Assessing the Performance of Phantom Routing on Source Location Privacy in Wireless Sensor Networks

Chen Gu, Matthew Bradbury, Arshad Jhumka, and Matthew Leeke
Department of Computer Science, University of Warwick,
Coventry, United Kingdom, CV4 7AL
{cspmaj, bradbury, matt, arshad}@dcs.warwick.ac.uk

***Abstract*—As wireless sensor networks (WSNs) have been applied across a spectrum of application domains, the problem of source location privacy (SLP) has emerged as a significant issue, particularly in safety-critical situations. In seminal work on SLP, phantom routing was proposed as an approach to addressing the issue. However, results presented in support of phantom routing have not included considerations for practical network configurations, omitting simulations and analyses with larger network sizes. This paper addresses this shortcoming by conducting an in-depth investigation of phantom routing under various network configurations. The results presented demonstrate that previous work in phantom routing does not generalise well to different network configurations. Specifically, under certain configurations, it is shown that the afforded SLP is reduced by a factor of up to 75.**

***Keywords*-Sensor networks; Source Location Privacy; Phantom Routing; Multiple Sources.**

I. Introduction

A wireless sensor network (WSN) consists of a number of small devices, known as sensor nodes or motes, that can sense the environment and use radio signals to communicate amongst themselves. WSNs have enabled the development of novel applications, including those in asset monitoring and tracking, with low levels of intrusiveness. As they are often deployed in safety-critical situations, including those in healthcare and military intelligence [1], the communication protocols used in WSNs must meet stringent security and privacy requirements.

Threats to privacy in monitoring applications can be considered along two dimensions: (i) content-based threats and (ii) context-based threats. Content-based privacy threats relate to use of the content of the messages broadcast by sensor nodes, such as gaining the ability to read an eavesdropped encrypted message. Much research has addressed the issue of providing content privacy, e.g., SPINS [2], with most efforts in this area focusing on the use of cryptographic techniques. On the other hand, context-based privacy threats focus on the context in which messages are broadcast and how information can be observed or inferred by attackers. Context is a multi-attribute concept that encompasses situational aspects of broadcast messages, including environmental and temporal information.

It is often desirable for the source of sensed information to be kept private in a WSN. Algorithms that protect this contextual information are said to provide source location privacy (SLP). SLP is important in many application domains, though it is of utmost concern in safety-critical situations. For example, in a military scenario the location of a source node may represent a soldier. In the case of healthcare, the location may refer to each patient or an ambulance. In each of these scenarios it is important to ensure that an attacker cannot find the location of the asset being monitored, whether it is a soldier or a patient. A WSN set up to forward the information collected about an asset would typically consist of: a dedicated node for data collection called a *sink*, many nodes that are involved in sending information from these assets called *sources*, and many other nodes in the network used to route messages over multiple hops from the sources to the sink. It has been shown that an attacker can backtrack message paths through the network to find the source node and capture the asset [3]. Thus, there is a need to develop algorithms to provide SLP.

A number of techniques have been proposed to provide SLP, such as using random walks [3], geographic routing, delays [4], dummy data sources [5, 6] and so forth. However, there has been no universal solution proposed to deal with all SLP issues. In general, the performance of SLP algorithms depends on the assumed attacker and network models. Various attacker models have been considered, from local attackers with a limited view of the network (but who can gain more information by moving), to global attackers with the ability to see and eavesdrop network traffic across the network. Some categories of techniques, like random walk methods, have good performance with attackers who have a local view of network traffic, but fail to provide SLP against the global attackers. There are further models that assume multiple local co-ordinating attackers [7]. These increasingly intelligent attacker models have improved SLP provision, as they have necessitated the reconsideration of existing approaches.

Network configurations can change the SLP provided by an algorithm, making it necessary to investigate a variety of different scenarios when developing an SLP algorithm. In a seminal paper on SLP, the authors of [3] proposed the concept of *phantom routing* to provide SLP. Their results showed that phantom routing provided a high level of SLP. However, the range of experiments conducted was restrictive,

such that little is known about the ability of phantom routing to provide high levels of SLP when conditions vary. In this paper, we address this shortcoming and evaluate the ability of phantom routing to provide SLP under different network configurations. Specifically, we identify three parameters, namely (i) message rate, (ii) number of sources and (iii) length of random walk, that impact on the performance of phantom routing. We vary these parameters to assess their impact, both individually and in combination, on phantom routing as a viable approach to the problem of SLP.

The main contributions of this paper are to:

- Identify three parameters that impact the performance of phantom routing.
- Derive expressions that capture the impact of the three identified parameters, as well as conducting a range of experiments to validate these findings.
- Demonstrate that, under varied network conditions, the performance of phantom routing can degrade by a factor of up to 75, confirming an initial conjecture that phantom routing works well under specific conditions but requires fine-tuning in order to realise optimal performance.

The overarching contribution of this paper is to demonstrate that much existing research in phantom routing does not generalise well to varied network configurations, particularly in the context of larger network sizes. More specifically, it is demonstrated that, under certain configurations, the SLP afforded by phantom routing is reduced by a factor of up to 75.

The remainder of this paper is organised as follows: Section II presents a survey of related work in SLP. Section III details the intention of this paper and phantom routing as a technique for SLP. The adopted system and attacker models are outlined in Section IV. Details of the experiments conducted are provided in Section V. Section VI presents the results of the experiments conducted, before Section VII concludes this paper with a summary of contributions.

II. Related Work

Wireless communication technologies and power-efficient sensors have been used in a broard spectrum of remote-sensing applications [8]. The SLP problem first emerged around 2005 [3]. Since then, many schemes have been proposed to provide SLP. For instance, sending dummy message can be considered a possible solution to the SLP problem [9, 10]. The principle of this approach is to mix real message broadcasts with dummy message broadcasts. When a node hears an event, it forwards the real message but, when it would otherwise be idle, it transmits dummy messages. By keeping the contents of the dummy and real messages indistinguishable, e.g., using padding and encryption, an attacker can not tell whether any message is real by means of comparison.

