2010 Fifth IEEE International Conference on Networking, Architecture, and Storage

# A Probabilistic Routing Protocol for Heterogeneous Sensor Networks

Yuefei Hu[1], Wenzhong Li[1], Xiao Chen[2], Xin Chen[1], Sanglu Lu[1], Jie Wu[3]
[1]State Key Laboratory for Novel Software Technology, Nanjing University
[2]Department of Computer Science, Texas State University, San Marcos, USA
[3]Department of Computer and Information Sciences, Temple University, Philadelphia, PA
Email: {huyuefei, lwz}@dislab.nju.edu.cn, xc10@txstate.edu, jiewu@temple.edu

***Abstract*—The past five years witnessed a rapid development in wireless sensor networks, which have been widely used in military and civilian applications. Due to different requirements in their application environment, sensors with different capacities, power, and so on are deployed. Data routing in such heterogeneous sensor networks is a challenging task. On one hand, the heterogeneous features bring about the diversity in their transmission ranges, which subsequently lead to asymmetric links in the communication graph. As a result, conventional routing strategies based on undirected graphs become unsuitable. On the other hand, sensors communicate with each other through intermittent asymmetric links. It is important to provide assurable delivery rate for mission critical applications. In this paper, we propose *ProHet*: a *Pro*babilistic routing protocol for *Het*erogeneous sensor networks, which can deal with asymmetric links well and work in a distributed manner with low overhead and assurable delivery rate. The ProHet protocol first produces a bidirectional routing abstraction by finding a reverse routing path for every asymmetric link. Then, it uses a probabilistic strategy to choose forwarding nodes based on historical statistics, which is shown to achieve assurable delivery rate by theoretical analysis. Extensive simulations are conducted to verify the efficiency of the proposed protocol.**

## I. Introduction

Recent advances in wireless communication technologies and electronics have paved the way for developing low-cost sensor networks. Sensor networks have a wide range of applications in military and the daily lives of individuals. In military environments, they can be used in command, control, communications, computing, intelligence, surveillance, reconnaissance, and target tracking systems [1], whereas in civilian applications, they can be used in environment monitoring [2], home health care [3], intelligent homes [4], disaster rescuing [5], and self-touring systems [6].

In sensor networks, 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 [7], [8]. Therefore, sensors may not have the same sensing capability and communication range. Here we just take their diverse transmission ranges brought about by their heterogeneity into account.

After the heterogeneous sensors have completed data collections, one major issue is how to route this data to the destination (mostly it is the sink in sensor networks) efficiently [9], [10], [11]. While these heterogeneous sensors have different transmission ranges, there will be asymmetric links in the communication graph. For example, if node A can reach node B, but B cannot reach A, the directed link from A to B is asymmetrical. Just because of asymmetry, the common undirected graph generated after abstraction turns into a directed graph, and then the off-the-shelf routing protocols for general wireless sensor networks cannot apply or work with higher overhead [12]. Designing efficient routing protocols for these heterogeneous sensor networks is challenging:

- In heterogeneous sensor networks, data is transmitted through asymmetric and unreliable links, therefore a reliable routing scheme that addresses assurable delivery rate and delay issues is important to guarantee the performance of the networks.
- Due to the resource limitation in wireless sensors, the routing protocol should be energy efficient and with low overhead.
- For scalability and robustness purposes, there should not be a central entity which computes the routing information. In other words, the routing scheme should be totally distributed.

In this paper, we propose **ProHet**: a **Pro**babilistic routing protocol for **Het**erogeneous sensor networks, which can handle asymmetric links well and work in a distributed manner with low overhead and assurable delivery rate. Specifically, we make the following contributions to deal with the heterogeneity and reliability issues:

- We adopt a strategy similar to [12] to find a reverse routing path for every asymmetric link and use a bidirectional routing abstraction for the off-the-shelf routing protocol.
- We propose a probabilistic strategy to choose forwarding nodes based on historical statistics, which is shown to achieve assurable delivery rate by theoretical analysis.
- We conduct extensive simulations to verify the efficiency of the proposed protocol.

The rest of the paper is organized as follows: Section II references the related work. Section III presents preliminaries and definitions used in this paper. Section IV proposes the ProHet protocol. Section VI evaluates the performance of the

978-0-7695-4134-1/10 $26.00 © 2010 IEEE
DOI 10.1109/NAS.2010.42

IEEE computer society

ProHet protocol. And conclusion is drawn in Section VII.

## II. Related Work

In this section, we give an overview of routing algorithms in heterogeneous sensor networks and probabilistic routing strategies.

### A. Routing in Heterogeneous Sensor Networks

Routing in homogeneous sensor networks has been well studied and many routing protocols have been proposed [9], [10], [11], [13], [14], [15], [16], [17], [18]. In these protocols, all sensor nodes have the same capabilities in terms of communication, computation, energy supply, reliability, etc. However, in applications such as aforementioned heterogeneous sensors with different capabilities may be deployed. It is reported in [7] that when properly deployed, heterogeneity can triple the average delivery rate and provide a five-fold increase in the network lifetime. Routing in heterogeneous sensor networks should be rethought about. Simply using the routing protocols in homogeneous sensor networks does not take advantage of the more capable sensors and does not work well.

