ProHet: A Probabilistic Routing Protocol with Assured Delivery Rate in Wireless Heterogeneous Sensor Networks

Xiao Chen, Zanxun Dai, Wenzhong Li, Yuefei Hu, Jie Wu, Hongchi Shi, and Sanglu Lu

***Abstract*—Due to different requirements in applications, sensors with different capacities are deployed. How to design efficient, reliable and scalable routing protocols in such wireless heterogeneous sensor networks (WHSNs) with intermittent asymmetric links is a challenging task. In this paper, we propose *ProHet*: a distributed probabilistic routing protocol for WHSNs that utilizes asymmetric links to reach assured delivery rate with low overhead. The ProHet protocol first produces a bidirectional routing abstraction by finding a reverse path for every asymmetric link. Then, it uses a probabilistic strategy to choose forwarding nodes based on historical statistics using local information. Analysis shows that ProHet can achieve assured delivery rate ρ if ρ is set within its upper-bound. Extensive simulations are conducted to verify its efficiency.**

***Index Terms*—Asymmetric links, heterogeneous sensor networks, routing, two-hop neighborhood information, wireless.**

I. Introduction

RECENT advances in wireless communication technologies and electronics have paved the way for developing low-cost wireless sensor networks (WSNs). WSNs have a wide range of military and civilian applications such as target tracking [3], environment monitoring [18], intelligent homes [11], disaster rescuing [22], and self-touring systems [23].

In WSNs, sensors gather information, such as temperature, humidity, light, etc. from the environment, process them locally, and then communicate with others or send the information to the sink for further processing. In various applications, different sensors may be used [17], [32]. Therefore, sensors may not have the same sensing capability and transmission range. Here we just take their diverse transmission ranges brought about by their heterogeneity into account. The WSN formed by heterogeneous sensors is referred to as the *wireless heterogeneous sensor network* (WHSN).

After the heterogeneous sensors have completed data collection, one major issue is how to route data to the destination (mostly it is the sink in WSNs) efficiently [15], [19], [33]. Since these heterogeneous sensors have different transmission ranges, there will be asymmetric links in the communication graph because if node A can reach node B, but B can not reach A, then the directed link from A to B is asymmetric. Thus, the common undirected graph generated after abstraction is turned into a directed graph, which makes the off-the-shelf routing protocols for general WSNs not applicable or work with higher overhead [24]. So the routing protocols for WHSNs need to be redesigned and should meet the following requirements: (1) Reliable with assured delivery rate and low overhead, which are important for mission critical applications; (2) totally distributed and use only local information for scalability and robustness purposes.

In this paper, we propose **ProHet**: a **Pro**babilistic routing protocol for **Het**erogeneous sensor networks, which can handle asymmetry links well and work in a distributed manner using local information with low overhead and assured delivery rate. It has two parts: the *preparation part* which includes identifying neighbor relationships and finding a reverse path for an asymmetric link, and the *routing part* which includes selecting nodes, forwarding messages and sending acknowledgement.

Other important issues in WSNs such as energy consumption and hot-spot are not discussed since the focus here is the usage of asymmetric communication links and assured delivery rate in WHSNs. Previous works have extensively studied the energy consumption and hot-spot problems in sensor networks [1], [5], [10], [21], [26], [35]. So we want to address the issues that are neglected by them.

The rest of the paper is organized as follows: Section II references the related work. Section III presents preliminaries. Section IV proposes the ProHet protocol. Section V does the analysis. Section VI evaluates the performance of ProHet by simulations. And the conclusion is drawn in Section VII.

II. Related Work

In this section, we give an overview of existing routing algorithms in WHSNs and probabilistic routing strategies.

A. Routing in Heterogeneous Sensor Networks

Routing in homogeneous sensor networks has been well addressed by many routing protocols [13], [15], [16], [19], [20], [25], [28], [29], [33]. In these protocols, all sensors have the same capabilities in terms of communication, computation, energy supply, reliability, etc. However, in applications such as

Manuscript received January 12, 2012; revised September 7, 2012; accepted December 31, 2012. The associate editor coordinating the review of this paper and approving it for publication was Y. Chen.

X. Chen, Z. X. Dai, and H. C. Shi are with the Department of Computer Science, Texas State University, San Marcos, TX, 78666 USA (e-mail: {xc10, zd1020, hs15}@txstate.edu).

W. Z. Li, Y. F. Hu, and S. L. Lu are with the State Key Laboratory for Novel Software Technology, Nanjing University, China (e-mail: {lwz, huyuefei, sanglu}@dislab.nju.edu.cn).

J. Wu is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA, 19122 USA (e-mail: jiewu@temple.edu).

Digital Object Identifier 10.1109/TWC.2013.013013.120007

aforementioned, heterogeneous sensors with different capabilities may be deployed. It is reported in [32] that when properly deployed, heterogeneity can triple the average delivery rate and provide a five-fold increase in the network lifetime. Routing in WHSNs should be rethought about: Simply using the routing protocols in homogeneous sensor networks does not take advantage of the diversity of the sensors.

In the literature, there are a few routing protocols designed for WHSNs [1], [7], [9], [10], [12], [34] where the sensors are categorized into powerful and less powerful ones. Sensors form clusters, with the powerful ones being the cluster heads. Two-level routing protocols are used in the network: The intracluster protocol is used to route messages between less powerful nodes and their clusterheads while the intercluster protocol is used to route messages between clusterheads. In these protocols, the capability of each individual sensor is not distinguished and the asymmetric links are not fully utilized. In [14], we proposed a protocol that differentiates the diverse transmission ranges of sensors and takes advantage of the asymmetric links to achieve assured delivery rate. But our preliminary work does not disclose the relationship between the assured delivery rate and the network parameters. In this paper, we enhance that in our analysis, give a more comprehensive description of ProHet, and conduct more simulations to justify our design idea and calculate overhead more accurately.

