Preserving Source-Location Privacy in Wireless Sensor Networks against a Global Eavesdropper

Zhiwen Zeng¹, Xiaoyan Hu¹, Zhigang Chen¹
¹School of Information Science and Engineering
Central South University
Changsha 410083, China

Hui Liu²*
²Department of Computer Science
Missouri State University
Springfield, MO 65897, United States

Abstract—While many works to date in wireless sensor networks (WSNs) security have focused on providing confidentiality for message contents, contextual information usually remains exposed. Thus the adversary especially the global eavesdropper can easily obtain the sensitive information such as the location of a target object in a monitoring application, which is critical to the mission of the sensor network. In this paper, we propose an energy-efficient scheme against global eavesdroppers for protecting location privacy and maximizing lifetime of WSNs. Our proposed technique builds fake source areas controlled by the Sink, which disseminates dummy traffic synchronized with real source area. Therefore, we create multiple candidate traces in the network to hide the real traffic generated by the real source. To ensure no impact on lifetime of WSNs, we minimize the energy consumption of hotspots. Analysis and simulation results show that the proposed scheme can significantly improve the security in source-location privacy preservation without reducing the network lifetime.

Keywords—source-location privacy; wireless sensor network; global eavesdroppers; fake source area;

I. INTRODUCTION

A wireless sensor network is composed of numerous cheap, small, energy-constrained, and spatially distributed autonomous sensors to monitor and study the physical world. It has gained more popularity in recent years and been widely used in lots of important applications such as environment monitoring, military surveillance, and target tracking. Sensors collaborate to gather data and disseminate the data to the sink. Since most of sensors are deployed in unattended or hostile environment, these applications are subject to a variety of security issues.

Among all of these security threats, source location privacy is of special interest to us since it cannot fully addressed by traditional security mechanisms, such as encryption and authentication. Consider a simple example of target tracking in WSNs. A sensor sends a message which includes event-related information to the sink when it detects an event. After this, the location of the event source has been exposed to the adversary who might passively monitor the network traffic, no matter how strong the data encryption key is. Further, the adversary may find out more sensitive information: whether, when and where a particular event occurred, for example, the appearing of a soldier in a target tracking sensor network [1, 2]. This can help the adversary in attacking the soldier, an unfortunate occurrence.

In the past two decades, a number of researches aiming at protecting source location privacy have been proposed. We can generally divide these works into two categories according to the capability of the adversaries: strategies against local eavesdroppers and those against global eavesdroppers. The local eavesdroppers have limited coverage, comparable to that of regular sensors. At any given time only a local area is under the adversary’s monitoring and the adversary tries to locate the source node hop-by-hop in a tracing back way. However, global eavesdroppers are much more powerful and formidable. They are able to monitor all the network traffic either by deploying their own cheap sensors that cover the whole area [3] or by employing a powerful site surveillance device with hearing range no less than the network radius.

Several methods have been proposed to preserving the source-location privacy against local eavesdroppers. For example, [4-7]. Kamat et al. [4] propose a classic location protecting protocol based on Phantom Routing. Firstly every message is randomly routed for $h$ hops to find a phantom source, and then the selected phantom source sends the message to the SINK node by flooding. However, both theoretical and practical results demonstrate that if the message is routed randomly for $h$ hops, the message will be largely within $h/5$ hops away from the actual source [5]. Xi et al. [6] propose a two-way greedy random walk named as GROW. In GROW, greedy random walk first creates a static random walk (path of receptors) from the sink node. Subsequently, messages are sent from the source node on a greedy random walk that will eventually arrive at a receptor node, from which the message will be forwarded to the SINK node following the established path. In order to well balance the energy consumption and privacy protection, J. Ren et al. [5] propose a two-phase routing scheme. The source node randomly determines an intermediate node from a pre-determined region around the SINK node called 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. In StaR, the entire network is divided into grids. One node in every grid is denoted as the head node. The source node randomly selects one grid in the constrained area, of which the head node becomes the random intermediate node. In [7], the authors propose a three-phase routing scheme that addresses the source-location privacy issue by using Radial Routing and Annular Routing based on the phantom source.

* The corresponding author: huiliu@missouristate.edu.
The strategies mentioned above all have outstanding performances in location privacy preservation against local eavesdroppers. However, in order to effectively defend powerful global eavesdroppers, we have to design more suitable schemes to preserve location privacy against them.

