SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks 2016; **9**:220–228
Published online 29 March 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/sec.527

RESEARCH ARTICLE

STaR: design and quantitative measurement of source-location privacy for wireless sensor networks

Leron Lightfoot[1], Yun Li[2] and Jian Ren[1]*

[1] Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, U.S.A.
[2] Microsoft, Redmond, WA 98052, U.S.A.

ABSTRACT

Wireless sensor networks (WSNs) can provide the world with a technology for real-time event monitoring. One of the primary concerns that hinder the successful deployment of WSNs is source-location privacy (SLP). The privacy of the source location is vital and highly jeopardized by the usage of wireless communications. Although message content privacy can be ensured through message encryption, it is much more difficult to adequately address the SLP. For WSNs, SLP service is further complex by the fact that sensors consist of low-cost and energy-efficient radio devices. Therefore, using computationally intensive cryptographic algorithms (such as public-key cryptosystems) and large-scale broadcasting-based protocols are not suitable for WSNs. In this paper, we analyze the quantitatively measure source-location information leakage in routing-based SLP protection schemes for WSNs. Through this model, we identify vulnerabilities of some well-known SLP protection schemes. We also propose a routing technique, called the Sink Toroidal Region (STaR), to provide adequate SLP with low energy consumption. With this routing technique, the source node randomly selects an intermediate node within a designed STaR area located around the sink node. Furthermore, this routing protocol ensures that the intermediate node is neither too close, nor too far from the sink node in relations to the entire network. While ensuring SLP, our simulation results show that the proposed scheme is very efficient and can be used for practical applications. Copyright © 2012 John Wiley & Sons, Ltd.

KEYWORDS
source-location privacy; source-location information leakage; quantitative measurement; wireless sensor networks (WSNs)

***Correspondence**
Jian Ren, Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, U.S.A.
E-mail: renjian@egr.msu.edu

1. INTRODUCTION

Wireless sensor networks (WSNs) have been envisioned as a technology that has great potential to be widely used in both military and civilian applications. Privacy has been an extensively studied topic in WSNs. One of the major and unsettled issues in privacy of WSNs is how to provide adequate routing-based source-location privacy (SLP). Sensor networks rely on wireless communications, which is, by nature, a broadcast medium and is more vulnerable to security attacks than its wired counterpart as a result of lack of a physical boundary. When messages are transmitted wirelessly in the open air, any compatible receiver within the transmission range of the sender is able to intercept the traffic. An adversary may be well-equipped with powerful transceivers to analyze the traffic patterns. They may be able to intercept traffic from one or multiple locations within the network environment. Without an adequate protection of the routing paths, an adversary may be able to determine the source location by using radio frequency localization techniques to trace back to the source in a hop-by-hop approach. Therefore, even if a powerful encryption algorithm is used to protect the source identity, the adversary may still be able to determine the location of the source by monitoring the traffic patterns and routing paths.

Privacy in a network consists of not only privacy of the message content but also the privacy of the source and destination locations. The focus of this paper is on SLP. The confidentiality of the message content can be protected by encryption, but the source location can be exposed in routing patterns. To be more concise, there may be different types of information, besides the message content, that are linked with a message transmission.

In providing adequate SLP, the sensor devices present major limitations. Sensors in the network are meant to be low-cost and energy-efficient devices. The sensors are designed to be deployed in environments where they can be damaged or destroyed; thus, the cost of these sensor nodes should be at a minimum. Clients can simply deploy many wireless sensor nodes into an environment and monitor the activities in the environment from one central location. Sensor nodes are also built to be placed in environments where they can be unattended for lengthy periods. These sensors may be deployed in areas where human attending and maintaining the sensors is impractical; thus, changing or recharging batteries in the sensor devices are infeasible. For the purpose of preserving battery life, using intensive cryptographic algorithms, such as public-key cryptosystems, and the usage of powerful transmitters are not suitable for WSNs. Therefore, energy consumption and SLP are two very vital components for the successful deployment of WSNs.

In this paper, we analyze the criteria to quantitatively measure source-location information leakage for routing-based SLP schemes. Through the measurement criteria, we are able to identify security vulnerabilities of some existing SLP schemes. We propose a scheme that can provide both content confidentiality and SLP. In the routing scheme, the message source node randomly selects an intermediate node within a designed Sink Toroidal Region (STaR) area located around the sink node. The STaR area is large enough to make it unpractical for an adversary to monitor the entire region. This routing protocol provides an energy-efficient routing technique that ensures that the intermediate node is neither too close, nor too far from the sink node in relation to the entire network. Through security analysis and performance results, local and global SLP is provided by the proposed scheme. Our results shows that the proposed scheme can adequately provide SLP and can achieve excellent performance.

