

RESEARCH ARTICLE

PSSPR: A Source Location Privacy Protection Scheme Based on Sector Phantom Routing in WSNs

Yuling Chen^{*1,4} | Jing Sun¹ | Yixian Yang² | Tao Li¹ | Xinxin Niu² | Huiyu Zhou³

¹State Key Laboratory of Public Big Data, College of Computer Science and Technology, Guizhou University, Guiyang, China

²School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing, China

³School of Informatics, University of Leicester, England, United Kingdom

⁴Guangxi Key Laboratory of Cryptography and Information Security

Correspondence

Yuling Chen, State Key Laboratory of Public Big Data, College of Computer Science and Technology, Guizhou University, Guiyang 550025, China.

Email: ylchen3@gzu.edu.cn

Abstract

Source location privacy (SLP) protection is an emerging research topic in wireless sensor networks (WSNs). Because the source location represents the valuable information of the target being monitored and tracked, it is of great practical significance to achieve a high degree of privacy of the source location. Although many studies based on phantom nodes have alleviated the protection of source location privacy to some extent. It is urgent to solve the problems such as complicate the ac path between nodes, improve the centralized distribution of Phantom nodes near the source nodes and reduce the network communication overhead. In this paper, PSSPR routing is proposed as a visible approach to address SLP issues. We use the coordinates of the center node V to divide sector-domain, which act as important role in generating a new phantom nodes. The phantom nodes perform specified routing policies to ensure that they can choose various locations. In addition, the directed random route can ensure that data packets avoid the visible range when they move to the sink node hop by hop. Thus, the source location is protected. Theoretical analysis and simulation experiments show that this protocol achieves higher security of source node location with less communication overhead.

KEYWORDS:

wireless sensor network, source location, center node, phantom routing, privacy protection

1 | INTRODUCTION

The Internet of Things (IoT) refers to real time collection of any objects or processes that need to be monitored, connected, and interacted through various information sensors. As a result of the rapid development of Internet of Things (IoT), wireless sensor networks (WSNs)¹⁻⁴ are vital components of the IoT has been applied in various domains. Unlike wired networks, WSNs are flexible to adapt to complex application scenarios, such as target detection⁵⁻⁷, military defense^{8,9}, medical assistance^{10,11}, etc. Sensor nodes in WSNs communicate with each other via wireless devices. These sensor nodes are used to sense and collect valuable information in the network and forward it to the sink node via multiple hops. Finally, the sink node stores and processes the collected information. However, due to the unattended and openness of WSNs, anyone with a relevant wireless receiver can detect and intercept messages among sensor nodes. The adversary may apply illegal means to communicate with powerful workstations or information sources and potential security issues will inevitably occur. Therefore, the security of the network has become a critical issue for the WSNs. Considering the practical significance of the source location, we focus on the SLP protection in this paper.

1.1 | Related works

Recently, there has been an increased focus on location privacy and many strategies attracted attentions as approaches to prevent adversaries from performing a backtracking strategy to obtain the source location. Ozturk et al.¹² first introduces his concept. Location privacy as a contextual-oriented problem has been widely applied in various domains, such as blockchain technology^{13,14}, smart cities, Intelligent technology and so on. Location privacy covers the source location privacy and the sink location privacy. This paper focus on the issue of Source location privacy (SLP) protection. Spachos et al.¹⁵ proposes a dynamic routing scheme (ADRS) based on the Angle. However, in this scheme, each hop can only randomly select the relay node to transmit packets within the fixed Angle range, which limits the protection of the source position to a certain extent. For the above reason, Liu Ya et al.¹⁶ utilized variable Angle and proposes a dynamic routing scheme (VADRS) to protect the source location privacy, which improves the safety performance of ADRS by selecting the optimal Angle for data transmission at each hop. Zhang Jiang nan et al.¹⁷ proposes a single virtual loop routing SVCRM protocol based on fake packets, and optimized it to put forward MVCRM protocol, which complicated the adversary's tracking path and extended the security time.

Considering a more powerful adversary, Wang et al.¹⁸ proposes an angle-based source location privacy protection protocol (PRLA), in which the concept of visible area was first proposed and defined a path through the visible area during phantom routing, called a failure path. Credit routing¹⁹ and RAPFPR protocol²⁰ both ensure the source location privacy, but they do not consider the influence of visible area. The PRLA protocol calculates the forwarding probability by the offset angle of the sensor nodes to reduce the possibility of routing through the visible area. In view of the shortcomings of PRLA scheme, Chen Juan et al.²¹ proposes two algorithms using phantom nodes to protect the source location privacy. By using limited flooding, the nodes away from the source were selected as phantom nodes, which significantly increased the location diversity of phantom nodes in Source Location Privacy Preservation Protocol in Wireless Sensor Network Using Source-Based Restricted Flooding (PUSBRF). Compared with the PUSBRF protocol, EPUSBRF avoids the visible area in the routing process and extends the network security time. Kong Xiangxue et al. proposes a source location privacy protection routing protocol based on random virtual ring (PRVR)²², which extended the routing path to the virtualization range of the source node, making it difficult for adversaries to implement efficient reverse tracking. Unlike the active attacks²³⁻²⁵, the adversaries in this scheme can only launch local passive attacks.

In addition, Kamat et al.²⁶ and Kang et al.²⁷ use directed random walk to protect the source location privacy. In their scheme, The phantom sources are far from the source node. According to the hops from a certain node to Sink, its neighbor nodes are classified into child and parent node sets respectively. When the forwarding node transmits the packets, it randomly selects a receiving node from the parent or child node set to send the packet to the Sink node. However, the phantom nodes obtained by the directed routing method will gather in some fixed areas, which cannot achieve the purpose of geographical diversity of the phantom nodes and complex transmission routes.

