Probabilistic Data Collection Protocols for Energy Harvesting Sensor Networks

Masaya Yoshida
Shizuoka University, Japan
Email: yoshida@aurum.cs.inf.shizuoka.ac.jp

Tomoya Kitani
Shizuoka University, Japan
Email: t-kitani@inf.shizuoka.ac.jp

Masaki Bandai
Sophia University, Japan
Email: bandai@sophia.ac.jp

Takashi Watanabe
Shizuoka University, Japan
Email: watanabe@inf.shizuoka.ac.jp

Pai Chou
University of California, Irvine, US
Email: phchou@uci.edu

Winston Khoon Guan Seah
Victoria University of Wellington, New Zealand
Email: winston.seah@ecs.vuw.ac.nz

***Abstract*—Energy harvesting has been studied as a candidate for powering next generation wireless sensor networks. The technologies that can harvest electric power from ambient energy sources include solar, vibration, heat and wind. However, sensor nodes powered by energy harvesting devices cannot always communicate with other nodes because the energy harvesting devices cannot provide a stable supply power. A node cannot know whether its neighboring nodes have enough energy to receive a data packet that it has transmitted. During the process of relaying the packet, each additional hop increases the overall probability of losing the packet. In this paper, we propose two data collection protocols for the energy harvesting wireless sensor networks called Probabilistic ReTransmission protocol (PRT) and PRT with Collision Consideration (PRT-CC). The idea is to derive the number of times to retransmit a packet based on the reception probability and the active intervals computed by the receivers themselves. In PRT-CC, each node computes the reception probability with packet collision consideration. The simulation results show that the proposed protocols achieve higher delivery ratio than the previous works (GR-DD and GR-DD-RT).**

I. Introduction

Wireless sensor networks (WSNs) are networks consisting of compact devices called sensor nodes that collect data such as temperature and humidity over the target area. Sensing data are relayed via wireless links between nodes using a multi-hop communication protocol. Most of studies on WSNs have been motivated by their wide range of applications, ranging from natural environmental monitoring and disaster data collection to security systems.

Since WSN nodes may be deployed in large numbers and mostly run on batteries, the maintenance of nodes is required, including replacing batteries that have been depleted of energy. However, largely-scale maintenance costs are high or even prohibitive. For example, it is very difficult, even if possible, to reach nodes frequently in remote locations such as jungle or underwater. Thus, to extend the lifetime of WSNs, one of the important issues is saving power while providing the desired quality of communications. To attack the issue, various protocols for power-efficient data collection have been proposed [1][2].

Recently, energy harvesting has received growing attention in next-generation WSNs [3] leading to the design and development of energy harvesting WSNs [4], [5]. Energy harvesting entails converting forms of ambient energy from the environment, such as light, thermal differential and vibration into energy. Such renewable sources are expected to reduce the need for frequent maintenance, thus enabling sustainable operation of WSNs. However, due to the unpredictable supply of harvested energy, it is difficult to determine the operating state of neighboring nodes. For this reason, conventional protocols that assume energy is provided by batteries is not applicable to WSNs utilizing energy harvesting for power. Therefore, protocols need to be designed with energy harvesting assumptions in mind.

This paper proposes two new protocols, called Probabilistic ReTransmission (PRT) and PRT with packet Collision Consideration (PRT-CC), with the goal of achieving high efficiency and reliability in data collection in the presence of unsteady power supplied by the energy harvesting.

II. Related Work

WSNs can potentially contain large numbers of nodes. The size of nodes are usually small for low cost and ease of deployment. Thus, energy harvesting devices, which must also be small, are usually unable to sustain the continuous operation of most sensor nodes. Fig. 1 shows an example of the energy level of a battery and an energy harvesting device of similar size.

A network with such unsteady power supply conditions cannot operate at all times. Furthermore, charging takes a variable amount of time depending on the environment, making it difficult to predict the charging state of the surrounding nodes. As a result, it is difficult to determine the best path to the sink node for data transmission. Such characteristics specific to WSNs powered by ambient energy harvesting are very different from those traditionally powered by batteries, making it difficult to adopt previously designed WSN protocols. Thus, it is necessary to develop new protocols that considers media access control, routing, topology maintenance, sleep control, and power management, including charging control, in order

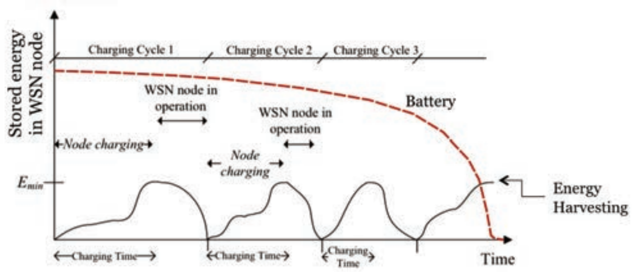


Fig. 1. Temporal change of energy level of a battery and an energy harvesting source

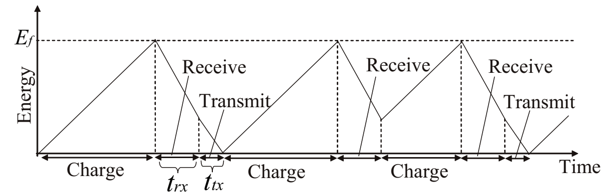


Fig. 2. The time history of the stored energy on the node

for WSNs to operate efficiently when powered by harvested energy sources.