Mehta et al. [2, 8] first presented the global eavesdropper model, and proposed two techniques, called periodic collection and source simulation to prevent the leakage of location privacy. The main idea of the periodic collection method is to make the traffic pattern independent of the presence of real objects. However, this approach consumes a substantial amount of energy for latency sensitive applications and largely reduces the lifetime of network since each node periodically sends packets at a reasonable rate regardless of whether it has real data to send or not. The source simulation method creates multiple candidate traces in the network to hide the traffic generated by real objects by simulating the movement patterns of real objects. However, this approach has two problems. First, it is challenging to model the movement patterns of real objects. Second, fake sources also bring much extra energy consumption to the hotspots of the network, which reduces the network lifetime.

The scheme based on proxy filtering is illustrated by Yang et al. [9], in which they select some sensors as proxies that proactively filter dummy packets on their way to the sink. Then, they proposed two mechanisms named as PFS (Proxy-based Filtering Scheme) and TFS (Tree-based Filtering Scheme). Because determining proxies (i.e. selection P elements out of V nodes) is an NP-hard problem as proved in [9], they adopted local search heuristics with no guaranteed maximal network lifetime.

Bicakci et al. [10, 11] investigate the network lifetime in various different proxy assignment strategies and different deployment scenarios. They propose a new filtering idea called OFS (Optimal Filtering Scheme) to maximize the network lifetime of wireless sensor networks while preserving event-unobservability against global eavesdroppers. Through a Linear Programming framework, they claimed that Linear Programming is an effective method to find the optimal locations of proxies under a set of linear constraints.

Recently, Ju Ren [12] adopts cluster structures to construct cyclic interference routing paths, in which the cluster heads will act as proxies to filter fake message generated by the fake source. They allow only real traffic flow to cross the hotspots and build cyclic diversionary routing paths in areas where the sensors have enough abundant energy to support them to maximize the network lifetime.

All of the above-mentioned researches do not solve the problem that the change of the real source location is ahead of that of the fake source location, therefore, the adversary easily detect the location of real source because of the abnormal operations of nodes. In this paper, we use the sink, the center of gathering data in the network, to guarantee the synchronization of data dissemination of real source and fake source, then ensure the preservation of location privacy. Our proposed scheme can achieve a significant improvement in network security without reducing the network lifetime since we take advantage of residual energy in non-hotspots. At the same time, we also consider to minimize the data transmission latency as one of our optimal goals.

The remainder of this paper is organized as follows. In Section II, the system model is described. Details of the proposed location privacy scheme against global eavesdroppers are illustrated in Section III. Section IV compares and analyzes the performances of our scheme based on the simulation studies. Section VI concludes the paper.

II. MODELS

A. Network Model

In this paper, we adopt a homogeneous network model, in which all of the sensors have roughly the same capabilities, power sources, and expected lifetimes. This is a common network architecture for many applications today not only because it is very simple for deployment and maintenance but also it has been well-studied and provides for relatively straightforward analysis. We make the following assumptions about our network model.

- A wireless sensor network is deployed with equal density throughout a circular region with the radius R. The whole network is fully connected through multi-hop communications [7]. The only SINK node is located at the center of the circular network that is the destination location that data packets will be routed to.
- The appearance of the object is randomly distributed in the entire network, so the probability that each sensor detects the information of the object as well as sends data to the SINK is equivalent. We suppose that the time that the message delivered by the farthest sensors in the entire network arrives the SINK is denoted as T, and the sensing radius of the sensor is r.
- We assume that a security infrastructure has already built in; that is, no information carried in the message will be disclosed. The key management, including key generation, key distribution, and key update, is beyond the scope of this paper.

B. Adversary Model

In this paper, we adopt the adversary model given in [12]. Adversaries are assumed to be external, passive and global. More precisely stated, adversaries cannot compromise or control any sensors. They do not conduct any active attacks such as traffic injection, channel jamming, or denial of service attack. However, adversaries can listen to all communication in the network, analyze the collect data and try to determine the location of each sensor node. We assume adversaries deploy their own sensors that cover the target WSNs in order to achieve eavesdropping all the communications of the whole target WSNs.

C. Energy Consumption Model

Energy consumption model [13] is adopted in this paper. We consider only the energy usage of transmitting and receiving messages. Energy consumption for transmitting
messages is shown in equation 1, and then equation 2 shows the energy spent for receiving a l-bit packet.

