ICIC Express Letters
Part B: Applications
Volume **7**, Number **7**, July **2016**

ICIC International ©2016 ISSN 2185-2766
pp. **1497–1503**

SOURCE-LOCATION PRIVACY PROTECTION ALGORITHM IN WIRELESS SENSOR NETWORK BASED ON PHANTOM SOURCE AND FAKE SOURCE

Leqiang Bai[1], Ling Li[1], Shiguang Qian[1] and Shihong Zhang[2]

[1]Information and Control Engineering Faculty
Shenyang Jianzhu University
No. 9, Hunnan East Road, Hunnan New District, Shenyang 110168, P. R. China
{ baileqiang; qsg }@sjzu.edu.cn; liling0ll@163.com

[2]Allwin Telecommunication Company
No. 6, Gaoge Road, Hunnan New District, Shenyang 110179, P. R. China
kengreen1987@163.com

Received December 2015; accepted March 2016

Abstract. *For the problem that the phantom source nodes generated by the phantom routing are near the real source node and algorithms based on flooding have excessive energy consumption, source-location privacy protection algorithm based on phantom source and fake source is put forward. With neighbor information and random walk, the improved algorithm randomly selects phantom source nodes and fake source nodes by intermediate nodes in the shortest phantom source-sink paths. Theoretical analysis shows that the algorithm can select intermediate node for each data packet. The experimental results show that the algorithm can decrease the energy cost and delivery latency. And safety period increases significantly with the increase of hops from source to sink node.*
Keywords: Wireless sensor network, Source-location, Phantom source, Fake source

1. **Introduction.** With the development of technology, information overload is common [1]. Wireless sensor network (WSN) can collect the information effectively. Sensor nodes form a multi-hop and self-organizing WSN in ad hoc manner. Privacy is categorized into two categories: data content and data location [2]. For the data content privacy, encryption is adopted to antagonize the adversary who can compromise network nodes and distort the content of packets. For the data location privacy, random routing and cyclic entrapment are utilized to antagonize the adversary monitored and analyzed traffic of the network.

Adversaries can be classified into two categories: global adversary and local adversary according to the monitor ability [3]. Global adversary is assumed to acquire a global view of the network traffic [4], based on nodes arranged in multiple places of the network. Local adversary randomly walks until overhearing a packet. The adversary can only trace the traffic flow by one hop during one packet transmission because the speed of a packet is far faster than the movement of adversary [5]. Based on attack pattern, there are two types of adversaries: active adversary and passive adversary. Active adversary distorts the content of packets. Passive adversary traces packets, without interfering with the normal communication of the network.

The Panda-Hunter Game and phantom flooding were put forward by Ozturk et al. [6]. Phantom single-path routing was proposed by Kamat et al. [7], while pure random cannot make phantom source node be away from real source node completely. Chen et al. proposed source-based restricted flooding protocol (PUSBRF) [8], while flooding consumes much energy. Mehta et al. proposed two methods to protect the privacy of the source-location: periodic data collection and the data source node simulation [4], while the delay

cannot be avoided. Tscha proposed boundary greedy stateless location privacy protection agreement (GSLP) [9], which is complex for application.

The algorithm based on phantom source and fake source (ABPSFS) presented here aimed at protecting effectively source-location privacy and reducing energy cost. According to neighbor information, node randomly selects phantom source nodes through random walk and fake source nodes by intermediate nodes. Using Panda-Hunter Game to conduct experimental verification, results show that the algorithm can effectively decrease the energy cost and delivery latency on this premise of approximate network safety period.

The rest of this paper is as follows. Section 2 is problem definition, including network model and the adversarial model. Section 3 is the flow and analysis of proposed ABPSFS. Section 4 shows experimental results and analysis. Section 5 is conclusion.

2. Problem Statement and Preliminaries.

2.1. System model.

Node types: sink node and general nodes [8]. The system is similar to the Panda-Hunter Game introduced in [6]. Once a panda appears, the corresponding node in the nearby area will observe and report data periodically to the sink. The illegal hunter, namely, adversary may try to track and locate source-location and capture the panda.

The paper makes the following assumptions about network.

1. The sink node is the only destination node of the network for collecting packets. The information of the sink node is public.
2. Each sensor node has three types of information: minimal hops to sink, neighboring nodes and the minimal hops of neighbor nodes to the sink. The nodes are stationary [10].
3. The content of each packet will be encrypted [11]. An adversary cannot read directly from the packet to obtain the location of the source node.

2.2. Adversary model.

Adversary is assumed to have the following characteristics.