There are various schemes for selecting nodes to send dummy messages. For example, some schemes select nodes in an attempt to entrap mobile attackers in a cycle [5], whilst others make direct attempts to lure attackers away from real sources using [6] or used tree-based broadcasts models [11]. These solutions typically differ in their attacker model, with some assuming mobile attackers that have limited network visibility and others assuming a single attacker that has a global view of the network.

Global solutions to SLP often involve all nodes broadcasting periodically. For example, in the Periodic Collection algorithm node will send messages that are either the real message or a dummy message generated to confuse the attacker [12]. Improvements to this technique have involved modifying the broadcast period to follow a statistical distribution [13]. The benefits of doing this are to decrease the latency of the real message transmission, whilst hindering statistical analyses of traffic patterns. An alternative technique against a global attacker is Source Simulation, where nodes broadcast dummy messages in a pattern around the network that matches the source's movement [12]. However, in order for this technique to be practical at scale, some level of privacy must be sacrificed in order to reduce energy consumption.

Perhaps the most significant disadvantage of the described SLP techniques is the volume of messages broadcast in order to provide SLP. This leads to increased energy consumption and an increased number of collisions, both of which result in a decreased packet delivery ratio. This means that a tradeoff between energy expenditure and privacy must be made [7]. For this reason, dummy message schemes may not be appropriate for many large-scale networks.

A variety of approaches that do not rely on the use of dummy messages have also been proposed to address the SLP problem [4, 14, 15, 16, 17]. In particular, Ozturk et al. [8] and Kamat et al. [3] proposed a two-phase solution. A message from the source would perform a walk through the network to a location where it would become a phantom source. At the phantom source the message would then be routed (by flooding, single-path routing, or some alternative scheme) to the sink. The author's initially investigated the issue by using a pure random walk scheme. Unfortunately, the scheme did not provide good SLP. This is because when the message randomly travels for h hops, it will tend to remain close to the source's location [3]. As the message doesn't finish the walk far from the source, it isn't effective in luring the attacker away as the phantom node will be created near the source. The solution the author's proposed to this problem was to use a directed random walk, where the message is either sent towards or away from a certain node in the network, e.g., the source, the west-most node, etc..

The phantom routing scheme was the first solution to use a random walk in the provision SLP [3]. Phantom routing has received a lot of attention in the literature, and many improved techniques [18, 19, 20] have been proposed based on it. The majority of these focus on improving how the random walk is performed. Wei-Ping et al. [18] proposed using location angles to construct the random walk. Similarly, Yao and Wen [20] used a directed random walk, whilst Zhang [19] had the directed walk adjust to an estimated source location. Xi

et al. [21] used a different approach in GROW, by recording neighbours in a bloom filter which informed the choice of the next node to be used in the random walk.

III. Problem Statement

The problem we address in this paper is as follows: Given a WSN topology $G = (V, E)$, where V is a set of wireless sensor nodes and E is a set of edges or links; a routing protocol $\mathbb{R}$ to transport data towards the sink, a set $\mathbb{L}$ of source locations, a safety period δ, evaluate the performance of phantom routing over δ in the presence of $|\mathbb{L}|$ sources using $\mathbb{R}$ in G.

The phantom routing technique can be explained as follows: The phantom routing consists of two phases. The first phase is the random walk phase and the second phase is the flooding phase. Instead of using pure random walk, an improved version (called the directed random walk) is used instead. In a directed random walk, each node divides it neighbouring nodes into two sets, which are opposite to one another. When a node starts to send a message, it randomly choose one set and sends the message to a neighbouring node chosen at random out of that set. When an intermediary node receives the message, the message will be transmitted to a random neighbour from the opposite set. For instance, if an intermediary node receive a message from a neighbour out of its South-West set, then it forwards the message to one of its neighbour in the North-East set. The forwarding message stops when it has travelled h hops or when it cannot be forwarded any more into the same direction. The final node is called phantom node. Then the message starts flooding throughout the network.

IV. System Model

In this section, we detail the system and attacker models assumed in this paper.

A. System Model

A wireless sensor node is a small computing device with communication and computation capabilities and a sensor network is a collection of such nodes, with a link between a pair of nodes. The link between node pairs may be unidirectional or bidirectional.

It is assumed that all nodes to have the same communication range. A node m that can directly receive a message from a node n is called a *neighbour* of n. Each node has a unique node ID. For ease of demonstration, this work focuses on grid topologies. It is assumed that a node will have knowledge of all of its neighbouring nodes.

A distinguished node, denoted by S, is responsible for collecting data and is called the *sink*. Other nodes, other than the sink, sense data from the environment and use a multi-hop route to carry data towards the sink. In this paper, we assume the existence of a single sink only. Message routing will generally use some data aggregation convergecast protocol [22]. When a node detects an event, it will route, in collaboration with other nodes, the message to the sink.

Multiple nodes called *sources* can exist in the network, these are regular nodes that have detected an asset and are broadcasting information about the asset's status.

B. Attacker Model

In this paper we assume a distributed eavesdropper attacker model [23]. The means that the only action the attacker performs is eavesdropping, while its location, hence knowledge, is distributed across the network, i.e., the attacker can move from one location to another in the network.

Device Strength: We assume that the adversary has a large energy source, i.e., we do not assume infinite energy source but rather that the amount of energy required for the task is much less than the amount of energy available. We also assume that the attacker has enough memory for data storage. The attacker has the ability to determine the source of a message that it overhears (for example, through the use of a directional antenna) and obtain the strength of the signal (for example, using spectrum analysers). However, the attacker does not have the keys to decipher the messages it overhears, so cannot obtain the contents of a message.

Attacker Network Knowledge: We assume that the attacker does not know the locations of nodes. Specifically, the attacker may know the topology of the network but not the specific locations of nodes. For example, the attacker may know that the topology is a grid, but not the placement of the nodes in the grid. WLOG, we also assume that the attacker knows the location of the sink (similar to the assumption made in the seminal work by Kamat et al. [3]). Since the attacker is a distributed eavesdropper, it will learn about the 1-hop neighbourhood of different nodes, depending on its location within the network.

V. Experimental Setup

In this section we outline the experimental configurations used to generate the results presented in Section VI.