The rest of this paper is organized as follows: In Section 2, related work is reviewed. The network models are described in Section 3. Section 4 presents the preliminary. The proposed SLP scheme is presented in Section 5. In Section 6, we provide theoretical security analysis of the proposed STaR scheme. Performance analysis and simulation results are provided in Section 7. We conclude in Section 8.

2. RELATED WORK

In the past 2 decades, originated largely from Chaum's mixnet [1], a number of protocols have been proposed to provide SLP [1–3]. The mixnet family protocols use a set of "mix" servers that blend the received packets so that the communication source (including the sender and the recipient) becomes ambiguous. They rely on the statistical properties of background traffic, also referred to as *cover traffic*, to achieve the desired anonymity. However, these schemes all require public-key cryptosystems and are not suitable for WSNs.

Broadcasting-based schemes provide SLP by mixing the valid messages with the dummy messages so that they become indistinguishable to the adversaries [4]. In a practical situation, the dummy messages can be significantly more than the valid messages, which not only consume a significant amount of the limited energy, but also increases network collisions and decreases the packet delivery ratio. Therefore, these schemes are not quite suitable for large-scale sensor networks.

Providing SLP through dynamic routing is, in our opinion, one of the most feasible approaches in WSNs [5–7]. The main idea is to prevent the adversaries from tracing back to the source location through traffic monitoring and analysis. A representative example of a routing-based protocol is the phantom routing protocol, which involves two phases: a random walk phase and a subsequent flooding/single-path routing phase. In the random walking phase, the message from the actual source will be routed to a phantom source along a random path or a designed directed path. The phantom source is expected to be far away from the actual source, which will make the actual source's location difficult to be traced back by the adversaries. However, theoretical analysis shows that if the message is routed h hops randomly, it is highly possible that the distance between the phantom source and the actual source is within $h/5$. To solve this problem, directed walk, through either a sector-based or a hop-based approach, was proposed. Take the section-based directed walk for example. The source node first randomly determines a direction that the message will be sent. This direction information is stored in the header of the message. Every forwarder on the random walk path will forward this message to a random neighbor in the same direction to ensure that the phantom source will be away from the actual source. Unfortunately, once the message is captured on the random walk path, the adversaries will be able to get the direction information stored in the header of the message. Therefore, exposure of the direction information decreases the complexity for adversaries to trace back to the actual message source in the order of 2^h.

In the schemes discussed in [8], messages are routed to a single randomly selected intermediate node (called totally random RSIN). While providing great SLP, it does not perform well because the routing paths tend to be long. To balance the need to SLP and efficiency, a constraint intermediate node selection scheme is proposed in [8]. In this scheme, the intermediate node is selected following a normal distribution [9] to be away from the source node for a minimum distance $d_{\min}$. It has been analyzed that the probability for the intermediate node to be within the radius $3d_{\min}$ to source node is 95.45%, and 99.73% for radius $4d_{\min}$, respectively. This scheme is called the (constraint) RSIN scheme. The (constraint) RSIN scheme can provide local SLP but not global SLP. The drawbacks for this scheme is that the possibility for the intermediate nodes to be selected is proportional to the distance between the intermediate nodes, and the source node are highly likely to be concentrated in an area surrounding the source node, but with minimum distance $d_{\min}$ away from the source.

3. NETWORK MODELS AND DESIGN GOALS

Source-location privacy is a key security requirement for military and many civilian applications. In the asset monitoring model, WSNs can be used to monitor the activities or presence of animals in a wild animal habitat. However, the information should be kept unavailable to illegal hunters. In military intelligence networks, to protect the message source, both the message source and the routing path have to be protected from adversarial attacks.

Before we describe our proposed SLP scheme in WSNs, we will provide the system model and adversarial model in this section to capture the relevant features of WSNs and potential adversaries in SLP applications.

3.1. The system model

The system is similar to the Panda–Hunter Game that was introduced in [10,5]. In the Panda–Hunter Game, a WSN is deployed in a habitat to monitor the location of a panda. The sensors are used to locate the general area of the panda. As soon as the panda is discovered, the corresponding source node will observe and report data periodically to the sink node. However, the source location should be kept secure from illegal hunters who may try to track and locate the panda. The goal is to make it infeasible for the hunters to determine the location of the panda by analyzing the traffic patterns in the network.

