R-STaR Destination-Location Privacy Schemes in Wireless Sensor Networks

Leron Lightfoot   Jian Ren
Department of Electrical and, Computer Engineering, Michigan State University
East Lansing, MI 48824
Email: {lightf13, renjian}@egr.msu.edu

***Abstract*—Wireless sensor networks (WSNs) can provide the world with a technology for real-time event monitoring for both military and civilian applications. One of the primary concerns that hinder the successful deployment of wireless sensor networks is providing adequate location privacy. Many protocols have been proposed to provide location privacy but most are based on public-key cryptosystems, while others are either energy inefficient or have certain security flaws. In this paper, after analyzing security weakness of the existing schemes, we propose an architecture that addresses the security flaw for destination location privacy in WSNs based on energy-aware two phase routing protocol. We call this scheme the R-STaR routing protocol. In the first routing phase of R-STaR routing, the source node transmits the the message to a randomly selected intermediate node located in a pre-determined region surrounding the source node, which we call the R-STaR area. In the second routing phase, the message is routed to the destination node using shortest path mix with fake message injections. We show that R-STaR routing provides a exceptional balance between security and energy consumption in comparison to existing well-known proposed schemes.**

I. Introduction

A Wireless sensor network is a network of sensor devices that are connected through shared wireless communication mediums. It has been envisioned as a technology with great potential to be widely used in military, health and event-monitoring applications, such as temperature, movement, sound, pressure, etc. Wireless sensor networks rely on wireless communications, which is by nature a broadcast medium and is more vulnerable to security attacks than its wired counterpart due to lack of a physical boundary. Security attacks, such as, privacy threats has been an extensively studied topic in Wireless Sensor Networks (WSNs). Privacy threats can be classified into two categories: (i) Content-based privacy threats, and (ii) Context-based privacy threats. Content-based privacy threats relates to protecting the content of the message and can be protected using cryptographic encryption algorithms. Context-based privacy threats relates to monitoring the transmission of the data (i.e. Eavesdropping and localization attacks). For instance, an adversary may be able to analyze the traffic patterns by monitoring the transmission of the data using radio frequency localization techniques. Unfortunately, using cryptographic techniques does not provide protection for context-based threats and presents a much greater challenge to solve. In this paper, we will focus on one particular context-based threat, routing-based destination location privacy.

In providing adequate destination-location privacy, the sensor devices present major limitations. Sensors in the network are meant to be low-cost and energy efficient devices, therefore, powerful transmitters are not suitable for WSNs. Using low transmission multi-hop routing approach is more desirable in WSNs because it requires less energy usage of the sensor devices and would prolong the network lifetime. The tradeoff in using multi-hop routing approach is that the network is more vulnerable to context-based localization attacks. An adversary can track the destination location sink node by analyzing the routing paths of messages transmitted from a source node. To prevent an adversary from determining the destination location in WSNs, we must create an energy-aware multi-hop routing algorithms that would make it infeasible for an adversary to determine the destination location by analyzing routing paths.

In this paper, we first analyze the security vulnerabilities of some existing destination location privacy schemes. Then we propose a two-phase destination-location privacy (DLP) routing scheme called R-STaR routing. R-STaR provides energy-efficient routing-based destination location privacy for WSNs.

The remainder of this paper is organized as follows: In Section II, we review existing studies. Section III, we describe the system model and design goals. The proposed R-STaR destination-location privacy scheme is presented in Section V. Section VI discusses security analysis of the R-STaR routing scheme. Evaluated simulation results is provided in Section VII. Section VIII concludes this paper.

II. Related Works

Many previous proposed privacy schemes [1]–[5] stem from the Chaum's mixnet [6] and DC-net schemes. These mixnet family schemes use a set of mix servers that blends received messages so that the senders and recipients become ambiguous. To accomplish the ambiguity between the senders and recipients, the schemes uses statistical properties of background traffic. They refer to the background traffic as cover traffic. Nevertheless, these schemes all rely on public-key cryptosystem and are not fitting for WSNs.