A. Tool and Configuration

The simulation environment was based on TOSSIM [24]. When it runs, it stores incoming events in the event queue ordered by the event time and executes them in order. The TOSSIM radio model is based on signal-strength, which accepts parameters including: the noise floor, the receiver sensitivity and a set of data that describes the propagation strengths. TOSSIM adopts the Closest Pattern Matching (CPM) algorithm to simulates the RF noise and interference a node hears. CPM creates a statistical model that captures bursts of interference and other correlated phenomena by a set of input of a noise trace. This model greatly improves the quality of the RF simulation and leads to better performance than traditional, independent packet loss models. The TOSSIM simulator was also extended to allow the capture ratio of sources to be monitored during the simulation. The radio and noise models used in all simulations were consistent with those used in [3].

Network topology and configuration: A square grid network layout of size $n \times n$ was used in all experiments, where n $\in \{11, 15, 21\}$, i.e., 121, 225, 441 nodes within the network respectively. The set of source nodes $\mathbb{L}$ and the sink node were considered independent components of the network. The performance of phantom routing was evaluated for message rates of 1, 2, 4 and 8 messages per second. The set of experiments for each network size were performed for three different network configurations: (i) the sink is located in one corner of the network, with the source node(s) in the middle in the network (referred to as the *SinkCorner* configuration); (ii) source node(s) in the corner and the sink node in the opposite corner (*FurtherSinkCorner*); (iii) the source node(s) in the corner and the sink in the middle (*SourceCorner*). Simulations for each combination of parameters were repeated 200 times. The parameters for the message communication were set such that messages only travel between nodes that are horizontal or vertical neighbours.

Sink-Source Distance: To aid in understanding how the different configurations are arranged it is useful to know the expected distance between the sink and source. Table I shows these distances for various network sizes and configurations. When comparing across configurations it is important to take these different distance into account.

Size	SinkCorner	FurtherSinkCorner	SourceCorner
11×11	10	20	10
15×15	14	28	14
21×21	20	40	20

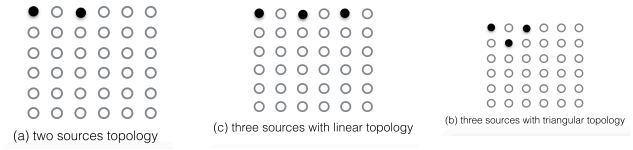
TABLE I: The sink-source distances (in hops) for the different network conditions and different sizes of a network

B. Sources Selection

Given our focus on multiple sources, we explain their distribution in the network. As mentioned, we consider up to three sources in the network. We have described three configurations that we consider of the network, namely SourceCorner, FurtherSinkCorner and SinkCorner. For each of these configurations, we focus on multiple sources distributions.

Two sources: In a network configuration with two source nodes, these are arranged in linear fashion, i.e., there is a distance of two hops between each other, as shown in Figure 1a.

Three sources: For three sources, we consider two types of source distributions: (i) a linear arrangement and (ii) a triangular arrangement, as shown in Figure 1b and Figure 1c respectively. The distance between adjacent sources is fixed to two hops. The application will generate messages at a frequency f_n, $f_n \in \{1, 2, 4, 8\}$. We now explain how this translates into individual message rates for nodes: Consider three source nodes as an example n_1, n_2 and n_3. Consider the application sending f_n messages per second. We split the period (1 sec) into f_n slots and the source nodes take it in turn to send a message. For example, if $f_n = 8$, then sources n_1 and n_2 will send three messages and source n_3 will send two messages, giving a total of 8 messages. In general, if there are n sources and the message rate is f_n, then there is a set of u sources that will transmit $\lceil \frac{f_n}{n} \rceil$ messages and and the rest (l sources) will transmit $\lfloor \frac{f_n}{n} \rfloor$ messages such that $u * \lceil \frac{f_n}{n} \rceil + l * \lfloor \frac{f_n}{n} \rfloor = f_n$.



(a) Linear 2 (b) Linear 3 (c) Triangular 3

Fig. 1: Network Layouts with Varying Number of Sources

Safety period: A concept called safety period was introduced in [3] that represented the level of privacy provided in terms of the number of messages that were sent by the source nodes. The greater the number of messages sent before capture, the greater the privacy provided. In this paper, we use an adapted definition of safety period introduced in [25], which was originally defined for each specific network size and source broadcast rate combination, the time taken by the attacker to detect the source nodes was doubled to establish a safety period. The aim was to ensure that an attacker had sufficient opportunity to detect a real source and also to bound simulation time.

In our case, the maximum simulation time, hence the upper bound of the capture time, of an experiment is computed as follows:

$$(1 + h_w/d_s) \times SP \tag{1}$$

This is where h_w is the length of the random walk, d_s denotes the sink-source distance and SP is the safety period when flooding alone is used[1]. The reason this definition is used is so that the safety period increases as the length of the random walk is higher.

Capture ratio: We define a metric called *capture ratio* (CR) as follows:

$$CR = \frac{\text{Number of experiments ending in a capture}}{\text{total number of experiments}} \tag{2}$$

When there are multiple sources in the network, a capture occurs when at least one of the sources are detected. We say that a source is detected when the attacker is co-located at the source.

VI. Results

In this section we conduct experiments to examine to impact of varied message rates, source numbers and random walk length, both in isolation and combination.

[1]It was shown that the flooding routing protocol provides no source location privacy, which we use as a base protocol.

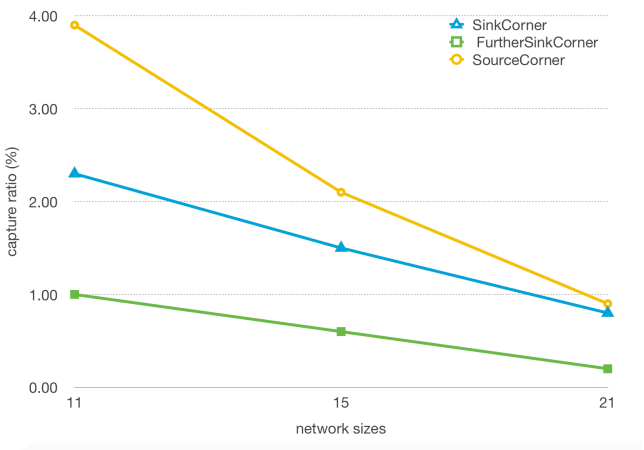


Fig. 2: Impact of varying network sizes and network configurations on SLP for 1 source broadcasting at 1 message/second

A. Base Case: SLP with one source broadcasting at 1 message/second

To establish a baseline, against which further experiments can be compared, a network with a single source transmitting messages at the rate of one message per second is used. The results, shown in Figure 2, demonstrate that as the network size increases, the capture ratio decreases. This indicated that SLP protection provided improves as network size increases. This relationship can be observed across all three network configurations.

