CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. (2013)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.3075

SPECIAL ISSUE PAPER

Protecting source–location privacy based on multirings in wireless sensor networks

Lin Yao [1], Lin Kang [1], Fangyu Deng [1], Jing Deng [2] and Guowei Wu [1,*,†]

[1] *School of Software, Dalian University of Technology, Dalian 116620, China*
[2] *Department of Computer Science, University of North Carolina at Greensboro, Greensboro, NC 27412, U.S.A.*

SUMMARY

Wireless sensor networks (WSNs) are expected to be widely deployed to collect data in military and civilian applications. Because of the open nature of WSNs, it is easy for an adversary to eavesdrop sensor communication and to trace packets, causing privacy concern for the sensor devices. The privacy issue, especially location privacy, can be critical for monitoring applications in WSNs. A unique case of location privacy is that of the sources, which are vulnerable of being captured and target attacks. In this paper, we propose a scheme to protect the source–location privacy based on a novel use of multiring topology. To achieve a uniformly distributed traffic pattern throughout the network, the source node selects two random rings each from its external rings and internal rings and a set of two random angles with a sum of 180 degrees for each packet. The packet is sent at one of the angles in each ring. Fake packets are also injected to provide path diversity and to increase attack time, which is defined as the time that the adversary takes to locate the source successfully. These techniques protect the source node from packet tracing attacks as well as traffic analysis attacks. Our analysis and simulations, performed in the NS2 simulator and MATLAB, demonstrate that our proposed scheme can provide better spatial traffic evenness and longer attack time, along with a modest increase of hop count and energy consumption.

Received 14 January 2012; Revised 28 November 2012; Accepted 23 May 2013

KEY WORDS: wireless sensor networks (WSNs); location privacy; multiring; attack time

1. INTRODUCTION

A wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion, or pollutants [1, 2]. The monitoring purpose of WSNs naturally leaves them operating in unattended or sometimes even hostile environments. As such, they are exposed to a variety of attacks such as communication eavesdropping, node compromising, and signal jamming.

The privacy issue is critical for monitoring applications in WSNs [2, 3]. Privacy in WSNs can be classified into data privacy and context privacy [4]. Data privacy concerns the privacy of data collected as well as queries posted in a WSN. Context privacy deals with sensor locations including the source and the destination. Among them, source–location privacy concerns the privacy of the locations from which the sensing data is generated and collected; destination location privacy is about the location where the collected data is delivered. Even though traditional security mechanisms such as encryption and authentication can protect data privacy in WSNs, sensor locations are hard to hide. For instance, in the panda hunter game [5], sensors are deployed to monitor pandas in a region and to send the monitoring information to the base station by multihop forwarding. Although

*Correspondence to: Guowei Wu, School of Software, Dalian University of Technology, Dalian 116620, China.
†E-mail: wgwdut@dlut.edu.cn

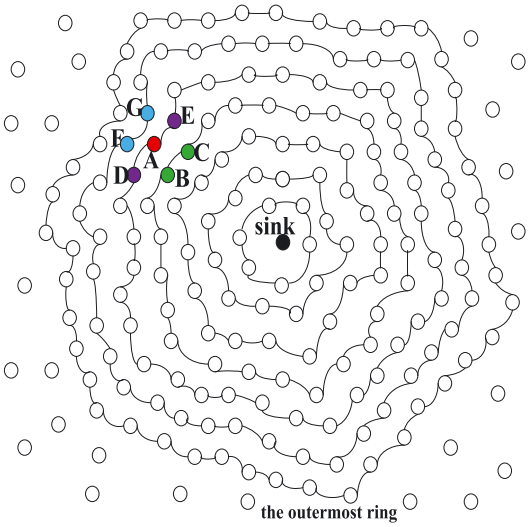

Figure 1. Multiring topology. The rings are numbered as 1 to L with ring 1 representing the innermost ring (closest to the data sink). Sensors on the same ring have the same hop count from the data sink. Each node maintains three lists: far-neighbor list, equidistant-neighbor list, and near-neighbor list. Each node also maintains a set of external rings and internal rings. For node A, nodes D, and E are equidistant-neighbors because they have the same hop count to the data sink as node A does. Nodes F and G are the far-neighbors because they are one-hop farther away from the data sink. Similarly, nodes B and C are near-neighbors.

the adversary cannot decrypt the encrypted messages sent among the sensors, it can easily track the traffic flow by traffic analysis and packet tracing. With these tools, the adversary can even locate the source sensor and find the panda through hop-by-hop back-tracing. The example highlights the importance of protecting source–location privacy in WSNs.

In general, protecting location privacy in WSNs is hindered by the sheer scale of the networks, the extreme resource constraints, and the powerful tools used by the attackers. Two of such attacking tools are traffic analysis and packet tracing [6, 7]. Traffic analysis takes advantage of the fact that sensors near the receiver forward a greater volume of packets than sensors that are further away. Thus, traffic analysis can be used to estimate or even pinpoint the receiver location by analyzing the network traffic. In packet tracing attack, a proactive attacker performs hop-by-hop backward traffic analysis to trace the data source [6].

To protect the source–location privacy, we propose a scheme based on a unique use of multiring topology. In our approach, multiple rings are formed after deployment and network initialization, with the data sink (receiver) located at the center of the rings (as shown in Figure 1). The sensors on the same ring are chosen such that they have the same hop count to the sink. The goal of our design is to generate seemingly uniformly distributed traffic. Our scheme takes advantage of the multiple layers of rings formed around the data sink. Instead of routing the data packet directly toward the sink as in the shortest-path scheme [5], we route it toward a ring that is outside the source's ring. Then, it will be routed on the ring for a randomized angle before it is routed toward a ring that is inside the source's ring. The data packet is then routed on the inner ring for another angle before it is finally routed toward the data sink. Our analysis and simulations, performed in the NS2 simulator and MATLAB, demonstrate that our proposed scheme can provide better spatial traffic evenness and longer attack time, defined as the time that it takes for the adversary to trace back to a source node, along with a modest increase of hop count and energy consumption.

