DATA COLLECTION ABOUT RARE PLANT SPECIES BY WIRELESS SENSOR NETWORKS THROUGH CONGESTION AWARE ROUTING

M. Masumuzzaman Bhuiyan

Department of Computer Science and Mathematics, Bangladesh Agricultural University
Mymensingh-2202, Bangladesh

ABSTRACT

Agricultural monitoring is an important application of wireless sensor networks (WSNs) where energy-efficiency and reliability are two essential performance measures. The extended lifetime of a network is of limited use if the acquired reliability does not meet the application requirement. In order to improve the reliability, it is necessary to reduce congestion because of the inherent correlation between reliability and congestion. Existing Congestion control algorithms, found in WSN literature, are reactive. After formation and detection of congestion, these algorithms try to reduce the congestion. In this paper, we present a congestion aware routing protocol that avoids congestion formation proactively and as a result, it improves data delivery success rate in WSNs deployed to gather agricultural data. Appropriate routes are selected based on location, energy and congestion information of neighbours as well as based on the location information of the base station. Experimental results show that the proposed method is energy-efficient and at the same time, it achieves high packet success.

Key Words: Agricultural data, Reliability, Congestion

INTRODUCTION

With the recent growing advancement of Micro-Electro-Mechanical System (MEMS), wireless sensor networks (WSNs) (Pottie, 1998) have attracted increasing attention. A typical WSN has little or no infrastructure and it consists of few tens to thousands of sensor nodes working together for a specific task. These sensors are tiny in their size and they are run by built-in batteries. As WSNs are typically deployed in inaccessible areas, it is difficult to replace these batteries after depletion. Therefore, existing research works have focused mostly on energy issue (Akyildiz *et al.*, 2002). But reliability remains as one of the vital issues in WSN. In this paper, we considered applications where WSNs are set up to provide data about the area where endangered plant species grow as well as about neighbouring areas where the plants are absent. Definitely, we want to get as many data as possible from all sensors for analysis and determination of conditions suitable to these plants. Here, the reliability (end-to-end success rate) of data transmitted by nodes is equally important as the energy efficiency. Again, there is an inherent correlation between data success rate and congestion in the network. Therefore, we need an effective congestion avoidance technique to achieve the required level of success rate.

A congestion-aware routing protocol is designed which works with data gathering WSNs where the base station (BS) periodically receives data from the sensor nodes. Lightly loaded nodes are chosen by all nodes during data forwarding towards the BS in order to avoid congestion and to increase transmission reliability thereby. In this paper, packet success rate improvement without sacrificing much on energy efficiency is our primary objective.

Existing congestion-control schemes (Wang *et al.*, 2006) designed to work in WSNs mainly mitigate congestion after its formation. CODA (Wan *et al.*, 2003) is a well known congestion-control technique. A congested sensor node broadcasts a back pressure message to the upstream source node. The upstream node reduces its data rate after receiving the backpressure message, and the node also propagates the message further based on its own congestion condition. An exact rate adjustment algorithm is applied to mitigate congestion in CCF (Ee and Bajcsy, 2004). In Fusion (Hull *et al.*, 2004), neighbouring nodes stop forwarding packets to a congested node as soon as congestion is detected and notified. Siphon (Wan *et al.*, 2005) redirected traffic to virtual sinks (VSs) when nodes become congested. These VSs have long-range radios that are used as a shortcut or *siphon* to mitigate congestion. If two or more data-carrying paths intersect at an intermediate node then that node may become congested. TARA (Kang *et al.*, 2007) utilized a detour path from this intersection node to the BS in order to reduce the load of the already congested path. In priority based congestion control protocol for wireless multimedia sensor networks (Zhou *et al.*, 2008), a node determines its data forwarding rate based on its congestion level and within this rate, the node forwards different traffic streams according to their priorities. Without trying to avoid congestion initially, all above schemes mitigate congestion after its formation. Data success rate is reduced because of the dropped packets until congestion is mitigated. Congestion avoidance reduces such packet drops.