B. The impact of message rates on SLP with one source

In real world scenarios it is expected that different applications will have different requirements with respect to how often messages are sent from the source. This section will present results under varying broadcast rates and analyse the effect different rates have on the provision of SLP.

Intuition: The intuition behind this investigation is the following: Denote the set of nodes h hops away from the source by N^h, where h is the length of the random walk, and denote the message rate by r per unit time. If all the nodes in N^h can be reached independently by a similar number of paths, then the expected number of messages received by any node $n \in N^h$ per unit time is approximately $\frac{r}{|N^h|}$. Thus, if an attacker has reached the node n, the higher r is, the higher is the likelihood that the attacker will hear a message at n, hence will move one hop closer to the source. Applying this reasoning over h means that a higher r can cause the attacker to capture the source before the safety period elapses. Therefore, we conjecture that *as the value of r increases it will result in a higher capture ratio*.

Result: To address this conjecture, experiments were conducted for each network size and configuration. The message rate of the application was varied such that the number of messages transmitted per second was 1, 2, 4 or 8 sent from a single source in the network.

From Figure 3, it can be observed that, across all network sizes and configurations, an increase in the message rate leads to a corresponding increase in the capture ratio, thereby confirming the conjecture. As more messages are being sent in the same period of time, the attacker has a greater number of chances to move towards the source in response to a message.

We make the two observations regarding these results:

1) For a given message rate, as the network size increased the capture ratio decreased.
2) For any message rate and any network size, the SourceCorner configuration always yielded the highest capture ratio, and the FurtherSinkCorner configuration yielded the lowest capture ratio. We conjecture that, for SourceCorner, the high capture ratio is due to the fact the attacker is "funnelled" towards the source. The lower capture ratio for FurtherSinkCorner is possibly due to the fact that the random walk tends to lead the attacker "away" from the source. It is also the case that the source-sink distance is higher in the FurtherSinkCorner configuration than in other configurations.

C. The impact of the number of sources on SLP broadcasting at 1 message/second

Having shown the impact of message rates on the level of SLP with only one source, we now analyse the impact of multiple sources on capture ratio.

Intuition: If the sources are scattered over the network, then the attacker may perform poorly as it is trying to capture all the different sources rather than focusing on a single location. This scenario is likely to lead to a low capture ratio.

On the other hand, if sources are clustered together the opposite is to be expected. Hence, we wish to analyze the impact when (i) two sources are located in the network within two hops of each other, and when (ii) three sources are clustered together such that the nodes are within two hops of each other. For each of these source distributions, we conjecture that, given that the sources are close to each other, the net effect will be that the attacker will be drawn towards either of the sources, resulting in a higher capture ratio.

Result: From Figure 4, it can be observed that the increasing number of sources leads to a higher capture ratio, supporting the conjecture. This is due to the distribution of nodes in the network. The distributions considered, are likely to be realistic. For example, the triangular array might model three army personnel walking close to each other during a surveillance operation.

We make the following observations:

1) With a linear arrangement of sources, the capture ratio increased at a slower rate compared to the increase