1. The hearing radius of local adversary is limited, equal to sensor transmission radius.
2. The adversary will not interfere with the proper communication of the network. The initial location of it is near the sink, eavesdropping packets which forwards to sink.
3. The adversary equips equipment such as GPS [9], analyzes the angle of arrival to track the nodes that sent packets. For fixed path routing of length n, the adversary is able to locate the message source node if it captures n messages [11].

3. Procedure of ABPSFS.

3.1. The overview of the ABPSFS.

In order to protect source-location privacy effectively, the algorithm selects intermediate node in the shortest phantom source-sink path by means of random number and hops information to sink node. And there must be an intermediate node during the transmission procedure of every packet. Intermediate node selects fake source node by random walk. Fake source node sends fake packet to sink to mislead adversary. Under the interference of fake packets, adversary cannot track true packets continuously. With the help of phantom source node and fake source node, the ABPSFS algorithm achieves the goal of protecting source-location privacy. Specifically, the ABPSFS is divided into four phases: network initialization, selecting phantom source node, selecting the intermediate node and intermediate node selecting fake source node as well as sending packets phase. The main symbols used in this paper are listed in Table 1.

TABLE 1. The main symbols used in the paper

Symbol	Meaning
S	Source node
B	Base station, the node collects packets, namely, sink node
h_S	The random walk hops of S selects phantom source node
SP	The phantom source node which S selected by h_S
h_{Q_B}	The minimal hops to the sink from Q, Q is random node, such as S
L_{Q_B}	The actual transmission path of packet which was sent by a random node Q to the sink node after the packet was transmitted h_{Q_B}
SPS	The intermediate node in the L_{Q_B} which generates random walk
P_{fake}	Judgment basis of selecting SPS in the L_{Q_B}, $P_{fake} = h_{Q\_B}/h_{S\_B}$, $Q \in L_{SP\_B}$
R	Random number generated by Q node ($Q \in L_{SP\_B}$), R in the interval of $[1 - P_{fake}, 1]$, Q node is selected as the SPS if $R \geq P_{fake}$
h_{SPS}	The random walk hops of SPS selects fake source node
$SPSP$	The fake source node which SPS selected by h_{SPS}

3.2. **Description of ABPSFS.**

3.2.1. *Network initialization phase.* Sink node broadcasts Sink_Init packet, Sink_Init = {sink_broadcast, sender_ID, sink_hop}. The sink_broadcast is message type; sender_ID is ID of sender node; sink_hop is the hop count to sink, and its initial value is zero. The node Q received packet stores sender_ID and the minimum sink_hop, forwards the packet after updating send_ID and increasing the sink_hop by one. In this way, node Q builds a neighbor list that contains its neighbor nodes at hops $h_{Q\_B} - 1$, h_{Q_B} and $h_{Q\_B} + 1$.

3.2.2. *Selecting the phantom source node phase.* S and B represent the real source and the base station node respectively as shown in Figure 1. The node with the minimal distance of panda becomes S. The S adds h_{S_B} to packet, sets the h_S as the hop count and generates random walk. S finds out neighbor nodes which are $h_{S\_B} + 1$ to sink, randomly selects a node as the forward node and sends the packet. If there is no node which is $h_{S\_B}+1$, randomly select a node from the neighbor nodes which are h_{S_B} to sink. Otherwise, randomly select a neighbor node. The node received packet decreases the hop count of the packet by one. If the hop count is 0, the node becomes SP. Otherwise, the node selects next forward node following above principle and sends the packet. Nodes that have been selected are excluded from process of random selecting next hop node, unless all the neighbor nodes have been selected. Repeat the process until the hop count is zero, and the process of selecting SP ends.

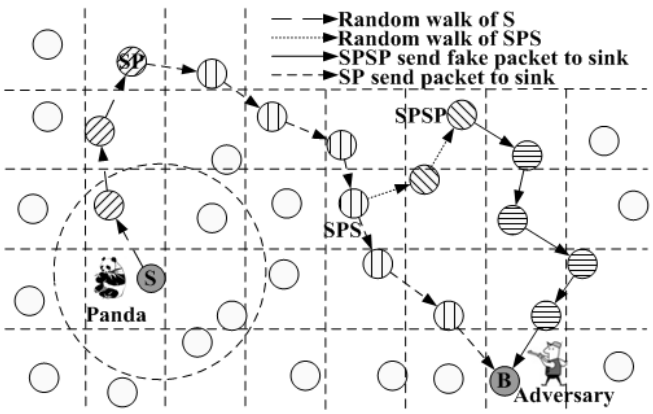

FIGURE 1. Illustration of ABPSFS

3.2.3. *Selecting the intermediate node phase. SP* checks neighbor list, randomly selects a neighbor node which has minimal hops to sink and sends packet. The node received the packet generates R. The node will be *SPS* if $R \geq P_{fake}$. Otherwise, the node sends the packet to a neighbor which has the minimal hops to sink. The node received the packet repeats the process until *SPS* was selected.

