ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **24**, Number 1, 2021, 37–57

Sink Location Privacy Protection Algorithm Based on Virtual Circle in WSNs

Leqiang BAI , Jizhe YU , and Keyan CAO

Information & Control Engineering Faculty, Shenyang Jianzhu University, China
E-mails: baileqiang@sjzu.edu.cn, yujizhe2019@outlook.com, caokeyan@gmail.com

Abstract. For the protection of sink location privacy in WSNs, sink location privacy protection algorithm based on virtual circle is put forward. A virtual circle is constructed with the sink as the center, and mathematical models are established according to the location of the source node to determine the expected phantom source node, which provides direction for selecting the first stage phantom source node. Taking the first stage phantom source node as the starting point, a second stage one is randomly selected on the circle in a clockwise direction. The selection of the two phantom source nodes provides the diversity of packet transmission path. At the same time, the two phantom source nodes can inject fake packets. The first one transmits fake packets to some random destinations and some fake sinks set in the network, inducing an adversary to trace away from the sink. The second one transmits fake packets along the circle to some random destinations in the network, which further increases the difficulty for an adversary to capture the sink and achieves the purpose of protecting sink location privacy. The analysis of setting to deploy fake sinks shows that the location of the fake sinks satisfies randomicity and dispersity, by randomly selecting the angle on the external region of the virtual circle and using the normal distribution mathematical model to determine the distance between the fake sink and the sink. The simulation results show that the algorithm can effectively induce adversaries to deviate from the real path and improve the safe time.

Key-words: *Wireless sensor networks* (WSN), sink location privacy, virtual circle, phantom source node, fake sink.

1. Introduction

With the rapid development of sensor technology and embedded technology, the microsensor has the ability of sensing, computing and communication. A network composed of a large number of sensors is called wireless sensor networks, which has attracted extensive attention in the fields of environmental monitoring, disaster warning, traffic management, animal protection, emergency rescue, national defense and military affairs. WSNs integrates technologies such as

sensor technology, embedded technology and communications technology, and it can transmit information of monitored objects to the sink in data packet in the multi-hop communication mode. However, it is easy for the shared transmission channel to be tracked and monitored by adversaries, which leads to serious privacy problems. WSNs privacy is mainly divided into content privacy and location privacy [1-3]. Concerning content privacy, the adversary by monitoring the link layer controls the sensor nodes in the network and steals or tampers the packet content transmitted by the sensor nodes, mainly using technologies of communication control and network anonymity to complete tasks of data query, access and control while hiding sensitive information. Location privacy mainly includes location privacy of source node and sink node. The adversary eavesdrops on the signals of the transmission data packet in the network, determining the location of source node or sink node by tracking source of the data packet hop by hop. Sink location privacy protection is very important in WSNs and has become a hot field of research today, since the sink [4] is the only node that receives data from all sensors in the sensor network.

Jing Deng *et al.* first raised the problem of sink location privacy in WSNs, and put forward the two tolerance strategies to protect sink location privacy against traffic analysis adversaries [5]. On this basis, Jing Deng *et al.* proposed four strategies against traffic analysis adversaries in order to protect the sink location privacy, by protecting the traffic from the aggregation node to the sink node [6]. Ying Jian *et al.* put forward a novel scheme of sink *location privacy routing* (LPR), dividing neighboring nodes into a closer list and a further list. When a real packet is transmitted to a node, the node randomly selects a further neighbor as the next hop with probability, which increases the length of the packet transmission path, and diversifies the transmission path, and improves the safe time of the sink [7]. To solve the problem of traditional random routing, Liu *et al.* proposed the Sink joint Ray Routing for data scheme. Sink node moves in a ring-shaped, non-random routing path, and packets are transmitted to mobile sink along the Ray Routing, making it difficult for adversaries to locate the sink [8]. Lin Yao *et al.* hide the location of the sink by injecting fake packets and fake sinks to induce the adversary to deviate from the real path, thus prolonging the time for adversaries to capture the sink [9]. Jun Long *et al.* developed a *ring based routing* (RBR) strategy, in which data packets are transmitted to the nearest routing ring via the shortest path and then to other routing rings via the routing line, increasing the difficulty for adversaries to locate the exact sink location [10]. To target global adversaries, Chai *et al.* put forward a KAS scheme, which proposed k-anonymity technology for the first time to protect the sink location privacy. KAS created k specific nodes to confuse the adversaries and hide the real sink node location, and designed the GAQO and APQO algorithms. Results show that the two algorithms can well hide the real location of sink node [11]. Di Ying *et al.* worked out the *anti-traffic analysis* (ATA) strategy, in which each node generated fake packets to balance network traffic and confuse powerful adversaries [12]. Changing network traffic pattern is a common way to protect sink location privacy. Ruben *et al.* devised a HISP-NC scheme, which injected many dummy packets to resist traffic analysis adversary. The node in this algorithm sent two packets, one to the node closer to the sink, and the other further away from the sink [13]. Some schemes can protect the location privacy of both the source node and the sink node. Lee *et al.* put forward a CRBS scheme, which transmited real and dummy packets to neighboring nodes. In order to resist the local adversary, this algorithm could eliminate traffic correlations and balance traffic in the sensor network [14]. Abuzneid *et al.* used false names to protect source nodes or sink location privacy, so adversaries could not find source or sink locations [15]. Honglong Chen *et al.* put forward a *location privacy algorithm* (LPSS) to protect

the location of the source node and sink node. In this algorithm, each node in the random routing strategy (FRW) transmited data packets to the neighbor whose hop number was less than its own hop number to the sink, but the algorithm could not provide location privacy protection. However, the algorithm of BT, DBT and ZBT could effectively improve the location privacy of source and sink, the latter two put forward on the basis of BT [16]. In order to resist the global attack, Juan Chen *et al.* formulated a strategy based on packet sending rate adjustment (SRA), which balanced the traffic of the whole network, thus hiding the location of the real sink by controlling the transmission rate of packets [17]. Nikolaos Baroutis *et al.* allocated traffic by injecting fake packets, making it difficult for adversaries to find the location of sink nodes [18]. Jian Wang *et al.* developed a sink location privacy protection against direction attack adversaries, who could measure the arrival angle of data packets and infer its transmission direction. Along the transmission direction, the adversary could capture the sink hop by hop [19]. On the schemes of source node location privacy protection, KONG *et al.* put forward a PRVR scheme, which avoided the increase of failure path and improved the randomness and diversity of routing path through directional random step, virtual loop routing and shortest path [20]. Zhou Chuang *et al.* put forward a RDPRPP protocol, which gave out the locations of phantom nodes with the random directed way and selected phantom nodes randomly within the locations. Packets were transmitted from the source node to the phantom source node, then from the phantom source node to the sink. There were great progresses of selected distance, costs and safe time [21]. Jia Zongpu *et al.* proposed a BRACRS scheme, which was routed through phantom routing stage, circular routing stage and shortest path routing stage. This algorithm improved not only safe time, but also source location privacy without increasing communication overhead [22]. Huang Bei-bei *et al.* devised a ABDRW scheme, which used the arctangent value of the slope of the adjacent nodes to determine the scope of the next phantom nodes on phantom path, making phantom source node far away from the source, and enhancing the source location privacy [23].

