Preserving Source-Location Privacy through Redundant Fog Loop for Wireless Sensor Networks

Preserving Source-Location Privacy through Redundant Fog Loop for Wireless Sensor Networks

Mianxiong Dong, Kaoru Ota
Department of Information and Electronic Engineering
Muroran Institute of Technology
Hokkaido, Japan
{mx.dong, ota}@csse.muroran-it.ac.jp

Anfeng Liu
School of Information Science and Engineering
Central South University
ChangSha, 410083, China
afengliu@csu.edu.cn

Abstract—**A redundant fog loop-based scheme is proposed to preserve the source node-location privacy and achieve energy efficiency through two important mechanisms in wireless sensor networks (WSNs). The first mechanism is to create fogs with loop paths. The second mechanism creates fogs in the real source node region as well as many interference fogs in other regions of the network. In addition, the fogs are dynamically changing, and the communication among fogs also forms the loop path. The simulation results show that for medium-scale networks, our scheme can improve the privacy security by 8 fold compared to the phantom routing scheme, whereas the energy efficiency can be improved by 4 fold.**

Keywords—wireless sensor networks, source-location privacy, redundant fog loop, performance optimization

I. INTRODUCTION

Wireless sensor networks (WSNs) are composed of a large number of sensor nodes that communicate with each other through multi-hop wireless links [1, 2]. Sensor networks rely on a wireless communication medium for broadcasting, which can be eavesdropped easily [3]. Adversaries may use expensive radio transceivers to intercept the networks and make use of the message flow patterns to trace the source of messages by moving along the reversed path [4, 5, 6, 7], even if strong data encryption is utilized. However, this ability to intercept such messages has become an issue of concern because in certain critical situations, such as endangered species or a vehicle with military officers, the privacy of the object of interest is important, and the location information should not be disclosed for safety reasons [4-7].

In this paper, we propose a redundant fog loop-based scheme (RFL scheme) that has a good privacy-preserving ability and energy efficiency. The major contributions of this paper are as follows:

- We propose a RFL scheme with strong privacy-preserving ability. This scheme creates many interference branch paths as well as many redundant fogs, which make it difficult for adversaries to determine where the real source node is and thus improves the privacy-preserving ability by many fold.
- The RFL scheme has high energy efficiency and lifetime performance. Through a detailed analysis of the energy consumption in the network, we use residual energy in non-hotspot regions to create more fake fogs. This step improves the network privacy status and optimizes the network energy resource utilization, hence maximizing the network lifetime.
- The RFL scheme has also been subjected to extensive simulations using Omnet++ [8], and the simulation results further strengthen the validity of our proposed scheme. When compared with other approaches for medium-scale sensor networks, our scheme can improve the privacy security by 8-10 fold and the energy efficiency by more than 4 fold.

The remainder of this paper is organized as follows: In Section II, the related work is reviewed. The system model is described in Section III. In Section IV, the details of the RFL scheme are presented. Section V offers an analysis and comparison of simulation results, and Section VI provides the conclusion and ideas for future work.

II. RELATED WORK

The privacy threats that exist for sensor networks can be broadly classified into two dimensions: (i) content-based privacy threats and (ii) context-based privacy threats [9]. Content-based threats are well understood [3] and are often addressed using cryptographic techniques. There are certain aspects of cryptographic techniques that cannot be widely used to solve context-based privacy threats [3, 9-11]. Context-based privacy is more challenging [3]. One aspect of context that is important in several applications is the preservation of source location privacy.

Many studies have addressed the preservation of source location privacy for WSNs [3, 4, 9-13]. The existing research can be divided into two categories based on adversary ability, namely, source location privacy-preserving protocols against local attacks [3, 6, 9-13] and source location privacy-preserving protocols against global attacks [14-17]. The flow sensed by adversaries is relatively minimal under local attacks. For example, adversaries can only sense wireless communication within one hop. In a global attack, adversaries can sense wireless communication within the entire network; thus, the global attack has strong attacking ability.

To withstand adversaries with global attack ability, [18] proposed the ConstRate protocol, which is based on the premise that all nodes in the entire network send data packets with a constant rate regardless of whether real data packets are received. This protocol effectively defends against global traffic analysis attacks, but the introduction of extensive pseudo-packets leads to a sharp decline in the network lifetime and an increase in the transport delay of actual data packets. As an enhancement, a proxy-based filtering protocol was proposed in [19], where a sensor node that serves as a proxy can filter out fake data packets, thereby reducing network traffic. Bicakci, Kemal et al. [16] proposed a filtering idea called the Optimal Filtering Scheme) to maximize the network lifetime and preserve event-unobservability against global eavesdroppers. Afterwards, [17] proposed a FitProbRate protocol which proved that, by adjusting the nodal data transmission rate, the source location privacy can be preserved and the transport delay can also be reduced.

However, more recent research, such as [17-19], has demonstrated that there are certain limitations for global eavesdroppers. Because all nodes are sending a large number of fake packets, it will greatly increase the nodal energy consumption, reduce the network lifetime, increase the packet collision probability and reduce the efficiency of packet transmission, in addition to increasing network delay. [11] noted that it is difficult to be a global eavesdropper in practice. Moreover, if the adversary has global ability, it is difficult to preserve the source location privacy. Therefore, adversaries only have local ability in practice, and research in this area has more practical significance.