B. Probabilistic Routing Strategies

The probabilistic routing strategy in WSNs is not a new topic and there are various studies about it. Paper [27] uses probabilistic routing to disseminate information in a wireless sensor network without maintaining routing table: the sensor nodes simply forward the received packets with some probability. Thus, it reduces traffic in the network and mitigates the broadcast storm problem. The authors in [4] propose Parametric Probabilistic Sensor Network Routing Protocols, a family of light-weight and robust multi-path routing protocols for sensor networks in which an intermediate sensor decides to forward a message with a probability that depends on various parameters, such as the distance of the sensor to the destination, the distance of the source sensor to the destination, or the number of hops a packet has already traveled. Probabilistic Flow-based Spread Routing Protocol in [30] makes the intermediate nodes forward packets with a probability based on neighboring nodes' traffic load and tries to achieve the balance of energy consumption when forwarding packets. In [8], the information obtained by sensors from the environment has different delivery probabilities according to their levels of importance to the end user. For example, the information of a potential chemical leak is more important than knowing that everything is fine and should have a higher delivery probability. The authors propose a new method for information delivery at a desired reliability using hop-by-hop schemes.

In the above works, the computation of probability has never been referred to a node's historical information of its delivery capability which may result in better performance. In this paper, we will explore historical statistics and propose a probabilistic routing protocol with assured delivery rate.

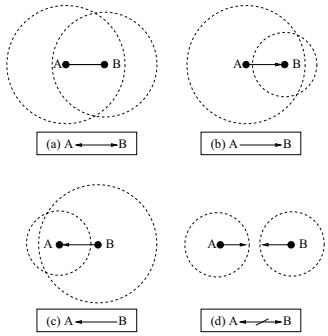


Fig. 1. The neighbor relationships between two nodes A and B. (a) A and B are In-out-neighbors of each other; (b) A is an In-neighbor of B and B is an Out-neighbor of A; (c) B is an In-neighbor of A and A is an Out-neighbor of B; (d) A and B are non-neighbors.

III. Preliminaries

A. Definitions of Nodes' Neighbor Relationships

A WHSN can be represented by a directed graph $G = \{V, E\}$, where V is the set of sensors (also called nodes), and E is the set of links (also called edges) in the network. For example, if sensor B is in the transmission range of sensor A, then there is a directed link from A to B. We assume graph G generated from the sensor network is a strongly-connected directed graph. Therefore, the sensor network is strongly-connected too.

We categorize the neighbor relationships of sensors into four categories: (1) In-out-neighbor; (2) In-neighbor; (3) Out-neighbor; and (4) Non-neighbor. For two nodes A and B, as shown in Figure 1(a), if $A \rightarrow B$ and $B \rightarrow A$, then A and B are In-out-neighbors of each other. If only $A \rightarrow B$ (or $B \rightarrow A$) as in Figure 1(b) (or 1(c)), then A (or B) is an In-neighbor of B (or A), and B (or A) is an Out-neighbor of A (or B). If neither $A \rightarrow B$ nor $B \rightarrow A$, they are non-neighbors of each other, as shown in Figure 1(d).

B. Definitions of One-hop and Two-hop Receivers

A node's one-hop receiver is the node's Out-neighbor or In-out-neighbor. A node's two-hop receiver is the one-hop receiver of the node's one-hop receiver.

C. Definition of Two-hop Neighborhood Information Model

In WSNs, the two-hop neighborhood information model, which means a node knows the information of its neighbors and the neighbors of its neighbors, is used by some researchers to guide routing [2], [6], [31]. This model is very attractive to large-scale WSNs because only local information is needed. This model is still helpful in WHSNs to steer the routing in the right direction. But because of the asymmetric links, the definition of the two-hop neighborhood information model in WHSNs should be changed to: A node knows its one-hop receivers and the one-hop receivers of its one-hop receivers. Still, the two-hop neighborhood information can be obtained by exchanging "Hello" messages between nodes in WHSNs.

Theorectically speaking, any k-hop ($k \geq 1$) neighborhood information model can be used. However, if one-hop neighborhood information is used, it is more like flooding which will cause large number of redundant data packets. If k-hop ($k \geq 3$) neighborhood information is used, it will

introduce much more communication overhead among neighbors without bringing much benefit comparing with the two-hop neighborhood information model. Our later simulation confirms these.

D. Definition of Delivery Probability

A node's delivery probability $P_{delivery}$ is defined as the ratio of the number of packets successfully delivered by the node denoted by N_d and the number of packets forwarded by it, denoted by N_f. It can be expressed as:

$$P_{delivery} = N_d/N_f \tag{1}$$

N_d and N_f for a node will be recorded in the routing process so that $P_{delivery}$ can be calculated locally and timely. At the beginning of routing when N_d and N_f do not exist, a routing protocol can work in a flooding manner for a while to establish these values. After some rounds of packet delivery, every node's delivery probability will become stable. Thus, the historical information of the network has been established.

IV. The ProHet Protocol

In this section, we present the ProHet protocol, which has two parts: the preparation part which includes identifying neighbor relationships and finding a reverse path for an asymmetric link, and the routing part which includes selecting nodes, forwarding messages and sending acknowledgement. The details are as follows:

A. Preparation Part

First each node needs to identify its In-out-neighbors and In-neighbors (if there is any) by sending each other "Hello" messages (see algorithm Identifying Neighbor Relationships). The identification of a node's Out-neighbors needs to wait until a reverse path is found.