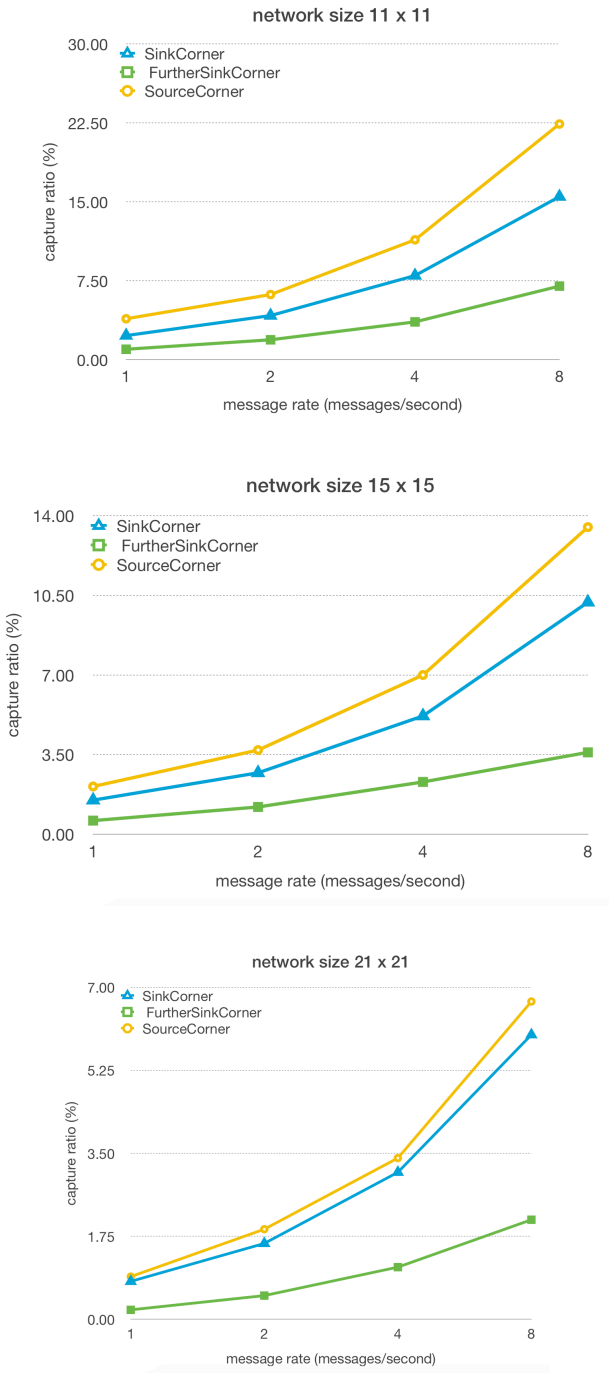


Fig. 3: The impact of varying message rates on SLP with 1 source

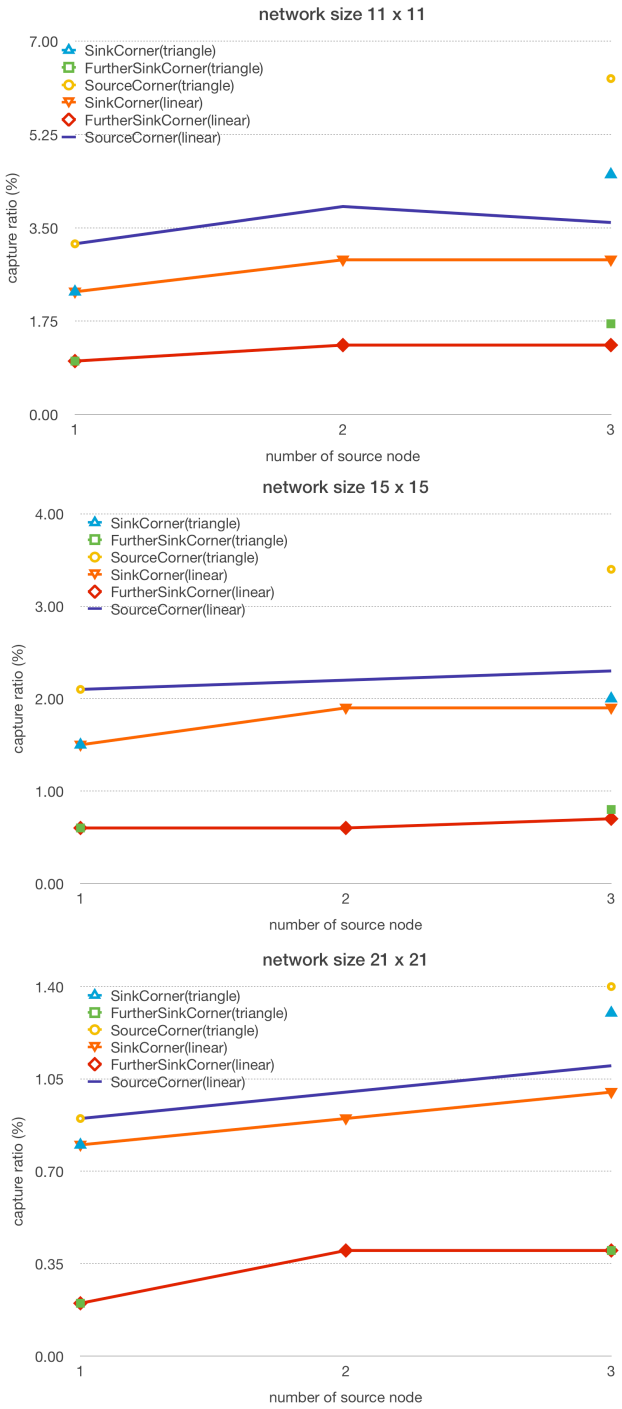


Fig. 4: The impact of varying the number of sources broadcasting at 1 message/second on SLP

seen for the triangular arrangement of the sources. The capture ratio of the linear arrangement remains the same (or sometimes lower) than two sources although both the sink-source distances are almost the same. But generally the capture ratio of such arrangement is higher than with a single source. The lower capture ratio is due to the fact that phantom nodes are scattered after random walk when the three nodes are distributed in a linear fashion rather than triangular arrangement.

2) Similar to the case with varying message rates, the larger the network, the lower the capture ratio.

D. The impact of the number length of random walk on SLP broadcasting at 1 message/second

The random walk is an important parameter that can be varied in Phantom Routing. This section examines the relationship between capture ratio and the length of the random walk.

Intuition: The intuition behind this investigation is the following: Denote the length of the random walk by h and denote the rate at which messages are transmitted per time unit by r. If the number of paths from a source to nodes in N^h is similar, then, we observed, the expected number of times a nodes will hear a message is per unit time if $\frac{r}{|N^h|}$. Thus, the mean time between two successive messages heard by an attacker is approximately $\frac{|N^h|}{r}$. Thus, the expected amount of time an attacker has to wait from reaching a phantom node h hops away from the source until it reaches the source, denoted by $\bar{T_c}$ is given by:

$$\bar{T_c} \approx \sum_{i=1}^{h} \frac{|N^i|}{r} \tag{3}$$

Therefore, we conjecture that the higher h is, the longer the attacker will have to wait to reach the source. This waiting time may then exceed the safety period, meaning that the source is not captured in time, thus reducing the capture ratio.

Result: To address this conjecture, experiments were conducted for each network size and configuration. The length of the random walk is varied to be either 3, 5, or 7 hops.

From Figure 5, it can be observed that, across all network sizes and configurations, an increase in the length of the random walk leads to a corresponding decrease in the capture ratio, thereby confirming the conjecture. The increase in SLP is not significant when the level of SLP is already high. We further notice that the SLP level provided is worse when there are three sources in the network than with one or two sources (Figure 6). However, the SLP level provided increases when the length of the random walk is increased.

E. The impact of multiple of source nodes and varying rate on SLP

So far results presented in this paper have observed that: (i) a higher message rate reduces the SLP level, (ii) a higher number of sources (and particularly their distribution) reduces the SLP, and (iii) a higher length of the random walk increases the SLP level. However, little is known regarding the importance of each individual parameters on the overall SLP level imparted and their interactions.

Intuition: When the message rate and number of sources are simultaneously increased, we expect the SLP level to decrease (i.e., capture ratio to increase). However, we are also interested in understanding the rate of decrease of SLP and which parameter contributes the greater decrease.