Few works in literature tried to avoid congestion. I2MR (Moghaddam and Adjeroh, 2008) tried to avoid congestion by minimizing interference. A set of physically separated paths with minimum interference is chosen by each node in order to ensure load balancing. The source node finds a new set of paths when existing paths face congestion during data transmission. But I2MR is not suitable for data gathering WSNs where all sensor nodes periodically generate small amount of data. This scheme is applicable where few nodes generate most of the traffic in the network. In reliable transport with memory consideration (RTMC) (Teo *et al.*, 2008), a node forwards a packet to another node if the receiver node has enough free memory to hold the incoming packet. Otherwise, it would wait until it gets a node with free memory. Although, the authors claimed to have congestion control, actually they avoided congestion without considering delay in data transfer. CAR (Kumar *et al.*, 2008) provided privileged routing to high priority data in the presence of low priority data. This scheme also is not suitable for data gathering WSN because of same priority level of all transmitted data. It is basically suitable to event detection system. It ensures delivery of data originated from few critical nodes without fair delivery of data generated by other nodes. Moreover, CAR never detects congestion. In this paper, our aim is to utilize congestion information of nodes during route selection

in order to avoid severe congestion which would lead to increased delivery of agricultural data sent by the nodes.

MATERIALS AND METHODS

In this section, parameters that accurately reflect congestion status of a node are determined first. Then analytical measurement of end-to-end success rates of some distance-based data forwarding techniques is computed. Finally, the proposed protocol is described that achieves high end-to-end success rate through congestion avoidance and thus improves reliability.

Congestion parameters

MAC layer buffer occupancy (BO) is a measurement of node level congestion. If the packet reception rate of MAC layer is higher than its packet sending rate then its buffer occupancy increases and eventually makes the node congested. Again, if all outgoing neighbours of a node are congested then the node incurs huge number of collisions during data transmission. That is why the Average Number of Transmission Attempts per Successful Transmission (ANTAST) reflects the degree of link level congestion. We consider BO and ANTAST parameters to perceive congestion level of a node. We also consider the distances of the neighboring nodes during route selection process. The next subsection shows that this distance factor ensures high packet success rate.

Distance-based routing schemes

We assume that each node knows its location as well as the locations of its neighbours and the BS. A node can forward its data to any neighbour that is within its transmission range and lies on the way towards the BS. We are interested to achieve high end-to-end success rate while not consuming much energy. In this subsection, we first compare success rates of three possible distance-based routing protocols: SMHR (Shortest Multi-Hop Routing), LMHR (Longest Multi-Hop Routing) and RMHR (Random Multi-Hop Routing). Each Routing considers sensor nodes with a fixed transmission power. In SMHR, when a node forwards a packet towards the BS, it chooses its closest neighbour on the way towards the BS. In LMHR, the node chooses the farthest neighbour and in RMHR, it chooses any neighbour randomly on the way towards the BS.

Let us assume that all sensor nodes are uniformly distributed with a node-density of λ^2 per unit area. On average, along a line of unit length, there will be λ number of nodes and two successive nodes will be $1/\lambda$ unit distance apart from each other. The transmission power of each node is fixed and can successfully reach another node that is at most x unit distance away from it. For simplicity, we assume $x = d/\lambda$. It means that when a node A transmits data, the farthest node B that can receive that data is the dth node from A along the fictitious line connecting nodes A and B. Therefore, when a node transmits data, any other node which is located inside the circle centred at the transmitter node with a radius d/λ can receive that data. In Fig. 1.a, node S sends its data directly to node R which is the

i^{th} node on the way from node S to the BS. The distance between S and R is i/λ where the maximum value of i may be d. Circles C_S and C_R are centred at nodes S and R respectively and the set N_S contains any node residing inside of C_S or C_R or both. When S sends data to R, any member node of N_S can cause collision to this transmission by sending data at the same time. Now

$$|N_S| = 4 \times Area\ of\ shaded\ region\ \times \lambda^2$$

$$= d^2(\pi + 2\sin^{-1}\frac{i}{2d}) + \frac{i}{2}\sqrt{4d^2 - i^2}$$
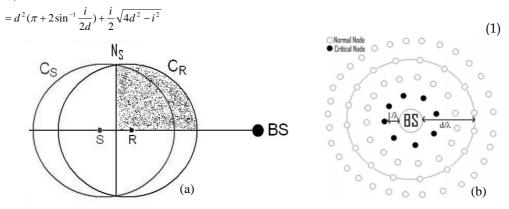
(1)

