This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

Preserving Source-Location Privacy in Wireless Sensor Network using STaR Routing

Leron Lightfoot, Yun Li, Jian Ren
Department of Electrical and Computer Engineering
Michigan State University, East Lansing, MI 48824
Email: {lightf13, liyun1, renjian}@egr.msu.edu

***Abstract*—In wireless sensor networks (WSNs), providing source-location privacy through secure routing is one of the most prosperous techniques. In this paper, we propose a routing technique to provide adequate source-location privacy with low energy consumption. We introduce this technique as the Sink Toroidal Region (STaR) routing. With this technique, the source node randomly selects an intermediate node within a designed 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. 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 source-location privacy, our simulation results show that the proposed scheme is very efficient and can be used for practical applications.**

I. Introduction

Wireless sensor networks 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 source-location privacy. The privacy of the source location is vital and highly jeopardized by the usage of wireless communications. When traffic is transmitted wirelessly in the open air, any compatible receivers 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 in network environment. Without an adequate protection of the routing paths, an adversary may be able to determine the source location by using RF 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 the privacy of the message content but also the privacy of the source and destination locations. The focus of this paper is on source-location privacy. 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 source privacy, the sensor devices present major limitations. Sensors in the network are meant to be low-cost and energy efficient devices. 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 the batteries in the sensor devices is 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 along with source-location privacy are two very vital components for the successful deployment of wireless sensor networks.

In this paper, we propose a two-phase routing scheme that addresses the source-location privacy issue by using a unique routing process. In the routing process, the source node randomly determines an intermediate node from a pre-determined region around the SINK node. We call this region the Sink Toroidal Region (STaR). From the random intermediate node, the message will then be routed to the SINK node through the shortest path routing. The STaR routing method is performed for every message the source node sends to the SINK node in the network. We analyze the performance of the proposed STaR routing method and existing methods with network simulations. Our simulation results show that the STaR routing scheme can provide performance comparable to or better than the existing schemes while enhancing source-location privacy.

The remainder of this paper is organized as follows. In Section II, the related works are reviewed. The system model and design goals are described in Section III. Section IV details the proposed source-location privacy scheme. Security and performance analysis are provided in Section V and Section VI, respectively. We conclude in Section VII.

II. Related Works

In the past two decades, originated largely from Chaum's mixnet [1] and DC-net [2], a number of source-location private communication protocols have been proposed [3]–[6]. The mixnet family protocols use a set of "mix" servers that mix the received packets to make the communication source (including the sender and the recipient) ambiguous. The DC-net family protocols [2] utilize secure multiparty computation techniques. However, both approaches require public-key cryptosystems and are not suitable for WSNs.

978-1-4244-5638-3/10/$26.00 ©2010 IEEE

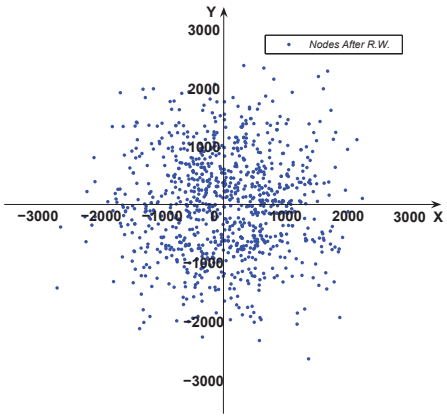


Fig. 1. Nodes distribution through random routing

In [7], base station location privacy based on multi-path routing and fake messages injection was proposed. In this scheme, every node in the network has to transmit messages at a constant rate. Another base station location privacy scheme was introduced in [8], which involves location privacy routing and fake message injection.

In [9], [10], source-location privacy is provided through broadcasting that mixes valid messages with dummy messages. The main idea is that each node needs to transmit messages consistently. Whenever there is no valid message, the node has to transmit dummy messages. The transmission of dummy messages not only consumes significant amount of sensor energy, but also increases the network collisions and decreases the packet delivery ratio. Therefore, these schemes are not quite suitable for large scale WSNs.

Routing-based protocols can also provide source-location privacy through dynamic routing so that it is infeasible for the adversaries to trace back to the source location through traffic monitoring and analysis. The main idea is to, first, route the message to a node away from the actual message source randomly, then forward the message to the SINK node using single path routing. However, both theoretical and practical results demonstrate that if the message is routed randomly for h hops, then the message will be largely within $h/5$ hops away from the actual source. An example is shown in Fig. 1, a source node is located at $(0, 0)$. This source node generates 1000 packets. Each packet is routed 50 hops until it reaches a randomly chosen node. The transmission range is 250 meters at most for one hop. We can see that most of the randomly selected nodes are located not far from the source. The average distance between the source and the randomly selected nodes is 4.2 hops, and the longest distance is just 12.2 hops.