The remainder of this paper is organized as follows. In Section 2, we discuss related works. Section 3 describes the system model and the details of our proposed scheme. Analysis is given in Section 4. Section 5 evaluates the performance of our scheme through simulations. Finally, we conclude our work in Section 6.

DOI: 10.1002/cpe

2. RELATED WORKS

Data source and base station are two crucial factors in the protection of location privacy in WSNs [4]. Our discussions of related work focus on surveying the current techniques concerning the privacy protection of data source. Existing techniques of preserving the source–location privacy can be categorized into four typical classes: flooding, random walk, dummy injection, and fake sources [4, 8].

2.1. Flooding

Information flooding is used to disguise the real data traffic, so that it is difficult for an adversary to track the source of data by analyzing network traffic. Flooding includes baseline flooding, probabilistic flooding, and phantom flooding [5].

In baseline flooding [5, 9], the message originator transmits its message to each of its neighbors, who in turn retransmits the message to each of their neighbors. Each node forwards the same message only once. The premise of this approach is that all sensors participate in the data transmission so that it is unlikely for an adversary to track a path of transmission back to the data source. However, the effectiveness of baseline flooding on privacy protection critically depends on the number of nodes in the transmission path between the data source and the base station. If the path is too short, the adversary can easily infer that the routing path of this packet is the shortest between the data source and the base station after it detects the arrival of the first packet at the base station. Then, the adversary can trace back from the last forwarding sensor along the routing path to the data source. Hence, it is easy for the source to be located. Furthermore, there are also energy consumption concerns.

Probabilistic flooding [5, 9] is an optimization of the baseline flooding technique to reduce energy consumption. In probabilistic flooding, a subset of nodes is chosen randomly, and only the subset forwards data. Each node forwards or broadcasts a packet with a predetermined probability after it receives packets. It is obvious that this scheme can not only significantly reduce energy consumption but also effectively prevent the adversary from tracing back to the data source. Nonetheless, the reception of data by the base station is not guaranteed owing to the randomness involved in this approach.

Phantom flooding [9] attempts to direct messages to different locations of the network so that the adversary cannot observe a steady stream of messages for its tracing purpose, which shares the same insights as probabilistic flooding. However, probabilistic flooding is not very effective to achieve this goal because shorter paths are more likely to deliver more messages. Therefore, a phantom source is used to draw the attacker away from the real source. In phantom flooding, when the real source needs to send out a message, it unicasts the message to a random node, which will send it out using baseline flooding.

The main problem of flooding is that it could still reveal source–locations quickly [10]. Moreover, flooding incurs a high energy consumption.

2.2. Random walk

The purpose of random walk approach is to guide the message through some random paths before sending toward the data sink.

As one of random walk approaches, the phantom source single path technique (PSSP) was proposed [5]. In PSSP, the real source unicasts the message to a random node, which will unicast the message to the data sink. Both unicasts are sent over the shortest-paths. However, the pure random walk approach is not statistically secure for protecting the location of the data source [10]. In particular, it can be shown that a pure random walk tends to stay around the real source.

Aiming at improving phantom technique, Yong *et al.* [10] proposed the greedy random walk technique, where both the source and sink perform random walks. Once both walks intersect, the message is routed alongside the opposite walk to the destination. At that point, the packet is forwarded to the base station through the path pre-established by base station.

DOI: 10.1002/cpe

Wang *et al.* [11] formulated the location privacy problem as an optimization problem on the average trace back time and minimal trace back time for the adversary to reach the message source starting from the sink. The weighted random stride scheme was proposed, allowing each individual sensor to make the routing decision locally and independently, with little deployment information. Nodes pick larger forwarding angles with a higher probability; therefore, more messages will be distributed to longer paths, deterring back-tracing by the adversary.

To ensure that phantom sources are far away from the real source, Kamat *et al.* [5] proposed directed walk. In the directed walk, the direction information is stored in the packet header and forwarding sensors send the packet along the direction. Yun *et al.* [12–16] addressed the problem of easy source-tracking with a random intermediate node (RRIN) scheme. In RRIN, the source node sends its packet to a randomly selected intermediate node with at least a minimum hop distance from itself. RRIN was further improved by protecting global (network-level) source–location privacy. This is achieved by mixing the packet with other packets through a network mixing ring [13, 14].

In particular, the random selected intermediate node with single path scheme (RSINSP) was proposed in [12]. Similar to RRIN, in RSINSP, the message source first randomly selects an intermediate node in the sensor domain. Then, the data packet is sent to the intermediate node before it is routed to the sink along the shortest-path. From this intermediate node, the message will be forwarded to the destination node. This scheme could provide local source–location privacy or global location privacy, depending on how the intermediate nodes are selected.

To reduce energy consumption, Yun *et al.* presented two routing schemes using multiple randomly selected intermediate nodes on the basis of angle and quadrant [15, 16]. In these schemes, packets are forwarded to the sink through multiple intermediate nodes based on a randomly distributed angle θ in the range of $(\beta, -\beta)$, where β is randomly selected from 0 and 180 degree. The selection of intermediate nodes is totally random, that is, every sensor node has the same chance of being selected as the intermediate node for any source node.

As will be demonstrated in Section 5, schemes using the random walk approach only provide limited protections for source–location privacy.

2.3. Dummy injection

Dummy packet injection schemes inject dummy packets to the network to defend from traffic analysis and packet tracing attack [17, 18]. In short-lived fake source routing [5], each sensor produces a fake packet and floods it to the network with a predetermined probability. Although easy to implement and being able to perturb local traffic observable to the adversary, the scheme becomes ineffective when it meets a global eavesdropper within a complete view of the network traffic. Two methods were introduced to address this problem: periodic collection and source simulation [19]. In the periodic collection method, every sensor node independently and periodically sends packets at a reasonable frequency regardless of its real data generation. In the source simulation method, every sensor node is a potential source node.