Fig. 1. (a) Node S sends data towards the BS through node R, (b) Critical nodes

Let us assume that the average probability of data transmission by a node is p ($0 < p < 1$). The probability that any node other than S in N_S does not send data when S sends its data to R is $(1-p)^{|N_S|}$. For simplicity, here we use $|N_S|$ instead of $|N_S|$-1. If S_i^h denotes the hop-by-hop success rate when data is directly sent to the ith node from the sender then $S_i^h = (1-p)^{|N_S|}$. If the BS is n-nodes away form the sender and S_i^e denotes the end-to-end success rate when every node forwards data to the ith node on the way towards the BS, then

$$S_i^e = (S_i^h)^{n/i} = (1-p)^{|N_s| \times \frac{n}{i}} = (1-p)^{f(i)} \text{ , here}$$

$$f(i) = |N_S| \times \frac{n}{i} = \frac{nd^2}{i}(\pi + 2\sin^{-1}\frac{i}{2d}) + \frac{n}{2}\sqrt{4d^2 - i^2}$$

(2)

We can write $f(i) = \phi(i) + \varphi(i)$ where $\phi(i) = \frac{nd^2}{i}(\pi + 2\sin^{-1}\frac{i}{2d})$ and $\varphi(i) = \frac{n}{2}\sqrt{4d^2 - i^2}$

Similarly, $S_{i+1}^e = (1-p)^{f(i+1)}$. It can be shown that $\phi(i) > \phi(i+1)$ and $\varphi(i) > \varphi(i+1)$ for $1 \leq i \leq d$. This means that $f(i) > f(i+1)$ and therefore, $(1-p)^{f(i)} < (1-p)^{f(i+1)}$ which leads to $S_{i+1}^e > S_i^e$ for $1 \leq i \leq d$. From this analytical finding, we can conclude that the data success rate at the BS will be the highest in case of LMHR as it forwards its data to the farthest neighbour and similarly, data success rate will be lowest in case of SMHR.

In addition to collisions, errors are introduced during signal propagation. In this paper, we consider ideal environment (i.e., no obstacle) where only path loss causes propagation error. As path loss increases with the increasing distance between the sender and the receiver, the probability of error due to path loss would be the highest in case of LMHR and the lowest in case of SMHR. Again, received signal power depends on path loss and transmitted signal power. But when any node S sends data directly to another node R which is d-nodes away from it, it means that the transmitted power is sufficient to be successfully received by R. Therefore, the effect of path loss on the success rate at the BS in these three models will not be very high. So, if we consider the probability of error due to path loss and collision, then the packet success rate at the BS should be the highest in LMHR and the lowest in SMHR.

To compare energy dissipation of sensor nodes in these protocols, we consider the maximum energy spent by any node in the network. All data generated by all nodes in SMHR have to be passed through the nodes within $1/\lambda$ unit radius of the BS (Fig. 1.b, we call them critical nodes). Highest amount (not all) of data have to be passed through critical nodes in LMHR with $d > 1$; here all nodes within d/λ unit radius of the BS will send their data directly to the BS by bypassing the critical nodes. As each node forwards its data to other nodes randomly in RMHR, all data will be uniformly distributed among all nodes within d/λ unit radius of the BS. Therefore, maximum energy spent by any of the critical nodes should be the highest in SMHR and the lowest in RMHR.