Algorithm: Identifying Neighbor Relationships

1: Every node in the network broadcasts a "Hello" message.
2: If two nodes A and B can receive each other's "Hello" message and the corresponding "Acknowledgment" of the "Hello" message, then each adds the other to its In-out-neighbor list.
3: If A receives B's "Hello" message, but not the "Acknowledgement" of its own "Hello" message, then A knows that B is its In-neighbor and adds it to its In-neighbor list. Then, A will find a reverse routing path to B.

Next, for each node that has an In-neighbor, it is necessary to find a reverse path using the Finding a Reverse Path algorithm. Finding a reverse path can fully utilize the asymmetric links in the HWSNs. The study of [24] shows that a significant percent of links in WHSNs are asymmetric and the connectivity of the network can be up to 97% when the maximum reverse routing path length (here "length" means the number of hops) is set to be 3. We set the expiration path length to 3 and most nodes can establish their reverse routing paths to their In-neighbors in our experiment. If a node receives more than one reverse routing path to an In-neighbor, it chooses the shortest one.

Algorithm: Finding a Reverse Path

1: Node A tries to find the reverse routing path to each of its In-neighbors by broadcasting a "Find" message containing the source ID ("A"), the destination ID (the ID of the In-neighbor (e.g. "B")), and an expiration length of 3 hops.
2: **if** some node C receives a "Find" message, **then**
3: **if** it is the destination node listed in the message, **then**
4: it will add the source node to its Out-neighbor list,
5: and send the identified reverse routing path to the source node by a "Path" message containing the reverse route.
6: **end if**
7: **if** it is not the destination node and the expiration length is greater than 0, **then**
8: it will rebroadcast the message after the following modifications:
9: decrease the expiration length by one;
10: append its own ID to the message.
11: **end if**
12: in all other cases, it will drop the message.
13: **end if**

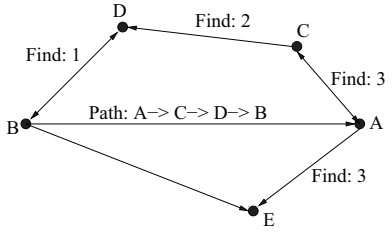


Fig. 2. An example of finding a reverse routing path.

Let's use the WHSN in Fig. 2 to explain the preparation part. In this network, A, B, C, D, E are sensors with different transmission ranges. The directed links in the graph represent their neighbor relationships. After broadcasting "Hello" messages, sensors B and D, A and C can receive each other's "Hello" messages and "Acknowledgements". Thus they identify each other as In-out-neighbors. However, sensor A gets sensor B's "Hello" message, but does not receive B's "Acknowledgement" to its own "Hello" message. It knows that B is its In-neighbor. Then, it starts to find a reverse routing path to B by broadcasting a "Find" message (A, B, 3). The number after "Find" in the figure represents the expiration length, initially set to 3. The "Find" message is received by sensors C and E. Sensor E's transmission range is so small that it cannot reach any other sensor in this example. Sensor C is not the destination node and the expiration length is 3, so it will rebroadcast the message by changing it to (A, C, B, 2). After sensor D receives the message, it is not the destination either and the expiration length is 2, so it will rebroadcast the message by changing it to (A, C, D, B, 1). When B receives the message, it sees that it is the destination. It knows by now that source A is its Out-neighbor and adds A to its Out-neighbor list. Also, it builds a "Path" message (A, C, D, B) and sends it to A. After A receives the "Path" message, it gets its reverse routing path to B: $A \rightarrow C \rightarrow D \rightarrow B$.

B. Routing Part

The nature of wireless communication is broadcasting. So the easiest and most reliable way to transmit a packet to the sink is flooding. However, flooding will cause serious communication overhead known as "flooding storm". In order to reduce overhead and achieve the assured delivery rate, we only choose a number of forwarding nodes based on historical statistics. Comparing to conventional routing protocols in WSNs, which ignore the existence of large numbers of asymmetric links, ProHet takes advantage of asymmetric links to route packets with high delivery ratio assurance.

In ProHet, two-hop neighborhood information model is used. Information in one-hop or more than two-hop neighborhood can also be used, we will justify why we adopt two-hop information in later simulations. Our basic idea is to choose a subset of two-hop receivers of a node which have high delivery probabilities as forwarding nodes, and choose the one-hop receivers that can cover the selected two-hop receivers to relay the message. The ProHet protocol contains three phases/algorithms: Selecting Nodes, Forwarding Messages, and Sending Acknowledgement. The Selecting Nodes algorithm chooses the subset of two-hop receivers and the corresponding one-hop receivers; the Forwarding Message algorithm forwards messages to the destination; and the Sending Acknowledgement algorithm sends back an "Acknowledgement" for a successful transmission and updates the delivery probabilities of forwarding nodes. The details are as follows:

Algorithm: Selecting Nodes

1: Source v calculates the probability threshold P_{th} using Condition (4) in Section V-C given assured delivery rate ρ.
2: v selects a subset of its two-hop receivers whose delivery probability $P_{delivery} \geq P_{th}$ into the set $SN_2(v)$;
3: v finds the minimal set of its one-hop receivers to cover all the nodes in $SN_2(v)$ by the following:
4: **repeat**
5: Add every $v \in N_1(v)$ to $SN_1(v)$, if there is a node in $SN_2(v)$ covered only by v;
6: Add $v \in N_1(v)$ to $SN_1(v)$, if v covers the largest number of nodes in $SN_2(v)$ that have not been covered;
7: If there is a tie, the choice is random;
8: **until** all the nodes in $SN_2(v)$ are covered.

