A Bayesian-based Energy Aware Routing Algorithm for Mobile WSNs

Sami S. Alwakeel and Agung B. Prasetijo
Department of Computer Engineering
College of Computer and Information Sciences, King Saud University
Riyadh, Kingdom of Saudi Arabia
swakeel@ksu.edu.sa, agungbp@ksu.edu.sa

Abstract— **Energy conservation is crucial in Wireless Sensor Networks (WSNs) for prolonging sensor node's life. This research attempts to gain benefits of Bayesian classifier through development of a Bayesian Classifier-Based Energy Aware Routing algorithm (BBEAR) for mobile WSNs. Through implementing a Bayesian classifier for routing process, a superior routing algorithm is designed with an intelligent node message rebroadcast decision. This paper discusses the proposed algorithm design and presents a comparison of its performance with flooding technique. BBEAR workflow is described, along with its classifier mechanism that allows the message routing judgment by each sensor node to be based on several node's and message's classifying attributes. In contrast to other routing schemes developed for WSN, BBEAR algorithm incorporates several node attributes, including: remaining energy that the nodes have, distance to the sink, number of messages being queued, node's success rate, delay of the received message, as well as the number of message duplicates received by the node. The BBEAR performance is also discussed and evaluated in terms of various measures such as: throughput, number of duplicates received, average delay, and the lifetime of the nodes. The simulation study of the proposed algorithm shows its superiority over flooding: as it provides 15% gain in message throughput, four times reduction of duplicates, and gives an extension of 60% nodes' lifetime while still maintaining the delay and delay jitter at low levels.**

Keywords— ***Wireless Sensor Network Routing; Quality of Service in WSN; Sensor network Energy Conservation; Bayesian Classifier in WSN***

I. INTRODUCTION

Ad-Hoc networks such as VANETs (Vehicular Ad-hoc Networks) and WSNs (Wireless Sensor Networks) route messages to destination based on a broadcast scheme. However, broadcasting messages blindly leads to broadcast storm, where massive redundant messages are generated by neighboring nodes (i.e. sensor/vehicle) and will in turn drain nodes' energy very quickly. Redundancy, contention and collision will also occur over the network broadcasting channels. The broadcast storm itself is highly likely to result severe congestion in the limited bandwidth WSNs. Therefore, it is a research challenge to find alternatives to WSN flood broadcasting. Redundancy can be minimized by applying P-Persistent approach where a probability of re-broadcast is set to every node. As a result, only prospective nodes with higher probability will conduct the rebroadcast [1]. BBEAR addresses those redundancy, contention and collision issues by introducing a small variable waiting (back-off) time, which is a function of the node Bayesian classifier likelihood that a node rebroadcasts in a mobile WSNs. Contention and collision can therefore be minimized, with a small penalty of introducing little delay prior to rebroadcasting. More details will be given in next sections. The remainder of this paper is organized as the following: Section 2 provides related researches in energy aware routings, and Section 3 presents the BBEAR design. Section 4 discusses the simulation model and experiment. The performance results and conclusions are covered in Section 5.

II. RESEARCH LITERATURE ON ENERGY AWARE ROUTING FOR WSNS

According to recent literature, there are two types of WSNs: event driven and continuous dissemination network according to whether messages are re-transmitted when sensing events occurred or according to pre-defined time schedule [2]. For both types, a mobile sensor network, cannot implement a static routing as the network topology is dynamic and changes continuously. Efficient routing algorithms using redundancy in routing paths are essential for increasing confidence that messages are delivered to the sink [3]. Most of the routing protocols discussed in current literature fall into one of the following: data-centric, hierarchical or location-based protocols [4]. Data-centric protocol is when routing is based on sink request for a specific data. Node data is labeled in order to allow network nodes to identify the data required by the sink. For instance, the sink may query nodes in certain regions to send information. Negotiation between nodes in such regions will reduce message redundancy. Two famous routing protocols using this data-centric mechanism namely: SPIN [5] and Directed-Diffusion [6]. Hierarchical protocols deal with nodes clustering. Efficiency is obtained from data aggregation and the selection of nodes within a cluster to be responsible for broadcasting. Protocols like LEACH [7] and PEGASIS [8] are examples of this protocol class. Location-based protocols employs geographic position information to forward information closer to the sink. Nodes that are located at a relatively longer distance from the sink, when compared to the sender-to-sink distance, are not allowed to forward the received messages. Location-based protocols like GAF [9] and GEAR [10] are normally equipped with GPS device.