Proposed routing protocol

The above analytical study shows that LMHR has better performance over other two models in terms of packet success rate but argument suggests that RMHR is better in terms of energy conservation of critical nodes near the BS. But all these models have no congestion aware component. To achieve high packet success rate, a node needs to choose the farthest node towards the BS during route selection and at the same time, it needs to reduce high energy consumption of critical nodes by choosing random routes. To achieve that we introduce a utility function f in (3) that would be applied to each neighbour of a transmitter node B. When B forwards a packet, it chooses the highest f-valued node among its neighbours those have a certain level of remaining energy. The function f has three components: (i) "distance of next node" for ensuring high packet success rate, (ii) Buffer Occupancy (BO) for reducing node level congestion and (iii) Average Number of transmission Attempts per Successful Transmission (ANTAST) for reducing link level congestion.

$$f_k = \alpha \times \frac{D_k}{D} + \beta \times (1 - \frac{BO_k}{Buff\_Size}) + \theta \times (1 - \frac{ANTAST_k}{Max\_ANTAST})$$

$$(3)$$

where D_k is the distance of next node k towards the BS from node B, D is the maximum distance that can be covered by the transmission power of each node, BO_k is the buffer occupancy of node k, $ANTAST_k$ is the average number of transmission attempts by node k per successful transmission, *Buff_Size* is the size of the buffer of each node, *Max_ANTAST*

is an estimated maximum value of *ANTAST* (which needs not to be accurate) and α + β + θ = 1. The last two terms in *f* help to reduce congestion formation by avoiding congested nodes. Moreover, during the choice of the next route, all three components of *f* together ensure randomness based on the surrounding congestion condition of the node. High energy dissipation of critical nodes is reduced by this randomness (as it is the case in RMHR). The algorithm used in the proposed protocol has been shown in Algorithm I.

To calculate *f* for each neighbour *k*, node *B* needs the values of D_k, BO_k and $ANTAST_k$. We assume static network and each node knows its location. A node *k* can broadcast its location, energy information, BO and ANTAST after receiving a fixed number of packets from other nodes or after a fixed interval whichever is earlier. To avoid creating oscillations, node *k* sends exponential averages of BO and ANTAST. The fact that we are only considering neighbours that have adequate remaining energy would enhance network lifetime.

Algorithm I. The proposed protocol

Procedure the proposed_ protocol (Neighbour-list *NL*, Distance- list *DL*, Buffer-list *BL*, TransmissionAttempt-list *TAL*)

local variables: Node *R*, Node *k*, real *f*

retain only 5 nodes in *NL* that have the highest remaining energy

for each node *k* in *NL*

$$f = \alpha \times \frac{DL(k)}{D} + \beta \times (1 - \frac{BL(k)}{Buff\_Size}) + \theta \times (1 - \frac{TAL(k)}{Max\_ANAST})$$

Let, node *R* has the highest value of *f* among all nodes in *NL*
return *R*

RESULTS AND DISCUSSION

The reliability of data transmission using the proposed protocol scheme was validated in a simulation using NS-2 (NS-2 website, 2006). The proposed protocol is not a congestion-mitigation scheme but is purely a routing protocol that uses MAC layer feedback. The results were compared with similar routing protocols like AODVR (Perkins and Royer, 1999) and SELAR (Lukachan and Labrador, 2004). The proposed routing is based on distances to neighbouring nodes. Similarly, the routing in AODVR is also based on distance. Routing scheme SELAR is location aware and uses the energy information of neighbouring nodes. The proposed protocol is also location aware and along with other information it also considers the energy levels of neighbouring nodes. Therefore, comparison with AODVR and SELAR assesses the proposed protocol with competing protocols in the same layer of protocol stack. The proposed protocol was not compared to

congestion avoidance techniques like I2MR (Moghaddam and Adjeroh, 2008) and CAR (Kumar *et al*., 2008) because they are not suitable to agricultural data gathering network.