The following assumptions are made about the system:

- The network is divided into grids. The sensor nodes in each grid are fully connected. In each grid, there is one header node responsible for communicating with other nearby header nodes. The whole network is fully connected through multi-hop communications [11–14].
- The sink node is the destination location where data messages will be routed. The information of the sink node is made public. On detecting an event, a sensor node will generate and send messages to the sink node through multi-hop routing.
- Each message will include a unique dynamic ID where the event was generated. Only the sink node can determine the source node location based on the dynamic ID.
- The sensor nodes are assumed to know their relative locations and the sink node location. We also assume that each sensor node has the knowledge of its adjacent neighboring nodes. The information about the relative location of the sensor domain may also be broadcasted through this network for routing information update [15–18].
- The key management, including key generation, key distribution, and key update, is beyond the scope of this paper. However, the interested readers are referred to reference such as [19–21] for more information.

3.2. The adversaries model

Motivated by the high profits related to panda hunting, the adversary would use the most advanced equipment, which means that they would have some technical advantages over the sensor nodes. In this paper, the adversary has the following characteristics:

- *Well-equipped*: The adversary does not need to worry about the energy consumption and has adequate computation capability. On detecting an event message, the adversary could determine the immediate sender of this message by analyzing the strength and direction of the signal received. The adversary is able to move to this sender's location without much delay and has enough memory to store any useful information. If needed, the adversary could compromise some sensor nodes in the network. We also assume that an will never miss the event, such as a panda, when they are in close proximity of the event.
- *Passive*: To prevent from being detected by the anti-hunting officials, the adversaries should not tamper any contents of the messages transmitted in the sensor network, or do any damage to the equipment, but only carry out some passive attacks, which only involve eavesdropping work.
- *Traffic monitoring*: The adversary is able to monitor the traffic in an area and receive all messages in this area. However, we assume that the adversary is unable to monitor the entire network. If the adversary can monitor all the traffic through the network, he or she can just monitor the events directly without relying on monitoring of the sensor network.

3.3. Design goals

Our design goal can be summarized as follows:

- Build a security evaluation model to facilitate the designing and analyzing of routing-based source-location protection schemes.
- The adversaries should not be able to get the source-location information by analyzing the traffic pattern.
- The adversaries should not be able to get the source-location information even if they are able to monitor a certain area of the sensor network and compromise a few network nodes.
- Only the sink node is able to identify the source-location through the messages received. The recovery of the source-location from the received message should be very efficient.
- The length of each message should be as short as possible to preserve the sensor nodes energy. This is because on average, transmission of one bit consumes about as much power as executing 800–1000 instructions [22].

4. PRELIMINARY: SOURCE-LOCATION PRIVACY EVALUATION MODEL

In [23], security analysis of SLP based on quantitative measurement on information leakage of the source-location has been proposed. The quantitative measurement divides information leakage analysis into three categories:

(1) *Correlation-based source identification attack*: Correlation-based attack is an ID-based source node determination. When an adversary receives a message with an ID whose location is already known, the location of this node is also known.
(2) *Routing traceback attack*: Routing traceback is an attack that when an adversary captures a message, he or she can identify the immediate message sender and quickly move to it. For fixed path routing of length n, if the adversary can capture n messages from this source, then he or she is able to locate the message source mode.
(3) *Reducing source space attack*: Reducing source space attack refers to the attack that the adversary can limit the source node to a proper subset/area in the networks when a message is captured. When multiple messages are captured, the subset/area may be further reduced so that the source-location can be limited to a subset/area that may lead to a relatively easy or complete source identification.

To prevent correlation-based source identification, a dynamic ID-based approach can be used to prevent adversaries from relating messages transmitted from each source [6]. This can be carried out by requiring that each node in the network be preloaded with an *ID-hash-chain* so that a different and uncorrelated ID is attached to each message. The adversaries are no longer able to get any useful information about the source node through correlation-based source identification.

For routing traceback and reducing source node space analysis, three criteria have been defined in [23].

Definition 1 *Source-location Disclosure Index (SDI)*
SDI measures, from an information entropy point of view, the amount of source-location information that one message can leak to the adversaries.

For a routing scheme, if we assume that the total privacy for a source node S is 1 and the SDI is fixed, then the adversary only needs to receive $\lceil 1/SDI \rceil$ messages initiated from S in order to successfully locate S. Therefore, for a good SLP scheme, SDI should be as small as possible.

Definition 2 *Source-location Space Index (SSI)*
SSI is defined as the set of possible network nodes, or area of the possible network domain, that a message can be transmitted from.

For a routing scheme, if SSI is large, it means that the message may be transmitted by many possible source nodes. On the contrary, if SSI is small, then the adversary can limit the possible source nodes to a small group. Therefore, for an SLP scheme, SSI should be as large as possible so that the complexity for an adversary to perform an exhaustive search of the message source is maximized.