In the literature, there are a few routing protocols designed for heterogeneous sensor networks [19], [20], [21], [22], [23]. The sensors in the heterogeneous networks are categorized into powerful and less powerful ones. Sensors form clusters, with the powerful ones being the cluster heads. There are two routing protocols used: intracluster and intercluster. The intracluster protocol is used to route messages between less powerful nodes and their clusterheads. And the intercluster protocol is used to route messages between clusterheads.

In this paper, we utilize the idea in [12] to deal with asymmetric links through establishing corresponding reverse routing paths. We also propose a routing protocol for heterogeneous sensor networks that is only based on local information.

### B. Probabilistic Routing Strategies

The probabilistic routing strategy in wireless sensor networks is not a new topic and there are various studies about it. The authors in [24] 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 [25] 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 [26], information has different delivery probabilities according to their criticality to end users. 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 as 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 assurable delivery rate.

## III. Preliminary

### A. Definitions of Nodes' Neighbor Relationships

A heterogeneous sensor network 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 the In-neighbor of $B$ (or $A$), and $B$ (or $A$) is the 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).

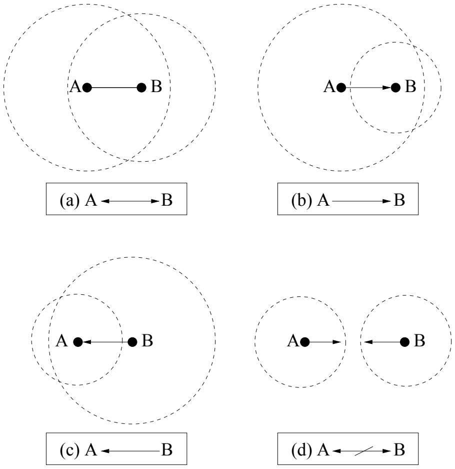


Fig. 1. The neighbor relationships between two nodes $A$ and $B$. (a) $A$ and $B$ are each other's In-out-neighbor; (b) $A$ is the In-neighbor of $B$ and $B$ is the Out-neighbor of $A$; (c) $B$ is the In-neighbor of $A$ and $A$ is the Out-neighbor of $B$; (d) $A$ and $B$ are non-neighbors

### B. Definitions of Probabilistic Delivery

Before presenting the ProHet protocol, we want to give three definitions related to a node: the *one-hop receiver*, the *two-hop receiver*, and the *delivery probability*. 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. Figure 2 gives an example of the one-hop and two-hop receivers of node $S$. 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:

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

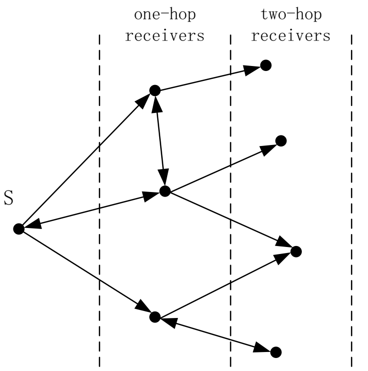


Fig. 2. One-hop and two-hop receivers of node $S$

## IV. The ProHet Protocol

In this section, we present the ProHet protocol, which has two parts: finding a reverse path for asymmetric links and building routing algorithms. The details are as follows:

### A. Finding A Reverse Path for Asymmetric Links

Finding a reverse routing path is the first step to handle asymmetric links in heterogeneous sensor networks. The study of [12] shows that a significant percent of links in heterogeneous sensor networks 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. Based on their observation, we can find a reverse path for each asymmetric link by tracing back three hops.

The process of finding a reverse routing path contains the following three stages:

1. Initialization
   a) Every node in the network broadcasts a "Hello" message.
   b) If two nodes $A$ and $B$ can receive each other's "Hello" message and the corresponding "Ack" of the "Hello" message, then each adds the other to its In-out-neighbor list.
   c) If $A$ receives $B$'s "Hello" message, but not the "Ack" 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 perform the next step to find the reverse routing path to $B$.
2. $A$ 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 to which it wants to find the reverse path (e.g. "$B$")), and an expiration length of 3 hops.
3. If some node $C$ receives a "Find" message,

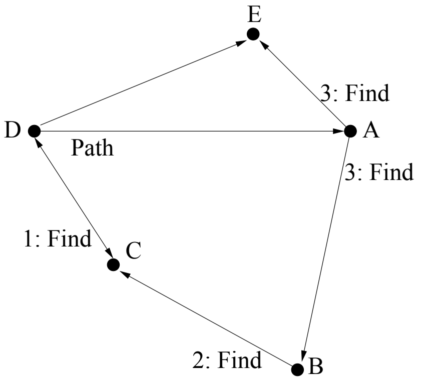


Fig. 3. An example of finding a reverse routing path

   a) if it is the destination node listed in the message, it will
      i) add the source node to its Out-neighbor list;
      ii) send the identified reverse routing path to the source node by a "Path" message containing the reverse route.
   b) if it is not the destination node and the expiration length is greater than 0, it will rebroadcast the message after the following modifications:
      i) decrease the expiration length by one;
      ii) append its own ID to the message.
   c) in all other cases, it will drop the message.