Providing location privacy through dynamic routing is, in our opinion, one of the most feasible approaches in WSNs. The main idea is to prevent the adversaries from monitoring the traffic to determine the location of a source or destination node. In [5], the DEEP (Differential Enforced Fractal Propagation) routing protocol is introduced to provide destination location privacy in WSNs. This scheme combines random walking to the destination and injecting of fake messages into the network to create hot spots to protect the destination location. The LPR (Location-Privacy Routing) with fake message injection is proposed in [7]. The LPR protocol randomizes the routing paths so that the forwarding direction of the real message is not always towards the destination. This scheme is similar to the DEEP routing protocol but does not create hot spots and try to solve the vulnerability of the DEEP routing.

Like destination-location privacy schemes, source-location privacy [8], [9] is also a vital component to context-based attacks. Many of the techniques used for protection of the source node can not be applied for protection of the destination node. Source-location privacy is beyond the scope of this paper but we can apply some of the attributes of source location privacy schemes for destination location privacy, such as using fake message injection and two-phase routing techniques. In this paper, we focus on providing routing-based destination-location privacy.

III. Network Models and Design Goals

To get a better understanding of the network, in this section we will provide the system model, adversarial model, and design goals to capture the relevant features of WSNs and potential adversaries in DLP applications.

A. The System Model

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.
- Every node in the network can become a source node on a detection of an event. On detecting an event, a sensor source node will generate and send messages to the destination node through a multi-hop routing.
- Each message will include a unique node ID where the event was generated. The SINK node can only determine source node location based off the node ID.
- The sensor nodes are assumed to know their relative location and destination 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.
- The key management, including key generations, key distribution and key update, is beyond the scope of this paper. However, the interested readers are referred to references such as [10].

B. The Adversaries Model

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 a transmitted message, the adversary could determine the receiver node by waiting to see which neighbor node retransmit the message. The adversary is able to move to this receivers 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.
- ***Passive:*** The adversaries 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.

C. Design Goals

Our design goals can be summarized as follows:

- The adversaries should not be able to get the destination-location information by analyzing the traffic pattern.
- The adversaries should not be able to get the destination-location information even if they are able to monitor a certain area of the sensor network and compromise a few network nodes.
- The length of each message should be as short as possible to save the previous sensor node power.

IV. Preliminary: Destination-Location Privacy Evaluation Model

In this section, we will provide security analysis of destination-location privacy based on quantitative measurement on information leakage of the destination-location. The quantitative measurement divides information leakage analysis into three categories:

1) *Correlation-based destination identification attack:* Correlation-based attack is an ID based destination 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 forward attack:* Routing forward is an attack that when a message is forward to the next hop node, an adversary captures a message and identify the immediate message receiver and quickly move to it.
3) *Reducing destination space attack:* Reducing destination space attack refers to the attack that the adversary can limit the destination 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 destination-location can be limited to a subset/area that may lead to the destination node.

To prevent correlation-based destination identification, a dynamic ID based approach can be used to prevent adversaries from relating messages transmitted from each source [9], [11].

For routing traceforward and reducing destination node space analysis, can be defined in three criteria as follow:

Definition 1 (Destination-location Disclosure Index (DDI))**.** DDI *measures, from an information entropy point of view, the amount of destination-location information that one message can leak.*

For a routing scheme, if we assume the total privacy for a destination node D is 1, and the *DDI* is fixed, then the adversary only needs to receive $\lceil \frac{1}{DDI} \rceil$ messages routed to D in order to successfully locate D. Therefore, for a good DLP scheme, *DDI* should be as small as possible.

Definition 2 (Destination-location Space Index (DSI))**.** DSI *is defined as the set of possible network nodes, or sub-area of the network domain, that can contain the destination node.*

For a routing scheme, if *DSI* is large, means that practically the entire network can contain the destination node. On the contrary, if *DSI* is small, then the adversary can limit the network to a small-sub areas of nodes that possible contain the destination node. To provide adequate destination-location privacy, *DSI* should be as large as possible.

Definition 3 (Normalized Destination-location Space Index (NDSI))**.** NDSI *is defined as the ratio of the* DSI *area over the total area of the network domain. Therefore, $NDSI \in [0, 1]$, and we always have $NDSI = 1 - \delta$ 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 DLP.

V. Proposed Scheme

In this section, we will present our proposed secure destination-location privacy schemes for wireless sensor networks.

