter (performance, energy, etc.), with any kind of software or hardware tools and at any granularity. This is not the case for public clouds (e.g., Amazon Web Services—AWS) where the users have access to their virtual resources through web or programmable interfaces that provide only specific information (e.g., state, performance), at least in the general case), and at specific granularities (e.g., every 5 minutes). Additionally, the main parameter of interest for the public clouds' users is the cost. The cost encountered by a user of public clouds depends mainly on the pricing policy of the public cloud provider.

In this work, the Cost and Utilization Optimization (CUO) algorithm, for minimizing the usage cost of Amazon EC2 instances and maximizing their utilization and the performance efficiency. The proposed algorithm collects information regarding the current set of AWS instances used (their type, number, utilization) and proposes a new set of instances that could be used for serving the same load. A number of experiments are performed, using input data that correspond to realist AWS configuration

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1728/006 doi:10.1166/jctn.2018.7368

scenarios. These experiments show that the proposed algorithm can increase the utilization of the resources used, and thus the efficiency in the use of the infrastructure resources, while lowering the accumulated resource costs. The remainder of the paper is organized as follows. Section 2 report on previous work. In Section 3 Amazon's computing and monitoring services are described. In Section 4 Cost and Utilization Optimization (CUO) algorithm is elaborated, while various extensions of it are investigated in Section 5. In Section 6, the Smart cloud Monitoring (SuMo) toolkit and its constituent modules are presented. Experimental results obtained using SuMo and the CUO algorithms are shown in Section 7.

2. RELATED WORK

IT monitoring has long been around, targeting the monitoring needs of physical servers and other devices (printers, switches, upc, etc.), clusters, grids and small or large data centers. However, clouds bring a completely new environment and introduce new requirements for IT monitoring tools, having a very large number of heterogeneous physical and virtual resources and producing a huge amount of raw monitoring information. Most works on the monitoring of cloud resources assume full access to the corresponding resources and are consequently more appropriate for private clouds. In particular, a number of works attempt to aggregate (or summarize) the raw monitoring data. In, the authors propose a scalable distributed data collection system that utilizes technologies from the semantic web in order to generate a machine readable overview of a cloud system without the need for an additional dedicated monitoring system. In, the authors propose a Runtime Model for Cloud Monitoring (RMCM), which provides an intuitive representation of a running cloud. Cloud Sense, described in, is a switch design that performs in-network compression of monitoring data streams via compressive sensing.

Monitoring can be categorized as state,[1] service[2,3] or performance[4] monitoring. Status monitoring is widely used for detecting critical events and abnormalities. Service monitoring relates to diverse aspects of different services, while performance monitoring evaluates the quality of the provided services. Quite recently, a number of products/services have appeared offering monitoring and analysis tools for public cloud resources, while the functionality of other already established products has been extended appropriately to meet the particular needs of private clouds. The authors in Ref. [5] discuss the design and implementation of a Private Cloud Monitoring System (PCMONS) and argue that it is possible to deploy a private cloud within an organization using only open source solutions and integrating it with traditional monitoring tools, such as Nagios;[6] significant development work, however, has to be carried out in order to actually integrate these tools. In Ref. [7] authors have presented an approach to configuration planning based on data refinement; correlating economic goals with sound technical data. The authors also present a proof of concept tool, called CloudXplor, which can be modularly embedded in generic resource monitoring and management frameworks.

Even though these products are very good in aggregating the collected information and presenting it in a user friendly manner to the users, there is significant room for adding more intelligence to the monitoring process and adding value to the analysis of public cloud monitoring data, in addition to simply forecasting virtual resource usage and associated costs.[8] In Ref. [9] this paper is applied social network analysis methods on virtual resources communication traces in order to identify important VMs and capture the time dynamicity of their interactions. In this work proceed a step further and instead of simply identifying over- or under-used cloud resources, this paper present an algorithm for optimizing the cost, the utilization and the performance of Amazon cloud resources. This mechanism provides specific suggestions regarding which instances type should be modified in order to minimize the cost to the user and maximize the utilization of the resources or increase performance by resolving capacity bottlenecks.

3. AMAZON WEB SERVICES

Amazon Web Services' (AWS)[10] public clouds is a collection of web services that provide access to Amazon's cloud infrastructure. Amazon Elastic Computing cloud-EC2[11] is one of the most important of these services, providing resizable compute capacity as a service. The basic unit of EC2 is the "instance." Each instance is a virtual resource with particular computational, storage and network characteristics, running a particular Operating System (OS) and located physically in one of Amazon's data-centers around the world. The above characteristics also determine the cost of the instance. There are two main instance types: (i) On-Demand Instances (ODI) and (ii) Reserved Instances (RI). With on-demand instances a user pays for compute capacity by the hour, with no long-term commitments, while with reserved instances the user make an one-time payment for reserving an instance and then pays again by the hour but at a significant discount. There are also Spot instances allow one to bid on spare Amazon EC2 instances and run them whenever his bid exceeds the current spot price.

For example, machine type "Extra Large Instance" (m1.xlarge) has the following characteristics:

- 15 GB memory
- 8 EC2 Compute Units—ECU: 4 virtual cores with 2 EC2 Compute Units each (according to Amazon,[1] one ECU provides the equivalent CPU capacity of a 1.0–1.2 GHz 2007 Opteron or 2007 Xeon processor.)
- 1,690 GB instance storage

RESEARCH ARTICLE

- 64-bit platform
- I/O Performance: High.

Amazon's Cloud Watch[12] provides monitoring for AWS cloud resources (such as Amazon EC2). Developers and system administrators can use it to collect and track various metrics. The basic Cloud Watch metrics are the following, though custom ones can also be added:

- CPU Utilization: The percentage of allocated EC2 compute units that are currently in use on the instance
- Disk Read/Write Ops: Number of completed read/Write operations from all ephemeral disks available to the instance
- Disk Read/Write Bytes: Number of bytes read/write from all ephemeral disks available to the instance
- Network In/Out: The number of bytes received/sent on all network interfaces by the instance.

Monitoring data are available automatically, at 1 or 5-minute interval steps, depending on the charging policy chosen. Only 1440 points of a particular metric can be provided by CloudWatch at time; e.g., the CPU utilization of an instance per 1-minute (step size) for the last 24 hours (period). Other step sizes and periods can also be selected.

4. CUO EXTENSIONS

The presented algorithm can also be used to support and optimize other resource characteristics, such as the memory (instead of the CPU) of the virtual instances. Similarly, the presented mechanism can maximize the utilization of the memory usage for all instances by suggesting the replacement of under-utilized (in terms of memory) instances with other instances of lower memory capacity and of lower cost. Again, by configuring properly parameter f, the Cost and Utilization Optimization (CUO) algorithm can also lead to the selection of high memory instances in order to replace instances with utilization near 100%, indicating memory bottleneck. In this implementation of the CUO algorithm in SuMo, it includes this extension because memory utilization is not among the metrics provided by default by CloudWatch (Section 3). Other resource characteristics, e.g., IO, could also be considered, separately or in conjunction with CPU and memory.

In the CUO algorithm presented, instance migration costs not considered, which are in any case difficult to calculate. Migration costs depend on a number of measurable (e.g., instance data need to be moved) and non (at least, not easily) measurable (e.g., the cost of disrupting the service offered by the particular instance) parameters. In any case if instance migration cost is known, then the CUO algorithm can easily be extended appropriately so as to include it in its optimization.

Moreover, many real world services (like Pinterest conten sharing service[13]) utilizing AWS, tend to use a mix of reserved—RI, on demand—ODI and spots EC2 instances to serve the fluctuating daily traffic, utilizing also the AWS auto-scaling feature. Using RI instances all the time they provide a baseline capacity, paying an initial fixed cost and a very small hourly cost. Normal ODI instances are used for serving the expected capacity for the day workload, with no upfront cost, only an hourly one. Additionally, peaks can be handled utilizing spot EC2 instances, which generally cost less than the ODI instances. AWS Auto scaling ensures that the number of instances used increases during demand spikes, and decreases automatically when demand is reduced.

5. SMART CLOUD MONITORING—SuMo

CUO is integrated in Smart cloud Monitoring (SuMo), an tool we develop for collecting monitoring data from Amazon Web Services (AWS) and analyzing them. SuMo includes and other intelligent mechanisms for resource profiling, performance spike detection, performance prediction, which are however out of the scope of the current work.

SuMo tool is open-source and available through github.[14] It is also is written in Python; utilizing the boto framework[15] for communicating with AWS. It also uses the SciPy[16] and NumPy[17] libraries for scientific computations. For the ILP technique SuMo interfaces with IBM ILOG CPLEX Optimizer,[18] which is free for noncommercial purposes, interfacing with other (more open) ILP solvers is also under way.

6. RESULTS

6.1. Configuration Scenarios and Metrics

The proposed Cost and Utilization Optimization (CUO) algorithm has been evaluated using synthetic data. Assumptions made on initial set S of M instances, and their workload signal WL for a period of $T = 24$ hours (step size equal to 1 hour), while CUO calculates a new set Snew of instances for serving the same load.

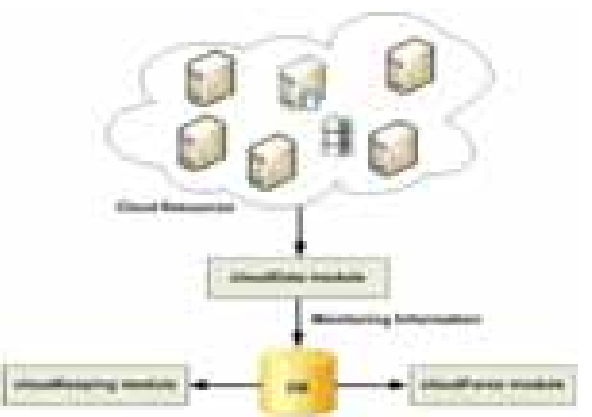


Fig. 1. SuMo—Smart cloud monitoring basic modules.

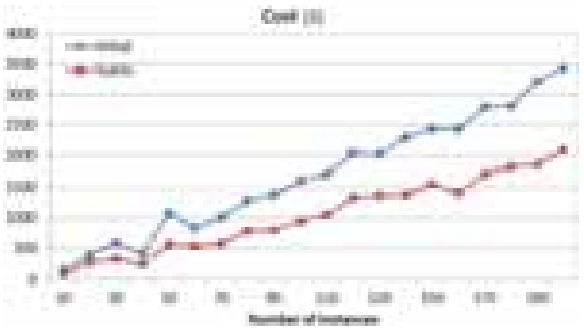


Fig. 2. The total cost (measured in $) of the instances of the initial set *S*, and the new set *S*new, selected by CUO.

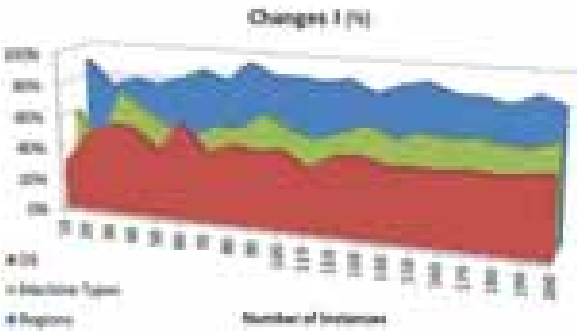


Fig. 4. The % of changes in the OS, Machine types and Regions in the new set of instances *S* new selected by SuMo, in relation to the original set *S*.

Even though real data could be more useful in the evaluation of the proposed algorithm, such data are not publically available. Nevertheless, this paper identifies a number of configuration scenarios that correspond (most of them) to real use cases, such as the Pinterest's[13]

- *Flat*, where every instance's OS and region of operation can change by CUO. This is not a very realistic scenario, however is merely used so as to get some baseline results to compare with. Additionally, the workload of each instance is selected uniformly in the range of [0, 100] percent of the instance's machine type capacity. Other workload ranges ([0, 20] and [80, 100] percent) are also evaluated.
- Region-constrained, where every instance's OS and region of operation cannot change. The workload scenarios of the Flat configuration are used.
- Cyclical-load, where the total workload changes over time causing the initiation of more instances, following a pattern similar to the one presented in Figure 1, for the Pinterest use case. Also, we assume that all instances' OS and region of operation cannot change.

Additionally, this work considers only On-Demand Instances (ODI) and their exact type is selected uniformly among the 196 different ODI instance types. The weighting coefficient W in the optimization function of was equal to 0.5. The value of parameter 'f' was set to 1. Unless stated otherwise, the aforementioned parameter values were the ones used in the results that follow. A number of experiments with different values for the parameters (WL, T, f, W) were also performed; for the sake of brevity, only some of the corresponding results are presented here.

The main metrics of interest are the capacity, the cost and the utilization of the instances selected by SuMo, which are compared to those of the initial set of instances. Furthermore, it is extended in tracking the changes of machine types in the new set of instances *S*new selected by SuMo.

6.2. Performance

6.2.1. Flat Configuration

Initially a number of experiments to study the role played by the number of running instances M was performed. Figure 2 shows the total cost of the instances in the original set S and in the new set *S*new selected by SuMo, assuming an observation period equal to T. The instances selected by SuMo result in a 40–50% reduction in the cost paid by the user. In both instance sets, S and *S*new, the total cost increases close to linearly with the number of running instances M, as expected.

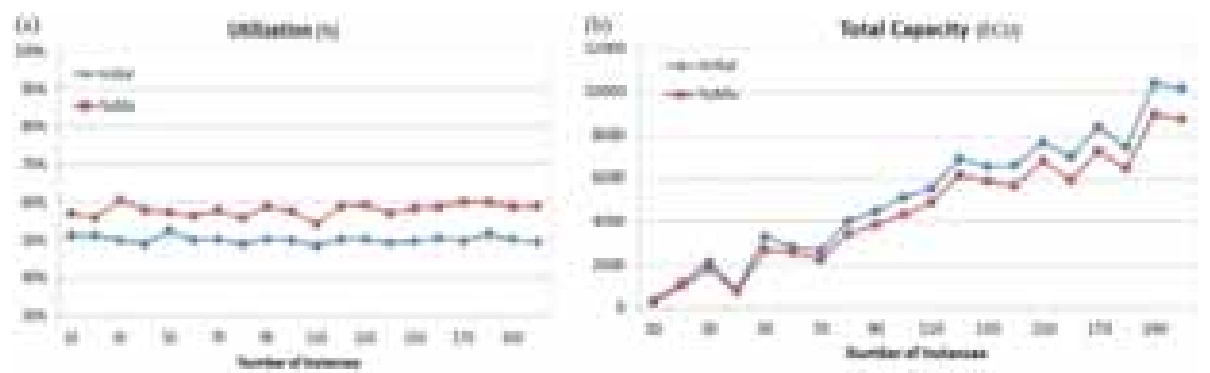


Fig. 3. The % utilization and the total capacity (measured in Elastic Compute Units—ECU) of the instances of the initial set *S*, and of the new set *S*new, selected by CUO.

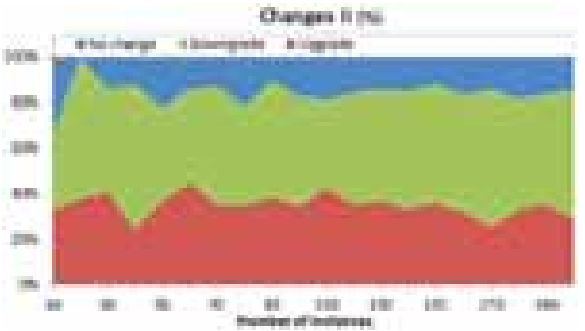


Fig. 5. Percentage of instance type upgrades, downgrades and no changes, in the new set of instances *S*new selected by SuMo, in relation to the original set *S*.

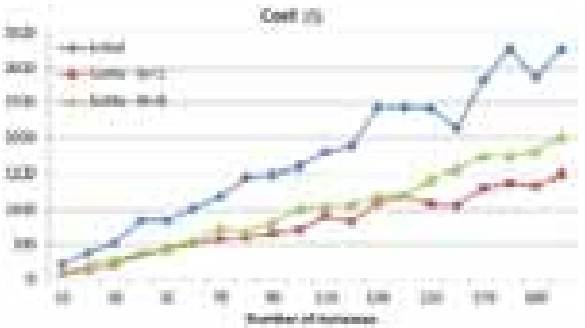


Fig. 6. The total cost (measured in $) of the instances of the initial set *S*, and the new set *S*new, selected by CUO, for $W = 0$ and $W = 1$.

Figure 3 shows the utilization of the cloud resources used and the total capacity (measured in EC2 Compute Units—ECU) of the instances in *S* and in *S*new. The instances' utilization is defined as the ratio of the instances average workload *WL* during the period of observation *T*, over the total capacity *P* of instances, defined. This paper observe that the instances selected by SuMo result in a higher utilization ratio (Fig. 3(a)), around 10% higher, due to the smaller total capacity instances chosen (Fig. 3(b)). However, the improvement in utilization shown in Figure 3(a) (∼10%) of *S*new in comparison to *S* is higher than the percentage of capacity reduction, shown in Figure 3(b), leading to the conclusion that the CUO algorithm also achieves a better matching between the requested workload and the offered capacity than the one randomly decided for the initial set *S*. The efficiency of the matching provided by CUO algorithm is also exhibited by the fact that the utilization achieved by the new set of instances *S*new, remains unaffected by the number of running instances *M*. One would expect that a larger number of running instances would leave room for more "errors" in the assignment of instances to resources and inefficient utilization of the available capacity, but this does not appear to be the case. It is also noted that the utilization of the instances in the initial set *S* remains unaffected by *M*, since the values of the related parameters (e.g., workload) were chosen, as already mentioned, from a uniform distribution.

Experiments with workload ranges [0, 20] and [80, 100] produce similar results.

In this paper Figure 4, illustrate the changes in the machine types, operating systems and regions in the new set of instances *S*new selected by SuMo, in relation to the original set *S*. It is observed that the CUO algorithm results almost always in regional changes, while the changes in the machine type and the operating systems are rather rare. Machine type changes relate both to utilization and cost factors, while regional and operating system changes relate only to cost parameters.

Figure 5 shows the percentage of upgrades, downgrades and the no changes of instances in the new set Snew, selected by SuMo. It's an upgrade (or downgrade or no change) when the workload of an instance in set *S* is executed in a higher (or lower or equal, respectively) capacity instance in the new set *S*new. From the experimental analysis, it is observed that the CUO algorithm results mostly (40%–50%) in downgrades, while fewer upgrades (30%–40%) and even fewer no changes (10%–30%). These percentages are highly affected by the chosen parameters, nevertheless, they demonstrate, generally the efficient operation of the CUO algorithm used in SuMo, since not only downgrades are performed,

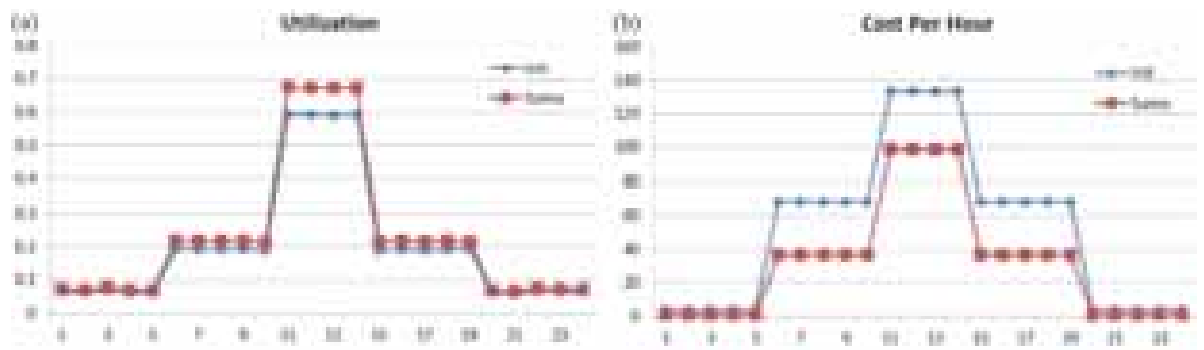


Fig. 7. The % utilization and the cost per hour of the instances of the initial set *S*, and of the new set *S*new, selected by CUO, for the cyclical-load configuration.

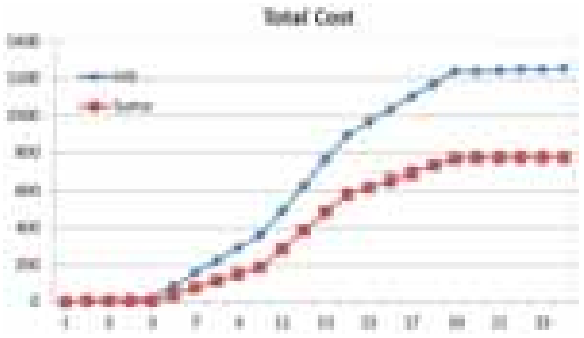


Fig. 8. The total cost (cumulative sum) of the instances of the initial set *S*, and of the new set *S*new, selected by CUO, for the cyclical-load configuration.

as one would normally expect so to increase resource utilization and decrease total cost, but more intelligent and less obvious choices, often resulting in upgrades, are also considered.

From the Figure 6, it is inferred that the instance cost of the new set Snew, selected by CUO becomes halved when compared to the initial set cost for the instance 140.

6.2.2. Cyclical-Load Configuration

In the experiments performed for the Cyclical-load configuration, this paper assumed a setting similar to the one presented for the Pinterest's use case.

In particular, during a 24-hour period there is a variable workload that causes the initiation of new instances, so as to serve the increasing load. In Figure 7, observe that the utilization and the cost per hour increase during peak hours. Even in this configuration, with time-varying characteristics, CUO achieves improved utilization and smaller cost. Figure 8 shows the cumulative sum of the cost.

6.2.3. Region-Constrained Configuration

The results obtained for the region-constrained configuration were similar to the Flat configuration, except the fact that only machine type changes occurred.

7. CONCLUSION AND FUTURE PLAN

This paper proposes a Cost and Utilization Optimization (CUO) algorithm for optimizing the cost and the utilization of a set of running Amazon EC2 instances. CUO receives information regarding the current set of instances used (their number, type, utilization) and proposes a new set of instances for serving the same load, so as to minimize cost and maximize utilization, or increase performance efficiency. CUO is integrated in Smart cloud Monitoring (SuMo), an open-source tool is developed for collecting monitoring data from Amazon Web Services (AWS) and analyzing them. To the best of our knowledge it is the first time that such a resource resizing algorithm is applied in an actual (in terms of resources' cost and characteristics) public cloud provider instead in a generic system, containing virtual resources. A number of experiments are performed under realist AWS configuration scenarios, where CUO increases the utilization of the resources used, and thus the efficiency in the use of the infrastructure resources, while lowering the accumulated resource costs.

References

1. S. Meng, Reliable state monitoring in cloud datacenters, *2012 IEEE Fifth International Conference on Cloud Computing* **(2012)**, pp. 951–958.
2. J. S. Ward and A. Barker, Semantic based data collection for large scale cloud systems, *Data-Intensive Distributed Computing (DIDC '12)*, ACM **(2012)**, pp. 13–22.
3. J. Weng, Event detection in twitter, *Proceedings of the Fifth International Association for the Advancement of Artificial Intelligence (AAAI) Conference on Weblogs and Social Media*, HP Labs Singapore **(2011)**, pp. 401–408.
4. A. Dastjerdi *SPE* 42, 501 **(2012)**.
5. Z. H. Wang, *Applied Mechanics and Materials* 48–49, 43 **(2011)**.
6. Newvem, http://www.newvem.com/.
7. Amazon Elastic Compute Cloud, http://aws.amazon.com/ec2/.
8. S. De Chaves, *IEEE Communications Magazine* 49, 130 **(2011)**.
9. S. Malkowski, CloudXplor: A tool for configuration planning in clouds based on empirical data, *ACM Symposium on Applied Computing* **(2010)**.
10. Amazon Web Services—AWS, http://aws.amazon.com/.
11. Amazon CloudWatch, http://aws.amazon.com/cloudwatch/.
12. Nagios, http://www.nagios.org/.
13. Pinterest use case, http://www.theregister.co.uk/2012/04/30/inside_pinterest_virtual_data_center.
14. SuMo-tool, https://github.com/SuMo-tool.
15. SciPy, http://www.scipy.org/.
16. NumPy, http://numpy.scipy.org/.
17. IBM ILOG CPLEX Optimizer, http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/.
18. On Demand Instances pricing, http://aws.amazon.com/ec2/pricing/pricing-ondemand-instances.json.

Received: 25 April 2018. Accepted: 14 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1734–1738, 2018

RESEARCH ARTICLE

Extracting the Edge Information from the Digital Images for Cloud Storage Using Fractals and Two Dimensional Discrete Wavelet Transforms

S. Sundaramurthy*, A. Kavitha, S. Daniel Madan Raja, and Amitabh Wahi

Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam, Erode Dt. 63401, Tamil Nadu, India

The increase in demand for the machine vision based approaches in the area of industrial automation and remote sensing application requires effective and efficient classifiers for the classification of images or objects. The extraction of the important feature becomes one of the challenging issues related to the classification problem. The key challenge is to extract the minimum number of important features from the large set of image data. In this paper, the two dimensional discrete wavelet transform (2D DWT) analysis is used for extracting the most useful information from the images. The Haar and Daubechies wavelets are used for extracting wavelet coefficients from the input images. The key advantage of Discrete Wavelet transform over Fourier transforms is temporal resolution characteristic. The extracted features can be stored in the cloud for further analysis.