One of related works is an energy-harvesting WSN for a railway monitoring system [6]. In [7], the authors introduce mobility to the power control. In [8], abstractions are developed to characterize the complex time varying nature of the harvesting sources, and it has been shown by [6], [7] and [8] that power control schemes can enable limited power to be used effectively. In [9], the authors study the rate assignment problem for rechargeable sensor networks and propose protocols to compute the optimal rate assignment. Two power control metrics have been developed for nodes powered by energy harvesting systems [10] which depend on the average queue length (the number of unsent messages) and average data loss rate (i.e., the amount of data that a node cannot receive during its sleeping/charging interval). In [11], a bridge monitoring application is studied with specific emphasis on the placement of the nodes and a data collection protocol optimized for the network topology and energy harvesting efficiency, while MAC protocols that can be used in energy-harvesting WSNs are studied and analyzed in [12].

As we will describe the problem in detail in section III, the instability of power supply arising from energy harvesting incurs packet loss resulting in low packet delivery rate. To overcome this problem, Geographic Routing with Duplicate Detection (GR-DD), and Geographic Routing with Duplicate Detection and Retransmission (GR-DD-RT) have been proposed [11]. In the GR-DD protocol, if a node receives the same data packet multiple times, then the latter packets are discarded to reduce unnecessary power consumption. GR-DD-RT is an extension of GR-DD, where if a node is fully charged and is ready to transmit but there is no unsent packet in the queue, it retransmits the previously sent packet (i.e., nodes repeatedly transmit as much as possible). These are simple protocols designed for use only in simple topologies like a linear topology, and therefore cannot achieve or provide higher performance in some scenarios. In this paper, we consider an arbitrary network topology and develop the data collection protocols that can achieve the higher delivery ratio than these protocols.

III. System Model

Our model is based on the repeating charging-and-transmitting model as proposed in [11]. Each sensor node is powered by an energy harvesting source, and it performs sensing and data transmission. The time history of the stored energy on a node is shown in Fig. 2, and each node operates according to the finite state machine shown in Fig. 3. During deployment, all nodes are pre-charged and therefore are able to execute a simple neighbor discovery process that involves broadcasting their locations to one another. Additional nodes that are deployed later will also broadcast their locations to their neighbors using the same neighbor discovery process.

Each node operates in three states: *charge*, *receive*, and *transmit*. After the receive or transmit state, a node returns to the charge state. At the charge state, node charges up to the minimum amount of energy, denoted by E_f, that should be sufficient for it to receive and transmit a packet. It then enters the receive state for a constant period of time. A node senses or receives data during the receive period. The amount of energy consumed by sensing is typically much lower than that for wireless communications. Therefore, we have accounted for the (small) amount consumed by sensing in the energy used by the node in the receive state. The receiving window, denoted by t_{rx}, is set to twice the width of the packet transmission time, denoted by t_{tx}, i.e. $t_{rx} = 2t_{tx}$ in common with [11]. The transmission time t_{tx} can be calculated by dividing the packet size s by the transmission rate α, namely

$$t_{tx} = \frac{s}{\alpha}.$$

Let P_{rx} and P_{tx} denote the receiving power consumption of a node and the transmitting power consumption of a node, respectively. Here E_f is set to the following equation as the energy required to perform one transmit and one receive as shown in Fig. 2,

$$E_f = P_{rx}t_{rx} + P_{tx}t_{tx}.$$

A received packet is put in the receiving node's queue. If there is a packet in the queue when the receive state ends and the channel is idle, then the node enters the transmit state and broadcasts the data packet at the head of the queue. At this time, any neighboring node that is currently in the receive state can receive the packet.After receiving and transmitting a packet, because the node has consumed its stored energy, it moves to the charge state. On the other hand, after the time the receiving state ends, if the queue is empty or the channel is not idle, then the node returns to the charge state until it charges up to E_f again. Each node repeats such cycles until the packet reaches the sink.

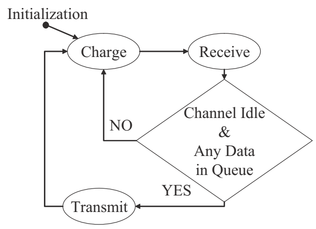


Fig. 3. The node operates according to the finite state machine

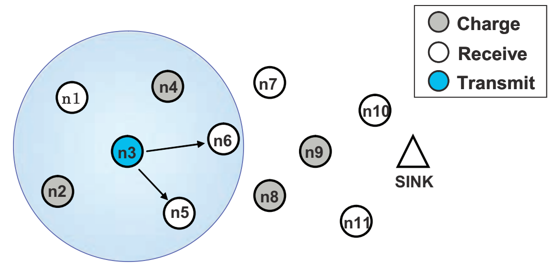


Fig. 4. Data collection example, where node n_3 transmits a packet

Even with the current state-of-the-art energy harvesting technology, if an energy harvesting device has the same (or even slightly larger) footprint as the wireless sensor node itself, the charging time can be significantly longer than the combined (receive and transmit) communication times. For instance, assuming we use MICAz nodes with the recharge rate of 10mW and the packet size of 800 bits, the receiving time t_{tx} will be 6.4ms and the transmitting time t_{tx} will be 3.2ms, whereas the charging time is about 77ms. This means at any moment, most neighboring nodes are expected to be in the charge state.

