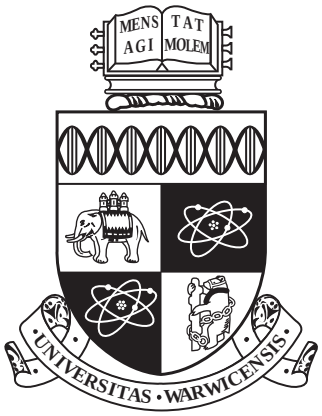


Source Location Privacy in Wireless Sensor Networks Under Practical Scenarios: Routing Protocols, Parameterisations and Trade-Offs

by

Chen Gu

A thesis submitted to The University of Warwick

in partial fulfilment of the requirements

for admission to the degree of

Doctor of Philosophy

Department of Computer Science

The University of Warwick

April 2018

Abstract

As wireless sensor networks (WSNs) have been applied across a spectrum of application domains, source location privacy (SLP) has emerged as a significant issue, particularly in security-critical situations. In seminal work on SLP, several protocols were proposed as viable approaches to address the issue of SLP. However, most state-of-the-art approaches work under specific network assumptions. For example, *phantom routing*, one of the most popular routing protocols for SLP, assumes a single source. On the other hand, in practical scenarios for SLP, this assumption is not realistic, as there will be multiple data sources. Other issues of practical interest include network configurations. Thus, thesis addresses the impact of these practical considerations on SLP. The first step is the evaluation of phantom routing under various configurations, e.g., multiple sources and network configurations. The results show that phantom routing does not scale to handle multiple sources while providing high SLP at the expense of low messages yield. Thus, an important issue arises as a result of this observation that the need for a routing protocol that can handle multiple sources. As such, a novel parametric routing protocol is proposed, called *phantom walkabouts*, for SLP for multi-source WSNs. A large-scale experiments are conducted to evaluate the efficiency of phantom walkabouts. The main observation is that phantom walkabouts can provide high level of SLP at the expense of energy and/or data yield. To deal with these trade-offs, a framework that allows reasoning about trade-offs needs to develop. Thus, a decision theoretic methodology is proposed that allows reasoning about these trade-offs. The results showcase the viability of this methodology via several case studies.

Acknowledgements

First, I would like to express my sincere gratitude to my supervisor Dr. Arshad Jhumka, whose guidance, encouragement and support have been invaluable to me during my time in the Department of Computer Science at the University of Warwick. I benefited greatly from his insightful advice and comments in finding and solving research problems. Your advice on my research has been priceless. I look forward to maintaining our collaboration in the future.

I want to thank my fellows in the SLP lab, particularly Matthew Bradbury and Jack Kirton, who have helped develop me academically and professionally. Also I would like to thank my lab mates: Dr. Bo Gao, Dr. Huanzhou Zhu, Dr. Bo Wang, Dr. Chao Chen, Peng Jiang, Dr. Zhuoer Gu, Junyu Li and Shenyuan Ren, for their stimulating discussions in current technology trends and for creating all the happy memories that we shared over the last four years.

I would like to thank the staff members in the Department of Computer Science for giving me tremendous help, both personally and academically over the past five years. In particular, I thank Dr. Roger Packwood for his support in solving technical issues and Sharon Howard who solves massive administrative issues for me.

Last but not the least, I want to express my eternal gratitude to my parents, who have always given me unconditional love and support through my life. I would not have undertaken higher education if it were not for my parents, who have always been there for me.

Declarations

This thesis is submitted to the University of Warwick in support of the authors application for the degree of Doctor of Philosophy. It has been composed by the author and has not been submitted in any previous application for any degree. The work presented was carried out by the author except where acknowledged.

For the purpose of reproduction, the source codes of protocols to run simulations are available at [52] and datasets used to generate results in this thesis can be found at [53]. More details can be found in the Appendix A.

Parts of this thesis have been previously published by the author in the following:

[54] C. Gu, M. Bradbury, A. Jhumka, and M. Leeke. Assessing the Performance of Phantom Routing on Source Location Privacy in Wireless Sensor Networks. In *IEEE 21st Pacific Rim International Symposium on Dependable Computing*, pages 99–108, 2015. ISBN 9781467393768. doi: 10.1109/PRDC.2015.9. URL `https://doi.org/10.1109/PRDC.2015.9`

[55] C. Gu, M. Bradbury, and A. Jhumka. Phantom walkabouts in wireless sensor networks. In *Proceedings of the ACM Symposium on Applied Computing*, pages 609–616, 2017. doi: 10.1145/3019612.3019732. URL `https://doi.org/10.1145/3019612.3019732`