Keywords: Digital Image, Cloud Storage, Wavelets, Fractals, Image Analysis.

1. INTRODUCTION

Over the recent couple of years, many approaches are developed to address the challenges in storing feature vectors used for data analysis in the cloud area. Many approaches are designed for showing considerable advances towards effective and efficient way for extracting useful information from the data. The motivation behind this research work blooms from the desire to extract the useful information from the image data set for efficient storage in Cloud environment for further analysis. It is also known that the data available in the real world are mostly fractal in nature. The statistical analysis of signals and images implies to detect the origin of redundancies like that of the power spectrum. The Authors in Refs. [1, 2] introduced a multifractal based method for reconstructing the images from the most singular fractal manifolds. The continuous wavelet transform is used for extracting the wavelet coefficients. The researchers used those most singular fractal components for reconstructing the images using Fourier Transform.

The importance of self-similarity of images in the area of digital image processing is examined.[3] The author outlined the evolution of the notion of self-similarity from Mandelbrot's generators to Iterated Function Systems and fractal image coding. The research has also examined the application of fractal based methods for image de-noising and super resolution. Finally, the author examined the non-local image processing based on fractal image coding.

The research work in Ref. [4] stated that the fractal geometry is the most widely used method for the image analysis. Most of the applications used various techniques for computing fractal dimensions of fractal images. The box counting technique is applied in this research to compute fractal dimension using lower bound. The study indicated that the importance of obtaining the minimum box size for fractal dimension computation. The researcher[5] proposed a novel technique based on fractal dimension for colour texture analysis. Colour texture information is very important feature for processing natural images. The research examined the red, blue and green components of the images for characterizing the textures present in the images. The author has also examined the correlation between the colour components. Bouligand–Minkowski Fractal Dimension method is applied by both the methods. It is concluded that the proposed method provided better performance over other colour texture analysis technique.

In general, most of the data in the form signals and images in the real world are complex objects and possess a high degree of redundant information. It is unnecessary store all those details for performing decision making

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1734/005 doi:10.1166/jctn.2018.7369

analysis. With help of few important information available in the data. The statistical properties of signals and natural scenes reveal that natural images can be viewed through different segments, which are most probably fractal components in nature. Such Fractal Components are very informative about the geometry of the images from which they were extracted. Those Components are usually equal to edges or boundaries present in the Image. In this proposed work, the discrete wavelet transform analysis is used for extracting the most revealing parts of the images using wavelets such as Haar and Daubechies. The statistical properties of the extracted fractal components such as mean, standard deviation, entropy and energy are determined to form a feature vector. Those feature vectors can be used for object recognition, signal analysis and image reconstruction in the multi environment digital images.

2. MULTI FRACTAL FRAMEWORK

The statistical analysis of signals and images involves the detection of the origin of redundancies from the given image. The multifractal framework provides an efficient framework than the conventional method of detecting the edges or contours. This framework splits any image into a collection of fractal sets, from which one of them is supposed to be the most informative, that set is usually edge-like.[1, 2]

The multifractal framework is mostly well adapted to certain types of natural images, but the real world images can also be well described by this framework. The fractal components are the structures of great geometrical relevance, determined by the statistical properties of natural images. Among the fractal components, certain components called as most singular fractal manifolds, provides the information about the contours and edges in the scenes.[7, 8] Those multifractals or fractal components are well used to code compress the images in the real world applications. The fractal figures generally have the following features like no characteristic length, self-similarity and non-integer dimension called as Fractal dimension.[9] The Figure 1 shows a simple fractal structure.

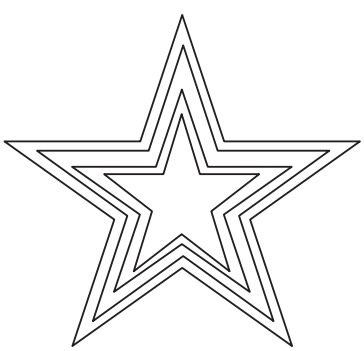


Fig. 1. Simple fractal image.

3. EDGE INFORMATION EXTRACTION

The 2-D discrete wavelet transform (DWT) is used to detect the fractal exponents from the image using wavelets like Haar and Daubechies (db2).[7] It is assumed that the images are free from Noise. The Figure 2 shows the flow diagram of the fractal decomposition of the images using Wavelet Transforms. The input image is resized and image gradient is determined by the Eq. (3.4). The 2-D DWT is applied on the gradient value to compute the four frequency subbands: approximate, horizontal, vertical and diagonal coefficients. Then, the fractal exponents are extracted from the four wavelet coefficients separately.[1, 2, 12] The threshold value is determined by taking the average of the 20, 60, 70 and 85 percentages of quantile.[11, 13, 14] The features like mean, standard deviation, entropy and energy values are computed from the four subbands to form a $[1 \times 16]$ feature vector. The procedure is applied to the entire training and testing data sets to extract the features.

Steps for extracting features from a scene using fractal decomposition

Step 1: Input image is resized into 128×128 image.
Step 2: Convert the resized image into a grayscale image.
Step 3: Determine the gradient magnitude of an image.
Step 4: Apply 2-D DWT on resulting image to obtain the four frequency bands obtained in the previous step.
Step 5: Determine the fractal components for each subband.

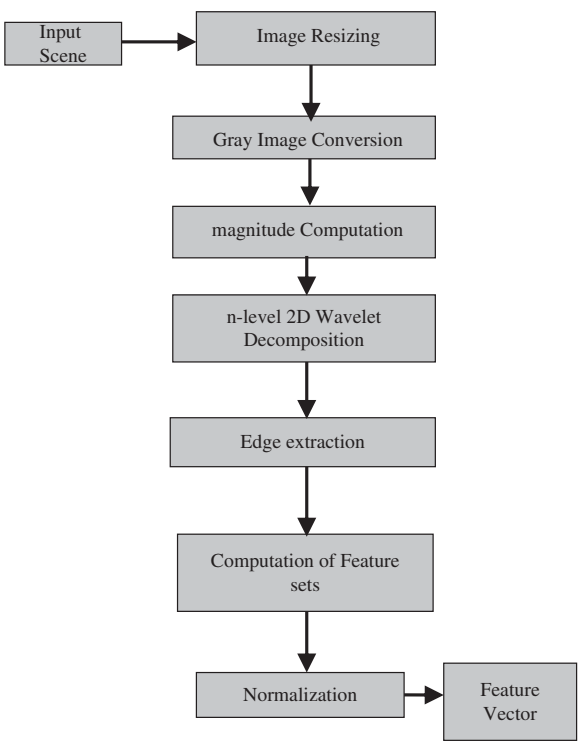


Fig. 2. Basic flow diagram of edge extraction.

RESEARCH ARTICLE

Step 6: Compute the features using fractal components and concatenate to form a feature vector.
Step 7: Normalize the feature vector using zero mean normalization technique.
Step 8: Apply the steps 1–7 for the entire training and testing data sets and store the feature vector for the classification purpose.

4. EXPERIMENT IMPLEMENTATION AND RESULTS

The proposed edge extraction method is tested using Matlab. The proposed procedure is tested on three classes of image datasets. The assumption is made that the images were free of Noise. The sample images are resized to 128×128 for the edge information extraction. The sample images were taken from the MIT database. The Haar and DB2 are the wavelets used for the determination of edge information. The feature vectors are extracted from the training and testing data sets of the Forest and Street scenes based on fractal decomposition using 2-D DWT. The images in the figures are not in original dimension and reduced for document purpose. The Figure 3 shows the sample forest image of size 256×256. The Figure 4 shows the grayscale image of 128×128. The Figure 5 shows gradient image. Figures 6(a) and (b) shows the level 1

Fig. 3. Sample forest image.

Fig. 4. Gray scale forest image.

Fig. 5. Grayscale forest image.

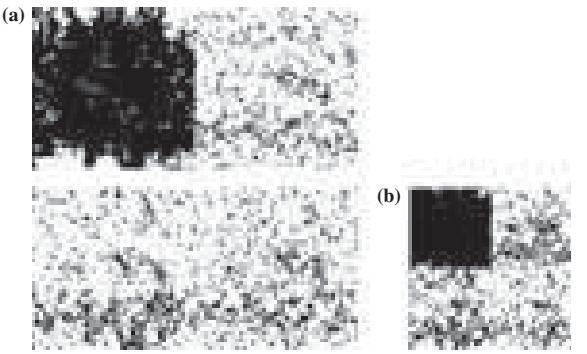


Fig. 6. Forest (a) Level 1 decomposition (b) level 2 decomposition using db2.

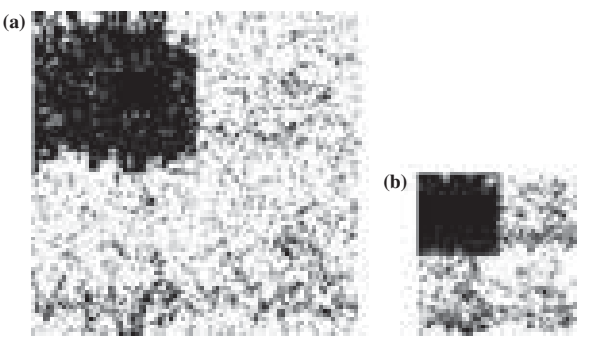


Fig. 7. Forest (a) Level 1 decomposition (b) level 2 decomposition using db2.

Fig. 8. Sample street image.

Fig. 9. Grayscale street image.

Fig. 10. Sample street gradient image.

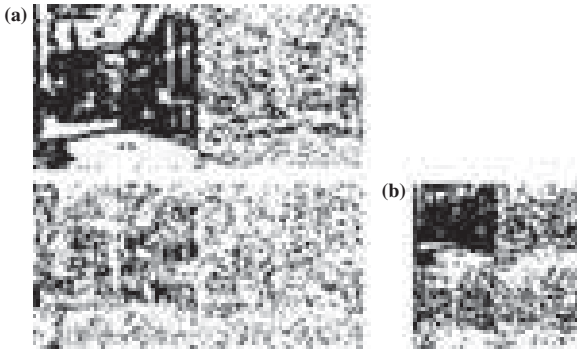


Fig. 11. Street (a) Level 1 decomposition (b) level 2 decomposition using db2.

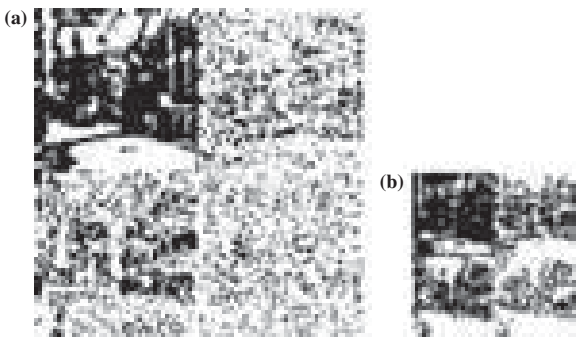


Fig. 12. Street (a) Level 1 decomposition (b) level 2 decomposition using db2.

and level 2 decomposition results using Haar. Figures 7(a) and (b) shows the level 1 and level 2 decomposition results using db2. Figure 8 is the forest image, Figure 9 is grayscale of it. Figure 10 is the gradient image of the forest image. Figures 11(a) and (b) shows the level 1 and level 2 decomposition results using Haar. Figures 12(a) and (b) shows the level 1 and level 2 decomposition results using db2.

5. CONCLUSION

This research work has examined the feature extraction process based on the multilevel fractal decomposition using Haar and Daubechies wavelets. The level-1 and level-2 wavelet decompositions are applied to extract the four frequency subbands. The coefficients such as approximation (LL), Horizontal (LH), Vertical (HL) and Diagonal (HH) are extracted for each dataset separately. The fractal components are extracted from each frequency band separately. Finally, the statistical features are computed by using only those fractal components. The fractal components in the image provide most useful information in the form of edges. Those Edge features of the images can be stored in the cloud for further analysis.

References

1. A. Turiel and A. del Pozo, *IEEE Transactions on Reconstructing Images from Their Most Singular Fractal Manifold Image Processing* 11, 345 **(2002)**.
2. A. Turiel and A. del Pozo, Understanding Multifractality: Reconstructing Images from Edges **(2008)**, pp. 1–13, aarXiv:cond-mat/9811372v1 [cond-mat.stat-mech].
3. M. Ebrahimi and E. R. Vrscay, Self-similarity in imaging, 20 years after fractals everywhere, *Proceedings of the International Workshop on Local and Non-Local Approximation in Image Processing*, LNLA, Lausanne, Switzerland **(2008)**, pp. 165–172.
4. Ajay Kumar Bisoi and J. Misra, On calculation of fractal dimension of images, Pattern Recognition Letters, Elsevier **(2001)**, Vol. 22, pp. 631–637.
5. A. RicardoBackes, D. Casanova and O. M. Bruno, *Pattern Recognition* 45, 1984 **(2012)**.
6. A. Arneodo, N. Decoster, and S. G. Roux, *The European Physical Journal B* 15, 567 **(2000)**.
7. J. Grazzini, A. Turiel, Hussein Yahia, and Isabelle Herlin, *Proceedings of the International Society for Photogrammetry and Remote Sensing* 35, 1125 **(2004)**.
8. T. Chen, Q. H. Wu, R. Rahmani-Torkaman, and J. Hughes, *Pattern Recognition* 35, 199 **(2002)**.
9. M. F. Barnsley, Fractals Everywhere, Academic Press, Cambridge, MA **(2014)**.
10. H. Pan and L.-Z. Xia, *IEEE Transactions on Neural Networks* 19, 2132 **(2008)**.
11. Maxvell Arulraj, Amir Nakib, Y. Cooren, and P. Siarry, *Applied Mathematical Sciences* 8, 131 **(2014)**.
12. Nadia M. G. Al-Saidi, Arkan J. Mohammed, and Adil M. Ahmed, *Applied Mathematical Sciences* 8, 117 **(2014)**.
13. S. Sundaramurthy and Amitabh Wahi, *Engineering and Technology International Journal of Computer, Information, Systems and Control Engineering* 8, 1381 **(2015)**.
14. Amitabh Wahi, S. Sundaramurthy, and C. Palanisamy, *Australian Journal of Basic and Applied Sciences* 2, 7 **(2014)**.

Received: 26 April 2018. Accepted: 14 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1739–1742, 2018

Deep Learning for Premature Divination of Diabetes

S. Naveena* and M. Gunasekaran

Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam 638401, Tamil Nadu, India

With the development of the science and technology, the infection or deficiency among the people of all age group is also increasing. Now a days, diabetes or blood pressure is becoming as one of the common characteristic possessed by each human being. There are so many research works carried out for predicting and diagnosing diabetes, whereas it predicts after the occurrence of diabetes. The proposed work concentrates on earlier prediction of diabetes with some symptoms which occur in advance and to help patient to prevent from being diabetic. Deep belief network (DBN) is used as a base classifier to split the given data set into diabetic and non diabetic. C4.5 algorithm is used to classify the diabetic dataset into type 1 and type 2 diabetic. The prediction provides higher accuracy as the algorithms handles both continuous and discrete values effectively.

Keywords: Missing Value, Outliers, Deep Belief Network, C4.5, ID3 Algorithm.

RESEARCH ARTICLE

1. INTRODUCTION

Medical data set involves huge amount of patient's information, data mining can be implemented for extracting knowledge from the data set given. It also helps to identify the relationship between different attributes in the medical data. Diabetes is a common health issue in which the plasma glucose plays a vital role. Even though the main reason is the glucose content, there are also some different factors like height, weight, genetic issues and immunity which contribute in diabetes. The early prediction and prevention, makes the patient free from diabetes.

Diabetes may also lead to much chronic disease like massive heart attack, blindness due to iris attack, kidney failure, etc. All these occur only because of the immunity failure and not because of diabetes. So the earlier prediction of diabetes based on some symptoms may prevent the patient from not only diabetes but also from many dangerous diseases.

Earlier prediction of diabetes includes various data mining techniques, machine learning algorithms and statistics that use current and past data sets to acquire information from it and using that knowledge to predict future occurrences. By applying machine learning algorithms on medical data, important decision and prediction can be made. In this work we are using C4.5 a machine learning algorithm to predict the type of diabetes (Type 1 or Type 2). With this result as a reference, the system can provide exact solution for the earlier diagnosis of patients' risk. This system can provide a valuable solution with higher reliability.

Before classifying diabetic patients into type 1 and type 2, the available data set has to be classified as diabetic and non diabetic. To perform this classification deep belief network has been used. Deep belief network (DBN) is a simple, unsupervised network in which each sub network's hidden layer is visible to the next. When trained on an unsupervised data set, DBN can learn to reconstruct its inputs. The layers act as feature detectors.[1] After this, DBN can be trained with supervision to perform classification.[2]

2. LITERATURE SURVEY

Bhat et al. proposed a regression approach which deals with the missing values and outliers in the diabetic data set and replaced the missing values with the corresponding attribute mean. They used classical neural network model for prediction and applied it on the preprocessed data set.[2]

Aiswarya Iyer et al. used naive Bayse and Decision Tree classification to find out patterns from the diabetes data sets. They compared the performance of both algorithms against Pima Diaetes Data sets. The results projected the effectiveness of each proposed classification model.[3]

Sabibullah et al. designed the prediction model based on soft computing to compute the risks of diabetic patients. They have used genetic algorithm for classification on real time medical data. From the results the risk level of patient and the risk of heart stroke are predicted.[4]

Rajesh and Sangeetha used C4.5 algorithm to find out patterns from the data set. The missing values in data set are not considered.[5]

Vaishnav and Patel made a comparative study on the different methods for handling missing data (K-Means, KNN,

*Author to whom correspondence should be addressed.

classification etc.).[7] Dai and Ji implemented a MapReduce based c4.5 decision tree algorithm. They incorporated the original algorithm into the mechanism of Map and Reduce process and conducted different tests on massive synthetic data sets and compared performance and accuracy on single v/s multi node Hadoop cluster. From the results it is observed that the algorithm provides time efficiency and scalability.[8]

3. PROPOSED WORK

In proposed system the features are preprocessed, selected and are applied to Deep belief network (DBN) for classification. The data set is classified into two (diabetic or not). Then the diabetic dataset is further applied to C4.5 algorithm to check whether it is Type 1 or Type 2 diabetes. The whole process is depicted in the Figure 1.

3.1. Data Preprocessing

Missing data create various problems when analyzing and processing data in the database. Missing data occurs due to typographical error, or irrelevant values or null fields. The missing values in a database can affect the accuracy and performance of the classifier which results in difficulty to extract the meaningful information, loss of efficiency. It may be very difficult to obtain the data quality mining results from the incomplete data sets, hence this missing gaps need to be treated.[7]

In the Pima Indian diabetes data set there are various attribute which have null values. It is not possible that any patient who has 0 blood pressure or 0 plasma glucose level in his body. Hence it is necessary to impute missing values for good classification results or it will cause wrong classification results. The missing values can be imputed by using different techniques such as classification clustering.

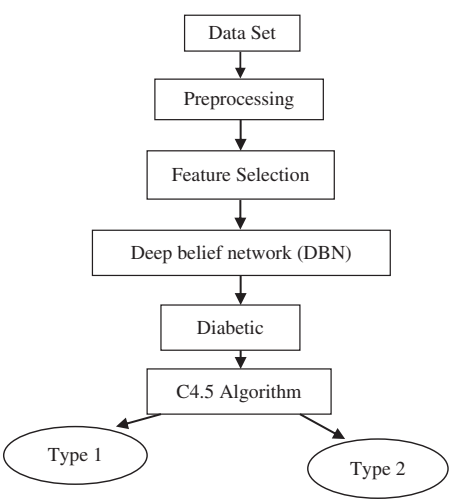


Fig. 1. Flow chart of proposed work.

Classification based technique has been used to replace missing values with attribute mean.

The following missing values imputation (MVI) algorithm is used to impute missing values in Pima Indian diabetes data set.

Input: Pima Diabetes Data set (X) having missing values
Output: Pima Diabetes Data set (X) after missing values imputation.

1. From attribute A_i, $i = 1, \ldots, N$ (where N is the total number of attribute) in data set X, retrieve attribute A of binary class variable.
2. According to the class variable divide the data set X in to X_d and X_{nd}.
   i. X_d = Sub data set having all diabetic records
   ii. X_{nd} = Sub data set having all non-diabetic records
3. Identify missing values from all attributes $A \in A_i = \{A1, A2, A3, \ldots, AN\}$, in each data set X_d and X_{nd} (consider 0 as missing value)
4. For each attribute $A \in A_i = \{A1, A2, A3, \ldots, AN\}$, in data set X_d, calculate attribute mean M_i, $i = 1, \ldots, N$ where N is the total number of attribute.
5. For each attribute $A \in A_i = \{A1, A2, A3, \ldots, AN\}$, calculate attribute mean M_{ni}, $i = 1, \ldots, N$ where N is the total number of attribute
6. Impute missing values in data set X_d with attribute mean M_i and impute missing values in data set X_{nd} with attribute mean M_{ni}.
7. Combine data set X_d and X_{nd}.
8. Return (X).

3.2. Deep Belief Network (DBN)

As Deep Belief Networks (DBN) name indicates, it is multi-layer belief networks. Each layer is Restricted Boltzmann Machine and they are stacked each other to construct DBN. The first step of training DBN is to learn a layer of features from the visible units, using Contrastive Divergence (CD) algorithm. Then, the next step is to treat the activations of previously trained features as visible unites and learn features of features in a second hidden layer. Finally, the whole DBN is trained when the learning for the final hidden layer is achieved.

This simple greedy learning algorithm works for training DBN. This is because that training RBM using CD algorithm for each layer looks for the local optimum and the next stacked RBM layer takes those optimally trained values and again look for the local optimum. At the end of this procedure, it is likely to get the global optimum as each layer consistently trained to get the optimum value.

Algorithm for training dataset with deep belief network is as follows:

1. Take the training data set and assign it directly to the visible unit.
2. Update the hidden unit states with the sigmoid activation function

$$\text{sigm}(x) = 1/(1 + \exp^{(-x)})$$

3. For ith hidden unit compute activation energy using the equation given below:

$$P(h_j \mid v) = \text{sigm}\Big(\sum_i w_{i,j} v_i + c_j\Big)$$

4. Assign the value of visible unit as 1 with the formula $P(v, h) = \exp(-E(v, h))/z$
5. Set the value of visible unit as 0 usnig the equation

$$P(v_i \mid h) = N\Big(\sum_j w_i, jh_j + b_i\Big)$$

6. Compute positive statistics for edge $(e_{ij}) = v_i * h_j$
7. Reconstruct the visible units using similar technique. For each visible unit, compute the activation energy and update the state.
8. Again update hidden units and compute $(e_{ij}) = v_i * h_j$ which is negative statistics for each edge.

After computing the sample probabilities of hidden layer, the logistic regression is applied to classify the data as diabetic or not diabetic. In the Logistic regression binary decision task function $0 \leq h(x) \leq 1$ is used to take decision whether data belongs to class 1 or class 0.

Sigmoid activation function is used to interpret result of logistic regression. If $h(x) \geq 0.7$ then we can say it is class 1 data point Else if $h(x) < 0.7$ then the data point belong to class 0.

3.3. C4.5 Algorithm

The diabetic data set obtained from deep belief network has to be classified based on its type. A supervised classification method called decision tree algorithm can be used for accurate classification. C4.5 is an improved version of ID3 algorithm with an additional feature of processing both continuous and discrete attributes. It also handles missing values and the attributes with minimum entropy is pruned after construction. The value of the information gain ratio is used to select the test characteristics at each node of the tree. The attribute with the highest information gain ratio is selected as the test feature for the current node. The following, is a pseudo-code for building C4.5 decision tree algorithm:

Input: Training data set (T); attributes (S).
Output: Decision Tree (T)
1. If tree is NULL then
2. Return failure
3. end if
4. set Tree $= \{\}$
5. for $a \in S$ do
6. setSplitInfo$(a, T) = 0$, Entropy$(a) = 0$
7. Compute entropy(a), SplitInfo(a, T), Gain(a, T), GainRatio(a, T)
8. end for
9. abest $=$ argmax{GainRatio(a, T)}
10. attach abest to Tree
11. for $v \in$ Values (abest, T) do
12. call C4.5(Ta, v)
13. end for
14. return Tree

Let c denote the number of classes, and $p(S, j)$ is the proportion of instances in that are assigned to jth class. Therefore, the entropy of attribute S is calculated as:

$$\text{Entropy}(S) = -\sum_{j=0} p(S, j) * \log p(S, j)$$

Accordingly, the information gain by a training data set T is defined as:

$$\text{Gain}(S, T) = \text{Entropy}(S) - \sum_{v \in \text{values}} (|r_{s,v}|/|r_s|) * \text{Entropy}(S_v)$$

where Values (Ts) is the set of values of S in T, Ts is the subset of T induced by S, and (Ts, v), is the subset of T in which attribute S has a value of v.

Therefore, the information gain ratio of attribute S is defined as:

$$\text{GainRatio}(S, T) = \text{Gain}(S, T)/\text{SplitInfo}(S, T)$$

The steps followed in C4.5 algorithm is described in Algorithm. The entropy, information gain and gain ratio can be calculated using above equations and recursively computed on sub tree.