Definition 3 *Normalized Source-location Space Index (NSSI)*
NSSI is defined as the ratio of the SSI area over the total area of the network domain. Therefore, NSSI $\in [0, 1]$, *and we always have NSSI* $= 1 - \delta$ *for some* $\delta \in [0, 1]$. *The* δ *is called the local degree.*

It is clear that the scheme with the local degree 0 provides the highest degree of SLP.

With these definitions, we have derived in [23] that fixed path routing schemes are the least secure SLP schemes.

Lemma 1. *Suppose that there is a fixed routing path between the source node S and the destination node D of length L hops. A is an adversary who can detect all messages transmitted to D. Then, after receiving L messages, A will be able to trace back to the source node S, that is,*

$$SDI = \frac{1}{L}$$

If there are n disjoint routing paths between the source node S and destination node D, and the length of the n paths are $L_1, L_2, \ldots, L_n$, respectively. For each message, the source node S will send it along path L_i with probability p_i, where

$$\sum_{i=1}^{n} p_i = 1$$

For path i, we have $SDI_i = \frac{p_i}{L_i}, i = 1, \ldots, n$. Define the overall SDI as

$$SDI = \sum_{i=1}^{n} p_i \cdot SDI_i$$

We will then have the following result [23].

Theorem 1. *Suppose there are n disjoint routing paths between the source node S and the sink node D. The lengths of the n routing paths are* $L_1, L_2, \ldots, L_n$. *Let* p_i *be the probability that messages will be transmitted along the path* L_i, *then when* $p_i = \frac{L_i}{L_1 + L_2 + \cdots + L_n}, i = 1, 2, \ldots, n$, *the SDI is minimized, which is*

$$SDI = \frac{1}{L_1 + L_2 + \cdots + L_n}$$

Corollary 1. *Suppose that there are n disjoint routing paths between the source node S and the destination node D.*

The length of the n routing paths are $L_1, L_2, \ldots, L_n$, respectively. The adversary then needs to receive on average

$$\frac{1}{SDI} = L_1 + L_2 + \cdots + L_n$$

messages to fully determine the location of the source node, that is, trace back to the source node.

As a result, to provide SLP in WSNs, we have to increase the total number of possible routing paths between the destination node and the source node. However, for a practical network configuration, the number of routing paths cannot be increased without limitation. This means that we will always have $SDI > 0$.

We can summarize the two defects of the SLP schemes through fixed routing paths as follows:

- Non-zero *SDI*: For fixed path routing, no matter how dedicated the scheme is designed, *SDI* is always larger than 0. In other words, for each message sent out by one source node, from a probability point of view, there is always a fraction of source information to be leaked to the adversaries. So, no matter how small the *SDI* is, when enough messages are received, the adversaries are always able to locate the source node.
- Limited *SSI*: Because the routing paths are fixed for the source node, the correlation between the messages transmitted on a particular path and the source node is high. In other words, *SSI* is small compared with the overall sensor network size.

Phantom routing is a dynamic routing scheme. In this scheme, the message is first routed to a phantom source through a random path before it is forwarded to the actual destination node. To make sure that the phantom source is away from the actual source node, the direction information must be stored in the message's header. In this way, the intermediate nodes on the routing path are able to select the next forward node on the routing path along the same direction.

In [23], each sensor node is assumed to have a unique ID that corresponds to a physical location. The problem of this design is that it is possible for the adversaries to monitor and link all messages from the same source node together, which may help the adversaries to identify the source-location because the IDs correspond to the grids' locations. Therefore, we have

$$SDI > 0$$

Whenever the adversaries discover a message sent from a grid with an ID that they already know, they can use this message to move closer to the message source. Fortunately, this problem can be easily solved if a dynamic ID is assigned for each message. In this case, the correlation between the source node and the message received in the random path can be viewed as zero, that is,

$$SDI \simeq 0$$

However, the direction information stored in the message header can facilitate the adversary to narrow the possible area of the source node. Take the section-based random walk as an example. Once a message is corrupted by an adversary on the random path, the adversary can determine to which direction of the current location the actual source node is located. Therefore, we have

$$NSSI < 1$$

When multiple adversaries collaborate in the target area *T*, the *NSSI* can be further reduced and the SLP is no longer well protected.

5. PROPOSED SCHEME

In this section, we will present our proposed secure SLP scheme for WSNs.