1.2 | Motivation and contributions

When sensor nodes are communicating with each other, with an appropriate wireless devices, a person can monitor the communication signals between pairs of nodes in the wireless sensor network. In spite of encryption techniques that protect the communication content exchanged between two sensor nodes, the adversaries mostly use powerful equipment or illegal means for locating the information sources. Therefore, many researchers have focused on source location privacy (SLP) protection in recent years. What's more, in the process of routing packets how to avoid visible area, prolong the safe time and reduce the adversary's detection probability are the current research focus. As a result, this paper proposes a source location privacy protection policy based on sector phantom routing in WSNs. In addition, the WSNs consist of many low-power wireless sensor nodes, it is a key consideration to improve the location security of source nodes while taking into account the performance of network nodes.

In this paper, we focus on SLP protection and propose a source location privacy protection scheme based on sector domain phantom routing in WSNs (PSSPR), which has a good performance in security. The main contributions of this paper are as follows:

- We propose a phantom routing based on sector domain. We use the coordinates of the central node V to select phantom nodes and determine the candidate phantom node area, which address defects of insufficient diversity of phantom node positions. And through directed routing to ensure that the selected phantom nodes are distributed in an area far from the source node.
- Due to the generated phantom node is located between the source node and the base station, our prouosed scheme uses a smaller communication overhead to achieve higher source location security.
- This scheme avoids the visible area and extends the path length of the adversary's backtracking data packet.

1.3 | Roadmap

The remainder of this paper is organized as follows. The network model and the adversary models are presented in section 2. Section 3 gives an overview of our proposed scheme. Section 4 evaluates the performance of our proposed scheme, whilst the security analysis is outlined in section 5. Finally, we conclude the paper and describe planned future studies in section 6.

2 | PRELIMINARIES

2.1 | Network model

In this paper, a typical Panda-Hunter model is used to study SLP protection. As shown in Fig.1. Pandas are assumed to inhabit a monitoring area with a large number of randomly and uniformly distributed sensor nodes. Once detecting a panda, the sensor node becomes the source and regularly reports the monitored message to the sink node in a multi-hop manner. The hunter can acquire the immediate sender node's location in the way of backtrace packets by analyzing the transmission signal. The operation is repeated until the hunter reaches the source node and stops. When the hunter enters the visible area of the source location, the source location is exposed. For the network, we make the following assumptions: 1) The sink node is regarded as the final destination of all data packets and it remains in the entire network center. The location of all sensor nodes remains unchanged after deployment. 2) Any two sensor nodes in the network communicate through one or multi-hops mode. 3) All sensor nodes in the network are randomly and uniformly deployed in the sensed area. In addition, all sensor nodes have the same characteristics, which means that they have the same computing ability, initial energy, and cache memory.

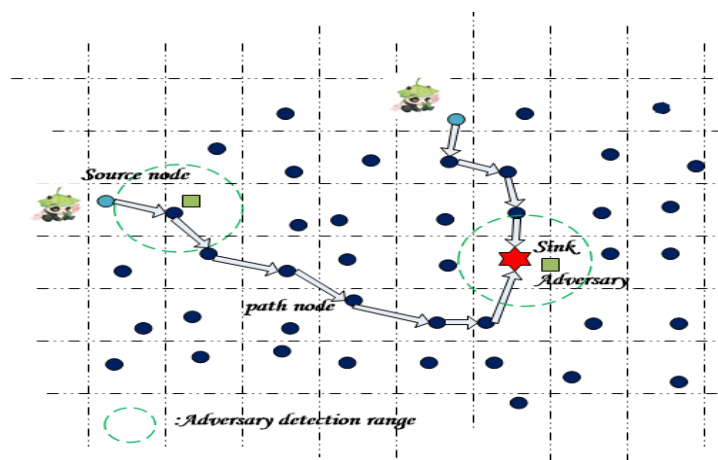


FIGURE 1 The Panda-Hunter model

2.2 | Attack model

In a typical SLP protection scenario, we assume a patient attacker mode with a small visibility of the network. The attacker can only eavesdrop any message for packets in sensor networks and cannot interfere with the normal operations of the sensors network. The data monitored by the source is regularly transmitted to the sink destination node according to a certain routing protocol. Therefore, under normal circumstances, if the attacker initially locates at the sink node, it can be guaranteed to successfully capture all packets. Driven by huge profits, the adversary is an illegal attacker that is equipped with powerful devices to eavesdrop any message from the source. When the attacker overhears a new message, it will measure the angle of arrival of the signal and the received signal strength to identify the immediate sender node, it will not have any functional impact on the network. Then, the adversary performs passive attack by starting back tracing the packet route by moving to the previous node towards the source until it reaches the source node. Algorithm 1 explains the strategy of the adversary.

2.3 | Security assumption

In the entire network, we assume that the sink node is absolutely secure. The communication between nodes adopts secure encryption and the adversary cannot obtain the location of the source through destroying the sink or intercepting data packets.

3 | PSSPR

Once the assets are detected, the nodes around the area will become source nodes and continuously monitor assets' activities and locations. The source periodically generates sensed data packets and forwards them to the sink node via a multi-hops routing manner. Before sending the packet, the source will first judge whether the distance $D_{S,B}$ between itself and the sink is within the communication radius r . If $D_{S,B}$ is less than or equal to r , the source will directly send the received packet to the sink. Otherwise, the source node transmits packets using the routing method we proposed from the current node to the sink node. In this section, we will introduce details of our proposed PSSPR scheme. The proposed PSSPR is based on our noticed source location privacy based on sector domain routing. It consists of four phases: network initialization, phantom routing based on sector domain, same-hop routing and variable Angle routing. An overview of the PSSPR scheme is shown in Fig.2.