In order to protect the sink location privacy, this paper proposes a sink location privacy protection *algorithm based on virtual circle* (ABVC). Mathematical models are used to determine the first stage phantom source node with uniformly distributed positions on the virtual circle, and the first stage phantom source node randomly determined the second stage phantom source node by an angle, which provides various packet transmission routing paths. Then, the first stage phantom source node transmits fake packets to fake sinks and some destinations, and the second stage phantom source node transmits fake packets to some destinations along the circle, which diversifies the packet transmission path, making it hard for an adversary to distinguish the real path from the fake path, thereby strengthening the protection of the sink location privacy.

2. System Model

2.1. Network Model

This paper makes the following hypotheses on the network model of WSNs:

(1) The network consists of a sink node and a large number of sensor nodes.

(2) The communication range of each sensor node is r. If each sensor is within the communication range of the other, they can communicate with each other directly.

(3) The node closest in distance to the monitoring target of the network is a source node, which transmits packets to the sink node in a specific period of time.

(4) This paper only studies location privacy protection. It assumes that the adversary can

only eavesdrop the signal transmitted by the packet, but cannot steal or modify the content of the packet. Thus, the network in this study has a basic security mechanism and the contents are encrypted for privacy purposes [24].

2.2. Adversary Model

This paper assumes that the adversary has the following characteristics, which are similar to the "hunter-panda" [25-26]:

(1) The only purpose of the adversary is to capture the sink. The adversary cannot interfere with the normal functions of the network. It does not have the aggressive ability to tamper with packets, alter routing paths, and destroy sensor nodes. The adversary in this paper can track and eavesdrop, and to some extent has the ability to attack in direction.

(2) The adversary is equipped with advanced equipment. It has strong analyzing power and storage capacity which can monitor the signal of a packet and estimate the location of the sending node through the device of antenna and spectrum analyzers.

(3) The adversary can choose to follow a packet or stay at the same location for a period of time to capture or analyze more packets. Meanwhile, the movement speed of the adversary is relatively slower than the transmission speed of a packet.

(4) The hearing radius of the adversary is equal to the communication radius of the sensor.

(5) If the sink appears in the detective communication range of the adversary, it can capture the sink directly.

3. Related Definitions and Mathematical Models

3.1. Related Definitions

(1) virtual circle: As shown in Fig. 1 and Fig. 2, it is with the sink *B* as the center and *R* its radius.

(2) The expected phantom source node P_1': (i) As shown in Fig. 1, if the source node S appears outside the virtual circle and makes a tangent to the virtual circle, the coordinate of the tangency point is regarded as the expected phantom source node P_1'. (ii) As shown in Fig. 2, when source node S appears inside the virtual circle, the coordinate of the point where the straight line between the source node S and the sink B intersects with the virtual circle is regarded as the expected phantom source node P_1'.

(3) The first stage phantom source node P_1: (i) As shown in Fig. 1, when the source node S appears outside the virtual circle, it transmits a real packet to P_1' along the tangent; if a node receiving the real packet has P_1' within its communication range, it is regarded as the first stage phantom source node P_1'.

(ii) As shown in Fig. 2, when the source node S appears inside the virtual circle, it transmits a real packet to P_1' along the straight line; if a node receiving the real packet has P_1' within its communication range, it is regarded as the first stage phantom source node P_1.

(4) The second stage phantom source node P_2: As shown in Fig. 1 and Fig. 2, the phantom source node P_1 randomly generates an angle $\theta \in (0°, 360°)$, and it transmits a real packet hop by hop along the virtual circle clockwise. If the angle formed by the node receiving the real packet, the sink B and the phantom source P_1 is greater than or equal to θ, the node is regarded as the second stage phantom source node P_2.

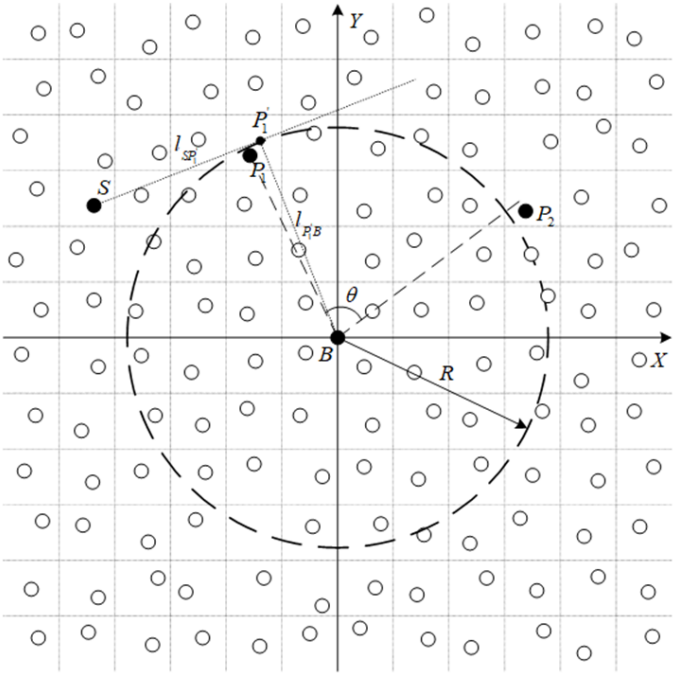

Fig. 1. Expected phantom source P_1' determined by source node outside the virtual circle.

3.2. The Mathematical Models of the Expected Phantom Source

3.3. The Mathematical Model of the Expected Phantom Source Node when the Source Node Appears outside the Virtual Circle

As shown in Fig. 1, when the sink node is at the origin of the rectangular coordinate system XOY, and if the coordinate of the source node S is (x_s, y_s), the expected phantom source node P_1' is (x_1, y_1), the slope of the straight line $l_{SP_1'}$ is k.

The standard equation of virtual circle in XOY is:

$$X^2 + Y^2 = R^2 \tag{1}$$

The equation of $l_{SP_1'}$ is:

$$Y - y_1 = k(X - x_1) \tag{2}$$

The equation of $l_{P_1'B}$ is:

$$Y = -\frac{1}{k}X \tag{3}$$

By Combining equations (2) and (3):

$$\begin{cases} x_1 = \frac{k}{k^2+1}(kx_s - y_s) \\ y_1 = -\frac{1}{k^2+1}(kx_s - y_s) \end{cases} \tag{4}$$

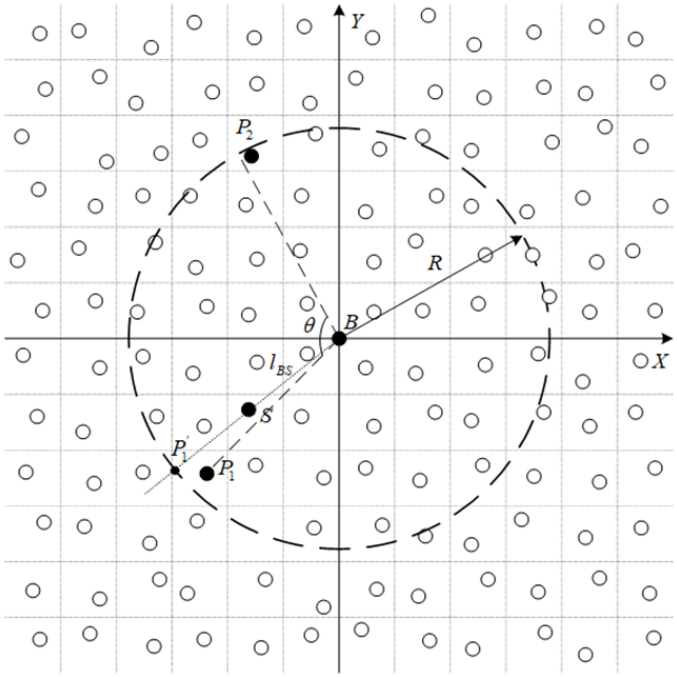

Fig. 2. Expected phantom source P_1' determined by source node inside the virtual circle.

