An Efficient N-to-1 Multipath Routing Protocol in Wireless Sensor Networks

Wenjing Lou
Department of Electrical and Computer Engineering
Worcester Polytechnic Institute, Worcester, MA 01609
Email: wjlou@ece.wpi.edu

Abstract— **A typical task in a wireless sensor network is that every sensor node senses its local environment and, upon request, sends the data of interest back to a base station. Based on this many-to-one communication pattern, we first propose a distributed *N-to-1* multipath discovery protocol which distinguishes from other multipath routing protocols in that it is able to find multiple node-disjoint paths from every sensor node to the base station simultaneously in one route discovery process. Then we propose a hybrid multipath data collection scheme which combines end-to-end multipath traffic dispersion and per-hop alternate path salvaging. Our simulation results show that the proposed *N-to-1* multipath discovery protocol is highly efficient and the hybrid data collection scheme based on it provides a seamlessly more reliable and more secure data collection service in wireless sensor networks.**

I. Introduction

Recent advancement in microprocessor, memory, and wireless networking and communication technologies have made the deployment of wireless sensor networks possible. A wireless sensor network typically is composed of a large number of low-cost sensor nodes which work collectively to carry out some real-time sensing and monitoring tasks within a designated area. This emerging technology has drawn growing attention recently since it provides a promising solution to some challenging tasks, such as the military sensing and tracking in the hostile ground, the remote sensing in nuclear plants, mines, and other hazardous industrial venues, real-time traffic monitoring, realtime weather monitoring, wild animal monitoring and tracking, etc.

Realization of a wireless sensor network faces many challenges. A close relative of the sensor network is the mobile ad hoc network (MANET). Although some of the wireless ad hoc networking techniques are applicable to sensor networks, a sensor network differs from an ad hoc network in many aspects [1]. The number of nodes in a sensor network is usually much larger than that in an ad hoc network. Sensor nodes are more resource constrained in terms of power, computational capabilities, and memory. Sensor nodes are typically randomly and densely deployed (e.g., by aerial scattering) within the target sensing area. The post-deployment topology is not predetermined. Although in many cases nodes are static, the topology might change frequently because wireless links are not stable and sensor nodes are prone to failure. An ad hoc network is typically infrastructureless, end-to-end communications are the common communication pattern. While a sensor network is typically formed around one (or more) *base station* (BS, a.k.a. *sink*). All the sensor nodes are usually designed to sense its local environment and, upon request, send the data of interest back to the base station which is generally several magnitudes more powerful than sensor nodes and serves as a concentration point of the sensor network and at the same time the nexus connecting the sensor network to the rest of the world.

Reliable and secure data collection is an important task in a sensor network. Reliability, defined as the successful end-to-end information delivery ratio, has been an issue in sensor networks since nodes are prone to failure and wireless transmission between nodes are susceptible to all kinds of interferences. Security is another issue since nodes, when deployed in hostile ground, are subject to security compromise. It is generally economically not feasible to make sensor nodes tamper-proof, which means that once a node is compromised, all the secrets stored in that node, including cryptographic keys, may be compromised too, which jeopardizes information relayed by that node. Multipath traffic dispersion has been known as an effective strategy to improve reliability in the face of path failures caused by unreliable links and frequent topological changes [18]. However, improved reliability can be achieved only at the cost of excessive redundancy, that is, sending more data than necessary along multiple paths such that reconstruction of original information can tolerate up to a certain amount of path failure/packet loss. In [13], we proposed a Secure Protocol for REliable dAta Delivery (SPREAD) for end-to-end message delivery in a mobile ad hoc network. In stead of using the single shortest path to route data from one node to the other, SPREAD splits a message into multiple shares using the secret sharing scheme and then delivers the message shares to the destination via multiple independent paths. The SPREAD idea was shown to be effective in improving security in the sense that it is more resistant to collusion attacks of up to a certain number of compromised nodes. However, from the security perspective, little or none redundancy should be added to the information transmitted. The amount of information redundancy added makes security and reliability a seemingly contradicting objectives for schemes based on multipath routing.

In this paper, we propose a distributed *N-to-1* multipath discovery protocol, based on which we propose a hybrid multipath scheme to achieve both more reliable and more secure data collection task in wireless sensor networks. While most of multipath routing protocols are source-initiated and

0-7803-9466-6/05/$20.00 ©2005 IEEE

aim to find multiple disjoint or partially disjoint paths between a single source-destination pair [4], [5], [15], [20], the distinct feature of our *N-to-1* multipath discovery protocol is that it is receiver-initiated (i.e., BS initiated) and at the end of one route discovery process, the protocol finds every sensor node a set of node-disjoint paths to the BS simultaneously. It is highly efficient, with an average overhead of less than one routing message per path. Then we propose a hybrid multipath data collection scheme, which combines end-to-end multipath data dispersion and per-hop alternate path routing to improve both reliability and security. The simulation results show that our hybrid scheme can achieve significantly better reliability and better security seamlessly with little or even none redundancy. The proposed scheme is extremely suitable for wireless sensor networks where the major task is for the base station to collect sensor readings from all the sensor nodes simultaneously.

The rest of the paper is organized as follows. Section II briefly reviews the SPREAD idea. Section III describes the distributed *N-to-1* multipath discovery protocol and its evaluation. In section IV, the hybrid multipath data collection scheme is proposed and evaluated. Finally, related work is reviewed in section V and conclusion is drawn in section VI.