4. CONCLUSION

Deep learning is a part of machine learning in Artificial Intelligence (AI) which has networks that are capable of learning from unstructured or unlabeled data. In this work deep belief network has been used as base classifier to split the large amount of dataset given as diabetic and non diabetic dataset. The C4.5 machine learning algorithm has been used to predict the type 1 and type 2 diabetes patients from the diabetic dataset. As C4.5 algorithm works on both continuous and discrete values, the algorithm provides maximum accuracy in prediction of diabetes.

References

1. Saravanakumar, Eswari, Sampath, and Lavanya, Predictive Methodology for Diabetic Data Analysis in Big Data, Elsevier, ISBCC **(2015)**.
2. V. H. Bhat, P. G. Rao, and P. D. Shenoy, An Efficient Prediction Model for Diabetic Database Using Soft Computing Techniques, Architecture, Springer-Verlag, Berlin, Heidelberg **(2009)**, pp. 328–335.
3. Aiswarya Iyer, S. Jeyalatha, and Ronak Sumbaly, Diagnosis of Diabetes Using Classification Mining Techniques, IJDKP, January **(2015)**, Vol. 5.
4. M. Sabibullah, V. Shanmugasundaram, and K. Raja Priya, *International Journal of Emerging Trends Technology in Computer Science* 2, 60 **(2013)**.
5. K. Rajesh and V. Sangeetha, *International Journal of Engineering and Innovative Technology (IJEIT)* 2, 224 **(2012)**.

6. Rajnik L. Vaishnav and K. M. Patel, *International Journal of Innovative and Emerging Research in Engineering* 2, 191 **(2015)**.

7. W. Dai and W. Ji, *International Journal of Database Theory and Application* 7, 49 **(2014)**.

8. Machine Learning Tutorials and Examples, https://www.toptal.com/machine-learning/machinelearningtheory-an-introductory-primer.

9. Anish Talwar and Yogesh Kumar, *International Journal of Engineering and Computer Science* 2, 3400 **(2013)**.

10. B. Brejova, T. Vina, and M. Li, Pattern discovery: Methods and software, Introduction to Bioinformatics, edited by S. A. Krawetz and D. D. Womble, Humana Press **(2003)**, Chap. 29, pp. 491–522.

11. Rajni Jain, Rule generation using decision trees, IASRI Md. Geaur Rahman, Md. Zahidul Islam, A decision tree-based missing value imputation technique for data pre-processing, *Proceedings of the 9-th Australasian Data Mining Conference (AusDM'11)*, Ballarat, Australia.

12. Gauri D. Kalyankar, Shivananda R. Poojara, and N. V. Dharwadkar, Weblog analysis using hadoop, *National Research Symposium on Computing-RSC 2016*, December **(2016)**, ISBN: 978-81-931456-1-8.

13. Sadhana and Savitha Shetty, *International Journal of Emerging Technology and Advanced Engineering* 4, 626 **(2014)**.

14. A. Ravishankar Rao, Atul Chhabra, Rajarshi Das, and Vikash Ruhil, A framework for analyzing publicly available healthcare data, *17th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Boston, MA, USA, IEEE Xplore, April **(2016)**, pp. 653–656.

Received: 17 April 2018. Accepted: 14 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1743–1749, 2018

RESEARCH ARTICLE

Comparative Analysis of Machine Learning Approaches for Twitter Sentiment Analysis

M. Nivaashini*, R. S. Soundariya, and P. Thangaraj

Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Erode 638401, India

Sentiment investigation is predominantly terrified with discovering and categorizing opinions or emotions that are expressed within a text. Nowadays, imparting sentiments and conveying emotions over social networking websites has turned out to be exceptionally normal. Therefore, a lot of data is made each day, on which mining can be feasibly performed to improve eminence of the data. Emotion analysis on such information can end up being instrumental in making a consolidated conclusion on particular yields. Twitter sentiment investigation deals with organizations ability to screen open feeling towards the items and occasions related to them dynamically. Twitter sentiment investigation regularly turns into a troublesome chore because of the existence of jargons and spelling mistakes. Twitter bounds the size of a tweet to 140 characters through new words which makes it harder to analyze and ascertain the sentiment as correlated with the typical opinion examination. Greatest prevailing studies about Twitter sentiment investigation are focused on the mining of new sentiment features. Yet, the selection of the pre-processing technique is unobserved. The proposed framework discusses the effects of text pre-processing method for efficient feature extraction and selection, and then classifies the tweets as positive, negative, neutral and irrelevant using machine learning approaches like Naïve Bayes, J48, Support Vector Machine (SVM), Random Forest Tree, Multilayer Perceptron, *k*-means etc. The best appropriate algorithm for twitter sentiment analysis is estimated by performance measurements like Precision, Recall, Accuracy, F1 measure, and so on.

Keywords: Twitter Sentiment Analysis, Tweets, Machine Learning Algorithms.

1. INTRODUCTION

Twitter, with more than 313 million[1] month to month dynamic clients and more than 500 million tweets for each day,[2] has now turned into a goldmine for organizations and people who have a strong political, social or monetary interest for keeping up and improving their clout and fame. Sentiment analysis compromises these organizations the capacity to monitor various real time social networking platforms. Sentiment analysis is the practice of inevitably spotting whether a text fragment contains opinionated or emotional substance, and it can besides decide the content's extremity. The main aim of Twitter sentiment classification is to order the sentiment extremity of a tweet as positive, negative, neutral or irrelevant.

Sentiment analysis methodologies can be comprehensively classified in two classes-dictionary based and machine learning based. Dictionary based approach is unsupervised as it proposes to perform investigation utilizing dictionaries and a scoring strategy to assess suppositions. Whereas machine learning approach includes the utilization of feature extraction and the corresponding feature set can be used to train the model and some dataset.

The elementary phases for performing sentiment analysis comprises data gathering, pre-processing of data, feature extraction, selecting features, sentiment detection and performing classification either by means of basic calculation or else machine learning approaches.

Twitter sentiment analysis is quite troublesome because of several reasons. Short messages inferable from as far as possible are an intense issue. Nearness of slangs, emoji's and incorrect spellings in the tweets makes an extra advance of pre-processing mandatory. There are a few strategies for feature extraction that can be utilized to gather related highlights from the content. These element extraction techniques can be adequately applied to tweets as well.

Tweets are typically made out of deficient, noisy and inadequately organized sentences, sporadic articulations, poorly framed words and non-lexicon terms. A sequence of pre-processing techniques (e.g., evacuating stop words, expelling URLs, swapping negations) are used before feature selection, in order to diminish the measure of noise

*Author to whom correspondence should be addressed.

in the tweets. Pre-processing is performed broadly in existing methodologies, particularly in machine learning-based methodologies.[3–6] Notwithstanding, few investigations focus around the impact of pre-processing technique on the execution of Twitter sentiment analysis.

A considerable measure of research has been done on Twitter data with a specific end goal to characterize the tweets and examine the outcomes. The proposed work plans to survey a portion of the inquires about in this area and concentrate how to perform different pre-processing and highlight extraction strategies keeping in mind the end goal to decide the best appropriate machine learning calculation for sentiment analysis on Twitter data utilizing *R*. The scope of this paper is constrained to that of the machine learning models and it additionally demonstrates the correlation of efficiencies of these models with each other.

The remainder of this paper is structured as follows. Related works and background are discussed in Section 2. The evaluation approach is presented in Section 3. The experimental results are presented in Section 4. Finally, Section 5 briefs, the conclusion and future work of the proposed framework.

2. RELATED WORKS

Sentiment Analysis is the examination of how feelings and perspectives can be identified with emotions and disposition appears in characteristic dialect regard to an event or occasion. Today occasions demonstrate that the sentiment analysis has achieved accomplishment which can go through positive and negative sentiments as well as manage conduct and feelings for various points and dialects. In the study of sentiment analysis, researcher utilize distinctive procedures for foreseeing the social feeling and emotion through the text and dialect.[7]

2.1. Pre-Processing Techniques

Most prevailing methodologies[3–6, 22–25] to find the sentiment polarity of tweets relate text pre-processing (e.g., POS, removing URLs, expanding acronyms, replacing negative mentions, stemming, removing stop words) to decrease the noise level in the tweets. The theory is that data pre-processing reduces the noise in the given content, and it should enhance the execution of the classifier and accelerate the classification procedure.

Haddi et al.[26] discovered the part of sentiment analysis in pre-processing of movie reviews. The tentative results show that the exactness of sentiment classification may be significantly improved using appropriate features and representation after pre-processing. Saif et al.[6] considered the impact of various stop words expulsion strategies for sentiment classification of tweets and in the case of evacuating stop words influences the execution of Twitter classifiers. They applied six different stop words identification methods to six different Twitter datasets and observed how removing stop words affects two supervised sentiment classification methods. They assessed the impact of removing stop words by observing fluctuations on the level of data sparsity, the size of the classifier's feature space and its classification performance. Using precompiled lists of stop words negatively impacted the performance of Twitter sentiment classification approaches.

Saif et al.[6] considered the impact of various stop words expulsion strategies for extremity arrangement of tweets and in the case of evacuating stop words influences the execution of Twitter feeling classifiers. They connected six diverse stop words distinguishing proof strategies to six distinctive Twitter datasets and watched how expelling stop words influences two directed opinion arrangement techniques. They evaluated the effect of expelling stop words by watching vacillations on the level of information sparsity, the span of the classifier's component space and its order execution. Utilizing precompiled arrangements of stop words contrarily affected the execution of Twitter assessment grouping approaches.

Saif et al.[22] found that pre-processing prompted a significant diminishment of the actual feature space. After pre-processing, the vocabulary estimate was decreased by 62%. However, they didn't talk about the impact on the execution of Twitter sentiment classifiers. Bao et al.[27] investigated the impact of pre-processing strategies on Twitter sentiment classification. They assessed the impacts of URLs, repeated letters, stemming, negation, and lemmatization. Exploratory outcomes on the Stanford Twitter Sentiment Dataset demonstrate that sentiment classification accuracy upturns when URL features reservation, repeated letters normalization and negation transformations are utilized, yet when stemming and lemmatization are applied, the accuracy diminishes.

Zhao[28] evaluated the accuracy of URLs, rehashed letters, negation, stop words, numbers and acronym in the binary Twitter sentiment classification task. The experiments demonstrate that the precision of sentiment classification increases after expanding abbreviation and swapping negation, although hardly shows variation when elimination of numbers, URL and stop words are used.

There are two essential approaches to distinguish sentiments from content. They are Symbolic methods and Machine Learning approaches.[29]

2.2. Symbolic Techniques

Much of the research in unsupervised sentiment classification using symbolic techniques makes use of available lexical resources. Turney[16] used bag-of-words approach for sentiment analysis. In that approach, relationships between the individual words are not considered and a document is represented as a mere collection of words. To determine the overall sentiment, sentiments of every word is determined and those values are combined with some aggregation functions. He found the polarity of a review based on

the average semantic orientation of tuples extracted from the review where tuples are phrases having adjectives or adverbs. He found the semantic orientation of tuples using the search engine Altavista.

Kamps et al.[17] used the lexical database WordNet[18] to determine the emotional content of a word along different dimensions. They developed a distance metric on WordNet and determined the semantic orientation of adjectives. WordNet database consists of words connected by synonym relations. Baroni et al.[19] developed a system using word space model formalism that overcomes the difficulty in lexical substitution task. It represents the local context of a word along with its overall distribution. Balahur et al.[20] introduced EmotiNet, a conceptual representation of text that stores the structure and the semantics of real events for a specific domain. Emotinet used the concept of Finite State Automata to identify the emotional responses triggered by actions.

One of the participant of SemEval 2007 Task No. 14[21] used coarse grained and fine grained approaches to identify sentiments in news headlines. In coarse grained approach, they performed binary classification of emotions and in fine grained approach they classified emotions into different levels. Knowledge base approach is found to be difficult due to the requirement of a huge lexical database. Since social network generates huge amount of data every second, sometimes larger than the size of available lexical database, sentiment analysis became tedious and erroneous.

2.3. Machine Learning Techniques

Machine Learning techniques use a training set and a test set for classification. Training set contains input feature vectors and their corresponding class labels. Using this training set, a classification model is developed which tries to classify the input feature vectors into corresponding class labels. Then a test set is used to validate the model by predicting the class labels of unseen feature vectors. A number of machine learning techniques like Naive Bayes (NB), Maximum Entropy (ME), and Support Vector Machines (SVM) are used to classify reviews.[11]

Joshi et al., applied the SVM (Support Vector Machine), Naïve Bayesian and the maximum entropy to compare the methods for twitter data analysis, for movie review from twitter for their dataset.[8] Rif'at et al., applied Several methods in Naïve Bayesian smoothing for analyzing the sentiment analysis and getting the optimum result 72.3% using the Laplace smoothing technique.[9] Brett Duncan et al., applied the neural network to classify the sentiment on tweets. The average accuracy (number of correctly classified tweets divided by the number of incorrectly classified tweets) was 74.15%.[10]

Spencer et al.,[13] has analyzed twitter data by adopting Naive Bayes approach. They have specified a tool for analyzing sentiments of tweets. They have shown the way to collect the data automatically and build the classifier to find negative and positive sentiments. They have concluded that the usage of Parts-of-Speech (POS) tags reduces the accuracy of classification process and we have included this in the proposed work. Han et al.,[14] has used character *N*-gram language and SVM classifier and has made a combination of both and came up with accuracy of 66.1% which is enhanced in the proposed work. Although their features don't perform well with their smaller data set, they could benefit from using a much bigger data set.

Xia et al.[15] used an ensemble framework for sentiment classification. Ensemble framework is obtained by combining various feature sets and classification techniques. In that work, they used two types of feature sets and three base classifiers to form the ensemble framework. Two types of feature sets are created using Part-of-speech information and Word-relations. Naive Bayes, Maximum Entropy and Support Vector Machines are selected as base classifiers. They applied different ensemble methods like Fixed combination, Weighted combination and Meta-classifier combination for sentiment classification and obtained better accuracy.

From the above reviews, there is a lack of proper and deep analysis of the impact of text pre-processing on Twitter sentiment classification. To fill this gap, this paper focus on evaluating the effects of text pre-processing on Twitter sentiment classification using different feature models and machine learning classifiers on Twitter datasets into four types of classification task.

3. PROPOSED METHODOLOGY

A dataset is created using twitter posts. Tweets are short messages with full of slang words and misspellings. So we perform a sentence level sentiment analysis as shown in Figure 1. This is done in three phases. In first phase pre-processing is done. Then a feature vector is created using relevant features. Finally using different classifiers, tweets are classified into positive, negative, neutral and irrelevant classes. Based on the number of tweets in each class, the final sentiment is derived.

3.1. Tweets Collection and Dataset Creation

Tweet collection involves gathering relevant tweets about the particular area of interest. The tweets are collected using Twitter's streaming API,[31,33] or any other mining tool (for example WEKA[32]), for the desired time period of analysis. The format of the retrieved text is converted as per convenience (for example JSON in case of Refs. [33, 34]). The proposed system uses the Sanders analytics dataset introduced by Go et al.[30,35] to perform necessary actions. Sanders analytics dataset consists of a total of 4685 tweets containing tweets of companies like Apple, Google and Microsoft. The collected tweets are annotated into four labels namely positive, negative, neutral and irrelevant. Table I shows the no. of tweets under each label.

Table I. Size of the twitter dataset for each emotion.

Positive	Negative	Neutral	Irrelevant
464	509	2153	1559

3.2. Pre-Processing of Tweets

The pre-processing of the data is a very important step as it decides the efficiency of the other steps down in line. It involves syntactical correction of the tweets as desired. The steps involved should aim for making the data more machine readable in order to reduce ambiguity in feature extraction. Keyword extraction is difficult in Twitter due to misspellings and slang words. So to avoid this, a pre-processing step is performed before feature extraction. Pre-processing steps include removing URL, avoiding misspellings and slang words. Misspellings are avoided by replacing repeated characters with 2 occurrences. Slang words contribute much to the emotion of a tweet. So they can't be simply removed. Therefore, a slang word dictionary is maintained to replace slang words occurring in tweets with their associated meanings. Domain information contributes much to the formation of slang word dictionary. Below are a few steps used for pre-processing of tweets.

3.2.1. Tokenization

Tokenization involves breaking of words in a tweet into tokens. The words of a tweet are separated by blanks. So, on encountering a space, a new token is formed. We need to consider several types of dictionaries when analysing the tweets.

- Lexical dictionary: This consists of maximum of English words that matches the word in a tweet with the list of words in dictionary when analysing tweets.
- Emoticon dictionary: Emoticons represent some meaning which are textual interpretations of the twitter. This dictionary helps to analyse tweets that consists of emoticons.
- Stop-words dictionary: Some words in a tweet don't have any kind of polarity and those words need not be considered for analysis. Those words are removed and considered as stop words. We conserve a dictionary containing all the stop words.

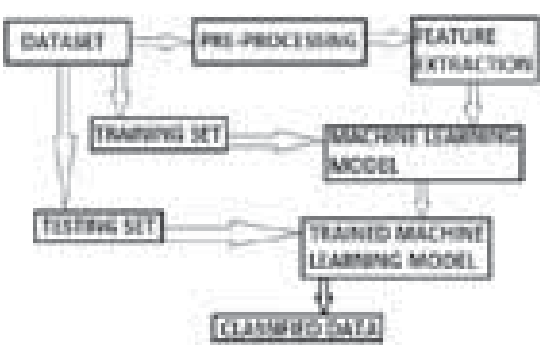


Fig. 1. Systematic design of the proposed system.

3.2.2. Normalization

Normalization of tweets improves accuracy which involves removal of noisy features. Data is normalized in these steps:

- Convert all the collected tweets into lower case
- Eliminate URLs by matching regular expression
- Remove punctuations and extra white spaces
- Remove stop words like 'a', 'the,' 'is' etc.
- Eliminate all "@username" and replace it with generic term AT USER
- Remove repeated letters in a word like 'hungryyyyyyyyyyyyyy,' 'huuuuuuuuuuuuuuungry' for 'hungry.' Although the lengthy words detect the strength of the sentiment, we are not concerned about this in the proposed approach.
- Remove comma, interrogation marks, double quotations at the starting and at the end of each word
- Remove those words that won't initiate with letters (like 20th, 7:35 am etc.).

3.3. Feature Vector Extraction

Feature extraction is done in two steps. In the first step, twitter specific features are extracted. Hashtags and emoticons are the relevant twitter specific features. Emoticons can be positive, negative, neutral or irrelevant. So they are given different weights. Positive emoticons are given a weight of '00,' negative emoticons are given a weight of '01,' Neutral emoticons are given a weight of '10' and Irrelevant emoticons are given a weight of '11.' There may be positive and negative hashtags. Therefore, the count of positive hashtags and negative hashtags are added as two separate features in the feature vector.

Twitter specific features may not be present in all tweets. So a further feature extraction is to be done to obtain other features. After extracting twitter specific features, they are removed from the tweets. Tweets can be then considered as simple text. Then using unigram approach, tweets are represented as a collection of words. In unigrams, a tweet is represented by its keywords. The proposed system maintains a negative keyword list, positive keyword list and a list of different words that represent negation. Counts of positive and negative keywords in tweets are used as two different features in the feature vector. Presence of negation contribute much to the sentiment.

So their presence is also added as a relevant feature. All keywords cannot be treated equally in the presence of multiple positive and negative keywords. Therefore, a special keyword is selected from all the tweets. In the case of tweets having only positive keywords or only negative keywords, a search is done to identify a keyword having relevant part of speech. A relevant part of speech is

adjective, adverb or verb. Such a relevant part of speech is defined based on their relevance in determining sentiment. Keywords that are adjective, adverb or verb shows more emotion than others. If a relevant part of speech can be determined for a keyword, then that is taken as special keyword. Otherwise a keyword is selected randomly from the available keywords as special keyword. If both positive and negative keywords are present in a tweet, we select any keyword having relevant part of speech. If relevant part of speech is present for both positive and negative keywords, none of them is chosen. Special keyword feature is given a weight of '1' if it is positive and '−1' if it is negative and '0' in its absence. Part of speech feature is given a value of '1' if it is relevant and '0' otherwise.

Thus feature vector is composed of 8 relevant features. The 8 features used are part of speech (pos) tag, special keyword, presence of negation, emoticon, number of positive keywords, number of negative keywords, number of positive hash tags and number of negative hash tags.

3.4. Tweets Classification

After the pre-processing and feature extraction steps classification is performed for training and validating the model's performance. The collected dataset is divided in two-training set (70%) and testing set (30%). The training set is used to train the classifier (machine learned model) while the testing set is the one on which the experimentation is performed. The classification task is performed using the algorithms like Naïve Bayes, SVM, Random Forest, Multi-Layer Perceptron, Ensemble Classifier, *K*-means as discussed in Refs. [36, 39]. In addition to the existing models, the proposed system uses J48 machine learning tree algorithm in Twitter sentiment classification.

J48 classifier is a simple C4.5 decision tree for classification. It creates a binary tree. The decision tree approach is most useful in classification problem. With this technique, a tree is constructed to model the classification process. Once the tree is built, it is applied to each tuple in the database and results in classification for that tuple.[37, 38]

Algorithm [1] J48:

```
INPUT:
    D //Training data
OUTPUT
    T //Decision tree
DTBUILD (*D)
{
T = φ;
T = Create root node and label with splitting attribute;
T = Add arc to root node for each split predicate and label;
For each arc do
  D = Database created by applying splitting predicate to D;
  If stopping point reached for this path, then
    T′ = create leaf node and label with appropriate class;
  Else
    T′ = DTBUILD(D);
  T = add T′ to arc;
}
```

While building a tree, J48 ignores the missing values i.e., the value for that item can be predicted based on what is known about the attribute values for the other records. The basic idea is to divide the data into range based on the attribute values for that item that are found in the training sample. J48 allows classification via either decision trees or rules generated from them.

Table II. Comparison of evaluation metrics for various machine learning models in twitter sentimental analysis.

Models	Accuracy (%)	Precision (%)	Recall (%)
Naïve Bayes	89.5	88.5	91
SVM	90	87.5	93
Random forest	87.5	85.5	81
Multilayer perceptron	89.9	88	90
Ensemble Classifier	90	87.5	93
K-means	75	55	36
J48	94	92.5	94.5

4. RESULT ANALYSIS

This section discusses about the performance analysis of J48 algorithm and the comparative analysis of other machine learning methods[36, 39] used in twitter sentiment analysis. Once a classifier for sentiment analysis is selected, the trained model classifier must be validated using cross fold validation. The performance of the model can be determined using the following measures:

(a) Accuracy: It is measured by the fraction of number of correct predictions over total number of predictions. The accepted accuracy is usually in the range 70% to 90%. If a model is 100% accurate then it usually depicts that model over fits the data.

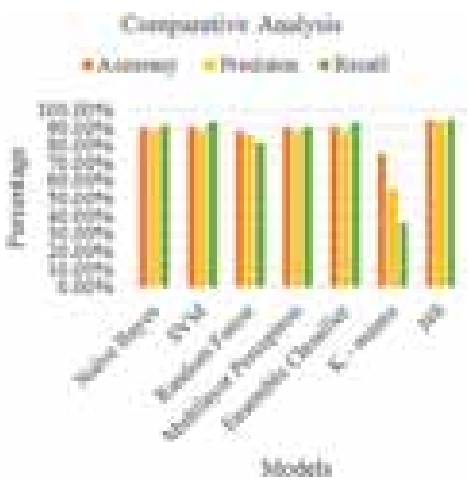


Fig. 2. Comparison of evaluation metrics for various machine learning models in twitter sentimental analysis.

(b) Precision: This measure shows how accurately the model makes predictions w.r.t each class. It is measured by number of correct predictions over total number of true positives and true negative examples.
(c) Recall: This measure shows the completeness of the model w.r.t each class. It is measured by number of correct predictions over total number of true positives and false negative examples.

Once the algorithms are trained, their performance is tested based on accuracy, precision and recall as shown in Table II and Figure 2. Based on the results, it is evident that J48 algorithm performs well than other machine learning algorithms in twitter sentimental analysis.

5. CONCLUSION AND FUTURE WORK

There are different Symbolic and Machine Learning techniques to identify sentiments from text. Machine Learning techniques are simpler and efficient than Symbolic techniques. These techniques can be applied for twitter sentiment analysis. There are certain issues while dealing with identifying emotional keyword from tweets having multiple keywords. It is also difficult to handle misspellings and slang words. To deal with these issues, an efficient feature vector is created by doing feature extraction in two steps after proper pre-processing. The proposed method of sentiment analysis on tweets involved several steps including tokenization of tweets, thorough preprocessing the tokens generated for removal of noise like URLs, punctuations, stop words etc. In the first step, twitter specific features are extracted and added to the feature vector. After that, these features are removed from tweets and again feature extraction is done as if it is done on normal text. These features are also added to the feature vector. Classification accuracy of the feature vector is tested using different classifiers like Naïve Bayes, SVM, Random Forest, Multi-Layer Perceptron, Ensemble Classifier, K-means and J48. All these classifiers have almost similar accuracy for the new feature vector, among which J48 algorithm achieves a bit higher accuracy than other machine learning algorithms. But it still lacks the dimension of diversity in the data. Along with this it has a lot of application issues with the slang used and the short forms of words. Many analysers don't perform well when the number of classes are increased. Also it's still not tested that how accurate the model will be for topics other than the one in consideration. All these shortcomings will be addressed in future.