3.1 | Network initialization

After the deployment of the wireless sensor nodes is completed, all sensor nodes have their own unique identity identification information ID of the whole network. In the network initialization phase, the sink establishes a horizontal Cartesian coordinate system with itself as the center and then broadcasts a *Sink-Msg* {ID, HopCount, Infor} message to all sensor nodes in flood mode. *ID* denotes sending node and *HopCount* denotes the hop count from the sink which initial value is set to zero. When a sensor node receives the *Sink-Msg* {ID, HopCount, Infor} message, it compares the current *HopCount* with the original *HopCount*. If the current *HopCount* is smaller than the original *HopCount*, the node updates *HopCount*, otherwise, the node discards it. Each node only records the minimum *HopCount*. *Infor* represents the coordinate information of the sensor nodes in the wireless network. If a node receives *Sink-Msg* for the first time, it records the hop count, upgrades the value of $\text{HopCount} = \text{HopCount} + 1$, and transmits the beacon message to its neighbor nodes. For each node that receives *Sink-Msg* information, the ID, Infor, and HopCount of the forwarding node should be stored in its neighbor table. After the sink completes the whole network broadcast in flood mode, each node records the coordinate information of the sink, itself and its neighbor nodes record and store the minimum HopCount from the sink. Table 1 lists the main notations used in this study.

TABLE 1 Summary of notations

Symbols	Meaning
S	Source node
Sink	Base station
P	Phantom node
$D_{S,B}$	The distance in hops between sink and source node
$D_{S,P}$	The distance in hops between phantom and source node
V	Center node
P_i	Pseudo phantom node
$A_{\lambda_i}, A''_{\lambda_i}$	Intermediate node
A''_{λ_i}	$A_{\lambda_i}, A''_{\lambda_i}$ are symmetric about node V
h_m	Same-hop routing hop count
r	eavesdropping radius of adversary; Communication radius
r_0	Radius of visible area
$N_{i,j}$	The distance in hops between node i and node j

Algorithm 1 Patient Adversary Algorithm

- 1: Adversary's initial position = Sink's position ;
- 2: When a packet is received at the sink;
- 3: Adversary's location = Immediate sender node's location;
- 4: **while** (Adversary's location != Source's location) **do**
- 5: Adversary's location = Immediate sender node's location;
- 6: **end while**
- 7: Adversary's location = Source's location;

3.2 | Phantom routing based on sector domain

The location information of the phantom node P is determined at this stage. In the network initialization phase, the source node obtains the coordinate information of itself and its list of neighbor nodes. Before forwarding the data packets, the source node first carries out R_{\max} hops directed flood routing. Hence, each node within the R_{\max} hops receives the beacon message of the source $S(S_X, S_Y)$, and feedback a response message $S-Mes$ to the source node. The $S-Mes$ contains the coordinate information of each node within R_{\max} hops from the source. Steps are as follows:

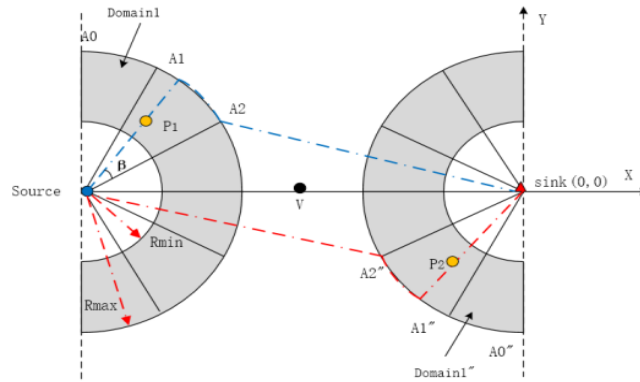


FIGURE 2 Selection of phantom nodes in sector domain

A) Before the source sends the data packets, it takes the sink node as the center to create a rectangular coordinate system. And the line between the source node and the sink node is the x-axis. Subsequently, the source node regards the central node between the sink node and itself as the central node $V(V_X, V_Y)$.

B) Establish the y-axis through the convergent node and perpendicular to the x-axis. On the left side of the y axis, the source chooses a closed semicircular ring area with itself as the center, the maximum distance R_{\max} as the outer radius, and the minimum distance R_{\min} as the inner radius, which is called the phantom area P_{Area1} . The phantom area P_{Area1} and the phantom area P_{Area2} are symmetrical about the central node $V(V_X, V_Y)$. P_{Area1} and P_{Area2} form the candidate domain.

C) Divide P_{Area1} into ω sector areas evenly. ω is an even number, then the angle θ of each sector area is $\frac{\pi}{\omega}$. These sector areas are designated as Domain1, Domain2, . . . , Domain ω .

D) Before transmitting packets, the source randomly selects a sector Domain λ_i , $\lambda_i \in [1, \omega]$ and λ_i randomly distributed. The angle $\beta \in [(\lambda_i - 1)\theta, \lambda_i\theta]$ of the line between the phantom node and the source and the line between the source node and the intermediate node is randomly distributed

The source randomly selects a node in Domain λ_i ($\lambda_i = 1, 2, \dots, \mu$) as the pseudo-phantom node P_1 during the k th ($k \geq 1$) packet transmission, and calculates the pseudo-phantom node P_2 according to the coordinate information of the central node V . P_1 and P_2 constitute a set of pseudo-phantom nodes.