II. A Brief Review of SPREAD

In [13], we proposed the SPREAD scheme as a complementary mechanism to enhance data confidentiality service in a MANET. The basic idea and operation of SPREAD is illustrated in Fig. 1. A secret message m is transformed into multiple shares, $S_1, S_2, \cdots$, by a secret sharing scheme, and then delivered to the destination via multiple independent paths. Due to the salient features of secret sharing and the distributed fashion of the multipath delivery, the SPREAD has been shown to be more resilient to a collusive attack by up to a certain number of compromised nodes, namely, even if a small number of paths/nodes/shares are compromised, the message as a whole is not compromised.

A number of coding schemes can be used to split the message for multipath routing in order to enhance reliability. Examples include well-known Reed-Solomon codes, diversity coding, multiple description coding, etc. In the SPREAD scheme [13], we used the threshold secret sharing scheme to split the information. A (T, N) threshold secret sharing scheme could transform a secret into N pieces, called *shares* or *shadows*. The nice property of the N shares is that form any less than T shares one cannot learn anything about the secret, while with an effective algorithm, one can reconstruct the secret from any T out of N shares. The generation of the shares is very simple - by evaluating a polynomial of degree $(T-1)$

$$f(x) = (a_0 + a_1 x + \cdots + a_{T-1} x^{T-1}) \bmod p$$

at point $x = i$ to obtain the $i-$th share:

$$S_i = f(i)$$

where coefficients $a_0, a_1, a_2, \ldots, a_{T-1}$ are secret bits while p is a large prime number greater than any of the coefficients and can be made public. Note when all the coefficients are used to carry secret bits, the fraction of redundant information is $\frac{N-T}{N}$. For $T = N$, there is no redundant information resulting from secret sharing[1].

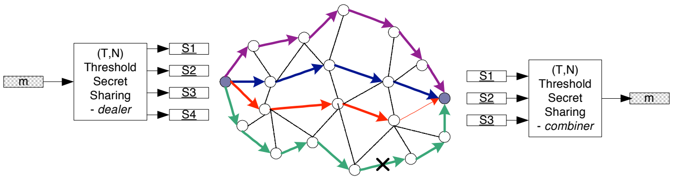


Fig. 1. Basic idea of SPREAD

According to the fundamental theorem of algebra, T values of a polynomial of degree $(T-1)$ can completely determine the polynomial (i.e., all its coefficients), while any fewer values cannot determine the polynomial (at least computationally difficult). Thus, any T shares can reconstruct the original secret bits, but any fewer shares cannot. Efficient $(O(T\log^2 T))$ algorithms have been developed for polynomial evaluation and interpolation [3]. In addition, the reconstruction is done in the base station, which is not computationally constrained very much. Therefore, in our new scheme for data collection in sensor network, we still choose secret sharing as the coding scheme.

III. The *N-to-1* Multipath Discovery Protocol

A challenging job in any multipath routing based schemes is the development of efficient and effective multipath routing protocols. In [13], we discussed multipath finding techniques between a single source-destination pair. In fact, most of current multipath routing protocols fall into this category. In response to the communication pattern in a sensor network, in this paper we propose a novel *N-to-1* multipath discovery protocol. Instead of finding multiple paths between a specific source and a specific destination, our proposed protocol takes advantage of flooding in a typical route discovery process and is able to find multiple node-disjoint paths from every sensor node to the common destination (i.e., the sink node) simultaneously. We present the distributed protocol and evaluate its path finding capability in this section.

A. Motivation and Overview

A typical task of a sensor network is data collection where the base station broadcasts the request for data of interest and every sensor node (or nodes that have the data of interest) sends its readings back to the base station. For this purpose, Berkeley's TinyOS sensor platform utilizes a flooding-based beaconing protocol. The base station periodically broadcasts a route update. Each sensor node when receiving the update for the first time rebroadcasts the update and marks the node from which it receives the update as its parent. The algorithm continues recursively till every node in the network has rebroadcasted the update once and finds its parent. What follows is that every node forwards the packets it received or generated to its parent until the packets reach the base station [11]. As illustrated in Fig. 2, the beaconing protocol

[1] The length of the coefficients needs to be one bit shorter than that of p.

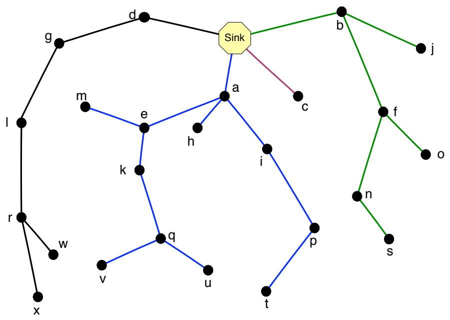


Fig. 2. Spanning tree created by flooding

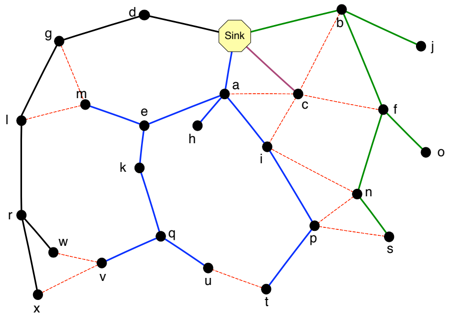


Fig. 3. A simple multipath extension of flooding

essentially constructs a breadth first spanning tree rooted at a base station. It finds every sensor node a single path back to the base station efficiently. However, both reliability and security suffer from the single path routing. The failure of a single node or link will disrupt the data flow from the node itself and all its children. Similarly, the compromise of a single node will cause the information leakage from the node and all its children.