Artificial Intelligence has also been employed in research on WSN routing. For example, Barbancho et al. [11] introduced an Artificial Neural Network on every node with selected QoS parameters: latency, throughput, error rate, and duty cycle. N. Kumar et al [12] introduced the Neural Network for selecting cluster heads as well as routing path selection based on node's energy and delivery ratio in a stationary WSN. K. Flouri and P Tsakalides [13] introduced Support Vector Machine (SVM) to conduct clustering. Arroyo-Valles et al. [14] used Bayesian decision for selecting rebroadcast nodes based on node distance to the sink.

Most of the protocols are designed for immobile (static) WSNs. Hence, for mobile nodes, they may face routing failure. Furthermore, whilst past research mostly rely only on limited parameters (remaining energy or node distance from the sink), our research uses many other parameters to reflect node's current status (e.g number of messages being queued, node's success rate, delay of the received message). This allows the re-broadcasting intelligent decision which reflects the overall instantaneous dynamic nature of the WSN and improves its performance.

III. Bayesian-based Energy Aware Routing Algorithm Design

BBEAR can be implemented in both of event driven or the continuous dissemination networks. Furthermore, it is implemented in mobile sensor nodes so that the routing mechanism is more challenging. To make routing forwarding decision, Bayesian classifier evaluates six attributes probabilities that impact the node re-broadcasting decision in WSNs:

1. *remaining energy*. The node's remaining energy is the most important. Having a lower energy level means that the node must be more selective when forwarding its messages.
2. *distance to the sink/destination*. When the distance to the sink is relatively short, the nodes have higher probability to rebroadcast messages.
3. *Number of messages queued in memory*. If the number of messages are high, the node's probability for rebroadcast is also high.
4. *number of duplicate message received*. If the number of message duplicates are high, the nodes will have lower probabilities for rebroadcast, so this will minimize other duplicates generated over the network.
5. *message delay*. If the messages delay time from the instant its generated by the sender to moment it is received is high, the nodes then has a higher like hood to re-broadcast and to deliver the messages. This will result a reduction of the overall total delay.
6. *node's success rate*. If the nodes have high success rate to reach the sink, it deserves high probability for rebroadcast messages.

At any moment, whenever a routing decision is to be made, the node's current status which belongs to a combination of above, is fed to the Bayesian classifier to estimate the joint probability of the node status. The estimated value is then translated to the node's waiting period before re-broadcasting.

The workflow of the proposed BBEAR algorithm is as the follows: When a message arrives, the node will check whether the message is new. If it is, the node classifier then estimates the Bayes probability value based on its current attribute's status for both: the node (remaining energy, distance to the sink, messages being queued, and past success rate delivering messages) and the message delay. The higher the value of the estimated probability, the shorter will be the waiting time to rebroadcast. If there are no duplicates of the same message received during the waiting period and if the waiting time counter reaches zero, the node will immediately rebroadcast the message to its neighbors. On the other hand, if during the waiting period, a duplicate of the queued message is received, the node classifier then recalculate again the probability and renew the waiting time for the respective message. If the number of duplicates reception occurred a given number of times, then the node discards the message. The assumption that many duplicates generated over the network, guarantee that the sink has received the broadcasted message.

Message acknowledgment is also introduced to BBEAR design. The sink will directly send an acknowledgment to every node immediately after it receives the message. Nodes that receive the acknowledgment must discard the associated message queued in the memory. With the acknowledgment introduction, redundancy of messages can be significantly reduced. The success rate of the node delivering messages is calculated as: the number of node acknowledged messages divided by the total number of broadcasted node messages. Message delay is the time difference between the arrival time at the node and the generation time-stamp of the message by the source node. Distance from the sink is the Euclidean distance, and we assume that every node and the sink are equipped with a Geographic Positioning System (GPS) device.

Bayesian classification needs well defined evidence to judge a new evidence that is not available in history. As the energy availability is the most important issue in WSN networks, the node routing classifier considers it as its major classification attribute. In BBEAR algorithm, to build evidence, the node's remaining energy is divided into two classes: high energy and low energy. When nodes are low in their energy, they must save the energy by delaying the node re-broadcast of the received messages and vice versa. Other attributes follow a similar pattern. The nodes must not rebroadcast messages received when node's conditions are not supportive: e.g. if distance from the sink is considered too long, messages in the queue are small, large number of message duplicates are received, low delay exists for message received, and small node's success rate delivering messages is calculated. In contrast, when all the other parameters are supportive for a rebroadcast, the nodes are set to rebroadcast messages. Table 1 below provides the summary of various attributes possible combinations and the expected forwarding decision.