3.2.4. *Intermediate node selecting fake source node as well as sending packets phase.*

1. *SPS* sends packet to sink along the shortest path.

The shortest path is a basic issue in network optimization [12]. *SPS* selects a neighbor node that has the minimal hops to sink, forwards the packet. The node received packet selects a neighbor node and forwards it based on the above conditions until packet arrived at sink.

2. *SPS* generates random walk and selects *SPSP*, and *SPSP* forwards fake packet to sink.

The process of *SPS* selecting *SPSP* by h_{SPS} is similar to the process of S selecting *SP* by h_S. *SPSP* sends fake packet to sink along the shortest path.

3.3. **Theoretical analysis.**

Theorem 3.1. *In any case, there is a* $Q \in L_{SP\_B}$ *that will be intermediate node SPS, and* $h_{SPS\_B} \leq h_{S\_B}$*, when* $h_{S\_B} \geq 2$ *and* $h_{SP\_B} \geq 2$*.*

Proof: The selection of *SPS* in phase 3 is presented in Figure 2:

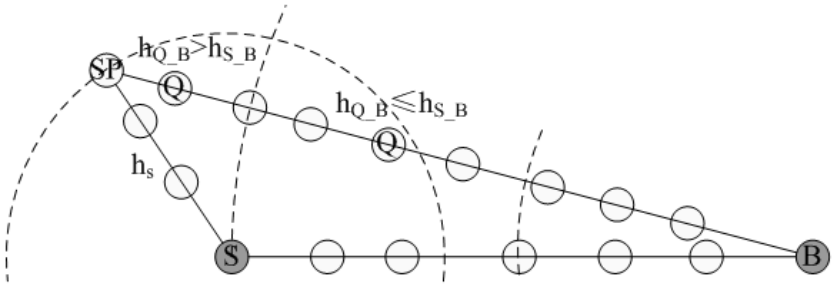

FIGURE 2. The inevitability and existence range of *SPS*

If $Q \in L_{SP\_B}$ and $h_{Q\_B} > h_{S\_B}$, $P_{fake} = h_{Q\_B}/h_{S\_B} > 1$. Because of $1 - P_{fake} \leq R \leq 1$, $R < P_{fake}$, the Q like this cannot be SPS.

If $Q \in L_{SP\_B}$ and $h_{Q\_B} \leq h_{S\_B}$, $P_{fake} = h_{Q\_B}/h_{S\_B} \leq 1$. With node Q away from *SP*, close to B, h_{Q_B} decreases, P_{fake} decreases, $1 - P_{fake}$ increases. When $h_{Q\_B} \leq h_{S\_B}/2$, $P_{fake} = h_{Q\_B}/h_{S\_B} \leq 1/2$, $1 - P_{fake} \geq 1/2$. Because of $1 - P_{fake} \leq R \leq 1$, $R \geq P_{fake}$, the Q will be selected to intermediate node *SPS*.

In summary, in any case, there is a $Q \in L_{SP\_B}$ that will be intermediate node *SPS*, and $h_{SPS\_B} \leq h_{S\_B}$, when $h_{S\_B} \geq 2$ and $h_{SP\_B} \geq 2$.

4. **Experiment and Analysis.** In order to verify the performance of the ABPSFS, experiments were conducted by Matlab platform to simulate phantom single-path routing, PUSBRF and ABPSFS from safety period, energy cost and transmission delay. Experiment configure is similar to [7,8]. 10000 sensor nodes were uniformly randomly distributed over a 6000×6000 (m^2) network which was evenly divided into 100×100 grids. The initial position of each node is grid center, with random disturbance, ensuring that there is only one node in each grid and the relative position is different. The communication radius of sensor is 100 m. The hearing radius of adversary is also 100 m. The average number of neighbors is 8.72. Sink is in the center of the network and S is selected randomly. Figures 3-5 provide the performance of these three algorithms, when h_S is fixed value. The results are the average of 50 times simulation experiments.

4.1. **Safety period.** The change trends of safety period as h_{S_B} changing were shown in Figure 3 when $h_S = 15$. The safety periods of three algorithms are increasing as the increase of h_{S_B}. The adversary needs to trace longer to overhear more packets to discover the source node as h_{S_B} increased. The PUSBRF selects phantom source by the hops of phantom source to S. When the value of h_{S_B} is small, PUSBRF has the maximal safety period. However, with the increase h_{S_B}, the safety period of ABPSFS surpasses it after h_{S_B} is close to 30.