5.1. STaR routing scheme

In STaR SLP scheme, to prevent the adversaries from getting any useful source-location information through *correlation-based source identification*, a dynamic ID proposed in [6] should be assigned for each message. STaR provides SLP through a two-phase routing protocol. In the first phase, the source node routes the message to an RSIN located in a pre-determine region around the sink node. We call this region the Sink Toroidal Region. The random intermediate node services as a fake source when the message is forwarded to the sink node. The intermediate node will forward the message to the sink node by single-path routing in the second phase. The combination of these two phases guarantees the local degree to be small, therefore, providing a high degree of SLP.

In our scheme, the network is evenly divided into small grids [6]. We assume that the sensor nodes in each grid are all within the direct communication range of each other. In each grid, the header node coordinates the communication with other header nodes nearby. We assume that the whole network is fully connected through multi-hop communications.

The goal of the proposed scheme is to provide local and global SLP with adequate energy-efficient routing. Local privacy is obtained by the fact that the intermediate node is expected to be neither too close, nor too far from the real source, for most cases. The STaR area would be a large area with at least a minimum radius distance *r* from the sink node to provide global privacy. Also, the STaR area guarantees that the intermediate node is at most a maximum distance *R* from the sink node to limit the energy consumption in the routing paths. This routing scheme is designed to give the illusion that the source node is sending messages to the sink node from all the possible directions. In this way, the STaR creates an effect that is similar to the totally random RRIN scheme [8] but with less energy consumption and shorter delays.

We assume that each sensor node only has knowledge of its adjacent nodes and has no accurate information of the sensor nodes more than one hop away. We also assume that each node has knowledge of the parameters that are shown in Figure 1. The descriptions of the parameters are as follows:

- x_0, y_0: The corresponding X and Y coordinates of the sink node location.
- R: The pre-determined radius from the sink to the outer-edge of the STaR area.
- r: The pre-determined radius from the sink node to the inner-edge of the STaR area.

From these parameters, $\{x_0, y_0, R, r\}$, the source nodes are able to generate random points within the STaR area. Because we assume that the sink node is located at the relative location (x_0, y_0), the source node selects the random location (x, y) according to the following two steps:

(1) Randomly select d uniformly from $[r, R]$.
(2) Randomly select θ uniformly from $[0, 2\pi]$.

In this way, we can calculate the coordinate of the intermediate node as $(x, y) = (x_0 + dcos(\theta), y_0 + dsin(\theta))$.

After obtaining the random location (x, y), the message can then be routed towards the grid at location (x, y). Because each node only knows its adjacent neighbor nodes' relative location, it can determine the direction that the message should be routed. Once the message is within the desired grid of the random location, the message is routed to the header node of the grid. The header node then becomes the random intermediate node. If the desired grid does not contain any nodes, then the last node in the routing path would become the desired location and the header node in that grid would become the intermediate node. The intermediate node then routes the received message to the sink node using single-path routing.

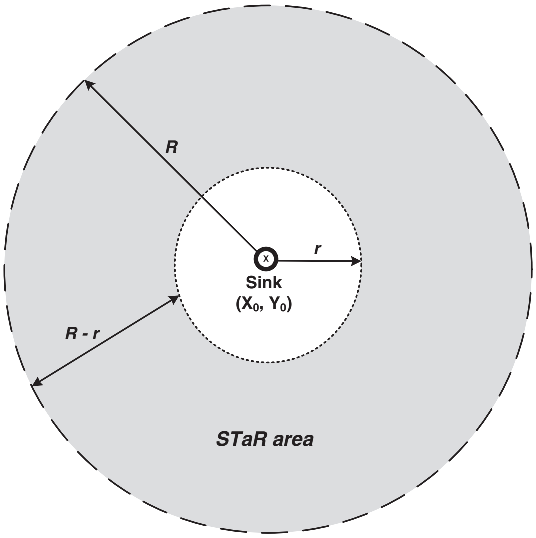


Figure 1. Distribution of the Sink Toroidal Region (STaR) area.

The proposed scheme will provide adequate SLP because it will repeat this procedure for every message sent out. In general, the source node will send out messages periodically. For every message, the source node will choose a new intermediate node within the STaR area using the procedure described earlier.

6. SECURITY ANALYSIS

We will analyze the SLP for STaR routing scheme. We assume that the adversary is unable to monitor the entire sensor area of the source node because, otherwise, it can monitor the actual event directly.

In the proposed STaR routing scheme, a random intermediate node is selected for each message from the STaR area shown in Figure 1. Unlike the directed walk of the phantom routing scheme, our protocol does not leak direction information to the adversaries.

Theorem 2. *For the proposed STaR scheme, if it is assumed that the STaR area is large enough so that the probability for multiple messages to be routed using the same intermediate node is negligible, then the amount of source-location information that can be leaked from one message is negligible, that is,*

$$SDI \simeq 0$$

and SLP with local degree 0.