To resist a global eavesdropper and reduce energy consumption, Yang *et al.* [20] proposed statistically strong source anonymity. In the FitProbRate scheme, the exponential distribution is used to control the rate of dummy traffic generation. Yang *et al.* [21] also introduced the notion of event source unobservability, aiming to hide the real event source in combination with mechanisms of dropping dummy messages. It can prevent the explosion of network traffic. They proposed two schemes, proxy-based filtering scheme (PFS) and tree-based filtering scheme (TFS). In PFS, some sensors are selected as proxies to collect and filter dummy messages from surrounding sensors. PFS greatly reduces system communication cost by dropping many dummy messages before they reach the base station. In TFS, proxies are organized into a tree hierarchy. Proxies closer to the base station filter traffic from proxies farther away, thus the message overhead could be further reduced.

Similar to flooding, dummy injection also has the drawback of introducing a large amount of message overload.

DOI: 10.1002/cpe

2.4. Fake source

The basic idea of fake data source is to choose one or more sensor node to simulate the behavior of a real data source to confuse the adversary. Fake packet generation was first introduced in [5]. The base station will create fake sources whenever it receives a signal indicating that a sensor wants to send data. These fake senders are far away from the real source and approximately at the same distance from the base station as the real sender. Both real and fake senders start generating packets at the same instance and with the same traffic pattern.

A significant challenge for the design of this technique is how to simulate the behavior of data sources without being detected, which is an open problem. Furthermore, such a technique will also incur more power consumption in WSN.

Summary: Although some of these location privacy techniques provide certain levels of protections for data sources in WSNs, the overall traffic pattern throughout the network is still tilted toward the sources, making it easier for the adversary to trace back. In this work, we propose a scheme to take advantage of the multiring approach and fake packet injection (FPI). The result is a more evenly distributed traffic in the network, making it more difficult for the adversary to trace back the sources.

3. THE PROPOSED SOURCE–LOCATION PRIVACY SCHEME

In this section, we first discuss our system model and adversary model, then we present the details of our proposed source–location privacy scheme. The frequently used notations are listed on Table I.

3.1. System model and adversary model

As discussed in Section 1, the hunters will trace and capture the pandas as soon as the source node is identified. Our goal is to make it difficult for the adversary to determine the source–location by either traffic analysis or packet tracing attack. In our system, sensor nodes are deployed as illustrated in Figure 1. We make the following assumptions regarding to our system model: The sink is the only gateway that collects data from the entire network. Multiple sinks are out of the scope of this work. Two-dimensional coordinates are adopted in our model. The coordinate origin is the sink. There exist multiple rings loosely centered at the sink. The sensors with the same hop count from the sink are distributed on the same ring. Message confidentiality and route selection are out of the scope of this work.

To locate a source node, the adversary is expected to possess sufficient computation and storage capability. We model the adversary as follows:

(1) The adversary only carries out passive attacks without interfering with the proper network functions, afraid of being caught.
(2) The adversary is capable of back-tracing to the original data source. It can record the entire trace back process and node IDs. In general, the traceback rate is proportional to packet transmission rate.
(3) The adversary knows the traffic pattern of the entire network.

Table I. Notations.

Symbol	Description
a	ath ring is where the data packet is generated
b	bth ring is the inner ring, where $b < a$
c	cth ring is the outer ring, where $a < c$
α	Random angle chosen between 0 and π. This is the angle of packet travel on the outer ring
β	$\beta = \pi - \alpha$. This is the angle of packet travel on the inner ring
p	The probability that a node generates fake packets
h	The hop count from the source to the data sink
TTL	Time-to-live (in hops) of an injected fake packet
L	The hop count of the outermost ring
N	The total number of nodes in the network

DOI: 10.1002/cpe

(4) The adversary can overhear a signal to estimate the sender location. Source ID of the packet will be disclosed to the adversary. The detection radius is assumed to be the same as the sensors' wireless transmission range.
(5) After arriving at a location, the adversary observes traffic for a period before deciding on where to trace back next. The duration of the observation time is called 'pause time'.

3.2. The proposed scheme

Our scheme has three steps: initialization, path diversification, and FPI. In the initialization step, rings are formed and each node chooses its far-neighbor list, equidistant-neighbor list, and near-neighbor list. The path diversification step will take place on the source node when it sends out a real data packet, whereas the packet injection step takes place on other rings at the same time. We discuss each of the three steps in the following subsections.

3.2.1. Initialization. The main goal of the initialization step is to establish the rings and to allow each sensor to choose the necessary lists for the path diversification step.

Ring formation depends on a beacon propagation, initiated by the data sink, with a hop count field. Every node overhearing this beacon message for the first time should record the hop count in every packet and rebroadcast the beacon after incrementing the hop count. Eventually, each node obtains to know its hop count from the sink. And it will generate its far-neighbor list, equidistant-neighbor list, and near-neighbor list. The far-neighbors are those neighbors that are at least one-hop further away from the sink. The near neighbors are those at least one-hop closer and the equidistant-neighbors are those with the same hop count to the sink [22–24].

3.2.2. Path diversification. The path diversification step ensures that traffic flow in network is evenly distributed. In the following discussions, we assume that a sensor node on the ath ring is going to send out a data packet.

(1) The source node selects three parameters: b, c, and α where $0 < b < a < c \leqslant L$. α is randomly chosen between 0 and π. These parameters, b, c, and α, are stored in the data packet so that forwarding sensors know the entry and exit points on the rings.
(2) The data packet is sent outward to the cth ring. This is through the far-neighbor lists of the forwarding nodes.
(3) Once the data packet arrives at the cth ring, it will travel counterclockwise on the ring for an α angle. This is through the equidistant-neighbors of the forwarding nodes.
(4) Once the data packet arrives at α angle of its entry point on the cth ring, it travels inward to the bth ring. This is through the near-neighbor lists of the forwarding nodes.
(5) Once the data packet arrives at the bth ring, it will travel counterclockwise on the ring for a $\beta = \pi - \alpha$ angle. This is through the equidistant-neighbors of the forwarding nodes.
(6) Once the data packet arrives at β angle of its entry point on the bth ring, it travels inward to the data sink. This is through the near-neighbor lists of the forwarding nodes.

