Designing an Application-Aware Routing Protocol for Wireless Sensor Networks

Mohammad Abdul Azim, M. Rubaiyat Kibria, and Abbas Jamalipour†
School of Electrical and Information Engineering
The University of Sydney, NSW 2006, Australia
Email: {azim,rkibria}@ee.usyd.edu.au, †a.jamalipour@ieee.org

Abstract—Routing protocols that can facilitate application-specific service guarantee in wireless sensor networks (WSN) constitute one of the key design objectives of current WSN research. Since energy-efficient routing protocols in literature do not offer a complete framework for service differentiation, a new approach is warranted. Formulating such routing approach requires the adoption of different cost metrics that can parameterize the application-specific requirements. To this effect, this paper proposes an application-aware routing protocol (AARP) that considers battery power, data transaction reliability and end-to-end delay for service differentiation. Two mathematical models namely Analytical Hierarchical Process (AHP) and Grey Relational Analysis (GRA) are incorporated for intermediate node selection (ranking and subsequent selection based on local weight calculation) for data transaction purposes. As demonstrated by the simulation results, the proposed routing approach offers service configurability across a range of applications.

I. INTRODUCTION

Although wireless sensor networks (WSNs) suffer from resource limitations (e.g., limited battery power, short transmission range etc.) and unreliable wireless connectivity [1], the cost effective nature of pervasive node deployment enables WSN to offer data transaction capability over a large area. So far such transaction capability has entailed the development of energy-efficient routing approaches in order to prolong the network lifetime. With the advent of new applications (e.g., environmental monitoring, industrial monitoring, military applications etc.) that exhibit inherent characteristics such as varied traffic pattern, data transaction reliability, and end-to-end delivery delay, conventional approaches (i.e., energy-efficient routing) are no longer capable of service differentiation, thereby warranting the development of newer approaches.

The routing proposal, presented in this paper, attempts to resolve the service differentiation of applications by incorporating the configurability of the abovementioned characteristics (alternatively cost metrics). Here, dynamic configuration of the cost metrics (in relation to each other) relates to different routing priorities for applications that govern the next hop node selection. Obviously such selection reflects the application-specific requirements of QoS guarantee for source-to-sink data transaction. To facilitate the selection process, two mathematical models namely Analytical Hierarchical Process (AHP) and Grey Relational Analysis (GRA) are incorporated. AHP [2] is a multi-criteria based decision-making algorithm that considers different available options or candidates, while GRA [3] selects the optimal option by comparing the similarity between each option and the ideal reference vector (governed by different cost metrics). Considering the QoS of majority of applications, four different cost metrics such as reliability (quantified by packet loss), delay, battery power and distance are incorporated for evaluating the routing path. Although more metrics could have contributed towards a more effective routing mechanism, it would eventuate at the cost of excessive energy expenditure as well as increased complexity (a very undesirable criteria for WSN), thereby justifying the inclusion of the most common WSN QoS parameters. The comparative weights (that dictates the importance of specific requirement or criteria) of the metrics are therefore considered as configurable parameters that can be tuned as per the application requirements. Hence, subsequent to the tuning, the packets are forwarded to the optimal next hop node.

The remainder of this paper is organized as follows. Section II highlights the related works in existing literature. In Section III, a detailed description of the proposed routing approach is presented, while Section IV presents the simulation results, followed by some concluding remarks.

II. RELATED WORKS

Routing decisions that are governed by the application-specific QoS requirements (such as delay, reliability, energy efficiency, and so on) can be referred to as application-awareness. In recent times, we have witnessed a significant amount of work in literature that addresses such awareness issues in WSN. However, for simplicity purposes and to reduce the computational complexity, majority of these works consider the QoS parameters either separately or in a combination of two or three. For example, in reliable information forwarding using multi-path (ReInForM) [4], end-to-end reliability over multi-hop routing is addressed by delivering data through redundant paths. The desired reliability is set according to the information content of the transacted data, whereas load distribution is incorporated through randomization in the next hop node selection. Although the algorithm provides accept-

This work is partly supported by the Australian Research Council. The corresponding author is Mohammad Abdul Azim {azim@ee.usyd.edu.au}. 

†This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.  

978-1-4244-2324-8/08/$25.00 © 2008 IEEE.
able performance regarding reliability and load distribution, it neglects the issues of timeliness (i.e., end-to-end delay) and transmission cost.