Table 1. The main and derived evidence for various combinations of BBEAR attributes

Evidence no	Energy remaining	Distance fr dest	Messages in queue	Message dups	Message delay	Success rate	Re-Broadcast?
1	**L**	**H**	**L**	**H**	**L**	**L**	**No**
2	L	L	L	H	L	L	No
3	L	H	H	H	L	L	No
4	L	H	L	L	L	L	No
5	L	H	L	H	H	L	No
6	L	H	L	H	L	H	No
7	**L**	**L**	**H**	**L**	**H**	**H**	**Yes**
8	L	H	H	L	H	H	Yes
9	L	L	L	L	H	H	Yes
10	L	L	H	H	H	H	Yes
11	L	L	H	L	L	H	Yes
12	L	L	H	L	H	L	Yes
13	**H**	**L**	**H**	**L**	**H**	**H**	**Yes**
14	H	H	H	L	H	H	Yes
15	H	L	L	L	H	H	Yes
16	H	L	H	H	H	H	Yes
17	H	L	H	L	L	H	Yes
18	H	L	H	L	H	L	Yes

For those attributes, empirical probability values for re-broadcast are defined. For the 18 rebroadcast evidence in the table, the conditional probabilities of all attributes can be set straight forward. Table 2 shows such assumed conditional probabilities for having rebroadcast or not.

Table 2. Probabilities of classifier input status settings

Status	Remaining energy		Distance from dest		Messages in queue		Message dups		Message delay		Success rate	
	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
Low	1/2	1	5/6	1/6	1/6	5/6	5/6	1/6	1/6	5/6	1/6	5/6
High	1/2	0	1/6	5/6	5/6	1/6	1/6	5/6	5/6	1/6	5/6	1/6

To clarify the above concepts consider the following sample for search to a new re-broadcast evidence, which is not listed in Table 1. Let the node have conditions as follows: *Energy Low, Distance High, Messages queued High, Duplicates Low, Delay High, Success rate Low*. So the conditional probability of the evidence (e.g waiting time to broadcast) is calculated as follows.

Pr[*Yes*|*Evidence*]
= Pr[*Enrg Low*|*Yes*] . Pr[*Dist High*|*Yes*] . Pr[*Queue High*|*Yes*] . Pr[D*up Low*|*Yes*] . Pr[*Delay High*|*Yes*] . Pr[*Success Low*|*Yes*]. Pr[*Yes*] / Pr[*Evidence*]
= 1/2 . 1/6 . 5/6 . 5/6 . 5/6 . 1/6 . 2/3 / Pr[*Evidence*]
= 0.0053584

Pr[*No*|*Evidence*]
= Pr[*Enrg Low*|*No*] . Pr[*Dist High*|*No*] . Pr[*Queue High*|*No*] . Pr[*Dup Low*|*No*] . Pr[*Delay High*|*No*] . Pr[*Success Low*|*No*] . Pr[*No*] / Pr[*Evidence*]
= 1 . 5/6 . 1/6 . 1/6 . 1/6 . 5/6 . 1/3 / Pr[*Evidence*]
= 0.0010717

Note that as Pr[*Evidence*] is the marginal likelihood or model evidence, its value is the same for all possible hypotheses, so it is omitted from calculation. Normalizing the value will result the following:

Pr[*Yes*|*Evidence*] = 0.0053584/(0.0053584+0.0010717)
= 0.8333307

Pr[*No*|*Evidence*] = 0.0010717/(0.0053584+0.0010717)
= 0.1666693

Therefore, for this sample, Bayesian classifier gives the decision that the node has a higher probability to rebroadcast messages.

IV. BBEAR Simulation Experiment

Simulation Experiments are carried with NS-2 Network Simulator, and each simulation trial was run for 1500 seconds. During simulation, messages are exponentially generated from all nodes with a rate of one message per second as long as the node has a sufficient energy level. The nodes are set to have 10 message memory buffer to store messages during the rebroadcast waiting period. Number of duplicates allowed is set to 20. When the number of duplicate for a given message exceeds this, the node will discard that particular message from the memory. The simulation experiment also assumes that there is an omni-directional antenna in each node with a circular coverage radius of 250 meters under IEEE 802.11 environment. Whenever a message is received by a given node, simulator collects the node attributes conditions and feed it to the Bayesian classifier to estimate the node probability to rebroadcast messages. This probability value is translated into node's rebroadcast waiting time. A highly probable rebroadcasting node (i.e higher Pr[*Yes*|*Evidence*]) – correlates to lower waiting time. Conversely, a node having fewer probabilities for a rebroadcast will be set to a longer waiting time. Our simulation experiments are carried with the waiting time in the range of 0 to 200 ms. The sink is located stationary at the middle of 1500 by 1500 network dimension area. 100 sensors are scattered and set to move randomly. Network performance measures are measured at the sink and include:

message throughput, number of duplicates, total delay, and delay-jitter, while other node performance measures are gathered from the node's status. These include congestion levels and nodes' lifetime the First Node to Die (FND), Half Node to Die (HND) and the Last Node to Die (LND). The node's low energy level is defined when the node has 30% of its initial energy or less. Node's distances from the sink, message queue, and message duplicates are similarly estimated at the 30% value or less of their respective maximum. The node's success rate for delivering the message is considered having a high level when 60% or more of the total message passed by the respective nodes are successfully delivered to the sink.

V. Results and Discussions

Figure 1 shows the performance comparisons between BBEAR and flooding technique in terms of throughput, number of duplicates, and congestion. It is shown that BBEAR performs superior over flooding technique as it gives 15% more message throughput and reduces message duplicates by almost four times from 852.6% to only 215.5%. The throughput performance of BBEAR is superior over all simulated network densities over the range of 60 to 160 nodes. Congestion has also been suppressed 28 folds, from 257.4% to only 9.13%. Note that congestion measure here is defined as the number of duplicates that arrived at a node 20 times or more.

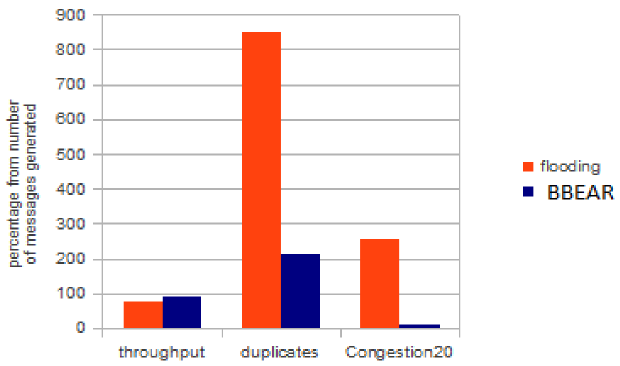

Fig. 1. Performance comparison between BBEAR and flooding scheme

It is shown that in terms of delay and delay jitter, BBEAR maintains both values at low level and relatively steady while flooding shows higher delay and delay jitter as the network density increases.

Experimenting with 100 nodes, BBEAR also extends the life time by about 60%, from 715.5 seconds to 1160 seconds for half of the nodes to be alive (HND). This HND and the other two measures – FND and LND – of the lifetime of the nodes within the network (shown in Figure 4) give us a clear understanding the nodes energy saving performance. It is also the reason why BBEAR gives more throughput over flooding. More dead nodes means having less dense living nodes and the radio coverage then cannot again reach those living neighbors.

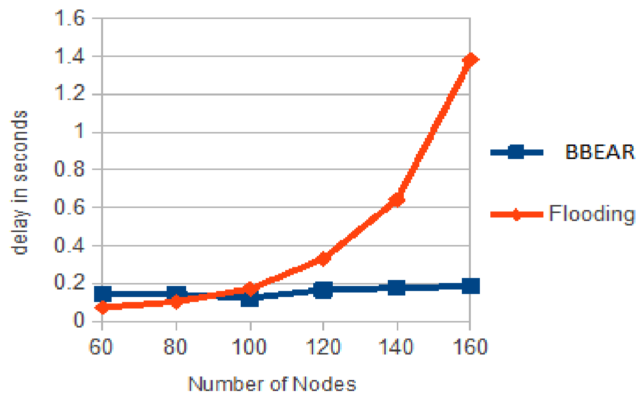

Fig. 2. Performance comparison of BBEAR over flooding scheme in terms of message delay for different network densities.