[56] C. Gu, M. Bradbury, J. Kirton, and A. Jhumka. A decision theoretic framework for selecting source location privacy aware routing protocols in wireless sensor networks. *Future Generation Computer Systems*, 2018. ISSN 0167739X. doi: 10.1016/j.future.2018.01.046. URL `https://doi.org/10.1016/j.future.2018.01.046`

Research was performed in collaboration during the development of this thesis,

but does not form part of the thesis:

[82] J. Laikin, M. Bradbury, C. Gu, and M. Leeke. Towards fake sources for source location privacy in wireless sensor networks with multiple sources. In *IEEE International Conference on Communication Systems*, pages 1–6, 2016. doi: 10.1109/ICCS.2016.7833572. URL `https://doi.org/10.1109/ICCS.2016.7833572`

Abbreviations

6LoWPAN	IPv6-based Low Power Wireless Personal Area Networks
CEM	Cyclic Entrapment Method
CLS	Cross Layer Solution
CPM	Closest Pattern Matching
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
CTP	Collection Tree Protocol
DAS	Data Aggregation Scheme
DBT	Dynamic Bidirectional Tree
DCS	Data-Centric Sensor Network
DROW	Directed Random Walk
DT	Decision Theory
FCF	Frame Control Field
GPS	Global Positioning System
GROW	Greedy Random Walk
IEEE	Institute of Electrical and Electronics Engineers
ILP	Integer Linear Programming
IoT	internet of Things
IP	Internet Protocol
LPSS	Location Privacy Support Scheme
MAC	Medium Access Control
NMR	Network Mixing Ring
PEM	Path Extension Method
PFS	Permanent Fake Source
PHR	Physical Header
PRESH	Probabilistic Reshaping
PRLA	Phantom Routing with a Locational Angle

PRS	Phantom Routing Scheme
PSRS	Phantom Single-Path Routing Scheme
PW	Phantom Walkabouts
RAM	Random-access Memory
RF	Radio Frequency
RRIN	Routing to a Randomly Intermediary Node
RRS	Random Routing Scheme
SADRW	Self Adjusting Directed Random Walk
SLFSR	Short-Lived Fake Source Routing
SLP	Source Location Privacy
SPIN	Sensor Protocols for Information via Negotiation
STaR	Sink Toroidal Region
TDMA	Time-Division Multiple Access
TFS	Temporary Fake Source
WSN	Wireless Sensor Networks
ZBT	Zigzag Bidirectional Tree

Symbols

Src	The Source node
$Sink$	The Sink node
msg	The normal message
$\mathcal{S}_{dir}$	The random walk set of a message
$\mathcal{M}_{dir}$	The random walk direction of a message
$\mathcal{B}_{dir}$	The biased random walk direction of a message
$\mathcal{P}_{biased}$	The probability of biased random walk
ψ	The safety factor
$\mathcal{TT}$	The time taken (seconds) of protectionless flooding
P_{safety}	The safety period (seconds)
M_s	The message with the short random walk
M_l	The message with the long random walk
Δ_{ss}	The distance in hops between the sink and the source
h_{walk}	The remaining hops of the random walk
$\mathcal{NC}$	The network configuration
$\mathcal{P}$	The name of a given routing protocol
$r_{\omega}^{\mathcal{NC},\mathcal{P}}$	The result of a attribute under $\mathcal{NC}$ and $\mathcal{P}$
$R_{\omega}^{\mathcal{NC},\mathcal{P}}$	The normalised result of a attribute under $\mathcal{NC}$ and $\mathcal{P}$
$r^{\mathcal{NC},\mathcal{P}}$	The result vector of all attributes under $\mathcal{NC}$ and $\mathcal{P}$
$R^{\mathcal{NC},\mathcal{P}}$	The performance vector of all attributes under $\mathcal{NC}$ and $\mathcal{P}$
$U_{\omega}^{\mathcal{NC},\mathcal{P}}$	The utility of a single attribute under $\mathcal{NC}$ and $\mathcal{P}$
$U^{\mathcal{NC},\mathcal{P}}$	The utility of the performance vector under $\mathcal{NC}$ and $\mathcal{P}$
u_a	The aspiration vector
λ_{ω}	The weight of a single attribute

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

The Internet is rapidly developing which gradually removes the digital barrier between the Internet and the physical world [123]. In the near future, not only the computing devices but objects such as cars or even people will be connected to the Internet. These objects equipped with tiny processors and radio transceivers known as sensor nodes or motes, can sense different attributes of the environment and use radio signals to communicate among themselves. The wireless sensor network (WSN) is one of such technologies and a vital component of the sensing technology, dealing with vast amounts of information for further processing and analysis [4]. WSNs have enabled the development of many novel applications [6], including asset monitoring [5], target tracking [41] and environment control [101] among others, with low level of intrusiveness. They are also expected to be deployed in the safety and security-critical systems including military [10] and medical services [97].