Substituting equation (4) into equation (1), we can get the value of *"k"*:

$$k = \frac{x_s y_s \pm R\sqrt{x_s^2 + y_s^2 + R^2}}{x_S^2 - R^2} \tag{5}$$

Substituting equation (5) into equation (4), we can find P_1'.

3.4. The Mathematical Model of the Expected Phantom Source Node when the Source Node Appears inside the Virtual Circle

As shown in Fig. 2, when the sink node B is at the origin of the rectangular coordinate system XOY, and if the coordinate of the source node S is (x_s, y_s), (x_1, y_1) is the expected phantom source node P_1'.

The equation of l_{BS} is:

$$Y = \frac{y_s}{x_s} X \tag{6}$$

Combined equations (1) and (6), we can get P_1':

$$\begin{cases} x_1 = \pm\sqrt{\frac{1}{x_S^2 + y_S^2}}\, x_S R \\ y_1 = \pm\sqrt{\frac{1}{x_S^2 + y_S^2}}\, y_S R \end{cases} \tag{7}$$

4. Procedure of ABVC

The ABVC is divided into four phases: the network initialization phase; the phantom routing phase based on source node location; the phase of the phantom source node P_1 injecting fake packets, and circumferential routing based on random angle; the phase of the phantom source node P_2 injecting fake packets, and shortest path routing. In other words, the sink B broadcasts a beacon packet in the network, and each node gets relevant information to generate a virtual circle with the sink B as the center and R as the radius. The source node appearing outside the virtual circle transmits real packets to the phantom source node along the tangent line. The source node S appearing inside the virtual circle transmits real packets to the phantom source node P_1 along the straight line. Phantom source nodes P_1 randomly generate a routing angle θ in a clockwise direction; starting from the phantom source node P_1, real packets are transmitted along the virtual circle. When the angle of a node receiving the real packet, the sink B, and the phantom source node P_1 is greater than or equal to θ, the node is phantom source node P_2. Then, the phantom source node P_1 transmits fake packets to fake sinks and random destinations. Phantom source node P_2 transmits real packets to the sink along the shortest path and fake packets to destinations along the virtual circle in a clockwise direction.

4.1. Network Initialization Phase

Main tasks of network security initialization phase: acquiring nodes information in the network and establishing nodes information list, getting neighbors to establish the neighbor list. Stored in the nodes information list are the nodes ID, coordinates and the minimum number of hops Hop to the sink of the nodes. The neighbor list stores the ID of the neighbor nodes, the coordinates of neighbor nodes and the minimum number of hops $sender\_Hop$ to the sink. Any node in the network obtains its coordinates through the location algorithm. Sink B broadcasts a beacon packet $Skin\_Init$ throughout the network.

Skin_Init = {InformationType, sink_coordinate, sender_ID, sender_coordinate, hop, fakesink_coordinates}: informationType represents the message types of packets to send; sink_coordinate means the coordinates of the sink, *sender_ID* refers to the *ID* of the sending node, *sender_coordinate* represents the coordinates of the sending node, *hop* refers to the number of hops from the sending nodes to the sink, with its initial value being 0, andf *akesink_coordinates* means the coordinates of the fake sink.

If *Q* is the node receiving packets *Sink_Init* in the network, the steps of processing the packet are as follows:

Step 1: *Q* reads *Sink_Init* and judges whether it is the first time to receive the packet. If it is the first time, store *sender_ID*, *sender_coordinate* and *sender_Hop* in the neighbor list, and then proceed with Step 2; otherwise turn to Step 3.

Step 2: *Q* judges whether it is itself the sink B. If Q is B, establish a virtual circle with the radius of R, and stop transmitting packets; otherwise, store the coordinates of the B and the fake sinks, update *Hop=hop+1*, and then proceed with *Step 4.*

Step 3: *Q* checks *whethersender_ID* is in the neighbor list. If *sender_ID* is in the neighbor list, update the minimum number of hops from this neighbor to the sink *sender_Hop =hop*; otherwise store *sender_ID, sender_coordinate and sender_Hop* in the neighbor list. Q judges the sizes of *hop+1* and *Hop*. If *Hop* is larger than *hop+1(Hop* $>$ *hop+1)*, renew the value of *Hop* as equal to *hop+1(Hop=hop+1)*, and then proceed with *Step 4*; otherwise, stop transmitting packets.

Step 4: Q updates *sender_ID* in *Sink_Init* to the *ID* of Q, *sender_coordinate* to the coordinate of Q, and *hop to Hop*. Packets transmission continues.

4.2. Phantom Routing Phase based on Source Node Location

As shown in Fig. 3 and Fig. 4, the source node S transmits a real packet to the phantom source node P_1. Assuming that Q is the node receiving the real packet. The following steps are taken for the source node S and the Q to process the packet:

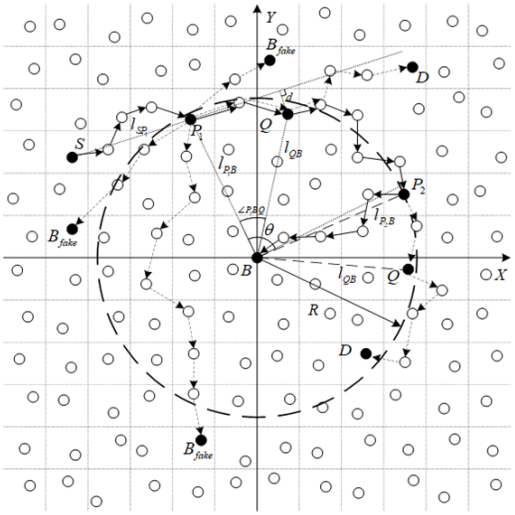

Fig. 3. An example of S outside the virtual circle transmitting packets to B.

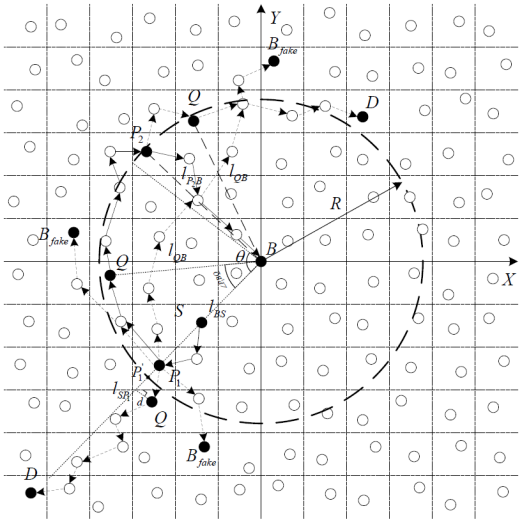

Fig. 4. An example of S inside the virtual circle transmitting packets to B.

Step 1: S calculates $P_1'(x_1, y_1)$, searches the neighbor list and transmits a real packet.