The proposed *N-to-1* multipath discovery protocol is based on the simple flooding initiated at the BS. Then, by carefully incorporating two other mechanisms into the protocol design, it is able to find every sensor node multiple node-disjoint paths back to the BS at the end of a distributed path discovery process. To facilitate the understanding, we present the multipath discovery procedure in two phases, with each phase implementing one of the two mechanisms. In fact, the second phase can be started at each individual node in a distributed fashion without considering the completion of phase one at other nodes. The mechanism used in phase one, *branch aware flooding*, takes advantage of the simple flooding technique. Without introducing additional routing messages, the mechanism is able to find a certain number of node-disjoint paths, depending on the density of the network topology. The mechanism used in phase two, *multipath extension of flooding*, helps to exchange the node-disjoint paths found in phase one among nodes on different branches. At the cost of some more message exchanges, it is able to increase the number of paths found at each sensor node.

We present the two phases of the proposed multipath routing protocol in the following two subsections.

B. Phase One: Branch Aware Flooding

The general form of the routing messages in both phases is $\{mtype, mid, nid, bid, cst, path\}$, where *mtype* indicates the type of message. We define *mtype*="RPRI" for phase one, which refers to "primary" because paths found by this type of messages are primary paths (on the shortest path tree); *mid* is the sequence number of the current routing update; *nid* is the identifier of the node sending out the message; *bid* is the identifier of the branch defined as *nid* of the node closest to the BS in the branch; *path* contains a sequence of nodes which the message has travelled; and *cst* is the cost of the *path*.

The propagation of the RPRI message follows exactly the same way as the TinyOS beaconing protocol. The BS initializes the routing update periodically (or on demand) by broadcasting message $\{RPRI, mid, Sink, \emptyset, 0, (Sink)\}$. Every node, say z, when hearing a message $\{RPRI, mid, nid, bid, cst, path\}$ for the first time, marks node nid as its parent, and it also learns the primary path back to the BS by following the reverse order of $p = path + (z)$. It then forms a new routing message $\{RPRI, mid, z, (bid == \emptyset)?z : bid, cst + cost(z, parent(z)), path + (z)\}$ according to the following rules: replacing nid field with its own ID; if bid field is Ø, replacing bid field with its own ID, otherwise keeping the original bid intact; updating cst field by adding the cost from z to the node from which this message is received; and updating $path$ field by appending its own ID at the end of the old path. Node z then rebroadcasts the new message in the neighborhood.

In the simple flooding protocol (such as the beaconing protocol), a node simply ignores the duplicate route update messages from other nodes. However, in our branch aware flooding, when a node z hears the same message (i.e., identified by the same mid) from a neighbor, it will check the content of the message and mark the neighbor accordingly. If the message has the same *bid* as node z itself, z will mark that neighbor as a *child* or *sibling*, according to the $path$ contained in the message; if the message has a different *bid*, which means the message is from another branch, z will mark that neighbor as a *cousin*. Node z maintains an alternate path set $\mathcal{Q}_z$. Once receiving a message from a cousin node, z will further examine the path contained in the message. If the new path $q = path + (z)$ is disjoint from the primary path p and any other alternate path with lower cost in $\mathcal{Q}_z$, the new path q will be included into the $\mathcal{Q}_z$, while at the same time, paths with higher cost than q that share common nodes with q will be removed from $\mathcal{Q}_z$. Same as the beaconing protocol, the propagation of the RPRI messages is terminated at the leaf nodes when each node has rebroadcasted the message once and only once.

The branch aware routing technique is actually based on the following observation. As show in Fig. 3, the number of branches a tree has depends on the number of immediate neighbors the base station has (e.g., 4 branches in the example where different branches are distinguished by different colors). The maximum number of node-disjoint paths from any node to the base station is thus bounded by the number of branches. We notice that while each node has a *primary* (in most cases also the shortest) path to the base station by following its tree links up, a link between two nodes that belong to two different branches will provide each node an alternate disjoint path to the base station through the other. For example, as shown in Fig. 3, while node w has the primary path $(w - r - l - g - d - Sink)$ back to the base station, it learns another alternate path $(w - v - q - k - e - a - Sink)$ from node v which is not in the same branch as w when overhearing v's broadcast. The branch aware flooding is therefore designed to allow nodes to go across a cousin link thereby finding disjoint paths in other branches. This mechanism takes advantage of the broadcast nature of the wireless communication. Without introducing extra routing messages, nodes that have cousin neighbors are able to find a few disjoint paths.

C. Phase Two: Multipath Extension of Flooding

The number of paths that could be found by the branch aware flooding is limited to the nodes that have cousin neighbors. In what follows, we present the second mechanism/phase proposed for our *N-to-1* multipath discovery protocol, a multipath extension to the flooding technique, which can find more node-disjoint paths at each sensor node at the cost of some more message exchanges.