Proof. For the proposed STaR scheme. The source-location disclosure index can be analyzed in two scenarios based on the location of the adversarial attacks: (i) the adversary monitors traffic between the RSIN in STaR area and the sink node, and (ii) the adversary monitors traffic between the source node and the RSIN in the STaR area.

For scenario (i), first, the STaR area is at least an r distance away from the sink node. Second, every node within the STaR area from the sink node has equal probability in being selected as the intermediate node for all messages by all possible source nodes. Therefore, the messages are being routed from the STaR area to sink from all directions with equal probability, as shown in Figure 2. Therefore, the messages to be transmitted from the STaR area to the sink node area bear no correlation with the actual message source, and it is impossible for an adversary to gain any information of source-location for the message source.

It is also impractical for the adversary to perform routing traceback to figure out the source location by only monitoring and analyzing traffic patterns around the sink node. In this scenario, the global SLP can be assured, and hence, SLP with a low local degree can be guaranteed.

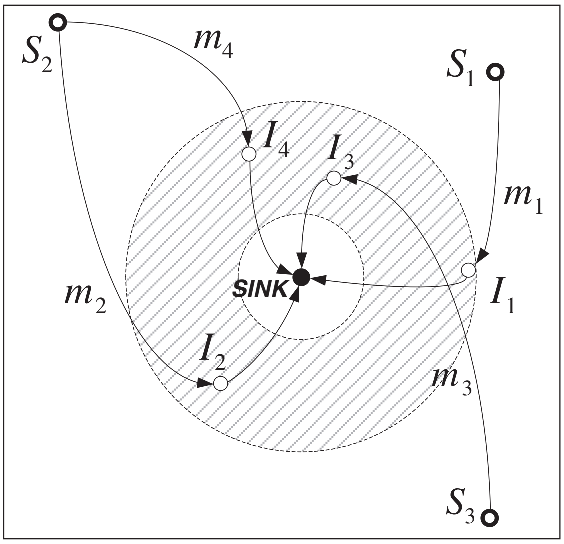


Figure 2. Routing illustration of the Sink Toroidal Region protocol.

For case (ii), the message source may be located anywhere in the network domain, and the intermediate node is expected to be far from the real source for most cases. The probability for the adversary to intercept a message is very low for large WSNs. The probability for an adversary to intercept multiple messages from the same source in the same location is negligible. In fact, if a dynamic ID is used for each message, then even if an adversary is able to intercept one message, he is still unable to link them together.

When only one message is intercepted, the adversary can only determine the immediate previous node on the basis of our assumption. The adversary can determine neither the direction of the next previous hop source node, nor the direction of the actual source node. In other words, the source-location disclosure index equals to 0 with a negligible exception for the nodes in the area of the adversary. Therefore, we have

$$SDI \simeq 0$$

and the SLP local degree is 0.

Theorem 3. *In the proposed STaR scheme, for any received message, the source node is either located in the area that the adversary can monitor, or the adversary has no information of the actual message source. That is*

$$NSSI \simeq 1$$

Proof. Similar to Theorem 2. The analysis can be divided into two scenarios: (i) the adversary monitors traffic between the RSIN in STaR and the sink node, and (ii) the adversary monitors traffic between the source node and the RSIN in the STaR area.

For case (i), when an adversary receives a message, if he is unable to find the message source in his area, then the source node can be in any direction and hop distance away from the location where the message is received.

For case (ii), as shown in Figure 3, when a message is received in location A, on the basis of our assumption, the adversary can find the immediate message node, say B. However, the only information that the adversary can get for the nodes prior to B are located in the shaded area I, $I+II$, and $I+II+III$, and so on. Because the adversary can determine neither the direction nor the hop distance of the actual source node, the actual message source node can be located anywhere of the sensor domain except the area centered at A. In this way, we have

$$NSSI \simeq 1$$

7. PERFORMANCE ANALYSIS AND SIMULATION RESULTS

To evaluate the performance of the schemes, extensive simulations have been conducted using ns-2 on Red Hat Linux system. The results of the simulations are shown in Figure 4. In the simulation, 400 nodes are randomly distributed in a square target area of size 3360 m × 3360 m, whereas the sink node is located at the center of the network. We set hop count of directed walking of phantom routing to be four, which on average the phantom source was found to be 526.12 m away from the real source.