In Figure 2, the data packet will be sent through the consecutive far-neighbors from the ath ring to the cth ring. Then, it travels on the cth ring for an α angle before going inside toward the bth ring. Then, it travels on the bth ring for a β angle before going toward the data sink directly. All transmissions on the rings are sent to those equidistant-neighbors (counterclockwise), and all transmissions between the rings are sent to those far-neighbors when the packet travels from inner ring to outer ring and near-neighbors from the outer ring to the inner ring.

In the example shown in Figure 2, $c = a + 2$ and $b = a - 2$. Therefore, the data packet travels from the ath ring to the cth ring in two hops and it travels from the cth ring to the bth ring in four hops.

It is easy to prove that a data packet will always reach the data sink given a non-partitioned network. Because each source will generate a set of b, c, and α for each data packet independently, the traffic on the rings are evenly distributed.

DOI: 10.1002/cpe

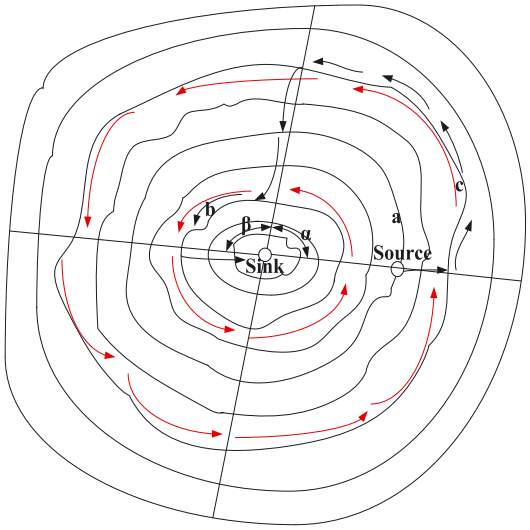

Figure 2. An operational example of our scheme. A source node is on the ath ring. It chooses the bth and the cth rings and the angles α and β. Then, the data packet travels outward from the ath ring to the bth ring. It travels counterclockwise for an α angle then goes inward to the bth ring, in which it travels another β angle. Then, it travels toward the data sink.

3.2.3. Fake packet injection. Fake packet injection is also adopted in our scheme to further protect source–location privacy. The goal of such transmissions is to hinder hop-by-hop trace back by the adversary. In our scheme, near-neighbors and far-neighbors of the forwarding nodes will inject fake packets with probability p. To limit the additional energy consumption caused by these fake packets, they are only forwarded for time-to-live hops.

In the example shown in Figure 2, nodes on the $b \pm 1$th rings and the $c \pm 1$th rings will generate fake packets with probability p.

4. PERFORMANCE ANALYSIS

In this section, we analyze the number of hop counts of data packet traveling from a source node to the sink in different schemes. We focus on two different schemes: the RSINSP [13] scheme and our proposed scheme. h is defined as the hop distance from the source to the sink.

In the RSINSP scheme, a source node randomly selects an intermediate node, and the data packet will be routed toward the intermediate node and then directly toward the data sink. We first estimate the probability density function (PDF) of nodes having a hop distance ℓ from the sink. We made a few simplifications to the network: we first assume that the network is circular by ignoring the border effect of the rectangle. The network's maximum hop count from the sink is L. There are N nodes in the network.

The number of nodes within ℓ hops from the data sink can be approximated as linearly proportional to ℓ^2. We denote this as

$$N(\ell) = \rho \ell^2 \,. \tag{1}$$

The expression for ρ is straightforward: the total number of nodes in the network can expressed as

$$N = N(L) = \rho L^2 \,. \tag{2}$$

Thus, we can estimate $N(\ell)$ as

$$N(\ell) = \frac{N \ell^2}{L^2} \,. \tag{3}$$

DOI: 10.1002/cpe

The probability that a chosen intermediate node has a hop count of ℓ from the sink can be simply expressed as

$$f(\ell) = \frac{N(\ell) - N(\ell-1)}{N} = \frac{2\ell - 1}{L^2} \quad \text{when } \ell = 1, 2, \cdots, L \tag{4}$$

The angular direction, θ, of the chosen intermediate node from that of the source node is a uniform distribution between 0 and 2π, that is, $\theta \in \mathcal{U}(0, 2\pi)$.

The hop distance from the source node to the chosen intermediate node is roughly

$$d(\ell, \theta) = \sqrt{h^2 + \ell^2 - 2h\ell\cos\theta}\,. \tag{5}$$

The hop count for the RSINSP scheme is then the expected value of the hop count toward the intermediate node and ℓ:

$$\begin{aligned} H^{(\mathrm{RSINSP})}(h) &= \sum_{\ell=1}^{L} \int_{\theta=0}^{2\pi} [\mathrm{d}(\ell, \theta) + \ell] f(\ell) \frac{1}{2\pi} \mathrm{d}\theta \\ &= \frac{1}{2\pi L^2} \sum_{\ell=1}^{L} \int_{\theta=0}^{2\pi} \left[\sqrt{h^2 + \ell^2 - 2h\ell\cos\theta} + \ell \right] (2\ell - 1) \mathrm{d}\theta \end{aligned} \tag{6}$$

If we take the lower bound of $\cos\theta$, which is -1, and replace it in (6). It becomes

$$\begin{aligned} H^{(\mathrm{RSINSP})}(h) &\leqslant \frac{1}{2\pi L^2} \sum_{\ell=1}^{L} \int_{\theta=0}^{2\pi} \left[\sqrt{h^2 + \ell^2 - 2h\ell \cdot (-1)} + \ell \right] (2\ell - 1) \mathrm{d}\theta \\ &= \frac{1}{2\pi L^2} \sum_{\ell=1}^{L} (h + 2\ell)(2\ell - 1) \cdot 2\pi \\ &= \frac{4L}{3} - \frac{1}{3L} + 1 + h\,. \end{aligned} \tag{7}$$