To solve this problem, several approaches have been proposed. In phantom routing protocol [11], the message from the actual source will be routed to a phantom source along a designed directed walk through either sector-based approach or hop-based approach. Take the section-based directed walk as an example, the source node first randomly determines a direction that the message will be sent to. The direction information is stored in the header of the message. Then every forwarder on the random walk path will forward this message to a random neighbor in the same direction determined by the source node. In this way, the phantom source can 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, the exposure of direction information decreases the complexity for adversaries to trace back to the actual message source in a magnitude of 2^h. Random walks from both the source node and the SINK node were also proposed in [12]. In this scheme, Bloom Filter was proposed to store the information of all the visited nodes in the network for each message to prevent the messages from hopping back. However, this design allows adversaries to recover significant routing information from the received messages. In addition, this design is "not realistic" for large scale WSNs.

We have analyzed that phantom routing will leak direction information to the adversaries while the messages are forwarded to the phantom sources. To prevent this, we proposed routing through a randomly selected intermediate node (RRIN) [13]. In this scheme, the source node first randomly selects an intermediate node at the sensor domain based on the relative location of the sensor nodes. The intermediate node is determined by two factors: 1) it must be outside the constrained region around the source and 2) it is normally distributed outside the constrained area. With this method, the selected intermediate node is expected to be away from the real source node, which provides local location privacy. In order to provide both local and global location privacy over the sensor network, the selection of intermediate nodes has to be totally random, i.e., every sensor node in the network has the same probability of being selected as the intermediate node for any source node. Unfortunately, the energy consumption for this design is quite high. In this paper, a design tradeoff has been made to balance these two needs. The intermediate nodes are evenly distributed in the STaR so that the messages can be routed to the SINK node from all possible directions.

III. Models and Design Goals

A. The System Model

The system is similar to the Panda-Hunter Game that was introduced in [11], [14]. In the Panda-Hunter Game, a wireless sensor network 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 [15]–[18].

978-1-4244-5638-3/10/$26.00 ©2010 IEEE

- The SINK node is the destination location that data messages will be routed to. 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 a 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 [19]–[22].
- 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 [23]–[25] for more information.

B. The Adversaries Model

Motivated by the high profits related to panda hunting, the adversary would use the most advanced equipments, which means 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. The adversary also has adequate computation capability. On detecting an event message, he could determine the immediate sender of this message by analyzing the strength and direction of the signal he received. He is able to move to this sender's location without much delay. The adversary also has enough memory to store any information useful to him. If needed, the adversary could compromise some sensor nodes in the network. We also assume that an adversary will never miss the event, such as a panda, when they are close to each other.
- ***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 equipments, 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 which is important in his opinion, and he could get all of the 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 though the network, he can just monitor the events directly without relying on monitoring of the sensor network.

C. Design Goals

Our design goals can be summarized as follows:

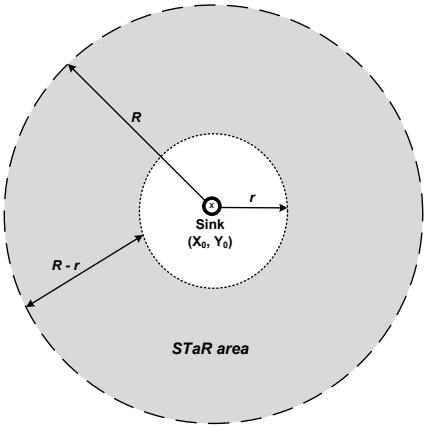


Fig. 2. Distribution of the STaR area

- 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 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.

D. Overview of the Proposed Scheme

When a normal node has a message to transmit, the message will first be transmitted to the header node in the grid. The header node will then forward this message to a randomly selected intermediate node, within a pre-determined region around the SINK node. The message is then forwarded to the SINK node from the intermediate node. The detailed description of the proposed scheme will be described in the subsequent sections.

IV. Proposed Routing-based Source-Location Privacy Scheme STaR

In this paper, we propose a two-phase routing protocol to provide source-location privacy. In the first phase, the source node randomly selects an intermediate node at the sensor domain and routes the message to the random intermediate node. The random intermediate node services as a fake source when the message is forwarded to the SINK node. In this scheme, the random intermediate node would be located in a pre-determine region around the SINK node. We call this region the Sink Toroidal Region (STaR). In the second phase, the intermediate node then forwards the message to the SINK node by single-path routing.