For local eavesdroppers, [6] introduced the Panda-Hunter game model for source location privacy. In this model, a large number of sensor nodes are deployed to monitor the wild habits of animals, such as pandas. Once the behavior of the panda is monitored, the sensor node closest to the panda will transmit the observed results to the base station. The hunter watching near the sink can locate the source node by tracing it in reverse hop-by-hop, eventually capturing the panda.

C. Ozturk et al. proposed a famous phantom routing protocol [11] to protect against eavesdropping attacks on the source location. However, the phantom node in the original proposed protocol is closer to the source node. As a result, the source location may still be easily found by the adversary. Both theoretical and practical results demonstrate that if the message is routed randomly for h hops, the message will largely be within h/5 hops away from the actual source. Several approaches have been proposed to solve this problem. [11] designed a directed walk through either a sector-based approach or hop-based approach to ensure the phantom node is as far from the actual source as possible, thereby reducing the threat to the source node when the adversary is traced to the phantom node.

There is considerable research on privacy-preserving protocols based on phantom routing. Yun Li et al. [4] developed two-phase dynamic routing-based schemes to provide source-location privacy. The main idea is to route the message to a node away from the actual message source randomly and then forward the message to the sink node using single-path routing.

The fake source idea was proposed to introduce more sources to the sensor network, which generates fake messages that are the same length as the real messages and encrypted as well so that an adversary cannot differentiate between the real and the fake message. In this scenario, it is expected that an adversary will be directed to a fake source [9], and the goal of privacy preservation is achieved.

Recent research has taken further measures to confuse adversaries; for example, a situation in which a phantom node is near the sink is not sufficiently confusing. M. E. Mahmoud and X. Shen proposed a cloud-based privacy-preserving scheme [7]. This scheme generates several fake source nodes around the real source node, and all source nodes send data and have routing paths. Therefore, a complex “cloud” area is formed in certain regions, whereby adversaries feel trapped in the cloud and are unable to recognize the real source node. Hence, it has stronger privacy-preserving ability than phantom routing.

In this paper, we are fully aware of the dynamics between energy consumption and network lifetime; thus, our goal is to not only preserve source node-privacy but also optimize network lifetime. We can fully use the remaining energy in non-hotspot regions to enhance the privacy-preserving ability without affecting network lifetime, thus maximizing network lifetime and greatly improving the preserving ability and energy efficiency.

III. The System Model

A. The System Model

1) Network model

We make the following general assumptions about our network model:

a) Our network model is similar to the explanatory Panda-Hunter Game introduced in [5, 6, 21]. In this Panda-Hunter Game, a sensor network is deployed with nodal density ρ to continuously monitor activities and locations of the animals in a wild animal habitat. As soon as a panda is discovered [6, 11], the corresponding source node in the nearby area will observe and report data to the sink node periodically [9, 11, 13]. The sink and sensor nodes are stationary. The sensor nodes are resource-constrained devices with low battery power and computation capacity.

b) The observed targets are randomly distributed in the network, i.e., the probability of each sensor node monitoring the target is equal, and thus, the probability of generating data to the sink is equal.

c) We consider that a security infrastructure, such as secure communication, has been built in. That is, no information carried in the message (e.g., packet head) will be disclosed. The key management, including key generation, key distribution and key update, is beyond the scope of this paper. However, interested readers are referred to such references as [3, 13].

2) *The Adversary Model*

Because of the high profits related to panda hunting, adversaries would try their best to equip themselves with advanced equipment, which means that they would have some technical advantages over the sensor nodes [5]. In this paper, the adversaries are considered to have the following characteristics:

a) The adversaries have sufficient energy resources, computation capability and memory for data storage [6, 11, 12, 13]. The adversary observes the wireless communication within a certain detection range. Upon detecting an event, they could determine the immediate sender by analyzing the strength and direction of the signal they receive. They can move to this sender's location without any delay. We also consider that the adversaries will never miss any event when the node sent packet is in the communication radius of adversaries.

b) The adversaries (such as hunters) are intelligent. They eavesdrop on the wireless transmissions and attempt to make use of the network traffic to determine the locations of pandas to hunt them. However, they cannot monitor the traffic of the entire network.

c) The adversaries will not interfere with the proper functioning of the network, such as modifying packets, altering the routing path, or destroying sensor devices, as such activities can be easily identified [12, 13]. However, the adversaries may carry out passive attacks, such as eavesdropping of the communications.

TABLE I. NETWORK PARAMETERS

Parameter	Value
Threshold distance (d_0) (m)	87
E_{elec} (nJ/bit)	50
e_{fs} (pJ/bit/m^2)	10
e_{amp} (pJ/bit/m^4)	0.0013
Initial energy (J)	0.5

B. *Energy Consumption Model and Related Definitions*

In this paper, we adopt the typical energy consumption model in [13, 21], where the transmission energy consumption E_t follows Eq. 1 and energy consumption E_r for receiving follows Eq. 2.

$$\begin{cases} E_t = lE_{elec} + l\varepsilon_{fs}d^2, \text{ if } d < d_0; \\ E_t = lE_{elec} + l\varepsilon_{amp}d^4, \text{ if } d > d_0. \end{cases} \tag{1}$$

$$E_r = lE_{elec} \tag{2}$$

where E_{elec} represents the transmission of circuit loss. Both the free space (d^2 power loss) and the multi-path fading (d^4 power loss) channel models are used. If the transmission distance is less than the threshold d_0, the power amplifier loss is based on the free-space model; in the opposite case, the multi-path attenuation model is used. ε_{fs} and ε_{amp} are the energy required by power amplification in the two models. l is the number of bits in a packet. The above parameter settings are given in *Table 1*, as adopted from [13, 21, 22].

