Adaptive Web-based Systems: From Framework to Implementation

Bujar Raufi, Juliana Georgieva

Abstract: The paper gives an insight of the work done so far in modelling and implementation of user adapted software systems or adaptive web sites. We introduce a five layer framework which gives to our adaptive web based system flexibility, expressiveness, modularity and interoperability. The work so far has been concentrated in the first four layers of the framework concerning data representation, representation of knowledge as well as development of adaptation algorithm for such systems. Initial experimental results on adaptation show good performance of the algorithm for both per-user based adaptation and adaptation for groups of users. The findings also indicate that the approach is suitable for generic adaptive web-based systems.

Keywords: Adaptive web based systems, adaptive hypermedia, user adapted software systems.

INTRODUCTION

Many of the traditional web pages used today suffer from a problem of representing the same set of data to each particular user [1]. This means that when a user accesses a web page of a library, it will be presented with the same set of information every time it visits it. A tourist that tries to find information about travel and hotel arrangements will be presented with the same order of information regardless of its destination preferences or browsing habits [1]. The same issue can be generalized for many, but not all, web applications in various domains like: e-commerce, corporate web sites and virtual education. The solution for overcoming this situation is the development of user-adaptive software systems [1] or Adaptive Web Sites. Adaptive Web Sites represent Web-Based Systems that tend to arrange their content, structure or both based on user access preferences.

Starting as a pioneering work of a few research communities in the middle of 90’s [2] today it represents a research milestone for many communities involving various research disciplines like: user modelling, Web Usage Mining, natural language processing, intelligent tutoring systems, cognitive sciences and semantic web. The adaptation process in adaptive web based systems is done by building a model of goals, knowledge and preferences as well as using those for interaction with the user during its browsing. In traditional adaptive web based systems such an interaction is done by building a user model and applying the same for adapting the content for the purpose of the user itself [3]. This relationship between system, user and adaptation is presented by Brusilovsky in the form of user modelling-adaptation loop [3]. In this approach, the system collects data about user, processes the same and finally results are presented in an adapted way. The adaptation process, as stated by Brusilovsky, is performed through the adaptation methods and adaptation techniques [3].

By adaptation methods we consider a more generalized and abstracted approach concerning adaptation process. This usually represents a description of what the adaptation should do rather than how to do it. Methods usually include a more abstracted representation of adaptation rules like inserting a certain fragment in some part of the page is some conditions are met etc. By adaptation technique on the other hand, we consider a more detailed and well specified approach for delivering adaptation and consequently, as such, they are a part of the implementation process. Each adaptation technique is characterized by specific knowledge representation, modelling, adaptation engine etc [3].
There are two distinctive levels where adaptation can be performed. The first one is known as content-level adaptation and the second is recognized as link-level adaptation. The former is known as Adaptive Presentation and the latter is accepted as Adaptive Navigation Support [3]. In adaptive presentation, the content of a page that is accessed by a particular user is adapted based on the current knowledge base and other user characteristics, goals and preferences, while in adaptive navigation support, the system guides users towards interesting and relevant information while keeps them away from non-relevant information without obstructing its overall navigation freedom.

The main goal of this paper is to give insight information regarding the work done so far in adaptive web based systems as well as to enlighten some new aspects by presenting a new framework and suggesting new adaptation algorithm which will be presented here. The rest of the paper is organized as follows: in section two we give the related work done so far in this area of research with a specific emphasis on frameworks built for these particular systems. Section three deals with work done on the layers of the framework proposed in [4] and section four concludes this paper.

LAYERED FRAMEWORK FOR ADAPTIVE WEB-BASED SYSTEMS

Most of the related work done concerning Adaptive Web-Based Systems is on the design of a suitable framework that will describe the complete functioning of such systems. For this purpose, many frameworks have been developed like: AHA! [6], SKILL [7], Multibook [8], ACE [9], ART-Web [10], MetaLinks [11], LAOS Framework for Authoring Adaptive Hypermedia Systems [12], GOMAWE [13] etc. In [5] are also identified three main design approaches that most of the adaptive web based systems fall in.