A. R-STaR Routing Scheme

The R-STaR DLP scheme is similar to the STaR routing scheme discussed in [12]. Instead of the STaR area being located around the SINK node, the STaR area is located around the source node, which we call it the R-STaR area. R-STaR routing scheme provides DLP through a two-phase routing protocol. In the first phase, the source node routes the message to a randomly selected intermediate node located in a pre-determine region around the source node. We call this region the Reverse Sink Toroidal Region (R-STaR). Opposite of the STaR routing scheme, the random intermediate node act as a fake destination node. In the second phase, the intermediate node in the R-STaR area will forward the message to the SINK node using single-path routing.

In our scheme, the network is evenly divided into small grids [11]. 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 DLP with adequate energy-efficient routing. DLP is obtained by using this scheme because the source node will not leak any information of the destination node location as the message is routed to the intermediate nodes in the first routing phase. Since the intermediate node serves as a fake location of the destination node in the first routing phase and can be located any direction of the source node with equal probability, an adversary will be unable to determine the destination node sub-area location by just monitoring messages transmitted near the source node. Therefore, the R-STaR routing scheme can provide adequate global destination-location privacy

Global privacy is obtained by the fact that the messages from the source node to the intermediate node will not provide of sub-area where the destination node may be located. We assume that the R-STaR area would be a large area with at least a minimum radius distance r from the source node. Also, the R-STaR area guarantees that the intermediate node is at most a maximum distance R from the source node to limit the energy consumption in the routing paths in the first phase. This routing scheme is designed to give the illusion that the destination node can be located in all the possible directions. In this way, the R-STaR creates an effect that is similar to the totally random RRIN scheme [8] but with less energy consumption and shorter routes.

We also assume that each node has knowledge of the parameters that are shown in Figure 1.

- x_S, y_S: The corresponding X and Y coordinates of the source node location,
- R_S: The pre-determined radius from the source to the outer-edge of the R-STaR area,
- r_S: The pre-determined radius from the source node to the inner-edge of the R-STaR area.

Since we assume that the source node is located at the relative location (x_S, y_S), the source node determines intermediate node location (x_i, y_i) according to the following steps:

1) Randomly select d_i uniformly from $[r_S, R_S]$.
2) Randomly select θ_i uniformly from $[0, 2\pi]$.
3) Calculate the coordinate of the intermediate node as $(x_i, y_i) = (x_S + d_i \cos(\theta_i), y_S + d_i \sin(\theta_i))$.

After obtaining the random location (x_i, y_i), the message can then be routed towards the grid that contains the location

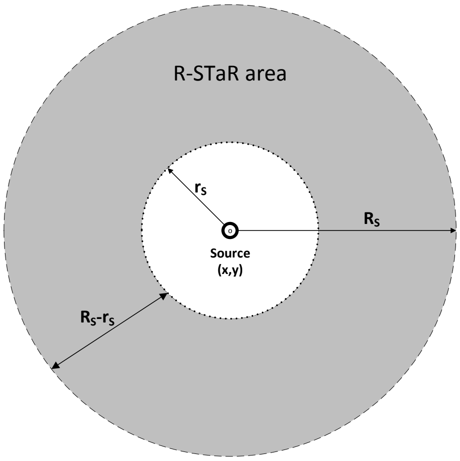

Fig. 1. Distribution of the R-STaR area

(x_i, y_i). Since 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 destination node using single-path routing.

To provide additional security for the second phase routing path, when the real message is received by the destination node, the destination node may inject a fake message into the network to give the illusion that the message is continued to be forward in the network to hide the identity of real destination node. Without this additional security feature, an adversary is able to locate the real destination node location by simply determining where the last hop node of the message routing path. This additional security feature will provide local network privacy for the destination-location.

For conserving the energy of the network, the fake message that is generated by the real destination node will have certain Time-To-Live (TTL) value for the number of hops the it is forward in the network. We will denote the TTL value for the fake message as N_{TTL} and can be determined by the following equation,

$$N_{TTL} = \lceil \alpha \cdot N_{hop} \rceil,$$

where α is a random generated value from $[0, 1]$ and N_{hop} is the hop count of the real message from the source node to the destination node. If the source node is located within 10 hops from the destination node, N_{TTL} will be derived as,

$$N_{TTL} = \lceil (\alpha \cdot N_{hop}) + 10 \rceil, N_{hop} < 10,$$

to ensure that the fake message will forward at least 10 hops away from the destination node when the source node is located is near the destination node.