Algorithm 2 From the source to the phantom nodes

-
- 1: Establish coordinate system and divide candidate domains into sectors;
 - 2: Randomly choose Domain $_{\lambda_i}$ ($\lambda_i = 1, 2 \dots \mu$) from candidate domains;
 - 3: Randomly choose P_i node as the phantom node P from candidate domains;
 - 4: Calculate angle β ;
 - 5: **if** $P_x \leq V_x$ **then**
 - 6: send the packet through h hops ($h \in [R_{min}, R_{max}]$) directed routing to reach the phantom node P , save $N_{S,P} = h$;
 - 7: Continue directed routing away from the source until reaches $N_{S,i} = R_{max}$;
 - 8: The position of A_{λ_i} is calculated from the coordinates of the phantom node P ;
 - 9: **else**
 - 10: Rolls back to the variable Angle routing path phase;
 - 11: Perform the same-hop routing phase;
 - 12: Refer to the Sink-Msg, perform h -hops directed routing to the sink and determine the position of the node P in this process;
 - 13: **end if**
-

3.3 | Same-hop routing

After the R_{max} hops directed routing or variable included Angle routing phase is completed, the forwarding node uses the angle β to calculate the number of same-hop routes h_m . As exhibited in formula (1):

$$h_m = \frac{\beta}{180^\circ} R_{max} \quad (1)$$

If the phantom node is selected from the left side of the center node, the angle β is calculated as:

$$\beta = \arccos \left(\frac{d_{S,A\lambda_i}^2 + d_{S,P}^2 - d_{A\lambda_i,P}^2}{2d_{S,A\lambda_i} \times d_{S,P}} \right) \quad (2)$$

Otherwise:

$$\beta = \arccos \left(\frac{d_{Sink,A\lambda_i''}^2 + d_{Sink,P}^2 - d_{A\lambda_i'',P}^2}{2d_{Sink,A\lambda_i''} \times d_{Sink,P}} \right) \quad (3)$$

Then the forwarding node performs h_m hops in the direction close to the x-axis, and transfer the data packet to the A_{λ_i} node or A_{λ_i}'' node. Such node is called the intermediate node in our proposed scheme. The rules for selecting the next-hop node in the forwarding process are as follows: the phantom node P selects a node $D_{S,B}$ hops away from the source or a node with the same hops as the sink in its neighbor table as the next hop forwarding node. After the h_m hop, the data packet is transmitted to the A_{λ_i} node or the A_{λ_i}'' node and then stops. This routing process is called the same hop routing.

The same-hop routing can be analyzed by two participants. On the one hand, for an attacker, performing multiple simultaneous jumps will make the attacker mistakenly think that he is trapped in a circular trap to confuse the attacker. On the other hand, for the source node, the same-hop routing extends the length of the attacker's reverse backtracking path and increases the security time.

3.4 | Variable angle routing path

We use the vector inner product method to calculate the angle φ_i . $\langle V_{i,M_i} \rangle$ and $\langle V_{S,Sink} \rangle$ respectively represent the vector from node i to M_i and the vector from source to the sink. The calculation formula of Angle is shown in formula (4).

$$\varphi_i = \arccos \left[\frac{\langle V_{i,M_i}, V_{S,Sink} \rangle}{\|V_{i,M_i}\| \times \|V_{S,Sink}\|} \right] \quad (4)$$

The forwarding node selects the node with the smallest angle φ_i as the next hop node. This stage is completed when the forwarding node is the sink or an intermediate node. The specific content is explained in Algorithm 3. In this stage, the forwarding direction of the next hop node is determined by the angle, so that the data packet is selected to the sink node in an approximately straight line. Under the condition that the source location is safe, this solution can control network communication consumption to a certain extent. In this stage, data packets are transmitted to sink nodes in an approximate straight line to save communication overhead

Algorithm 3 Variable Angle Routing Path

```

1: Utilize  $\langle V_{i,M_i} \rangle$  and  $\langle V_{S,Sink} \rangle$  to calculate the angle  $\varphi_i$ ;
2: Select the neighbor node  $i$  with the smallest included Angle  $\varphi_i$  as the receiving node of the packet;
3: while the forwarding node is common node do
4:   Go back to step one;
5: end while
6: if forwarding node is the sink then
7:   Stop forwarding packets;
8: else
9:   End the variable routing path;
10: end if

```

4 | SECURITY ANALYSIS

4.1 | Safety analysis

In this section, we evaluate the safety of our proposed PSSPR protocol through theoretical analysis from four metrics. Random directed path, the distance from the phantom node to the source $D_{S,P}$, the number of random phantom nodes N and the failure path are the four influencing metrics. Namely, by comparing the three factors with the the HBDRW scheme and the PUBRF scheme, we come to the conclusion that the influence on safety of PSSPR is stronger. $D_{S,P}$ refers to the hop count between the source node and the phantom node. If the filtered phantom nodes are clustered near the source, adversaries can capture the source node by tracing back a short routing path. Thus, the method does not protect the privacy of the source location. And N affects the diversity of the path from the source to the phantom nodes. Therefore, the greater the number of phantom nodes, the more effective the security of the source location can be protected. Eventually, The transmission path of the source forwarding packets to the sink passes through the path of the visible area, which is called the failure path. Avoiding the failure path is equivalent to extending the effective path.

4.1.1 | Random directed path

The path of a data packet from the source node to the phantom source node after a directed random h -hop is defined as a random directed path. When the attacker reaches a certain phantom source node through reverse tracking of the data packet, the source location privacy protection protocol generates a random path. The more paths there are, the more difficult it is for an adversary to trace the true source node. In a large-scale sensor network with uniformly distributed nodes, the phantom source nodes in the source location privacy protection protocol proposed in this paper are evenly distributed in a closed ring area centered on S , the maximum distance R_{max} is the outer radius, and the minimum distance R_{min} is the inner radius. Compared with the HBDRW protocol, the PUBRF protocol increases the random directed path generated by the HBDRW protocol by $\xi = 1 - 4\gamma/2\pi = 1 - 2 \arccos \frac{h-1}{h} / \pi$. In addition, compared with the PUBRF protocol, the random directed path generated by the PSSBR protocol has increased: $\xi = 1 - h / (R_{min} + R_{min} + 1 + \dots + R_{max})$.