The first one is the data-centric approach where the main focus is the data, its organization and relationship. This approach is mostly used in developing model-driven context aware web applications. The second approach is data and web mining approach where data is exploited and manipulated with no clear emphasis on knowledge or in the best case; the knowledge discovered needs additional explanation from a domain expert. Finally the third approach is completely knowledge-centric, where data is abstracted and generalized in higher levels of application logic. This approach usually focuses on semantic web and the creation of the ontology that describe adaptive web based systems.

The overhead approaches suffer from several drawbacks due to one way treatment of the problem i.e. they are either completely data centric or completely knowledge centric. In [4] we presented a top-down five layer framework that offers flexibility, expressiveness, interoperability and modularity to address the above mentioned problems in the sense of generalization of the approaches. We coined it as top-down because it starts with the data layer at the top and narrows down to the adapted content by taking into consideration other layers that lay in between. The layers are considered in such a way to take into account both data, knowledge (concepts about data) as well as the user. This framework consists of five layers as depicted below in Fig. 1.

The first layer called the data layer is consisted of loosely coupled atomic data units. The units can be texts, multimedia files (audio, video or animations) or any other describable content. These units represent the building blocks of a web page. The second layer represents a layer of concepts and concept relationships gathered from atomic units with a complete ontology designed to function in this layer [5]. The third layer takes care of the information gathered from users that accesses the page. The information acquired here reside on the user information repository, which stores user access on the page’s content in the form of web logs, as well as pattern repository that uses data mining techniques for extracting and accumulating user access behaviour. The adaptation layer consists of an original adaptation algorithm on how to perform adaptation and finally presentation layer rearranges the web site’s content or structure based on the algorithm from adaptation layer.
In the following section we will focus on the work and approaches done on the first four layers of the above depicted framework.

**FROM DATA TO ADAPTATION**

Tackling the problems mentioned in the overhead section requires a separated approach on solving challenges in every layer. In data layer for example, a specific data definition that will match the atomic unit concept has been devised. For this purpose, during the authoring phase we have included a special module that stores such units in a database as well as RDF units which can be reusable. The data has been divided into proper types and subtypes to match the ontology defined in concept layer.

In concept layer on the other hand, a complete ontology for generic adaptive web based systems has been developed that supports the atomic unit idea [5]. The main aim of the ontology is to envision a web site that will be consisted of loosely coupled atomic data units. Based on user preferences, these atomic units can be gathered to create more composite unit (pages) that will be presented in an adapted way. The ontology consists of concept domain and user domain. In concept domain we store atomic units as concepts as well as concept relationships. These concepts have their own information weights (ratings) and can be aggregated into pages. Ratings and weights can be generated in an offline data mining procedure (indexing and similarity matrix generation). Data and concept layer have to be designed in such a way in order to ensure proper linking between concept and content as presented in [4]. The benefit of the designed ontology is twofold: first one is to “glue” data layer, concept layer and user layer by representing them in more abstracted way and the second is the reasoning mechanism that ontology usage offers through semantic web technologies.

In user layer, browsing behaviour and access pattern of the visitors have been analyzed. In this layer, we used more non invasive methods in order not to obstruct the user’s navigational freedom. Therefore, as main source of information gathered from users were taken web servers logs out of which sessions and navigation patterns were extracted. With sessionization, a particular user request is aggregated into a group of session that belonged to the same user while with navigation pattern user behaviour as well as its next probable visit is being obtained and observed.

In adaptation layer, a more structured approach on developing an effective adaptation algorithm has been presented. In [14] several adaption approaches are
analyzed and it is concluded that they suffer from several drawbacks that can be outlined as follows:

- In all the approaches mentioned above there is no clear focus on adaptation formalism and algorithm on how the adaptation is performed?
- Many of the data mining techniques used in overhead contributions focus extensively on user navigation patterns (the importance of which it should not be overlooked) rather than what the documents itself represents and to whom they are related.
- Some of the approaches use definition of rules how and when the adaptation should occur which sometimes it represents a tedious authoring task. Like in [15] where the author decides about relevant information based subjective judgment from the web site administrator, rather than real user preferences. Our proposed approach, as is going to be presented in sections below, completely omits this step.