(1) When S appears outside the virtual circle, substitute equation (5) into equation (4) to calculate P_1'. A real packet is transmitted to the neighbor nodes nearest in distance to P_1'.

(2) When S appears inside the virtual circle. According to equation (7) and calculates P_1'. S transmits a real packet to the neighbor nodes nearest in distance to P_1'.

Step 2: Q judges whether P_1' exists within its communication range. If P_1' does exist, stop transmitting the real packet, and take Q as P_1; otherwise, Q searches the neighbor list, calculates the distance from each neighbor node to P_1', and transmits the real packet to the neighbor nodes nearest in distance to P_1'.

4.3. Phase of the Phantom Source Node P_1 Injecting Fake Packets, and Circumferential Routing based on Random Angle

As shown in Fig. 3 and Fig. 4, the phantom source node P_1 transmits fake packets to the fake sink B_{fake} and random destination D, while transmitting a real packet to phantom source node P_2. Assuming that Q is the node receiving the real packet, and its coordinate is (x_Q, y_Q), and the coordinate of P_1 is (x_1, y_1). The steps of the phantom source node P_1 and the node Q to process the packets are as follows:

Step 1: P_1 transmits a real packet to P_2, fake packets to D and B_{fake}.

(1) P_1 sends a fake packet of $TTL = 5$ along the straight line l_{SP_1}. The equation of l_{SP_1} is shown in equation (8).

$$Y = \frac{y_1 - y_S}{x_1 - x_S} X + \frac{x_1 y_S - x_S y_1}{x_1 - x_S} \tag{8}$$

(2) P_1 transmits fake packets to the neighbor nodes nearest in distance to B_{fake}.

(3) P_1 randomly generates an angle $\theta \in (0°, 360°)$ in the clockwise direction, and divides neighbor nodes into a clockwise list and a counterclockwise list, with line $l_{P_1 B}$ as the boundary. P_1 calculates the distance from each node in the clockwise list to the virtual circle along the clockwise list, and transmits a real packet to the neighbor nodes nearest in distance to the virtual circle.

Step 2: Q judges whether the packets received are fake ones. If fake, carry out Step 3; otherwise, do Step 4.

Step 3: Q judges whether the packets received carry TTL .

(1) If the packets carry TTL, and the TTL is $0 (TTL = 0)$, Q discards the fake packet, and is regarded as destination D; otherwise, $TTL - 1$, calculate the vertical distance d from Q to line l_{SP_1}, as shown in equation (9). Q transmits the fake packet to the neighbor node which has the minimum distance d .

$$d = \left| \frac{(y_1 - y_S) x_Q - (x_1 - x_S) y_Q + x_1 y_S - x_S y_1}{\sqrt{(y_1 - y_S)^2 + (x_1 - x_S)^2}} \right| \tag{9}$$

(2) If Q does not receive packets with TTL, it is necessary to judge if Q is B_{fake}. If it is, discard the fake packets; otherwise, transmit the fake packets to the neighbor nodes nearest in distance to the fake sinks.

Step 4: Q judges whether $\angle P_1 B Q$ is greater than or equal to θ. If $\angle P_1 B Q$ is greater than or equal to θ, Q is regarded as P_2; otherwise, Q divides the neighbor nodes into a clockwise

list and counterclockwise list with line l_{QB} as the boundary, Q calculates the distance from each neighbor node in the clockwise list to the virtual circle, transmits the real packet to the neighbor nodes nearest in distance to the virtual circle.

4.4. Phase of the Phantom Source Node P_2 Injecting Fake Packets, and Shortest Path Routing Phase

As shown in Figs. 3 and 4, the second stage phantom source node P_2 transmits a real packet to the sink B along the shortest path and a fake packet to random destination D. Assuming that Q is the node receiving real packets, the steps of the second stage phantom source node P_2 and the Q to process the packets are as follows:

Step 1: P_2 transmits a real packet to B and a fake packet to D.

(1) P_2 transmits a real packet to the neighbor nodes nearest in distance to B.

(2) P_2 divides the neighbor nodes into a clockwise list and a counterclockwise list with line l_{P_2B} as the boundary. P_2 calculates the distance from each neighbor node in the clockwise list to the virtual circle and transmits a fake packet of $TTL = 5$ to the neighbor nodes in the clockwise list nearest in distance to the virtual circle.

Step 2: Q judges whether the packet it receives is fake. If the received packet is the fake packet, then go to Step 3; otherwise, turn to Step 4.

Step 3: Q judges whether the TTL is 0. If $TTL = 0$, Q discards the fake packet, and is regarded as the *D*; otherwise, $TTL-1$, Q divides the neighbor nodes into a clockwise list and a counterclockwise list with line l_{QB} as the boundary. Q calculates the distance from each neighbor node in the clockwise list to the virtual circle, transmits the fake packet to the neighbor nodes in the clockwise list nearest in distance to the virtual circle.

Step 4: Q judges whether it is itself the sink node B. If Q is B, it stops sending the real packet; otherwise, it transmits the real packet to the neighbor nodes nearest in distance to B.

5. The Analysis of Setting to Deploy Fake Sinks

In ABVC, the deployment of fake sinks plays a vital role in confusing the adversary. In order to deploy the fake sinks randomly and dispersedly, an analysis of setting to deploy fake sinks is done, and the main parameters used to deploy the fake sinks are shown in Table I.

Table 1. Parameters for deploying fake sinks and implications

Parameter	Implication
n	The number of subregions outside the virtual circle ($n > 3$)
F_j	The $j-th$ fake sink ($1 \leqslant j \leqslant m$)
α	A random angle of $\left[0, \frac{2\pi}{n}\right]$
m	The number of fake sinks ($1 < m \leq \lfloor \frac{n}{2} \rfloor$)
T	The period of deploying fake sinks $T(T = \lfloor \frac{n}{m} \rfloor)$ means that, F_{j-1} is deployed in region S_i, and F_j should satisfy the Angle $\beta : ((i-1)+(j-1)T)\gamma \leqslant \beta_j \leqslant (i+(j-1)T)\gamma$

As shown in Fig. 5. A rectangular coordinate system XOY is established with the sink as the center of the circle, rotating counterclockwise along the x-axis. The shaded region S in the figure is divided into n equal sectors. The angle γ of each sector is $\frac{2\pi}{n}$. When deploying m fake sinks in

S during the period of T, the result of establishing fake sinks in random subregions satisfies the requirement of dispersiveness. First, the subregion S_i is randomly selected, then an angle α in $S_i(1 \leq i \leq n)$ is selected to determine the location of F_1, meanwhile, F_1 meets the angle $\beta_{j=1}$ in the XOY coordinate system. F_2 is deployed on S_{i+T}, satisfying the period T and the angle β_2, so that the location of F_1 and F_2 satisfies dispersity. Similarly, F_j is deployed on $S_{i+(j-1)T}$ and meets the angle β_j, from which we can conclude that the location of all contiguous fake sinks satisfies dispersity. Therefore, there are blank $k\,(k = T-1)$ subregions between any two sequential contiguous fake sinks. After deploying all m fake sinks during the period of T, there are continuous g empty subregions from F_m to F_1 counterclockwise. The $g\,(g \geq k)$ is shown in equation (10).