In the Selecting Nodes algorithm, notation $N_1(v)$ denotes v's one-hop receivers and $N_2(v)$ denotes v's two-hop receivers. Node u *covers* v if u is an In-out-neighbor or In-neighbor of v. $SN_2(v)$ and $SN_1(v)$ denote v's selected two-hop and one-hop receivers, respectively.

Let's use an example to explain the Selecting Nodes algorithm in Figure 3. Suppose V (marked in red) has a packet to send. We use the algorithm to select v's two-hop (will be marked in black) and one-hop receivers (will be marked in blue). If there is a directional link $A \rightarrow B$ or a bidirectional link $A \leftrightarrow B$, it means A covers B. First, suppose six of V's two-hop receivers H, J, K, M, N, P are selected into $SN_2(v)$ because their delivery probabilities are no less than P_{th} given ρ. Next, we select the minimal set of V's one-hop receivers to cover all of the nodes in $SN_2(v)$ as follows: Node H is only

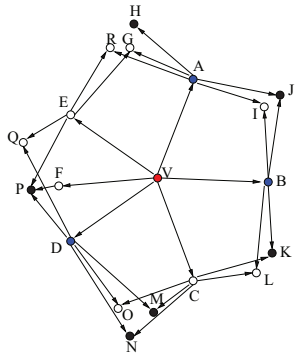


Fig. 3. An example of the Selecting Nodes Algorithm.

covered by one one-hop receiver A. So, A is selected into $SN_1(v)$. Node A also covers J. Next, the one-hop receiver that covers the most of the remaining nodes in $SN_2(v)$ is node D. So, it is also put into $SN_1(v)$. Now, the only node left in $SN_2(v)$ is K. It is covered by both B and C. Since neither B nor C covers any other remaining node in $SN_2(v)$, we can choose either one of them to cover K. Suppose we choose B, so finally $SN_1(v) = \{A, B, D\}$.

Algorithm: Forwarding Messages

1: The current forwarding node v broadcasts the packet P containing $SN_1(v)$, $SN_2(v)$, and the message to be delivered to the sink; the forwarding number N_f of v is increased by one;
2: If a receiver $u \in N_1(v)$ is in $SN_1(v)$, it will rebroadcast P, increase its forwarding number N_f by one and attach u's ID in P as a forwarding node in the path;
3: **repeat**
4: Set node $t \in SN_2(v)$ as the new source and apply the Selecting Nodes and Forwarding Message algorithms;
5: **until** P reaches the sink.

Algorithm: Sending Acknowledgement

1: When a packet P reaches the sink, the sink sends an acknowledgement P_{ack} to all the forwarding nodes on the path. The later arrived copies of P are dropped.
2: When a forwarding node m receives P_{ack}, it increases its N_d by one, and
3: **if** its previous node t is its In-out-neighbor **then**
4: it sends P_{ack} directly to t;
5: **else if** m has a reverse path to t **then**
6: m sends P_{ack} to t via the reverse path of the asymmetric link $t \rightarrow m$;
7: **else**
8: m simply drops P_{ack}
9: **end if**

Next, any forwarder will run the Forwarding Messages algorithm, where the forwarding number N_f is recorded.

After the message reaches the sink, the sink will send back an acknowledgement P_{ack} to all the forwarding nodes on the path using the Sending Acknowledgement algorithm. Due to

the asymmetric links, the reverse paths may be used. On the way to send back P_{ack}, the delivery number N_d is recorded and the node's delivery probability $P_{delivery}$ is obtained using Formula (1). Every forwarding node refreshes its $P_{delivery}$ each time a message is sent from a source to the sink and the acknowledgement comes back.

V. Analysis

The key point for a node v to select its two-hop receivers is the value of the probability threshold P_{th} given an assured delivery rate ρ ($0 \leq \rho \leq 1$). In this section, we first give an upper-bound of ρ, then prove that ρ can be achieved if it is within its upper-bound and present a method to calculate P_{th}.

A. Upper-bound of ρ

Obviously, if the delivery rate ρ is set too high and the delivery probabilities of nodes in the network are too low and the network is sparse, then ρ can not be achieved. So we need to find the upper-bound of ρ to make it possible to achieve.

Suppose node v has a total of m two-hop receivers whose delivery probabilities obtained by Formula (1) in non-increasing order are $p_1, p_2, \cdots, p_m$. The highest delivery rate v can achieve is when $P_{th} = p_m$, which means all of its m two-hop receivers are selected into the forwarding set. Then the following is true:

$$\begin{aligned} &1-(1-p_1)(1-p_2)\cdots(1-p_m) \geq 1-(1-p_{min})^m \\ &\geq 1-(1-p_{min})^{out\text{-}d_{min}} \geq \rho \end{aligned} \tag{2}$$

In the above, p_{min} is the minimum delivery probability of nodes in the whole network. Thus, $p_1, p_2, \cdots, p_m \geq p_{min}$. The value $out\text{-}d_{min}$ is the minimum m in the whole network. So $m \geq out\text{-}d_{min}$. The values of p_{min} and $out\text{-}d_{min}$ can be known after a network has been set up and several rounds of packet delivery have been done. So ρ is upper-bounded by

$$\rho \leq 1-(1-p_{min})^{out\text{-}d_{min}} \tag{3}$$

That means, the delivery rate ρ that can be achieved depends on the nodes' delivery probabilities and the network density.

B. Condition for assured delivery rate ρ

Theorem 1: ProHet guarantees the assured delivery rate ρ if Condition (3) holds.