References

1. Twitter, Available: http://en.wikipedia.org/wiki/Twitter.
2. The 2014 #YearOnTwitter, Available: http://blog.Twitter.com/2014/the-2014-yearontwitter.
3. E. Kouloumpis, T. Wilson, and J. Moore, Twitter sentiment analysis: The good the bad and the omg! *Proc. 5th Int. AAAI Conf. Weblogs Social Media* **(2011)**, pp. 538–541.
4. D. Terrana, A. Augello, and G. Pilato, Automatic unsupervised polarity detection on a Twitter data stream, *Proc. IEEE Int. Conf. Semantic Comput.*, Newport Beach, CA, USA, September **(2014)**, pp. 128–134.
5. H. Saif, Y. He, M. Fernandez, and H. Alani, Semantic patterns for sentiment analysis of Twitter, *Proc. 13th Int. Semantic Web Conf., Apr.* **(2014)**, pp. 324–340.
6. H. Saif, M. Fernandez, Y. He, and H. Alani, On stopwords, filtering and data sparsity for sentiment analysis of Twitter, *Proc. 9th Lang. Resour. Eval. Conf. (LREC)*, Reykjavik, Iceland **(2014)**, pp. 80–81.
7. R. Joshi and R. Tekchandani, Comparative analysis of Twitter data using supervised classifiers, *2016 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore **(2016)**, pp. 1–6. DOI: 10.1109/INVENTIVE.2016.7830089.
8. I. S. Jacobs and C. P. Bean, Fine particles, thin films and exchange anisotropy, Magnetism, edtited by G. T. Rado and H. Suhl, Academic, New York **(1963)**, Vol. III, pp. 271–350.
9. M. Kumar and A. Bala, Analyzing Twitter sentiments through big data, *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi **(2016)**, pp. 2628–2631.
10. R. A. Ramadhani, F. Indriani, and D. T. Nugrahadi, Comparison of Naïve Bayes smoothing methods for Twitter sentiment analysis, *2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, Malang **(2016)**, pp. 287–292.
11. G. Vinodhini and R. Chandrasekaran, *International Journal of Advanced Research in Computer Science and Software Engineering* 2, 282 **(2012)**.
12. E. Boiy, P. Hens, K. Deschacht, and M.-F. Moens, Automatic sentiment analysis in on-line text, *Proceedings of the 11th International Conference on Electronic Publishing* **(2007)**, pp. 349–360.
13. J. Spencer and G. Uchyigit, Sentimentor: Sentiment analysis of twitter data, *1st International Workshop on Sentiment Discovery from Affective Data* **(2012)**, pp. 56–66.
14. Q. Han, J. Guo, and H. Schtze, Codex: Combining an svm classifier and character n-gram language models for sentiment analysis on twitter text, *Second Joint Conference on Lexical and Computational Semantics*, Atlanta, Georgia **(2013)**, pp. 520–524.
15. R. Xia, C. Zong, and S. Li, *Information Sciences: An International Journal* 181, 1138 **(2011)**.
16. P. D. Turney, Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews, *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Association for Computational Linguistics **(2002)**, pp. 417–424.
17. J. Kamps, M. Marx, R. J. Mokken, and M. D. Rijke, Using Wordnetto measure semantic orientations of adjectives, *4th International Conference on Language Resources and Evaluation* **(2004)**.
18. C. Fellbaum, Wordnet: An Electronic Lexical Database (Language, Speech, and Communication) **(1998)**.
19. D. Pucci, M. Baroni, F. Cutugno, and A. Lenci, Unsupervised lexical substitution with a word space model, *Proceedings of EVALITA workshop, 11th Congress of Italian Association for Artificial Intelligence*, Citeseer **(2009)**.
20. A. Balahur, J. M. Hermida, and A. Montoyo, *IEEE Transactions on Affective Computing* 3, 88 **(2012)**.
21. C. Strapparava and R. Mihalcea, Semeval-2007 task 14: Affective text, *Proceedings of the 4th International Workshop on Semantic Evaluations*, Association for Computational Linguistics **(2007)**, pp. 70–74.
22. H. Saif, Y. He, and H. Alani, Alleviating data sparsity for Twitter sentiment analysis, *Proc. CEUR Workshop*, Sepember **(2012)**, pp. 2–9.
23. H. G. Yoon, H. Kim, C. O. Kim, and M. Song, *J. Informetrics* 10, 634 **(2016)**.
24. F. H. Khan, U. Qamar, and S. Bashir, *Appl. Soft Comput.* 39, 140 **(2016)**.
25. A. Agarwal, B. Xie, and I. Vovsha, Sentiment analysis of Twitter data, *Proc. Workshop Lang. Social Media, Assoc. Comput. Linguistics* **(2011)**, pp. 30–38.

26. E. Haddi, X. Liu, and Y. Shi, *Procedia Comput. Sci.* 17, 26 **(2014)**.
27. Y. Bao, C. Quan, L. Wang, and F. Ren, The role of pre-processing in Twitter sentiment analysis, *Proc. 10th Int. Conf. (ICIC)*, Taiyuan, China, April **(2014)**, pp. 615–624.
28. Z. Jianqiang, Pre-processing boosting Twitter sentiment analysis? *Proc. IEEE Int. Conf. Smart City/SocialCom/SustainCom (SmartCity)*, Sepember **(2015)**, pp. 748–753.
29. C. J. V. Rijsbergen, Information Retrieval, Butterworth-Heinemann, 2nd edn., Newton, MA, USA **(1979)**.
30. A. Go, R. Bhayani, and L. Huang, Twitter Sentiment Classification Using Distant Supervision, Stanford, Stanford, CA, USA, Project Rep. CS224N **(2009)**.
31. D. Zimbra, M. Ghiassi, and S. Lee, Brand-related twitter sentiment analysis using feature engineering and the dynamic architecture for artificial neural networks, *49th IEEE International Conference on System Sciences (HICSS)*, Koloa, HI, USA **(2016)**, pp. 1530–1605.
32. Varsha Sahayak, Vijaya Shete, and Apashabi Pathan, Sentiment Analysis on Twitter Data, (IJIRAE), January **(2015)**, ISSN: 2349-2163.
33. P. Barnaghi, J. G. Breslin, and P. Ghaffari, Opinion mining and sentiment polarity on twitter and correlation between events and sentiment, *IEEE Second International Conference on Big Data Computing Service and Applications* **(2016)**.
34. Nehal Mamgain, Ekta Mehta, Ankush Mittal, and Gaurav Bhatt, Sentiment Analysis of Top Colleges in India Using Twitter Data, (IEEE) ISBN-978-1-5090-0082-1 **(2016)**.
35. http://cs.stanford.edu/people/alecmgo/trainingandtestdata.zip.
36. Bhumika Gupta, Monika Negi, Kanika Vishwakarma, Goldi Rawat, and Priyanka Badhani, *International Journal of Computer Applications* 165 **(2017)**.
37. M. H. Danham and S. Sridhar, Data Mining, Introductory and Advanced Topics, Person Education, 1st edn. **(2006)**.
38. Aman Kumar Sharma and Suruchi Sahni, A Comparative Study of Classification Algorithms for Spam Email Data Analysis, *IJCSE* **(2011)**, Vol. 3, pp. 1890–1895.
39. Hima Suresh and S. Gladston Raj, An unsupervised fuzzy clustering method for twitter sentiment analysis, *IEEE International Conference on Computational Systems and Information Systems for Sustainable Solutions* **(2016)**.

Received: 17 April 2018. Accepted: 14 May 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1750–1753, 2018

RESEARCH ARTICLE

Assessment of Beer-Quiché Algorithm with Minmax-Q for Stochastic Multipath Routing in Mobile Ad-Hoc Network

D. Sathiya[1,*] and B. Gomathy[2]

[1] *Department of Computer and Science Engineering, Vivekanandha College of Technology for Women, Tiruchengode 637205, India*
[2] *Department of Computer and Science Engineering, Bannari Amman Institute of Technology, Sathyamangalam 638401, India*

In an Adhoc, the attackers try to interrupt the data packets being broadcasted in the network between the nodes, which travels from source to destination. To solve this problem, analysis has been made between Beer-Quiche and Minmax-Q theoretical routing model utilized for stochastic multipath routing in MANETs. In Beer-Quiche initially, the path setup and route selection is regulated using proposed approach. During each stage, the source node tracks the available paths, path residual bandwidth and policy of the attackers, who gathers information from previous stage to provide an attack. Depending on such selection, the optimal paths are selected by the source node for broadcasting the packets. Further, an optimal routing selection policy is proposed to broadcast the packet to its destination node during each stage of the Beer-Quiche game. In Markov routing method multiple paths has been computed between source–destination pairs and selects an energy-efficient path stochastically from those paths to forward the data packets. In addition, this protocol also secures data flow in the network as the packets are forwarded through random paths from the source node to the destination node. The performance analysis and numerical results shows that Beer-Quiche achieves significant performance gains in terms of energy consumption, delay, throughput and security than of Minimax-Q learning based routing.

Keywords: MANET, Beer-Quiche, Minmax-Q, Adhoc.

1. INTRODUCTION

Mobile ad-hoc network (MANET) can be installed easily and it is an economically feasible communication media for various applications. Although MANETs can easily be installed, the nodes of MANETs have limited battery power and it cannot efficiently serve the purpose of sensing and processing data during communications once the onboard battery power of the constituent nodes get exhausted. Therefore, it is[5,6] pricing-based methods[7,8] and game theoretic approaches[9,10] have been proposed in the past few years. The cryptographic approaches are usually computationally intensive and therefore difficult to implement in MANETs. The trust based mechanisms are application specific and not generic to be used in any kind of MANETs. These approaches demand trusted or tamper proof hardware in every node of a MANET. A few Game theoretic approaches have been introduced for strategic interactions among the nodes under a mathematical model. Game theoretic based security, defense mechanisms and decision making schemes may be used in the resource constrained networks. However, the performance of these techniques degrades when applied to an unfriendly MANET's environment and also when the attackers are adaptive in nature. We investigated and found that in ad hoc networks stochasticity in routing can be a good solution for making data paths secure and at the same time energy efficient. Stochastic routing is a mechanism that can be used to select the next hop in a path according to a probability distribution. This routing scheme can also secure the routing path as it explores multiple neighbor nodes for selecting the next hop in a path thereby forcing packets to use paths randomly.[11] In order to utilize the benefits of stochastic routing, various approaches have been proposed in the literature.[11–14] A Markov chain based framework is developed in Ref. [12] for balancing the network load and studying the routing performance in wireless sensor networks. The convergence of the framework is highly questionable while the network is very dense. The authors of Ref. [13] suggested a stochastic semi definite programming to deal with the location uncertainty of the nodes and obtain a clear understanding of tradeoff between message passing overhead and latency in the route discovery process. In Ref. [14], using a probabilistic local broadcast transmission model, the authors have showed that an index policy is not only optimal for stochastic routing but is also optimal for control transmission in an

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1750/004 doi:10.1166/jctn.2018.7372

ad hoc network. Though all these works have utilized the stochastic routing to optimize the performance of routing in network, we investigated and found that it can also be used to ensure the routing security at the same time. An optimal centralized stochastic routing algorithm for randomly duty-cycled wireless sensor network is proposed in Ref. [15]. In game theoretical model, a mathematical model uses an interaction strategy among the nodes. Models with individual enforcing distributed framework,[3] which uses equilibrium criteria to measure the optimal outcome of the game under adhoc scenario. Models with cooperation and interaction strategies,[4] interacts among the nodes and *rejects* the adaptation of the various attacks by the intruders. Models with decision making strategy[6] makes better provisions for resource constraints in the adhoc network. However, if the attackers are high adaptable inside the network, the performance of such game strategies degrade design hostile MANET's. Conventionally, the technique used in Ref. [2] meets the compatible routing strategy using path variation technique, which is a receiver centric model. Likewise, the technique in Ref. [7] uses the same strategy in static scenario. The technique in Refs. [2 and 8] uses routing auction scheme under multiple paths and this encourages better packet forwarding. Other packet routing technique uses two-hop relay protocol using a Markov chain theoretical framework.[9] The individual node decision to deliver the traffic is used in a forwarding game model.[8]

2. STOCHASTIC GAME FRAMEWORK

A stochastic game is a tuple (n; S; A1:::n; T; R1:::n), where n is the number of agents, S is a set of states, Ai is the set of actions available to agent i (and A is the joint action space A1:::An), T is a transition function S A S ! [0; 1], and Ri is a reward function for the ith agent S A! R. This looks very similar to the MDP framework except we have multiple agents selecting actions and the next state and rewards depend on the joint action of the agents. It's also important to notice that each agent has its own separate reward function. The goal for each agent is to select actions in order to maximize its discounted future rewards with discount factor. Stochastic games are very natural extension of MDPs to multiple agents. They are also an extension of matrix games to multiple states.

3. ALGORITHM 1: ITERATIVE MINIMAX-Q LEARNING

Step 1 (Initialization):
At each state s t and time t = 0, 1, . . . ,T; List the action set As(s t) of the source node and the action set AJ(s t) of the attackers; Set Q(s t, as, aJ) ← 1; Set V(s t) ← 1; Set θ (s t, a) ←1/|As(s t)|, ∀a ∈ As(s t);
Step 2 (Action a t s selection at time t):
for i := 1 to T do if (state st is not observed previously) then s t add into s h; else if (state st is observed previously)
then use previously listed As(s t), AJ(s t), Q(s t, as, aJ), V(s t) and θ (s t); an action a t s is chosen uniformly with probability p cur;
else an action a t s is chosen with probability θ (s t, as) in current state s t;
end
end
Step 3 (Online policy improvement and value estimation):
if (attackers action = a t J && source action = a t s && state transit from st to st + 1 && reward = u(s t, a t s, a t J))
then Q(s t, a t s, a t J) ← (1 − γt)Q(s t, a t s, a t J) + γt [u(s t, a t s, a t J) + αV(s t + 1)]; θ (s t) ← arg maxθ (s t) minθ (s t) a θ (s t, as)Q(s t, as, aJ); V(s t) ← min θJ (s t) 抗 a θ (s t, as)Q(s t, as, aJ); γ t + 1 ← $\mu \cdot \gamma$ t;
end
Step 4 (Termination): if (routing strategy is optimal)
then stop;
else Go to step 2;
end

4. IMPROVED BEER-QUICHE GAME MODEL FOR VERIFICATION OF NODES

The resource sharing is regarded as an indispensable task and considered as an important task for security purpose.

The participation of nodes leads to the problem of integrity and it avoids the malicious nodes. It reduces the network overhead using stochastic game theory model. In the present study, one player (player 1) is used for transmission of packets and another player (player 2) is to prevent the packets from malicious attack. This is achieved with the help of providing agreement between the MANET nodes. It considers the adjacent two nodes, say, *a* and *b*, which are authenticated for proper communication. In addition to this, a strategy is established for message transmission between the nodes *a* and *b*. The strategy given to node *a* is penalty or reward and it requests to establish communication with *b*. The node *a*, then holds two actions, namely, *reject* or *accept*. The action is put hold if the node *a* is *wimpy* and the action is allowed to take place, if the node *a* is *surly*. The node *a* with action *surly* is used for verifying the signature, if parser succeeds. However, if parser fails, the node cannot verify the signature, since the node is *wimpy*. The *surly* and *wimpy* node selection is done based on joining time of the node (either *a* or *b*) and the broadcasted message in a periodical manner through a central server. The node *a* can hold the public or private key information of itself, however, it cannot hold the public or private key information of node *b*. The Beer-Quiche model is used to verify the nodes based on the given set of rules and provides better message transmission between *a* and *b* nodes to improve the data integrity. The Beer-Quiche theoretical Game model is shown in Figure 1 and this represents the strategy of multi

RESEARCH ARTICLE

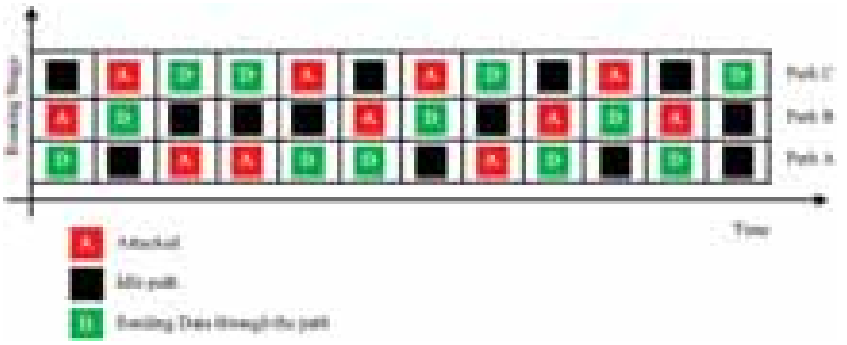


Fig. 1. Stochastic multipath routing game model.

path routing between the nodes a and b. The source node S with probability y, selects the node a as *surly* type and with probability $1-y$, the node a is selected as *wimpy*. The node a transmits the message with either of two strategies, namely, penalty or reward. The strategies of node a is unworthy, if the node is *wimpy*. The message is transmitted to the node b with probability y and it *accept*s the received message, once a proper authentication is made with *surly* node with reward action. On contrary, the node b loses the message, when it *reject*s the request from the node a (*surly*) after authentication with *penalty* action. The authentication of message request by the sending node is -1 and the receiver node is $+1$.

4.1. Notations

$\beta 1(a \mid c)$ Probability of c related to node *player* 1 after selecting a $\beta 2(b \mid c)$ Probability of c related to node *player* 2 after selecting b

$\mu(c \mid a)$ Guessing made by *player* 2 from the *player* 1 by viewing c

$ui(a \mid b \mid c)$ Pay-off factor associated with the message sent through the *player* 2 or the actions based on the *player* 1.

5. NUMERICAL RESULTS

Consider a MANET network infrastructure of node N, deployed in an area of 1500×1500 m^2 and implemented using NS-2 Simulation tool. The normalized node density is given as $x = (N-1)A/S$ and it defines the average neighbours surrounding each node under the transmission is A within the range of 300 m. The IEEE 802.11b and UDP is chosen as MAC standard and transport layer, respectively. The packet size is made fixed with a rate of 512 bytes and a constant bit rate source is used for traffic supply. The disjoint multiple paths is considered for simulation between the source (player 1) and destination (player 2) pair. The total packets transmitted are 300 and the residual bandwidth for multiple path is 4. The residual bandwidth is taken into consideration to compare the strategies of various routing from the source node. The attackers have the ability to train themselves and it leads to formation of optimal strategy against source node. From, the Figure 2 and Table I, it is seen that the bandwidth utilization is compared against proposed stochastic Beer-Quiche multipath routing (SBQMR), Stochastic multipath routing (SMR). From the results, it is seen that the average utilization of residual bandwidth against the five schemes shows that the SBQMR attains better or maximum utilization of residual bandwidth than the conventional scheme. The proposed strategy has a better caring on future utilization of residual bandwidth, which considers the maximum payoff during the present stage. The simulation is carried out with a total of 200 packets for transmission. The numerical results of SBQMR against and SMR is estimated against routing overhead, total packets delivered and average delay. It is assumed that nodes in the established path tend to forward the packets to meet the requirements. However, the nodes in other paths behave in a malicious way. From the Figure 2, it is found that the routing overhead of proposed method is found slightly higher than SMR.

This is due to the fact that proposed method takes into consideration the establishment of path and strategy of attackers. From the Figure 3, it is found that during

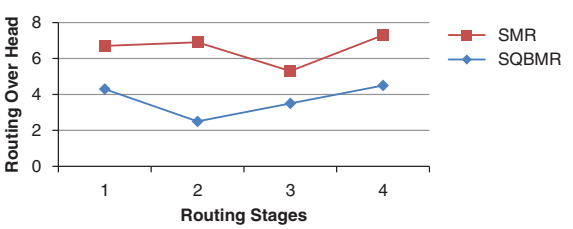


Fig. 2. Routing overhead versus stages.

Table I. Maximum utilization of residual bandwidth.

Bandwidth values	SMR utilization	BQMR utiization
0	0	0
5	0.141	0.163
10	0.262	0.292
15	0.291	0.303
20	0.339	0.343
25	0.360	0.391
30	0.370	0.384

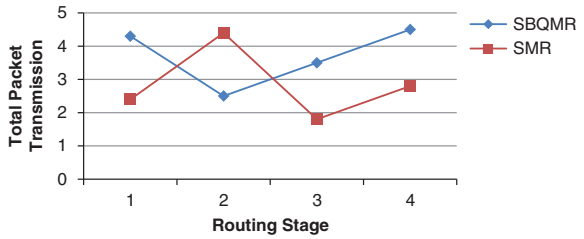


Fig. 3. Total packet transmissions.

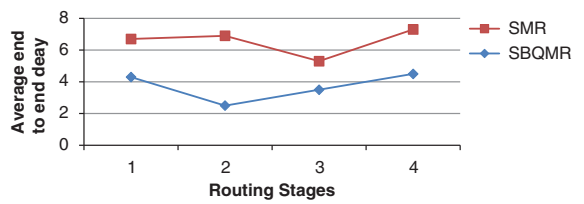


Fig. 4. Average end to end delay.

each stage of routing, the SBQMR delivers higher packets than SMR. The proposed multipath routing transmits maximum packets using stochastic paths during each routing stage. It is noticed that SBQMR uses less routing stages to transmit the packets than SMR. Finally; Figure 4 shows the average end to end delay during packet transmission. It defines the average time taken by the packets to reach the destination during each routing stage. It includes delays due to the propagation of packets, route discovery and retransmission in physical and routing layer. The results shows that the proposed method attains lesser average delay SMR. Also, it is seen that the delay increases as the stages increases. This is due to the fact that as the routing stages increases, paths are required to route the packets increases. Therefore, the proposed SBQMR method performs well with improved security ‘’/concerns reduced routing overhead, increased packets delivery and reduced average end-to-end delay.

6. CONCLUSION

The paper presents the analysis between stochastic Beer-Quiche with Minmax-Q game theoretical models for multipath routing in MANETs. The selection of routing strategy is done in a strategic manner and the route is switched among the multi paths by learning the strategy of attackers. Thus, the packets are broadcasted using stochastic Beer-Quiche game theoretical model by finding whether the node is *surly* or *wimpy*. Such data flow occurs in a stochastic way between the source and destination to evading the attackers. The numerical simulation verifies the performance of stochastic Beer-Quiche game theoretical model than the conventional method. The results shows that the Beer-Quiche method has a significant gains in terms of reduced routing overhead, increased packet transmission, reduced average end-to-end delay and increased utilization of residual bandwidth. The results proves that the proposed method finds the optimal route to ensure attacks, which are not present in the established path and it attains significant performance gain.

References

1. D. Sathiya and B. Gomathy, Improved security and routing path learning in MANETs using Beer-Quiche game theoretical model in cloud computing, Cluster Computing, Springer, Heidelberg, Germany **(2017)**.
2. S. Sarkar and R. Datta, A game theoretic framework for stochastic multipath routing in self-organized MANETs, Pervasive and Mobile Computing, Elsevier, Heidelberg, Germany **(2017)**.
3. J. W. Huang, I. Woungang, H.-C. Chao, M. S. Obaidat, T. Y. Chi, and S. K. Dhurandher, Multi-path trust-based secure AOMDV routing in ad hoc networks, *Proc. Global Telecommunications Conference, GLOBECOM 2011* **(2011)**, pp. 1–5.
4. S. Sarkar and R. Datta, *Int. J. Commun. Netw. Distrib. Syst. (IJCNDS)* 11, 92 **(2013)**.
5. S. Sarkar and R. Datta, QSRec: An algorithm for quick and secure recovery from link failures in mobile ad hoc networks, *Proc. International Conference on Communications, Computers and Devices, ICCCD* **(2010)**, pp. 1–6.
6. W. Yu and K. J. R. Liu, *IEEE J. Sel. Areas Commun.* 23, 2260 **(2005)**.
7. L. Anderegg and S. Eidenbenz, Ad Hoc-VCG: A truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents, *Proc. ACM MobiCom* **(2002)**, pp. 245–259.
8. H. Shen and Z. Li, A hierarchical account-adied reputation manageme ny system for MANETS, *IEEE/ACM Trans. Netw.* 23, 70 **(2015)**.
9. S. Sarkar and R. Datta, A game theoretic model for stochastic routing in self-organized MANETs, *IEEE Wireless Communications and Networking Conference* **(2013)**, pp. 1962–1967.
10. M. J. Osborne and A. Rubinstein, A Course in Game Theory, The MIT Press, Cambridge, Massachusetts **(1994)**.
11. Z. Ji, W. Yu, and K. J. R. Liu, *IEEE J. Sel. Areas Commun.* 26, 1204 **(2008)**.
12. F. Li, Y. Yang, and J. Wu, *IEEE Trans. Syst. Man Cybern.-B* 40, 612 **(2010)**.
13. B. Wang, Y. Wu, K. J. R. Liu, and T. C. Clancy, *IEEE J. Sel. Areas Commun.* 29, 877 **(2011)**.
14. L. Anderegg and S. Eidenbenz, Ad hoc-VCG: A truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents, *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*, ACM **(2003)**, pp. 245–259.
15. M. J. Osborne and A. Rubinstein, A Course in Game Theory, MIT Press, Cambridge **(1994)**, p. 4; F. Li, Y. Yang, and J. Wu, *IEEE Trans. Syst. Man Cybern.* 40, 612 **(2010)**.
16. X. Chen, H. M. Jones, and D. Jayalath, *IEEE Trans. Mob. Comput.* 10, 108 **(2011)**.
17. W. Yu and K. R. Liu, *IEEE Trans. Mob. Comput.* 6, 507 **(2007)**.
18. N. Nisan and A. Ronen, *J. Artif. Intell. Res.* 29, 19 **(2007)**.
19. J. Liu, X. Jiang, H. Nishiyama, R. Miura, N. Kato, and N. Kadowaki, *IEEE J. Sel. Areas Commun.* 30, 2169 **(2012)**.
20. J. Liu, X. Jiang, H. Nishiyama, and N. Kato, *IEEE/ACM Trans. Netw. (TON)* 20, 1950 **(2012)**.