IV. RFL SCHEME DESIGN

A. *Overview of the Proposed Scheme*

The overall structure of the RFL scheme proposed in this paper is composed of two main parts: (1) fogs with branches, which serve as routing paths within and around fogs, and (2) the larger loop routing path among fogs leading to the sink. This novel structure gives our RFL scheme good privacy-preserving ability and network lifetime, as discussed below.

1) The RFL scheme achieves good privacy-preserving ability through two key components.

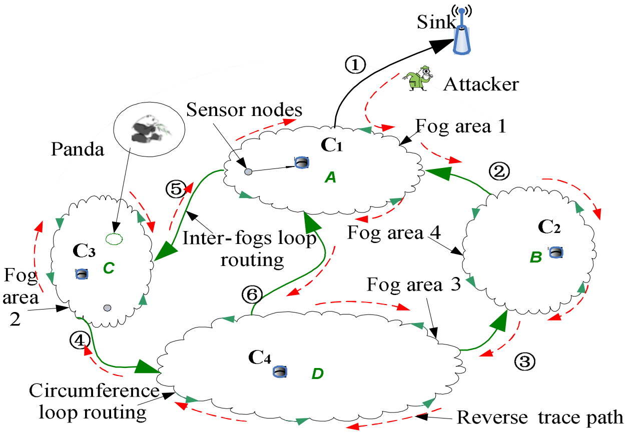

Fig. 1. Illustration of RFL-based routing

a) The fog: multiple branch routing paths are generated within the fog, but only the routing path branch with the source node sends real data; the other branches are interference. All branches route away from the fog center and generate routing branches continuously. All branches gather in a circular routing loop called the fog. In such a structure, the adversary traces back to the fog and finds that there are many branch paths, which is similar to entering into a fog area, and the real source path cannot be distinguished. Therefore, we call it the fog area. Unlike previous studies, this paper not only creates fog where the real source node is but also creates multiple fogs where there is no real source node. Most of the proposed approaches only seek to create interference routings in the area where the real source node resides to confuse the adversary. Some good examples are the famous phantom routing [6] and the cloud-based scheme that only creates a cloud around the source node [7]. In this paper, we refer to the scheme that sends data to the sink with a traditional routing strategy (such as the shortest route) as a near-source interference scheme because it only creates interference routing around the real source node. The drawback of this type of scheme is that the adversary can approximately compute the range of a real source node when it traces to the phantom node or the cloud, which can pose a great threat to the protected object. However, in this paper, as shown in Fig. 1, the RFL scheme creates interference fogs both in the real source node and other non-source node vicinities. For example, in Fig. 1, fogs A, B, C, D are created. Because they are connected via each other, the source node can be located in any of them. Thus, if the same attack method is adopted, the probability of the source node being attacked with our scheme is only 1/4 of that with phantom routing or a cloud-based routing scheme. In other words, our RFL scheme improves the privacy-preserving ability by 4 fold.

b) The loop routing: This is another important component for improving the ability to preserve privacy. There are two types of loop routings. One is around fog peripherals. The benefit of this type of loop routing is that it is effortless and easy for adversaries to discover and trace back. Hence, it is more likely to mislead adversaries only to be trapped in a routing loop along fog peripherals. Thus, the probability of an adversary gaining access to the fog is minimal, which can dramatically reduce the probability of discovering the source node. This type of loop routing provides the local source-location privacy. The other type of loop routing is routing among fogs. When the adversary traces back from the sink, it can easily be trapped into the larger routing loops among fogs. As shown in Fig. 1, when the adversary traces back from the sink along ⊖, which is marked by a red arrow, it can be easily entered into the routing loop among fogs along the route ⊜→⊛→④→⑤→⊜ or the routing loop route ⑥→④→⑤→⑥. In addition, the following two reasons pose great difficulty for the adversary. The first reason is that it takes a long time for the adversary to return back to the paths that have already been traced because of longer loop routing among fogs; at worse, it may never even realize the path, consequently making it difficult to infer the region of the real source node. The other reason is that fogs are dynamically generated, and they disappear after some time. Therefore, after a period of trace time, the original routing loop is completely replaced by a new one. Thus, it always appears as a non-repeating and non-terminating path for the adversary. In other words, if the dynamical fog creation time is shorter than the time required for an adversary to trace back a loop, the adversary will encounter new routing loops unabated. This offers network-level source-location privacy.

2) The RFL scheme achieves energy efficiency and network lifetime optimization through the following mechanism.