We use this model to evaluate the simplest case where nodes can opearate with small energy. It is possible to use the more complex model such as a node transmits and receives some packets in a cycle or a node charges and transmits/receives simultaneously. However, such a model makes the design of the sensor node more complex. In this paper, we assume that our model does not use ACK as [11] to keep the model simple. To use ACK, the collision avoiding schemes for ACK are needed.

A. Data Collection Example

Let us consider an example of data collection (see Fig. 4) where each node has a different charge level. Node n_3 generates a data packet to be transmitted. Within the communication range of n_3, there are nodes n_1, n_2, n_4, n_5 and n_6. When a node is physically further away from the sink node than a sender, it does not enqueue its received packet but discards it instead, because it is unlikely for the packet to be sent closer to the sink. Therefore, in this case, nodes $\{n_1, n_2\}$ do not relay the packet from n_3. After node n_3 generated data, and recharged, if the channel is idle at that time, then n_3 broadcasts the packet of the data. Suppose that node n_4 is being recharged and nodes $\{n_5, n_6\}$ are in the receive state, then nodes $\{n_5, n_6\}$ can receive the packet, as shown in Fig. 4. In the same manner, nodes $\{n_5, n_6\}$ in turn broadcast the packet. The sink is assumed to have steady power and is always in the receive state. In the manner described above, nodes relay the packets to the sink.

In another scenario, node n_3 generates another data packet and transmits it by broadcasting in the same manner as above. However, this time, suppose that all nodes $\{n_4, n_5, n_6\}$ are in charge state, and no node can receive the packet; consequently, the packet will be lost. This model assumes one-way communication where nodes transmit only data packets without transmitting ACK, or handshaking packets like Request-To-Send (RTS) and Clear-To-Send (CTS). Therefore, the sender node n_3 cannot know whether any of its neighboring nodes is able to receive the data packets it transmitted. As the packets are lost during relaying, the overall data collection rate deteriorates.

As the example shows, because the neighboring nodes are not always able to receive a packet during relaying, each additional hop increases the overall probability of losing the packet. Similarly, WSNs that take traditional approaches to the sleep control will suffer from the same problem, although a possible solution is to synchronize their sleep schedule. However, the synchronized approach cannot apply to energy harvesting WSNs due to the variable and unpredictable charging times of each node.

IV. Technical Approach

In conventional networking protocols, if a sender does not receive an ACK, then the sender interprets it to be a communication error and retransmits the packet to ensure high reliability. However, as shown in Section III, our communication model is broadcast-based without using ACK, and thus a sender cannot check whether any of its neighboring nodes actually receives the packet that it sent. By retransmitting the same packet, a sender's neighbors have a higher probability of receiving the packet, but it needs additional energy for the retransmission and longer time to recharge. In order to decide how many times a sender retransmits a packet, we propose the Probabilistic ReTransmission (PRT), by having each node calculates the probability of receiving a packet based on its own operating time and use this probability to decide whether to retransmit a packet or not. In addition, we propose PRT with Collision Consideration (PRT-CC), an extension of PRT which also considers the collision probability from hidden terminals of the sender.

A. Probabilistic ReTransmission (PRT)

The idea of PRT is to derive the number of times to retransmit a packet based on the reception probability and the operating time computed by the receivers themselves. Let N_i denote the set of nodes within the communication range of node n_i, and let S_i denote the subset of N_i whose nodes are closer to the sink than n_i. For example, in the case of Fig. 4, $N_3 = \{n_1, n_2, n_4, n_5, n_6\}$, and $S_3 = \{n_4, n_5, n_6\}$. Each node can measure its own operating time with a timer. We assume that a node n_i can keep measuring a total charging time T_{chi}, a total receiving time T_{rxi}, and a total transmitting time T_{txi} during the period $[t-\tau, t]$, where t denotes a current time and $\tau = T_{chi} + T_{rxi} + T_{txi}$ denotes a measurement period. If there is a node $n_h \in N_i \setminus S_i$ that broadcasts a packet to neighbor n_i in its communication range, the reception probability p_i can be approximated as follows:

$$p_i = \frac{T_{rxi}}{T_{chi} + T_{rxi} + T_{txi}} \cdot \frac{t_{rx} - t_{tx}}{t_{rx}}. \quad (1)$$

In Equation (1), t_{tx} and t_{rx} denote the time for transmitting and receiving one packet, respectively, as defined in Section III. The $\frac{t_{rx} - t_{tx}}{t_{rx}}$ part refers to the case of the sender's transmit (Tx) mode not overlapping with the receiver's receive (Rx) mode, and hence, the receiver cannot receive the packet. The entire equation implies that if a node has a high charging rate, then it also has high reception probability. If $p_i = 0$, then it means the node is unable to charge and therefore unable to operate at all; whereas if $p_i = 1$, then it means it is powered by a stable source and is always ready to receive, such as the sink node. Node i sets the initial value of neighbor node j's reception probability p_j to 1 and updates p_j when node i receives a packet from node j. τ should be longer than one cycle (charge, receive and transmission), but it becomes increasingly harder to adapt to a change of environment as τ become larger; consequently, the optimal τ depends on the network environment, such as, the type of energy harvest device that node uses and its charging rate.