The fake generated message must take the same path trajectory as the real message as it was routed from the intermediate node to the destination node. And to do so, we must determine the following values: d_f, θ_f, X_f, Y_f. Where (X_f, Y_f) are the coordinates of a random point on the trajectory path for the fake message route. To determine this random point, we must first determine d_f, which is the distance from the destination node to the random fake point (X_f, Y_f). The destination node will determine d_f as follow,

$$d_f = T_r \cdot N_{TTL},$$

where T_r is the transmission range (in meters) of wireless sensor devices. This equations ensure that the random fake point, (X_f, Y_f), will be at least N_{TTL} hops from the destination node.

Second, we must determine the θ_f, which is the angle of message path trajectory from the destination node to the fake random point. To do so, we must determine the slope of trajectory path of the real message from the intermediate node to the destination node, which we will denote this slope as m_{slope}. Let (X_i, Y_i) and (X_d, Y_d) represent the coordinates of the intermediate node and destination node, respectively. Therefore, θ_f can be determine as follow,

$$\theta_f = tan^{-1}(m_{slope}) = tan^{-1}((Y_d - Y_i)/(X_d - X_i)).$$

With d_f and θ_f, we can determine the random point (X_f, Y_f) as follow,

- $X_f = cos(\theta_f) \cdot d_f + X_d$
- $Y_f = sin(\theta_f) \cdot d_f + Y_d$.

Now with the random point (X_f, Y_f), the destination node can send the fake generated message on the same trajectory path as the real message. The fake message will only be forward in the the direction of the random (X_f, Y_f) point for N_{hop} hops. If the fake message reaches a node on the edge of the network, this node in the forwarding process will discard the fake message and act as the destination node of the message.

VI. Security Analysis for R-STaR Routing Schemes

In this section, we will analyze the security for the R-STaR routing. We will analyze the security all three proposed schemes by analyzing the security of each routing phase.

A. *First Routing Phase*

In the first routing phase, the source node route the message to a random intermediate node located in R-STaR area using the shortest routing path. The intermediate node that can be located in any direction not associated with the location of the destination node. Therefore, in the first routing phase, no information is leaked about the destination node location by analyzing traffic in the first routing phase.

Theorem 1. *For the proposed schemes, assume that the R-STaR area is large enough so that the probability for multiple messages to be routed using the same intermediate node is negligible and it is equally likely for the intermediate node to be selected from any directions. Then, the amount of destination-location information that can be leaked from a message in the first routing phase is negligible, i.e.,*

$$DDI \simeq 0,$$

and DLP with local degree 0.

Also, DSI will include all nodes in the network as possible destination nodes. Therefore, $DSI = N$, where N represent the number nodes in the entire network environment, and $NDSI = \frac{DSI}{N} = 1$.

B. Second Routing Phase

To analyze the security for the second routing phase, we assume that the adversary has identified the entire R-STaR area dimensions and is trying to monitor messages in the second routing phase. For simplicity, we further assume that the network environment is a circle region with a radius R_N, as seen in figure 2. To help analyze the security strength of the proposed routing techniques, we will break the entire network environment into 3 regions.

- Region 1: The source node region of the network. This region have a radius r with the source node being the center location of the region.
- Region 2: The R-STaR region of the network. This region have outer radius of R and inner radius of r. It is located around the source node location.
- Region 3: Outside the R-STaR region. This region is the remaining area of the network that is outside of region 1 and region 2.

The second routing phase transmit the real message from the random intermediate node to a destination node. Recall, upon the destination node retrieving each real message, the destination node will generate and transmit a fake message into the network. The fake generated message will be sent in the network for random number of hops, N_{TTL}, on the same trajectory as the real message help conceal the destination node location and identity. To further help analyze the routing techniques, we will break the routing path into three routing phases.

- $S \to I$: The routing path of the real message, m_r, from the source node, S, to the random intermediate node, I.
- $I \to D$: The routing path of the real message, m_r, from the random intermediate node, I, to destination, D.
- $D \to D_f$: The routing path of the fake message, m_f, from the real destination, D, to the fake destination, D_f.