$$g = n - T\,(m-1) - 1 \tag{10}$$

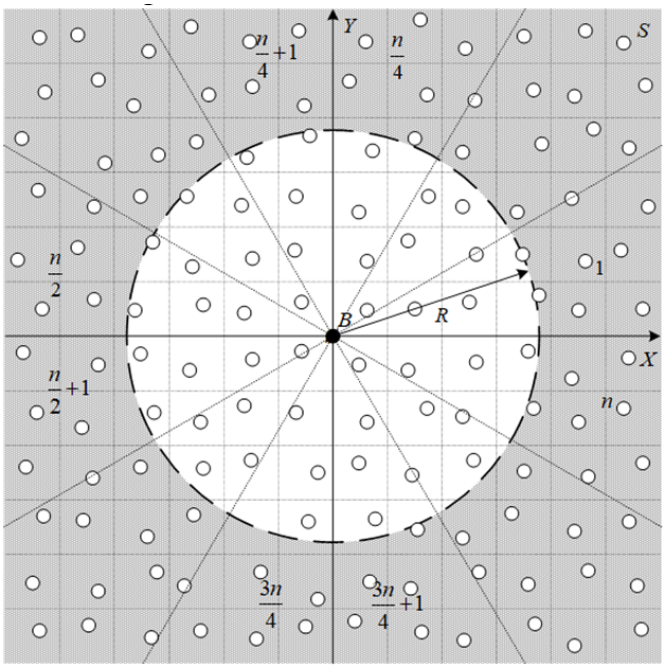

Fig. 5. The rectangular coordinate system and separate sectors.

The above analysis shows there is a minimum distance d_{min} between any two sequential contiguous fake sinks as follows:

$$d_{min} = R\sqrt{2 - 2cosk\gamma}$$

Assuming that the distance between any two sequential contiguous fake sinks is d_{rand}, then $d_{rand} \geq d_{min}$.

In order to make it difficult for the adversary to capture the sink node, the fake sinks should not be too close to the sink. However, to avoid excessive energy consumption, the fake sinks should not be too far from the sink. So, in this paper, the mathematical model of normal distribution [9],[27] is used to deploy fake sinks, and the following analysis is made.

Suppose the random number x is normally distributed with mean 0 and variance σ^2, that is $x \sim N\left(\mu, \sigma^2\right)$. When $\mu = 0, \sigma = 1, x$ is said to follow a standard normal distribution, and its distribution function $\phi(x)$ is defined as follows:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \tag{11}$$

Assuming that R_{fake} is the distance from a randomly selected fake sink to the sink:

$$R_{fake} = R(|x| + 1) \tag{12}$$

Assuming that ρ is a real number greater than 0, namely $\rho > 0$. Substitute ρ into equation (12), and we have the value of R_{fake}:

$$R_{fake} = (\rho + 1) R \tag{13}$$

Substitute ρ into equation (11), and we can get the probability of R_{fake} falling within the interval $[R, (\rho + 1) R)$ as follows:

$$P(-\rho < X \leq \rho) = \phi(\rho) - \phi(-\rho) = 2\phi(\rho) - 1 \tag{14}$$

When $\rho = 1$, according to equation (14), the probability of R_{fake} falling within the interval [R,2R) is:

$$P(-1 < X \leq 1) = 2\phi(1) - 1 = 0.68268$$

Similarly, when $\rho = 2$, the probability of R_{fake} falling within the interval [R,3R) is:

$$P(-2 < X \leq 2) = 2\phi(2) - 1 = 0.95449$$

When $\rho = 3$, the probability of R_{fake} falling within the interval [R,4R) is:

$$P(-3 < X \leq 3) = 2\phi(3) - 1 = 0.99730$$

The above analysis shows, the probability of R_{fake} falling within the interval [R,2R) is 0.68268, the probability of R_{fake} falling within the interval [R,3R) is 0.27181, the probability of R_{fake} falling within the interval [R,4R) is 0.04281. In other words, using the mathematical model of the normal distribution to deploy the fake sinks, the probability of the fake sinks appearing near the circle is much larger than that of it appearing far from the circle.

As stated in the above two analyses, based on the period T of the fake sink deployment, an angle is randomly determined outside the virtual circle for selection of a subregion for the fake sink. Then, using the mathematical model of normal distribution, the distance between the fake sink and the sink is determined, so that the fake sink can appear as close to the virtual circle as possible. In the end, the specific location of the fake sink is determined by the angle and distance.

6. Performance Evaluation

In this paper, we evaluate the performance of ABVC according to three criteria: delivery time, strength of privacy protection and energy consumption. LPR [7], MRF [19] and ABVC algorithms are simulated using the Matlab R2017b simulation platform. To realize uniform distribution of sensor nodes, the $6000 \times 6000\ m^2$ network is divided into 100×100 grids, where the 10000 sensor nodes are uniformly and randomly distributed at the center of each grid. Each sensor node appears anywhere in each grid with random disturbance $\varepsilon\left(\varepsilon \sim N\left(\mu, \sigma^2\right)\right)$. In our simulations, the sink is located at the center of the network, and source nodes are randomly selected. The source period is defined as the time between the transmissions of two packets from a source node [7].

6.1. Selection of Virtual Circle Radius

The ABVC chooses the different sizes of R, which corresponds to the different results of the algorithm of safe time and energy consumption. The safe time of the sink refers to the number of real packets received by the sink before being captured by the adversary. The energy consumption refers to the total number of hops it takes for all packets to be transmitted under a certain routing protocol. We compare the changes in R values ranging from 100m to 3000m in terms of safe time and energy consumption. Fig. 6 and Fig. 7 respectively reflect the curve of safe time and energy consumption produced by different *R* values; R400, R800, R1200, R1600, R2000, R2400 and R2800 represent 400m, 800m, 1200m, 1600m, 2000m, 2400m, 2800m respectively.

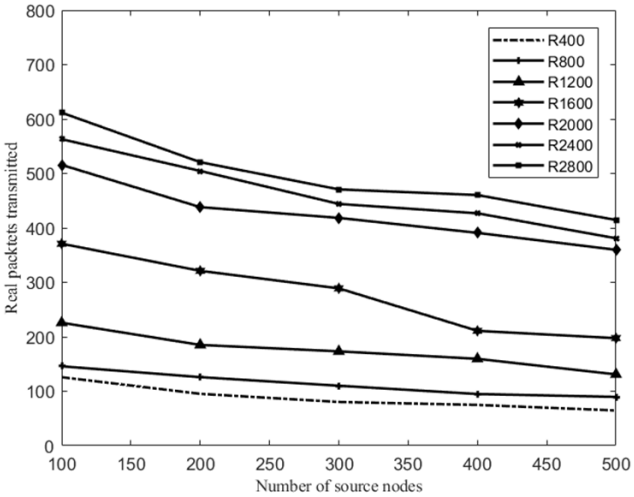

Fig. 6. Safe time provided by different values of R.

As shown in Figs. 6 and 7, when R is 2000m, the safe time is significantly improved compared with that of a smaller radius. With a larger radius, little difference in safe time is observed, and the increase of energy consumption is within an acceptable range. Therefore, the virtual circle radius of this paper is set at 2000m.