Proof: According to ProHet, a node will select k out of its m two-hop receivers whose delivery probabilities are greater or equal to P_{th} as its forwarding nodes. Assume the delivery probabilities of the m two-hop receivers are $p_1, p_2, \cdots, p_m$ in non-increasing order. In the worst case, P_{th} takes the minimum value p_m and $k = m$, which means all of its m two-hop receivers are selected into the forwarding set. Even when that happens, we know that Condition (2) is true. So ρ can be achieved. Better than that, ρ can be achieved with a larger P_{th} and fewer two-hop receivers as forwarders. ■

Therefore, in this paper, when we set the assured delivery rate ρ in ProHet, we make sure that Condition (3) is satisfied.

C. Determining the value of P_{th}

Now we show how to determine the value of P_{th}. Suppose there are m two-hop receivers in v's two-hop neighborhood whose delivery probabilities are $p_1, p_2, \cdots, p_m$. And there are k ($k \geq 1$) out of m two-hop receivers whose delivery probabilities $p_1, p_2, \ldots, p_k$ are no less than P_{th}. They will be selected by v to forward a packet. Without loss of generality, we assume $p_1 \geq p_2 \geq \cdots \geq p_k \geq P_{th}$. In order to reach the delivery rate ρ, the following must be satisfied:

$$1-(1-p_1)(1-p_2)\cdots(1-p_k) \geq 1-(1-P_{th})^k \geq \rho$$

$$\text{So, } P_{th} \geq 1-(1-\rho)^{\frac{1}{k}} \tag{4}$$

Also from the analysis, we know that P_{th} has a maximum value of p_1 and a minimum value of p_m. Setting $P_{th} = p_1$ means that only the two-hop receiver with the highest delivery probability is selected into the forwarding set to reach the delivery rate. And setting $P_{th} = p_m$ indicates that all of the two-hop receivers need to be selected to reach ρ. In all other cases, P_{th} should have a value between the two. The procedure to obtain P_{th} from Condition (4) is as follows: First, the delivery probabilities of v's m two-hop receivers are ordered non-increasingly in the list $p_1, p_2, \cdots, p_m$. Then, we initialize P_{th} to p_1 and $k = 1$, that is, only the node with the highest delivery probability is considered. We check if Condition (4) is true or not. If the condition is true, $\{P_{th} = p_1 \text{ and } k = 1\}$ is our solution. Otherwise, we add the node with the second highest delivery probability, and set $k = 2$ and $P_{th} = p_2$. Again, we check if the condition is satisfied. This process continues by adding the next node in the list, increasing k by one and setting P_{th} to p_k until Condition (4) is satisfied. Finally, P_{th} is set to p_k and the number of two-hop receivers chosen into the forwarding set is k.

VI. Simulations

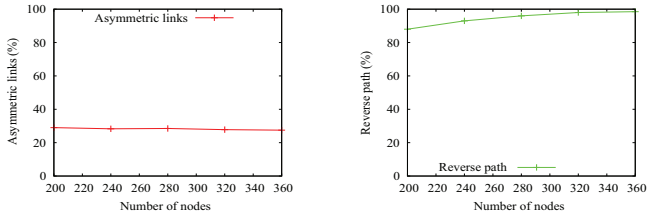
In this section, we justify several design choices in ProHet and evaluate its performance by comparing it with the following three protocols using a self-written simulator in Java.

- Flooding, the conventional algorithm.
- Random-K, in which random K one-hop receivers are selected to forward packets.
- TopRatio-K, in which K one-hop receivers with the highest delivery probabilities are selected to relay packets.

A. Simulation Setup

We used the following metrics to evaluate the performance of the protocols:

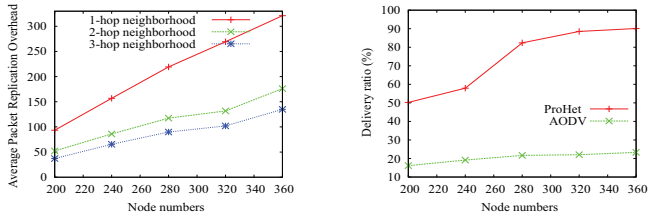
- Delivery ratio: the ratio of the number of packets successfully delivered to the total number of packets generated.
- Average hops: the average hops of a packet successfully sent from a source to a sink.
- Average packet replication overhead: the average number of packet replications used to successfully deliver a packet.
- Average control message overhead: the average number of control messages which include all of the communication messages (except the main packet) to identify neighbors, find reverse paths and update nodes' delivery probabilities needed to successfully deliver a packet.



(a) The percentage of asymmetric links in the network

(b) The percentage of reverse paths found within 3 hops

Fig. 4. Percentages of asymmetric links and reverse paths found with the node transmission diversity of $20m$.



(a) Comparison of replication overhead using one-hop, two-hop, and three-hop models in ProHet

(b) Comparison of ProHet with AODV in delivery ratio

Fig. 5. Comparison of delivery ratio of ProHet using different information models and with that of AODV.

In our experiments, nodes were deployed in a $500m \times 500m$ area. To diversify nodes' transmission ranges, we used the idea in [24] to let a node have one of the three transmission ranges: the *minimum*, the *normal*, and the *maximum*. The normal transmission range is the average of the minimum and the maximum transmission ranges. We set the normal transmission range, also the default, to $50m$. *Node transmission diversity* is the difference between the maximum and the minimum ranges. The link loss rate of each link was randomly set between 0% and 20% initially. In both Random-K and TopRatio-K algorithms, the value of K was set to 5. To implement message sending and receiving, a virtual concept of *time slots* was used. In each time slot, we randomly chose a sensor to generate a new message and let it send the message to the sink. Each node used a buffer to cache packets from other nodes. We assumed that all of the packets in the buffer can be transmitted to the next-hop node within one time slot. The simulation time was set to $10,000$ time slots. During the experiments, we randomly generated 20 different deployments of heterogeneous sensor nodes and calculated the average performance in the simulation results.