Nodes in WSNs are simple and inexpensive with limited energy. Preserving privacy is achieved at the cost of increased energy expenditures. For example, all of the network nodes broadcast regularly in the privacy-preserving scheme proposed by [15, 16]. Although it has stronger privacy-preserving ability against a global attack, its larger energy consumption decreases network lifetime greatly and makes it impractical. In contrast, when privacy preservation is not considered, the routing protocols have a high network lifetime but are insecure. For instance, in phantom routing, privacy is achieved by generating a phantom node around the source node, and in cloud-based routing, privacy is improved by creating more fake source node clouds. These approaches deplete a certain amount of energy in generating phantoms and creating fake source clouds to enhance privacy, thus reducing the network lifetime. Unlike previous studies, the RFL scheme in this paper has the same lifetime as a normal unprotected routing protocol but possesses a stronger privacy-preserving ability than current secured schemes, as the area near the sink in WSNs is a hotspot region and the lifetime of an entire network is determined by the node lifetime of a hotspot region. In contrast, more than 90% of the node energy remains unutilized in non-hotspot areas after the network dies [13, 21]. Thus, redundant fogs and routings loops are created in the non-hotspot region for preserving the source node privacy of the network. The hotspot receives only real data from the non-hotspot region for transmission to the sink without any fogs created, and thus, there is no further energy burden on the hotspots. Therefore, the network lifetime in our scheme is optimal. Our scheme fully uses the energy of the non-hotspot region such that when the network dies, nearly all of the energy of the entire network is fully utilized, thus enhancing energy efficiency.

B. Construction of a Single Fog

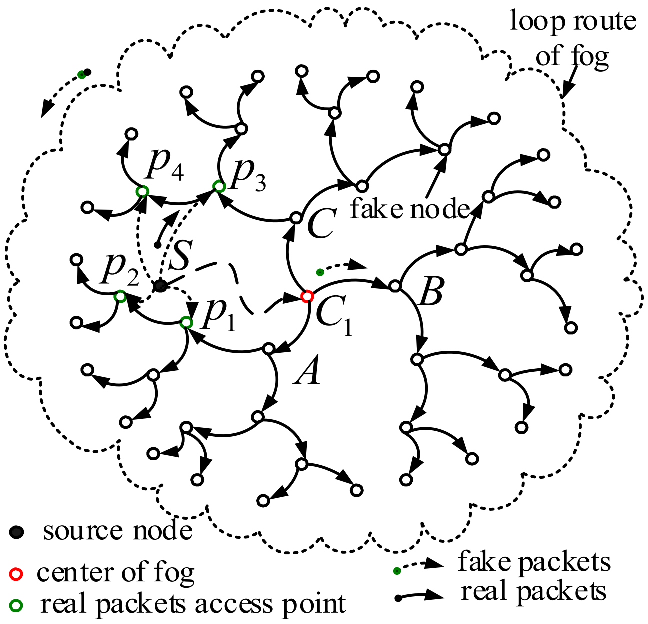

Fig. 2. Construction of a single fog

Fog is the basic unit composition of the RFL scheme; its structure is shown in Fig. 2. The end node C_1 after random walk ς hops of source node S is the fog center. There are two reasons why the sensor node S is not at the center of the fog: (a) The source node at the center can be easily inferred by the adversary, making it vulnerable to attacks. (b) Through a random walk mechanism, although the adversary may know the algorithm of this paper, it cannot estimate the distance from the source node to the fog center or the approximate location of the source node. When C_1 is certain, σ_1 branch routing paths are issued uniformly away from C_1, and then, each routing path branch issues σ_2 routing paths; this process is continued until the hops from the fog center to the current node are δ. Then, connect all leaf nodes whose hops to the fog center are δ. The detailed process is as follows. First, C_1 is the end point where source node S randomly walks ς hops; set C_1 as the center. C_1 selects σ_1 neighbor nodes as the first layer of branch routing paths (nodes A, B, and C in Fig. 2); these σ_1 nodes must be uniformly distributed; that is, consider the angle to C_1 to be $\alpha_1 = 2\pi / \sigma_1$; C_1 selects a neighbor node randomly, for example, node A in Fig. 2, and then selects one node at each α_1 degree direction, such as nodes B and C in the figure. Each node in the first layer selects σ_2 neighbor nodes in the same manner as C_1 ; similarly, these σ_2 nodes must be distributed uniformly, and the angle between these σ_2 nodes can be calculated as $\alpha_2 = 2\pi / (\sigma_1 \sigma_2)$. Similarly, each node in the second layer selects σ_2 nodes as the third layer of the fog, and the branch routing path generation is completed until the δ th layer node is selected. Nodes in the outermost layer (δ th layer) are called leaf nodes; after the leaf nodes are selected, connect them, and then the routing loop paths are completed, and thus, the construction of the entire fog is finished.

C. Construction of Multiple Fogs and Formation of Loop Routing Paths among Fogs

This section mainly discusses how to create multiple fogs after the construction of a single fog. The selection of multiple fogs must meet the following principles: (1) The created fogs must balance the network energy consumption to enhance energy efficiency and optimize network lifetime. (2) The location of these fogs must be diverse and random such that the adversary cannot infer fog in which the real source node is located.

First, the selection of a fog location must be determined by energy consumption to achieve balanced energy consumption. Because energy consumption varies considerably in different regions of the network, the energy consumption can be balanced when the rate of creating fogs is proportional to the remaining energy.

Consider the probability of selecting these $\vartheta - 1$ locations to be $\omega = \{\omega_1, \omega_2, ...\omega_{\vartheta-1}\}$. These locations indicate how far away from the sink the fogs should be created. For any selected location ϖ_i , nodes on the circumference of ϖ_i from the sink meet the requirements. Therefore, it is flexible for the RFL to select the fog center, which makes the fogs diverse and random. The process of creating fogs is as follows.