As WSNs have been applied across a spectrum of application domains, they increase the complexity and challenges in both academia and industry, especially on data security and privacy. Specifically, the problem of source location privacy (SLP) has emerged as a significant issue, particularly in security-critical situations. There is a need to provide SLP because it has been shown that a malicious attacker can trace the monitored assets by eavesdropping messages in the WSNs. In many situations, it is very important to hide the physical location of objects which originally send messages. Mastering the above aspects with the awareness of the different potential issues becomes one of the most critical research topics in the further advance of WSNs.

1.1 Wireless Sensor Networks

1.1.1 Overview

WSNs are highly distributed systems consisting of a number of tiny devices, known as sensor nodes or motes, that can sense physical phenomena and use radio signals to communicate among themselves (see Figure 1.1). The device responsible for generating data is called the *source*. There is another device called the *sink* (or the base station): a powerful device that gathers and processes all the information collected by the sensor nodes. The sink serves as an interface between the sensor nodes and the users. Sensor nodes are fitted with a large variety of physical sensors (e.g., temperature, pressure, humidity and radiation), for the purpose of monitoring, tracking and controlling environments and assets. In fact, WSNs have enabled the development of many novel applications in agriculture and farming [81, 98], environmental monitoring [93, 101, 104], medical services [27], and military applications [10].

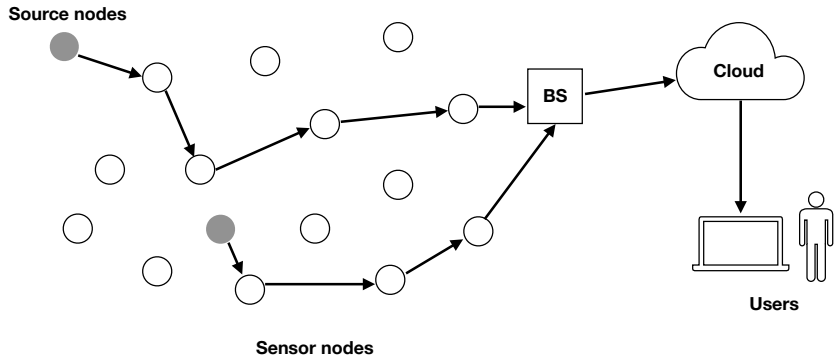


Figure 1.1: Demonstration of wireless sensor networks

1.1.2 Components of Nodes

The sensor nodes in WSNs normally consist of four essential components [5]: the sensing unit, the processing unit, the transceiver and the power unit (see Figure 1.2). The sensing unit consists of a series of physical sensors that provides the node with the ability to sense different environmental conditions.

The processing unit is composed of simple 32/64-bit microprocessors which have limited computational capabilities (typically between 8 and 25 MHz) and memory space (typically between 4 and 10 kB for RAM). The transceiver (or radio interface) allows the sensor node to send and receive messages at a low data rate (between 70 and 250 kB/s) usually in the 2.4 GHz unlicensed industrial, scientific, and medical (ISM) radio band of the radio spectrum. Choosing between one band or another depends on the application scenario. Communications in higher frequency bands have a longer range but find it difficult to overcome obstacles. Lastly, the power unit provides energy to all the other components to ensure they operate well. The power unit usually uses two AA batteries (i.e., 3V) as an energy supply, thus it is regarded as the most limiting component in sensor nodes as they cannot be replaced or recharged (without other powers that recycle energy) once the network has been deployed. In addition, sensor nodes may be equipped with other optional components depending on the practical scenarios, such as localisation systems (e.g., GPS chips), power scavengers (e.g., solar panels) and external flash memories.

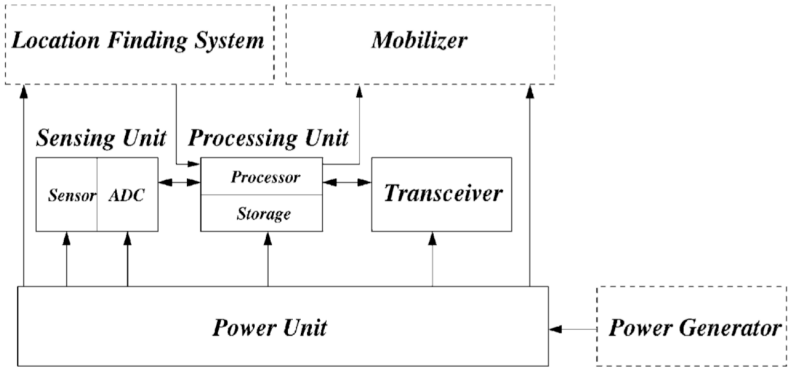


Figure 1.2: The components of a sensor node [6]