Result: To investigate the relationships between these parameters, we conduct experiments with (i) one and two sources,

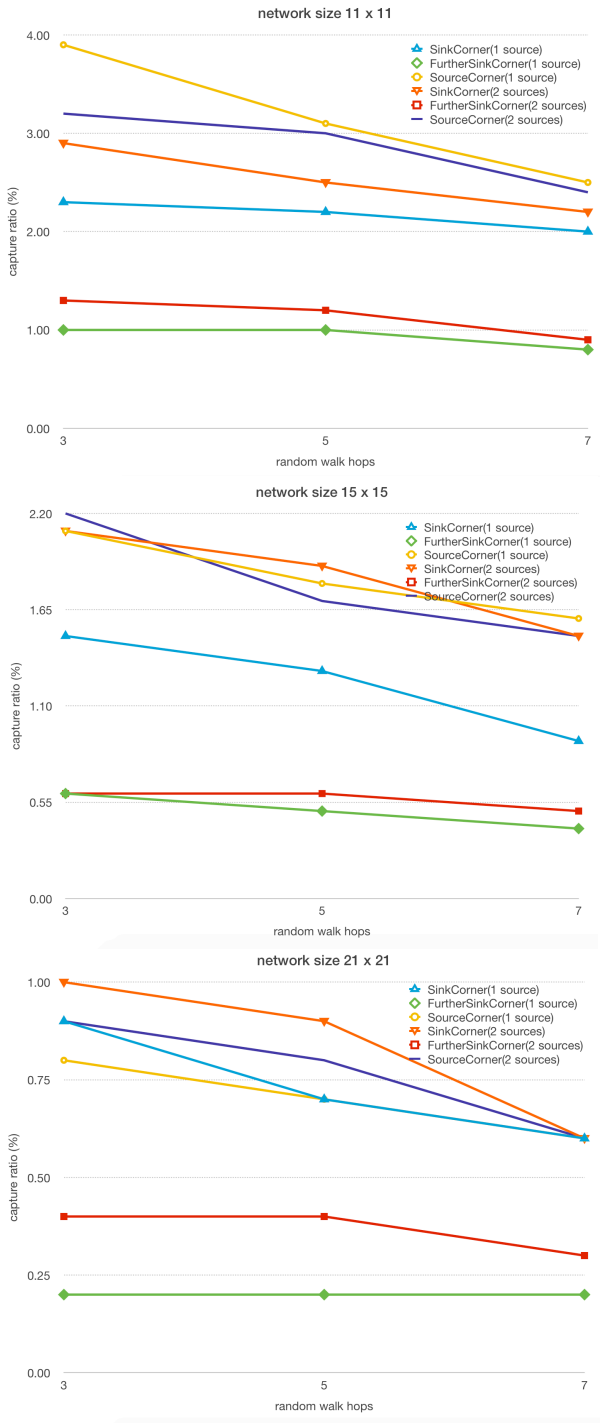


Fig. 5: The impact of varying random walk length on SLP with a single source broadcasting messages at 1 message/second

and (ii) three sources (for linear and triangular distributions). Figure 7 shows the results for the case of one and two sources and Figure 8 shows the result when three sources are present for multiple message rates.

From Figure 7, it can be observed that SLP decreases (increasing capture ratio) with two sources and higher message rates, which matches the intuition. It can also be observed that

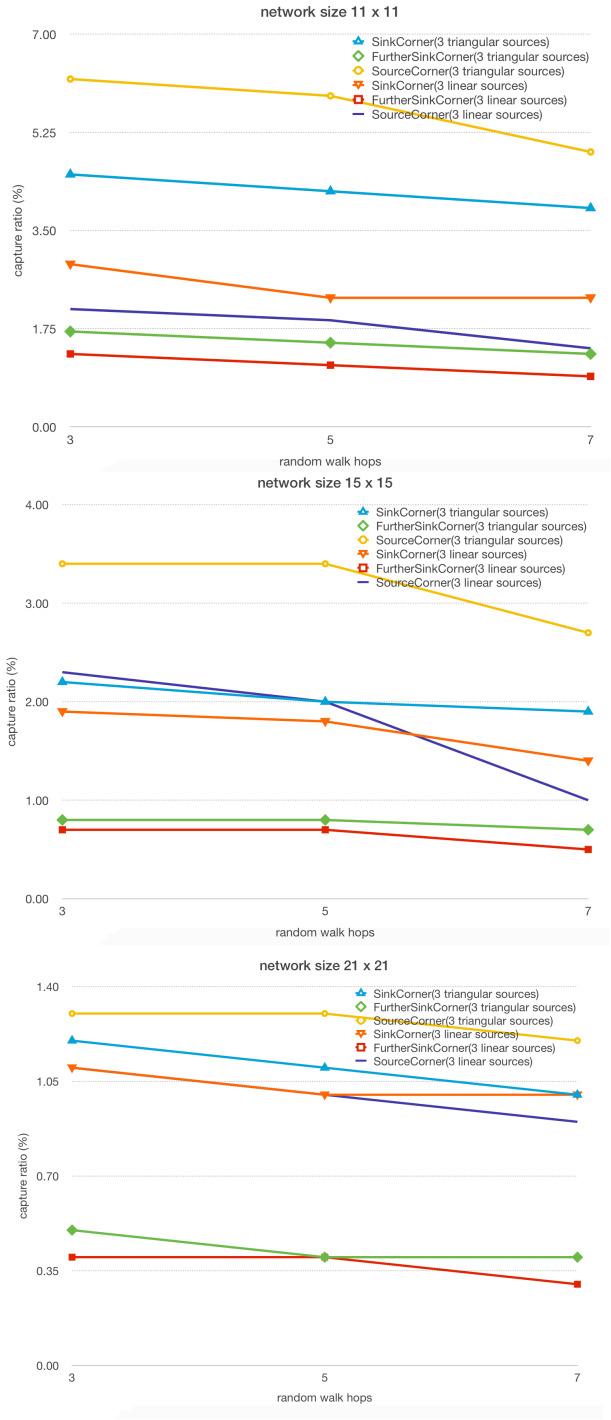


Fig. 6: The impact varying random walk length on SLP with 3 sources broadcasting messages at 1 message/second