After this whole process ends, most nodes will establish their reverse routing paths to their In-neighbors. If a node receives more than one reverse routing paths to an In-neighbor, it chooses the shortest one.

We use the heterogeneous network in Figure 3 to explain this process. 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 the initiation stage, sensor $A$ gets sensor $D$'s "Hello" message, but it does not receive $D$'s "Ack" to its own "Hello" message. It knows that $D$ is its In-neighbor. Then, it starts to find a reverse routing path to $D$ by broadcasting a "Find" message ($A$, $D$, 3). The number before "Find" in the figure represents the expiration length, initially set to 3. The "Find" message is received by sensors $E$ and $B$. Sensor $E$ matches the case in 3(b) and will rebroadcast the message after decreasing the expiration length by one and appending its own ID to the message. But, $E$'s transmission range is so small that it cannot reach any other sensor in this example. Sensor $B$ is not the destination node and the expiration length is 3, so it will rebroadcast the message by changing it to ($A$, $B$, $D$, 2). After sensor $C$ receives the message, it is not the destination and the expiration length is 2, so it will rebroadcast the message by changing it to ($A$, $B$, $C$, $D$, 1). When $D$ 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 ($A$, $B$, $C$, $D$) and sends it to $A$.

After $A$ receives the "Path" message, it gets its reverse routing path to $D$: $A \to B \to C \to D$.

### B. Routing Algorithms

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 a "flooding storm". In order to reduce overhead and achieve assurable delivery rate, we only choose a number of forwarding nodes based on historical statistics. Comparing to conventional routing protocols in wireless sensor networks, which ignore the existence of large numbers of asymmetric links, ProHet takes the advantages of asymmetric links to route packets with high throughput and delivery ratio assurance.

Considering asymmetric links, using 2-hop information is helpful to design a reliable routing strategy. Information in more than a 2-hop neighborhood can also be used, which will require more message exchanges. Therefore, in this paper, we use 2-hop neighborhood information. The basic idea is as follows: we choose a set of two-hop receivers of a node, with high delivery probabilities as forwarding nodes, and then 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 Acknowledgement. The Selecting Nodes algorithm chooses the set of two-hop receivers and one-hop receivers; the Forwarding Message algorithm forwards messages to the destination; and the Acknowledgement algorithm sends back the "Ack" message for a successful transmission. The details are given in the following:

**Algorithm: Selecting Nodes**

```
Node $v$ calculates the probability threshold $P_{th}$ according to Eq. (2) in subsection IV-C, given assurable delivery rate $\rho$.
$v$ selects a fraction of its two-hop receivers whose delivery probability $P_{delivery}$ is higher than or equal to $P_{th}$ into the set $SN_2(v)$;
$v$ finds the minimal set of its one-hop receivers to cover all the nodes in $SN_2(v)$ by the following:
repeat
    Add every $v \in N_1(v)$ to $SN_1(v)$, if there is a node in $SN_2(v)$ covered only by $v$;
    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;
    If there is a tie, the choice is random;
until all the nodes in $SN_2(v)$ are covered.
```

In the Node Selection algorithm, notation $N_1(v)$ denotes $v$'s one-hop receivers and $N_2(v)$ denotes $v$'s two-hop receivers. Here, $v$ can get $N_1(v)$ by broadcasting a "Hello" message, and get $N_2(v)$ by collecting its one-hop neighbors' one-hop neighbor information. $u$ *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.

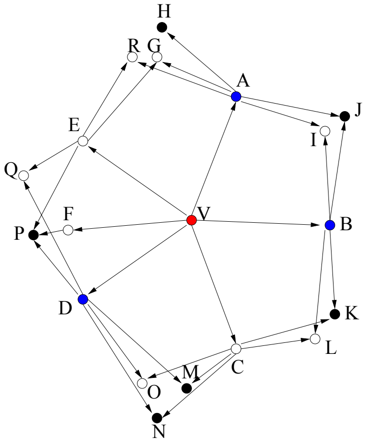


Fig. 4. An example of the Selecting Nodes Algorithm

We use an example to explain the Selecting Nodes algorithm. In the heterogenous sensor network in Figure 4, suppose $v$ is the node that has the packet (marked in red). We are going to 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 $A \to B$ or a bidirectional link $A \leftrightarrow B$, it means $A$ covers $B$. Suppose six $V$'s two-hop receivers $H, J, K, M, N, P$ are selected according to $P_{th}$ and put into $SN_2(v)$. Next, we select the minimal set of $V$'s one-hop receivers to cover $SN_2(v)$. Node $H$ is only 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**

```
The current forwarding node $v$ broadcasts the packet $P$ containing $SN_1(v)$, $SN_2(v)$, and the message needed to deliver to the sink; the forwarding number $N_f$ of $v$ is increased by one;
Any node $u \in N_1(v)$ rebroadcasts $P$ if it is in $SN_1(v)$, then increases its forwarding number $N_f$ by one and attaches $u$'s ID in $P$ as a forwarding node in the path;
repeat
    Set node $t$ in $SN_2(v)$ as the new "source" node "$v$" and apply Selecting Nodes and Forwarding Message algorithms;
until $P$ reaches the sink.
```