When node n_i broadcasts a packet once, the probability q_i that at least one node can receive the packet is expressed as:

$$q_i = 1 - \prod_{j}^{S_i} (1 - p_j). \quad (2)$$

According to equation (2), there are nodes with high reception probability within range of node n_i, or if the number of nodes within range is large, then the probability that at least one node can receive is high. Assuming the sender (n_i) repeats transmitting a given packet a times, then the probability that at least one node receives the packet, r_{ai}, is calculated as:

$$r_{ai} = 1 - (1 - q_i)^a.$$

In PRT, let Th denote the reliability threshold on q_i, and we assume that $r_{ai} \geq Th$ is reliable. Then, a sender calculates an optimal number of retransmission, denoted by a', which is the minimum value such that the probability $r_{a'i}$ remains above the reliability threshold, as follows:

$$\begin{aligned} r_{a'i} &\geq Th \\ 1 - (1 - q_i)^{a'} &\geq Th \\ (1 - q_i)^{a'} &\leq 1 - Th. \end{aligned}$$

Taking $\log_{(1-q_i)}$ on both sides,

$$a' \geq \log_{(1-q_i)} (1 - Th). \quad (3)$$

After the sender calculates a', it then transmits a packet a' times. The Th value should be chosen based on the reception probability of the neighboring nodes. As Th increases, so does the optimal number of repeated transmissions a' and energy consumption. Therefore, it is expected that an optimal Th value exists for different settings.

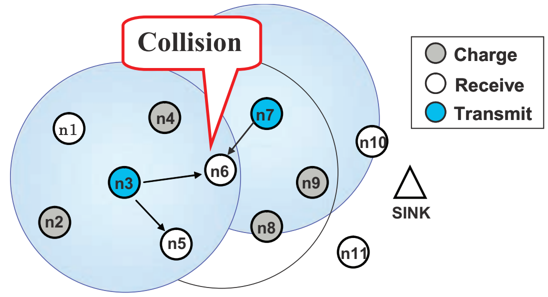


Fig. 5. The nodes n_7, n_8 and n_9 are hidden terminals of node n_3

B. PRT with Collision Consideration (PRT-CC)

In PRT, the reception probability p_i of node n_i is computed without considering collisions. Therefore, a' can be estimated to be smaller than the optimal value when collisions occur frequently. Then, in PRT-CC, each node computes the reception probability with collision consideration.

Figure 5 shows an example that node n_3 broadcasts a packet. Nodes in $N_3 = \{n_1, n_2, n_4, n_5, n_6\}$ cannot enter the transmit state because the nodes do carrier sensing before starting transmission as shown in Section III. However, nodes $\{n_7, n_8, n_9\}$ that are outside communication range of node n_3, cannot know that node n_3 is transmitting a packet. At this time, if a node in $N_6 \setminus N_3 \setminus n_3$ starts transmission, n_6 cannot receive the packet sent from n_3 correctly (i.e., the nodes $\{n_7, n_8, n_9\}$ are hidden terminals of node n_3).

In PRT-CC, given a node n_i and a neighboring node n_j of n_i, the reception probability p'_{ji} that node n_j can receive the packet sent from n_i can be recalculated as follows:

$$p'_{ji} = p_j \cdot \prod_{k}^{N_j \setminus N_i \setminus \{n_i\}} (1 - c_k). \quad (4)$$

In Equation (4), the c_k expresses the probability that hidden terminals $n_k \in N_j \setminus N_i \setminus \{n_i\}$ starts transmission when node n_i transmits a packet, and it is approximated as follows:

$$c_k = \frac{T_{txk}}{T_{chk} + T_{rxk} + T_{txk}}. \quad (5)$$

Equations (4) and (5) imply that if there is a node that transmits packets frequently in $N_6 \setminus N_3 \setminus n_3$, then the collision probability will increase, and the reception probability will decrease. Moreover, the probability that at least one node can receive the packet when node n_i broadcasts a packet once is expressed as Equation (6):

$$q'_i = 1 - \prod_{k}^{S_i} (1 - p'_{ki}). \quad (6)$$

Computing a' by Equation (3) with q'_i instead of q_i, nodes can estimate the optimal number of times for retransmission

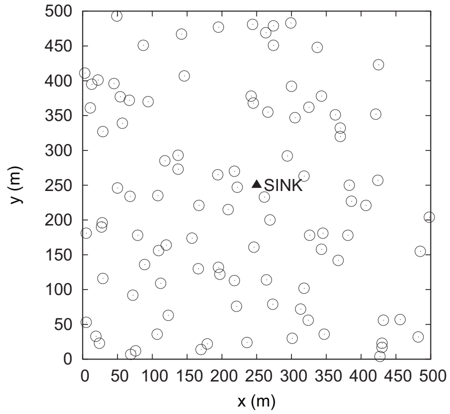


Fig. 6. Random Topology

with collision consideration. In order to ensure that a packet is transmitted a' times, PRT and PRT-CC use the duplicate detection technique of GR-DD, where if a node receives the same data packet multiple times, then the latter packets are discarded. In PRT-CC, more overhead is incurred than PRT, due to the use of information of the neighboring nodes' transmission probability, in addition to the reception probability.