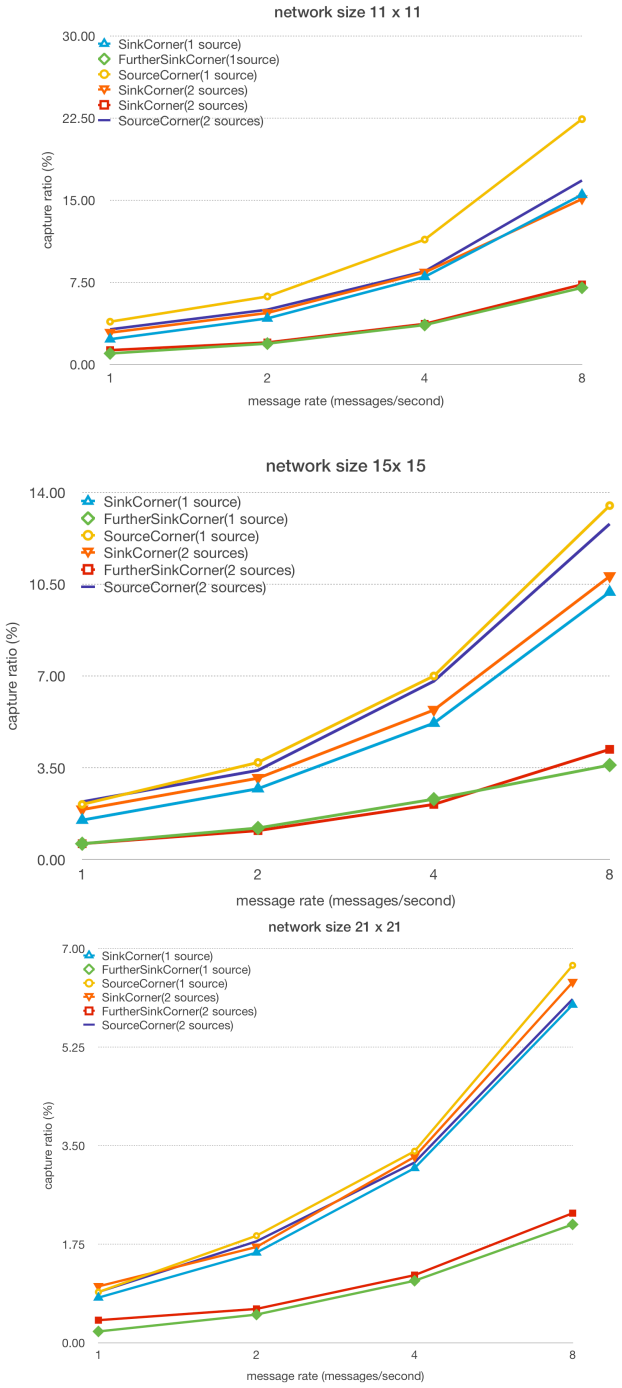


Fig. 7: The impact of 1 or 2 sources and varying broadcast rates on SLP

the SLP level matches the worst SLP level imparted between message rate and number of sources.

When three sources are considered in a triangular arrangement, it can be observed that the overall SLP level is worse than the SLP level of either higher message rate or two sources (see Figure 8). This is due to the fact that the triangular arrangement ”funnels” the attacker towards the source. It was also noticed that, in general, the SourceCorner configuration experienced the worst provided SLP levels among all configurations. This is likely to be due to the fact that it has the shortest sink-source distance of all three configuration tested with (see table I).

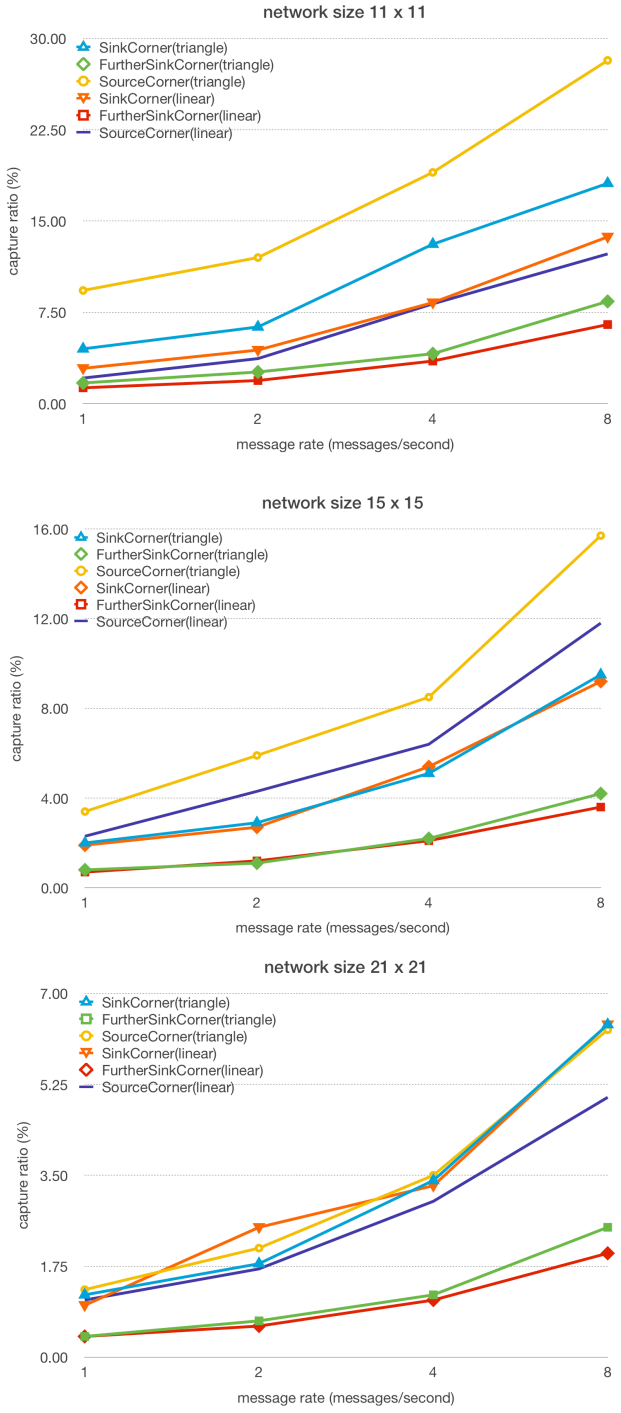


Fig. 8: The impact of broadcast rates on SLP with 3 sources

VII. Conclusion

In this paper, we have investigated the performance of phantom routing, a well-known algorithm that provides SLP in WSN, under various network scenarios. We have considered three application parameters: (i) message rates, (ii) number of sources and (iii) the length of the directed random walk. Our results show that (i) an increase in the message rates causes a decrease in the SLP level provided, i.e., the capture ratio increases, (ii) the number of sources also caused an increase in the capture ratio, while the triangular arrangement of sources caused the highest increase and (iii) the length of the random walk causes an increase in SLP level. We also looked at the combined effect of some of these parameters. Overall, our results show that the SLP levels of phantom routing can drop by up to a factor of 75, under some parameterisation. Our conclusion is that phantom routing is not as effective as initially claimed, as it was evaluated under a restricted set of circumstances and network configurations.