We can also take the upper bound of $\cos\theta$, which is $+1$, and replace it in (6). It becomes

$$\begin{aligned} H^{(\mathrm{RSINSP})}(h) &\geqslant \frac{1}{2\pi L^2} \sum_{\ell=1}^{L} \int_{\theta=0}^{2\pi} \left[\sqrt{h^2 + \ell^2 - 2h\ell \cdot (+1)} + \ell \right] (2\ell - 1) \mathrm{d}\theta \\ &= \frac{1}{2\pi L^2} \sum_{\ell=1}^{L} (|h - l| + \ell)(2\ell - 1) \cdot 2\pi \\ &= \frac{1}{L^2} \left[\sum_{\ell=1}^{h-1} + \sum_{\ell=h}^{L} \right] (|h - \ell| + \ell)(2\ell - 1) \\ &= \frac{4L}{3} - \frac{1}{3L} + 1 - h + \frac{h(2h-1)(h-1)}{3L^2}\,. \end{aligned} \tag{8}$$

Therefore, we have these simple upper and lower bounds for the hop count in the RSINSP scheme:

$$\frac{4L}{3} - \frac{1}{3L} + 1 - h + \frac{h(2h-1)(h-1)}{3L^2} \leqslant H^{(\mathrm{RSINSP})}(h) \leqslant \frac{4L}{3} - \frac{1}{3L} + 1 + h\,. \tag{9}$$

Now we will calculate the hop count of data packet transmission in our proposed scheme. Assume that the inner and outer rings are chosen randomly between $1, 2, \cdots, (h-1)$ and $(h+1), (h+2), \cdots, L$, respectively, and α is simply chosen randomly between 0 and π, that is, $\alpha \in \mathcal{U}(0, \pi)$, and $\beta = \pi - \alpha$.

The total hop of message transmissions is

$$c - h + c\alpha + c + b(\pi - \alpha)\,, \tag{10}$$

DOI: 10.1002/cpe

where the $c - h$ term represents the transmission from source to the outer ring, $c\alpha$ represents the transmission on the outer ring, $b(\pi - \alpha)$ represents the transmission on the inner ring, and c represents the overall transmission from outer ring toward the data sink.

The hop count is then the expected value of the earlier equation (for $h \in \{2, 3, \cdots, (L-1)\}$)

$$H^{(\text{Proposed})}(h) = \sum_{b=1}^{h-1} \sum_{c=h+1}^{L} \int_{\alpha=0}^{\pi} [c - h + c\alpha + c + b(\pi - \alpha)] \frac{1}{h-1} \cdot \frac{1}{L-h} \cdot \frac{1}{\pi} \mathrm{d}\alpha \,. \quad (11)$$

We evaluate the accuracy of our analysis in Section 5.2.

5. PERFORMANCE EVALUATION

We evaluate our proposed scheme in NS2 [24] and MATLAB simulator. Our MATLAB simulations are presented in Section 5.1. In MATLAB simulations, we concentrate on the traffic patterns as well as attack time of different schemes. In NS2 simulations, we focus on hop count and energy consumption of different schemes (Section 5.2).

As described in Section 2, we compare our method with the methods PSSP in [5] and RSINSP in [12]. In PSSP, every packet will be transmitted to the sink in its shortest-path after it is routed toward a fixed phantom source. Moreover, in RSINSP, the real source randomly selects an intermediate node in the network. Then, the data packet is sent to the intermediate node before it is routed to the sink along the shortest-path.

5.1. Spatial traffic evenness and attack time

We evaluate the traffic pattern of our scheme and compare it with two other schemes: shortest-path scheme [5] and the RSINSP scheme [13]. Our evaluation is based on the metric of *spatial traffic evenness* [25], in which we measure the difference of numbers of packet transmissions. Such an evaluation is performed on all nodes and some neighborhoods. The neighborhoods selected for evaluation are those with one-hop closer to the sink than the source node. We count the numbers of packet transmissions in these neighborhoods and compute the standard deviation divided by the mean of these numbers (i.e., coefficient variance). Our comparison is based on a transmission of a total of 100 packets and various number of sources. This evaluation does not include FPIs. We randomly distribute $N = 2,000$ sensor nodes in an area of size 1000×1000 m^2. The data sink is

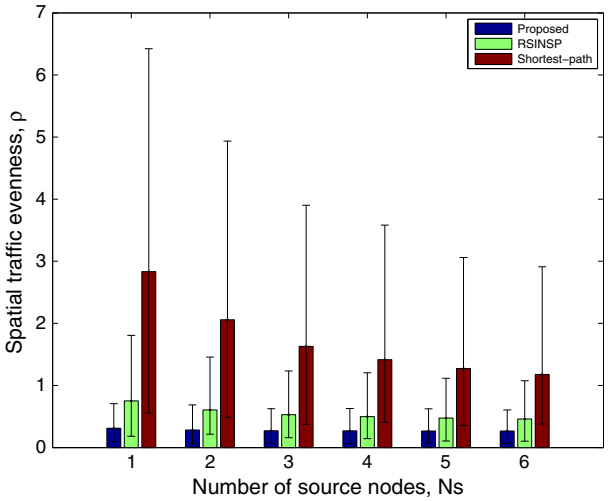

Figure 3. Spatial traffic evenness comparison of our proposed scheme, random selected intermediate node with single path (RSINSP), and shortest-path scheme in the regions one-hop closer to the sink than the source nodes. The 95% confidence intervals are shown in this figure.

DOI: 10.1002/cpe

assumed to be at the center of this region. The wireless transmission range (as well as the adversary observing range) is $R = 75$ m.