C. Exchanging Packet Reception Probability

In both the PRT and PRT-CC protocols, a sender node uses its neighboring nodes' packet reception probability to compute the number of times for retransmission. Since each node computes its own packet reception probability, there needs to be a way for a node to obtain this information from its neighboring nodes.

One way of exchanging the information of probability among nodes is to schedule a time interval for information exchange. During this interval, all nodes update the probability of their neighbors simultaneously. However, in energy-harvesting WSNs, not all nodes have the adequate amount of charged energy to perform this exchange task at the scheduled time, making this scheme difficult to realize. Furthermore, time synchronization is difficult, if not impossible, to achieve when nodes run out of energy.

Another method is to piggyback the packet reception probability onto regular data packets. The advantages are that this does not require a special time interval dedicated to just information exchange, and it does not require the model of operation to be modified. Being a broadcast model means that multiple receivers can receive simultaneously, and each node just needs to update its local information. The overhead results in higher energy consumption and make charging time longer. However, the additional data overhead required is only the reception probability information. Depending on the desired accuracy of the information of the probability, e.g. if we assume it is 8 bits, then PRT-CC requires 32 bits of extra information for the reception probability and collision probability.

TABLE I
SIMULATION PARAMETERS

Simulation Time	1000 (s)
Transmission Power Consumption P_{tx}	76.2 (mW)
Receiving Power Consumption P_{rx}	83.1 (mW)
Transmission Power	-3.0 (dbm)
Receiving Power	-85.0 (dbm)
Transmission Antenna Gain	0 (dbi)
Receiving Antenna Gain	0 (dbi)
Frequency f	2.4 (GHz)
Data size	800 (bits)
Transmission Rate α	250 (kbps)

V. EVALUATIONS

We evaluate our proposed protocols by simulation. In this simulation, n nodes are deployed randomly in the area as shown in Fig. 6. The area size is 500m $\times$ 500m. There is a sink node and it is located at the center of the area. This sink is powered by a steady power supply so that it can be always ready to receive. In the simulation, we assume there is no packet loss within the communication range, but collision means no correct packet can be received. The communication range of node is computed by Friis transmission equation [13]. The decay factor (m) of the Friis transmission equation is 2, giving a communication range of about 125m. At all nodes, data packets are generated according to a Poisson distribution with an arrival rate of 0.01 packet/s. We assume that the original packet size s is 800bits. To provide the necessary information for out protocols, the size of a packet for PRT is 808 bits including 8bits of reception probability, and that for PRT-CC is 832bits including 32bits of reception probability and collision probability, respectively.

Since the charging rate of each node varies with time, we denote the average of charging rate by Ch and the variation of Ch by v. Given a node, the charging rate of the node is generated by the uniform distribution within $[Ch-v, Ch+v]$ for each instance of time. In the simulation, Ch is 10mW similar to [11] and v is fixed at a value of 2mW.

In PRT and PRT-CC, the reliability threshold Th is set to 0.9, the measurement period τ is set to ∞ because the functions using τ is only for the network where the average charging rate varies greatly. The transmission/receiving power, data rate, and various parameters of the node are based on those of the MICAz platform. The simulation parameters are shown in Table I.

The evaluation metrics are delivery ratio and delay. The delivery ratio is defined to be the unique packet count as received by the sink divided by the total generated packet count. The delay is the time from the instant a packet is generated till the instant it reaches the sink node. To evaluate the performance of the PRT and PRT-CC protocols, we compare our results with those for GR-DD and GR-DD-RT. We evaluate the proposed protocols, with different parameter values, namely, the number of nodes n, the average of charging rate Ch, the decay factor m and the reliability threshold Th.

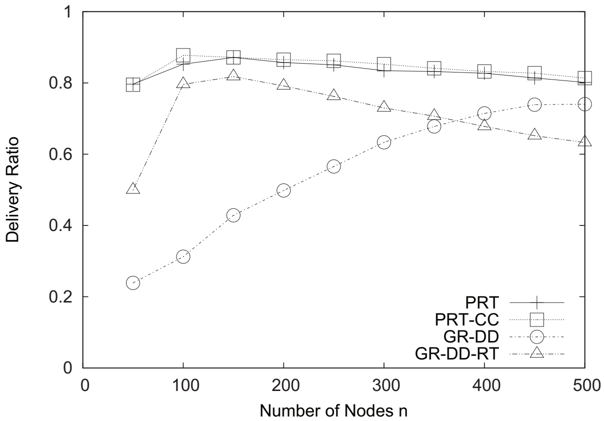


Fig. 7. Number of Nodes *n* vs. Delivery Ratio

A. Number of Nodes (n) vs. Delivery Ratio

As the number of nodes *n* changes, it affects the node density and the delivery ratio. We evaluate the delivery ratio of each protocol by varying the number of nodes.

The simulation results in Fig. 7 show that PRT and PRT-CC achieve a higher data delivery ratio than the GR-DD and GR-DD-RT.