After Selecting Nodes phase is done, any source node and forwarding node will run the Forwarding Messages algorithm, where the forwarding number $N_f$ is recorded.

After the message reaches the sink, it will send back an

**Algorithm: Acknowledgement**

```
When the first copy of a packet $P$ reaches the sink node, the sink generates an acknowledgement $P_{ack}$ of $P$ to all the forwarding nodes on the path back to the source. The later arrived copies of $P$ are dropped.
When an intermediate node $m$ receives $P_{ack}$, it increases its $N_d$ by one, and
if its previous node $t$ is its In-out-neighbor, then
    it sends $P_{ack}$ directly to $t$;
else if $m$ has a reverse path to $t$, then
    $m$ sends $P_{ack}$ to $t$ via the reverse path of the asymmetric link $t \rightarrow m$;
else
    $m$ simply drops $P_{ack}$
end if
```

acknowledgement $P_{ack}$ to all the forwarding nodes on the path. Because of the asymmetric links, the reverse path found in the last section is used. On the way to send back $P_{ack}$, the delivery number $N_d$ is recorded and the node's delivery probability $P_{delivery}$ can be obtained using Equation (1). The $P_{delivery}$ is refreshed in every forwarding node each time a message is sent from a source to the sink, then the sink sends back an acknowledgement to the source.

At the initial stage of running the routing protocol, every node's delivery probability does not exist. So, the ProHet protocol will work in a flooding manner like epidemic routing. After some rounds of packets delivery, each node's delivery number $N_d$ and forwarding number $N_f$ have values, so every node's delivery probability can be computed locally and timely. After the routing protocol has been running for a long time in the network, every node's delivery probability will become stable. Thus, the historical information of the network has been established and used.

### C. Analysis

In this section, we explain how to calculate $P_{th}$, which is used by a node to select its two-hop receivers in the Selecting Nodes algorithm. Suppose we set the assurable delivery rate to be $\rho$. There are $k$ two-hop receivers whose delivery probabilities $p_1$, $p_2$, $\cdots$, $p_k$ are greater than or equal to $P_{th}$. In other words, they will be selected by a node to forward a packet later. In order to reach the assurable delivery rate $\rho$, the following must be satisfied:

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

So,

$$P_{th} \geq 1-(1-\rho)^{\frac{1}{k}} \tag{2}$$

Next, we show how to get $P_{th}$ using Eq. 2. Suppose $\rho$ is set to a number in $(0,1)$ and there are $m$ two-hop receivers in a node's two-hop neighborhood. We order the delivery probabilities of all the two-hop receivers of a node in a non-increasing order ($p_1$, $p_2$, $\cdots$, $p_m$). Starting from $k=1$, we first set $P_{th}$ to $p_1$, that is, only the node with the highest delivery probability is chosen. Then, we use this value of $P_{th}$ to check if it satisfies Eq. 2 or not. If the equation is satisfied, $P_{th}=p_1$ and $k=1$ are our solution. Otherwise, we look to the first two nodes with the highest delivery probabilities by setting $k=2$ and $P_{th}=p_2$. Again, we need to check if the equation is satisfied. This process continues by increasing $k$ by one and setting $P_{th}$ to $p_k$ in each step, until Eq. 2 is finally satisfied. Then, $P_{th}$ is set to $p_k$ and the number of two-hop receivers chosen is $k$. Next, we prove that the $P_{th}$ chosen this way exists:

*Theorem 1:* $P_{th}$ calculated by the above method exists.

*Proof:* We define $out\text{-}d$ as the summation of one node's Out- and In-out-neighbors. $out\text{-}d_{min}$ is the minimum $out\text{-}d$ in the network and $out\text{-}d_{max}$ is the maximum $out\text{-}d$ in the network.

Therefore,

$$out\text{-}d_{min} \leq k \leq out\text{-}d_{max}$$

If we replace $k$ in Eq. (2) by $out\text{-}d_{min}$ and $out\text{-}d_{max}$, the following is true:

$$\begin{aligned} 1-(1-\rho)^{\frac{1}{out\text{-}d_{min}}} &\geq 1-(1-\rho)^{\frac{1}{k}} \\ &\geq 1-(1-\rho)^{\frac{1}{out\text{-}d_{max}}} \end{aligned}$$

According to Eq. 2,

$$\begin{aligned} p_{th} &\geq 1-(1-\rho)^{\frac{1}{k}} \\ p_{th} &\geq 1-(1-\rho)^{\frac{1}{out\text{-}d_{max}}} \end{aligned} \tag{3}$$

Therefore, $P_{th}$ has a minimum value of $1-(1-\rho)^{\frac{1}{out\text{-}d_{max}}}$. Since $out\text{-}d_{max}$ exists, $P_{th}$ exists.