The results observed on those nodes one-hop closer to the sink than the source node are shown in Figure 3. We can see that the spatial traffic pattern is much more even in our scheme compared with RSINSP and the shortest-path scheme. As the number of source nodes increases, the spatial traffic pattern becomes more even, as these source nodes spread the transmissions in the entire network region. However, our proposed scheme's advantage can still be seen clearly. Another interesting observation is that our scheme offers similar spatial traffic evenness for different numbers of sources, a clear indication of a rather evenly distributed transmission pattern in these observed regions. The RSINSP offers a distributed traffic between our proposed scheme and the shortest-path scheme, with its randomly chosen phantom sources.

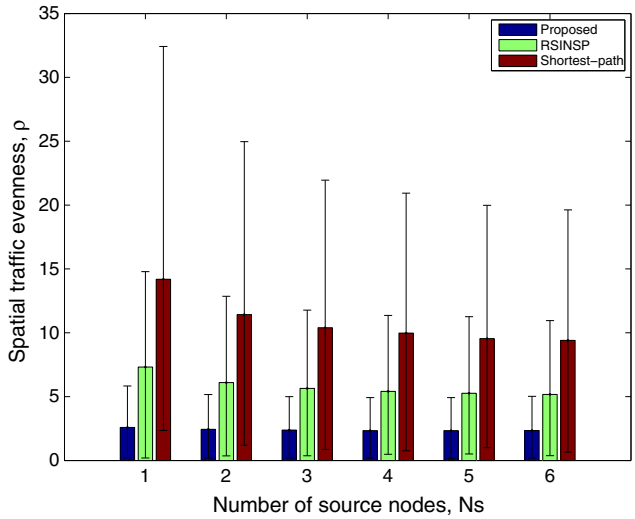

Figure 4. Spatial traffic evenness comparison of our proposed scheme, random selected intermediate node with single path (RSINSP), and shortest-path scheme among all nodes in the entire network. The 95% confidence intervals are shown in this figure.

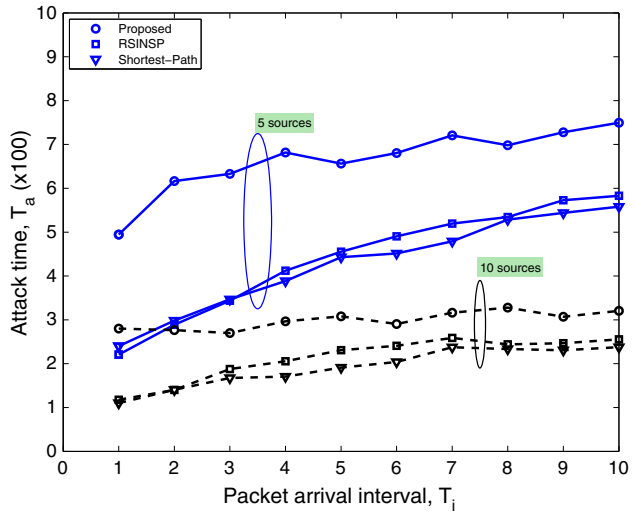

Figure 5. Attack time comparison of our proposed scheme, random selected intermediate node with single path (RSINSP), and the shortest-path scheme for different packet arrival intervals. The pause time was assumed to be 10 units and FPI probability $p = 0.1$.

DOI: 10.1002/cpe

Figure 4 presents the spatial traffic evenness of different schemes in the entire network region. In this evaluation, we count the packet transmissions of every node in the network and compute the spatial traffic evenness. The same conclusions can be drawn, except that the traffic is less evenly distributed when we consider the entire network.

We also evaluate the attack time of different schemes. The attack time is defined as the time for the adversary to trace back to any one of the source nodes. Our simulations use a FPI probability $p = 0.1$ and a pause time of 10 time units (pause time is defined as the time during which the adversary stays in one location and observes packet transmissions before deciding to move on its next location). Note that the actual values of the attack time are of no importance, but the relative values are. The results are presented in Figure 5. The attack time for our scheme is significantly longer than the RSINSP and the shortest-path schemes. Interestingly, the RSINSP scheme only offers a slight increase of attack time compared with the shortest-path scheme. Another observation from Figure 5

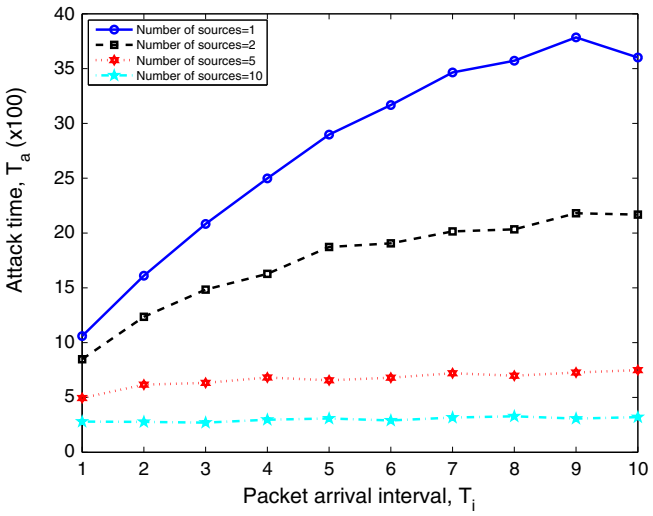

Figure 6. Attack time comparison of our proposed scheme for networks with different number of sources. The pause time was assumed to be 10 units and fake packet injection probability $p = 0.1$.