1.1.3 Communication Methods

In WSNs, data reporting methods could be time-driven, query-driven, event-driven, or hybrid of these methods [7]. An event-driven approach is the most

usual one because of energy efficiency. In the event-driven model, a sensor node starts reporting data to the sink immediately after an event has been detected (e.g., a sudden change of environment). Instead of establishing a direct communication link to the sink due to high transmission power consumption, the source uses multi-hop communications to deliver data which are transmitted through multiple intermediate nodes. If no events or transmission tasks are detected, the node becomes inactive (sleep) mode to save energy.

To fulfil these multi-hop communications, there are two basic protocols adopted to meet the demands of data transmission: flooding and single-path routing (SPR). Flooding is the simple routing algorithm where the node forwards data to all its neighbouring nodes except the one that sent it. The intermediate nodes repeat the process until data visits all the nodes in the network. The advantage of the flooding protocol is that it is very reliable due to massive data redundancy. However, it is also very energy inefficient because all the nodes are involved in transmitting data. On the contrary, the single-path routing protocol is intended to minimise the number of relaying nodes used to reach the sink. In the single-path routing, whenever a node has event data to transmit, it sends the message to a neighbouring node which is closer to the sink than itself. This operation is repeated for each of the nodes until the data is finally delivered to the sink. Additionally, some sensor networks may take advantage of data-aggregation protocols [46, 61, 67, 152] to further reduce network traffic on its way to the sink. Data aggregation consists of a set of operations (e.g., counting, average, maximum, minimum) that are performed at some intermediate points of the network to combine data originating from different sources. Processing data at intermediate nodes results in more energy-efficient data communication, thereby increasing the overall lifetime of the network [123].

1.1.4 Protocol Stack

The protocol stack in WSNs has five layers: the physical layer, the data link layer, the network layer, the transport layer and the application layer [5]. The physical

layer addresses the needs of simple but robust modulation, transmission, and receiving techniques. It is responsible for frequency selection, carrier frequency generation, signal detection, signal processing and data encryption. The data link layer is responsible for the multiplexing of data streams, data frame detection, medium access control (MAC) and error control. It ensures reliable point-to-point and point-to-multipoint connections in a communication network. The network layer takes care of routing the data supplied by the transport layer. It is responsible for specifying the assignment of addresses and how packets are forwarded. The transport layer helps to maintain the flow of data if the sensor networks application requires it. This layer is especially needed when the system will be accessed through the Internet or other external networks. Depending on the sensing tasks required, different types of application software can be built and used on the application layer.

1.1.5 Standardisation

The standardisation has been become a key requirement for the development of wireless sensor networks. The standards define the functions and protocols necessary for sensor nodes in the sensor networks [99].

IEEE 802.15.4 Low Rate WPANs

IEEE 802.15.4 is the proposed standard for low rate wireless personal area networks (LR-WPAN's) [19, 58]. This standard is designed for wireless sensor applications that require short-range communication and low power consumption. IEEE 802.15.4 standard asks devices following the agreed physical and data-link layer protocols. For instance, the physical layer supports 868/915 MHz low bands and 2.4 GHz high bands; The MAC layer controls access to the radio channel using the CSMA/CA mechanism. The IEEE 802.15.4 standard also allows the formation of the star and peer-to-peer topology for communication between network devices. In the star topology, the communication is performed between network devices and a single central controller, called the PAN coordinator.

A network device is either the initiation point or the termination point for network communications. The PAN coordinator is in charge of managing all the star PAN functionality. In the peer-to-peer topology, every network device can communicate with any other within its communication range. The PAN coordinator acts as the root in the network. The peer-to-peer topology allows more complex network formations to be implemented such as ad hoc networks.

ZigBee

The ZigBee standard was publicly available in 2005 [92]. It defines the higher layer (i.e., above network layer) communication protocols upon on the IEEE 802.15.4 standards for LR-PANs. On the physical and data link layer, ZigBee adopts the IEEE 802.15.4 standard for LR-WPANs. ZigBee also defines mesh, star and cluster tree network topologies with data security features and application profiles [99]. ZigBee meets the unique needs of sensors and control devices, typically with low bandwidth, low latency and very low energy consumption for long battery lives and for large device arrays.