In PRT and PRT-CC, each node can adapt to the changes in the node density by computing the optimal number of times to retransmit a packet depending on the number of nodes within its communication range. The results show slight improvement in PRT-CC compared with PRT. This is because collision consideration is not very effective when collisions do not occur frequently. Moreover, increasing packet size degrades the the delivery ratio of PRT-CC. On the other hand, PRT is effective even though its operation is simple.

As the number of nodes in the area increases, the delivery ratio also increases in GR-DD while its in GR-DD-RT decreases. Since the number of nodes in the communication range of each node increases, the number of nodes which can receive a packet when sender broadcasts it also increases. In GR-DD-RT, each node retransmits the packets as many as possible and consumes excessive energy regardless of the change of the node density. As a result, the collisions in the network occur more frequently and the charging time becomes longer, and adversely affects the delivery ratio.

B. Charging Rate (Ch) vs. Delivery Ratio

The charging rate of a node is also a critical factor which affects the system performance directly. The simulation results are illustrated in Fig. 8 for different values of *Ch*. When the number of nodes in the area is small, GR-DD-RT achieves higher delivery ratio than GR-DD (Fig. 8(a)). When more nodes are deployed in the area, GR-DD provides superior performance compared with GR-DD-RT if *Ch* is higher than 5mW (Fig. 8(c)). Our proposed protocols PRT and PRT-CC achieve high delivery ratio consistently regardless of the number of nodes in the area and the charging rate.

Every protocol achieves the higher delivery ratio with the higher average charging rate *Ch*. As the charging rate increases, nodes can charge quickly and can operate with higher duty cycle. Nodes have more opportunities for receiving/transmitting packets, and can deliver more packets to the sink.

According to [14], the charging rate on a 10cm^2 energy harvesting material (which is about the same size as the mote) is from 0.032mW (indoor) to 37mW (direct sunlight) when energy is harvested from solar or 5mW (piezoelectric) when energy is harvested from vibration. The charging rate is expected to improve as the energy harvesting technology advances.

C. Decay Factor (m) vs. Delivery Ratio

In PRT and PRT-CC, each node uses the information of neighboring nodes in its communication range. In this simulation, we vary the decay factor *m* of Friis transmission equation to change the communication range of nodes and the number of nodes in the range.

Figure 9 shows the results of the packet delivery ratio of the four protocols for different values of *m*. Every protocol provides low delivery ratio with large *m*. In the case with *m* larger than 3, nodes cannot deliver most of the packets in every protocol.

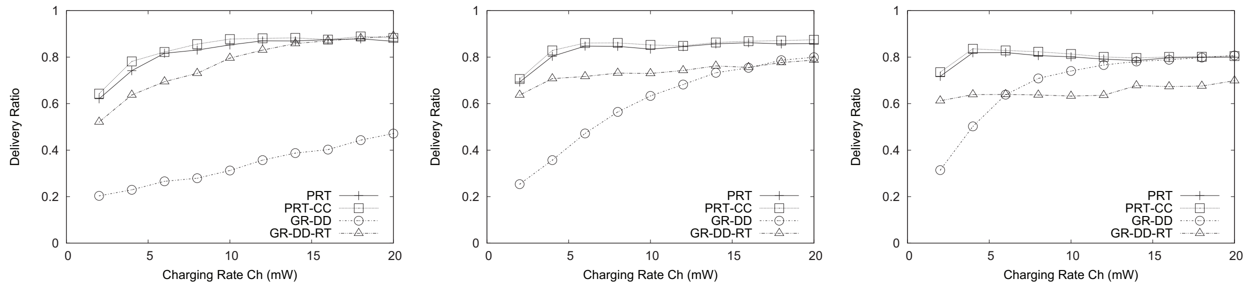
The communication range of a node depends on the quality of the wireless link. At the same transmission power, a node can transmit a signal further in obstacle-free space than one with obstacles. In the Friis transmission equation, the decay factor *m* represents the path quality, where larger *m* denotes worse condition for signal transmission. *m* is measured empirically, $m = 2$ refers to free (outdoor) space and $m = 2.5$ means the indoor space while $m \geq 3$ means the environment where it is not good for wireless communication, approximately. At shorter ranges, the number of neighbor nodes become fewer, and consequently, the delivery ratio decreases. To overcome this problem, nodes should use higher power to transmit packets at the cost of longer charging times.

D. Reliability Threshold (Th) vs. Delivery Ratio

In PRT and PRT-CC, we can control the nodes by setting the different reliability threshold *Th*. Figure 10 shows the results for different values of *Th*. The results show that PRT and PRT-CC achieve higher delivery ratio.