Received: 23 April 2018. Accepted: 24 May 2018.

RESEARCH ARTICLE

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1754–1761, 2018

RESEARCH ARTICLE

Design and Implementation of Logical Sensing Algorithm Based Smart Home Security

Y. Adline Jancy[1,*], B. Gomathy[2], and A. Benuel Sathish Raj[3]

[1] *Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore 641022, India*
[2] *Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam 638401, India*
[3] *Department of Electrical and Electronics Engineering, Karunya Institute of Technology and Sciences, Coimbatore 641114, India*

The paper proposes the logical sensing algorithm to improve the security in the existing home automation systems and explain the various security issues in a smart home. The work deals with unusual behavior at natural access points to a home as primary and secondary access points depending on their use. Logic sensing algorithm is implemented by identifying normal user behavior at these access points and requesting user verification when necessary. User behavior identified when various access points changed states. The combination of sensors, microcontrollers, Raspberry Pi and ZigBee communication are used to identify user behavior at various access points and implement the logical sensing algorithm. Moreover, the algorithm also verifies the legitimacy of a fire alarm by measuring the change in temperature, humidity, and carbon monoxide levels, thus defending against manipulative attacks. In the experiment, the proposed logical sensing algorithm successfully implemented for a month in a studio apartment. During the course of the experiment, the algorithm understands all the state changes of the primary and secondary access points and user identity was verified 55 times by generating 14 warnings and 5 alarms.

Keywords: Home Automation, Smart Homes, Wireless Sensor Networks, Access Control, ZigBee.

1. INTRODUCTION

Researchers have experimented and improved the concept of smart home since the late 1970s. As technology advanced with time, the concept of home automation and people's expectation from a smart home has changed dramatically. Modern smart homes are a sophisticated combination of various Ubiquitous Computing Devices and Wireless Sensor/Adhoc Networks.[1] All these new user expectations, complicates electronics, and unpredictable user behavior have brought new security challenges to the home automation front. In the past, an average home had to deal with common slash and grab criminals, while a modern home has to deal with sophisticated and tech-savvy attackers who know how to find vulnerabilities and manipulate the security devices to gain access or cause distress to the inhabitants.[2] Despite smart home security being critical, there are some vulnerabilities in the existing systems.[3,4] Over the years researchers demonstrated various security issues associated with the devices and technology used in modern smart homes. The wireless sensor networks deployed in modern smart homes for the device to device communication is vulnerable to various Routing[5] and Wormhole attacks.[6] Popular communication technologies like ZigBee and 802.15.4 used in smart homes are susceptible to Replay attacks.[7] All these factors contributed to the rapid rise in home burglaries over the past decade[8,9] and demonstrate the importance of Home Security in the modern world. The previous works in smart home security[10,11] explains the changing role of modern home security systems and defines the role of a modern home automation system as, one capable of identifying, alerting and preventing intrusion attempts in a home at the same time preserving evidence of the intrusion or attempted intrusion so that the perpetuators or perpetuators can be identified and prosecuted. An ideal way to improve home security and defend against intrusion is to recognize a home's authorized inhabitants and identify their position inside a home at all times without inconveniencing its inhabitants. This is extremely challenging and complex, given the unpredictable nature of human behavior and home being occupied by guests and other trusted people. Identifying access points to a home and regulating access to them is the next logical step towards securing a home. The paper proposes that, normal user behaviour at access points to a home adhere to a set of predictable

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1754/008 doi:10.1166/jctn.2018.7374

behaviors. These user behaviors when analyzed by our novel logical sensing algorithms can differentiate between normal and attack behaviors.

1.1. Objectives of the Work

- Distinguish between primary and secondary access points in a home based on how they are used. Detect all user actions at these access points.
- Understand user behaviour after the change in state of an access point.
- Identify and isolate attack behavior by analyzing the user behavior at various access points using our logical sensing algorithm. Trigger warning or raise alarms depending on the situation.

The proposed work uses Raspberry Pi, microcontrollers and sensors to detect and distinguish between normal and attack user behaviors at various access points. Section 2 explains the methodology used to identify intrusions at primary and secondary access points. Section 3 describes the experiment and hardware setup. Section 3 states the result of the experiment and discusses the experiment results along with the advantage of the work. The paper concludes by stating future directions the research could take.

2. METHOD

In this paper, we analyzed various access points in a home to identify different improbable scenarios within a smart home during its operation. In a typical home, these natural access points are the front door, back door, balcony doors, and windows. Even though the window is not a normal access point it can be used as one; most likely by an intruder depending on the situation. The paper proposes that, irrespective of the number and type of access points in a home, the behavior of a legitimate user at these access points can be broken down into a set of possible events which can be predicted. Based on the purpose of the access points, the paper classifies access points into primary and secondary. In a home, when an access point is used by its inhabitants as a primary means to enter and exit from their home, it is categorized as the primary access point like the front door, back door etc. On the other hand, secondary access points like the window, balcony door etc. also provide entry/exit to a home but they are rarely used for that purpose because there are other convenient ways in and out of a home for a legitimate user.

2.1. Primary Access Point

The front door is the primary access point to any home. This paper proposes the use of motion and proximity sensors to detect user behavior at primary access points. When a user leaves an occupied home, the motion and proximity sensors placed near the access point inside the home are triggered before the door is opened. Once the user stepped out and closes the door the motion and proximity sensors will not be triggered. When someone enters an empty home, they are entering from outside so, the motion and proximity sensors will not be triggered before the door is opened. Once the door is opened and the user enters the home the motion and proximity sensors placed inside the home will be triggered. Figure 1 shows the shows the flowchart of the door state changes and sensor operations of the primary access point. Table I, shows all the possible initial states, intermediate states, final states and motion and proximity sensor triggers before and after the state changes of the main door when the home is occupied and empty. Table II represents the special cases the algorithm could take during its operation. States in Table II are only triggered from a few particular previous states which are explained in this section (Section 2.1). The sensor placements should also ensure that, anyone using the door to enter and exit the home cannot do so without triggering the motion and proximity sensors. The algorithm works by analyzing multiple proximity and motion sensor values before and after the door is opened or closed. Once the door state is changed the algorithm considers the number of proximity and motion sensor values before the door state is changed to identify if the door was opened from the inside or outside.

After the initial state change, the algorithm keeps observing the door for a specific interval of time called "door observation time." The door state during this time is called intermediate state of the door. The algorithm observes the motion and proximity sensor values during the door observation time to identify user actions at an access point. States 1 to 16 from Table I are triggered only when the home is occupied and states 17 to 32 occurs only when the home is empty.

When the home is occupied and the user opens a closed door from the inside, the proximity and motion sensors are triggered on the way to open the door. After opening the door, the user can either:

(a) Leave the door open and come back into the house by triggering the motion and proximity sensors after opening the door (state 1 in Table I).

(b) Leave the door open and step outside the home without triggering the motion and proximity sensors after opening the door (state 13 in Table I). This leaves the home vulnerable to intruders, so after a fixed amount of state change time, the home state is changed to empty, so a user will have to verify his identity upon re-entry into the home. Before changing the state of the home the algorithm issues a warning, informing the user about the impending state change.

(c) Step out and close the door behind him within door observation time allowed by the algorithm without triggering the motion and proximity sensors after the door is closed (state 4 in Table I). When state 4 occurs in a single person occupied home, the state of the home changes empty after the door observation timer has expired.

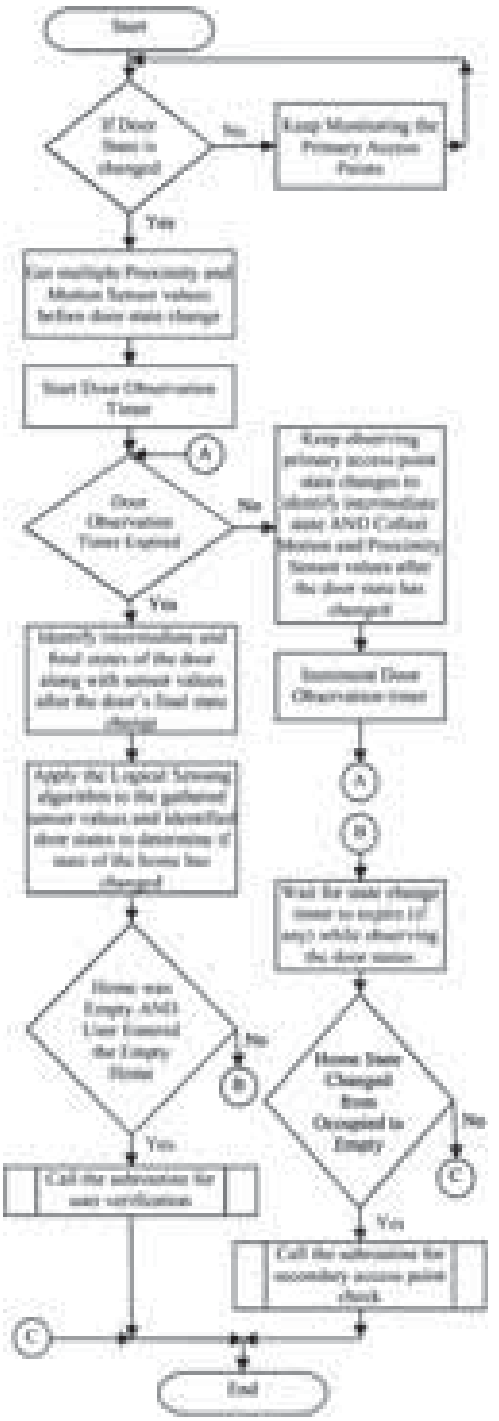


Fig. 1. Flowchart showing door state changes and sensor operations of the primary access point.

(d) Close the door from the inside within the door observation time allowed by the algorithm and comes back in, triggering the motion and proximity sensors after the door is closed (state 6 in Table I). When the home is occupied and the door is open, it can be closed from the inside or from the outside. These states are discussed below.

(e) User closes the door from the inside and comes back into the home. The motion and proximity sensors are triggered before and after the door is closed (state 7 in Table I).

(f) Sensors are not triggered after the final state, as the user executes the final state from outside the home. Since the user has stepped outside the home leaving the door open, the state change timer is started, which upon expiry changes the home state to empty (state 11 in Table I).

(g) The user closes an open door from the inside and opens it within the door observation time and comes back into the home. Motion and proximity sensors are triggered

Table I. Possible states a main door could take.

State no	Initial state intermediate state	Final state	Motion sensor trigger		Proximity sensor trigger		Home empty
			Before	After	Before	After	
1	C→O	O	✓	✓	✓	✓	F
2	C→O	O	×	✓	×	✓	F
3	C→O	O	×	×	×	×	F
4	C→O	C	✓	×	✓	×	F
5	C→O	C	×	×	×	×	F
6	C→O	C	✓	✓	✓	✓	F
7	O→C	C	✓	✓	✓	✓	F
8	O→C	C	✓	×	✓	×	F
9	O→C	C	×	×	×	×	F
10	O→C	O	×	×	×	×	F
11	O→C	O	✓	×	✓	×	F
12	O→C	O	✓	✓	✓	✓	F
13	C→O	O	✓	×	✓	×	F
14	O→C	C	×	✓	×	✓	F
15	C→O	C	×	✓	×	✓	F
16	O→C	O	×	✓	×	✓	F
17	C→O	C	×	✓	×	✓	T
18	C→O	O	✓	✓	✓	✓	T
19	C→O	O	×	✓	×	✓	T
20	C→O	O	×	×	×	×	T
21	C→O	C	✓	×	✓	×	T
22	C→O	C	×	×	×	×	T
23	C→O	C	✓	✓	✓	✓	T
24	O→C	C	✓	✓	✓	✓	T
25	O→C	C	✓	×	✓	×	T
26	O→C	C	×	×	×	×	T
27	O→C	O	×	×	×	×	T
28	O→C	O	✓	×	✓	×	T
29	O→C	O	✓	✓	✓	✓	T
30	C→O	O	✓	×	✓	×	T
31	O→C	C	×	✓	×	✓	T
32	O→C	O	×	✓	×	✓	T

Table II. Special cases.

State no	Initial state intermediate state	Final state	Motion sensor trigger	Proximity sensor trigger	Home empty	Algorithm action
33	O→C	O	✓	✓	F	Reset/start state change timer
34	C→O	O	✓	✓	T	Activate IVM

before the door is closed (initial state) and after the door is open (final state), as the user initially came from inside to close the door and went back into the home after leaving the door open (state 12 in Table I).

(h) The proposed algorithm keeps monitoring the door and sensor values for different state changes. Some states like state 2, 3, 5 and 15 in Table I are only triggered when their previous states are either state 5 or state 9. Figure 2 demonstrates the states the proposed system goes through before states 2, 3, 5 and 15 are triggered. Figure 2(a) shows state 2 transition; state 2 is triggered when the home is occupied and a closed door is opened without triggering the motion and proximity sensors before opening the door. In a single person occupied home, such an event will only take place, when the user steps out of a home leaving the door open (state 11 and state 13 in Table I) and within the state change timer expiry period states 5 or 9 are triggered. When state 2 is triggered it means the user re-entered the home leaving the door open, so the state change timer can be reset. Figure 2(a) also shows the transition of states 9 and 10.

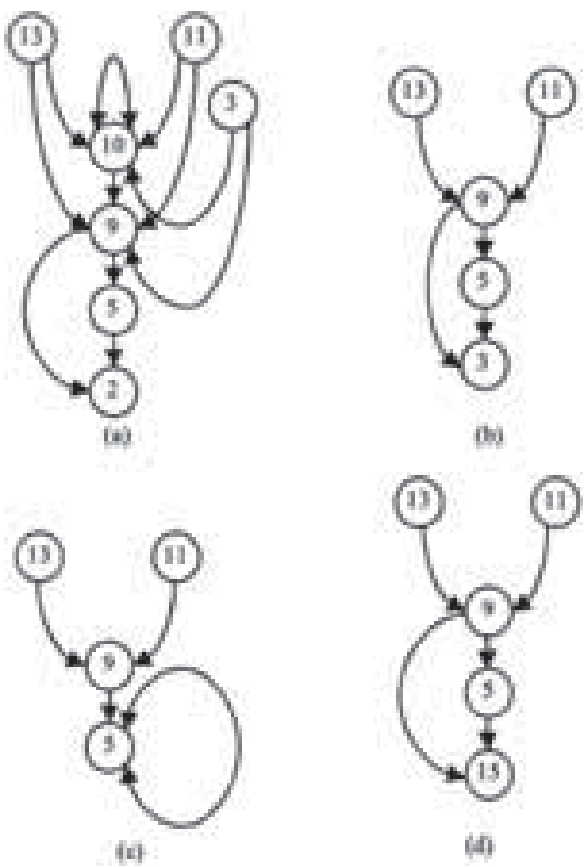


Fig. 2. (a) Shows state 2 transition along with state 9 and 10 transition; (b) state 3 transition; (c) state 5 transition (d) state 15 transition during the operation of the home.

Figure 2(b) shows state 3 transition; state 3 happens when the home is occupied and a closed door is opened without triggering the motion and proximity sensors before and after opening the door. In a single person occupied home, like state 2, state 3 will only be triggered if states 5 or 9 are triggered before the state change timer has expired. Figure 2(c) shows state 5 transition; state 5 transpires when the home is occupied and a closed door is opened and closed without triggering the motion and proximity sensors before or after closing the door. In a single person occupied home, similar to states 2 and 3, state 5 is only triggered if states 5 or 9 are triggered before the state change timer has expired. Figure 2(d) shows the transition of the system to state 15; state 15 is triggered when the home is occupied and a closed door is opened and closed. The motion and proximity sensors are not triggered before opening the door but the sensors are triggered after closing the door which indicates the user stepped back into the home after closing the door. So after triggering state 15, in a single person occupied home the state change timer is reset. A single person occupied home becomes empty when states 11, 13, 4 and 8 are triggered. When states 11 and 13 are triggered the algorithm utilizes a state change timer as mentioned above in (b) and (g). When states 4 and 8 are triggered the home state is changed when door observation timer expires as mentioned above in (c) and (e). Door observation time is a relatively short period of time compared to the state change time. When a particular state is triggered the previous state of the door is also considered to accurately determine the possible next states and the occupied status of the home. Table III shows the possible previous states for a particular state and timer actions in the proposed system when the home is occupied and empty. State 10 is triggered when an open door is first closed then opened from outside without triggering the motion and proximity sensors before or after the initial and final states. State 10 is triggered when the previous states are states 11, 13, 10 or 3 and with state change timer still running. Similarly, states 9, 14 and 16 are only triggered when their previous states are either are 3, 10, 11 or 13. In a single person occupied home, triggering states 14 and 16 means the user has re-entered the home so the state change timer is reset.

Whenever states 3, 10, 11 and 13 are triggered (the door remains open and the state change timer is running) the algorithm keeps monitoring the front door and the sensors until another door state is triggered or until the state change timer is expired. When the user re-enters the home by triggering the motion and proximity sensors before the state change timer has expired, state 33 from Table II is triggered. As the user re-entered the home, the algorithm stops and resets the state change timer. When a home changes its state from empty to occupied, the identity of the person causing the state change has to be verified. The techniques used for verifying user identity[12] can vary from

Table III. Possible next state for a particular state.

State no	Possible previous states	IVM trigger	Timer status
0	17, 19, 20, 27, 31, 32	No	None
1	0, 5, 6, 7, 9, 14, 15	No	None
2	5, 9	No	Reset state change timer
3	5, 9	No	Continue state change timer
4	0, 5, 6, 7, 9, 14, 15	No	None
5	5, 9	No	Continue state change timer
6	0, 5, 6, 7, 9, 14, 15	No	None
7	0, 1, 2, 10, 12, 16	No	None
8	0, 1, 2, 10, 12, 16	No	None
9	3, 10, 11, 13	No	Continue state change timer
10	3, 10, 11, 13	No	Continue state change timer
11	0, 1, 2, 10, 12, 16	No	Start state change timer
12	0, 1, 2, 10, 12, 16	No	None
13	0, 5, 6, 7, 9, 14, 15	No	Start state change timer
14	3, 10, 11, 13	No	Reset state change timer
15	5, 9	No	Reset state change timer
16	3, 10, 11, 13	No	Reset state change timer
17	4, 8, 17, 26, 22	Yes	Start identity verification timer
18	Irrelevant	Alarm	Alarm triggered, so no timer
19	4, 8, 22, 26	Yes	Start identity verification timer
20	4, 8, 22, 26	No	None
21	Irrelevant	Alarm	Alarm triggered, so no timer
22	4, 8, 22, 26	No	None
23	Irrelevant	Alarm	Alarm triggered, so no timer
24	Irrelevant	Alarm	Alarm triggered, so no timer
25	Irrelevant	Alarm	Alarm triggered, so no timer
26	11, 13, 20, 27	No	None
27	11, 13, 20, 27	No	None
28	Irrelevant	Alarm	Alarm triggered, so no timer
29	Irrelevant	Alarm	Alarm triggered, so no timer
30	Irrelevant	Alarm	Alarm triggered, so no timer
31	11, 13, 27	Yes	Start identity verification timer
32	11, 13, 27	Yes	Start identity verification timer
33	3, 10, 11, 13, 33	No	Reset/Start state change timer
34	20, 27	Yes	Start identity verification timer

simple 4 digit pin, facial recognition, retinal scan verification, fingerprint verification to sophisticated biometric gait recognition, vein recognition, and voice recognition. The user identification should be done within a fixed time period called the verification time period. The verification timer stars once the door to an empty home is opened and someone enters, it stops when a valid user identity is confirmed. When the timer exceeds the verification time period an intruder alert is triggered. Depending upon the level of security and user preference the intruder alert can be an audible alarm to scare off intruders and alert the neighbors, silent alarm to alert the authorities or any other defensive measures.

Figure 3 shows the identity verification process in the algorithm. When the home is empty and the closed, main door is opened from outside without triggering the motion and proximity sensors. After opening the door the user can either:

(i) When the home is empty and the user comes in after opening the door he will have to confirm his identity within the verification time period.

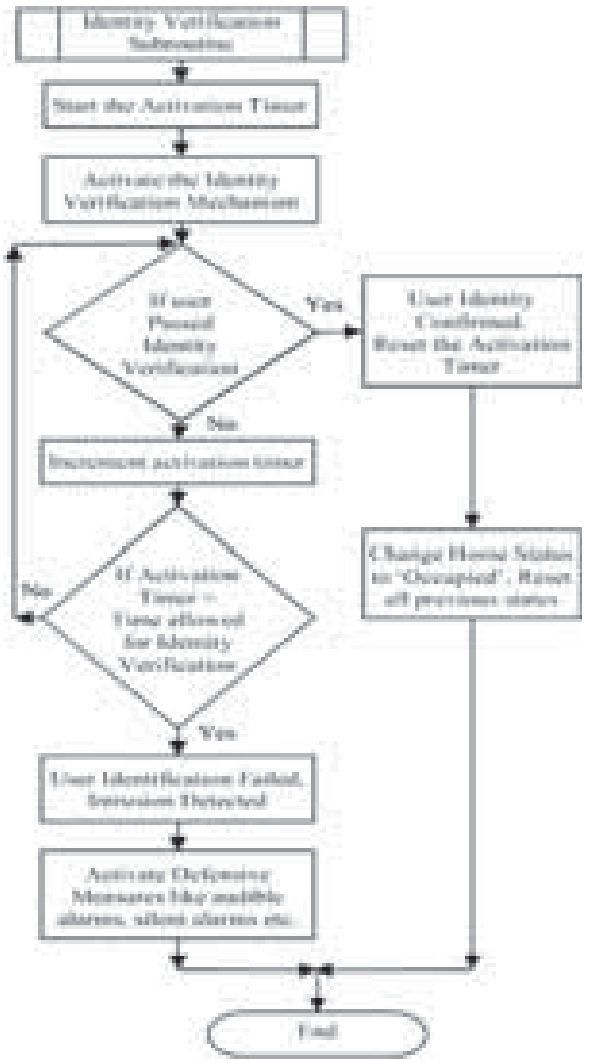


Fig. 3. Flowchart showing identity verification process.

(j) Leave the door open and enter the home triggering the motion and proximity sensors while moving into the home (state 19 in Table I). The verification timer is triggered as soon as the door is opened so the user has to verify his identity using the IVM.

(k) Leave the door open and decides to stand at the door or go out of the house, without triggering the motion and proximity sensors (state 20 in Table I). After triggering state 20 the door remains open, so the algorithm keeps monitoring the motion and proximity sensors until another door state is triggered; whenever motion and proximity sensors are triggered immediately after state 20, it means someone entered the home through the open door triggering state 34 from Table II, so IVM is activated and user identity is verified.

(l) Closes the door within the sensing time allowed by the algorithm and steps out of the home without triggering

the motion and proximity sensors after the door is closed (state 22 in Table I). So sensors are not triggered before the initial state or after the final state of the door and no one entered the home, so user identity is not verified, home remains empty and the door is closed. Similarly Table III describes the different state of IVM activation. The IVM is placed inside the home so whenever the user identity is successfully verified the home becomes occupied and all the previous states of the door are cleared and the current state is set to state 0.

The algorithm can be implemented at each of the primary access points independently. The algorithm observes motion and proximity sensor readings at each of the open primary access points because the user can enter the home through any of the open primary access points. If the proximity and motion sensors are triggered at any of the open access points state change timer is reset. The time difference between the motion and proximity sensor triggers from each of the access points used by the user to exit and enter the home is considered along with the distance between access points to distinguish between normal user behavior and sneaky attack behavior. The algorithm takes into account the minimum time required by the user to move between the access points to identify attack behaviors.

2.2. Secondary Access Points

The balcony door and windows form the secondary access points in a home. Usually balcony door opens into a relatively secure and private area, sometimes even a few floors up. So, these balcony doors can remain open for long periods of time when the house is occupied. When the home becomes empty an observant, resourceful and proficient intruder can use this door to gain access to the home, in order to avoid that, balcony doors must be closed when the home becomes empty. The algorithm keeps monitoring the state of the balcony door, by placing motion and proximity sensors near the window inside the home, so in an empty home when the balcony door is opened the system triggers intrusion defense mechanisms without waiting for any identity verifications. So when the home is empty and the window is opened the system triggers the intrusion defense mechanisms without waiting for identity confirmation. In addition, when the home is occupied and the window is opened from the outside, the system triggers a warning and asks the user to confirm his identity because under normal circumstances windows are rarely opened from the outside. Whenever a home changes state from occupied to empty the algorithm checks if the secondary access points to the home are secure. If not it issues a warning to the user to secure the secondary access points. Figure 4 shows the flowchart of the secondary access point checking when the home becomes empty.