We placed 196 sensor nodes in an area of 425 metre × 425 metre with uniform distribution. Each sensor node has a fixed transmission range of 80 metres. We employed 1 Mbps IEEE 802.11 distributed coordination function (DCF) MAC protocol. All sensor nodes periodically generated and forwarded data towards the BS. In simulation, we set α = 0.50, β = 0.25 and θ = 0.25 for the proposed protocol. Fig. 2 compares packet success rate of the proposed protocol, AODVR and SELAR where Fig. 3 shows maximum energy dissipation of critical nodes.
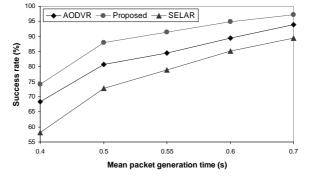

Fig. 2. Success rates of the proposed protocol, AODVR and
SELAR with different packet generation rates

The proposed protocol performs better than AODVR and SELAR in terms of packet success rate as shown in Fig. 2. AODVR chooses only the shortest route towards the BS. The proposed protocol chooses shortest but lightly loaded route. Thus the proposed protocol yields higher success rate as packet drop is reduced by congestion avoidance. On the other hand, SELAR attempts to make the energy consumption of all nodes uniform during route selection by choosing the neighbour with the highest energy. Since SELAR primarily targets at higher network lifetime by lowering energy consumption without any concern of packet delivery chance in chosen route, it demonstrates lower success rate than the proposed protocol. For different data generation rates, the proposed protocol yields success rate 11.5%-26.5% higher than SELAR and 4.3%-8.3% higher than AODVR.

An interesting finding observed in Fig. 2 is that AODVR has higher success rate than SELAR. We have analytically proved that the highest end-to-end success rate can be achieved if maximum distant nodes are chosen during route selection. AODVR chooses the shortest route which is equivalent to choosing the farthest node towards the destination during transmission. But SELAR does not use the shortest route.

The maximum node energy spent in the network limits the network's lifetime. AODVR does not consider energy dissipation of nodes during route selection where SELAR considers only energy dissipation. That is why the maximum node energy spent in the

network is the highest in AODVR and the lowest in SELAR as observed in Fig. 3. Beside the success rate, energy dissipation is considered in the proposed protocol. During route selection, the proposed protocol chooses one node from a set of neighbours that have sufficient amount of remaining energy. Therefore, the maximum node energy dissipation in the proposed protocol is lower than that found in AODVR, but slightly higher than SELAR. Slightly higher energy consumption in the proposed protocol is acceptable as it attains higher success rate than SELAR which brings increased reliability of WSNs. Additional control packets are broadcast at regular intervals in the proposed protocol describing nodes' energy and congestion levels. Energy dissipation due to the transmission of these control packets have also been considered during simulation.
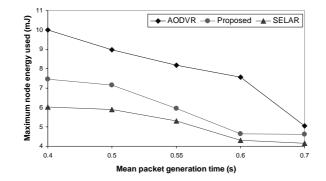

Fig. 3. Maximum node energy dissipation in the proposed protocol,
AODVR and SELAR with different packet generation rates

Fig. 4 shows the standard deviation of node energy spent expressed as the percentage of mean node energy spent across the network. Lower standard deviation means more uniform dissipation of energy among nodes. Node energy dissipation is the least uniform in AODVR but it is more uniform in both proposed protocol and SELAR. Moreover, the performances of the proposed protocol and SELAR are comparable in this case. SELAR chooses neighbour with the highest amount of remaining energy during data forwarding. As energy levels of nodes change continually, route selection in SELAR is somewhat random over time. This results more uniform energy dissipation of nodes. Route selection in the proposed protocol is also random based on neighbouring nodes' congestion status. Therefore, similar energy dissipation is also observed in the proposed protocol. But AODVR uses the same shortest paths over time which results non-uniformity of node energy consumption across the network.

Data success rates of AODVR, proposed protocol and SELAR at a fixed data generation rate under different transmission ranges of nodes are compared in Fig. 5. Network topology is same as the previous experiments. Mean packet generation time is 0.55 second and transmission range of nodes varies from 80 metres to 140 metres with a gradual increment of 20 metres. Behaviour similar to Fig. 2 is illustrated at all transmission ranges. In all cases, the proposed protocol has the highest success rate and SELAR has the lowest one.
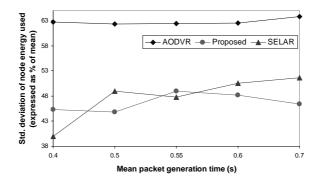
Fig. 4. Uniformity of node energy dissipation with different packet generation rates
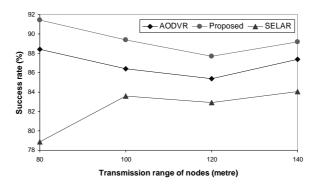