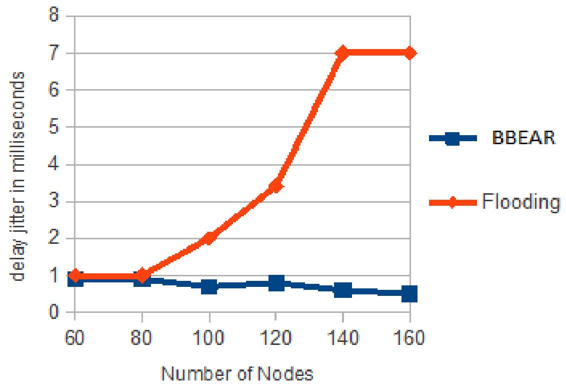

Fig. 3. Performance comparison of BBEAR over flooding scheme in terms of delay-jitter for different network densities

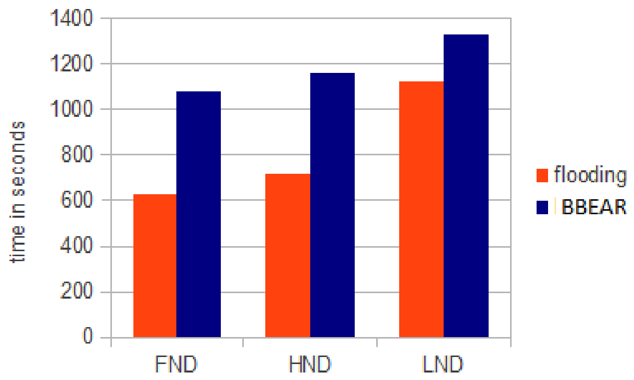

Fig. 4. Performance comparison of BBEAR over flooding scheme in terms of node's lifetime.

VI. CONCLUSIONS

Results suggest that BBEAR shows its superiority over flooding technique for all performance measures. The node's lifetime is roughly prolonged by 60%. An additional gain of 15% message throughput is also achieved. Besides, message duplicates and congestions are heavily suppressed as the number of duplicates is reduced to one fourth when compared to the flooding scheme. Average delay and delay jitter are maintained low. With this in mind, BBEAR performance is clearly superior in all performance measures mentioned above.

ACKNOWLEDGMENT

This project work is supported by the Research Center for Sensor Network and Cellular Systems, University of Tabuk, Tabuk – Kingdom of Saudi Arabia.

REFERENCES

[1] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "The Broadcast Storm Problem in a Mobile Ad Hoc Network," in *Proceedings of The 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking - MobiCom'99*, 1999, pp. 151–162

[2] L.B. Ruiz, A.A.F, Loureiro, and J.M. Nogueira, "Functional and information models for the MANNA architecture," in *Colloquia Francophone sur la Gestation de Reseaux et de Services*, pages 455–470, February 2003

[3] F. Stann and J. Heidemann, "Bard: Bayesian-assisted resource discovery in sensor networks," INFOCOM 2005. in *24th Annual Joint Conference of the IEEE Computer and Communication Societies. Proceedings of the IEEE*, vol. 2, pp. 866–877, 2005

[4] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks*," Journal of Ad Hoc Networks*, vol. 3, issue 3, pp. 325-349, 2005

[5] W. Heinzelman, J. Kulik, and H. Balakrishnan, Adaptive protocols for information dissemination in *wireless sensor networks, MobiCom'99*, 1999

[6] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *MobiCom'00*, 2000

[7] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy efficient communication protocol for wireless sensor networks," in *Hawaii International Conference System Sciences*, 2000

[8] S. Lindsey, and C.S. Raghavendra, PEGASIS: Power Efficient Gathering in *Sensor Information Systems, IEEE Aerospace Conference*, 2002

[9] Y. Xu, J. Heidemann, and D. Estrin, "Geography Informed Energy Conservation for Ad hoc Routing," in *MobiCom'01*, 2001

[10] Y. Yu, D. Estrin, and R. Govindan, GEAR: A Recursive Data Dissemination Protocol for Wireless Sensor Networks, UCLA Computer Science Department Technical Report, 2001

[11] J. Barbancho, C. León, F.J. Molina, and A. Barbancho, "Using artificial intelligence in routing schemes for wireless networks", Journal of Computer Communications, vol. 30, pp. 2802-2811, 2007

[12] N. Kumar, M. Kumar, and R.B. Patel, "Neural Network Based Energy Efficient Clustering and Routing in Wireless Sensor Networks," First International Conference on Networks and Communications, pp. 34-39, 2009

[13] K. Flouri and P Tsakalides, "Training a SVM-based Classifier in Distributed Sensor Networks," in *14th European on Signal Processing Conference*, 2006

[14] R. Arroyo-Valles, A.G. Marques, J.J. Vinagre-Diaz, and Jesus Cid-Sueiro, "A Bayesian Decision Model for Intelligent Routing in Sensor Networks," in *3rd International Symposium on Wireless Communication Systems*, pp. 103-107, 2006