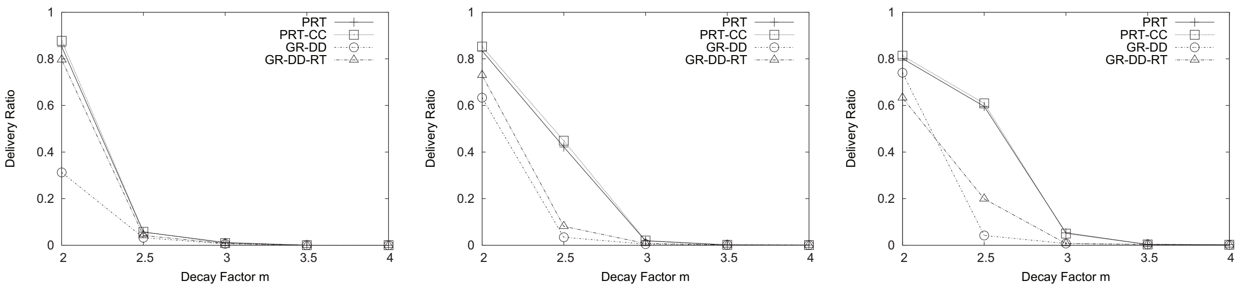
When *Th* is set to a large value, nodes retransmit the same packet many times to achieve high reliability. However, in the case where large amount of data are generated at a node or charging rate of node is too small, if *Th* is set to a large value, the node cannot transmit all the packets in its queue. Therefore, *Th* should be appropriately chosen depending on the node's status.

Moreover, *Th* also depends on applications. If the system is used in critical applications like disaster information networks and structural health monitoring [15], *Th* should be set to a large value to provide high accuracy and reliability. On the other hand, large *Th* is not needed when the system is used in applications which do not need high reliability, such as environmental monitoring [16].



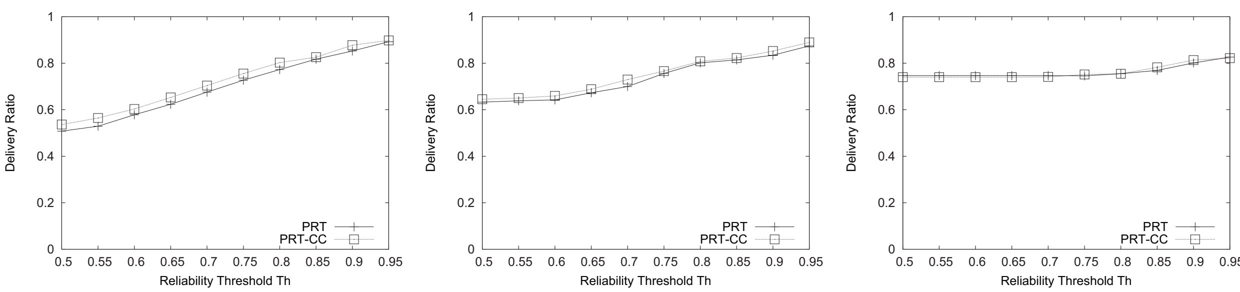
(a) $n = 100$ (b) $n = 300$ (c) $n = 500$

Fig. 8. Charging Rate *Ch* vs. Delivery Ratio



(a) $n = 100$ (b) $n = 300$ (c) $n = 500$

Fig. 9. Path Loss Coefficient *m* vs. Delivery Ratio



(a) $n = 100$ (b) $n = 100$ (c) $n = 500$

Fig. 10. Reliability Threshold *Th* vs. Delivery Ratio

E. Delay

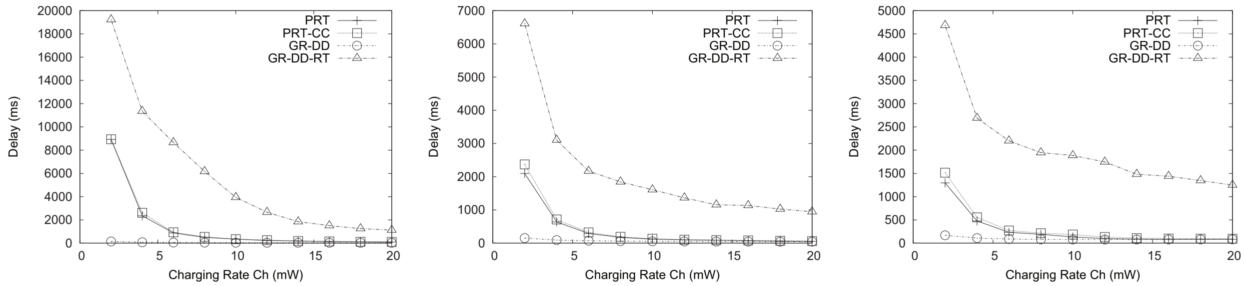
Some urgent sensor network applications need to collect data quickly. Therefore, we evaluate the delay with different number of nodes in the area n and average charging rate Ch.

Figures 11 and 12 show the delay of the packets for different number of nodes and average charging rates, respectively. As the average charging rate Ch increases, the delay decreases in every protocol because of the shorter charging time. All the results show that GR-DD-RT takes a long time to deliver the packets to the sink. This is because in GR-DD-RT some packets reach the sink node after a node broadcasts many times.

The delay in the proposed PRT and PRT-CC are larger than GR-DD especially when Ch is small, for example when the sun is obscured by clouds. The reason is that the delivery ratio in GR-DD is lower than PRT and PRT-CC. where many packets from nodes far away from the sink cannot reach the sink node actually and computation of the delay is based on few packets that actually reached the sink. If the charging rate is more than 10mW, regardless the number of nodes, the delay is less than 1 second. It can be short enough to use for most applications including not only environment monitoring but also most sensor networks.