■

## V. Simulations

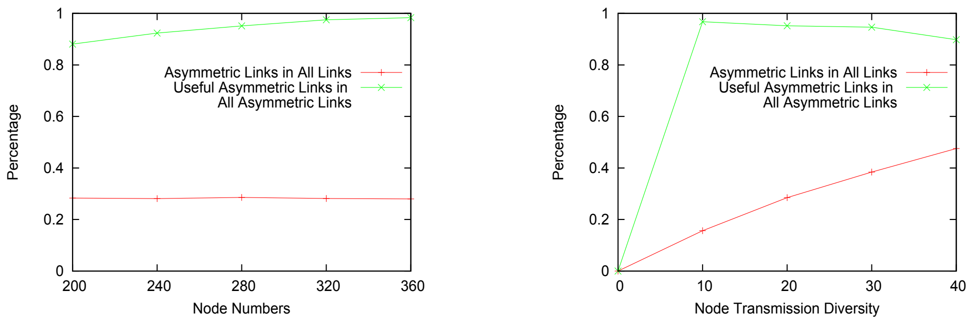
In this section, we evaluate the performance of ProHet protocol using a self-written simulator in the Java language. The following protocols are included for comparison:

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

### A. Simulation Setup

We use the following metrics to evaluate the performance of the proposed protocols:

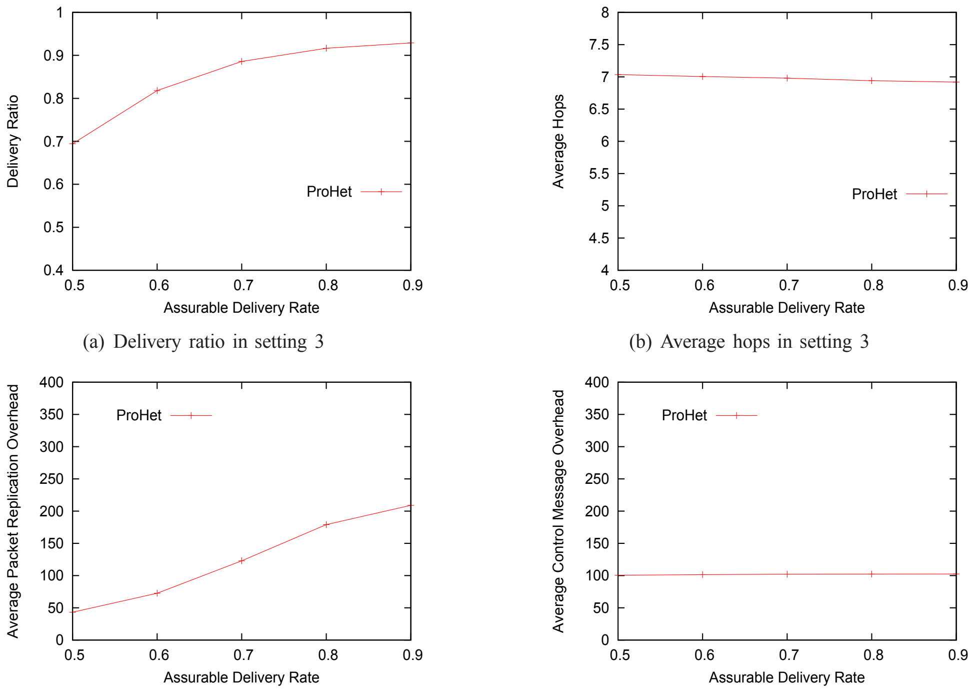
- Delivery ratio: the ratio of the number of packets successfully delivered to the total number of packets generated.
- Average hops: the average hops of a packet successfully sent from a source to a sink.



(a) Asymmetric links in setting 1

(b) Asymmetric links in setting 2

Fig. 5. Results of the usefulness of asymmetric links



(a) Delivery ratio in setting 3

(b) Average hops in setting 3

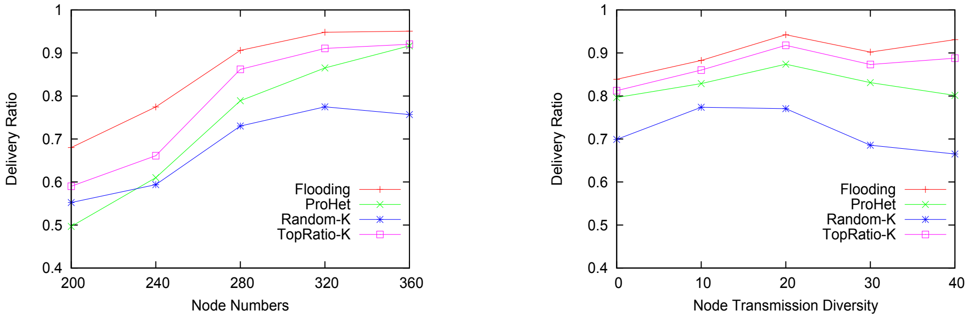
(c) Average packet replication overhead in setting 3