Therefore, in order to address the above mentioned problems, a complete well defined algorithm based on mathematical formalism has been derived. The overall adaptation process proposed in our approach is consisted of three main constituents as depicted in fig 2.

\[ L = \begin{pmatrix} l_{1,1} & \cdots & l_{1,n} \\ \vdots & \ddots & \vdots \\ l_{n,1} & \cdots & l_{n,n} \end{pmatrix} \]

From this matrix, by using similarity measures between documents, a new matrix can be generated represented as:

\[ S = \begin{pmatrix} s_{1,1} & \cdots & s_{1,n} \\ \vdots & \ddots & \vdots \\ s_{n,1} & \cdots & s_{n,n} \end{pmatrix} \]
For every element of the matrix $L$ it applies that $l_{i,j} = l_{j,i}$ where $i = 1, \ldots, n$ and $j = 1, \ldots, n$. Form the values of indexes $i$ and $j$ it can be retrieved two distinct vectors that comprises documents that are similar to $l_{i,j}$. The two retrieved vectors are $v_1[p, i]$ where $p = 1, \ldots, n$ and $v_2[q, j]$ where $q = 1, \ldots, n$ and $l_{i,j} = l_{j,i} \in L$ and $v_1[p, i], v_2[q, j] \in S$.

Our approach firstly sorts the vectors $v$, which are columns in matrix $S$ and presents them in the form of links or short texts. The complete algorithm is given in the form of a pseudocode depicted as below.

**Algorithm 1:** Pseudocode of the proposed adaptation algorithm

Require: $L[], S[]$

Ensure: $v_1, v_2$

for $i = 1$ to $n$ do
  for $j = 1$ to $n$ do
    print $l_{i,j}$
    for $k = 1$ to $n$ do
      $v_1[k] \leftarrow S_{k,j}$
      $v_2[k] \leftarrow S_{j,k}$
    end for
    sort($v_1, v_2$)
  end for
return $v_1, v_2$ {Return a sorted value from $S$}

The presented pseudocode in algorithm 1, returns a sorted similarity vector for each $l_{i,j}$ values form link matrix $L$. The above proposed algorithm is suitable for both personal adaptation or personalization where similar content related to $l_{i,j}$ is presented to a single user as well as adaptation to a group of users that share the same interest. The latter can be achieved by clustering the similarity matrix for discovering the most interesting groups. The performance of the algorithm on both cases is depicted in two figures below.

**CONCLUSION AND FUTURE WORK**

In this paper, a complete review on work done on first four layers of our top-down framework has been presented. In data layer we presented a compliant data representation that is compatible with other three layers of the framework especially the concept layer. In concept layer a complete ontology for general purpose adaptive web
based systems is presented as well as a new adaptation approach in delivering adaptive content to users or groups of users with different preferences has been introduced. The delivery of such adaptive content is achieved through the use of two auxiliary matrices that are utilized by newly presented adaptation algorithm. The approach presented here uses a straightforward algorithm that retrieves and delivers adaptive content to visitors and omits the use of rules defined by web site administrator which in many cases can be unimportant and subjective to them.

Future work would involve development of system that will revive the idea presented here as well as evaluation of the degree of help that such systems give to web site visitors.

REFERENCES


ABOUT THE AUTHORS

Bujar Raufi, M.Sc, Faculty of Contemporary sciences and Technologies, Department of Computer Science, South East European University, Phone +389 44 356 168, E-mail: b.raufi@seeu.edu.mk.

Prof. Dr. Juliana Georgieva, PhD, Faculty of Computers and Control Systems, Department of Programming and Computer Technologies, Technical University of Sofia, e-mail: july@tu-sofia.bg.