VI. Conclusions

In this paper, we discussed data collection for energy harvesting wireless sensor networks and showed that frequent packet loss as an intrinsic problem. In order to overcome the problem and to improve data collection efficiency, we proposed two data collection protocols named Probabilistic ReTransmission (PRT) and Probabilistic ReTransmission with Collision Consideration (PRT-CC), where a sender node calculates the probability of receiving a packet based on its own active intervals. We evaluated PRT and PRT-CC by simulation, and the results show that PRT and PRT-CC achieve higher delivery ratio than the previous works (GR-DD and GR-DD-RT). For the future work, we are considering mathematical analysis of the models and protocols and how to determine the optimal Th and τ for different network scenarios. We are



(a) $n = 100$ (b) $n = 300$ (c) $n = 500$

Fig. 12. Charging Rate *Ch* vs. Delay

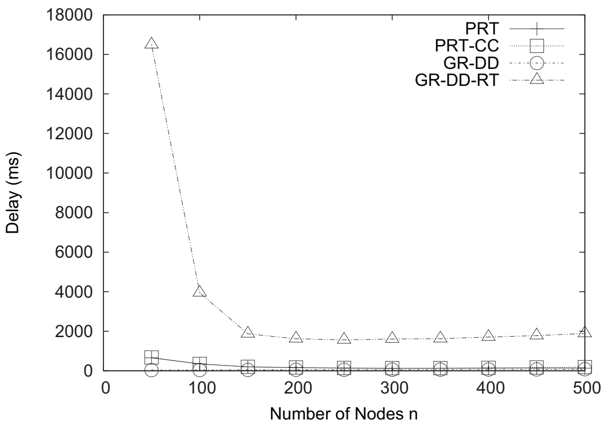


Fig. 11. Number of Nodes *n* vs. Delay

also studying other approaches including explicit acknowledgments, and comparing them with our protocols.

REFERENCES

[1] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. of the Hawaii Int'l Conf. on System Sciences*, 2000, pp. 1–10.

[2] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficent mac protocol for wireless sensor networks," in *Proc. of the IEEE Computer and Communications Societies(INFOCOM)*, 2002, pp. 1567–1576.

[3] A. Kansal and M. Srivastava, "An environmental energy harvesting framework for sensor networks," in *Proc. of the 2003 International Symposium on Low Power Electronics and Design*, 2003, pp. 481–486.

[4] C. Park and P. Chou, "AmbiMax: Autonomous energy harvesting platform for multi-supply wireless sensor nodes," in *Proc. of SECON*, 2006, pp. 168–177.

[5] K. Lin, J. Hsu, S. Zahedi, D. C. Lee, J. Friedman, A. Kansal, V. Raghunathan, and M. B. Srivastava, "Heliomote: Enabling long-lived sensor networks through solar energy harvesting," in *Proc. of ACM Sensys*, November 2-4 2005.

[6] H.-P. Tan, P. Lee, W. K. G. Seah, and Z. Eu, "Impact of Power Control in Wireless Sensor Networks Powered by Ambient Energy Harvesting (WSN-HEAP) for Railroad Health Monitoring," in *Proc. of the Second International Workshop on Applications of Ad hoc and Sensor Networks (AASNET09)*, Bradford, UK, May 26-29 2009, pp. 804–809.

[7] M. Rahimi, H. Shah, G. Sukhatme, and J. Heideman, "Studying the feasibility of energy harvesting in a mobile sensor network," in *Proc. of IEEE Int'l Conf. Robotics and Automation*, 2003, pp. 19–24.

[8] A. Kansal, S. Z. J. Hsu, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Transactions on Embedded Computing Systems*, 2006.

[9] K.-W. Fan, Z. Zheng, and P. Sinha, "Steady and fair rate allocation for rechargeable sensors in perpetual sensor networks," in *Proc. of the 6th ACM conference on Embedded network sensor systems*, 2008, pp. 239–252.

[10] V. Joseph, V. Sharma, and U. Mukherji, "Optimal sleep-wake policies for an energy harvesting sensor node," *PEDIATRICS*, vol. 115, no. 1, pp. 250–256, 2005.

[11] Z. A. Eu, H.-P. Tan, and W. K. G. Seah, "Routing and Relay Node Placement in Wireless Sensor Networks Powered by Ambient Energy Harvesting," in *Proc. of the IEEE Wireless Communications and Networking Conference (WCNC)*, Budapest, Hungary, Apr 5-8 2009, pp. 5–8.

[12] ——, "Design and performance analysis of mac schemes for wireless sensor networks powered by ambient energy harvesting," *Ad Hoc Networks*, vol. 9, no. 3, pp. 300–323, 2011.

[13] H. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, May 1946.

[14] W. K. G. Seah, Z. A. Eu, and H.-P. Tan, "Wireless Sensor Networks Powered by Ambient Energy Harvesting (WSN-HEAP) – Survey and Challenges," in *Proc. of the 1st International Conference on Wireless Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE)*, May 17-20 2009.

[15] C. Park, Q. Xie, P. Chou, and M. Shinozuka, "DuraNode: wireless networked sensor for structural health monitoring," in *Proc. of the IEEE International Conference on Sensors 2005*, 2005, pp. 277–280.

[16] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *ACM International Workshop on Wireless Sensor Networks and Applications*, 2002.