(d) Average control message overhead in setting 3

Fig. 6. Results of the impact of assurable delivery rate on defined metrics

- Average packet replication overhead: the average number of packet replications needed to successfully deliver a packet.
- Average control message overhead: the average number of control messages needed to successfully deliver a packet, which is much smaller than a packet's size.

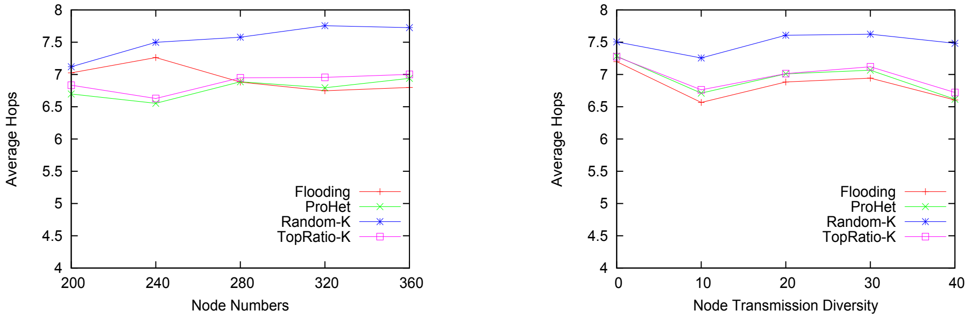
In our experiments, nodes are deployed in a $500m \times 500m$ area. To diversify the transmission ranges of nodes, we use the idea in [12], then a node can have one of the three transmission ranges: the *minimum*, the *normal*, and the *maximum* transmission ranges. The normal transmission range is the average of the minimum and the maximum transmission ranges. Here, we set the normal transmission range, which is also the default transmission range to $50m$. *Node transmission diversity* is defined as the difference between the maximum and the minimum ranges. We also consider the link loss and randomly set the link loss rate between 0% and 20%. In both Random-K and TopRatio-K algorithms, the value of K is set to 5. To implement message sending and receiving, a virtual concept of *time slots* is used. In each time slot, we randomly choose a sensor to generate a new message and let it send the message to the sink. Each node uses a buffer to cache packets from other nodes. Assume all packets in the buffer can be transmitted to the next-hop node within one time slot. The simulation time is set to 10000 time slots. During the experiments, we randomly generate 10 different deployments of heterogeneous sensor nodes and calculate the average performance in the simulation results.



(a) Delivery ratio in setting 1

(b) Delivery ratio in setting 2

Fig. 7. Results of delivery ratio



(a) Average hops in setting 1

(b) Average hops in setting 2

Fig. 8. Results of average hops

To study the performance of the ProHet protocol under different network parameters, we use the following system settings:

- Setting 1: node number is a variable. We set the number of nodes from 200 to 360 with a step of 40 and set the node transmission diversity to $20m$.
- Setting 2: node transmission diversity is a variable. We set the transmission diversity from $0m$ to $40m$ with a step of $10m$ and set the node number to 280.
- Setting 3: assurable delivery rate $\rho$ is a variable. We set the network's required assurable delivery rate from 0.5 to 0.9 with a step of 0.1, and set the node number to 360 and node transmission diversity to $20m$, respectively.

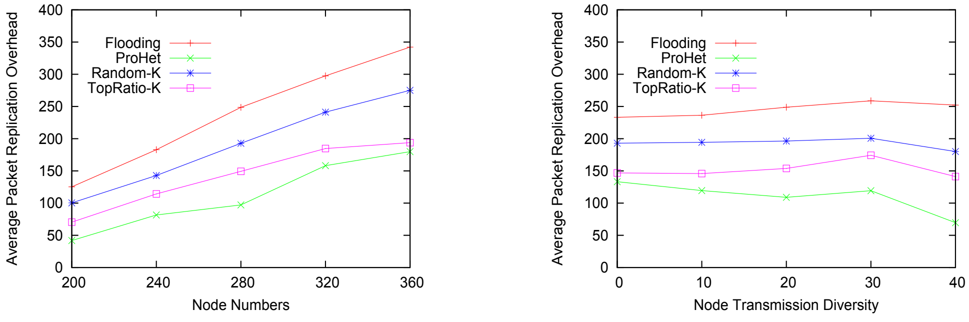
### B. Experimental Results

The impact of asymmetric links is shown in Figure 5. In this figure, the red line indicates the ratio of asymmetric links to all the links in the network, which shows that about 30% of the total links are asymmetric links. The green line shows the percentage of asymmetric links, which have a reverse path to their In-neighbors within 3 hops. We call them "useful links". From this figure, we can see that over 90% asymmetric links are useful links. This justifies that setting the expiration length to 3 in the algorithm is good enough for a node to find a reverse path to its In-neighbor in most conditions.