Phase two message exchange uses the same message format but with $mtype$ field set to "RALT", which refers to "alternate" because paths found by this type of messages form the alternate paths in the path set. The RALT messages are used to further propagate the alternate paths found at one node to its parent and sibling/cousin neighbors[2]. The propagation of the RALT messages is initiated distributively and independently at each node where an alternate disjoint path(s) is found during the branch aware flooding. For each alternate path q, node z forms a RALT message $\{RALT, mid, z, q.bid, q.cst, q\}$ and broadcasts it in its neighborhood[3].

Upon receiving a RALT message {*RALT, mid, nid, bid, cst, path*}, node z will ignore it if it is from its parent. Otherwise, it will check and see if itself is already in the path contained in the message. If not, node z learns about a new path $q = path + (z)$. Again, node z includes the new path q into its alternate path set $\mathcal{Q}_z$ if q is disjoint from any other paths in $\mathcal{Q}_z$ of lower cost. If q is included, node z excludes from the path set $\mathcal{Q}_z$ paths of higher cost and intersecting with q. Whenever a new path q is added to $\mathcal{Q}_z$, node z forms a new RALT message $\{RALT, mid, z, q.bid, q.cst, q\}$ and broadcasts it in the neighborhood.

The propagation of RALT messages terminates when no new disjoint path is added to any path set. At this time, each node has found a set of disjoint paths to the BS.

TABLE I

PARAMETERS OF NETWORKS SIMULATED

Transmission Range (TR)	15 m	20 m	25 m	30 m
Average node degree (*d*)	6.05	10.19	15.29	21
Average network diameter (*D*)	10.56	6.09	4.72	3.85

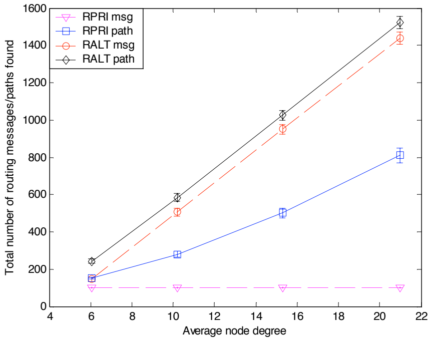


Fig. 4. Path finding capability

The rational behind the design of the phase two mechanism is to maximize the number of disjoint paths at each node by further propagating alternate paths found at phase one across multiple branches. Using the same example as shown in Fig. 3, notice that if w further propagates the disjoint paths it learned to its neighbors, its parent or siblings/cousins might learn a new disjoint path as well. For example, node r has the primary path $(r - l - g - d - Sink)$. When it hears a disjoint path $(w - v - q - k - e - a - Sink)$ from w and it does not yet know a path through branch a, it learns a new disjoint path $(r - w - v - q - k - e - a - Sink)$. The tradeoff of the second phase is that it finds more disjoint paths with additional routing messages.

D. Performance Evaluation

We use simulations to evaluate the performance of the proposed *N-to-1* multipath routing protocol. We simulate a sensor network consisting of 100 nodes randomly deployed in a field of 100m × 100m square area. The base station is located in the middle of one edge. Nodes have same transmission range in one experiment. In order to evaluate the impact of the edge density on the performance, we vary the transmission range in different experiments to adjust the edge density in the network. We tried four different transmission ranges, 15, 20, 25, and 30 meters[4]. Table I summarizes some topological parameters of the networks simulated when using different transmission ranges, including average node degree d (i.e., average number of neighbors a node has) and average network diameter D (i.e., maximum hop count from any sensor node to the BS based on shortest path routing). The simulation results are averaged over 60 random network deployments. The 95% confidence intervals are shown in the figure.

Fig. 4 shows the total number of routing messages and the total number of disjoint paths found in the simulated networks. We observe that the branch-aware flooding mechanism could find disjoint paths without incurring any extra message

[2] Not intended for the children because the parent node must be in the primary path of the children node.

[3] A short delay might be introduced here to allow multiple paths to be broadcasted in one message.

[4] Randomly generated networks sometimes are not connected if the edge density is not high. In our simulation, 70% of networks are not connected when average node degree is 7 (e.g., TR=15m) and 5% of networks are not connected when average node degree is 11 (e.g., TR=20m). Only the results from connected networks are considered here.

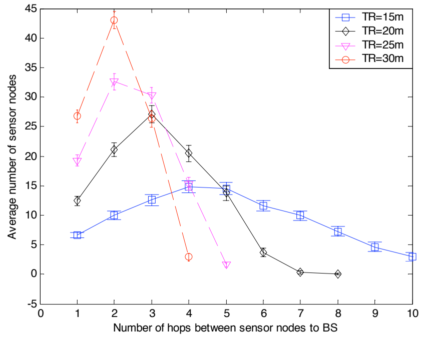


Fig. 5. The distribution of nodes in terms of distance

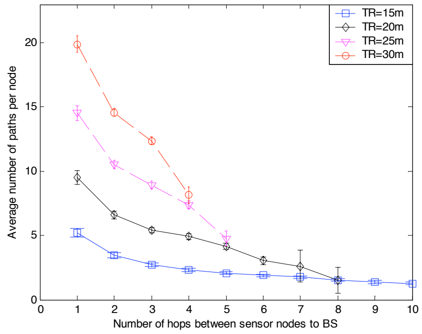


Fig. 6. The distribution of paths in terms of distance

exchanges. When the edge density is high, say when the average node degree is 22, this simple modification could find an average of 8 node-disjoint paths per node. Our multipath extension of flooding mechanism, although requiring more message exchange, is able to find more paths. The results show that, in general, the routing protocol is highly efficient in terms of path finding - the per path cost is less than one message.