As shown in Fig. 3, C_1 is the center of the fog where the source node is located; given its connection to the sink, $C_1 S$ is the Y axis, and the vertical direction of $C_1 S$ is the X axis. Sort $\{\omega_1, \omega_2, ...\omega_{\vartheta-1}\}$ as $\{\upsilon_1, \upsilon_2, ...\upsilon_{\vartheta-1}\}$ according to $C_1.h$ ($C_1.h$ is the distance from C_1 to the sink). Construct the first fog on the right side of this fog, and select an angle α' randomly from $[\alpha_1, \alpha_2]$. If $\upsilon_1 > C_1.h$, then α' is an angle in the first quadrant and in a counterclockwise direction to the X axis. If $\upsilon_1 < C_1.h$, then α' is an angle in the fourth quadrant and in a clockwise direction to the X axis. Start from the α' direction, and route forward until the distance from the current routing node to the sink is υ_1 ; set this point as the new fog center, namely, node C_2 in the figure, and then construct the fog as described in the previous section. Similarly, construct the fog on the right side of the source node; that is, randomly select an angle α'' from $[\alpha_1, \alpha_2]$ (see in Fig. 3) and route along this direction to the fog center until point C_3 is reached, and the distance from C_3 to the sink is υ_2 ; then, construct the second fog. The third fog starts from the center of the fog where the source node is; construct a fog until route C_4 is reached. The construction principle of subsequent fogs can be described as follows. (a) To the right, construct a fog on the rightmost of the fogs already constructed, such as fog 5 in Fig. 3; (b) then, to the left, construct a fog on the leftmost of fogs; (c) then, construct fogs starting from the non-outermost fogs; (d) repeat (a) and (c) until the number of constructed fogs is $\vartheta - 1$.

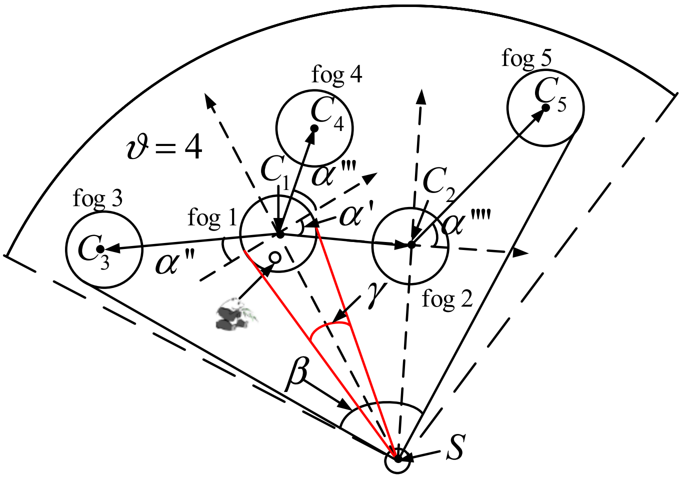

Fig. 3. Selection of fog centers

The routing loops among fogs are formed as follows. First, start from the rightmost fog of the source node, select the nearest fog according to the right-hand rule, connect the first found fog with the shortest routing strategy, and then find the next fog forward in the same manner. Repeat this process until a larger routing loop is formed by connecting all fogs. To increase the complexity of the routing loops, randomly select two fogs and connect them with the shortest routing strategy (from the fog away from the sink to the fog near the sink). Finally, select a node dynamically and randomly from the routing path near to the sink and route it to the sink with the shortest routing (for security, selection in this manner can make the shortest routing to the sink vary considerably, which can confuse the adversary and disperse node flow to enhance the ability against hotspot attacks). The final network structure is shown in Fig. 1.

D. Routing protocol of RFL

We have discussed the network structure of the RFL scheme, and the routing protocol of RFL is described below.

1) The routing protocol within fogs (for example, the source node fog). The protocol starts from the fog center C_1, C_1 routes to nodes in the first layer, and when the data are received, each node pauses for a random period of time τ. Data are then sent from nodes in the first layer to nodes in the second layer, and finally, all data are sent to the routing loop around the fog. During this process, if nodes receive real data, *the fake data will be discarded, and only real data are sent to the sink.*

2) The routing protocol of loop routing. Nodes in the routing loop send data according to the following rule: If no real data are received during the data transmission cycle, one dummy message will be generated and stored. However, if real data are received, then fake data are discarded and the real data are stored. In terms of the data transmission time, the stored data are sent to the next route node; the routing direction is similar to the fog construction method in the previous section.

3) The data transmission protocol of real source node S. Node S randomly selects one node p_1 nearby, as shown in Fig. 2, and then sends data to the sink via the access point p_1; after a period of time, it selects another access point p_2, continues in this manner, and selects p_3, p_4. When these access points receive real data, they will send real data forward within a predetermined period of time.

V. Analysis of the Simulation Results And Performance

OMNET++ is used for simulation verification. OMNET++ is an open network simulation platform that provides open source, component-based, modular simulation platform for large networks; it has been widely recognized by academics [8]. Without special instructions, the simulation scene is set as follows: 4,000 nodes are deployed in the simulations, the network radius is set to $R = 600$ and the node of communication radius r is 40. Some simulation parameters are as follows: $\vartheta = 4$, $\delta = 5$, $\sigma_1 = 3$, $\sigma_2 = 2$.