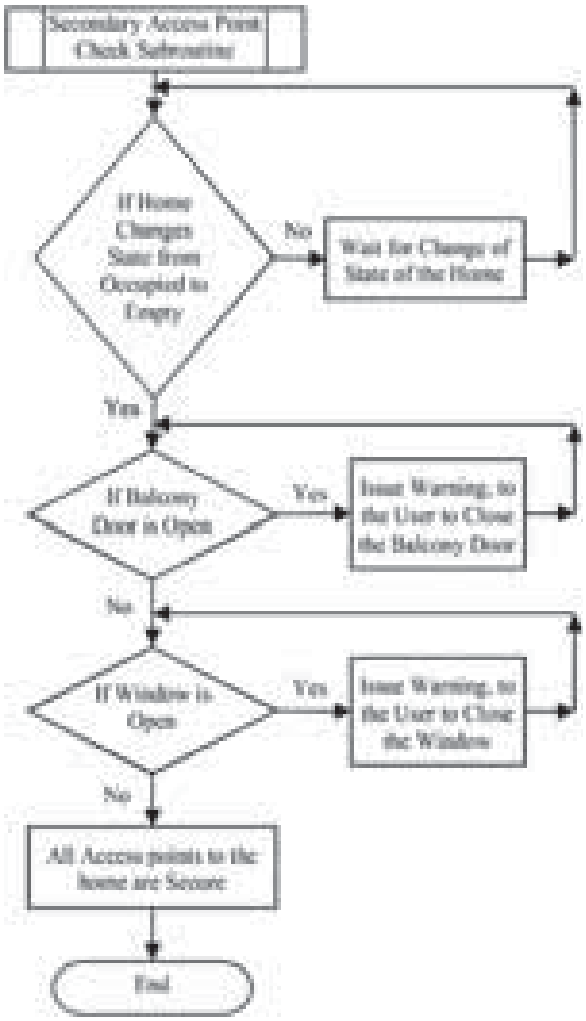


Fig. 4. Flowchart demonstrating secondary access point check when home becomes empty.

2.3. Fire Alarm

The work of Fouladi[2] discussed the weakness in the existing smart home architecture and demonstrated how an attacker will compromise various networked elements in a home. The easiest way to get the inhabitants out of a home is to trigger an emergency alarm like the fire alarm. When a fire alarm is triggered all the automatic locks of a home are disabled. During the home fire the carbon monoxide and the ambient temperature levels in the area of the fire will go up and inversely the humidity in and around the area will go down. Warnings are triggered as a means to get user attention when something out of the ordinary happens. In the proposed system warnings are generated when: (a) user leaves the door open to a home and steps outside, also the state change timer is about to expire. (b) The home becomes empty and at least one of the secondary access points is unsecured; the algorithm

warns the user about the open secondary access point. Upon receiving the warning the user comes back into the home and secures the secondary access point. The warning will not stop until all the secondary access points are secured. (c) One or more of the secondary access point windows are opened from the outside when the home is occupied. Identity verification timer is activated and the user is asked to confirm his identity because under normal operating conditions windows are rarely opened from the outside. (d) The fire alarm is triggered without sufficient change in humidity, ambient temperature or carbon monoxide levels. The warning informs the user about a possible fire alarm manipulation which user can verify on sight.

During the one month period the main access point changed state 305 times. The algorithm was able to detect all these and reduce them to 190 state changes by identifying and eliminating the intermediate state changes mentioned in Table I. The most common state triggered was state 4 in Table I, it was triggered 46 times. While state 17 was triggered 33 times, making it the second most popular. State 13 was triggered 20 times making it the third most triggered state. State 6 and state 1 were triggered 12 and 11 times respectively. State 31 happened 6 times, State 19 was triggered 13 times, States 9 and 16 were triggered five times; states 7, 14, and 15 were triggered four times; states 5, 11, 20, 22 and 26 were triggered thrice; states 2, 8 and 12 were triggered twice; states 3, 10, 18, 21, 27 and 32 were triggered only once during the one month time period. States 23, 24, 25, 28, 29 and 30 were not triggered during the one month time period.

3. RESULTS

The algorithm generated 14 warnings, 8 regarding the open primary access point, 3 for not securing the secondary access point before leaving the home and another 3 warnings for opening the balcony door during day when the home was occupied. Intruder alarm was triggered 5 times during the experiment, 4 were related to primary access points while one was related to secondary access points. IVM was triggered 59 times and user successfully verified his identity 55 times. Alarm was killed using the 12 character password five times. The state change timer was activated 23 times while the user re-entered the home before the state change timer expired 15 times, so the home state changed due to state change timer expiry 8 times. The graph in Figure 5 shows the number of state changes for primary access point, total number of warnings generated, IVM triggers, number of user identity verifications and number of alarms generated.

Secondary access points changed states 56 times; balcony door changed state 27 times and window state was changed 29 times. All of the 27 times balcony door changed state the user was at home but 3 of them happened when the user was in home, so these 3 times warnings

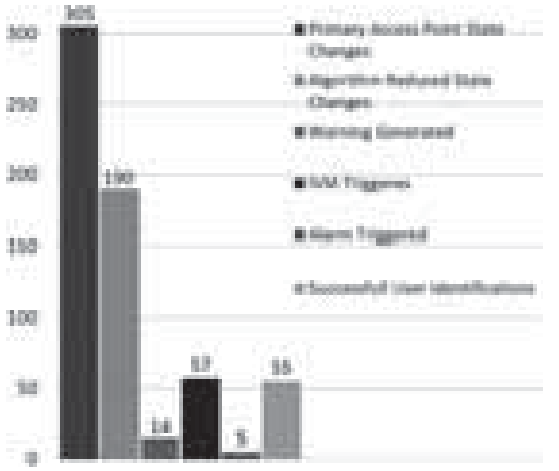


Fig. 5. Graph showing the number of state changes for primary access point, algorithm reduced state changes, total number of warnings generated, IVM triggers, number of user identity verifications and number of alarms generated.

were generated and IVM was activated along with the identity verification timer, which was reset when the user conformed their identity. Once the open balcony door was closed due to wind when the user was in home, so no identity verification mechanism was initiated. The home was occupied all 29 times when the state of the balcony window was changed. Once the window was opened from the outside when the user was in bed during night; the intrusion defense mechanisms (audible alarm) were triggered without waiting for any user identity verifications.

4. CONCLUSION

The paper detects user actions at primary and secondary access points in a home using different sensors. These detected user actions and behaviors are compared with normal user behavior at various access points to identify intrusions or intrusion attempts. In the experiment, our proposed algorithm was able to successfully identify all 305 state changes of the main access point and reduce them to 190 user behaviors while the secondary access point changed state 56 times. The alarm was triggered five times when the user failed to confirm his identity. Six of the fourteen warnings generated were regarding secondary access points while the other eight were relating to primary access point when the home became empty. In addition to identifying intrusions in home, the algorithm also warns user about imminent and live potential security vulnerabilities by identifying the status of various access points, user position and behaviors.

For future works, we plan to improve user behavior prediction by analyzing various user actions inside the home to further improve smart home security.

References

1. C. Suh and Y.-B. Ko, *IEEE Transactions on Consumer Electronics* 54, 1177 **(2008)**.
2. B. Fouladi and S. Ghanoun, Security Evaluation of the Z-Wave Wireless Protocol, Black Hat USA, August **(2013)**.
3. W. Wang and Z. Lu, *Computer Networks* 57, 1344 **(2013)**.
4. N. Komninos, E. Philippou, and A. Pitsillides, *IEEE Communications Surveys and Tutorials* 16, 1933 **(2014)**, Fourth quarter.
5. C. Karlof and D. Wagner, *Ad Hoc Networks* 1, 293 **(2003)**.
6. Y. Hu, A. Perrig, and D. Johnson, *IEEE Journal on Selected Areas in Communications* 24, 370 **(2006)**.
7. Y. Mo and B. Sinopoli, Secure control against replay attacks, *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL **(2009)**, pp. 911–918.
8. D. Deadman, *International Journal of Forecasting* 19, 567 **(2003)**.
9. A. C. Jose, R. Malekian, and N. Ye, *IEEE Access* 4, 5776 **(2016)**.
10. A. C. Jose and R. Malekian, *Smart Computing Review* 5, 269 **(2015)**.
11. M. O. Oloyede and G. P. Hancke, *IEEE Access* 4, 7532 **(2016)**.
12. M. Ryan, Bluetooth: With low energy comes low security, *WOOT'13 Proceedings of the 7th USENIX Conference on Offensive Technologies* **(2013)**, pp. 4–4.
13. V. Namboodiri, V. Aravinthan, S. N. Mohapatra, B. Karimi, and W. Jewell, *IEEE Systems Journal* 8, 509 **(2014)**.
14. B. Al Baalbaki, J. Pacheco, C. Tunc, S. Hariri, and Y. Al-Nashif, Anomaly behavior analysis system for ZigBee in smart buildings, *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, Marrakech **(2015)**, pp. 1–4.
15. R. Malekian and Abdul Hanan Abdullah, *International Journal of Innovative Computing, Information and Control* 8, 6445 **(2012)**.

Received: 23 April 2018. Accepted: 6 June 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1762–1768, 2018

RESEARCH ARTICLE

Multi-Context Based Optimal Resource Provisioning in Mobile Cloud Environments

S. Durga[1,*], S. Mohan[2], and J. Dinesh Peter[1]

[1] *Karunya Institute of Technology and Sciences, 641114, India*
[2] *CCIS, Al Yamamah University, Riyadh, 11512, KSA*

Mobile Cloud computing (MCC) has become a promising paradigm for extending computing services to the mobile clients. However dynamic and cost-efficient resource provisioning remains challenging due to its limited battery, intermittent connections, and mobility. This paper presents an optimal resource provisioning and management model that includes a two-stage process. In the first stage, the cost, waiting time of the mobile clients and the utilization of the cloud server are jointly optimized using cuckoo search-based optimization algorithm. In the second stage, context-aware cloud resource management algorithm is proposed to improve the efficiency of the system. The results of simulation experiments show that the proposed framework can provide better performance as compared to the existing state of art approaches.

Keywords: Mobile Clients, Cuckoo Search, Context-Aware, Optimal Resource Provisioning, VM Migration.

1. INTRODUCTION

The recent evolution of internet technologies, mobile computing, and popularity of smart mobile devices brings a new paradigm shift from conventional mobile computing and services to the universal interface to online services and applications in the real world. Furthermore, BYOD (Bring Your Own Device) extending the business workflows into mobile or personal devices. As a result, developers worldwide are building increasingly complex and scalable mobile applications that require ever-increasing amounts of processing capability and energy. So, these applications are developed with cloud backend support in terms of storage capacity and/or heavier processing abilities. Recently, Mobile Cloud Computing (MCC) Environments have witnessed that, the inherent natures of mobile devices lead to various fluctuations in response time and performance results in perspective of cloud service. Therefore the mobile client requests are taken into consideration using an efficient resource provisioning algorithm, as they may lose their connection at any time to the cloud service provider. The dynamic allocation of suitable resources based on the client side as well as CSP side context parameters is an important aspect that needs to be considered in mobile cloud computing. To address this issues, we propose an optimal resource provisioning and management framework.

The main contributions of this paper are as follows.
(1) This paper makes use of cloud infrastructure which provides suitable resources to process mobile client requests. An optimal resource provisioning framework is proposed which uses mobile device battery status and connection type as part of mobile client contexts and energy expenditure and resource availability as part of CSP contexts to make resource allocation decision.
(2) Resource allocation problem is formulated as an optimization problem and solved using cuckoo search optimization algorithm.

The rest of the paper is organized as follows. In Section 2, the review of the related work is presented. Section 3 describes the optimal resource provisioning framework. In Section 4 and 5, we present the optimal resource provisioning model. Section 6 shows the performance evaluation of the proposed system. Finally, Section 6 concluded the paper.

2. RELATED WORK

Past research has investigated a variety of dimensions in the field of cloud resource provisioning. The inspiration for this research originates from how the mobile clients can productively make use of a cloud server to optimize the resource utilization and mobile user's QoS. Verma and Kaushal[6] propose a deadline constraint heuristic based genetic algorithm for workflow scheduling in the cloud.

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1762/007 doi:10.1166/jctn.2018.7375

Their algorithm achieves reduced cost while completing the task execution before its deadline. The major issue with their heuristic algorithm is that the mono objective property of the problem.

Jin et al.[7] propose a truthful incentive mechanism, to coordinate the resource auction between mobile devices and cloudlets. However, the price is being fixed between the client and the server and will not be changed based on the demand and supply these interns affect the MCC system efficiency. The joint optimization of resource provisioning model proposed by Chase and Niyato[8] has a sensitivity analysis of prices at both initial and provisioning stages. However, the internal cost information in cloud networking may not be easy to obtain due to random network delay and VM migrations. Mireslami et al.,[9] consider simultaneous optimization of deployment cost and QoS as a problem and solved using the branch and bound technique to obtain realistic discrete solutions. Their algorithm yields an optimal combination of resources by considering the customer's requirements but can handle only the workloads from the web.

Gai et al.[10] present an energy aware heterogeneous task assignment algorithm for mobile cloud. Their model was designed to solve the energy minimization problem in heterogeneous computing environments. The optimized resource allocation and task scheduling algorithm for vehicular cloud has been proposed by Midya et al.[11] They proposed a three-tier architecture consisting of vehicular cloud, local cloudlet, and the centralized cloud. Their algorithm was designed using Hybrid Particle Swarm optimization to efficiently handle multiple requests from on road users while maintaining QoS. However, the performance of the architecture was affected due to the slow convergence time of the hybrid algorithm.

Hassan et al.[12] studied the resource provisioning problem for mobile media traffic in a cloud environment. The aim of this paper is to maintain high system utilization by avoiding over-provisioning of resources to services. They solved the resource provisioning problem using mixed integer linear programming model and the heuristics. Liu et al.[13] propose a multi-resource allocation strategy as a semi Markov decision process under average cost and system throughput.

Li et al. (2017) propose multiple context-based service scheduling for balancing cost and benefits of mobile users and cloud server. They modeled the resource provisioning problem as the network utility maximization problem and solved using Lagrangian multipliers method. Under this model, authors claim the improvement in mobile user's QoS experiences by considering the multiple system parameters.

The paper contribution is different from other related works. All the new coming requests are analyzed and based on the current context situations, optimal allocation of the suitable resource to the services or applications is proposed. Secondly, if there is a context change, based on the current state of the request, dynamic provisioning of resources is introduced.

3. SYSTEM MODELLING AND PROBLEM FORMULATION

In this section, multi-context based optimal resource provisioning and management system model is proposed. First, we define the proposed system model. Secondly, we formulate a resource allocation problem to balance the benefits of mobile clients and the cloud server.

3.1. Optimal Resource Provisioning Model (ORPM)

We consider that each request has its client context information along with its Quality of Service (QoS) requirements. The client device context information such as Battery Power Status (B_s), Connection Type (C_T)[4] and timestamp (t_s) and the client request's QoS requirements such as VM_{type} in terms of its CPU, RAM and storage capabilities and No. of VMs are collected through an Application Programmable Interface (API). All the application requests are characterized earlier by the application developer and are sent along with original computing request through the mobile API. We assume that the service request sent by the mobile clients are workloads to the Cloud. Each workload has its CPU, RAM and storage requirements. Figure 3 shows the proposed optimal resource provisioning and management framework. The framework consists of four main entities: Request manager unit (RMU), Context manager unit (CMU), Resource Provisioner unit (RPU) and resource manager unit (RM). Each mobile client request furnishes the current context information such as battery status, Connection type, timestamp, in addition to the computing requirements. The functional entities along with their Interactions are listed below.

— *Request manager unit* (*RMU*): The client request is first intercepted by the RMU where the request specifications are analyzed. The context information is segregated and analyzed by the RMU. The current mobile device battery status could be estimated in the cloud because mobile device consumes energy while transmitting thus the energy will decrease after the context is transferred to the cloud.

The residual energy is calculated as follows

$$\text{Residual Energy} = E_r = B_s - E_t \tag{1}$$

Where, E_t is the energy consumed by the mobile device while transmission and is calculated based on the transmission bandwidth b_{MD} and the size of data d_{MD} transmitted.

$$E_t = (d_{\text{MD}}/b_{\text{MD}}) * E_p \tag{2}$$

Where, E_p is the percentage energy consumed by mobile device while transmission. Based on E_r, RMU is authorized to admit or reject the request. As shown in

RESEARCH ARTICLE

Algorithm 1, the request will be rejected when the residual energy is less than the predefined threshold. Also, if the signal strength is low or if the MD's battery status is low then the data cannot be sent. Being idle connection the connection will be automatically closed.

— *Context Manager Unit* (*CMU*): Receives context information from RMU and Stores, monitors context data until the request is fulfilled or the request is timed out. When there is serious context change in the client context, CMU notifies the context change to the RPU.

— *Resource Provisioner Unit* (*RPU*): RPU aims to improve the QoS experience of mobile clients as well as the benefits of cloud service provider. RPU considers device energy, budget and deadline constraints of mobile clients and also the benefits in terms of resource utilization and profits of cloud service providers to solve the optimization problem. Finally optimal allocation of resources to the requests are made by the RPU.

When there is a context change notification for the admitted requests, the optimal resource Provisioner immediately sends the allocation update to the resource manager, consequently, the corresponding VM is migrated according to the new context information. The problem formulation and proposed solution is described in Section 4.

— *Physical Server Unit* (*PSU*): PSU is comprised of a bunch of computing servers which are used to process customer's requests using their CPU, RAM, Storage and network resources.

ALGORITHM 1 (REQUEST MANAGER ALGORITHM).

```
Input: C_T, B_s, t_s, D_l, VM_type (CPU, RAM,
Storage), No. of VMs)
Output: Admission status Ad_s
For Each SR_i
    Extract C_Ti, B_si, t_si of SR_i
    Compute E_r using Eq. (1)
    If E_r < Th then
        Ad_s ⟵ Admit
        Context DB[i] ← CT_i, Bs_i, ts_i
        sub ORPM ( )
    else
        Ad_s ⟵ reject
        Return
    End if
End of Each
```

— *Resource Manager* (*RM*): Resource manager entity is responsible for managing, creating, destroying and migrating the VMs on the physical servers. VM migration can be triggered by two types of events: server overloaded event and server under loaded event.

PROPOSITION 1. *Infrastructure based Mobile Cloud can be modeled as the composition of M/M/1/∞ and M/M/c/n: FIFO queues.*

Justification—Cloud has enormous resources which can satisfy the legislative requirements of numerous real-time applications. This made the leap to envision the ubiquitous/urban computing environments. Because of its scalability nature, at one time, enormous requests are hitting the cloud server. Hence the arrival rate of requests and the service rate of cloud providers are exponentially distributed and follow markovian property of exponential distribution. The Markovian property deals with memory less property. i.e., the current state of the system is independent of its previous state. For example, service time is exponentially distributed. This property states that the probability that a customer currently in service has t units of remaining service is independent of how long it has already been in service. Due to the uncertainty conditions of the cloud for instance system load, large no. of customers and requests, it is difficult to predict the arrival and service times from the previous times.[17] *Hence, M/M/1/∞: FIFO is used to model the resource allocator and M/M/c/n: FIFO is adopted to model the cloud servers.*

Based on Proposition 1, the RPU is designed using M/M/1 queue with an arrival and service rate modeled as an exponential random variables λ and μ, respectively. where $\lambda < \mu$.

The average service time of the queue is found using

$$T_{\text{RPU}} = \frac{1/\mu}{1-\lambda/\mu}$$

Since there is no request loss in the previous system, the arrival of requests at also follows Poisson process and the service time is assumed to be exponentially distributed with mean service time MS_i^{-1}. The response time at each computing server is $T_{\text{PSU}} = \sum_{i=1}^{n} 1/PM_i/(1-\lambda/PM_i)$. After processing the requests at *MS*, the service results are sent back to the customers. Average time a customer spends in the system (E_{est}) is defined as

$$E_{\text{est}} = T_{\text{RPU}} + T_{\text{PSU}} = \frac{1/\mu}{1-\lambda/\mu} + \sum_{i=1}^{n} \frac{1/PM_i}{1-\lambda/PM_i} \quad (3)$$

The main objectives of the proposed Optimal Resource Provisioning Model are,

1. To identify and analyze the workload to find its QoS requirements (includes its residual energy, connection type, computing resource requirements-CPU, RAM, storage).
2. To provision of resources without SLA violation.

For application task requests, the SLA is made between the application provider and the cloud service provider. In order to fulfill the SLA, CSP needs to take its best effort.

3.2. Problem Formulation

This section details the resource provisioning problem and also elaborates on the optimization equations and constraints. Mobile client contexts are often changing.

To improve the QoS, service providers supposed to take necessary actions either by increasing or decreasing the no. of VMs to complete the client requests on time. Also, based on the current context of the client, the resources can be revoked in order to reduce the cost.

Thus the resource provisioning problem is considered as an NP-complete problem. And the resource provisioning problem is that of providing suitable physical/virtual resources to the requests according to its current state. The request state aware resource provisioning aims to maximize the overall utility function of the system subject to the resource utilization constraint CSPs, energy, deadline and cost constraints of the mobile clients. The request state aware resource provisioning problem is formulated as follows,

$$U_{\mathrm{MCS}} = \sum_{j=1}^{n} \mathrm{util}_{\mathrm{cdc}}^{j} + (Dl_m^i - E_{\mathrm{est}}) + \left(E_m^{\mathrm{threshold}} - \sum_{j=0}^{n}(e_m j)\right) + \left(\mathrm{Budget}_m - \sum_{j=0}^{n}(P_{mj}^{\mathrm{cpu}} + P_{mj}^{\mathrm{ram}} + P_{\mathrm{mj}}^{\mathrm{bandwidth}})\right) \quad (4)$$

Objective Function is

$$\mathrm{Max}\ \ [U_{\mathrm{MCS}}] \quad (5)$$

$$\text{Subject to,}\ \ \mathrm{Budget}_m \geq \sum_{j=0}^{n}(P_{mj}^{\mathrm{cpu}} + P_{mj}^{\mathrm{ram}} + P_{mj}^{\mathrm{bandwidth}})$$

$$j = \text{no. of jobs} \quad (6)$$

$$DL_m \geq E_{\mathrm{est}} \quad (7)$$

$$E_m^{\mathrm{threshold}} \geq \sum_{j=0}^{n}(e_m \times j) \quad (8)$$

$$A_{\mathrm{cdc}}^{\mathrm{resource}} \geq \sum_{j=0}^{n}(R_{mj}^{\mathrm{cpu}} + R_{mj}^{\mathrm{ram}} + R_{mj}^{\mathrm{bandwidth}})$$

$$\text{available resource} > \text{required resource} \quad (9)$$

In Eq. (1), $\mathrm{util}_{\mathrm{cdc}}^{j}$ is the average resource utilization of the physical machine j in the CDC.

$(Dl_m^i - \mathrm{TAT}_{\mathrm{est}}^i)$ is the slack time available for the mobile clients request.

$(E_m^{\mathrm{threshold}} - \sum_{j=0}^{n}(e_m j))$ is the mobile device residual energy

$$\left(\mathrm{Budget}_m - \sum_{j=0}^{n}(P_{mj}^{\mathrm{cpu}} + P_{mj}^{\mathrm{ram}} + P_{mj}^{\mathrm{bandwidth}})\right)$$

is the budget amount spent by the mobile clients for CPU, RAM and bandwidth resources.

4. OPTIMAL RESOURCE ALLOCATION SCHEME

The proposed optimization problem is solved using a Meta heuristic cuckoo search Algorithm 2.[15, 16]

ALGORITHM 2 (CUCKOO BASED ALLOCATION ALGORITHM).

```
Initialize Allocation Map
While (RPU Queue!=Empty)
  For Each new Virtual machine request,
   Extract Deadline, Initial arrival
     time, Budget, QoSrequirements
   Call CSA()
   Update Allocation map
   return
  End For Each
End While
Sub CSA (Fitness Function)
  Max_iter ← 350
Generate initial populations of n host
  nests xi (i = 1,2,...,n) Choose best
nest
  While (t < Max_iter) or (stop criterion)
     Get a cuckoo randomly by Levy
       Flights
     Perform new nest
     The fitness function F(X) as
       per Eq. (11)
        Evaluate its fitness/quality
     Choose a nest among n (say j)
       randomly
     If (F_i > F_j) Replace j by the new
       solution End if
      A fraction (p) of worse nets are
        abandoned and new ones are built
     Keep the best solutions (or nests
      with quality solutions)
     Rank the solutions and find the
       current best
  End while
Save the best-achieved nest and its
cost
```

For a new solution for cuckoo i, a levy flight is performed by the following stochastic equation

$$X_i^{t+1} = X_i^t + \propto \otimes \mathrm{levy}(u) \quad (10)$$

where the step size $\propto > 0$

$$\mathrm{levy}(u) = t^{-\lambda}, \quad 1 < \lambda \leq 3$$

The first population is initialized randomly with 100 solutions. Fitness Function of each solution

$$F(X) = \sum_{i=1}^{j}\sum_{j=1}^{n} U_{ij} + C_{ij} + E_{\mathrm{est}} \quad (11)$$

Where the Server Utilization is computed using

$$\rho = \lambda_e / C\mu \quad (12)$$

5. SIMULATION RESULTS

This section presents simulation studies to evaluate the performance of the proposed framework. In the following, we first explain the simulation set up and then investigate the parameters of the algorithm performance.