6LoWPAN

IPv6-based low power wireless personal area networks (6LoWPAN) enables IPv6 packets communication over an IEEE 802.15.4 based network [107]. With the benefit of the standard, low power devices can communicate directly with IP devices using IP-based protocols. As the IPv6 packet size is much larger than the frame size of IEEE 802.15.4, an adaptation layer, new packet format, and address management are used in the 6LoWPAN standard. 6LoWPAN is designed for applications with low data rate devices that require Internet communication.

1.2 Source Location Privacy in Wireless Sensor Networks

1.2.1 Classification of Privacy Issues

WSNs have earned acceptance, and extensive work has been done on their development [28, 57]. However, privacy protection has received a lack of attention, and it is necessary to consider and address all potential privacy risks that may arise from the adoption of this technology. Threats to privacy in WSNs can be considered along two dimensions, content privacy and context privacy (see Figure 1.3) [85]. Content privacy threats relate to use of the content of the messages broadcast by sensor nodes, such as gaining the ability to read an encrypted message. Content privacy thus focuses on providing integrity, freshness, non-repudiation and confidentiality of the messages exchanged in the WSN. In particular, content privacy includes data aggregation privacy and query privacy. Data aggregation is designed to substantially reduce the volume of traffic in the WSN by fusing or compressing data in the intermediate sensor nodes. However, if intermediate sensor nodes are compromised, an adversary may decrypt the transmitted data, inject bogus data or tamper with raw data, thus compromising the content privacy. There are serval techniques for privacy-preserving data aggregation [59, 119, 158]. In addition, an adversary can also infer client interests with enquire leak, causing query privacy. Anonymity techniques [20, 157] are mainly used to address this privacy issue.

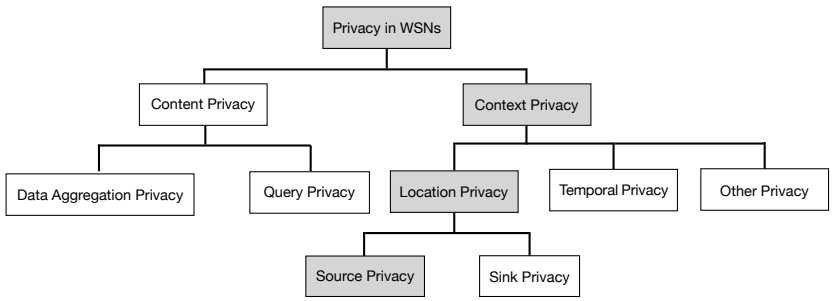


Figure 1.3: Classification of privacy issues in the wireless sensor networks

On the other hand, context privacy threats focus on the context in which messages are broadcasted and how information can be observed or inferred by attackers. Context privacy comprises of hiding the identity, location of nodes and traffic flow in the WSN. Context is a multi-attribute concept that encompasses the situational aspects of broadcasted messages, including environmental and temporal information. Location privacy may arise for such special sensor nodes such as the source [74, 105, 126, 144, 148] and the sink [32, 33, 72]. It is often desirable for the source of sensed information to be kept private in the WSN. In addition, temporal privacy concerns the time when sensitive data is created at the source, collected by a sensor node and delivered to the sink [75]. This type of privacy is also very important, because an adversary with knowledge of such timing information may be able to pinpoint the location of the tracked target without having the knowledge of data being transmitted in the WSN [75].

1.2.2 Why Provide Source Location Privacy?

For the location privacy, let us use a panda-hunter game [115] as an example. In a WSN, a node that senses a panda (e.g., temperature changes) informs the sink that a change has occurred by sending messages that travel through intermediate nodes to the sink. Poachers attempt to identify the location of the data source to find the panda. Poachers often with a local vision of network communications can act in the following way to find the panda: They start from any point of the network[1] and move around. They are equipped with devices capable of measuring the arrival angle of received signals, which can estimate the location where the messages sent from. Then poachers move on to the nodes and repeat the process until they reach the location of panda. As the movement follows the path of communication, this is usually referred to as traceback attack. Similar problems occur in other applications. For example, in a military application, a soldier transmitting messages can unintentionally

[1] Usually adversaries are assumed to start from the sink, as they can observe any incoming communication.

disclose their location, even when encryption is used. Other real-world examples include monitoring badgers [41] and the WWF's Wildlife Crime Technology Report [1], both of which would likely benefit from context security measures. In this thesis, the context this thesis focuses on protecting is that of *source location*.