A. Energy and network lifetime

Fig. 4 is a screenshot of the simulation results after implementing the RFL scheme. As shown in Fig. 4, the simulation results demonstrate the successful implementation of the RFL strategy proposed in this paper. To allow for a sufficient simulation comparison, we implemented the phantom route and a protocol that only constructs one fog near the source, which represents the near source interference policy (NSIP), such as the cloud-based protocol [7]. Fig. 5 shows the energy consumption under the phantom route, NSIP and RFL. The simulation setup is as follows: First, randomly select 200 source nodes, and execute the following three instructions for each node: (A) The source collects data according to the phantom route; (B) Construct only one fog around the source; and (C) Construct ϑ=4 fogs according to the RFL. After 10 rounds of data collection for one source node, change to another source and repeat the same process until all 200 nodes have been routed 10 times according to these three policies; finally, calculate the energy consumption. The results are shown in Fig. 5; under the phantom route, the energy consumption near the sink is extremely high, whereas it is low away from the sink, which shows that the energy efficiency is not high for regions with a significant amount of energy remaining. However, in the RFL scheme, the energy consumption away from the sink is high as well; in addition, this will not affect the network lifetime because the energy consumption of the far away regions is not higher than the hotspots. In this manner, there is only one route to the sink through the hotspots near the sink, and we construct as many interference routes as possible by fully using the remaining energy in regions away from the sink. In the NSIP scheme, the energy consumption in the non-hotspot area is higher than that in the phantom route, but there is still considerable energy remaining compared with our RFL scheme; the energy consumption is balanced when $\vartheta = 4$. The results in Fig. 5 indicate that although our scheme constructs many interference fogs and branch routes, it has improved the energy efficiency greatly without affecting network lifetime.

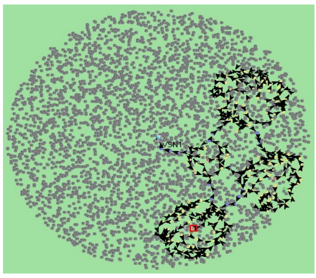

Fig. 4. Simulation screenshot in the RFL scheme

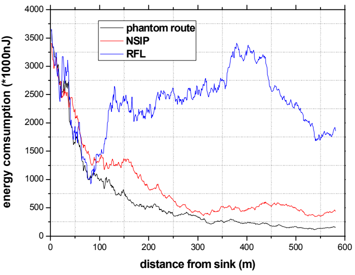

Fig. 5. Energy consumption under different source privacy schemes

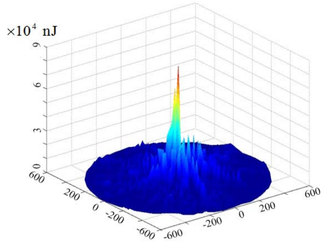

Fig. 6. Energy consumption under the phantom route

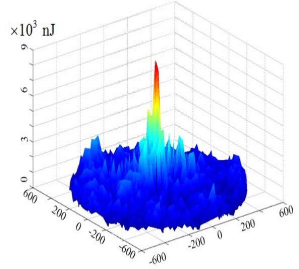

Fig. 7. Energy consumption under the loop fog route with one fog

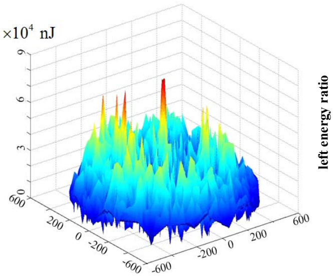

Fig. 8. Energy consumption under the loop fog route

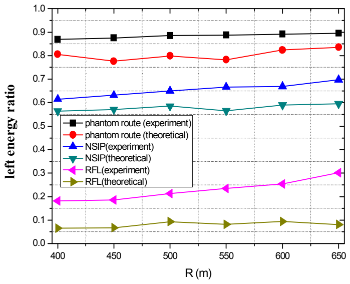

Fig. 9. Energy remaining under different policies

Figs. 6, 7 and 8 show the three-dimensional map of energy consumption under the phantom route, NSIP and RFL, respectively, with $\vartheta = 4$. As shown in Fig. 6, under the phantom route, the energy consumption near the sink is high, and there is significant energy remaining in regions away from the sink (the left energy ratio is supposed to be 90% in [9]). Under NSIP, the energy consumption in the non-hotspot area increases, but there is still considerable energy remaining. Under the RFL with $\vartheta = 4$, the energy consumption is balanced in different regions, and thus, the network energy is fully used with a relatively high energy efficiency.

Fig. 9 compares the energy remaining among RFL, NSIP and the phantom route. The energy consumption under RFL is nearly balanced for the entire network. The remaining energy ratio is approximately 20% when the network dies, which is relatively low. Under phantom route, because the energy consumption near the sink is much higher than other regions, the percentage of energy remaining is as high as 80% when the network dies, which is much higher. Under NSIP, the percentage of remaining energy is also as high as 60%. As seen from the simulation results, by selecting the appropriate parameters and number of fogs, the energy can be effectively utilized in RFL, which improves the energy efficiency by 4 fold compared to the other policies.