B. Experimental Results

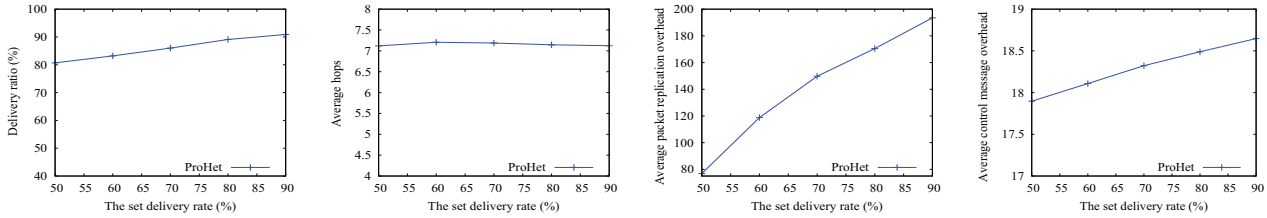
We first studied the percentage of the asymmetric links in the network and the percentage of these asymmetric links having a reverse path within 3 hops. Fig. 4(a) shows that about 30% of the total links in the network are asymmetric when the node transmission diversity is set to $20m$. Fig. 4(b) indicates that over 90% of the asymmetric links can find their reverse paths within 3 hops using our algorithm, which justifies setting the expiration length to 3 is good enough to find most of the reverse paths for the asymmetric links.

In order to explain that using the two-hop neighborhood information model is reasonable, we compared the performance of the one-hop, two-hop, and three-hop information models. We used the same three transmission ranges for the nodes, set the node transmission diversity to $20m$ and set the assured delivery rate to 80%. We found that the delivery ratio of the one-hop information has a marginal improvement over those of the two-hop and three-hop information models because it is more like flooding. However, the packet replication overhead of the one-hop information is significantly higher than those of the two-hop and three-hop information models as shown in Fig. 5(a). Considering the significant replication overhead in the one-hop model and the communication overhead among neighbors in the k-hop ($k \geq 3$) model, we think using two-hop neighborhood information model is appropriate.

In order to illustrate the improvement of delivery ratio, we compared ProHet with AODV. Though they have several differences: AODV is for ad hoc wireless networks while ProHet is for heterogeneous sensor networks; AODV assumes symmetric communication links while ProHet deals with asymmetric ones, both of them use reverse path in routing and have some similarity in design methodology. We used the same three transmission ranges for the nodes, set the transmission diversity to $20m$ and set ProHet's assured delivery rate to 80%. From the results in Fig. 5(b), we can see that the delivery ratio of AODV cannot be guaranteed because it does not use asymmetric links and does not set achieving assured delivery rate as its design goal whereas in ProHet, with the increase of node numbers and thus more connections, it can reach the assured delivery rate and exceed.

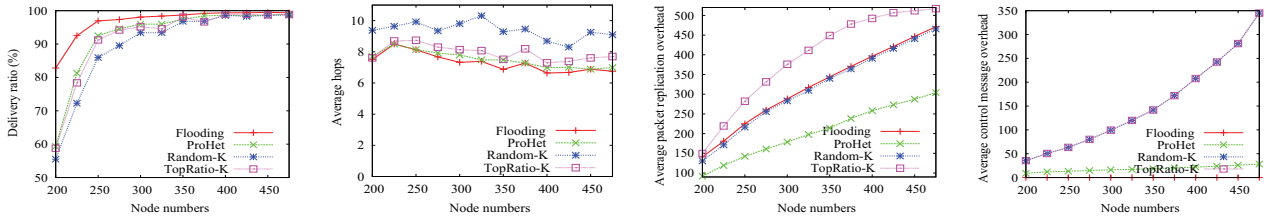
We also conducted simulations to reflect the impact of ρ on the performance of ProHet (see Fig. 6). We set the node number to 350 and the node transmission diversity to $20m$. Fig. 6(a) shows that with ρ going from 50% to 90%, the actual delivery ratio achieved by ProHet is greater than the set value. This proves that ProHet can guarantee the assured delivery rate. In Fig. 6(b), with the increase of ρ, the average hops remain almost a constant, implying that the latency of ProHet is under control. In Fig. 6(c), the average packet replication overhead increases when ρ increases, which indicates more duplications are produced to achieve the assured delivery rate. But the control message overhead is only slightly increased in Fig. 6(d), which means ProHet does not generate a lot more control overhead to reach a higher delivery ratio.

Finally we compared ProHet with the three protocols (see Fig. 7). The node transmission diversity was set to $20m$ and the assured delivery rate was 95%. Fig. 7(a) verifies that Flooding has the highest delivery ratio. ProHet's higher delivery ratio than those of TopRatio-K and Random-K means that a careful selection of forwarders based on P_{th} is better than selecting the top K or randomly. Also, the increase of delivery ratios of all the strategies with the increase of node numbers implies more connections between nodes can result in more successful deliveries. Fig. 7(b) confirms that Flooding has the lowest hops to deliver packets. ProHet's near-lowest-hops indicate its low latency. Fig. 7(c) shows that ProHet has the least average packet replication overhead, which proves that the probabilistic strategy to choose forwarding nodes in the two-hop neighborhood is effective to removing many redundant packets in the delivery process. Flooding has no



(a) Delivery ratio (b) Average hops (c) Average packet replication overhead (d) Average control message overhead

Fig. 6. Impact of ρ on ProHet's performance with 350 nodes and node transmission diversity of $20m$.