The impact of assurable delivery rate used in the ProHet protocol is illustrated in Figure 6. It is shown that with the increase of the assurable delivery rate from 0.5 to 0.9, the delivery ratio increases. However, the average hops per packet remains constant, which means it is not sensitive to the assurable delivery rate. The average packet replication overhead increases when the assurable delivery rate increases, which means more duplications are generated to achieve the assurable delivery rate. In the mean time, the control message overhead is near constant, which means the delivery ratio can be increased without increasing the control message overhead.

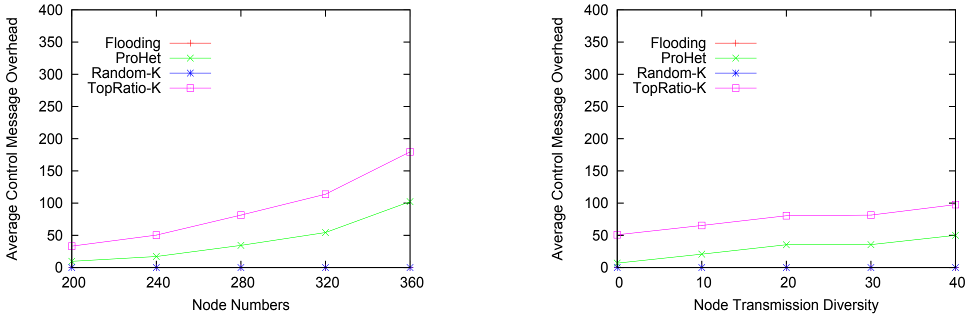
Comparisons of the ProHet protocol with the other three strategies are shown in figures 7, 8, 9, and 10. Figure 7 shows the delivery ratios of these strategies. We can see that Flooding has the highest delivery ratio. Both TopRatio-K and ProHet have lower delivery ratios than Flooding, but their performance is close to each other. Random-K has the lowest delivery ratio among all strategies. Figure 8 reports that Flooding, TopRatio-K, and ProHet have similar average hops to deliver packets, but Random-K uses more hops. Figures 9 and 10 show the average packet replication overhead and average control message overhead of these strategies. We can see that ProHet has the lowest packet replication overhead in both settings, which is about 30% to 50% lower than that in the other three strategies.

In summary, the ProHet protocol guarantees a high delivery ratio with low communication overhead.



(a) Average packet replication overhead in setting 1 (b) Average packet replication overhead in setting 2

Fig. 9. Results of average packet replication overhead



(a) Average control message overhead in setting 1 (b) Average control message overhead in setting 2

Fig. 10. Results of average control message overhead

## VI. Conclusion

In this paper, we proposed ProHet: a probabilistic routing protocol for heterogeneous sensor networks. The ProHet protocol is designed to deal with two challenging issues in heterogeneous sensor networks: asymmetric communication links and reliability. It addresses the asymmetric links by finding reverse routing paths and improves the reliability by choosing forwarding nodes based on historical statistics. We showed that ProHet can achieve assurable delivery rate by theoretical analysis. And its efficiency was verified by our extensive simulations.

In our future work, on one hand, we will address the issue of efficient routing discovery in wireless sensor networks with asymmetric links, aiming at a further reduction in communication cost. On the other hand, we will compare our protocol with other heterogeneous protocols using theoretical analysis and simulation experiments.

## Acknowledgment

This work was partially supported by NSF grant CNS 0835834 and REU supplement grant. National Natural Science Foundation of China under Grant No. 90718031, 60721002, 60803111; the National Basic Research Program of China (973) under Grant No.2009CB320705; The JiangSu Natural Science Foundation (No. BK2009100).

## References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayicrci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–116, 2002.

[2] A. M. Mainwaring, D. E. Culler, J. Polastre, R. Szewczyk, and J. Anderson, “Wireless sensor networks for habitat monitoring,” in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications(WSNA)*, Atlanta, Georgia, USA.

[3] A. Sixsmith and N. Johnson, “A smart sensor to detect the falls of the elderly,” *IEEE Pervasive Computing*, vol. 3, pp. 42–47, 2004.

[4] I. A. Essa, “Ubiquitous sensing for smart and aware environments,” *IEEE Personal Communications*, vol. 7, pp. 47–49, 2000.

[5] D. A. Patterson, “Rescuing our families, our neighbors, and ourselves,” *Commun. ACM*, vol. 48, no. 11, pp. 29–31, 2005.

[6] 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 Computer*, vol. 33, no. 7, pp. 42–48, 2000.

[7] M. D. Yarvis, N. Kushalnagar, H. Singh, A. Rangarajan, Y. Liu, and S. Singh, “Exploiting heterogeneity in sensor networks,” in *24th Annual Joint Conference of the IEEE Computer and Communications Societies(INFOCOM)*, Miami, Florida, USA, Mar. 2005, pp. 878–890.