TABLE 2 Percentage of random directeg path

h	R_{min}	R_{max}	HBDRW/PUBRF	PUBRF/PSSPR
5	4	6	40.97	33.33
10	8	12	28.71	20.00
15	12	18	23.38	14.29
20	16	24	20.22	11.11
25	22	28	18.07	14.29
30	26	32	16.48	14.78

With the increase of h , the average number of random directed paths increased by the PSSPR protocol will increase. Table 2 shows that the PSSPR protocol can significantly increase the number of random directed paths, thereby effectively improving the security of the source location privacy.

4.1.2 | Failure path

The failure path refers to the transmission path of the phantom node forwarding data packets to the sink node passes through the path of the visible area. The SLP protocol in WSN based on locational angle prove the probability of a failure path in a sensor network with uniformly distributed nodes is:

$$(\arcsin(r_0/H) + \arcsin(r_0/h))/\pi \quad (5)$$

It shows that the larger the r_0 , the smaller the H and h , the greater the probability of the transmission path passing through the visible area. In our proposed scheme, on the left side of the y axis, the source chooses a closed semicircular ring area with itself as the center, the maximum distance R_{max} as the outer radius, and the minimum distance R_{min} as the inner radius, which is called the phantom area P_{Area1} . The phantom area P_{Area1} and the phantom area P_{Area2} are symmetrical about the central node $V(V_x, V_y)$. We know that a certain node u in the visible area, $HopCount_{u,s} \leq r_0$. So the phantom node selected by this method completely bypasses the visible area in the routing path of transmitting data packets, avoiding the generation of failure path.

4.1.3 | The distance from the phantom nodes to the source $D_{S,P}$

- HBDRW protocol selects the receiving node according to the hop count between the neighbor nodes of the forwarding node and the sink node. Phantom nodes are mainly distributed on a circle with S as the center, radius h , and arc 4γ where γ is $\arccos \frac{h-1}{h}$. For comparison, let $h = R_{min} + h_x$. For the HBDRW protocol, after random h hops, the average distance from the phantom nodes to the source is:

$$\overline{D_{S,P} HBDRW} = r(R_{min} + h_x) (h \in [0, R_{min} - R_{max}]) \quad (6)$$

- Due to the source node performs h hop directed random routing, the selected phantom nodes lies in the circumference with S as the center and $h = R_{min} + h_x$ as the radius. Correspondingly, the average distance from the phantom node to the source node in the PUSBRF protocol is:

$$\overline{D_{S,P} PUSBRF} = r(R_{min} + h_x) (h \in [0, R_{min} - R_{max}]) \quad (7)$$

- In order to achieve a high level of phantom node distribution, it is evident from Fig. 2. that The phantom nodes are distributed in the shaded area. The distance between the phantom node and the source node is between R_{min} and R_{max} , so the average distance from the phantom nodes to the source is expressed as:

$$\overline{D_{S,P} PSSPR} = \frac{R_{min} + R_{max}}{4} + \int_0^{\frac{\pi}{2}} \frac{\sqrt{H^2 + (R_{min} + R_{max})^2 - 2(R_{min} + R_{max})H \cos \alpha}}{\pi/4} \quad (8)$$

As shown in Table 3, in order to ensure the same simulation conditions of the three protocols, the corresponding values of R_{min} and R_{max} are set for different directed routing hop count.

TABLE 3 Values of R_{max} and R_{min}

h	R_{min}	R_{max}	$D_{S,P} PSSPR$
5	4	6	31.40
10	8	12	32.03
15	12	18	33.25
20	16	24	34.64
25	22	28	36.10
30	26	32	37.42

4.1.4 | The number of random phantom nodes N

- Since the number of phantom nodes affects the random path diversity from source nodes to phantom nodes, the number of phantom nodes plays an important role in safety performance. The more phantom nodes there are, the less likely an adversary is to locate the source by backward and hop-by-hop tracking packets. In PUSBRF protocol, the phantom nodes are concentrated on the circle with S as the center of the circle and the radius is $(R_{min} + h_x)$ hops. And the number of phantom nodes is denoted as N_{PUSBRF} .

$$N_{PUSBRF} = 2\pi(R_{min} + h_x) (h \in [0, R_{max} - R_{min}]) \quad (9)$$

- In HBDRW protocol, the phantom nodes are distributed on the arc with S as the center of the circle, a radius of $(R_{min} + h_x)$ hops, and a center angle of 4γ where γ is $(R_{min} + h_x - 1)/(R_{min} + h_x)$. The number of phantom nodes is designated as N_{HBDRW} .

$$N_{HBDRW} = 4\gamma \times (R_{min} + h_x) = 4 \arccos \frac{R_{min} + h_x - 1}{R_{min} + h_x} \times (R_{min} + h_x) (h \in [0, R_{min} - R_{max}]) \quad (10)$$

- In our proposed scheme, the phantom node is concentrated in the ring area with S as the center of the circle and $(R \in [0, R_{min} - R_{max}])$ as the radius. We denote the number of phantom nodes in this scheme is N_{PSSPR} . Then, the number of phantom nodes in PSSPR scheme is approximate:

$$N_{PSSPR} = 2\pi(R_{min} + h_x) \times \left(\frac{h_x + h_x + 1 + \dots + R_{max} - R_{min}}{h_x} \right) \quad (11)$$

The number of phantom nodes produced by the three protocols is compared as shown in Table4.