The transmission cost and successful delivery probability are in fact addressed in [5] which provides an optimal power allocation and routing algorithm that maximizes the network lifetime and minimizes the heuristic routing cost. Although [5] offers a fair solution, it does not provide any mechanism for ensuring end-to-end delay for data delivery. Instead, such data delivery framework is proposed in [6] that defines a stateless protocol for real-time communication in sensor networks (SPEED). The algorithm attempts to enforce a minimum data delivery time for packets by maintaining a soft real-time delay set and uniform speed. In case the speed cannot be achieved, the maximum allowable delay is evidently not attained, leading to packet dropping. The fact that SPEED is only concerned about the timeliness of data delivery and load distribution without considering the reliability, reduces its efficiency in offering application-specific service differentiation. To redress this limitation, Multipath multi-SPEED (MMSPEED) (an extension of SPEED) is proposed in [7]. MMSPEED offers QoS provisioning in the context of reliability and timeliness by employing probabilistic multi-path forwarding in the reliability domain whereas multiple speeds are adopted to meet the end-to-end data delivery deadlines. Unfortunately, MMSPEED does not consider the transmission cost over multi-path selection which can severely affect the network efficiency. A solution for addressing the transmission cost is presented in [8] and [9]. While the former considers several cost metrics and delivers data over the minimum cost path using Dijkstra’s algorithm [10], the latter presents a packet duplication mechanism where energy and delay are balanced over multi-paths. To this effect, the former can be regarded as the closest counterpart to the algorithm proposed in this paper. However, the centralized approach restricts its deployment to hierarchical networks only. The latter, on the other hand, is not very suitable at handling a dynamic environment as the packet loss rate varies over the time and is not a constant.

The fact that service diversity requires the incorporation of a range of critical QoS parameters such as energy efficiency, end-to-end delay, reliability etc., warrants a routing protocol design that can concurrently offer lifetime extension as well as application-specific service differntiability. To the best of our knowledge (following extensive literature review) no such work has been identified. This not only limits the range of service classes that the network provider can offer as part of the service level agreement (SLA), but it also restricts the QoS guarantee profile under dynamic network conditions (e.g., limited buffer space, variable delay requirements etc.). In an effort to resolve this, a configurable routing protocol is therefore proposed in the next section that utilizes the mathematical techniques of AHP and GRA, and provides concurrent lifetime extension and optimal routing path selection in accordance to the application level service requirements. For a detailed description of the AHP and GRA algorithms, interested readers may refer to [2] and [11].

III. APPLICATION-AWARE ROUTING PROTOCOL (AARP)

The proposed AARP provides a mechanism to derive routing paths that support the QoS requirements of applications while maintaining the energy efficiency of the network. A categorial description of the protocol is given below which is preceded by the set of assumptions considered.

A. Assumptions

The protocol considers three fundamental assumptions: (i) nodes can access their battery level and can adjust their transmit power (currently available in nodes as well as in many existing protocols), (ii) nodes are densely deployed (dense deployment is a common notion and provides robustness to the system) and their location information is known a priori (achievable by GPS or GPS-free location algorithms), and (iii) clocks at nodes are synchronized so as to offer global synchronization (using existing literature) for delay measurements (a valid assumption since the sensed data has no significance without time).

B. Design Objectives

The primary objective of the proposal is to design an energy-efficient and application-aware routing protocol. To address the QoS issue of respective applications, two metrics namely reliability and delay are defined, while two other metrics namely greedy distance (i.e., distance from source to sink) and battery usage are identified for energy-efficiency. A brief description of the design objectives as well as the metric definitions are included below.

Energy Efficiency: It is widely acknowledged that energy efficiency is the most important issue in WSN routing and relates to the lifetime extension of the network. As such, by utilizing shorter paths instead of detours, it is possible to limit the energy usage thereby extending the network lifetime. However, with static sinks (as is commonly considered), repetitive usage of the same path could lead to premature node deaths, necessitating the protocol to address both energy efficiency and load distribution in an effective manner. To redress the former issue, the algorithm adopts a greedy distance (distance between the candidate node and the destination sink) based selection mechanism where nodes having the smallest greedy distance is regarded as the desirable next hop nodes. The selection subsequently brings the routing path closer to the shortest path (invisible line between source and sink) and saves energy (hence the term “energy-efficient”). Fig. 1 illustrates the distance measurements for the candidate nodes in the forwarding path. In terms of fair load distribution, nodes having more remaining battery energy participate more in data forwarding so as to prevent the network from early partitioning.

QoS Support: Two fundamental issues, namely reliability and timeliness, that govern the QoS awareness of WSN applications are identified here.