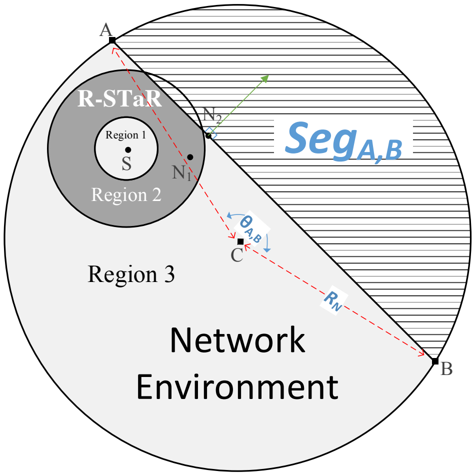

Fig. 2. Illustration of R-STaR routing

With the assumption that the adversary knows location and the dimensions of the R-STaR area, we will analyze how much information an adversary can gain from observing one message transmitted from node N_1, in region 2 (R-STaR area), to node N_2, in region 3. If an adversary observes a communication between a node in the R-STaR area and one outside the R-STaR area, he is able to assume that the message is transmitting on the second or third routing phase. For analyzing purposes, the adversary assumes that the message is being routed $I \to D$. With the described assumptions, we must analyze how much information is leaked to adversary when a message is observed one hop outside the R-STaR area. Figure 2 helps illustrate the communication between node N_1 and node N_2. In Figure 2, the arrow from N_2 shows the assume trajectory of the message path based off the physical location of node N_1 and node N_2. $Seg_{A,B}$ represents the sub area of that an adversary can assume the destination node can be located based off only observing of the communication between node N_1 and node N_2 while assuming the message is on the routing path $I \to D$.

The segment area, $Seg_{A,B}$, can be determine as follows:

$$Seg_{A,B} = \frac{R_N^2}{2} \cdot (\frac{\pi}{180} \cdot \theta_{A,B} - sin(\theta_{A,B})).$$

When using the original R-STaR routing technique for communication between N_1 and N_2, the amount of information that can be leaked in route $I \to D$ can be evaluated as follow:

$$DSI = Seg_{A,B},$$

and

$$NDSI = \frac{DSI}{\text{The Entire Network Area}}.$$

Therefore,

$$NDSI = \frac{Seg_{A,B}}{\pi \cdot R_N^2} = \frac{1}{2 \cdot \pi} \cdot (\frac{\pi}{180} \cdot \theta_{A,B} - sin(\theta_{A,B})).$$

Since messages will be routed out of the R-STaR area using different trajectory routing paths, by simply monitoring the traffic out of the R-STaR area is infeasible for the adversary to pinpoint the destination node location. $Seg_{A,B}$ area can be a relatively large area that an adversary would need to

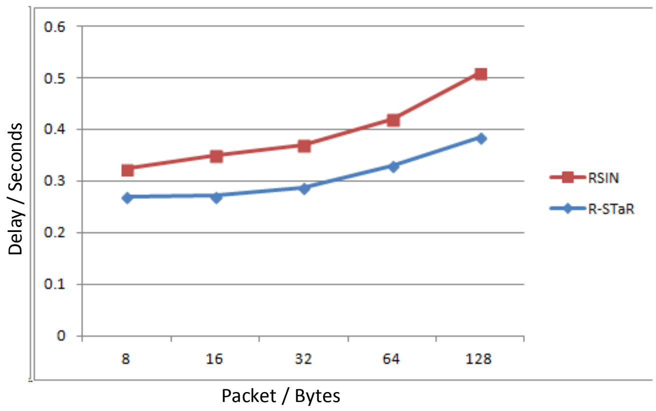

Fig. 3. Message Latency

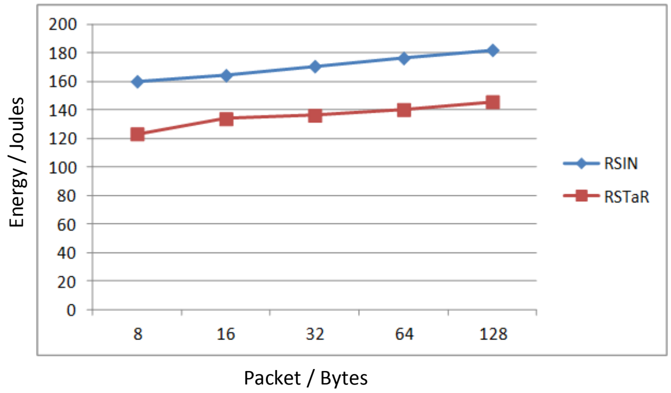

Fig. 4. Energy Consumption

monitor to trace forward to destination node location. Even if an adversary is able to trace near the location of the destination node, the fake packet injection from $D \rightarrow D_f$ still conceals the real destination node location for local network security.