TABLE 4 Comparison of the number of phantom nodes

h	R_{min}	R_{max}	N_{HBDRW}	N_{PUSBRF}	N_{PSSPR}
5	4	6	12.87	31.42	94.24
10	8	12	18.04	62.83	282.74
15	12	18	22.03	94.25	565.48
20	16	24	25.40	125.66	942.47
25	22	28	28.38	157.08	942.47
30	26	32	31.07	188.50	706.86

4.2 | Communication overhead analysis

In the PUSBRF strategy, the source needs to perform $(R_{min} + h_x)$ hops flooding, which is expressed as the $(R_{min} + h_x)$ hops finite flooding hop value from the source to the phantom nodes. Then the data packet is forwarded from the phantom node to the sink in the shortest route, so the average communication overhead of the PUSBRF protocol is:

$$\overline{E}_{PUSBRF} = R_{min} + h_x + \int_0^\pi \frac{\sqrt{H^2 + (R_{min} + h_x)^2 - 2(R_{min} + h_x)H \cos \alpha}}{\pi} d\alpha \quad (12)$$

\overline{E}_{PUSBRF} is the average communication overhead of the PUSBRF protocol; $(R_{min} + h_x)$ is the hop length from the source node S to the phantom node P; H is the hop distance between the source and the sink. Phantom nodes generated by HBDRW are distributed on the circumference of a circle with S as the center, $(R_{min} + h_x)$ as the radius, and 4γ as the radian, where γ is $(R_{min} + h_x - 1)/(R_{min} + h_x)$ then the average communication overhead of HBDRW protocol is:

$$\begin{aligned} \overline{E}_{HBDRW} = R_{min} + h_x + & \int_0^\gamma \frac{\sqrt{H^2 + (R_{min} + h_x)^2 - 2(R_{min} + h_x)H \cos \alpha}}{2\gamma} d\alpha + \\ & \int_\pi^{\pi+\gamma} \frac{\sqrt{H^2 + (R_{min} + h_x)^2 - 2(R_{min} + h_x)H \cos \alpha}}{2\gamma} d\alpha \end{aligned} \quad (13)$$

For this protocol, the abscissa of the selected phantom node is smaller than the central node as an example. The data packet arrives at the phantom node after $(R_{\min} + h_x)$ hops from the source node. After the phantom node is selected, it continues to the direction away from the source node to reach R_{\max} , and then the same Jump, and finally in the variable angle routing path stage. Among them, the communication cost of the sector domain phantom routing is $\overline{E}_1 = R_{\max}$, the communication cost of the same hop number routing is:

$$\overline{E}_2 = \overline{h_m} = \int_0^{\pi} \frac{\beta\pi}{180^\circ} R_{\max} d\theta \quad (14)$$

And the communication cost of the variable angle routing path stage is:

$$\overline{E}_3 = 2[H - R_{\max} + \sum_{i=1}^{\frac{\mu}{2}-1} \sqrt{H^2 + R_{\max}^2 - 2R_{\max}H \cos(i\theta)}] / \mu \quad (15)$$

Then, the average communication overhead of PSSPR scheme is:

$$\overline{E}_{PSSPR} = R_{\max} + h_m + 2(H - R_{\max} + \sum_{i=1}^{\frac{\mu}{2}-1} \sqrt{H^2 + R_{\max}^2 - 2R_{\max}H \cos(i\theta)}) / \mu \quad (16)$$

5 | SIMULATION

5.1 | Safety performance

In this section, we simulate the performance of our proposed PSSPR protocol in MATLAB platform. The simulation environment and parameter configuration are as follows. There are 10000 sensor nodes Random and evenly distributed over the monitored area of 6000m \times 6000m. And the communication radius r of each node is 100m and is one third of the the visible area. The eavesdropping radius of the adversary is equal to the communication radius r . In addition, the sink node is deployed in the center of the network. In the experiment, $\omega = 6$, and the result is shown in Fig.3.

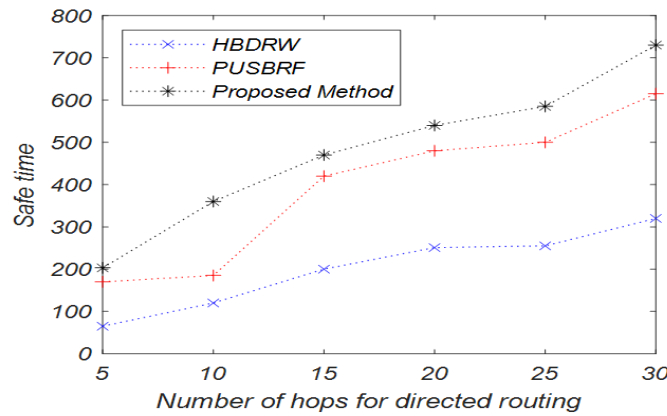


FIGURE 3 Safe time VS directed routing

The safety time is defined as the total number of data packets sent by the source before being captured by the adversary. Fig.3 illustrates the results of 50 simulations of PSSPR scheme under the condition that hop count H is fixed at 60. In the figure, the safe time of this protocol is extended as the directed routing h increases. It can be seen that the safety performance of PSSPR scheme is the best. Therefore, the adversary spends more time trying to capture the source node. As shown in Fig.3, the safety time of the PSSPR scheme is 2.38 times that of the HBDRW protocol and 1.22 times that of the PUSBRF protocol. As the increases of h , the average distance between the source node and the phantom node is prolonged, which increases the diversity of the directed random path and enhances the complexity of the transmission path at the same time. The complexity of the routing path makes it more difficult for the adversary to track backwards. This results in the longest safe time for the PSSPR protocol among the three protocols.

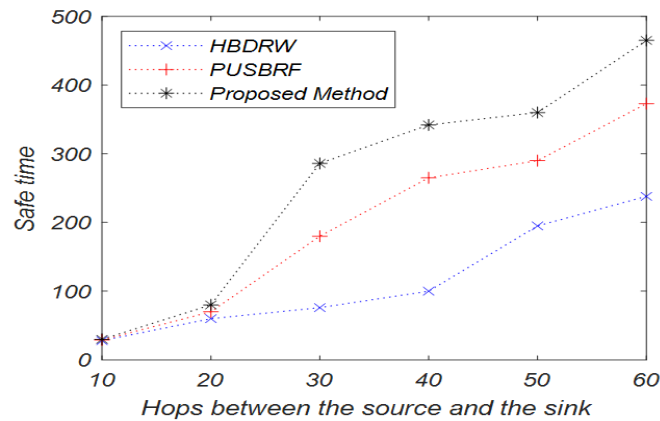


FIGURE 4 Safe time VS distance to sink

Fig.4 displays the results of 50 simulations of PSSPR scheme under the condition that random directed routing h is 15. This result is attributed to the increase of H . As the distance to the sink node increases, the routing paths between the source node and the sink node become longer on average. In addition, the length of the transmission path increases. Therefore, the adversary spends more time on the backtracking process due to the longer routing path. Meanwhile, our propose PSSPR protocol completely avoids the visible area in the process of transmitting packets, and thus, the safe time increases. As shown in Fig.4, the safety time of PSSPR scheme is 2.42 times of HBDRW scheme and 1.29 times of PUSBRF scheme respectively. So our proused scheme performs clearly better than the other schemes.