[8] L. Lazos, R. Poovendran, and J. A. Ritcey, “Probabilistic detection of mobile targets in heterogeneous sensor networks,” in *Proceedings of the 6th International Conference on Information Processing in Sensor Networks(IPSN)*, Cambridge, Massachusetts, USA, April 2007, pp. 519–528.

[9] C. Intanagonwiwat, R. Govindan, and D. Estrin, “Directed diffusion: a scalable and robust communication paradigm for sensor networks,” in *Proceedings of the 6th annual international conference on Mobile computing and networking(MobiCom)*, Aug., Boston, Massachusetts, USA 2000, pp. 56–67.

[10] 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 *Proceedings of the 8th Annual International Conference on Mobile*

*Computing and Networking(MOBICOM)*, Atlanta, Georgia, USA, Sept. 2002, pp. 148–159.

[11] A. Manjeshwar and D. P. Agrawal, "Teen: A routing protocol for enhanced efficiency in wireless sensor networks," in *Parallel and Distributed Processing Symposium, International*, Los Alamitos, CA, USA, April 2001, pp. 2009–2015.

[12] V. Ramasubramanian and D. Mossé, "Bra: a bidirectional routing abstraction for asymmetric mobile ad hoc networks," *IEEE/ACM Transactions on Networking(TON)*, vol. 16, no. 1, pp. 116–129, 2008.

[13] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol forwireless microsensor networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences(HICSS)*, Jan. 2008.

[14] D. Tian and N. D. Georganas, "Energy efficient routing with guaranteed delivery in wireless sensor networks," in *IEEE Wireless Communications and Networking Conference(WCNC)*, New Orleans, Louisiana, USA, Mar. 2003, pp. 1923–1929.

[15] B. Karp and H. T. Kung, "Gpsr: greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th annual international conference on Mobile computing and networking(MobiCom)*, Aug., Boston, Massachusetts, USA 2000, pp. 243–254.

[16] A. Rao, C. H. Papadimitriou, S. Shenker, and I. Stoica, "Geographic routing without location information," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking(MobiCom)*, San Diego, CA, USA.

[17] J. Newsome and D. X. Song, "Gem: graph embedding for routing and data-centric storage in sensor networks without geographic information," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems(SenSys)*, Los Angeles, California, USA, Nov. 2003, pp. 76–88.

[18] G. Q. Wang, Y. C. Ji, D. C. Marinescu, and D. Turgut, "A routing protocol for power constrained networks with asymmetric links," in *Proceedings of the 1st ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks(PE-WASUN)*, Venezia, Italy, Oct. 2004, pp. 69–76.

[19] X. Chen, W. Y. Qu, H. L. Ma, and K. Q. Li, "A geography-based heterogeneous hierarchy routing protocol for wireless sensor networks," in *Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications(HPCC)*, Dalian, China, 2008, pp. 767–774.

[20] X. Du and F. Lin, "Designing efficient routing protocol for heterogeneous sensor networks," in *Proceedings of the 24th IEEE International Performance Computing and Communications Conference(IPCCC)*, Phoenix, Arizona, USA, 2005, pp. 51–58.

[21] V. Paruchuri, A. Durresi, and L. Barolli, "Energy aware routing protocol for heterogeneous wireless sensor networks," in *16th International Workshop on Database and Expert Systems Applications(DEXA Workshops)*, Copenhagen, Denmark, Aug. 2005, pp. 133–137.

[22] K. X. Xu and M. Gerla, "A heterogeneous routing protocol based on a new stable clustering scheme," in *IEEE Military Communications Conference(MILCOM)*, Anaheim, California, USA, Oct. 2002, pp. 838–843.

[23] Q. Zhang and W. G. Chang, "A power efficiency routing protocol for heterogeneous sensor networks," in *4th International Conference on Wireless Communications, Networking and Mobile Computing(WiCOM)*, Dalian, China, Oct. 2008, pp. 1–4.

[24] C. L. Barrett, S. Eidenbenz, L. Kroc, M. V. Marathe, and J. P. Smith, "Parametric probabilistic sensor network routing," in *Proceedings of the Second ACM International Conference on Wireless Sensor Networks and Applications(WSNA)*, San Diego, California, USA, Sept. 2003, pp. 122–131.

[25] N. Wang and C. H. Chang, "Performance evaluation of geographic probabilistic flow-based spreading routing in wireless sensor networks," in *Proceedings of the 4th ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks(PE-WASUN)*, Chania, Crete Island, Greece, Oct. 2007, pp. 32–38.

[26] B. Deb, S. Bhatnagar, and B. Nath, "Information assurance in sensor networks," in *Proceedings of the Second ACM International Conference on Wireless Sensor Networks and Applications(WSNA)*, San Diego, California, USA, Sept. 2003, pp. 160–168.