The next three figures reveal some more characteristics of the paths found. Fig. 5 shows the distribution of nodes in terms of their distances from the BS. Correspondingly, Fig. 6 depicts the average number of node-disjoint paths found per node as a function of distance between sensor nodes and the BS. It is observed that the closer the node to the BS, the more paths from that node to the BS. This is reasonable because it is harder to find node-disjoint paths when nodes are far away from the BS as each alternate path has to find more unused nodes to reach the destination. This property is actually desirable for the tasks we are considering. Since the typical task of the network is to collect data from all sensor nodes, packets are travelling from everywhere toward the BS. Nodes that are closer to the BS would be used more for forwarding traffic thus it is more desirable for those nodes to be more reliable. More available alternate paths gives a node more choices in the face of node or link failures thus inferring better reliability. Fig. 7 shows the average hop count per path correspondingly. The dotted line is plotted as the reference, indicating the shortest distance to the BS. It is observed that the average path length is typically 1 or 2 hops longer than the shortest path, regardless the length of the shortest path.

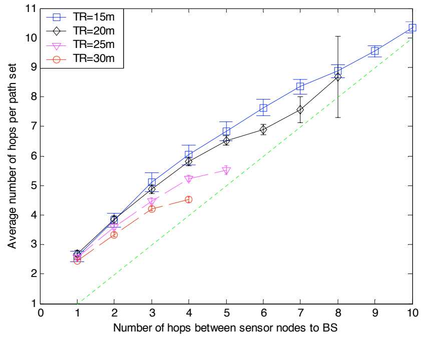


Fig. 7. The quality of the alternate paths

The typical data collection in a sensor network involves the following communication patterns: (a) broadcast from the base station to sensor nodes (e.g., requests of data of interest); (b) from sensor nodes to the base station (e.g., sending back the sensor readings); and (c) node to node communication (e.g., if aggregation of sensor readings are applied). In this section, we described our multipath discovery protocol which, similar to any on-demand routing protocol, starts with a route update initialized at the base station. This route update is a network wide broadcast thus can be used to fulfill the above mentioned type-(a) communication. Then at the end of the discovery, each node will be able to find a set of multiple node disjoint paths to the base station with which our hybrid data collection scheme can be implemented for the type-(b) communication. We did not consider data aggregation explicitly. However, if data aggregation center is applied, a hierarchical routing structure can be constructed: from each sensor node to the aggregation center forms the lower layer and from each aggregation center to the base station forms the higher layer. Our proposed algorithm could be made applicable to each layer acoordingly.

IV. The Hybrid Multipath Data Collection Scheme

A. *Per-hop Alternative Path Packet Salvaging*

In SPREAD [13], the communication we considered is end-to-end and we actually implemented a concurrent multipath data dispersion scheme, meaning, information is split at source and segments are spread onto multiple paths between the source and destination pair. However, single path routing is assumed for the delivery of each packet/segment. This is true for most of the multipath routing approaches proposed in the literature [15], [18], [20] where the proposed multipath routing protocols aim to find multiple disjoint paths between a single source and destination pair on-demand. As shown in the paper [13], the SPREAD scheme is effective in improving the data security as it is harder for the adversaries to compromise multiple nodes/paths.

In fact, with respect to reliability, multipath routing is effective for persistent errors, such as node failures or persistent link errors (i.e., pure erasure channel model), while it does not help in the case that the lost of packets is due to intermittent link

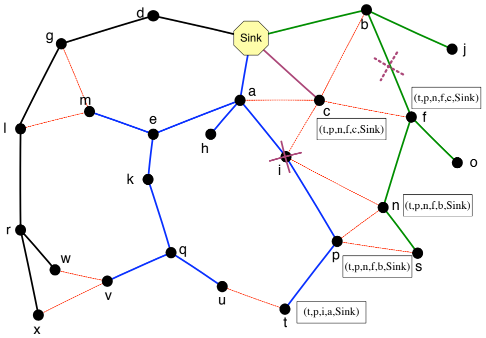


Fig. 8. Alternate path packet salvaging

failures. In addition, the end-to-end multipath routing approach essentially relies on redundant paths to improve reliability, while the unreliability of each path remains unimproved. Due to the use of alternate paths, which are likely less reliable than the primary path, reliability of end-to-end data delivery becomes even worse with multipath routing, given the same amount of information redundancy. The results therefore make security and reliability two seemingly contradicting design objectives - reliability requires more redundancy while security demands less or no redundancy.