5.2 | communication overhead

To balance the communication overhead and maximize the safety time, we conduct a theoretical analysis of the communication consumption. In this experiment, the communication overhead refers to the average hops count by transmitting one data packet from the source node to the sink node. we can analyze the simulation results of communication consumption based on four parts: the whole network flood consumption, sector-based routing path consumption, the same-hop routing consumption and the variable angle routing path consumption. Since the three protocols have the overhead of the whole network flood, our paper only considers the latter three parts. As demonstrated in Fig. 5, When the hop count H between the source node and the sink node is fixed as 60, under the condition of different hops of directed routing, the average communication overhead result obtained by transmitting 50 packets from the source node. As the random directed routing hops increases, the communication overhead of the three routing protocols become higher on average. This occurs because the larger the h , the longer the transmission path between the source node and the phantom node is. In addition, this stage does not greatly reduce hop count for the data packets to be routed to the sink, thus increasing the communication Overhead. It can be seen from Fig. 5 that the communication overhead of this scheme changes slightly with directed routing hops. Fig. 5 shows that when $h=30$, the communication overhead of the three routing protocols differs the most. The communication overhead of PSSPR scheme is 7.73% and 10.04% less than that of HBDRW and PUSBRF respectively.

In Fig. 6, we shows the changes in the average communication overhead for different distances from the source nodes to the sink node. The simulation of the safety time is performed in terms of the hop count of random directed routing is fixed at $h=20$ and the source node sends 50 data packets to the sink node. It is observed that communication overhead of the three routing protocols exhibits a decreasing trend with the increasing of H . This is because with the increase of H , more sensor nodes participate in transmitting packets, which results in an growth in transmission path. Therefore, the communication overhead increases. We also observe that the communication overhead of the three routing protocols is not much different under the condition of the same value of h . From an overall point of view, our proposed PSSPR scheme contributes only slightly to communication overhead, which is 9.14% less than HBDRW and 11.57% less than PUSBRF.

6 | CONCLUSION

The application of wireless sensor network in the field of monitoring needs to design an effective SLP routing scheme. The existing SLP protection schemes mainly protect the SLP by changing or increasing the routing path length, which greatly increases the communication overhead. In this paper, we presented the PSSPR scheme for source location privacy, which decreases communication consumption in the interest of maintaining

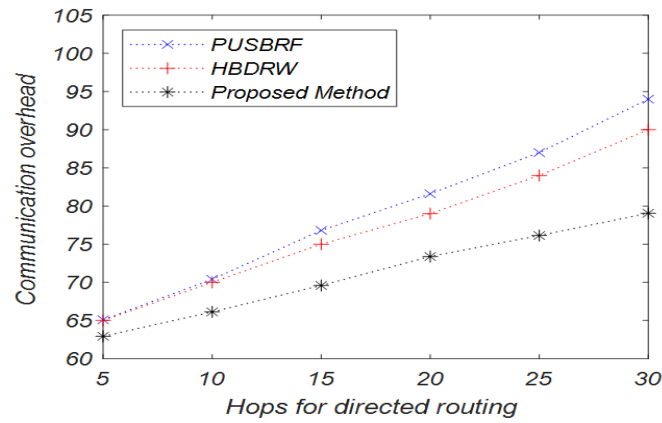


FIGURE 5 Communication overhead VS directed routing hops

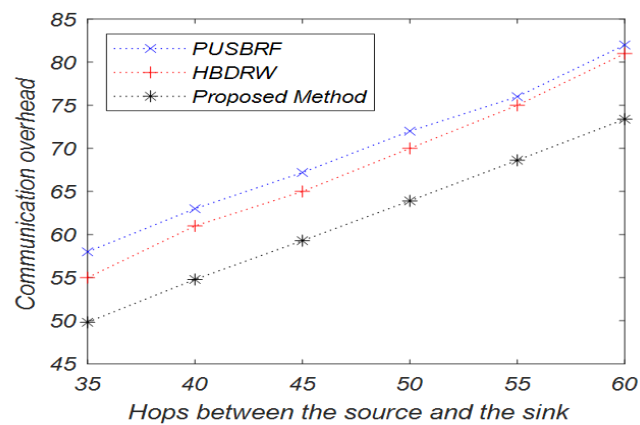


FIGURE 6 Communication overhead VS distance to sink

safety time. The scheme provides a dynamic route generation process with randomly selected phantom nodes. In the PSSPR, initially, the source node and sink node jointly establish several virtual sectors domains. Then, the phantom nodes are selected from the different sectors and the packet is transmitted through different phantom nodes to construct the dispersed routing paths to achieve high security performance of the source. Along with the PSSPR protocol completely avoids the failure path and generates enormous phantom nodes while improving the geographic diversity of the phantom nodes. Therefore, this scheme adds random directed routes, implements multiple paths and reduces overlapping paths. The overall comparative analysis shows that PSSPR protocol proves itself to be an effective scheme in the considered performance elements.

In our future work, we plan to investigate the times each node participates in routing. Then, we will comprehensively consider communication overhead and balance the energy consumption of the entire network, so that the network has a longer security time.