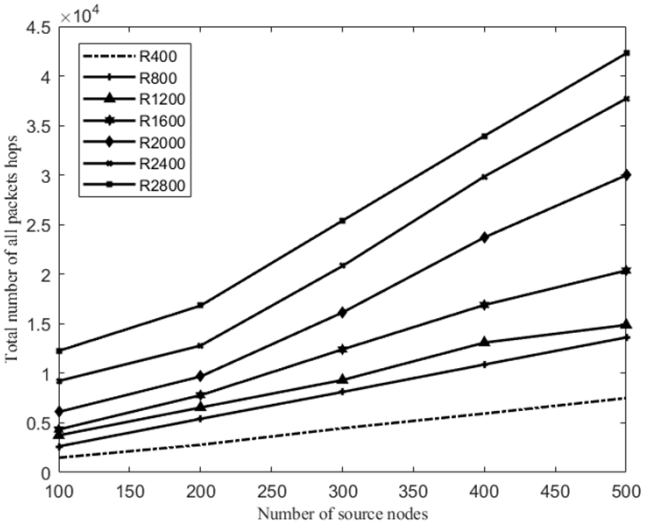

Fig. 7. Energy consumption provided by different values of R.

6.2. Delivery Time

Delivery time refers to the average number of hops it takes for a real packet to travel from the source node to the sink under a certain routing protocol. As shown in Fig. 8, with the hops from the source node to the sink increasing, the delivery time of the three algorithms increases. Since the transmission of fake packets is not considered, the transmission path of real packets in the MRF algorithm is divided into the shortest routing path and random routing path. The main transmission path of the MRF algorithm is the shortest routing path. However, the nodes on the random routing path transmit fake packets to the closer or further list in order to cover the path of real packet transmission, which is the same as the routing path of LPR algorithm. To reduce excessive energy consumption, the random routing path is short, and the packet transmission is always toward the sink. As a result, the delivery time is short. In the LPR algorithm, packets are transmitted along random routing paths to the closer or further list, and the randomness of packet transmission increases the hops. Therefore, the delivery time of this algorithm increases rapidly. In the ABVC algorithm, when the hops from the source node to the sink are less than 25, the delivery time is higher and the growth rate faster. The reason is that the radius of the virtual circle is 2000m, and the sink is about 25 to 27 hops away. Source nodes appear mostly inside the virtual circle, which has experienced the phantom routing phase based on source node location, the circumferential routing phase based on random angles and shortest path routing phase, increasing the hops of packet transmission. As a result, the delivery time of the ABVC algorithm is slightly higher than that of the LPR algorithm. When the hops between the source node and the sink are greater than 25, source nodes appear outside the virtual circle, are relatively close to the virtual circle, and there are fewer packet hops in the phantom routing phase based on source node location. So, the delivery time is shorter and the growth rate slower. As to the LPR algorithm, however, the nodes always randomly transmit packets to the closer list or further list, the transmission being always random, whether the source node is near or far away from the sink. Especially when the source nodes are far away from the sink, the length of the transmission

path obviously increases, and continuously. Therefore, after more than 25 hops, the delivery time of the LPR algorithm is larger than that of the ABVC algorithm.

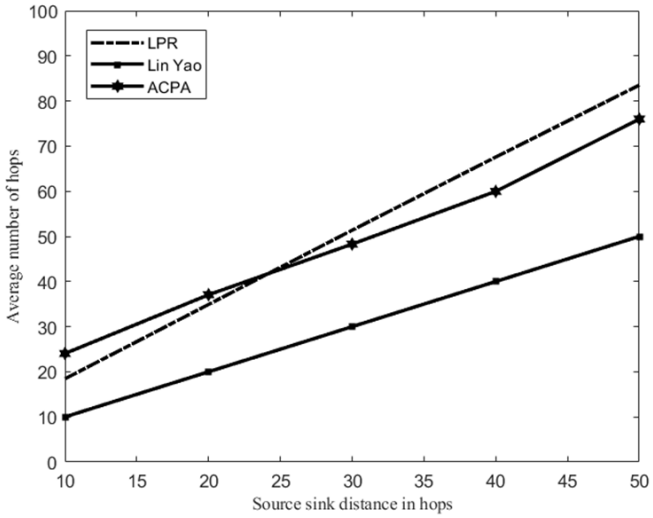

Fig. 8. Delivery time during transmission.

6.3. Strength of Privacy Protection

The safe time of the sink and the attack time of the adversary are used in evaluating the protection intensity of location privacy. The safe time of the sink refers to the number of real packets received by the sink before being captured by the adversary. The attack time of the adversary refers to the number of moving steps the adversary takes before capturing the sink.

6.3.1. Safe Time

As shown in Fig. 9, a comparison is made among the safe times of the three algorithms, without the injection of fake packet and fake sinks. In the MRF algorithm, packets are transmitted mainly along the shortest routing path, thus the shortest safe time. In the LPR algorithm, the nodes transmit packets randomly to the closer list or further list, that is, the transmission path of packets is random. Compared with the MRF algorithm, while it can improve the safe time of the sink, packets are always transmitted toward the sink, which reduces the protection of the sink location privacy, shortening the safe time of the algorithm. In the ABVC algorithm, a virtual circle is selected with high safe time and an acceptable energy consumption. The source node transmits packets to the virtual circle, applying a variety of routing mechanisms, improving the sink location privacy. So, the safe time of this algorithm is longest.

As shown in Fig. 10, as the number of source nodes increases, the safe time of all three algorithms decreases. In the LPR algorithm, the packet transmission path cannot effectively confuse the adversary without fake packet injection. While the number of source nodes increases, the packets are always transmitted toward the sink, which cannot protect the sink location privacy, so this algorithm has the shortest safe time. In the MRF algorithm, the real packet transmission

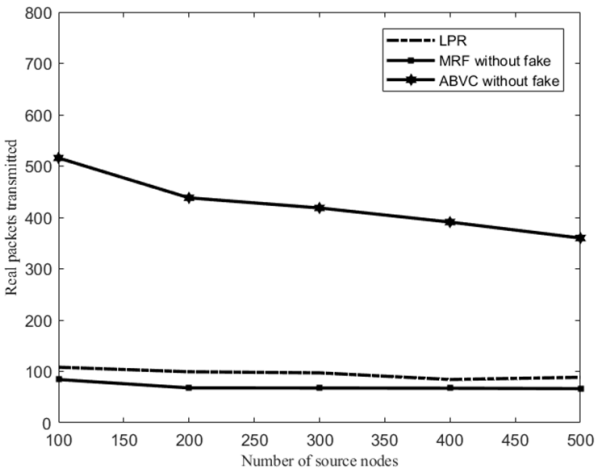

Fig. 9. Safe Time provided by different algorithms without fake packets injection and fake sinks injection.

path is mainly the shortest routing path; large numbers of fake packets are transmitted to the random destinations or fake sinks from the intersection node formed by the shortest routing path, and from nodes on the random routing path, which can protect sink location privacy. However, nodes injecting fake packets are distributed on the random routing path, which is short, and with the increase of the number of source nodes, the fake packets generated are close to the sink, which makes it less effective in protecting the sink location privacy. So the safe time of this algorithm is relatively long. In the ABVC algorithm, a virtual circle with long safe time is first selected, then the phantom source node on the virtual circle transmits fake packets to the random destination, and the phantom source node P_1 transmits fake packets to fake sinks deployed outside the virtual circle and closer to the virtual circle. In this way, the sink location privacy is well protected and the safe time of the sink is improved. This algorithm has a longer safe time.