A distinct feature of our multipath discovery protocol is that it finds multiple node-disjoint paths at each sensor node. If the sensor network uses a reliable MAC protocol, such as IEEE 802.11 which acknowledges the successful transmission of each frame, each node knows whether the transmission is successful or not before it removes the frame from its transmission buffer. Therefore, taking advantage of the multiple paths available at each hop, we adopt an active per-hop packet salvaging strategy so that reliability of each packet delivery (or each path) can be greatly improved. It works as follows. Each packet carries with it the source routing option. At an intermediate node z, if the transmission to the next hop is not successful, z actively salvages the packet by sending it to another randomly selected route to the destination rather than dropping the packet. Only when all the next hops from node z to the BS fail should the packet be dropped. Fig. 8 shows an example that a packet originated at node *t* is salvaged twice at nodes *p* and *f* respectively and finally reaches the destination. One potential problem here is when a node salvages a packet with a new path but that new path consists of a node that the packet has already travelled. In this case, a routing loop would result. This problem can be easily solved by the source routing option. Notice that our multipath discovery protocol guarantees the loop freedom for all the paths selected. Each packet carries the source routing option when it is sent out. At an intermediate node, when salvaging needs to be done, the node makes sure no loop would form by comparing the partial route the packet already travelled and the candidate path it would use to salvage the packet. Only when there is no common node would the candidate path be selected. Then the intermediate node modifies the source routing option carried in the packet by replacing the rest of the source route by the newly selected salvaging path. When a node reaches the BS, what it carries is the actual path it travelled through.

The per-hop alternate path packet salvaging is an effective and efficient way to improve reliability. Particularly, it improves the reliability on a per packet/path basis without imposing redundant information. Assume a packet is at node z and to be routed to its next hop, in the single path routing case, there is only one known next hop, therefore the probability that the packet might be dropped at this node is $1 - P_n P_l$, where P_n is a probability that a node can reliably relay a packet and P_l is the probability that a link can reliably deliver a packet. While if node z knows m possible next hops and it applies active alternate path packet salvaging, the probability that the packet might be dropped at the node can be reduced to $(1 - P_n P_l)^m$. Further, if the path consists of h hops, the packet delivery ratio of that path would be $(P_n P_l)^h$ without salvaging, while with the active alternate path packet salvaging, the packet delivery ratio becomes $(1 - (1 - P_n P_l)^m)^h$. This calculation is not accurate since salvaging does not guarantee independence of paths used. However, it does provide some intuitive understanding and show potential improvement that alternate path packet salvaging could achieve. We will use simulations to evaluate security and reliability with or without salvaging in the next section.

B. Simulation Studies

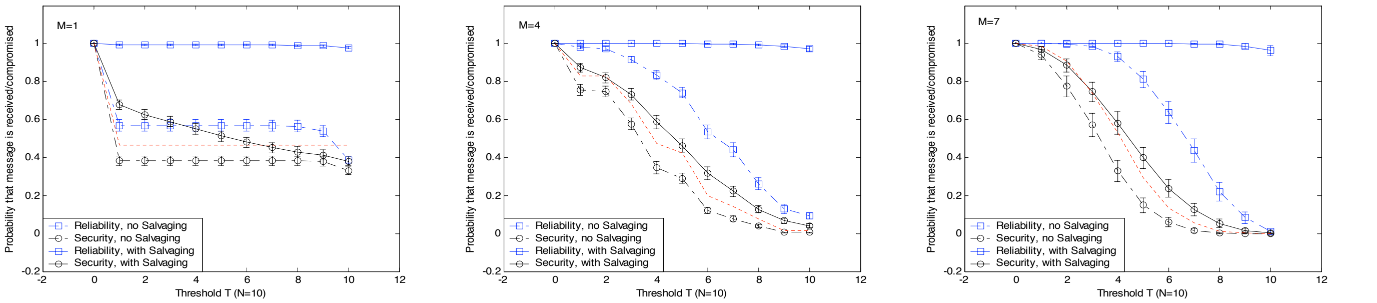
In this section, we evaluate the overall security and reliability performance of the proposed hybrid multipath data collection scheme, namely, the combination of the concurrent multipath routing on the end-to-end data collection task and the alternate path routing on each packet delivery along the designated path.

We run the *N-to-1* multipath discovery protocol we proposed in section III. Then we consider the impact of node failure, link failure, as well as compromised node problem. We assume that node failure is persistent. Once a node fails, it cannot be used to forward packets. Link failure is intermittent and is independent of each packet transmission. When a link error occurs to a packet, no retransmission is performed for the same packet. Node compromise is persistent too. If a node is compromised, all the shares/packets relayed by that node are considered compromised.

Due to space limitation, we only report results in networks where the transmission range is 20m (refer to section III-D for network parameters). Each node which is at least 2 hops away from the sink node initiates 100 messages. Each message is divided into $N = 10$ shares and spread onto M paths ($M = 1, \cdots, 7$). For $M = 1$, the shall allocation vector is $\underline{n} = [10]$, namely, all the 10 shares go through the primary path. For $M = 2$, $\underline{n} = [5\ 5]$, namely, 5 shares take the 1st path and 5 shares take the 2nd path. Similarly, for $M = 3$, $\underline{n} = [4\ 3\ 3]$; $M = 4$, $\underline{n} = [3\ 3\ 2\ 2]$; $M = 5$, $\underline{n} = [2\ 2\ 2\ 2\ 2]$; $M = 6$, $\underline{n} = [2\ 2\ 2\ 2\ 1\ 1]$; $M = 7$, $\underline{n} = [2\ 2\ 2\ 1\ 1\ 1\ 1]$.

The simulation results are averaged over 300 randomly generated networks. The 95% confidence intervals are shown in the figures.