Techniques that protect source location are said to provide source location privacy (SLP). The SLP problem focuses on ensuring that the location of a source node or asset can only be observed or inferred by those intended to observe or decipher it [74]. SLP is important in many application domains, though it is of utmost concern in security-critical situations. The importance of SLP is not in the protection of hardware of itself, but the need to hide the presence of events in the field. In each of these scenarios, it is important to ensure that an attacker cannot find or deduce the location of the asset being monitored, whether it is an endangered animal or a soldier. In the panda protection example, poachers have the local view of the network, meaning that they can monitor a limited range of messages transmitted. On the other hand, a more powerful adversary called a global adversary who has a global view of the network uses its sniffers to eavesdrop all communication. It has been shown that in a non-SLP protected network, even a weak attacker such as a distributed eavesdropping attacker [71] can backtrack along message paths through the network to find the source node and capture the asset [74]. Thus, there is a need to develop SLP-aware routing protocols.

1.2.3 Formalisation of Source Location Privacy Problem

The SLP problem was first formalised based on the panda-hunter game [74]. In the WSN, the purpose of the network is to monitor the source, while the purpose of the routing strategy is two-fold, to deliver messages to the sink and to enhance the location privacy of the asset in the presence of an adversarial attacker following a movement strategy. This model is formalised containing six-tuple $(G, Sink, Src, \mathcal{P}, \mathcal{A}, \mathcal{M}_{\mathcal{A}})$, where:

- $G = (V, E)$ defines the network graph where V represents the set of sensor

nodes, and E is a set of communication links connecting two distinct nodes.

- $Sink$ is the network sink, to which all communication in the sensor network must ultimately be routed to. Typically there is only *one* sink in the WSN.
- Src is an asset (i.e., the source) that the sensor network monitors.
- $\mathcal{P}$ is the routing protocol employed by the sensors to protect the asset from being acquired or tracked by the attacker $\mathcal{A}$.
- $\mathcal{A}$ is the attacker, or hunter, who seeks to acquire or capture the asset Src through a set of movement rules $\mathcal{M}_{\mathcal{A}}$.

The following chapters will expand on this representation and explain aspects of the panda-hunter game further in the literature. Section 2.1 will detail the network model including G, $Sink$ and Src. The threat model including $\mathcal{A}$ and $\mathcal{M}_{\mathcal{A}}$ will be described in Section 2.2. The routing protocols $\mathcal{P}$ will be reviewed in Section 2.3.

1.3 Problem Statement and Research Contributions

Privacy is a key issue in WSNs. Aimed at providing SLP, a number of techniques have been proposed such as phantom routing using random walks [74, 140], message delay [64], fake sources [69, 115] and others [31, 113, 124, 127]. Most existing researchers mainly focus on proposing protocols to provide SLP but are constrained by a lack of in-depth investigation of SLP in practical scenarios such as multiple sources and different network configurations.

This thesis aims to handle the SLP issue with multiple sources and different network configurations. The problem statement is: **Could routing protocols protect SLP under multiple sources and various network configurations in the WSN?** To address the SLP issue, the thesis investigates it in

terms of routing protocols, parameterisations and trade-offs. Specifically, phantom routing is evaluated under multiple sources and various configurations. The results show some shortcomings of phantom routing such as low SLP with multiple sources. Then a novel parametric routing protocol is proposed, called *phantom walkabouts*, for SLP in WSNs. Phantom walkabouts provides high level of SLP with multiple sources at the expense of data yield. Finally, a decision theoretic methodology that allows reasoning about these trade-offs is proposed. This thesis makes the following main contributions:

1. **Assessing the performance of phantom routing on source location privacy under practical scenarios**

   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 scenarios, omitting simulations and analyses with multiple sources and different network configurations. These shortcomings above are addressed by conducting an in-depth investigation of phantom routing under multiple sources and two different network configurations. Simulations are conducted by varying four parameters: (i) random walk length, (ii) source period, (iii) network size and (iv) number of sources. The results demonstrate that previous work in phantom routing does not provide a high level of SLP with multiple sources and does not generalise well to different network configurations.

2. **Developing phantom walkabouts to achieve high level of SLP**

   Because recent work has shown some limitations of phantom routing such as poor performance with multiple sources, phantom walkabouts is proposed, a novel and more general version of phantom routing, which performs routes of variable lengths. Phantom walkabouts addresses several shortcomings of phantom routing such as unexpected termination of the random walk and poor SLP under a certain network configuration. Parameterisations are varied in phantom walkabouts to analyse the impact on SLP and other

performance attributes including receive ratio, messages sent and message latency. Through extensive simulations, the results show the viability of phantom walkabouts. For example, under certain parameterisations, phantom walkabouts achieves extremely high SLP with acceptable decrease in other attributes.