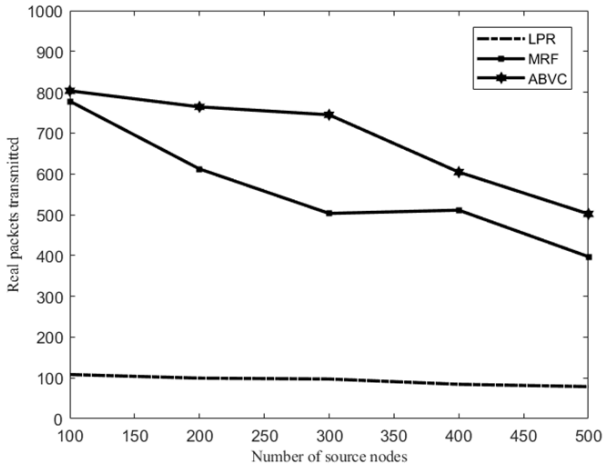

Fig. 10. Safe Time provided by different algorithms with fake packets injection and fake sinks injection.

6.3.2. Attack Time

As shown in Fig. 11, with the increase of the number of source nodes, the attack time of the three algorithms varies. In this paper, the adversary can appear anywhere. Concerning the LPR algorithm, with no fake packet and fake sinks injection, it is only possible to delay the adversary's capturing the sink by increasing the length of packet transmission path. However, with the increase of the number of source nodes, the overall transmission trend of packets is toward the sink, which makes it easy for the adversary to capture the sink. So this algorithm is not effective enough in increasing the attack time, and the attack time is the shortest. Both the MRF algorithm and the ABVC algorithm inject fake packets and fake sinks, which can increase the adversary's difficulty in capturing the sink. In the MRF algorithm, the adversary can easily trace the intersection node, because packets are transmitted to the intersection node along the shortest path at the beginning. In the random routing stage, the fake packet path makes the adversary deviate from the real packet transmission path. However, with the increase of the number of source nodes, the fake packet path generated by the intersection nodes and the nodes on the random routing path is close to sink, and cannot effectively induce the adversary to the destination nodes and the fake sinks. The attack time of this algorithm is relatively long. In the ABVC algorithm, the phantom source node P_1 transmits fake packets to the destination node and the fake sink, which can effectively make the adversary deviate from the real packet transmission path. In the circumferential routing phase, packets are transmitted on the virtual circle and far from the sink, which increases the difficulty of the adversary to capture the sink. The fake path transmission is transmitted by P_2 can effectively hide the shortest path, causing adversary to deviate from real packet path transmission path, and delaying attack time. With delayed attack time, the protection of the sink location privacy is enhanced.

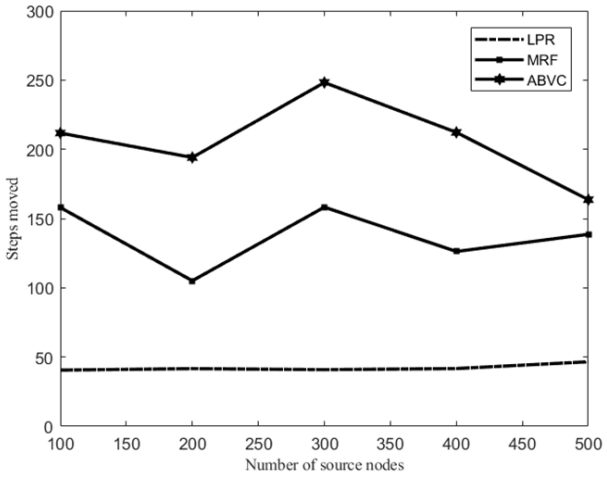

Fig. 11. Attack time provided by different algorithms with fake packets injection and fake sinks injection.

6.4. Energy Consumption

The energy consumption refers to the total number of hops it takes for all packets (both fake and real packets) to be transmitted under a certain routing protocol. As shown in Fig. 12, a

comparison is made among the energy consumption of the three algorithms without injection of fake packet and fake sinks. In the MRF algorithm, the real packet is mainly transmitted along the shortest path, so the energy consumption is the least. In the LPR algorithm, packets are randomly transmitted to the closer list or further list. In this algorithm, the attacks by the adversary is delayed by increasing the length of the packet transmission path, making the energy consumption higher than that of MRF. In the ABVC algorithm, the source node transmits packets to the virtual circle far away from the sink through the phantom routing phase based on source node location; the phantom source node P_1 transmits packets on the virtual circle clockwise at a random angle through the circumferential routing phase based on random angles; the phantom source node P_2 transmits packets from the virtual circle to the sink through the shortest path routing phase. The above three stages generally increase the number of packets forwarding. However, the transmission path of packets in the LPR algorithm is always toward the sink. Compared with the ABVC algorithm, the packet is not transmitted to a place farther from the sink. Thus, the ABVC algorithm makes the energy consumption higher than that of the other two algorithms.

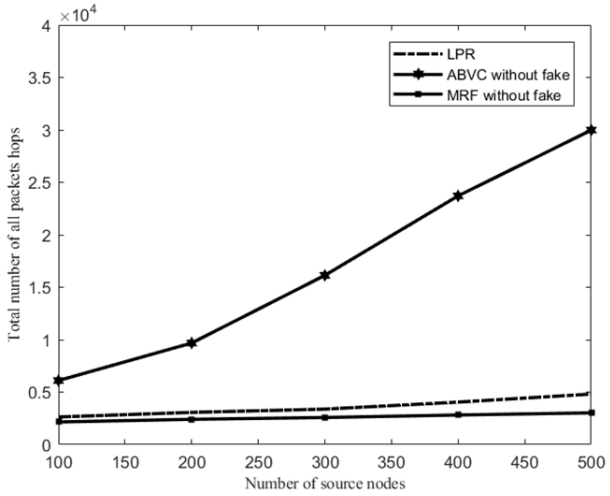

Fig. 12. Energy consumption provided by different algorithms without fake packets injection and fake sinks injection.

As shown in Fig. 13, when fake packets and fake sinks are injected and the number of source nodes increases, the energy consumption of the three algorithms increases. LPR only increases the length of the transmission paths of the real packets without fake packet injection, so the growth of energy consumption is lowest. In the MRF algorithm, the real packet is mainly transmitted along the shortest path, which consumes less energy than the LPR algorithm. However, a large number of fake packets generated in the random routing path from the intersection node, including those transmitted by the intersection node and nodes on the random routing path, the algorithm consumes extra energy. Therefore, the energy consumption of this algorithm is higher than that of the LPR algorithm. In the ABVC algorithm, the real packet path consumes the most energy among the three algorithms. The large number of fake packet paths generated by the two phantom source nodes greatly increases the energy expenditure, so the energy consumption of this algorithm is higher than that of the other two algorithms.

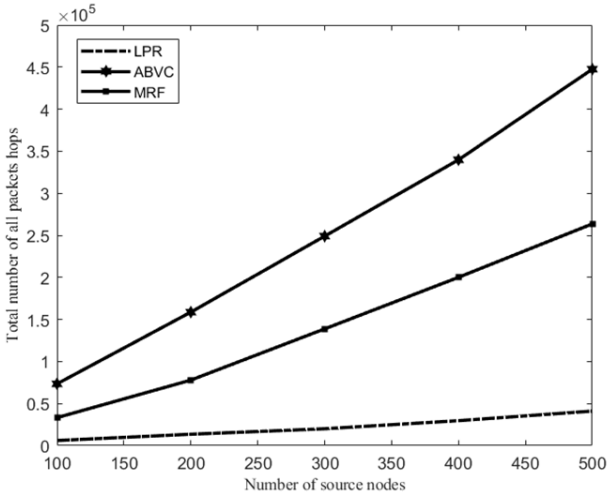

Fig. 13. Energy consumption provided by different algorithms with fake packets injection and fake sinks injection.