B. *Security performance*

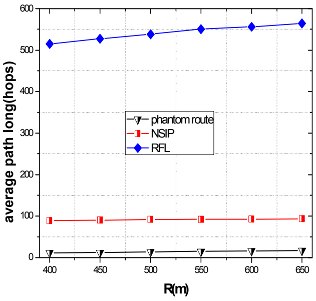

Fig. 10. Comparison of the route length under different policies with different R values

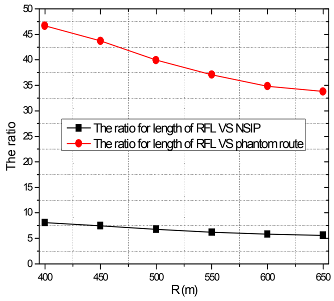

Fig. 11. Improvement as a result of RFL

Fig. 10 compares the route lengths between RFL, NSIP and the phantom route under different network scales of R. Fig. 10 illustrates that the total route length is approximately 500 hops under RFL, 100 hops under NSIP and only 10 hops under the phantom route. Fig. 11 shows the ratio for the length of RFL against NSIP and the phantom route. The total routing path length of RFL is approximately 33.79 to 46.68 fold higher than that of the phantom route and 5.59 to 8.07 fold higher than that of NSIP. The previous analysis demonstrates that a longer route length makes it more difficult for the adversary to trace to the source. Thus, the route length refers to the ability to protect against attacks, and a longer route length implies higher security. The simulation results in Figs. 10 and 11 show that the security performance of RFL is more than 5-fold better than NSIP and the phantom route.

Next, the program simulates the performance of an adversary attack under RFL, NSIP and the phantom route; the simulation scenario in Fig. 12 is similar to the one shown in Fig. 4. There is only one adversary in our simulation. The adversary waits near the sink and traces back once the data have been detected. Because the adversary cannot determine the actual direction of the source, when there are multiple source directions, the adversary will randomly select one direction to trace back. If it reaches the end of the path, which infers an unsuccessful attack, the adversary will turn back to the last branch and trace back along another branch and repeat this process until it reaches the pre-set attack hops. Then, we calculate the success probability (the statistical results). In addition, we view the adversary as highly intelligent and can return to the last un-traced route to continue the next attack without time cost; namely, the adversary can trace back with no cost in time or tracing. The simulation is repeated 200 times, and statistical results are obtained. Fig. 12 illustrates that under the phantom route, the success probability of an attack can reach 100% when the number of hops traced is less than 20; under NSIP, it can be 100% when 50 hops are traced, whereas under RFL, the probability of a successful attack is less than 50% when the number of hops traced reaches 100, which shows that RFL improves the ability to protect against attacks by more than 8 fold.

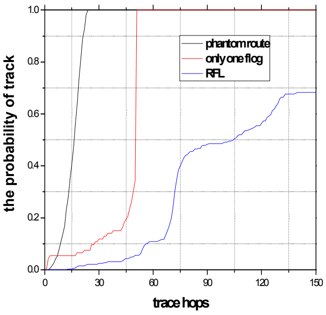

Fig. 12. Success probability with different trace hops

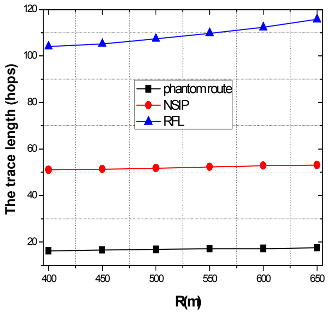

Fig. 13. Trace hops under a success probability of 50% for different R

Fig. 13 shows the traced hops for a success probability of 50% under different values of R. Two conclusions can be drawn from Fig. 13. (1) Under different R, in the RFL scheme, the number of hops traced is 5 fold greater when the success probability reaches 50%, which demonstrates that RFL has a stronger ability to protect against attacks than other policies. (2) As R grows, the number of traced hops increases rapidly, whereas in other policies, it increases slowly, which shows that a larger network scale yields a higher security performance in RFL, and RFL clearly has stronger scalability.

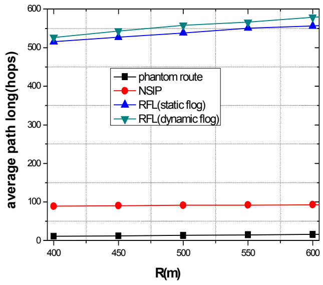

Fig. 14. Average trace length under static and dynamic fogs

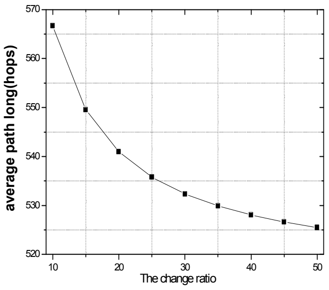

Fig. 15. Average trace length under different change ratios

Another feature that makes RFL superior to existing privacy-preserving policies against traceback attacks is that in RFL, multi-fogs are constructed regardless of whether the source is moving, and fogs are dynamically changing and moving. Therefore, when the adversary traces back, it always traces to the un-traced routes, which is equivalent to an increase in the average trace length; thus, it is more difficult for the adversary to trace to the source. Fig. 14 shows the average trace length in the situation in which one fog is revoked and a new fog is constructed once data collection is processed for 30 rounds. As shown in Fig. 14, the RFL with the dynamic scheme has a longer trace length than other policies, which indicates a better performance by RFL. Fig. 15

shows the relationship between the average route length and fog changing rate. With data collection ranging from every 50 rounds to every 10 rounds, as the fog area changes more rapidly, the average trace length becomes longer, indicating an improved security performance. However, because of the energy consumption and routing rebuilding costs, dynamic changing requires trade-off optimization between security and costs. Because other policies do not have this feature, their privacy-preserving ability is relatively weak.

VI. Conclusion

In this paper, we have proposed a redundant fogs loop (RFL) scheme for preserving source-location privacy and optimizing energy utilization which maximizes both the network security and lifetime. The RFL scheme constructs multi-fogs by fully using extra energy in non-hotspot regions. Each fog creates multiple fake packets around the source node, which provides local source-location privacy. Besides fogs are connected by routing loops and thus it offers global-level source-location privacy. At the same time, RFL enhances the security by dynamically constructing and revoking fogs, which improves energy efficiency by more than 5 times, and also maximize network security without decreasing the lifetime of WSNs.