3. **Developing a decision theoretic framework for selecting SLP-aware routing protocols**

   Routing protocols such as phantom routing and phantom walkabouts have been proposed that provide SLP, all of which provide a trade-off between SLP and other performance attributes. Experiments have been conducted to gauge the performance of the proposed protocols under different network parameters such as network sizes. As there exists a plethora of protocols which contain a set of possibly conflicting performance attributes, it is difficult to select the SLP protocol that will provide the best trade-offs across them for a given application with specific requirements. For example, the phantom walkabouts provides high level of SLP at expense of the receive ratio. However, the decrease of the receive ratio may be not acceptable for some scenario such as military applications. Therefore, a decision theoretic procedure is proposed for selecting the SLP-aware routing protocol that achieves the best trade-offs for the applications and network configurations. The results show the viability of the approach through different case studies.

More detailed summaries of the contributions are given at the end of Chapter 3, Chapter 4 and Chapter 5.

1.4 Thesis Organisation

The rest of this thesis is organised as follows. **Chapter 2** reviews various topics related to the thesis including system models, threat models, existing SLP-aware

routing protocols with different techniques, other existing context privacy issues, simulators and performance attributes.

Chapter 3 presents an in-depth investigation of the phantom routing protocol under practical scenarios. Phantom routing is implemented with multiple sources and various network configurations and then assessed by conducting a range of experiments.

Chapter 4 presents phantom walkabouts, a routing protocol that with variable random walk lengths. Phantom walkabouts aims to lead an adversary roaming around in the network, hence keeping the source location safe. Simulations are conducted by varying parameterisations of short and long random walks in phantom walkabouts, and the results will show better SLP performance than phantom routing.

In **Chapter 5**, a methodology is proposed where routing protocols are first profiled to capture their performance according to a desired set of attributes, and then a decision theoretic procedure is used for selecting the most appropriate SLP-aware routing protocol for the type of network and application under study. The results demonstrate the viability of the approach through various case studies.

Chapter 6 summarises the thesis and discusses further work.

CHAPTER 2

Literature Review

As one aim of the thesis is to assess and develop SLP-aware routing protocols, this chapter first reviews various system models and threat models used in such protocols in Section 2.1 and Section 2.2. System models investigate the contents including the sensor nodes, network configurations and the message structure. For threat models, they focus on the capability of an adversary in WSNs. Under the definition of system models and threat models, Section 2.3 then explores some existing methodologies that solve the SLP problem. Some other aspects of privacy issues in WSNs are also considered as complementary knowledge of SLP in Section 2.4. Furthermore, Section 2.5 presents popular simulations and testbeds used to test WSNs. Finally, performance attributes mostly used in experiments are reviewed in Section 2.6.

2.1 System Models in Wireless Sensor Networks

To investigate the SLP problem, there is a need to specify and model wireless sensor networks. To present a clear understanding of a network, the system models are considered from three aspects: (i) sensor nodes, (ii) network and (iii) the message structure.

2.1.1 Sensor Nodes Modelling

In much of the literature, nodes are randomly deployed in WSNs [25, 29, 42, 94, 122, 136, 138]. They collect data from the environment and send data to the sink(s). Specifically, when an object appears at a location monitored by a sensor node, the node becomes the source node and will send messages destined for the sink. Any sensor can become a source node as long as it has something

to report to the sink [72, 129]. When the object moves to a new location, it may trigger another sensor node to send messages, and that node then becomes the source [113, 141]. However, some authors assume that the source location is stationary, i.e., the source does not move in the network [69, 70, 71]. All the nodes have a limited radio range and nodes within the range can either send or receive from each other [42, 64, 72, 96, 109, 113, 114]. Because of their limited radio range, nodes send their data to the sink using multi-hop communication. Some energy-efficient MAC protocols (e.g., IEEE 802.11) allow nodes to detect packets while in idle mode [95, 108]. There is only one sink in the network [30, 76], but it can be either static or mobile [72]. Some authors assume that the sink may have other capabilities. For instance, the sink knows the network configuration and is able to monitor the energy consumption and remaining battery power of every node [141]. In the WSN, nodes know their location and have knowledge of the location of their adjacent neighbouring nodes through GPS [3, 88, 127, 148]. On the other hand, some authors assume that nodes do not have GPS capability as localisation services consume too much energy [15, 17, 69, 70, 71]. Instead, the knowledge of their relative location can be obtained by broadcasting beacon packets sent from the sink [72].