(a) Comparison of delivery ratio (b) Comparison of average hops (c) Comparison of average packet replication overhead (d) Comparison of average control message overhead

Fig. 7. Comparison of ProHet with Flooding, Random-K and TopRatio-K with node transmission diversity of $20m$.

control overhead in Fig. 7(d) due to not keeping neighbor information in routing. ProHet's control overhead is much lower than those of Random-K and TopRatio-K since it establishes neighborhood information every two hops instead every one hop in the routing process. In summary, ProHet can achieve similar performance of delivery ratio and latency to those of Flooding, but with a much lower replication overhead and a low control overhead.

VII. Conclusion

In this paper, we proposed ProHet, a probabilistic routing protocol, to deal with asymmetric links, reliability and scalability issues in WHSNs. It addresses asymmetric links by finding reverse paths and improves reliability and scalability by choosing forwarders based on historical statistics using local information. We showed that ProHet can achieve assured delivery rate by theoretical analysis if the assured delivery rate is set within its upper-bound. The efficiency of ProHet was evaluated by our extensive simulations. In our future work, we will design more efficient routing protocols in WHSNs.

Acknowledgment

This work was partially supported by NSF grant CNS 0835834; National Natural Science Foundation of China grants 61073028 and 61021062; and National Basic Research Program of China (973) grant 2009CB320705.

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

References

[1] S. Ben Alla, A. Ezzati, A. Beni Hssane, and M. L. Hasnaoui, "Hierarchical adaptive balanced energy efficient routing protocol (HABRP) for heterogeneous wireless sensor networks," *2011 International Conf. Multimedia Comput. Syst.*

[2] C. Adjih, P. Jacquet, and L. Viennot, "Computing connected dominated sets with multipoint relays," Technical Report, INRIA, Oct. 2002.

[3] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayicrci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–116, 2002.

[4] C. L. Barrett, S. Eidenbenz, L. Kroc, M. V. Marathe, and J. P. Smith, "Parametric probabilistic sensor network routing," in *Proc. 2003 ACM International Conf. Wireless Sensor Netw. Appl.*, pp. 122–131.

[5] A. Behzadan, A. Anpalagan, and B. Ma, "Prolong network lifetime via nodal energy balancing in heterogeneous wireless sensor networks," in *Proc. 2011 IEEE ICC*, pp. 1–5.

[6] X. Chen and J. Shen, "Improved schemes for power-efficient broadcast in ad hoc networks," *International J. High Performance Comput. Netw.*, vol. 4, no. 3/4, p. 198–206, 2006.

[7] X. Chen, W. Y. Qu, H. L. Ma, and K. Q. Li, "A geography-based heterogeneous hierarchy routing protocol for wireless sensor networks," in *Proc. 2008 IEEE International Conf. High Performance Comput. Commun.*, pp. 767–774.

[8] B. Deb, S. Bhatnagar, and B. Nath, "Information assurance in sensor networks," in *Proc. 2003 ACM International Conf. Wireless Sensor Netw. Appl.*, pp. 160–168.

[9] X. Du and F. Lin, "Designing efficient routing protocol for heterogeneous sensor networks," in *Proc. 2005 IEEE International Performance Comput. Commun. Conf.*, pp. 51–58.

[10] B. Elbhiri, R. Saadane, and D. Aboutajdine, "Stochastic distributed energy-efficient clustering (SDEEC) for heterogeneous wireless sensor networks," *ICGST International J. Comput. Netw. Internet Research*, vol. 9, no. II, pp. 11–17, 2009.

[11] I. A. Essa, "Ubiquitous sensing for smart and aware environments," *IEEE Personal Commun.*, vol. 7, pp. 47–49, 2000.

[12] D. L. Guidoni, A. Boukerche, L. A. Villas, F. S. H. Souza, R. A. F. Mini, and A. A. F. Loureiro, "A framework based on small world features to design HSNS topologies with QoS," *2012 IEEE Symp. Comput. Commun.*

[13] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 2008 Hawaii International Conf. Syst. Sci.*

[14] Y. F. Hu, W. Z. Li, X. Chen, X. Chen, S. L. Lu, and J. Wu, "A probabilistic routing protocol for heterogeneous sensor networks," in *Proc. 2010 IEEE International Conf. Netw., Archit., Storage*.

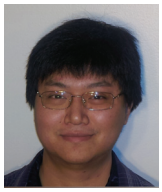
[15] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *Proc. 2000 International Conf. Mobile Comput. Netw.*, pp. 56–67.

[16] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *Proc. 2000 International Conf. Mobile Comput. Netw.*, pp. 243–254.

[17] L. Lazos, R. Poovendran, and J. A. Ritcey, "Probabilistic detection of mobile targets in heterogeneous sensor networks," in *Proc. 2007 International Conf. Inf. Process. Sensor Netw.*, pp. 519–528.