\[
\begin{align*}
E_{\text{member}} &= lE_{\text{elec}} + l\frac{d^2}{d_0} \quad d < d_0 \quad (1) \\
E_{\text{member}} &= lE_{\text{elec}} + l\frac{d^4}{d_0^4} \quad \text{if} \quad d > d_0
\end{align*}
\]

\[E_{R}(l) = lE_{\text{elec}} \quad (2)\]

where \(E_{\text{elec}}\) is transmitting circuit loss. When the distance \(d\) between transmitter and receiver is less than the threshold \(d_0\), the free space (\(d^2\) power loss) channel model is considered. Otherwise, the multi-path fading (\(d^4\) power loss) channel model is adopted. \(\frac{d^2}{d_0}\) and \(\frac{d^4}{d_0^4}\) are the energy required by power amplification in these two models, respectively. The above parameter settings are given in Table 1 [13].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold distance ((d_0)) (m)</td>
<td>87</td>
</tr>
<tr>
<td>Sensing range (r_s) (m)</td>
<td>15</td>
</tr>
<tr>
<td>(E_{\text{elec}}) (nJ/bit)</td>
<td>50</td>
</tr>
<tr>
<td>(\frac{d^2}{d_0}) (pJ/bit/m²)</td>
<td>10</td>
</tr>
<tr>
<td>(\frac{d^4}{d_0^4}) (pJ/bit/m⁴)</td>
<td>0.0013</td>
</tr>
<tr>
<td>Initial Energy (J)</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\(b = \log_2 \frac{|S_r|}{|S_f|}\) is used to measure the level of privacy preservation in [2]. Depending on the users and applications, this can be easily modified to support different kinds of privacy measurement models. For example, we can define high, medium and low privacy levels using appropriate values of \(b\). \(S_r\) is defined as a set of sensors that represent the set of possible locations for the objects sensed by the target network. Let \(S_f\) be the number of real sources which the adversary finds, then the adversary knows the location of target object. For example, the periodic collection method [2] has every sensor node independently and periodically send packets at a reasonable frequency regardless of whether there is real data to send or not, thus \(S_f\) is a set of all the sensors in the network while \(S_r\) is the set of potential real sources determined by the adversary. The source simulation method [2] randomly selects a set of sensors and pre-load each of them with a different token. These tokens will be passed around between sensors to simulate the behavior of real objects, thus \(S_f\) is a set of sensors which simulate the real objects while \(S_r\) is the number of real objects which is 1 in most scenarios.

### III. The Routing Scheme Based on Fake Source Area (FSA)

In this section, we illustrate our proposed routing scheme based on Fake Source Area (FSA) for location privacy preservation and lifetime maximization in wireless sensor networks. The principles of FSA can be summarized as the following three aspects. Firstly, FSA builds a number of fake source areas to generate interference data gathering areas by using the abundant energy in areas far from the sink, because the sensors in these areas always remain much energy when the network dies. Secondly, since only the center in every fake source area and real source area transmits data to the SINK, the number of transmitted packets as well as the energy consumption in hotspots is reduced, therefore our method increases the network lifetime. Finally, the SINK has knowledge of the movement pattern of real source sensors through data collections, it broadcasts the change to all the sensors and asks to change the locations of centers. Thus, from the view of the adversary, all the centers change at the same time which enhances the difficulty to detect the real source location. The main idea of FSA can be detailed as the following five phases.

#### A. Deployment and initialization of network

Before deployment, we randomly select a set of sensor nodes and pre-load each of them with a different token. Every token has a unique ID. For convenience, we call the node holding a token the token node. Token nodes are considered as fake sources and centers of fake source areas. The change of token nodes means the change of fake source areas. The random deployment of token nodes can guarantee that when the real object appears anywhere in the network, it will be treated as same as the random fake sources. The forming of real source area and fake source areas are described as the following steps. 1) Each token node and real source act as the centers by broadcasting one data packet, which includes the location of centers and the number of hops, to all the neighboring nodes. For example, if we define the size of area is three hops within the centers, the number of hops will be set as 3 initially. Therefore, the number of hops starts from 3, it decreases by 1 once the data packet is forwarded one more hop. When the number of hops reaches zero, the data packet will be discarded. Only sensors within the area can receive this data packet. 2) When the neighboring nodes receive the broadcasted data packets, they know they are members of this area centered at the corresponding token node. Definitely it is possible that one sensor receives multiple data packets sent by different token nodes, the sensor will select the closest token node among them as its center. So far, fake source areas and real source area are established.

#### B. Data gathering within areas

Only real source transmits the real packet while all the other nodes send fake packets. As shown in Fig. 1, the data transmission is initiated by the most outside nodes, and continues to the centers within all the areas. Every sensor node checks whether the packet includes real data or dummy traffic when it receives one packet. If the packet is just dummy traffic, the sensor node directly discards the dummy traffic and generate
a new fake one to send. Otherwise, the sensor node forwards real data to its neighbors until data packet finally arrives the center of its belonged area.