- Reliability: Reliability here is defined as the ability to deliver data to the destination with minimum packet loss.
For example, applications such as forest fire detection may require that the packets reach the destination or the monitoring station without any loss. Guaranteeing such loss-less data transaction therefore would entail careful routing mechanism with prioritized forwarding. Since unreliability of the wireless link in WSN is largely due to interference and congestion, reliability metric is often considered as reciprocal to packet loss rate (varying over time).

- Delay: Some applications (e.g., a battlefield application reporting a moving target) may have the obligation to deliver data within a constrained delay period since the importance of the information has little impact after certain time period. In WSN, such end-to-end delay measurements are facilitated through timestamps across a priori synchronized network environment, and prefers packet forwarding over least delay paths.

**Combination:** In addition, an application may need a routing protocol that supports a combination of these criterion.

**Configurations:** Since applications can have varying priority levels or weights for cost metrics that govern the delivery of the sensed data, different configurations can be simultaneously deployed to facilitate application-specific data transaction.

### C. Proposed Protocol Operations

Subject to the periodic control message exchange between the neighboring nodes, the proposed AARP builds and maintains a localized routing table, and subsequently utilizes the AHP and GRA algorithms to select the next hop node, as per the application requirements. A categorical description of different steps of the algorithm are briefly presented below.

**Routing Table Maintenance:** In a straightforward manner, nodes collect their location and power metric information, and conveys the information to their single hop neighbors. Since the routing metrics, in particular delay and reliability vary with time, to offer reasonable measurements the metrics are updated quite frequently and can result in considerable overhead. As such, exponential moving average technique is utilized to get a current view of the metrics, and is represented as follows, 

\[ EMA_t = EMA_{t-1} + x \cdot (y_{t-1} - EMA_{t-1}) \]

where \( x \) is the smoothing factor (between 0-10%), \( y \) is the observation, \( t \) is the time, while \( EMA \) denotes the exponential moving average.

**Consistent Configurations:** Following the detection of current metric values, the different configurations for application-specific data delivery is derived. To achieve this, AHP is incorporated where it breaks down a multicriteria-decision making problem into a series of decision factors. Mathematically denoted as a square matrix \( M \), it represents the different configuration of applications as per their QoS requirements. Here each element of \( M \) represents the relative weight ratio of the decision factors with respect to their parent choices, and are graded from 1 to 9 (from equal importance to absolutely more important). As such, different weight ratios correspond to different configuration profiles for different applications. Table I illustrates an instance of such configuration profile. Because of the random selection of the relative weight ratio, \( M \) is often not perfectly consistent and need to be evaluated for consistency (using consistency ratio [2]).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Reliability (r)</th>
<th>Delay (d)</th>
<th>Power (p)</th>
<th>Distance (dist)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability (r)</td>
<td>1</td>
<td>1/(wd_p)</td>
<td>1/(wd_r)</td>
<td>1/(wd_{dist}d)</td>
</tr>
<tr>
<td>Delay (d)</td>
<td>(wd_d)</td>
<td>1</td>
<td>(wd_p)</td>
<td>(wd_{dist}d)</td>
</tr>
<tr>
<td>Power (p)</td>
<td>(wd_r)</td>
<td>1/(wd_p)</td>
<td>1</td>
<td>(wd_{dist}p)</td>
</tr>
<tr>
<td>Distance (dist)</td>
<td>(wd_{dist}d)</td>
<td>1/(wd_{dist}d)</td>
<td>1/(wd_{dist}p)</td>
<td>1</td>
</tr>
</tbody>
</table>

**Forwarding:** Next hop forwarding decision involves selecting the optimal node from the candidate nodes in the forwarding path (as demonstrated in Fig. 1). The optimization is carried out by utilizing the GRA method. In GRA method, the Grey Relational Coefficient (GRC) ranks the available choices by quantifying the similarity of their attributes in emulating the ideal option (reference vector \( E_o \)). Because of their disparity, prior to calculating the GRC, the attributes are normalized. Note that the normalization is carried out under two situations: smaller-preferred and larger-preferred. While smaller preferred is assumed for minimizing the attributes,
larger-preferred governs the maximization effort. These normalized values are then considered in deriving the ideal option \( E_o = \max(r \min(d) \max(p) \min(dist)) \) and the GRC values. GRC values can therefore be calculated as follows [11],

\[
GRC_i = \frac{\Delta_{\text{min}} + \psi \Delta_{\text{max}}}{\Delta_i + \psi \Delta_{\text{max}}}
\]

where \( i \)-candidate options, \( \Delta_{\text{min}} \) and \( \Delta_{\text{max}} \) are the minimum and maximum of all \( \Delta_i \)'s (\( \Delta_i \) represents the absolute difference between the normalized attributes of the \( i \)-th option and \( E_o \)), and \( \psi = [0,1] \) is the distinguishing coefficient to weaken the effect of \( \Delta_{\text{max}} \) if it gets too big. The candidate node with the largest GRC is therefore regarded as the optimal selection. Fig. 2 shows the functional block diagram of the proposed protocol.