ACKNOWLEDGMENT

This study is supported by Foundation of National Natural Science Foundation of China (Grant Number: 61962009); Major Scientific and Technological Special Project of Guizhou Province (20183001); Science and Technology Support Plan of Guizhou Province ([2020] 2Y011); Foundation of Guizhou Provincial Key Laboratory of Public Big Data (2018BDKFJJ013); Foundation of Guangxi Key Laboratory of Cryptography and Information Security (GCIS202118).

References

1. Peng H, Chen H, Zhang X, Fan Y, Li C, Li D. Location privacy preservation in wireless sensor networks. *J Software*. 2015;26(3):617–639.
2. Qian P, Wu M. Overview of the research and methods of privacy protection in the Internet of Things. *Comput Appl Res*. 2013;30(1):13-20.
3. Wang Y, Yang G, Li T, Li F, Tian Y, Yu X. Belief and fairness: a secure two-party protocol toward the view of entropy for IoT devices. *J Network Comput Appl*. 2020;161(1):102641.
4. Li F, Wang D, Wang Y, et al. Wireless communications and mobile computing blockchain-based trust management in distributed internet of things. *Wireless Commun Mobile Comput*. 2020;2020(5):1-12.
5. Tellez M, El-Tawab S, Heydari H. Improving the security of wireless sensor networks in an IoT environmental monitoring system. In: Gardner D, eds. *IEEE Systems and Information Engineering Design Symposium (SIEDS)*. IEEE; 2016:72-77.
6. Lazarescu M. Design of a WSN platform for long-term environmental monitoring for IoT applications. *IEEE J Emerging Sel Top Circuits Syst*. 2013;3(1):45-54.
7. Yang J, Zhou J, Lv Z, Wei W, Song H. A real-time monitoring system of industry carbon monoxide based on wireless sensor networks. *Sensors*. 2015;15(11):29535-29546.
8. Ismail B, Salvatore Domenic M, Ravi S. A survey of intrusion detection systems in wireless sensor networks. *IEEE Commun Surv Tutor*. 2014;16(1):266-282.
9. Furtak J, Zieliński Z, Chudzikiewicz J. Security techniques for the WSN link layer within military IoT. In: *World Forum on Internet of Things (WF-IoT)*. 3rd ed. IEEE; 2016:233-238.
10. Dinesh R, Marimuthu RA. Survey about WSN and IoT based health care applications and ADPLL contribution for health care systems. In: *International Conference on Awareness Science and Technology (iCAST)*. 10th ed. IEEE; 2019:1-8.
11. Adame T, Bel A, Carreras A, Melia-Segui J, Oliver M, Pous R. CUIDATS: an RFID-WSN hybrid monitoring system for smart health care environments. *Future Gener Comput Syst*. 2018;78:602-615.
12. 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*. New York, NY, USA: Association for Computing Machinery; 2004:88-93.
13. Wang Y, Wang Y, Wang Z, Yang G, Yu X. Research cooperations of blockchain: toward the view of complexity network. *J Ambient Intell Hum Comput*. 2020;11:1-14. <https://doi.org/10.1007/s12652-020-02596-6>
14. Li P, Li K, Wang Y, et al. A systematic mapping study for blockchain based on complex network. *Concurrency Comput: Pract Exper*. 2020:e5712. <https://doi.org/10.1002/cpe.5712>
15. Spachos P, Toumpakaris D, Hatzinakos D. Angle-based dynamic routing scheme for source location privacy in wireless sensor networks. In: *Vehicular Technology Conference (VTC Spring)*. 79th ed. IEEE; 2014:1-5.
16. Liu Y, Xu Y, Song L. Variable-angle based dynamic routing scheme for source-location privacy in wireless sensor network. *Comput Appl Res*. 2018;35(1):257-260.
17. Jiangnan Z, Chunliang C. Research on the privacy protection scheme of source node location in wireless sensor network. *J Sens Technol*. 2016;29(9):1405-1409.
18. Wang WP, Chen L, Wang JX. A source-location privacy protocol in WSN based on locational angle. In: *International Conference on Communications*. IEEE; 2008:1630-1634.
19. Lu Z, Wen Y. Credit routing for source-location privacy protection in wireless sensor networks. In: *International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012)*. 9th ed. IEEE; 2012:164-172.
20. Zhao Z, Liu Y, Zhang F, Zhou J, Zhang P. Research on source location privacy routing based on angle and probability in wireless sensor networks. *J Shandong Univ (Sci Ed)*. 2013;48(9):1-9.

21. Chen J, Fang B, Yin L, Su S. Source location privacy protection protocol based on limited flooding of source nodes in sensor networks. *Chin J Comput.* 2010(9):1736-1747
22. Kong X, Yuan S, Chen M. Source location privacy protection routing protocol based on virtual ring. *SensMicrosyst.* 2018(1):66-69.
23. Li T, Wang Z, Yang G, Cui Y, Chen Y, Yu X. Semi-selfish mining based on hidden Markov decision process. *Int J Intell Syst.* 2021;36(7):3596-3612. <https://doi.org/10.1002/int.22428>
24. Li T, Chen Y, Wang Y, et al. Rational protocols and attacks in blockchain system. *Secur Commun Networks.* 2020. <https://doi.org/10.1155/2020/8839047>
25. Wang Y, Yang G, Bracciali A, et al. Incentive compatible and anti-compounding of wealth in proof-of-stake. *Inf Sci.* 2020;530:85-94.
26. Kamat P, Zhang Y, Trappe W, Ozturk C. Enhancing source-location privacy in sensor network routing. In: *International Conference on Distributed Computing Systems (ICDCS'05)*. 25th ed. IEEE; 2005:599-608.
27. Kang L. Protecting location privacy in large-scale wireless sensor networks. In: *International Conference on Communications*. IEEE; 2009:1-6.