In the simulations, we compared our proposed STaR with several existing schemes: (i) phantom routing; (ii) routing to a single RSIN with the intermediate node selected following the normal distribution [9]; (iii) routing to a single RSIN anywhere in the sensor domain (Total RSIN); and (iv) routing

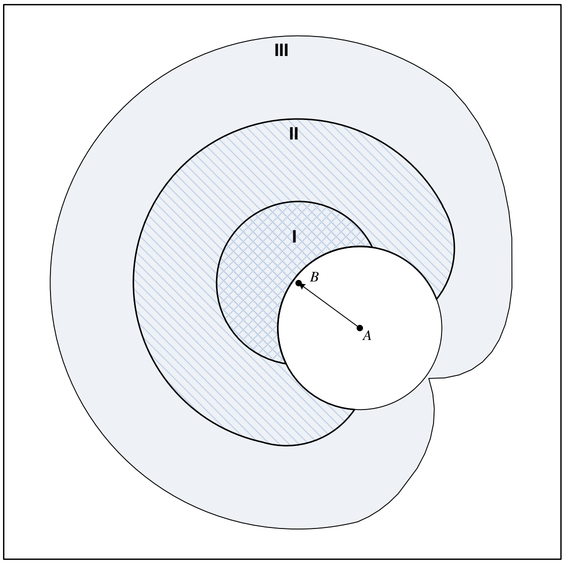
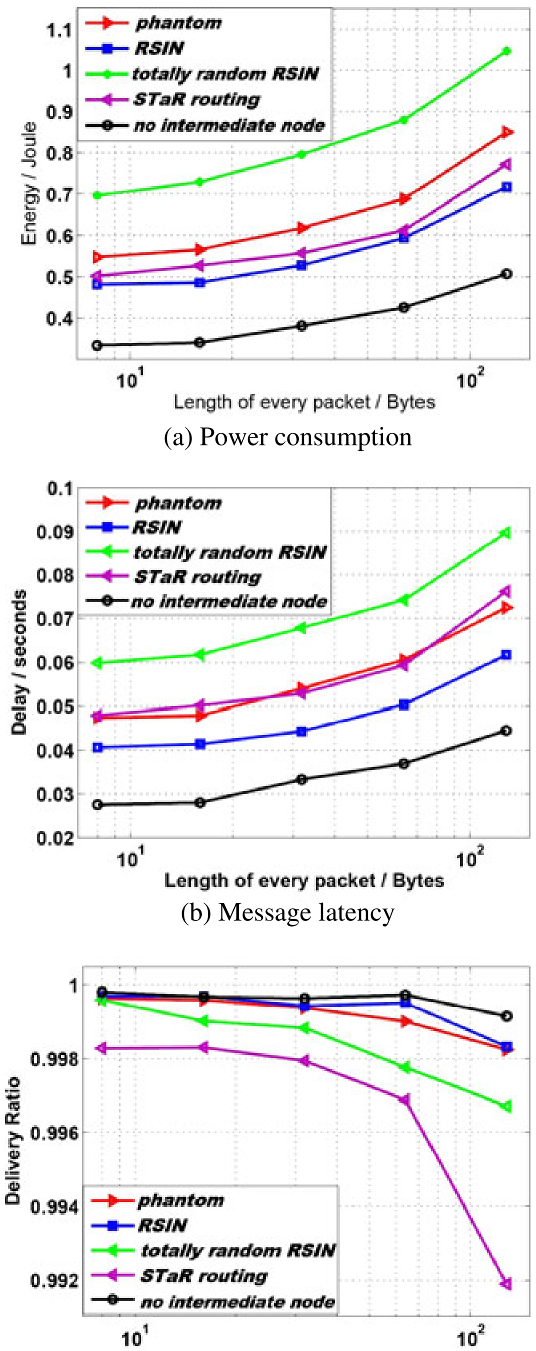


Figure 3. The source-location analysis of Sink Toroidal Region routing scheme.



(a) Power consumption

(b) Message latency

(c) Message delivery ratio

Figure 4. Performance of routing by single-intermediate node. STaR, Sink Toroidal Region; RSIN, randomly selected intermediate node.

from the source node to the sink node directly without an intermediate node (no intermediate node).

For the RSIN scheme, the minimum distance between the source node and the intermediate nodes was set to 480 m, and the average distance turned out to be 529.14 m. For our proposed STaR routing, the inner radius, r, was set to 480 m, whereas the outer radius, R, was set to 640 m.

Figure 4 gives the compared simulation results on the power consumption, message latency, and message delivery ratio. Through analysis and simulation results, we find that direct routing without intermediate node has the best performance, whereas totally random RSIN has the worst performance. The performance of the RSIN scheme is better than phantom routing for comparable security because the average routing paths in phantom routing is longer than the RSIN because of the more curved routing paths. The performance of STaR is between the totally random RSIN and RSIN. The delivery ratio for STaR is slightly lower than the two RSIN schemes because of the possible higher collisions ratio.