Acknowledgment

This work is supported in part by JSPS KAKENHI Grant Number 26730056, 15K15976, and JSPS A3 Foresight Program, the National Natural Science Foundation of China (61379110), the National Basic Research Program of China (973 Program) (2014CB046305).

References

[1] Y. Liu, A. Liu, Z. Chen, Analysis and Improvement of Send-and-Wait Automatic Repeat-reQuest protocols for Wireless Sensor Networks, Wireless Personal Communications, vol. 81, no. 3, pp. 923-959, 2015.

[2] L. Jiang, A. Liu, Y. Hu, et al. Lifetime Maximization through Dynamic Ring-Based Routing Scheme for Correlated Data Collecting in WSNs. Computers & Electrical Engineering, vol. 41, pp. 191-215, 2015.

[3] N. Li, N. Zhang, S.K. Das, B. Thuraisingham, Privacy preservation in wireless sensor networks: A state-of-the-art survey, Ad Hoc Networks, vol. 7, no. 8, pp. 1501-1514, 2009.

[4] Y. Li, J. Ren, J. Wu. Quantitative Measurement and Design of Source-Location Privacy Schemes for Wireless Sensor Networks. IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 7, pp. 1302-1311, 2012.

[5] S.B. He, J.M. Chen, Y.X. Sun, et al. On Optimal Information Capture by Energy-Constrained Mobile Sensors, IEEE Transactions on Vehicular Technology, vol. 59, no. 5, pp. 2472-2484, 2010.

[6] P. Kamat, Y. Zhang, W. Trappe. Enhancing source-location privacy in sensor network routing. In: Proceedings of 25th IEEE Int. Conf. Distrib. Comput. Syst., Columbus, OH, USA, November, 2005, 599–608.

[7] M.M.E.A. Mahmoud, X. Shen. A Cloud-Based Scheme for Protecting Source-Location Privacy against Hotspot-Locating Attack in Wireless Sensor Networks, IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 10, 1805-1818, 2012.

[8] OMNet++ Network Simulation Framework, http://www.omnetpp.org/. 2013.

[9] A. Jhumka, M. Leeke, S. Shrestha, On the Use of Fake Sources for Source Location Privacy: Trade-Offs Between Energy and Privacy, The Computer Journal, vol. 54, no. 6, pp. 860-874, 2011.

[10] K. Pongaliur, X. Li, Maintaining source privacy under eavesdropping and node compromise attacks, In: Proceedings of Shanghai, IEEE INFOCOM 2011,1656-1664.

[11] C. Ozturk, Y. Zhang, W. Trappe, Source-location privacy in energy constrained sensor network routing, In: Proceedings of SASN '04, (New York, NY, USA), 88–93, ACM, 2004.

[12] J. Long, A. Liu, M. Dong, et al. An energy-efficient and sink-location privacy enhanced scheme for WSNs through ring based routing. Journal of Parallel and Distributed Computing, vol. 81, pp. 81: 47-65, 2015.

[13] J. Long, M. Dong, K. Ota, et al. Achieving source location privacy and network lifetime maximization through tree-based diversionary routing in wireless sensor networks. IEEE Access, vol. 2, pp. 633-651, 2014.

[14] H. Wang, B. Sheng, Q. Li. Privacy-aware routing in sensor networks, Computer Networks, vol. 53, no. 9, pp. 1512-1529, 2009.

[15] K. Bicakci, H. Gultekin, B. Tavli, et al. Maximizing lifetime of event-unobservable wireless sensor networks, Computer Standards and Interfaces, vol. 33, no. 4, pp. 401-410, 2011.

[16] K. Bicakci, I.E. Bagci, B. Tavli. Lifetime bounds of wireless sensor networks preserving perfect sink unobservability, IEEE Communications Letters, vol. 15, no. 2, pp. 205-207, 2011.

[17] M. Shao, Y. Yang, S. Zhu, G.H. Cao, Towards statistically strong source anonymity for sensor networks, Phoenix, AZ, IEEE INFOCOM 2008, 51 – 55.

[18] K. Mehta, D. Liu, and M. Wright, Location privacy in sensor networks against a global eavesdropper, In: Proceedings of IEEE International Conference on Network Protocols (ICNP), Oct. 2007,314-323.

[19] Y. Yang, M. Shao, S. Zhu, B. Urgaonkar, G.Cao,Towards event source unobservability with minimum network traffic, ACM Conference on Wireless Network Security,77-88,2008.

[20] Y. Ouyang, Z. Le, G. Chen, et al. Entrapping adversaries for source protection in sensor networks,Buffalo-Niagara Falls, NY, WoWMoM 2006, 23-34, 2006.

[21] A. Liu, X. Jin, G. Cui, Z. Chen, Deployment guidelines for achieving maximum lifetime and avoiding energy holes in sensor network. Information Sciences, 2013; 230, 197-226.

[22] A. Liu, D. Zhang, P. Zhang, et al. On mitigating hotspots to maximize network lifetime in multi-hop wireless sensor network with guaranteed transport delay and reliability. Peer-to-Peer Networking and Applications, vol. 7, no. 3, pp. 255-273, 2014.