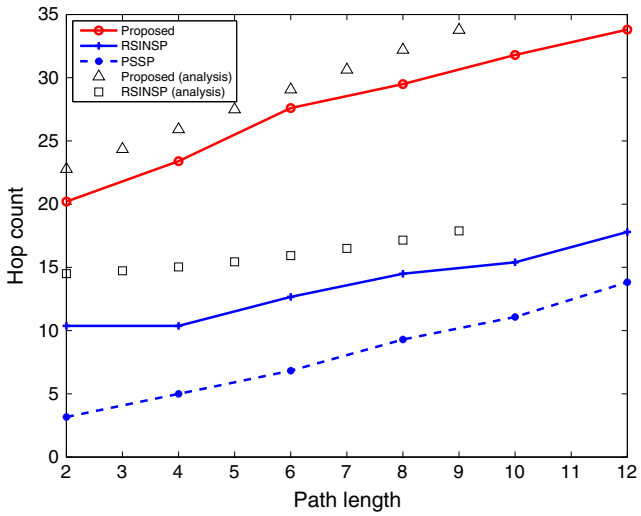

Figure 7. Hop count comparison of the proposed scheme, phantom source single path (PSSP), and random selected intermediate node with single path (RSINSP).

DOI: 10.1002/cpe

is that the attack time increases with packet arrival interval. This can be explained by the decreased overall traffic as packet arrival interval increases.

In Figure 6, we compare our proposed scheme's attack time for different number of sources. When the number of sources in the network is larger, the attack time is shorter. This is because it is more likely for the adversary to bump into a source traffic when there are more sources in the network.

5.2. Energy consumption

Our NS2 simulations are performed to focus on hop count and energy consumption as well as the effect of FPI. In a network region of $400 \times 400\ \mathrm{m}^2$, a total of $N = 440$ nodes are randomly distributed. The sink is deployed at the center of the region. The wireless transmission range is $R = 25\ \mathrm{m}$. We use hop count as the metric to evaluate energy consumption.

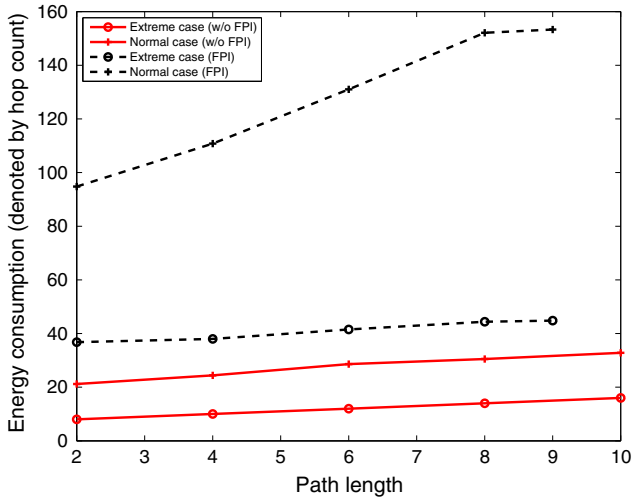

Figure 8. Energy consumption of the proposed scheme, with or without fake packet injection (FPI). Extreme cases are plotted with the average/normal cases.

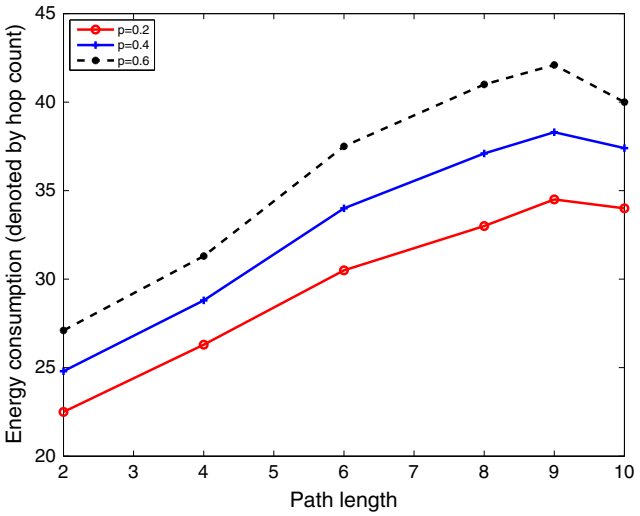

Figure 9. Energy consumption of the proposed scheme with different fake packet injection probability. The results level off as path hop reaches the network boundary, causing the boundary effect.

DOI: 10.1002/cpe

The hop count results are presented in Figure 7. Path length is defined as the hop count from the source node to the data sink. As path length increases, the hop counts in all schemes increase. Due to the additional travel on the rings, our proposed scheme has a larger hop count as compared with the RSINSP and the PSSP schemes. The RSINSP scheme has a larger hop count than the PSSP scheme because of its randomly selected phantom sources. Numerical results from our analysis, (6) and (11), are also presented in Figure 7. Our numerical results match simulation results nicely, except that numerical results are slightly higher than the simulation results. This can be explained by the simplification that we made in our analysis, as we measure each distance with hop count instead of actual distance.

We also evaluate the energy consumption of our scheme with or without FPI. The results are presented in Figure 8. As can be seen, the energy consumption denoted by hop count increases with path length, the hop count distance between the source node and the data sink. When FPI is turned on, our scheme consumes more energy. The energy consumption levels off when the path length reaches network border, causing a boundary effect.

In Figure 9, we compare the energy consumption of different FPI probability, p. As p increases, energy consumption increases as more and more nodes overhearing data packet transmissions send out fake packets to trick the adversary. The extreme cases are those using a shorter route toward the data sink. Therefore, they have lower energy consumption. The boundary effect of a large path length is present again.

6. CONCLUSION

Source–location privacy is critical to the successful deployment and operation of WSNs. In this paper, we have proposed a new scheme to protect the source–location privacy in WSNs through two methods: randomly selected rings and FPI. Multiple rings, centered at the data sink, are formed by the sensor nodes. When a source sends packets, two random rings will be chosen. Fake packets serving as decoys will be injected by some of the neighboring nodes on the adjacent rings of these two rings. Thus, a more uniform spatial traffic pattern can be formed .

We have presented our analysis as well as performance evaluation in NS2 and MATLAB. Our evaluations focused on spatial traffic evenness, attack time, hop count, and energy consumption. Our results show that the proposed scheme provides much better spatial traffic evenness and attack time compared with other privacy protection schemes, with a modest increase of energy consumption.