VII. Simulation Results

To evaluate the performance of the schemes proposed, extensive simulations have been conducted using ns-2 on RedHat Linux system. The results of the simulations are shown in Fig. 3 and Fig. 4. In the simulation, 400 nodes are randomly distributed in a square target area of size 2100×2100 meters. We illustrate the performance of the totally randomly selected intermediate nodes (RSIN) and the R-STaR routing protocol. For RSIN scheme, the source node can select any node in the network as the intermediate node with equal probability. For R-STaR routing, the inner radius, r, was set to 395 meters, while the outer radius, R, was set to 605 meters. Through analysis and simulation results, we find that RSTaR provide the best results. R-STaR provides similar level of security as the RSIN scheme but with less energy consumption and shorter delays.

VIII. Conclusions

For a successful deployment of wireless sensor networks, location privacy is vital. In this paper, we address the issue of destination-location privacy. We proposed a scheme called R-STaR routing, which we believe solves some of the security vulnerabilities of some existing schemes with small cost to energy consumption. We believe this scheme is practical due to the simplicity of the design. In the future, we will provide some simulation and performance results to show the security strength in protecting the destination-location.

Acknowledgment

This work was supported in part by the NSF under grants CNS-0845812, CNS-1117831, CNS-1217206, and ECCS-1232109.

References

[1] L. von Ahn, A. Bortz, and N. Hopper, "k-anonymous message transmission," in *Proceedings of CCS*, Washington D.C., USA., 2003, pp. 122–130.

[2] S. Goel, M. Robson, M. Polte, and E. Sirer, "Herbivore: A Scalable and Efficient Protocol for Anonymous Communication," Cornell University, Ithaca, NY, Tech. Rep. 2003-1890, February 2003.

[3] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous connections and onion routing," *IEEE J. on Selected Areas in Coomunications*, vol. 16, no. 4, pp. 482–494, 1998.

[4] M. Reiter and A. Rubin, "Crowds: anonymity for web transaction," *ACM Transactions on Information and System Security*, vol. 1, no. 1, pp. 66–92, 1998.

[5] J. Deng, R. Han, and S. Mishra, "Countermeasures against traffic analysis attacks in wireless sensor networks," *Security and Privacy for Emerging Areas in Communications Networks, 2005. SecureComm 2005. First International Conference on*, pp. 113–126, Sept. 2005.

[6] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, February 1981.

[7] Y. Jian, S. Chen, Z. Zhang, and L. Zhang, "Protecting recieiver-location privacy in wireless sensor networks," *INFOCOM 2007. Twenty-six Annual Joint Conference of the IEEE Computer and Communications Societies*, pp. 1955–1963, May 2007.

[8] Y. Li, L. Lightfoot, and J. Ren, "Routing-based source-location privacy protection in wireless sensor networks," in *IEEE International Conference on Electro/Information Technology, IEEE EIT 2009)*, June 2009, pp. 29–34.

[9] L. Lightfoot, Y. Li, and J. Ren, "Star: Design and quantitative measurement of source-location privacy for wireless sensor networks," *Wiley: International Journal on Security and Communication Networks 2012*, 2012.

[10] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and J. Tygar, "SPINS: Security protocols for sensor networks," in *Seventh Annual International Conference on Mobile Computing and Networks (MobiCOM 2001)*, Rome, Italy, July 2001.

[11] Y. Li and J. Ren, "Preserving source-location privacy in wireless sensor networks," in *SECON'09: Proceedings of the 6th Annual IEEE communications society conference on Sensor, Mesh and Ad Hoc Communications and Networks*. Piscataway, NJ, USA: IEEE Press, 2009, pp. 493–501.

[12] L. Lightfoot, Y. Li, and J. Ren, "Preserving source-location privacy in wireless sensor networks using star routing," in *IEEE Global Telecommunications Conference, IEEE GLOBECOM 2010*, December 2010, pp. 1–5.