5.1. Simulation Setup

We consider an MCC environment comprising of mobile clients and datacentre architecture providing IaaS as depicted in Figure 1. As the targeted system includes generic cloud infrastructure and heterogeneous mobile client requests, it is extremely difficult to conduct repeatable large-scale experiments on a real infrastructure, which is required to evaluate and compare the proposed algorithms. Also, it is not feasible to measure the performance of the proposed framework in an exact way using a specific test bed as it limits the experiments to the scale of the test bed. Therefore, simulations are used to ensure the reliability of the experiments.

Cloud Sim toolkit has been chosen as it is the modern framework for modeling and simulation of cloud infrastructure. Cloud Sim Tool Kit was extended to support the simulation of mobile client environment using java. The mobile client simulation was carried out to extract the current battery status of the client under random waypoint mobility model. For experimental results, the extracted battery context have been included with heterogeneous cloud workloads. Google cloud workload traces[20] of 12000 machines for a period of one month of May 2011 are considered for the workload resource requirements. We have considered a cloud data center having 10 to 12 nos. of physical machine servers. The experimental parameter settings of our simulation are shown in Table I.

Table I. Simulation parameters.

Simulation parameters	Value
No. of data center	1
No. of physical machines	[5, 15]
Processing elements (PEs) per machine	[Core 2 duo, core i3, i5 and i7]
PEs capacity, memory, storage	(100–3000) MIPS
Memory	(1024–12,276) MB
No. of the mobile device	[5, 50]
The mobility model	Random waypoint mobility
Battery capacity	[500, 1000] joule
Bandwidth capacity	Wi-Fi [6–54 Mbps], 3G [75 Mbps]
Minimum bandwidth threshold	Wi-Fi 1 Mbps, 3G [0.1 Mbps]
Input data size	[1–10] Mb
Power consumed on idle	Wi-Fi [0.5, 1] W, 3G [0.6, 1.2] W
Power consumed on data transmission	Wi-Fi [1, 1.8] W, 3G [1.5, 2.2] W
Deadline	[50, 1500] ms
Budget	5–1000 dollars
Request arrival rate	[50–250] requests/sec

The request of mobile clients arrives at the cloud based on a Poisson arrival process. The power capacity of the mobile device battery is varied based on the internal chemistry of the battery type such as alkaline, lithium, lead-acid etc. Hence the battery power status from the mobile phone is retrieved in percentage. The residual battery level of the mobile device is found based on the amount of data transferred and the network bandwidth.

5.2. Performance Evaluation

In the following figures, the performance of the proposed model is examined. First, the performance of Cuckoo based optimal resource provisioning algorithm with respect to the maximum number of iterations is examined in Figure 2. Next, the experiments aim to compare ORPM with Flexible context constraint-based service provisioning model (FCCSP) proposed by Chunlin et al.,[14] SMDP based multi-resource allocation strategy (SMRAS) proposed by Hassan et al.[12] and Semi markovian decision processing and with Efficient Resource Provisioning (ERP) proposed by Liu et al.[13]

Suppose that the number of resources in the network (in terms of physical machines) is [0, 50], and the total

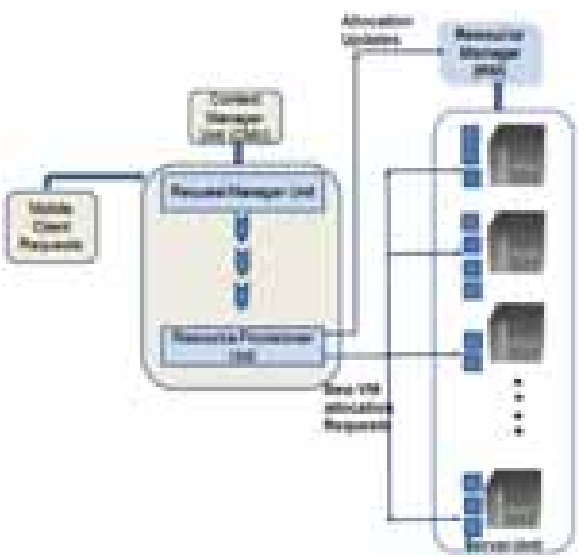


Fig. 1. Proposed ORPM.

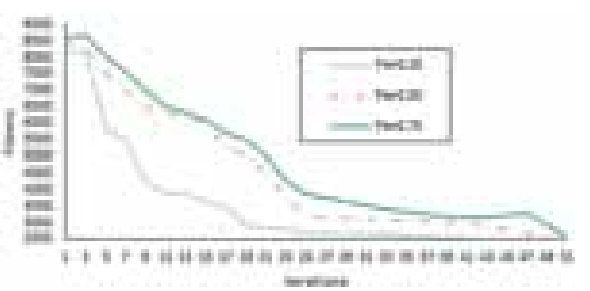


Fig. 2. Execution time by varying P_a.

number of jobs is 50–500. Figure 2 shows the execution time for this scenario when the probability of abandon eggs is 0.25, 0.5 and 0.75.

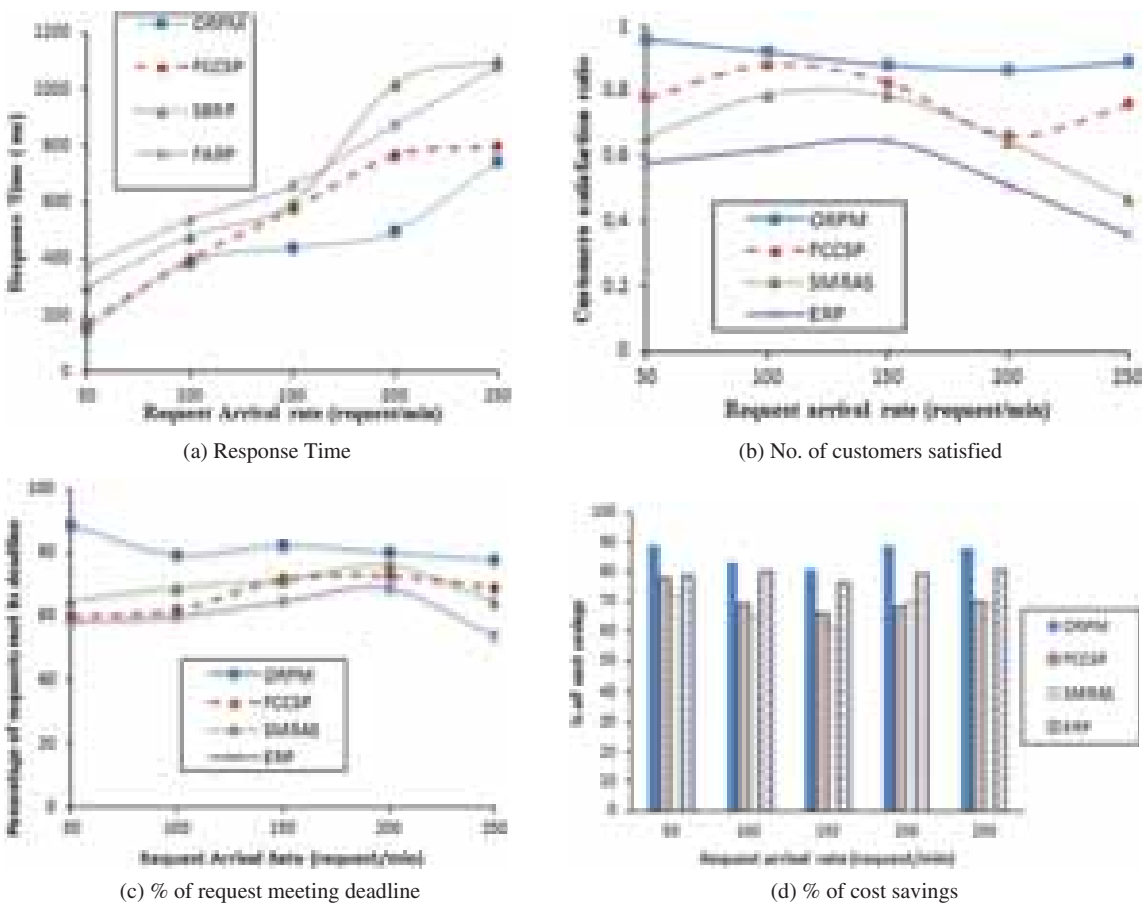
It is evident from the figure that when the abandon probability is low, the convergence speed of the algorithm becomes high.

5.2.1. Impacts of Varying Arrival Rate

This section presents the evaluation of our proposed ORPM based on service response time, customer satisfaction ratio, percentage of requests meeting its deadline, resource utilization by varying the request arrival rate. The impacts of request arrival rate on service response time, customer satisfaction ratio, and percentage of requests met the deadline and percentage of mobile clients cost savings are illustrated in Figures 3(a)–(d) respectively. Figure 3(a) shows that the service response time increases when the request arrival rate increases. When $\lambda = 250$, the response time is as much as 16% higher than that by $\lambda = 50$. When $\lambda = 100$ requests/minute, the service response time of the proposed ORPM is close to the response time of the request using FCCSP. The performance differences among ORPM, FCCSP, SBRP, and FARP obviously appear when λ is over 100 requests/minute. As shown in Figure 3(a), the proposed ORPM provides satisfactory services at a much lower response time than the other three relative approaches. As seen in Figure 3(b), as the request arrival rate is high, the customer satisfaction ratio of ERP drops more quickly than ORPM, FCCSP, and SBRP. ORPM has greater execution success ratio than the other three compared algorithms as λ becomes high. When the arrival rate is 150 requests per minute, the customer satisfaction ratio of ORPM becomes similar to FCCSP and SBRP.

Figure 3(c) illustrates the percentage of requests met the deadline with varying request arrival rates. The percentage drops dramatically right after the λ reaches 200 requests/min. This phenomenon is caused by cloud system performance degradation due to too many request processing. When the λ is 200, the percentage of requests met the deadline is similar to FCCSP and SBRP and is outperforms ERP by 11%.

Figure 3(d) plot the mobile client cost savings under various arrivals. It can be seen from the figure that the percentage of cost savings of ORPM decreases up to the



(a) Response Time

(b) No. of customers satisfied

(c) % of request meeting deadline

(d) % of cost savings

Fig. 3. Impacts of arrival rate on system performance.

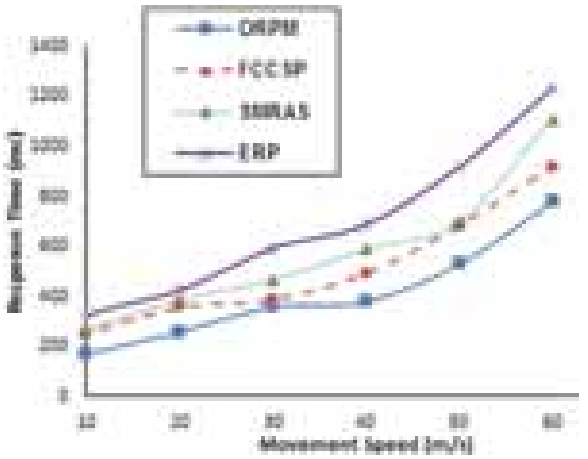


Fig. 4. Impacts of response time by varying movement speed.

arrival rate 150 request/min and increases when arrival rate is over 150. Since the request state base resource allocation and deallocation algorithm have been used to minimize the cost of ORPM system, cost savings of mobile clients get increased compared to FCCSP, SMRAS, and ERP.

5.2.2. System Performance When Varying Movement Speed

Figure 4 shows the varying response time under different mobile client's movement speed. From the figure, we can see that, as the movement speed increases, so is the response time of the request. The response time of the ORPM has a steady increase with a speed. This phenomenon is caused by the increase in disconnections due to increased mobility and quick battery drain. When the movement speed is 5 m/s, the response time of ORPM is similar to the other three related algorithms.

The results prove that ORPM performs better when compared to FCCSP, SMRAS, and ERP.

6. CONCLUSION

The goal of this paper is to design and solve the optimization problem. The proposed model aims to find the optimal allocation of resources as well as to ensure efficient management of resources in MCC environments. In the simulation, the proposed model is compared with other three associated algorithms. The performance evaluation of this model shows that the proposed framework achieves, significant performance improvements in terms of service response time, customer satisfaction ratio, and energy consumption. We intend to extend this work on more varied parameters, under multi-cloud environments.

References

1. D. Huang, T. Xing, and H. Wu, *IEEE Network* 27, 6 **(2013)**.
2. Atta ur Rehman Khan, Mazliza Othman, Sajjad Ahmad Madani, and Samee Ullah Khan, *IEEE Commun. Surveys Tuts* 16, 393 **(2014)**.
3. E. Ahmed, A. Gani, M. K. Khan, R. Buyya, and S. U. Khan, *Journal of Network and Computer Applications* 52, 154 **(2015)**.
4. Atta ur Rehman Khan, Mazliza Othman, F. Xia, and Abdul Nasir Khan, *IEEE Cloud Computing* 2, 42 **(2015)**.
5. S. Durga and S. Mohan, Mobile cloud media computing applications: A survey, *Proceedings of the 4th International Conference on Signal and Image Processing 2012 (ICSIP 2012)*, edited by S. Mohan and S. Kumar Suresh, Lecture Notes in Electrical Engineering, Springer, India **(2013)**, Vol. 222, pp. 619–628.
6. A. Verma and S. Kaushal, *International Journal of Grid and Utility Computing* 5, 96 **(2014)**.
7. A. L. Jin, W. Song, P. Wang, D. Niyato, and P. Ju, *IEEE Transactions on Services Computing* 9, 895 **(2016)**.
8. J. Chase and D. Niyato, *IEEE Trans. Services Comput.* 10, 396 **(2017)**.
9. S. Mireslami, L. Rakai, B. Homayoun, and M. Wang, *IEEE Trans. Network and Service Management* 14, 676 **(2017)**.
10. K. Gai, M. K. Qiu, and H. Zhao, *Journal of Parallel and Distributed Computing* 111, 126 **(2017)**.
11. S. Midya, A. Roy, K. Majumder, and S. Phadikar, *Journal of Network and Computer Applications* 103, 58 **(2018)**.
12. M. M. Hassan, M. Al-Qurishi, B. Song, and A. Alamri, Efficient resource provisioning for mobile media traffic management in a cloud computing environment, *Algorithms and Architectures for Parallel Processing, ICA3PP 2014*, edited by X. Sun, et al., Lecture Notes in Computer Science **(2014)**, Vol. 8630, pp. 352–363.
13. Y. Liu, M. Lee, and Y. Zheng, *IEEE Transactions on Mobile Computing* 15, 2398 **(2016)**.
14. L. Chunlin, Y. Xin, Z. Yang, and L. Youlong, *Journal of Computer Networks* 122, 138 **(2017)**.
15. X. S. Yang and S. Deb, Cuckoo search via levy flights, *World Congress on Nature and Biologically Inspired Computing* **(2009)**, pp. 210–214.
16. X. S. Yang and S. Deb, *Int. J. Math. Model. Numer. Optim.* 1, 330 **(2010)**.
17. J. Vilaplana, F. Solsona, I. Teixido, J. Mateo, F. Abella, and J. Rius, *The Journal of Supercomputing* 69, 492 **(2014)**.
18. A. Beloglazov and R. Buyya, *Concurrency and Computation: Practice and Experience* 24, 1397 **(2012)**.
19. A. Beloglazov and R. Buyya, *IEEE Transactions on Parallel and Distributed Systems* 24, 1366 **(2013)**.
20. J. Wilkes and C. Reiss, Details of the Cluster Data-2011-1trace **(2011)**, https://code.google.com/p/.

Received: 24 April 2018. Accepted: 7 June 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1769–1773, 2018

An Intruders Attack Towards Each Layer in Wireless Networks

D. N. Aditya, M. Gowtham*, and C. Preetham

Department of Computer Science & Engineering, Rajeev Institute of Technology, Hassan Karnataka 573201, India

The movement to wireless system from wired system has been a worldwide pattern in these previous decades. The versatility and adaptability brought by wireless network made it conceivable in numerous applications. Among all the contemporary remote net-works, WSN is a standout amongst the most vital and novel applications. On the in spite of customary system engineering, MANETs does not require a settled system foundation; each and every hub (node) functions as both a transmitter and a beneficiary. Hubs discuss specifically with each other when they are both inside a similar correspondence extend. Else, they depend on their neighbors to transfer messages. The self-arranging capacity of hubs in MANETs made it famous among basic mission applications like military utilize or crisis recuperation. Lately, security has turned into a most essential administration in wireless sensor networks. Contrasted with different systems, WSN is powerless against different kinds of assaults. In this paper, a relative investigation of threats assaulted in each layer of the system is discussed.

Keywords: MANET's, Data Authenticity, WSN, Attacks on Sensor Networks, Denial of Service, Reliability, Availability, Hub.

RESEARCH ARTICLE

1. INTRODUCTION

Nowadays, fame of synergistic and portable applications in remote systems particularly in MANETs prompted making them appropriate for some security sensitive applications (e.g., military ones).

Early proposed arrangements were for the most part in view of the assault situated methodologies. More accurately, they have endeavored to recognize or keep a few sorts of assaults.[1–10] Afterwards, Cryptography has been utilized to give a general plan structure.[11] In fact, the nature of asset compelled hubs is one of the noteworthy issues that authorized the designers to propose lightweight and less asset expending crytographic secure plans in WSN. In any sort of remote systems security and secrecy are basic issues. In wireless sensor network (WSN) the significance of security is important. Wireless sensor network (WSN) and its applications are assaulted by Intrusions and different assaults to intrude on the qualities it serves. Remote wireless sensor network is generally given to the remote regions. Wireless sensor network is utilized to decide the distinctive natural observing applications, for example, temperature, pneumatic stress, moistness discovering.[11] Since sensor systems are often utilized as a part of remote territories and in those zones, the activity of the system is left unattended, and unmonitored routinely that assistance the interlopers to make a simple focus for considerable assaults, unlawful access and treating. Sensor hubs need to bargain with assaults because of asset decreased exercises and working in unaware surroundings likewise hard to separate security infringement from hub disappointments. Issues are likewise made in supporting connection characteristics and general close down of the system. These asset restrictions require security components that are intended for WSN applications, with the goal that the constrained assets can be utilized capably. This paper shows the diverse verities of normal assaults which are debilitating this system.

2. RELATED WORK

Thinking about any kinds of the systems while interfacing any two PCs by utilizing wired or wireless sensor network, gatecrasher may attempt assault in any layer. This paper investigation the kinds of assault may happens in every layer while exchanging information between any two gadgets. Diverse dangers at each layer in OSI show are abridged as underneath.

*Author to whom correspondence should be addressed.

Table I. Attacks in layers.

Layers	Types of attack in each layers
Application layer	Attacks on reliability
Presentation layer	Achieving secure communications, performing authentication, performing authorization, using code access security
Session layer	Session fixation, session side jacking, cross-site scripting, Malware and unwanted programs
Transport layer	Injects false messages, energy drain attacks
Network layer	Spoofing or replaying information, selective forwarding or black holes, sink holes, sybil attacks, node replication attacks, wormholes flooding, attacks against privacy
Data link layer	Jamming, collision
Physical layer	Jamming, tampering

2.1. Kinds of Attacks in Each Layer

2.1.1. Sensor Network Security in Physical Layer

The target of physical layer is to build the dependability by decreasing way misfortune impact and shadowing. This layer is in charge of built up association, information rate, regulation, information encryption, flag location, and recurrence age and flag identification. The most widely recognized assaults on the physical layer are jamming and tampering.

1. Jamming

The radio flag transmission can meddle with the radio frequencies utilized by the WSN, which is called jamming. As the enemy ability expands, it can influence bigger bits of the system by sending other radio signs. The enemy can utilize couple of hubs to obstruct the whole system. This condition is called Jamming at the physical layer and henceforth bringing about denial of service. In this circumstance the enemy won't have the capacity to get any information yet will be capable to hinder a few hubs.

2. Tampering

Now and then the hubs are physically altered by a foe. Such condition is called altering. An altering assailant may harm, supplant, and electronically question the hubs to obtain information.[10]

2.1.2. Sensor Network Security Issues at Data Link Layer

The target of Data interface layer is to protect interoperability among correspondence between hubs to hubs. This layer is in charge of mistake location, multiplexing, avoidance of impact of bundles, rehashed transmission and so forth. The information interface layer is helpless because of the reason that the information is transmitted in an open unreliable medium. Subsequently it is helpless to the assaults on the realness, trustworthiness and privacy of the information being directed.[1] The primary assaults at information interface layer are crash and sticking.

1. Collision

For WSNs the exchange of an information parcel may fizzle if the radio channel was presently involved by another sensor hub. This outcomes in event of an excessive number of impacts on the radio channel. In this way, we can't stand to set up an asset requesting correspondence between a base station and the sensor hubs. Crashes can be maintained a strategic distance from with the particular schedule vacancy task to every sensor hub.

2. Jamming

Jamming can happen when the information get stuck with radio signs from different transmissions.

2.2. Sensor Network Security Issues at Network Layer

The goal of Network layer is to discover best way for proficient steering system. This layer is in charge of directing the information from hub to hub, hub to sink. Hub to base station, hub to group head and the other way around. Vulnerabilities at organize layer are,

- Spoofing or replaying information
- Selective forwarding or black holes
- Sink holes
- Sybil attacks
- Node replication attack
- Wormholes
- Flooding
- Attacks against privacy

1. Spoofing

A Spoofing assault is a circumstance in which one individual or program effectively masks as another by misrepresenting information and along these lines picking up an ill-conceived advantage. Through ridiculing, or replaying the directed data the system activity can be broadly undermined. Consistent modifications in messages which result in the bundle misfortune amid transmission may require the individual hubs to retransmit parcels persistently. Hence the hubs may turn out to be dead significantly sooner than their normal life because of energy depletion. Likewise some of the time replaying messages brings about making tremendous measure of movement stream on the system.

2. Selective forwarding or black holes

Ordinarily the sensor systems are multi-hop frameworks. In this way, the sensors pass data from one end to the base station by steering them through middle of the road nodes. Sometimes a malevolent hub might be available inside the system way. In a flooding based convention, the aggressor (pernicious hub) tunes in to demands for courses at that point answers to the objective hubs that it contains the high caliber or most limited way to the base station.[9] Then the target may pick the course which contains the malignant hub. This malignant hub exhibit in the course may specifically forward the information parcels, i.e., advances a few parcels to the following hub, and drops others. The outcome is loss of tremendous measure of information, amid the multibounce data trade process. For another situation it might happen that the vindictive hub drops every one of the parcels it gets, subsequently

no data is sent. This makes a dark gap. Such assaults are compelling when the assailant is expressly incorporated into the information way of sensor organize.

3. Sinkholes

In this assault the aggressor baits the majority of the sensor arranges movement to go through the pernicious hub hence making a sinkhole with noxious hub at its middle. Since now the vast majority of the information is being steered through the malevolent hub, the aggressor/malignant hub can play anything with the sensor information.[5] Many different assaults, for example, wormhole, specific sending or listening in can be started through this sinkhole assault.

4. Sybil Attacks

By and large, the sensors in a remote sensor system may need to cooperate to achieve an errand, thus they can utilize dissemination of subtasks and excess of data. In Sybil assault,[6] the aggressor/malevolent hub demonstrate different characters. Since each real hub in a sensor organizes has a solitary personality, thus various dangers can be watched. Since foe has different personalities, the pure hubs might defeat multi-way information through the same vindictive hub. Sybil assault can be performed for assaulting the circulated stockpiling, steering system, information collection, voting, reasonable asset portion and mischief recognition. Essentially, any shared system (particularly remote specially appointed systems) is helpless against Sybil assault. Be that as it may, as WSNs can have a type of base stations or portals, this assault could be counteracted utilizing effective conventions. Location of Sybil hubs in a system isn't so natural. Newsome et al. utilized radio asset testing to recognize the nearness of Sybil node(s) in sensor arrange and demonstrated that the likelihood to distinguish the presence of a Sybil hub.

5. Wormholes

Two noxious hubs may make a shrouded channel (route) between them.[8] This is known as wormhole assault. The two pernicious hubs might impart over intense information interface when contrasted with the connection between the genuine sensor hubs. For this situation likewise there numerous conceivable outcomes of assaults.[1] The two malignant hubs might be available in various areas in the system, and the assailant records the bundles (or bits) at one area.

6. Flooding

Now and then the pernicious hub can cause massive movement of pointless messages on the system. This is known as the flooding. A few, times noxious hubs replay some actua communicate messages, and subsequently creating pointless movement on the system. This can cause blockage, and may in the long run prompt the depletion of finish hubs. This is a type of Denial of Service attack.[1]

2.2.1. Sensor Network Security Issue at Transport Layer

The goal of Transport Layer is to build up correspondence for outside systems i.e., Sensor organize associated with the web. This is most testing issue in remote sensor systems. Once in a while the assailant may be sufficiently solid to reach up to the vehicle layer, because of the assault being undetected at the lower layers. Transport layer assaults are infusion of false messages and vitality deplete assaults[11] and are named takes after:

1. Data Integrity Attack

Data Integrity assaults trade off the information going among the hubs in WSN by changing the information contained inside the parcels or infusing false information. The aggressor hub must have all the more preparing, memory and vitality than the sensor hubs. The objectives of this assault are to misrepresent sensor information and by doing as such trade off the casualty's examination. It likewise distorts steering information keeping in mind the end goal to disturb the sensor system's ordinary task, perhaps making it pointless. This is thought to be a kind of foreswearing of administration assault. This assault can be safeguarded by adjusting halter kilter key framework that is utilized for encryption or we can utilize computerized marks, however this requires a ton of extra overhead.