Fig. 9 shows significant improvement in reliability when salvaging is used. Due to space limitation, we only present



(a) M=1

(b) M=4

(c) M=7

Fig. 9. Security and reliability performance with or without packet salvaging (10% faulty nodes, 10% compromised nodes, link failure probability 1%)

figures for $M = 1, 4, 7$, figures for other M values show the same trend. The X axis is the threshold T (with N set to 10 in all simulations) which can be interpreted as level of redundancy. Here reliability is represented by the probability that a messages is successfully delivered, which is calculated as the total number of messages received at the Sink node over the total number of messages initiated from all the sensor nodes. A message is received when at least T shares of the message reach the Sink. Similarly, security is represented by the probability that a message is compromised, which is calculated as the total number of messages compromised over the total number of messages initiated by all the sensor nodes. A message is compromised when at least T shares are compromised by the compromised nodes collectively. Therefore, $T = 10$ means no redundancy and either the BS or adversaries must receive/intercept all the 10 shares to recover a message. It is observed that without salvaging, the packet loss ratio is sensitive to redundancy level and is unacceptably high even with excessive redundancy (small T values). However, our alternate path packet salvaging effectively maintains a very high (close to 100%) delivery ratio at all redundancy levels, even with zero redundancy. On the other hand, we observe that security is very sensitive to the redundancy - the less redundancy, the more secure the scheme is. The dotted line is drawn as the reference. It indicates the security achieved when all the nodes and links are reliable therefore no salvaging is performed[5]. As expected, salvaging weakens security a little bit because of possible overlapping of the paths. However, the impact is not significant compared with significantly improved reliability. This is the most desirable property that enables our proposed scheme to improve both security and reliability at the same time.

Fig. 10 and Fig. 11 plot security and reliability as a function of the number of paths used respectively. Due to space limitation, only the result for a very challenging situation is shown where 20% of nodes are compromised/faulty. It is clear that our scheme is effective in reducing the probability that a message might be compromised. We observe that although the active packet salvaging breaks the independence of paths, the probability that a message might be compromised decreases with the increase of the number of paths used to spread the information. In fact, in less challenging situations (i.e., less number of compromised nodes), the improvement is more significant (i.e., curves dropping more steeply). The results confirms the effectiveness of the proposed scheme - it is more resistant to the collusive attacks of compromised nodes. Correspondingly, the reliability performance shows that the proposed scheme is able to maintain pretty good message delivery ratio in the face of both link and node failures. Therefore, we conclude that the proposed scheme with active alternate path salvaging is more robust to node/link failure problem too.

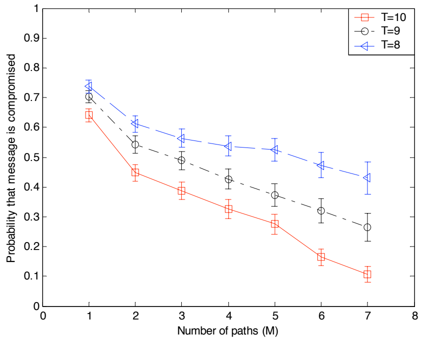


Fig. 10. Security performance (20 compromised nodes)

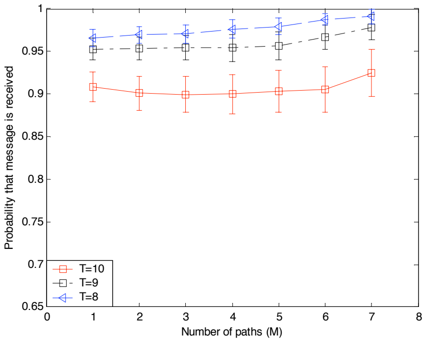


Fig. 11. Reliability performance (20 faulty nodes, 2% link failure rate)

[5] The lower compromise probability for the no salvaging case does not indicate more security. It is lower than the reference because of loss of packet.

V. Related Work

Efficient data delivery in sensor networks is a challenging task. Direct diffusion [6], [9] and SPIN [10] are two exemplary data dissemination paradigms. As a data-centric approach, direct diffusion employs low rate flooding to establish gradients and uses gradual reinforcement of better paths to accommodate certain levels of network and sink dynamics. SPIN adopts meta-data negotiation to eliminate the redundant data transmission and is suitable for scenarios where an individual sensor disseminates its observations to all sensors in a network. Some other approaches for data dissemination in sensor networks include flooding based Gossiping [8], probabilistic-based flooding [2], [16], geometry-based flooding [14], cluster-based LEACH [7], hierarchical-based TTDD [19], etc.

Multipath routing has been a promising technique in mobile ad hoc networks in order to aggregate limited bandwidth, to smooth traffic burstiness, to alleviate network congestion, and to improve fault tolerance, and most importantly, to improve reliability [4], [18]. Several multipath routing protocols have been proposed to find multiple disjoint or partially disjoint paths between a single source and destination pair [5], [6], [15], [20]. These multiple paths can be used in different ways. One way is to use them alternatively, namely, use the primary path first, when the primary one fails, switch to the secondary one, and so on. The other type of usage is to use the multiple paths simultaneously.

Our approach distinguishes from previous work in that (a) Our *N-to-1* multipath discovery protocol is receiver-initiated (in contrast to the common source-initiated route discovery) and the protocol is efficient in that it finds multipath from every sensor node to the base station, which fits the special communication pattern (i.e., multiple senders to a single receiver) in the sensor network very well; (b) We adopt a hybrid multipath approach for data delivery. We use concurrent multipath scheme to spread traffic onto multiple disjoint paths for end-to-end data delivery. Meanwhile, taking advantage of the multiple paths available at each node, the per-hop alternate path packet salvaging uses the multiple paths alternately and helps to improve the reliability of each packet delivery/path significantly; and (c) The overall scheme improves both security and reliability.