In future work, we plan to optimize our algorithm on the basis of energy consumption and implement it on more practical networks.

ACKNOWLEDGEMENTS

This research is sponsored in part by the National Natural Science Foundation of China and the Fundamental Research Funds for the Central Universities (contract/grant number: No. 61173179, No. 61202441 and No. DUT13JS10). J. Deng's research was supported in part by the Kohler Fund at the University of North Carolina at Greensboro.

REFERENCES

1. Yick J, Mukherjee B, Ghosal D. Wireless sensor network survey. *Computer Networks* 2008; **52**(12):2292–2330.
2. Wang Y, Attebury G, Ramamurthy B. A survey of security issues in wireless sensor networks. *IEEE Communications Surveys and Tutorials* 2006; **8**(2):2–23.
3. Chen X, Makki K, Yen K, Pissinou N. Sensor network security: a survey. *IEEE Communications Surveys Tutorials* 2009; **11**(2):52–73.
4. Li N, Zhang N, Das S, Thuraisingham B. Privacy preservation in wireless sensor networks: a state-of-the-art survey. *Ad Hoc Networks* 2009; **7**(8):1501–1514.
5. Kamat P, Zhang Y, Trappe W, Ozturk C. Enhancing source-location privacy in sensor network routing. *Proceedings of the 25th Distributed Computing Systems (ICDCS'05)*, Columbus, OH, USA, 2005; 599–608.
6. Jian Y, Chen S, Zhang Z, Zhang L. A novel scheme for protecting receiver's location privacy in wireless sensor networks. *IEEE Transactions on Wireless Communications* 2008; **7**(10):3769–3779.

DOI: 10.1002/cpe

7. Li X, Wang X, Zheng N, Wan Z, Gu M. Enhanced location privacy protection of base station in wireless sensor networks. *Proceedings of the 5th International Conference on Mobile Ad-hoc and Sensor Networks (MSN '09), IEEE*, Fujian, China, 2009; 457–464.
8. Shao M, Hu W, Zhu S, Cao G, Krishnamurth S, La Porta T. Cross-layer enhanced source location privacy in sensor networks. *Proceedings of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON'09)*, Rome, Italy, 2009; 1–9.
9. Ozturk C, Zhang Y, Trappe W. Source-location privacy in energy-constrained sensor network routing. In *Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN'04)*. ACM: New York, NY, USA, 2004; 88–93.
10. Xi Y, Schwiebert L, Shi W. Preserving source location privacy in monitoring-based wireless sensor networks. *Proceedings of the 20th International Conference on Parallel and Distributed Processing Symposium (IPDPS'06)*, Rhodes Island, Greece, 2006; 1–8.
11. Wang H, Sheng B, Li Q. Privacy-aware routing in sensor networks. *Computer Networks* 2009; **53**(9):1512–1529.
12. Li Y, Lightfoot L, Ren J. Routing-based source-location privacy protection in wireless sensor networks. *Proceedings of the IEEE International Conference on Electro/Information Technology (EIT'09), IEEE*, Ontario, Canada, 2009; 29–34.
13. Li Y, Ren J. Mixing ring-based source-location privacy in wireless sensor networks. *Proceedings of the Internatonal Conference on Computer Communications and Networks (ICCCN'09)*, San Francisco, CA, USA, 2009; 1–6.
14. Li Y, Ren J. Preserving source-location privacy in wireless sensor networks. *Proceedings of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON'09), IEEE*, Rome, Italy, 2009; 1–9.
15. Li Y, Ren J. Providing source-location privacy in wireless sensor networks. *Wireless Algorithms, Systems, and Applications (WASA'09)*, Boston, MA, USA, 2009; 338–347.
16. Li Y, Ren J. Source-location privacy through dynamic routing in wireless sensor networks. *Proceedings of the IEEE INFOCOM'10*, San Diego, CA, USA, 2010; 1–9.
17. Luo X, Ji X, Park M. Location privacy against traffic analysis attacks in wireless sensor networks. *Proceedings of the International Conference on Information Science and Applications (ICISA'10), IEEE*, Seoul, South Korea, 2010; 1–6.
18. Ouyang Y, Le Z, Liu D, Ford J, Makedon F. Source location privacy against laptop-class attacks in sensor networks. In *Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks (SECURECOMM'08)*. ACM: New York, NY, USA, 2008; 1–10.
19. Mehta K, Liu D, Wright M. Location privacy in sensor networks against a global eavesdropper. *Proceedings of the IEEE International Conference on Network Protocols (ICNP'07)*, Beijing, China, 2007; 314–323.
20. Shao M, Yang Y, Zhu S, Cao G. Towards statistically strong source anonymity for sensor networks. *Proceedings of the IEEE INFOCOM'08, IEEE*, Phoenix, AZ, USA, 2008; 51–55.
21. Yang Y, Shao M, Zhu S, Urgaonkar B, Cao G. Towards event source unobservability with minimum network traffic in sensor networks. In *Proceedings of the First ACM Conference on Wireless Network Security (WISEC'08)*. ACM: New York, NY, USA, 2008; 77–88.
22. Zeng Y, Zhang S, Guo S, Li X. Secure hop-count based localization in wireless sensor networks. *Proceedings of the International Conference on Computational Intelligence and Security (CIS'07), IEEE*, Harbin, China, 2007; 907–911.
23. Huang X, Deng J, Ma J, Wu Z. Fault tolerant routing for wireless sensor grid networks. *Proceedings of 2006 IEEE Sensors Applications Symposium (SAS '06)*, Houston, TX, USA, 2006; 66–70.
24. (Available from: http://www.isi.edu/nsnam/ns/) [Accessed on 2011].
25. Kim J, Park D, Theocharides T, Vijaykrishnan N, Das C. A low latency router supporting adaptivity for on-chip interconnects. In *Proceedings of the 42nd Annual Design Automation Conference (DAC '05)*. ACM: New York, NY, USA, 2005; 559–564.

DOI: 10.1002/cpe