2. Energy Drain Attack

WSN is battery controlled and powerfully sorted out. It is troublesome or difficult to supplant/revive sensor hub batteries. Since there is a constrained measure of vitality accessible, assailants may utilize traded off hubs to infuse created reports into the system or produce expansive measure of activity in the system. Created reports will cause false cautions that waste genuine reaction endeavors, and deplete the limited measure of vitality in a battery fueled system. However the assault is conceivable just if the interloper's hub has enough vitality to transmit bundles at a consistent rate. The point of this assault is to annihilate the sensor hubs in the system, corrupt execution of the system and at last split the system matrix and therefore take control of part of the sensor organize by embeddings another Sink hub.[11] To limit the harm caused by this assault manufactured reports ought to be dropped on the way as right on time as would be prudent.

2.2.2. Sensor Network Security Issue at Session Layer

1. Session Fixation

Where the aggressor sets a client's session id to one known to him, for instance by sending the client an email with a connection that contains a specific session id. The assailant now just needs to hold up until the point when the client signs.

2. Session Side

Jacking Where the assailant utilizes bundle sniffing to peruse organize movement between two gatherings to take the session treat. Numerous sites utilize SSL encryption for login pages to keep aggressors from seeing the watchword, yet don't utilize encryption for whatever is left of

the site once validated. This permits assailants that can read the system movement to block every one of the information that is submitted to the server or website pages saw by the customer. Since this information incorporates the session treat, it enables him to imitate the casualty, regardless of whether the watchword itself isn't compromised.[1] Unsecured Wi-Fi hotspots are especially defenseless, as anybody sharing the system will by and large have the capacity to peruse the majority of the web activity between different hubs and the entrance point.

3. Cross-Site

Ccripting Where the assailant traps the client's PC into running code which is dealt with as reliable on the grounds that it seems to have a place with the server, enabling the aggressor to get a duplicate of the treat or perform different activities.

4. Malware

Malware and undesirable projects can utilize program commandeering to take a program's treat records without a client's learning, and after that perform activities (like introducing Android applications) without the client's knowledge.[2] An assailant with physical access can essentially endeavor to take the session key by, for instance, getting the document or memory substance of the suitable piece of either the client's PC or the server.

2.2.3. Sensor Network Security Issue at Presentation Layer

1. Accomplishing Secure Communications Trading Information

Over uncovered systems, for example, the Internet, acquaints dangers with the security of your interchanges. In the event that your application forms delicate information, for example, charge card subtle elements or therapeutic records—you should execute systems to ensure the mystery and respectability of the information as it goes between application segments.

2. Performing Authentication

The reason for confirmation is to safely set up the character of a man who needs to utilize your application. Approval and inspecting require verification. The most well-known verification method on the Windows stage is the utilization of client names and passwords (utilizing worked in Windows validation or a system, for example, Microsoft .NET Passport). However, different instruments, for example, savvy cards, biometrics, and advanced declarations, are picking up in fame as they turn out to be more available and simpler to execute. This is particularly valid in applications that require more elevated amounts of security.

3. Performing Authorization

Approval is the way toward deciding if a client has consent to get to a specific asset or bit of usefulness. To perform approval, you should have a confirmation component set up to build up the character of the client, and that instrument must decide the personality of the client precisely and dependably. Utilizing Code Access Security Code get to security is a security include that applies to all .NET. Framework oversaw code to shield PC frameworks from malignant code and to give an approach to enable versatile code to run securely.

2.2.4. Sensor Network Security Issue at Application Layer

The goal of Application Layer is to exhibit last yield by guaranteeing smooth data stream to bring down layers. This layer is in charge of information gathering, administration and handling of the information through the application programming for getting dependable outcomes. Primary assault at application layer is assaults on unwavering quality

1. Attacks on Reliability

In the event that a foe changes the information in one way then it puts a question mark on the unwavering quality of the information. In this assault assailant needs to distinguish the way of correspondence and place foe in that way to change the information. A foe can create false information or inquiry by joining the system. At the point when a hub reacts to these wrong information or inquiry, drives them to experience the ill effects of the vitality deplete assault. As a rule to guarantee unwavering quality affirmation is normal for each effective information conveyance.[11]

3. CONCLUSION

Wireless system is utilized as a part of various kinds of uses because of huge development. Thus security and unwavering quality turns into the primary worry in each remote sensor arrange applications. This paper presents study on various kinds of assaults conceivable on various layers of system. Security defenses must be sufficiently solid to maintain a strategic distance from foe to influence the system, as remote system is more inclined to assaults, it is important to deal with the information with full privacy and with high security.

References

1. G. Bianchi, *International Journal of Advanced Science and Technology* 17, 31 **(2010)**.
2. J. Sen, *Journal of Communication Networks and Information Security (IJCNIS)* 1, 55 **(2009)**.
3. J. Rehana, Security of wireless sensor network, *TKK T-110.5190 Seminar on Inter Networking* **(2009)**.
4. M. Sharifnejad and M. Sharifi, A survey on wireless sensor networks security, *4th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisia **(2007)**.

5. C. Karlof and D. Wagner, *Ad Hoc Networks* 1, 293 **(2003)**.
6. A. Singla and R. Sachdeva, Review on security issues and attacks in wireless sensor networks, *IJARCSSE* **(2013)**, Vol. 3.
7. Lovepreet Kaur, *International Journal of Computer Applications (0975–8887)* 100 **(2014)**.
8. B. Parno, A. Perrig, and V. Gligor, Distributed detection of node replication attacks in sensor networks, *Proceedings of IEEE Symposium on Security and Privacy* **(2005)**.
9. A. D. Wood and J. A. Stankovic, *IEEE Computer* 35 **(2002)**.
10. Asif Habib, Sensor network security issues at network layer, *2nd International Conference on Advaificances in Space Technologies*, National Engineering and Scientific Commission, Islamabad, Pakistan **(2008)**, pp. 58–63.
11. A. Lapidoth and P. Narayan. *IEEE Transactions on Information Theory* 44, 2148 **(1998)**.
12. A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, Wireless Networks, SPINS: Security Protocols For Sensor Networks, Springer **(2002)**.

Received: 22 April 2018. Accepted: 1 July 2018.

Journal of
Computational and Theoretical Nanoscience
Vol. 15, 1774–1778, 2018

RESEARCH ARTICLE

Evaluating UDP Traffic and Machine Learning Based Intrusion Detection in Software-Defined Network

D. Saranya* and S. Kuppuswami

Department of Computer Science, Kongu Engineering College, Erode, Tamil Nadu 638052, India

In the modern era, the growth of the Internet is getting wider with enormous impact on traffic and security issues. To ensure high availability of storage, data centers are evolving. In order to maintain the complex operations over the highly interconnected network, data center administrators run multiple traffic management tasks. Software Defined Networking is the new technology paradigm shift which reduces the network management complexity. OpenFlow is the communication protocol which is the enabler of software-defined network. The system allows the network administrator to classify a variety of traffics in a network to enhance the security. The projected algorithmic program reduces the number of packets out of order by increasing the hard timeout time of flow entries once the CPU's usage will increase and by gathering statistics from the switches. The system allows the network administrators to classify varied traffics in a network to enhance the security. The svm classifier will classify the normal and attack packets. On classifying attack packets the system will produce the alert to ensure the network security. This work enhances the security in software defined networks.

Keywords: SVM, Machine Learning, Software-Defined Networking, Intrusion Detection, Packets Out of Order.

1. INTRODUCTION

Software-defined network (SDN) is an emerging networking architecture in which control plane is decoupled from the forwarding plane. In this architecture network control and the forwarding functions are separated in such a way that the network controlling can be programmed directly. SDN architecture emphasizes dynamic, manageable and adaptable network, which is ideal for the high bandwidth requirements. SDN, in other words, is a framework which separates control logic and provides the centralized control to the network. SDN architecture comprises of a controller, northbound Application Program Interfaces (APIs) and southbound APIs. SDN provides virtualization, which runs software separately from the underlying hardware.

The problem with the SDN is communication delay between the Controller and switch (RTT) which is address in this project by varying the ideal time of the flow entries.

The control plane realizes the flexible service access and the unified hardware control through its two interfaces, i.e., open southbound interface (SBI) and northbound interface (NBI), respectively. However, the centralized control plane makes it vulnerable to numerous intrusions, as once the controller is hijacked by the attacker, the whole network will be faced up with an out of order circumstance. Therefore, the secure operations of both the controller and the network are compromised. The existing intrusion detection techniques mostly focus on examining break-ins at compute level. However, SDON control plane requires detection occurring at the network level. What is also worth noticing is that there are increasingly more researchers focusing on analyzing optical network by means of machine learning techniques. Figure 1 represents the basic SDN controller function.

2. LITERATURE SURVEY

The most noteworthy of various jounals relevant to the current study are reviewed, below.

Chen, Ding et al.[1] proposed the concept of collaborative controllers. SDN-based applications can be deployed on large-scale networks only with east/west bound interfaces. However, there is still a long way to standardize SDN east/westbound due to the high complexity and comprehensive optimization goals. To make it attainable to deploy SDN within the internet to support each ancient and rising network applications, author projected a lightweight collaborative controller mechanism that has higher

*Author to whom correspondence should be addressed.

 1546-1955/2018/15/1774/005 doi:10.1166/jctn.2018.7388

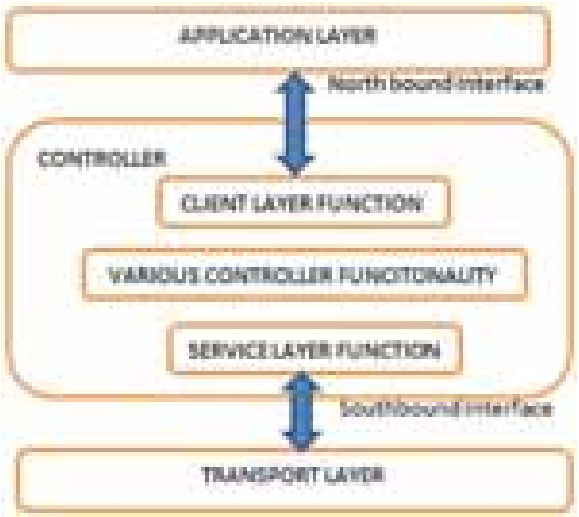


Fig. 1. Function of SDON controller.

characteristics such as application independence, controller independence, minimum functions, reliability and etc. The author also designed its software program named LCM-SCware running on each controller and its application layer protocol enabling corporation among the controllers. The prototype system is implemented and the experimental results demonstrate that LCMSC enables efficient cooperation among heterogeneous controllers in multiple SDN domains connected by Internet and support both existing and emerging services. The contributions of this paper include: (i) A lightweight collaborative mechanism for SDN controllers (LCMSC) and its model are proposed; (ii) The software program of LCMSC named LCMSCware running on the controllers and the LCMSC protocol for controller communication are designed; (iii) The LCMSC prototype is implemented, and the effectiveness of the mechanism is verified. This lightweight mechanism should possess the following properties such as application independence, function minimization, controller independence, and reliable communication.

Song, Liu et al.[2] surveyed Using Software-Defined Networking (SDN), the flexibility and programmability of networks can be significantly increased through the decoupling of the control and data planes. However, network scale-up in large-scale datacentres can rapidly increase the computational complexity of operations such as the shortest path calculation on the network topology or Quality-of-Service (QoS) routing, which, in turn, can cause quality problems in current SDN controllers. ParaFlow that is projected in this paper represents a multithreaded SDN controller which supports fine grained and utilizing multi-/many-core resources by exploiting application to accelerate event methods. ParaFlow also provides a flow-based programming interface that allows application developers to program with network flows rather than various types of low-level events. To solve this scalability problem, this paper proposes ParaFlow, a multithreaded controller that supports fine-grained parallelism in event processing by exploiting parallelism in applications and utilizing multi-/many-core resources. Another characteristic of ParaFlow is its use of a flow-based programming interface that allows application developers to write programs using network flows instead of assorted low-level events.

The main contributions of this paper include 1. Fine-grained parallelism support 2. Flow-based programming interface. The major components of the system include 1. Events, 2. flows, 3. states, 4. event_dispatcher, 5. flow_dispatcher and 6. state_dispatcher. This paper concludes that, an SDN controller that supports fine-grained parallelism and extends event-based SDN programming by abstracting flows and network states. ParaFlow preferentially exploits multicore/many-core resources to solve the computational complexity growth problem via fine-grained parallelism. Our experimental evaluation confirmed that ParaFlow can achieve satisfactory performance and scalability with the number of computing resources. Our future work will focus on extending ParaFlow to current distributed SDN control planes.

Oktian, Lee et al.[3] In this paper, author surveys and investigates those approaches. Author has discovered that they can classify the methods into several design choices that are trending among SDN adopters. Each design choice may influence several SDN issues such as scalability, robustness, consistency, and privacy. Thus, the author further analyses the pros and cons of each model regarding these matters. Author concludes that every design begets some characteristics. The author additionally presents the planning decisions to make distributed controller that's scalable, robust, consistent and secure.

The various design choices are 1. Switch to controller connection strategy, 2. Network information distribution strategy, 3. Controller coordination strategy, 4. In-band versus out-of-band connection strategy. The author concludes that each design has its own characteristic. They may excel in solving one issue but suffer from the others. For instance, the flat model is excellent in controlling failure problems but has several disadvantages in scalability, privacy, and consistency. Second, the author summarizes the combined pros and cons from each design to construct a guideline for designing future distributed SDN controllers. When one wants to build the most scalable distributed controller, then consider building with a combination of Master/Slave, hierarchical, leader-based and out-of-band design. Similarly, one can also generate other combinations such as the most robust controller with Master/Slave, flat and leaderless combination. The SDN controller with the best privacy is the combination of the hierarchical and out-of-band. Lastly, the controller with the best consistency is the combination of IP alias, hierarchical, leader-based and in-band design. Several related works

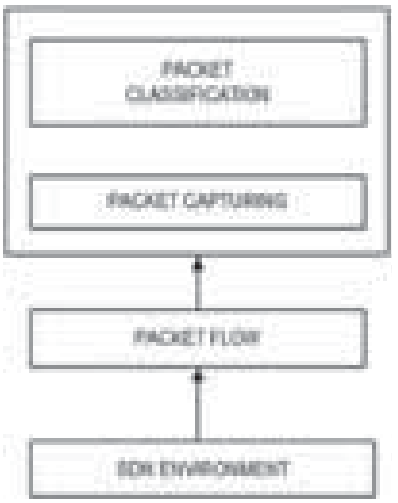


Fig. 2. Packet classification.

are capable of discovering and solving some fundamental issues of distributed SDN controller. By combining those works with our previous analysis, one can set up a desired distributed SDN controller with a custom trade off among scalability, robustness, consistency and privacy.

3. PROPOSED SYSTEM

A novel technique for analysing UDP traffic in SDN network is proposed that works in two phases: topology construction and analysis of traffic. OpenFlow-based SDN has two shortcomings that affect the network's performance. First, the communication between the controller and the switch introduces significant delay in the round-trip time (RTT). Second, the RTT delay ends up in packet loss or packets inward at the destination out of order. These two major issues are more likely to occur with UDP traffic due to its inherent characteristics. The proposed work examines the reasons behind these issues, hypothetically discusses some potential solutions and implements the proposed algorithm. The proposed method reduces the number of packets out of order by increasing the hard timeout time of flow entries when the CPU's usage increases and by gathering statistics from the switch more often. During the demonstration of OpenFlow-based SDN, a POX controller, Jperf is employed for the analysis purpose diagrammatically and Mininet is employed to emulate the infrastructure required. Here, testing is performed on two different topological networks and analysis of the UDP packets using the Jperf tool is done with the number of bytes being transmitted per second. Traffic from SDN is forwarded and the classification is done through the SVM classifier. The SVM classifier efficiently classifies the two varied specified classes i.e., normal and attack packets. On classifying the attack packets the alert will be generated in order to provide the secured flow in the SDN.

4. IMPLEMENTATION

4.1. Topology Construction

This topology deals with pox controller, OVSK switch and hosts connected to the switch. Figure 3 represents the Linear Topology which consists of a single switch connected to two hosts. The Switch used here is the default OVSK Switch. The controller used is the POX controller which is available default in Mininet Software and its control is coded using the Python.

4.2. Bandwidth Analysis

Performance analysis is finished using Jperf tool. JPerf is intended to run as a client/server application. To run a check we are going to got to found out a JPerf server on your network. Then we are able to run a JPerf client from another location on the network which can connect to the remote server. To begin the JPerf server choose the radio button labeled server then click Run IPerf. By default, JPerf runs in communications protocol mode and listens on port 5001.

There are several options that can be adjusted to modify the parameters of the test. In this project UDP buffer size and bandwidth has been varied. Throughput analysis has been done for the flow entries and the graph will update in real time to reflect the results. Figure 4 represents the variation of the bandwidth with increases in number of packets.

4.3. Detection of Intrusion

The SDN test bed is the module which emulates the SDN environment with hosts and switches. The controller divides the program logic, filters and aggregates the results in efficient way. This module will insert and forward the flows from the SDN. The packets from the SDN environment is captured and preprocessed into dataset by this module. This module performs the packet inspection

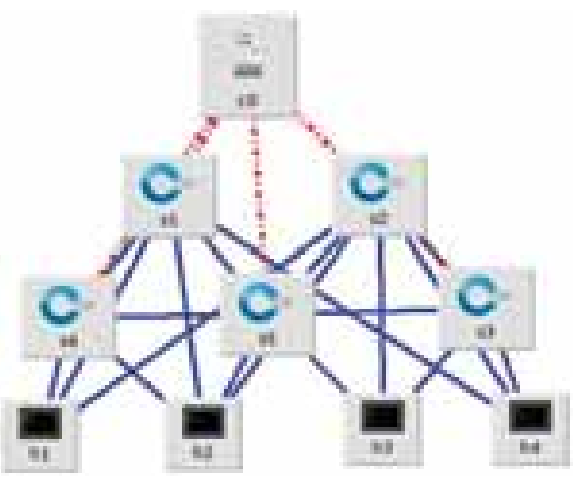


Fig. 3. Topological network.

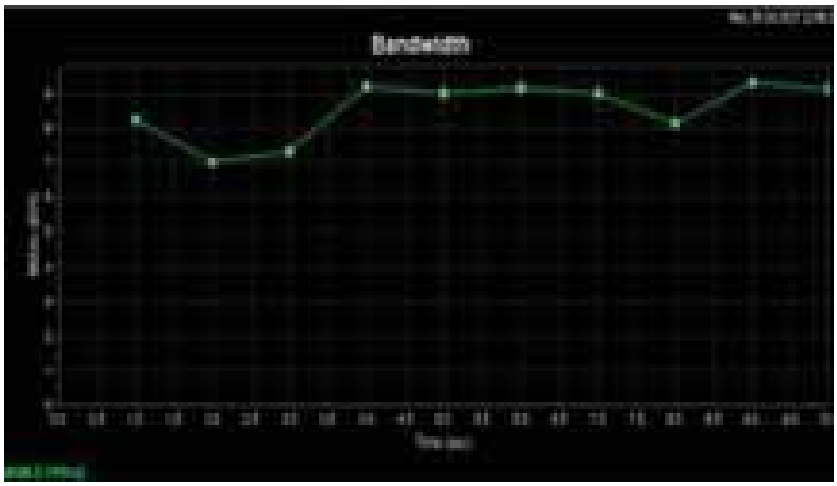


Fig. 4. Bandwidth variation.

which is forwarded from the controller of the SDN testbed. This module is SVM classifier which classifies the traffic into normal and attack packets.The hardware required for the experiment is Intel Core i3, 4 GB RAM and 500 GB hard disk drive. The software tools required are the Ubuntu 14.04 and Eclipse LUNA. The language used for the implementation of the project is java. The controller used is FloodLight. The Instant Virtual Network used is Mininet. IDS application used is SNORT IDS. For classification the WEKA datamining tool is used.

4.4. Performance Analysis

The parameters considered for evaluating the performance of the proposed idea are Detection rate, False positive rate and Accuracy.

The detection rate is defined as the number of intrusion instances detected by the system (True Positive) divided by the total number of intrusion instances present in the test set. The formula is given in the Eq. (1).

Reducing the false positive rate improves the performance of the system. The formula is given in the Eq. (2).

Accuracy rate is the percentage of test set samples that are correctly classified by the model. The formula is given in the Eq. (3).

$$\text{Detection Rate} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{1}$$

$$\text{False Positive Rate} = \frac{\text{FP}}{\text{FP}+\text{TN}} \tag{2}$$

$$\text{Accuracy} = \frac{\text{TN}+\text{TP}}{\text{TN}+\text{TP}+\text{FN}+\text{FP}} \tag{3}$$

5. RESULT OBTAINED

The analysis obtained from the jperf tools is that the bandwidth varies with the different networking topologies.

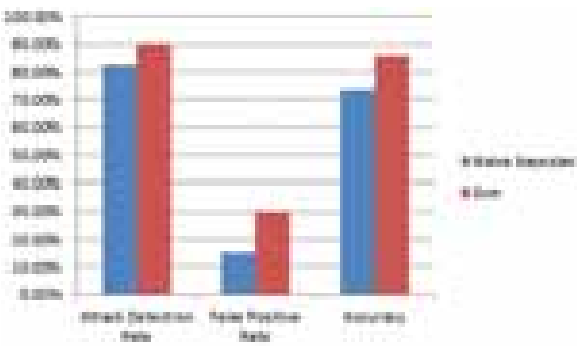


Fig. 5. Comparison of classifiers.

By controlling up of ideal and hard timeout time, the out of order packets can be controlled by increasing the timeout time depending upon the arrival of the packets. From the above Figure 5, it is concluded is that the SVM classifier produce more accuracy than the Naive bayes classifier. Thus the system provides better accuracy in classifying attacks which shows the enhancement in SDN security.

6. CONCLUSION

The growth of network is being tremendous and is in need of traffic management and security. The analysis has been done with the various networking topologies and with more numbers of switches, and the traffic is controlled by varying the timeout time of the flow. the traffic classification technique based on SVM classifier which in turn enhance the security in SDN. This system focuses both on traffic classification and security in SDN.

7. FUTURE ENHANCEMENTS

The future work can be dealt with automated system and more efficient classifiers. Rule based network policies can

be incorporated into the existing system which make the system more efficient and automated. The rule based network policies are set of rules that will allow and deny the traffic flow accordingly. This rule based policies can be dealt in the future work.

References

1. M. Chen, K. Ding, et al., *Computer Networks* 121, 65 **(2017)**.
2. P. Song, Y. Liu, C. Liu, and D. Qian, *Journal of Network and Computer Applications* 87, 46 **(2017)**.
3. Y. E. Oktian, S. Lee, H. Lee, and J. Lam, *Computer Networks* 121, 100 **(2017)**.
4. B. A. A. Nunes, M. A. S. Santos, B. T. de Oliveira, C. B. Margi, K. Obraczka, and T. Turletti, *IEEE Communications Magazine* 52 **(2004)**.
5. D. Zats, Tathagata Das, Prashanth Mohan, and R. H. Katz, DeTail: Reducing the Flow Completion Time Tail in Datacenter Networks, Technical Report No. UCB/EECS-2011-113 **(2011)**.
6. F. Hu, Q. Hao, and K. Bao, *Journal of IEEE Communication Surveys and Tutorials* 16, 2181 **(2014)**.
7. FloodLight, http://floodlight.openflowhub.org/ **(2011)**.
8. H. Jin, T. Cheocherngngarn, D. Levy, Alex Smith, D. Pan, J. Liu, and N. Pissinou, Joint Host-Network Optimization for Energy-Efficient Data Center Networking **(2012)**.
9. H. Deng, Q.-A. Zeng, and D. P. Agrawal, *IEEE Networks Magazine* 3, 2147 **(2003)**.
10. Linux Advanced Routing and Traffic Control, http://www.lartc.org/ **(2000)**.
11. M. Yu, L. Jose, and R. Mia, University of Southern California Princeton University, software defined traffic measurement with OpenSketch, *Proceeding of 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI'13)* **(2013)**.
12. Open vSwitch. http://openvswitch.org/ **(2011)**.
13. R. Sherwood, G. Gibby, K.-K. Yapy, G. Appenzellery, M. Casado, N. McKeowny, and G. Parulkar, FlowVisor: A Network VirtualizationLayer'http://OpenFlowSwitch.org/downloads/technicalreports/openflow-tr-2009-1-flowvisor.pdf' **(2009)**.
14. Mininet, Mininet: An instant virtual network on your laptop (or other PC)—Mininet, [online] Available at: http://mininet.org/ [Accessed June] **(2016)**.
15. OvS, Open vSwitch, [online] Available at: http://openvswitch.org/ [Accessed 30 June 2016] **(2014)**.
16. iPerf, iPerf—The TCP, UDP and SCTP network bandwidth measurement tool, [online] Available at: http://iperf.fr/ [Accessed June] **(2016)**.
17. VMWarevCenterSuite.http://www.vmware.com/products/datacenter-virtualization/vcenter-operationsmanagement/overview.html **(2013)**.
18. S. Scott-Hayward, G. O'Callaghan, and Sakir Sezer, SDN Security: A Survey **(2012)**.
19. Stuart Andrews, Ioannis Tsochantaridis, and Thhomas Hofmann, Support Vector Machines for Multiple-Instance Learning **(2000)**.
20. T. Benson, Aditya Akella, and David A. Maltz, *ACM-Journal* 23, 390 **(2010)**.

Received: 20 April 2018. Accepted: 16 May 2018.