VI. Conclusion

Data collection is an important task in a wireless sensor network. Reliable and secure techniques are desired to perform the task efficiently. In this paper, we consider a wireless sensor network where the typical task is to disseminate data requests from a base station to all sensor nodes and to collect sensor readings from every sensor node back to the base station. We first propose an efficient *N-to-1* mulitpath discovery protocol which initiates a route update periodically or on demand at the base station and at the end of each discovery process, finds every sensor node a set of node-disjoint paths back to the base station. Then based on the availability of multiple paths at each node, we propose a hybrid multipath scheme for secure and reliable data collection task. The proposed scheme applies the secret sharing scheme at the source to split information into multiple shares and then spread the shares onto multiple paths available at the source for concurrent delivery. For each message share, while travelling along one path, an alternate path packet salvaging strategy is further adopted so that reliability of each packet delivery is improved significantly. The simulation results show that the proposed multipath discovery protocol is highly efficient, with less than one message per path found. The proposed hybrid multipath data collection scheme is more resilient to node/link failures and collusive attacks of compromised nodes. It is effective in improving both reliability and security at the same time.

References

[1] A. F. Akyildiz, W. Su, Y. Sankarasubramainiam, E. Cayirci, “A survey on sensor networks”, *IEEE Communications Magazine*, August 2002

[2] C. Barrett, S. Eidenbenz, L. Kroc, “Parametric probabilistic sensor network routing”, *Proc. 2nd International Workshop on Wireless Sensor Networks and Applications (WSNA 2003)*, San Diego, Sept. 2003

[3] T. Cormen, C. Leiserson, R. Rivest, *Introduction to algorithms*, MIT Press, 1990

[4] S.K. Das, A. Mukherjee, et al, “An adaptive framework for QoS routing through multiple paths in ad hoc wireless newtorks”, *J. Parallel Distributed Computing*, 63(2003)141-153

[5] S. De, C. Qiao, H. Wu, “Meshed multipath routing: an efficient strategy in sensor networks”, *IEEE Wireless Communications and Networking Conference (WCNC’03)*, New Orleans, LA, Mar 2003

[6] D. Ganesan, R. Govindan, S. Shenker, D. Estrin, “Highly-resilient, energy-efficient multipath routing in wireless sensor networks”, *Mobile Computing and Communication Review*, 5(4):10-24, Oct 2001

[7] W. Heinzelman, A. Chandrakasan, and H. balakrishnan. “Energy-effcient rotuing protocols for wireless microsensor networks”. *Proc. 33rd hawaii Intenational Conference on System Sciences (HICSS’00)*, Hawaii, Jan. 2000

[8] Z. Hass, J. Halpern, and L. Li. “Gossip-based ad hoc routing”. *proc. of IEEE INFOCOM02*, New York, June 2002

[9] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, “Directed diffusion for wireless sensor networks”, *IEEE/ACM Transactions on Networking*, 11(1):2-16, Feb 2003

[10] J. Kulik, W. Heinzelman, H. Balakrishnan, “ Negotiation-based protocols for disseminating information in wireless sensor networks”, *Wireless Networks*, 8(2-3):169-185, March/May 2002

[11] C. Karlof, D. Wagner, “Secure routing in wireless sensor networks: Attacks and countermeasures”, *Elsevier’s AdHoc Networks Journal, Special Issue on Sensor Network Applications and Protocols*, 1(2-3):293-315, September 2003

[12] W. Lou, Y. Fang, “A multipath routing approach for secure data delivery”, *IEEE MILCOM*, McLean, VA, Oct 2001

[13] W. Lou, W. Liu, Y. Fang, “SPREAD: Enhancing data confidentiality in mobile ad hoc networks, *IEEE INFOCOM 2004*, HongKong, China, March 2004

[14] W. Liu, Y. Zhang, W. Lou, Y. Fang, “Scalable and Robust Data Dissemination in Wireless Sensor Networks, *IEEE GLOBECOM 2004*, Dallas, TX, Dec 2004

[15] M. K. Marina, S. R. Das, “On-demand multipath distance vector routing in ad hoc networks”, *9th International Conference on Network Protocols*, Riverside, CA, November, 2001

[16] Y. Sasson, D. Cavin, and A. Schiper. “Probabilistic broadcast for flooding in wireless mobile ad hoc networks”. *Proc. of IEEE Wireless Communications and Networking Conference (WCNC’03)*, New Orleans, Louisiana, Mar. 2003

[17] D. Sidhu, R. Nair, S. Abdallah, “Finding disjoint paths in networks”, *Proc. of ACM SIGCOMM*, 1991

[18] A. Tsirigos, Z.J. Haas, “Multipath routing in the presence of frequent topological changes”, *IEEE Communication Magazine*, Nov 2001

[19] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang. “A Two-tier Data Dissemination Model for Large-scale Wireless Sensor Networks”. *Proc. of the Ninth Anuual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom02)*, Atlanta, Georgia, Sept. 2002

[20] Z. Ye, S. V. Krishnamurthy, S. K. Tripathi, “A framework for reliable routing in mobile ad hoc networks”, *IEEE INFOCOM 2003*, Sanfrancisco CA, Mar 2003