8. CONCLUSIONS

Source-location privacy is vital to the successful deployment of WSNs. In this paper, we have analyzed the quantitative measurements SLP for routing-based schemes. With criteria of quantitative measure SLP, we have proposed a scheme that can achieve SLP in WSNs through a STaR routing scheme for local and global SLP protection. We carried out theoretical analysis to evaluate the security and the performance of the proposed schemes and compared it with other existing schemes. Our simulation results demonstrate that the proposed STaR routing scheme can achieve excellent performance in energy consumption, delivery latency, and message delivery ratio.

ACKNOWLEDGEMENTS

This research was supported in part by the NSF under grants CNS–0845812 and CNS–1050326, and CNS–1117831.

REFERENCES

1. Chaum D. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM* 1981; **24**(2):84–88.
2. Reed M, Syverson P, Goldschlag D. Anonymous connections and onion routing. *IEEE Journal on Selected Areas in Communications* 1998; **16**(4): 482–494.
3. Reiter M, Rubin A. Crowds: anonymity for web transaction. *ACM Transactions on Information and System Security* 1998; **1**(1):66–92.
4. Shao M, Yang Y, Zhu S, Cao G. Towards statistically strong source anonymity for sensor networks. In Proceedings of IEEE INFOCOM 2008. April2008; 51–55.
5. Kamat P, Zhang Y, Trappe W, Ozturk C. Enhancing source-location privacy in sensor network routing. In Proceedings of IEEE ICDCS 2005. June2005; 599–608.

6. Li Y, Ren J. Preserving source-location privacy in wireless sensor networks. In Proceedings of IEEE SECON 2009. Rome, Italy, June 22–26,2009.
7. Li Y, Ren J. Source-location privacy through dynamic routing in wireless sensor networks. In Proceedings of IEEE INFOCOM 2010. San Diego, USA, March 15–19,2010.
8. Li Y, Lightfoot L, Ren J. Routing-based source-location privacy protection in wireless sensor networks. In IEEE EIT 2009. Windsor, Ontario, Canada, June 7–9,2009.
9. Wikipedia. Normal distribution. http://en.wikipedia.org/wiki/Normal_distribution.
10. http://www.panda.org/.
11. Ye M, Li C, Chen G, Wu J. EECS: an energy efficient clustering scheme in wireless sensor networks. In Performance, Computing, and Communications Conference, 2005. IPCCC 2005. 24th IEEE International. April2005; 535–540.
12. Heinzelman WB. Application-specific protocol architectures for wireless networks. PhD thesis, Supervisor-Anantha P. Chandrakasan and Supervisor-Hari Balakrishnan, 2000.
13. Neander J, Hansen E, Nolin M, Bjorkman M. Asymmetric multihop communication in large sensor networks. International Symposium on Wireless Pervasive Computing 2006, Jan. 2006.
14. Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing* 2004; **3**:366–379.
15. Zhang Y, Liu W, Fang Y, Wu D. Secure localization and authentication in ultra-wideband sensor networks. *IEEE Journal on Selected Areas in Communications* 2006; **24**:829–835.
16. Zhang P, Martonosi M. LOCALE: Collaborative localization estimation for sparse mobile sensor networks, *International Conference on Information Processing in Sensor Networks (IPSN)* April 2008; 195–206.
17. Srinath T. Localization in resource constrained sensor networks using a mobile beacon with in-ranging. IFIP International Conference on Wireless and Optical Communications Networks, 2006.
18. Chen D-min, Zhang Y. Research of WSN localization algorithm based on entropy function. *First International Workshop on Education Technology and Computer Science (ETCS)* 2009; **1**: 229–233.
19. Chan H, Perrig A. PIKE: peer intermediaries for key establishment in sensor networks. *Proceedings IEEE INFOCOM 2005* 2005; **1**:524–535.
20. Perrig A, Szewczyk R, Wen V, Culler D, Tygar J. SPINS: Security protocols for sensor networks. In MobiCOM 2001. Rome, Italy, July2001.
21. Traynor P, Kumar R, Choi H, Cao G, Zhu S, La Porta T. Efficient hybrid security mechanisms for heterogeneous sensor networks. *IEEE Transactions on Mobile Computing* 2007; **6**:663–677.
22. Hill J, Szewczyk R, Woo SHA, Culler D, Pister K. System architecture directions for networked sensors. In Proceedings of ACM ASPLOS IX. November2000.
23. Li Y, Ren J, Wu J. Quantitative measurement and design of source-location privacy schemes for wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems* accepted to appear.