[18] A. M. Mainwaring, D. E. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. 2002 ACM International Workshop Wireless Sensor Netw. Appl.*
[19] A. Manjeshware and D. P. Agrawal, "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks," in *Proc. 2001 International Parallel Distributed Process. Symp.*, pp. 2009–2015.
[20] J. Newsome and D. X. Song, "GEM: graph embedding for routing and data-centric storage in sensor networks without geographic information," in *Proc. 2003 International Conf. Embedded Netw. Sensor Syst.*, pp. 76–88.
[21] M. Perillo, Z. Cheng, and W. Heinzelman, "An analysis of strategies for mitigating the sensor network hot spot problem," in *Proc. 2005 IEEE MobiQuitous*, pp. 474–478.
[22] D. A. Patterson, "Rescuing our families, our neighbors, and ourselves," *Commun. ACM*, vol. 48, no. 11, pp. 29–31, 2005.
[23] J. M. Rabaey, M. J. Ammer, J. L. da Silva Jr., D. Patel, and S. Roundy, "Picoradio supports ad hoc ultra-low power wireless networking," *IEEE Comput.*, vol. 33, no. 7, pp. 42–48, 2000.
[24] V. Ramasubramanian and D. Mossé, "BRA: a bidirectional routing abstraction for asymmetric mobile ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 1, pp. 116–129, 2008.
[25] A. Rao, C. H. Papadimitrious, S. Shenker, and I. Stoica, "Geographic routing without location information," in *Proc. 2003 Annual International Conf. Mobile Comput. Netw.*
[26] H. Rivas, T. Voigt, and A. Dunkels, "A simple and efficient method to mitigate the hot spot problem in wireless sensor networks," in *2006 Workshop Performance Control Wireless Sensor Netw.*
[27] R. R. Rout, S. K. Ghosh, and S. Chakrabarti, "A network coding based probabilistic routing scheme for wireless sensor network," in *Proc. 2010 Intertional Conf. Wireless Commun. Sensor Netw.*, pp. 1–6.
[28] D. Tian and N. D. Georganas, "Energy efficient routing with guaranteed delivery in wireless sensor networks," in *Proc. 2003 IEEE Wireless Commun. Netw. Conf.*, pp. 1923–1929.
[29] G. Q. Wang, Y. C. Ji, D. C. Marinescu, and D. Turgut, "A routing protocol for power constrained networks with asymmetric links," in *Proc. 2004 ACM International Workshop Performance Evaluation Wireless Ad Hoc, Sensor, Ubiquitous Netw.*, pp. 69–76.
[30] N. Wang and C. H. Chang, "Performance evaluation of geographic probabilistic flow-based spreading routing in wireless sensor networks," in *Proc. 2007 ACM International Workshop Performance Evaluation Wireless Ad Hoc, Sensor, Ubiquitous Netw.*, pp. 32–38.
[31] J. Wu, "An enhanced approach to determine a small forward node set based on multipoint relays," in *Proc. 2003 IEEE Veh. Technol. Conf. – Fall*, vol. 4, pp. 2774–2777.
[32] M. D. Yarvis, N. Kushalnagar, H. Singh, A. Rangarajan, Y. Liu, and S. Singh, "Exploiting heterogeneity in sensor networks," in *Proc. 2005 IEEE INFOCOM*, pp. 878–890.
[33] F. Ye, H. Y. Luo, J. Cheng, S. W. Lu, and L. X. Zhang, "A two-tier data dissemination model for large-scale wireless sensor networks," in *Proc. 2002 International Conf. Mobile Comput. Netw.*, pp. 148–159.
[34] Q. Zhang and W. G. Chang, "A power efficiency routing protocol for heterogeneous sensor networks," in *Proc. 2008 International Conf. Wireless Commun., Netw. Mobile Comput.*, pp. 1–4.
[35] W. Y. Zhang, X. J. Du, J. Wu, S. D. Soysa, and Y. Liu, "Near-minimum-energy routing in heterogeneous wireless sensor networks," in *Proc. 2010 IEEE GLOBECOM*, pp. 1–5.

Xiao Chen is an associate professor of Computer Science at Texas State University. She received her Ph.D. Degree in Computer Engineering from Florida Atlantic University. Her research interests are in the areas of sensor and ad hoc wireless networks. She has served as an associate editor, program committee member, session chair, and reviewer of numerous international journals and conferences.



Zanxun Dai is a master student of Computer Science at Texas State University-San Marcos. He received his Bachelor's Degree in Software Engineering from Harbin Institute of Technology, China. His research interests are in the areas of sensor networks and ad hoc wireless networks.

Wenzhong Li received his B.S. and Ph.D. degree from Nanjing University, China, both in computer science. He is now an associate professor of Computer Science in Nanjing University. His research interests include wireless networks, pervasive computing, and social networks. He has published over 30 papers at international conferences and journals. Dr. Li was also the winner of the Best Paper Award of ICC 2009. He is a member of IEEE and ACM.

Yuefei Hu received his B.S. degree from Jilin University and his M.S. degree from Nanjing University, China, both in computer science. He is now working in the Zhengzhou Commodity Exchange, one of the four futures exchanges in China. His research interests include wireless sensor networks and delay tolerant networks.

Jie Wu is Chair and Laura H. Carnell Professor in the Department of Computer and Information Sciences at Temple University, USA. Prior to joining Temple, he was a program director at the National Science Foundation and Distinguished Professor at Florida Atlantic University. His research interests include wireless networks, mobile computing, routing protocols, fault-tolerant computing, and interconnection networks. Dr. Wu's publications include over 600 papers in scholarly journals, conference proceedings, and books. He has served on several editorial boards, as general chair and co-chair of many top journals and conferences. He was an IEEE Computer Society Distinguished Visitor and the chair for IEEE Technical Committee on Distributed Processing. He is an ACM Distinguished Speaker and a Fellow of the IEEE.

Hongchi Shi is a Professor and the Chair of the Computer Science at Texas State University. Prior to joining Texas State University, he was an Assistant/Associate/Full Professor of Computer Science and Electrical and Computer Engineering at the University of Missouri. He obtained his Ph.D. in Computer and Information Sciences from University of Florida. His research interests include parallel and distributed computing, wireless sensor networks, neural networks, and image processing. He has served on many organizing and/or technical program committees of international conferences. He is a member of ACM and a senior member of IEEE.

Sanglu Lu received her B.S., M.S., and Ph.D. degrees from Nanjing University in 1992, 1995, and 1997, respectively, all in computer science. She is currently a professor in the Department of Computer Science and Technology and the State Key Laboratory for Novel Software Technology. Her research interests include distributed computing, pervasive computing, and wireless networks. She is a member of IEEE and ACM.