7. Conclusions

To protect the sink location privacy, the paper proposes an algorithm of sink location privacy protection based on virtual circle in WSNs. In ABVC, the phantom source node is determined by using the equations of line and virtual circle, not only randomizing its own location, but also generating enough long and complex packet transmission paths on the virtual circle, which can effectively induce adversaries to deviate from the real path and improve the safe time. The analysis of setting to deploy fake sinks shows that the selection of an angle can be made randomly in the external region of the virtual circle with the radius of ; by using the mathematical model of normal distribution, the distance between the fake sink and the sink can be determined, enabling the fake sink to appear near the virtual circle with higher probability. To sum up, the coordinates of the fake sinks are determined according to the random angle and distance. The simulation results show that the algorithm can effectively improve the strength of protection of the sink location privacy.

Acknowledgements. This work was supported by National Natural Science Foundation of China(No.61602323), National Postdoctoral Foundation of China(No.2016M591455), Youth Seedling Foundation of Liaoning Province(No.lnqn201913).

References

[1] H.J. SMITH, T. DINEV, AND H. XU., *Information privacy research: an interdisciplinary review*, Social Science Electronic Publishing **35**(4), pp. 989–1015, 2011.

[2] A. PROANO, L. LAZOS AND M. KRUNZ, *Traffic Decorrelation Techniques for Countering a Global Eavesdropper in WSNs*, IEEE Transactions on Mobile Computing **16**(3), pp. 857–871, 2017.

[3] M. BRADBURY, A. JHUMKA, *A Near-Optimal Source Location Privacy Scheme for Wireless Sensor Networks*, IEEE Trustcom/bigdatase/icessm, pp. 409–416, 2017.

[4] KUSDARYONO A, LEE KO AND LEE Y. A clustering protocol with mode selection for wireless sensor network, JIPS7(1), pp. 29-42, 2011.

[5] J DENG, HAN R AND MISHRA S., *Intrusion tolerance and anti-traffic analysis strategies for wireless sensor networks*, Proceedings of 2004 International Conference on Dependable Systems and Networks,2004.

[6] J DENG, HAN R AND MISHRA S., *Countermeasures Against Traffic Analysis Attacks in Wireless Sensor Networks*, Proceedings of First International Conference on Security and Privacy for Emerging Areas in Communications Networks, 2005.

[7] J YING, CHEN S, ZHANG Z AND ZHANG L., *A novel scheme for protecting receiver's location privacy in wireless sensor networks*, IEEE Transactions on Wireless Communications **7**(10), pp. 3769–3779, 2008.

[8] A LIU, X LIU AND Z TANG., *Preserving Smart Sink-Location Privacy with Delay Guaranteed Routing Scheme for WSNs*, Acm Transactions on Embedded Computing Systems **16**(3), pp. 1–25, 2017.

[9] LIN YAO, LIN KANG, PENGFEI SHANG AND GUOWEI WU., *Protecting the sink location privacy in wireless sensor networks*, Personal and Ubiquitous Computing **17**, pp. 883–893, 2013.

[10] JUN LONG, ANFENG LIU, MIANXIONG DONG AND ZHI LI., *An energy-efficient and sink-location privacy enhanced scheme for WSNs through ring based routing*, Journal of Parallel and Distributed Computing, pp.47–65, 2015.

[11] G. CHAI, M. XU, W. XU AND Z. LIN., *Enhancing Sink-Location Privacy in Wireless Sensor Networks through k-Anonymity*, International Journal of Distributed Sensor Networks **8**(4), pp. 1–16, 2013.

[12] B.D.YING, D. MAKRAKIS AND H.T. MOUFTAH, *Anti-traffic analysis attack for location privacy in WSNs*, Eurasip Journal on Wireless Communications & Networking, 2014.

[13] R. RIOS, J. CUELLAR AND J. LOPEZ, *Probabilistic receiver-location privacy protection in wireless sensor networks*, Elsevier Science Inc **321**(10), pp. 205–223,2015.

[14] S. LEE, J. KIM AND Y. KIM, *Preserving source- and sink-location privacy in sensor networks*, Computer Science & Information Systems13(1), pp. 115–130, 2015, **13**(00):40–40.

[15] A.S. ABUZNEID, T. SOBH, M. FAEZIPOURF, A. MAHMOOD AND J. JAMES, *Fortified Anonymous Communication Protocol for Location Privacy in WSN: A Modular Approach*, Sensors **15**(3), pp. 5820–5864, 2015.

[16] HONGLONG CHEN, WEI LOU, *On protecting end-to-end location privacy against local eavesdropper in Wireless Sensor Networks*, Pervasive and Mobile Computing, pp.36–50, 2015.

[17] J. CHEN, Z. LIN, Y. LIU, Y. HU AND X. DU, *Sink location protection protocols based on packet sending rate adjustment*, International Journal of Distributed Sensor Networks12, 2016.

[18] N. BAROUTIS, M. YOUNIS, *Load-conscious maximization of base-station location privacy in wireless sensor networks*, Computer Networks **124**(4), pp. 126–139, 2017.

[19] JIAN WANG, FENGYU WANG, ZHENZHONG CAO, FENGBO LIN AND JIAYAN WU, *Sink location privacy protection under direction attack in wireless sensor networks*, Wireless Networks **23**(2), 2017.

[20] KONG XIANG-XUE, YUAN SHAO-QING, CHEN MENG, *Routing protocol of source-location privacy protection based on virtual ring*, Transducer and Microsystem Technologies **37**(1), pp. 66–69, 2018.

[21] ZHOU CHUANG, HU XIAOHUI, *Phantom routing privacy protocol based on random directed in WSN*, Application Research of Computers **35**(11), pp. 689–697, 2018.

[22] JIA ZONGPU, WEI XIAOJUAN AND PENG WEIPING, *Privacy protection strategy about source location in WSNs based on random angle and circumferential routing*, Application Research of Computers **33**(3), 2016.

[23] HUANG BEI-BEI, FENG YONG, LI XIU-QI AND HUANG Qi, *Angle-based directed random walk phantom routing protocol for WSNs*, Transducer and Microsystem Technologies **35**(11), pp. 123–127, 2016.

[24] HAN G, ZHOU L, WANG H, ZHANG W AND CHAN S., *A source location protection protocol based on dynamic routing in WSNs for the Social Internet of Things*, Future Generation Computer Systems **82**, 2018.

[25] OZTURK C, ZHANG Y AND TRAPPE W, *Source-Location privacy in energy-constrained sensor network routing*, In: Proc. of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN 2004).

[26] YUN LI, JIAN REN, *Mixing Ring-Based Source-Location Privacy in Wireless Sensor Networks*, Proceedings of 18th International Conference on Computer Communications and Networks,2009.

[27] YUN LI, JIAN REN, *Mixing Ring-Based Source-Location Privacy in Wireless Sensor Networks*, Proceedings of 18th International Conference on Computer Communications and Networks,2009.