Fig. 5. Success rate at varying transmission ranges with a fixed
mean packet generation time of 0.55 s

CONCLUSION

The proposed routing algorithm reduces congestion by avoiding congested nodes during route selection. As a result, the BS receives a great number of data packets about the environment which are sent by the sensor nodes. In achieving its success, the proposed protocol utilizes congestion parameters into routing decision and at the same time, it is energy efficient and works in a distributed manner. Therefore, it is scalable, i.e., it can be used to gather agricultural data over a large geographic area. Simulation results show that the proposed protocol provides significantly high success rate with slight reduction in network lifetime. Future study will focus on analysing the performance of the proposed protocol in case of specific endangered plants rather than all plant species as common.

REFERENCES

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y. and Cayirci, E. 2002. A Survey on Sensor Networks. IEEE Communications Magazine: 102-114.

Ee, C. T. and Bajcsy, R. 2004. Congestion Control and Fairness for many-to-one Routing in Sensor Networks. Proceedings of ACM Sensys. Baltimore, MD, USA: 184-161.

Hull, B., Jamieson, K. and Balakrishnan, H. 2004. Mitigating Congestion in Wireless Sensor Networks. Proceedings of ACM Sensys. Baltimore, MD, USA: 134-147.

Kang, J., Zhang, Y. and Nath, B. 2007. TARA: Topology-Aware Resource Adaptation to Alleviate Congestion in Sensor Networks. IEEE Transaction on Parallel and Distributed Systems, 18(7): 919-931.

Kumar, R. 2008. Mitigating Performance Degradation in Congested Sensor Networks. IEEE Transactions on Mobile Computing, 7(6): 682-697.

Lukachan, G. and Labrador, M. A. 2004. SELAR: Scalable Energy-Efficient Location Aided Routing Protocol for Wireless Sensor Networks. Proceedings of 29th Annual IEEE International Conference on Local Computer Networks: 694-695.

Moghaddam, M. H. Y. and Adjeroh, D. 2008. A New Priority Based Congestion Control Protocol for Wireless Multimedia Sensor Networks. Proceedings of 9th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks. Newport Beach, CA, USA: 1-8.

NS-2 website, URL http://www-mash.cs.berkeley.edu/ns/.

Perkins, C. E. and Royer, E. M. 1999. Ad Hoc On-Demand Distance Vector Routing. Proceedings of Second IEEE Workshop Mobile Computing Systems and Applications.

Pottie, G. 1998. Wireless Sensor Networks. Proceedings of Information Theory Workshop. San Diego, CA, USA: 139-140.

Teo, J.-Y., Ha, Y. and Tham, C.-K. 2008. Interference-Minimized Multipath Routing with Congestion Control in Wireless Sensor Network for High-Rate Streaming. IEEE Transactions on Mobile Computing, 7(9): 1124-1137.

Wan, C.-Y., Eisenman, S. B. and Campbell, A. T. 2003. CODA: Congestion Detection and Avoidance in Sensor Networks. Proceedings of ACM Sensys. Los Angeles, California, USA: 266-279.

Wan, C.-Y., Eisenman, S. B., Campbell, A. T. and Crowcroft, J. 2005. Siphon: Overload Traffic Management using Multi-radio Virtual Sinks in Sensor Networks. Proceedings of ACM SenSys. San Diego, CA, USA.

Wang, C., Sohraby, K., Li, B., Daneshmand, M. and Hu, Y. 2006. A Survey of Transport Protocols for Wireless Sensor Networks. IEEE Network, 20(3): 34-40.

Zhou, H., Guan, X. and Wu, C. 2008. Reliable Transport with Memory Consideration in Wireless Sensor Networks. Proceedings of IEEE International Conference on Communications. Beijing, China, 2819-2824.