IV. SIMULATION RESULTS

The proposed AARP is simulated over a probabilistic wireless network simulator (Prowler) [12]. The simulator considers a WSN testbed where 100 sensors are randomly distributed over an area of 200m x 200m, with a centralized sink node. The wireless interface is modeled as a Rayleigh fading channel with decaying signal strength function and experiencing random disturbances. In fact, the probabilistic link model defined in [13] is set as the channel model whereas CSMA-CA is chosen as the MAC protocol. Note that the deterministic decaying function considers the following parameters: path loss exponent=3, output power= - 105 dBm, frame length=50 bytes and a shadowing standard deviation of 3.8 dB. For generality purposes, \( \psi \) is set to 0.5 in the GRC calculation, while the simulation is carried out for total of 50 seconds (provides adequate service differentiation).

Three types of protocol configurations are considered here representing three types of service functionalities. Denoted as \( C_1 \) and \( C_2 \), the configurations are translated into AHP matrices (as shown below) where relative priorities of cost metrics (as defined in Table I) are incorporated to emulate the service delivery functionality. Note that \( C_1 \) and \( C_2 \) are consistent matrices.

\[
C_1 = \begin{bmatrix}
1 & 0.2 & 0.333 & 0.125 \\
5 & 1 & 2 & 0.333 \\
3 & 0.5 & 1 & 0.5 \\
8 & 3 & 2 & 1
\end{bmatrix} \quad C_2 = \begin{bmatrix}
1 & 0.167 & 0.5 & 0.125 \\
6 & 1 & 5 & 2 \\
2 & 0.2 & 1 & 0.2 \\
8 & 0.5 & 5 & 1
\end{bmatrix}
\]

From the above definition of \( C_1 \) and \( C_2 \), it is evident that \( C_1 \) prioritizes delay, power and distance, while \( C_2 \) prioritizes delay and distance. To that effect, \( C_1 \) corresponds to applications that require uniform energy dissipation and lifetime extension (e.g., military applications) whereas \( C_2 \) corresponds to applications that demand timely delivery guarantee, higher success rate and throughput (e.g., industrial monitoring).

Accordingly, the independent impacts of \( C_1 \) and \( C_2 \) on different service profiles and energy dissipation are illustrated in Fig. 3 and Fig. 4 respectively. The results demonstrate that different configurations are suitable for different kinds of applications and can be configured by the network provider or operator. For example, in cases where the network is resource limited and requires energy efficiency, \( C_1 \) offers the better configuration profile (as can be seen from the energy variance in Fig. 4(b) and lifetime extension in Fig. 4(c)). This however comes at the expense of longer delivery delay (shown in Fig. 3(a)) and less throughput availability (shown in Fig. 3(c)). Conversely, for environments that demand lower end-to-end delay, higher success rate and larger throughput, \( C_2 \) offers the better solution among the two. Similarly, other types of functionalities can also be derived by configuring the AHP matrix. Note that the AHP matrix derivation is subjective in nature which conforms to the application-specific requirements. Obviously the trade-off among the cost metrics will depend on the operator defined service requirements as well as the existing network conditions. Translated into tabular form, these configurations can therefore be stored within the sensor nodes and evaluated in accordance to the varying network conditions and requirements.

V. CONCLUSIONS

One of the most vital issues in WSN is the lifetime of the network and is fundamental to any design and development effort. The proposed protocol in this paper addresses this issue effectively by limiting the energy dissipation as well as promoting load distribution across the sensor bed. Load distribution is enforced by the participation of higher energy nodes in data forwarding process whereas reduced energy dissipation is achieved by incorporating the greedy distance based selection method that ensures fewer number of hops and short distances. Obviously the selection process is governed by the configuration profile of the applications that utilizes relative weights of different QoS metrics to meet the service requirements. Different configuration profiles therefore result in different next hop node selections. In addition, the proposed protocol also exhibits scalability (requires only a local map of neighboring nodes) and low overhead (i.e., lack of ID maintenance, small routing table, and low signalling cost), thereby making the routing approach suitable for long term and large scale deployments.

REFERENCES

Fig. 3. Performance impact on timeliness, reliability and throughput.

(a) Latency.

(b) Success rate.

(c) Throughput.

Fig. 4. Performance impact on various energy measures.

(a) Energy per packet delivery.

(b) Energy variance.

(c) Predicted lifetime.