---

C. Data transmission from centers to the SINK

After the centers gather packets from their areas, the centers determine whether they obtain real packet. If the center has the real packet, it will transmit the real packet to the SINK along the shortest path between itself and the SINK. Otherwise, it discards gathered dummy packet, generates one new fake packet, and forwards the fake packet to the SINK also along the shortest path between itself and the SINK. Since the size of every area is fixed and every center sends packets to the SINK almost at the same time after they have gathered data from their area, all the packets sent by the centers arrive the SINK simultaneously. The adversary cannot tell the difference between real source area and random fake source areas because they are synchronized with each other in the aspect of data transmission. Eventually the SINK receives the real data packets which include data related to the tracking target such panda and location information of real source. Then, the SINK will decide whether the fake source areas need to move in next round by analyzing the change information of real source location.

D. The changes of fake source areas

After the first round of data collection of the SINK as shown in Fig. 2, the SINK obtains the real source location. Our method continues the second round of data collection without making any changes of fake source areas. However, after the second round of data collection, the SINK conducts two possible actions according to the changes of the real source location. First, the real source location remains the same as before, which means the target object does not moves at all or maybe only moves within the monitoring area of the previous real source sensor. FSA continues next round of data collection without changes of fake source areas. Second, the real source location changes, which means the target object already moves out of the monitoring area of the previous real source node. So we can predict the real source area definitely changes in next round of data collection. But we can conclude that the next real source should be in one hop area away from the current real source under our assumption that the speed of target object is not very fast and the target object should still be in the range of one hop to two hops distance from the current real source. In a summary, the current fake source randomly select one node within one hop area from itself as the new fake source as the new center of fake source area in order to synchronize the movement patterns of fake source areas and real source area. The token will then be passed to the new selected fake source from the previous fake source. The delivery of such taken between sensor nodes will be always protected by the pairwise key established between them.

E. New fake source areas

After the selection of centers of new fake source areas, FSA continues the next round of data collection according to the above phases A, B, C, D, and E.

IV. SIMULATION AND PERFORMANCE ANALYSIS

We use the discrete event-based simulator-OMNET++ [14] to simulate the two protocols, FSA and Source Simulation, comparing their energy consumption and network lifetime in order to achieve a certain level of location privacy $b$.

A. Simulation Environment

Assume there are 4000 sensor nodes uniformly distributed in a circular area with the radius of 600 meters, $R$. Every node can communicate with those nodes no more than 50 meters far from it given the assumption that the sensing radius is as the same as the transmission radius $r = 50$ meters. The node density of the network is 0.0035 sensors per square meters. We also suppose the size of source areas is 2 hops and the number of fake source areas is 10. Only one target object appears in the network. The initial location of the target object is random. If the target object moves from coordinates $(x,y)$
to \((x \pm \Delta x, y \pm \Delta y)\), and \(\Delta x^2 + \Delta y^2 \leq r^2\). That means within the interval of two rounds of data collection, the moving distance of the target object is less than that of one hop.

B. Performance Analysis

It is known that the better the level of privacy preservation, the more the number of fake source areas or the larger the radius of fake source areas. Our experiments study the communication cost and network lifetime in order to achieve a certain level of location privacy preservation.

![Fig. 3](image1.png)

Fig. 3 the comparison of energy consumptions corresponding to different level of location privacy preservation.

From Fig. 3 we observe that the energy consumption of source simulation method [2] is 4 or 5 times as high as our proposed scheme FSA in order to achieve the same level of privacy preservation. Because most nodes only transmit data in short distance in FSA while fake sources in [2] need to send data all the way to the SINK which involve more intermediate nodes and consume more energy compared with FSA.

![Fig. 4](image2.png)

Fig. 4 the comparison of network lifetime corresponding to different level of location privacy preservation.

FSA has two ways to enhance the location privacy preservation, one is to increase the number of fake source areas, and the other is to increase the size of fake source areas. As we know that the network lifetime depends on the energy consumption of the hotspot and has no direct relationship to the whole network energy consumption. Therefore, increasing the size of the fake source areas do not reduce the network lifetime while the increasing the number of the fake source areas directly reduce the network lifetime. In order to prove the advantages of FSA, we adopt to increase the number of fake source areas to enhance the privacy preservation in our simulation experiments. As shown in Fig 4, the network lifetime of FSA is almost 80 to 90 times as long as the network lifetime in source simulation method since FSA fake sources gather packets within areas before sending packets to the SINK, which largely decreases the number of packets sent to the SINK, reduces the number of data transmission in hotspots, and then improves the network lifetime.

V. Conclusion

Source-location privacy is significantly important to the successful deployment of wireless sensor networks. In this paper, we propose a novel routing scheme based on the fake source areas for protecting source-location privacy against global eavesdroppers named as FSA. We carry out extensive simulation experiments to compare our proposed method with other existing schemes. Our simulation results demonstrate that the proposed FSA can effectively enhance the privacy preservation as well as maximize the network lifetime.

References