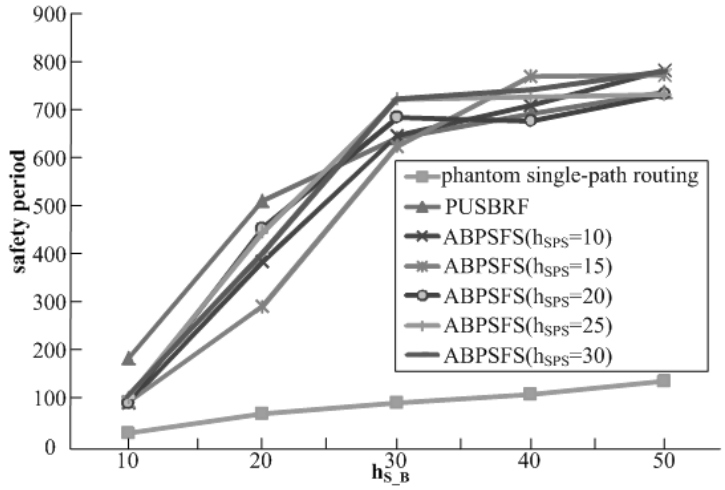

FIGURE 3. Safety periods of phantom single-path routing, PUSBRF and ABPSFS

4.2. **Energy cost.** The change trends of energy cost as h_{S_B} changing were shown in Figure 4 when $h_S = 15$. The energy costs of three algorithms are increasing as the increase of h_{S_B}. The phantom single-path routing has the minimal energy cost, because phantom source sends the packet directly to the sink along the shortest path. The PUSBRF has the maximal energy cost, because flooding sends large amounts of packets, which increases energy cost.

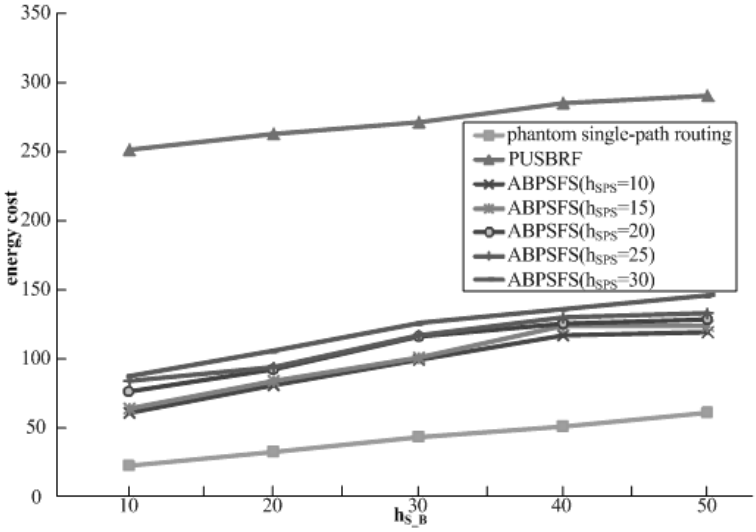

FIGURE 4. Energy costs of phantom single-path routing, PUSBRF and ABPSFS

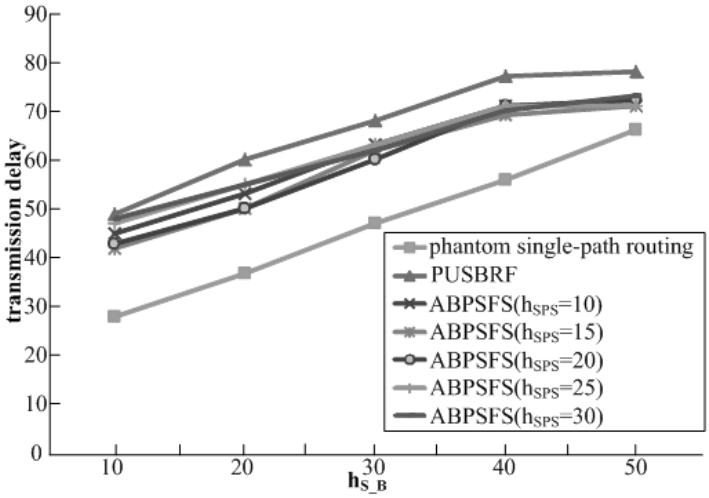

FIGURE 5. Transmission delays of phantom single-path routing, PUSBRF and ABPSFS

4.3. **Transmission delay.** The change trends of transmission delay of three algorithms as h_{S_B} changing were shown in Figure 5, when $h_S = 15$. The transmission delays of three algorithms are increasing as the increase of h_{S_B}. The phantom single-path routing and ABPSFS have lesser transmission delay, because phantom source sends the packet directly to the sink along the shortest path. The fake packets generated by *SPSP* have no effect on delay. The PUSBRF has the maximal transmission delay because of flooding.