The goal of the proposed scheme is to provide local and global source-location privacy 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 away 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 away 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 [13] but with less energy consumption.

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 perimeters that is shown in Fig. 2. The description of the perimeters 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 perimeters, $\{x_0, y_0, R, r\}$, the source nodes are able to generate random points within the STaR area. Since 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 $(x, y) = (x_0 + d\cos(\theta), y_0 + d\sin(\theta))$.

After obtaining the random location (x, y), the message can then be routed towards the grid at location (x, y). Since each node only knows its adjacent neighbor nodes' relative location, it can determine the direction that the message should be routed to. 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.

The proposed scheme will provide adequate source-location privacy since 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 above.

V. Security Analysis

In this section, we will analyze that the proposed STaR routing scheme can provide source-location privacy.

In our scheme, a random intermediate node is selected from the STaR area. We assume that the STaR area is large enough that it would be unpractical for an adversary to monitor entirely. From the probability point of view, for a large network, the chances that the messages will be routed using the same path and the same intermediate node are extremely low.

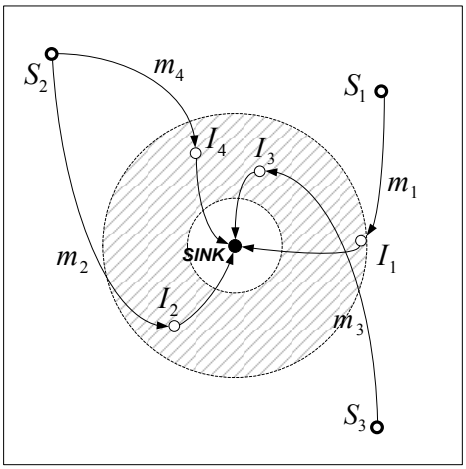


Fig. 3. Routing illustration of the STaR protocol

Unlike the directed walk of the phantom routing scheme, our protocol does not leak direction information to the adversaries.

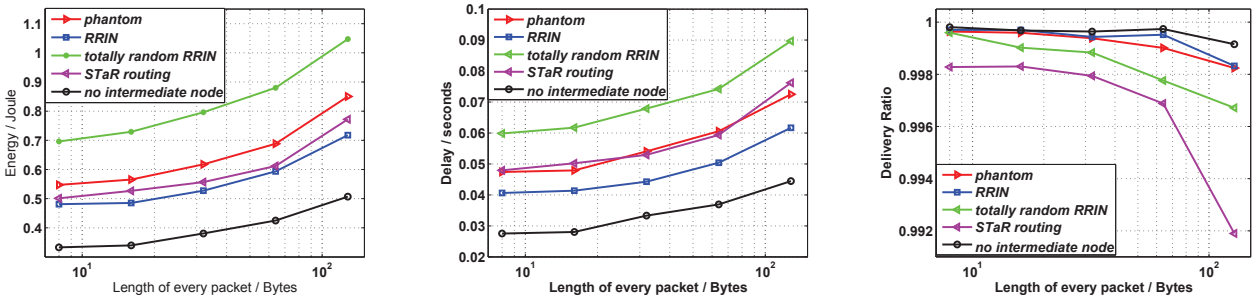
The security of the proposed STaR routing scheme can be analyzed based on the location of the adversarial attacks: i) the adversary monitors traffic between the source node and the randomly selected intermediate node, and ii) the adversary monitors traffic between the randomly selected intermediate node in STaR and the SINK node. For case i), the message source may be located anywhere 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. It is virtually impossible for an adversary intercept multiple messages from the same source, as shown in Fig. 3. Therefore, STaR provides local source location privacy.

For case ii), the STaR area is at least a minimum radius distance, r, from the SINK node and traffic will be transmitted to the SINK node from all possible directions with equal probability. Therefore, it is quite impossible for an adversary to predict the direction of the source node. 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 way, global source location privacy can be assured.

VI. Performance Analysis and 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. 4. In the simulation, 400 nodes are randomly distributed in a square target area of size 3360×3360 meters, while 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 meters away from the real source. For RRIN scheme, the minimum distance between the source node and the intermediate nodes was set to 480 meters, and the average distance turned out to be 529.14 meters. We also illustrate the performance of the totally randomly selected intermediate nodes. For STaR routing, the inner radius, r, was set to 480 meters, while the outer radius, R, was set to 640 meters.

978-1-4244-5638-3/10/$26.00 ©2010 IEEE



(a) Power consumption (b) Message latency (c) Message delivery ratio

Fig. 4. Performance of routing by single-intermediate node

Through analysis and simulation results, we find that direct routing without intermediate node has the best performance while totally random RRIN has the worse performance. The performance of the RRIN scheme is better than phantom routing for comparable security since the average routing paths in phantom routing is longer than the RRIN due to the more curved routing paths. The performance of STaR is between the totally random RRIN and constrained RRIN. The delivery ratio for STaR is slightly lower than the two RRIN schemes due to the possible higher collisions ratio.