References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393 – 422, 2002.

[2] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "Spins: Security protocols for sensor networks," *Wirel. Netw.*, vol. 8, no. 5, pp. 521–534, Sep. 2002.

[3] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing source-location privacy in sensor network routing," in *Distributed Computing Systems, 2005. ICDCS 2005. Proceedings. 25th IEEE International Conference on*, June 2005, pp. 599–608.

[4] X. Hong, P. Wang, J. Kong, Q. Zheng, and jun Liu, "Effective probabilistic approach protecting sensor traffic," in *Military Communications Conference, 2005. MILCOM 2005. IEEE*, Oct 2005, pp. 169–175 Vol. 1.

[5] Y. Ouyang, Z. Le, G. Chen, J. Ford, and F. Makedon, "Entrapping adversaries for source protection in sensor networks," in *World of Wireless, Mobile and Multimedia Networks, 2006. WoWMoM 2006. International Symposium on a*, 2006, pp. 10 pp.–34.

[6] A. Jhumka, M. Bradbury, and M. Leeke, "Fake source-based source location privacy in wireless sensor networks," *Concurrency and Computation: Practice and Experience*, 2014.

[7] A. Jhumka, M. Leeke, and S. Shrestha, "On the use of fake sources for source location privacy: Trade-offs between energy and privacy." *Comput. J.*, vol. 54, no. 6, pp. 860–874, 2011.

[8] C. Ozturk, Y. Zhang, W. Trappe, and M. Ott, "Source-location privacy for networks of energy-constrained sensors," in *Software Technologies for Future Embedded and Ubiquitous Systems, 2004. Proceedings. Second IEEE Workshop on*, May 2004, pp. 68–72.

[9] Y. Yang, M. Shao, S. Zhu, B. Urgaonkar, and G. Cao, "Towards event source unobservability with minimum network traffic in sensor networks," in *Proceedings of the First ACM Conference on Wireless Network Security*, ser. WiSec '08. New York, NY, USA: ACM, 2008, pp. 77–88.

[10] M. Shao, Y. Yang, S. Zhu, and G. Cao, "Towards statistically strong source anonymity for sensor networks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, April 2008.

[11] J. Long, M. Dong, K. Ota, and A. Liu, "Achieving

source location privacy and network lifetime maximization through tree-based diversionary routing in wireless sensor networks," *Access, IEEE*, vol. 2, pp. 633–651, 2014.

[12] K. Mehta, D. Liu, and M. Wright, "Location privacy in sensor networks against a global eavesdropper," in *IEEE International Conference on Network Protocols, 2007. ICNP 2007.*, October 2007, pp. 314–323.

[13] M. Shao, Y. Yang, S. Zhu, and G. Cao, "Towards statistically strong source anonymity for sensor networks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, April 2008, pp. –.

[14] R. A. Shaikh, H. Jameel, B. J. DAuriol, H. Lee, S. Lee, and Y.-J. Song, "Achieving network level privacy in wireless sensor networks," *Sensors*, vol. 10, no. 3, pp. 1447–1472, 2010.

[15] C.-Y. Chow, M. Mokbel, and T. He, "A privacy-preserving location monitoring system for wireless sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 10, no. 1, pp. 94–107, Jan 2011.

[16] X. Luo, X. Ji, and M.-S. Park, "Location privacy against traffic analysis attacks in wireless sensor networks," in *Information Science and Applications (ICISA), 2010 International Conference on*, April 2010, pp. 1–6.

[17] M. Shao, W. Hu, S. Zhu, G. Cao, S. Krishnamurth, and T. La Porta, "Cross-layer enhanced source location privacy in sensor networks," in *Sensor, Mesh and Ad Hoc Communications and Networks, 2009. SECON '09. 6th Annual IEEE Communications Society Conference on*, June 2009, pp. 1–9.

[18] W. Wei-Ping, C. Liang, and W. Jian-xin, "A source-location privacy protocol in wsn based on locational angle," in *Communications, 2008. ICC'08. IEEE International Conference on*. IEEE, 2008, pp. 1630–1634.

[19] L. Zhang, "A self-adjusting directed random walk approach for enhancing source-location privacy in sensor network routing," in *Proceedings of the 2006 international conference on Wireless communications and mobile computing*. ACM, 2006, pp. 33–38.

[20] J. Yao and G. Wen, "Preserving source-location privacy in energy-constrained wireless sensor networks," in *Distributed Computing Systems Workshops, 2008. ICDCS'08. 28th International Conference on*. IEEE, 2008, pp. 412–416.

[21] Y. Xi, L. Schwiebert, and W. Shi, "Preserving source location privacy in monitoring-based wireless sensor networks," in *Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International*, April 2006, pp. 8 pp.–.

[22] A. Jhumka, "Crash-tolerant collision-free data aggregation scheduling for wireless sensor networks," in *Reliable Distributed Systems, 2010 29th IEEE Symposium on*, Oct 2010, pp. 44–53.

[23] Z. Benenson, P. M. Cholewinski, and F. C. Freiling, *Wireless Sensors Networks Security*. IOS Press, 2008, ch. Vulnerabilities and Attacks in Wireless Sensor Networks, pp. 22–43.

[24] P. Levis, N. Lee, M. Welsh, and D. Culler, "Tossim: accurate and scalable simulation of entire tinyos applications," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, ser. SenSys '03. New York, NY, USA: ACM, 2003, pp. 126–137.

[25] A. Jhumka, M. Bradbury, and M. Leeke, "Towards understanding source location privacy in wireless sensor networks through fake sources," in *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*, June 2012, pp. 760–768.