According to the above analysis, the safety period increases with the increase of h_{S_B}, but the energy cost and transmission delay also increase. Energy cost and transmission delay of ABPSFS are between PUSBRF and phantom single-path routing. Safety period of ABPSFS is greater than PUSBRF when h_{S_B} is large. Under the condition of similar network safety period, energy cost and transmission delay of ABPSFS are significantly lower than PUSBRF. In order to balance safety period, energy cost and transmission delay, ABPSFS algorithm makes the h_S and h_{SPS} in the range of $10 \leq h_S \leq 15$ and $10 \leq h_{SPS} \leq 20$, and specific values can be adjusted by the network security requirements.

5. **Conclusions.** The paper proposed source-location privacy protection algorithm based on phantom source and fake source. This paper demonstrates inevitability and existence domain of intermediate node. Based on Panda-Hunter Game, experiments verify the effectiveness of the algorithm. The results indicate that the algorithm can effectively resist the backtracking attack of local adversary, meet the requirement of source-location privacy protection when the source node is away from sink and prolong network lifetime by reducing energy cost and transmission delay. For future work, that the privacy protection algorithm combined data content and location, algorithm resisted more skilled adversary and algorithm fitted multiple monitor objects are all hot research directions.

Acknowledgment. Project is supported by the National Natural Science Foundation of China under Grant (60973022/F020202).

REFERENCES

[1] M. Zhang, S. Q. Yin, C. Liu, Y. Zeng and J. Y. Liu, Using a modified similarity measure in collaborative filtering, *ICIC Express Letters, Part B: Applications*, vol.4, no.6, pp.1755-1761, 2013.

[2] R. D. Pietro and V. Alexandre, Location privacy and resilience in wireless sensor networks querying, *Computer Communications*, vol.34, no.3, pp.512-523, 2011.

[3] M. Raj, N. Li, D. G. Liu, M. Wright and S. K. Das, Using data mules to preserve source location privacy in wireless sensor networks, *Pervasive and Mobile Computing*, vol.11, pp.244-260, 2014.

[4] K. Mehta, D. Liu and M. Wright, Protecting location privacy in sensor networks against a global eavesdropper, *IEEE Trans. Mobile Computing*, vol.11, no.2, pp.320-336, 2012.

[5] H. Chen and W. Lou, On protecting end-to-end location privacy against local eavesdropper in wireless sensor networks, *Pervasive and Mobile Computing*, vol.16, pp.36-50, 2015.

[6] C. Ozturk, Y. Zhang and W. Trappe, Source-Location privacy in energy-constrained sensor network routing, *Proc. of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks*, pp.88-93, 2004.

[7] P. Kamat, Y. Zhang, W. Trappe and C. Ozturk, Enhancing source-location privacy in sensor network routing, *Proc. of the 25th IEEE Int'l Conf. on Distributed Computing Systems*, pp.599-608, 2005.

[8] J. Chen, B. X. Fang, L. H. Yin and S. Su, A source-location privacy preservation protocol in wireless sensor networks using source-based restricted flooding, *Chinese Journal of Computer*, vol.33, no.9, pp.1736-1747, 2010.

[9] Y. Tscha, Routing for enhancing source-location privacy in wireless sensor networks of multiple assets, *Journal of Communications and Networks*, vol.11, no.6, pp.589-598, 2009.

[10] M. M. E. A. Mahmoud and X. Shen, A cloud-based scheme for protecting source-location privacy against hotspot-locating attack in wireless sensor networks, *IEEE Trans. Parallel and Distributed Systems*, vol.23, no.10, pp.1805-1818, 2012.

[11] Y. Li, J. Ren and J. Wu, Quantitative measurement and design of source-location privacy schemes for wireless sensor networks, *IEEE Trans. Parallel and Distributed Systems*, vol.23, no.7, pp.1302-1311, 2012.

[12] F. Luo, X. G. Zhang and D. J. Wei, A bio-inspired algorithm for solving shortest path problem with fuzzy arc lengths, *ICIC Express Letters, Part B: Applications*, vol.5, no.3, pp.627-632, 2014.