VII. Conclusions

Source-location privacy is vital to the successful deployment of wireless sensor networks. In this paper, we introduced a STaR routing scheme for local and global source-location privacy protection. We carried out theoretical analysis to evaluate the security and the performance of the proposed scheme and compared it with other existing schemes. Our simulation results demonstrate that the proposed STaR routing scheme can achieve excellent performance in energy consumption and delivery latency. Message delivery ratio is slightly lower than the other schemes but it is still satisfying overall.

Acknowledgements

This research was supported in part by the NSF under grants CNS-0845812 and CNS-0848569.

References

[1] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, February 1981.

[2] D. Chaum, "The dinning cryptographer problem: Unconditional sender and recipient untraceability," *Journal of Cryptology*, vol. 1, no. 1, pp. 65–75, 1988.

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

[4] 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.

[5] 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.

[6] R. D. G. Danezis and N. Mathewson, "Mixminion: design of a type iii anonymous remailer protocol," *IEEE Symposium on Security and Privacy*, pp. 2–15, 2003.

[7] J. Deng, R. Han, and S. Mishra, "Countermeasures against traffic analysis attacks in wireless sensor networks," *SecureComm 2005*, pp. 113–126, Sept. 2005.

[8] Y. Jian, S. Chen, Z. Zhang, and L. Zhang, "A novel scheme for protecting receiver's location privacy in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 7, pp. 3769–3779, October 2008.

[9] Y. Yang, M. Shao, S. Zhu, B. Urgaonkar, and G. Cao, "Towards event source unobservability with minimum network traffic in sensor networks," in *WiSec '08: Proceedings of the first ACM conference on Wireless network security*, (New York, NY, USA), pp. 77–88, ACM, 2008.

[10] M. Shao, Y. Yang, S. Zhu, and G. Cao, "Towards statistically strong source anonymity for sensor networks," *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pp. 51–55, April 2008.

[11] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing source-location privacy in sensor network routing," *Distributed Computing Systems, 2005. ICDCS 2005. Proceedings. 25th IEEE International Conference on*, pp. 599–608, June 2005.

[12] Y. Xi, L. Schwiebert, and W. Shi, "Preserving source location privacy in monitoring-based wireless sensor networks.," in *IPDPS*, IEEE, 2006.

[13] Y. Li, L. Lightfoot, and J. Ren, "Routing-based source-location privacy protection in wireless sensor networks," in *IEEE EIT 2009*, (Windsor, Ontario, Canada), June 7-9 2009.

[14] http://www.panda.org/.

[15] M. Ye, C. Li, G. Chen, and J. Wu, "Eecs: an energy efficient clustering scheme in wireless sensor networks," *IPCCC 2005*, pp. 535–540, April 2005.

[16] W. B. Heinzelman, *Application-specific protocol architectures for wireless networks*. PhD thesis, 2000. Supervisor-Anantha P. Chandrakasan and Supervisor-Hari Balakrishnan.

[17] J. Neander, E. Hansen, M. Nolin, and M. Bjorkman, "Asymmetric multihop communication in large sensor networks," *Wireless Pervasive Computing, 2006 1st International Symposium on*, pp. 7 pp.–, Jan. 2006.

[18] O. Younis and S. Fahmy, "Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 3, pp. 366–379, Oct.-Dec. 2004.

[19] Y. Zhang, W. Liu, Y. Fang, and D. Wu, "Secure localization and authentication in ultra-wideband sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 24, pp. 829–835, April 2006.

[20] P. Zhang and M. Martonosi, "Locale: Collaborative localization estimation for sparse mobile sensor networks," pp. 195 –206, april 2008.

[21] T. Srinath, "Localization in resource constrained sensor networks using a mobile beacon with in-ranging," pp. 5 pp. –5, 0-0 2006.

[22] D. min Chen and Y. Zhang, "Research of wsn localization algorithm based on entropy function," vol. 1, pp. 229 –233, march 2009.

[23] H. Chan and A. Perrig, "Pike: peer intermediaries for key establishment in sensor networks," *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 1, pp. 524–535 vol. 1, March 2005.

[24] 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.

[25] P. Traynor, R. Kumar, H. Choi, G. Cao, S. Zhu, and T. La Porta, "Efficient hybrid security mechanisms for heterogeneous sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 6, pp. 663–677, June 2007.