Nodes in the networks are not only classified into three categories: the source, normal nodes and the sink [32, 59, 63, 69, 70]. Instead, there are some other special types of nodes in the networks. Ekici et al. [42] assume that there are a small number of verifier nodes, which have the responsibility of verifying the location of sensor nodes and a small number of malicious nodes which possess the same properties as regular sensor nodes. Another authors [43] assume a hybrid wireless sensor network with anchor, trusted, and untrusted nodes. Trusted nodes can utilise standard encryption algorithms to hide an anchor nodes' positional information where both anchor nodes and trusted nodes share required common information. Untrusted nodes use the same radio hardware used by anchor nodes and trusted nodes. Li et al. [86] assume there is another type of nodes called data mules in the network. They are mobile agents, moving independently which

do not communicate with each other. They perform random walk movements on the grid, whereby each transition produces a move of equal probability to a horizontally or vertically adjacent cell.

2.1.2 Network Modelling

Most authors presume that a WSN is composed of finite two-dimensional grids (cells) [3, 17, 69, 71, 82, 86, 87, 88, 89, 127, 148]. In the grid network configuration, Bradbury et al. [17] and Laikin et al. [82] consider the *SourceCorner* configuration where the single source locates in the corner and the sink is in the centre. Jhumka et al. [71] consider other configurations called the *SinkCorner* configuration and the *FurtherSinkCorner* configuration. In the SinkCorner configuration, the sink is at the corner of the grid, while the source is at the centre. The FurtherSinkCorner configuration is similar to the SinkCorner configuration, except that the source is slightly offset from the centre.

On the other hand, other authors presume a tree configuration that of a topological tree rooted at the sink [34, 35, 149]. They also presuppose the sink cannot be compromised and it has a secure mechanism. In a tree communication model, the root is the sink receiving data from the leaf nodes which simply act as routers. Besides, Wadaa et al. [136] and Yang et al. [150] assume that a network is partitioned into a number of clusters through a training process. In this configuration, a high-end device is deployed into each cluster, acting as the cluster head. In contrast to sensor nodes, high-end cluster heads have relatively higher computation capabilities, larger storage sizes, and longer radio ranges. For the communication in the network, authors [87, 88, 89, 108] assume bidirectional links only, meaning two nodes are considered neighbours if they can hear each other and the whole network is fully connected through multi-hop communications. However, authors [15, 17, 69, 70, 71] do not claim that links are bidirectional, i.e., links may disappear intermittently.

Authors [69, 70, 134] formalise the network as an undirected graph $G = (V, E)$, where the set of vertices V represents the set of N wireless sensor nodes and

the set of edges E represents the set of links between the nodes. Two nodes $m \in V$ and $m^{'} \in V$ are said to be 1-hop neighbours iff $\{m, m^{'}\} \in E$, i.e., m and $m^{'}$ are in each other's communication range. The graph $G = (V, E)$ defines the topology of the network with network of size $n \times n = N$.

2.1.3 Message Structure

For the scope of context privacy, most authors assume that the source encrypts messages and messages are decrypted at the sink [30, 37, 64, 69, 73, 87, 88]. As a consequence, the contents of messages will not leak out to prevent an adversary decrypting or modifying the contents. The encryption procedure can be achieved by using a shared secret key between the nodes and the sink. However, the contents of key management [21, 38, 44, 146, 155] including key generation, distribution and update are beyond the scope of the thesis.

Messages continue to be sent periodically for a certain period, and will stop when the object leaves the sensor's monitoring area [72, 113, 129, 141]. They contain information both in the header and the payload. The header information is used at every hop for the routing purpose and thus contain information about the sender and recipient of the message. The payload contains the information of the monitored object reported by the source. Figure 2.1 is an example of the CC2420 packet header structure which can be explained as follows [2]: The PHR contains frame length information; The FCF is the frame control field defined in the IEEE 802.15.4 specifications and the CC2420 data sheet; The $\#seq$ is the data sequence number, which is incremented for each packet sent by a particular node. This is used in acknowledging that packet, and also filtering out duplicate packets; The destination PAN ensures the network can sit side by side with another TinyOS network and not interfere; The destination is the address of a packet; The source is the local source ID; The 6LowPAN is the TinyOS network ID for the 6LowPAN TinyOS Network layer; The AM type defines the type of a packet.

In the network, all messages are transmitted in the same format and have the

Figure 2.1: TinyOS 2.x header format [121]

same length [37, 64, 141] and Sheng and Li [128] only consider one-dimensional data. Each message includes a unique ID where the source event is generated [17, 89, 137]. Therefore, normal nodes can understand whether the messages have been received and the sink can determine the source node location based on the ID.

2.2 Threat Models in Wireless Sensor Networks

This section provides an overview of adversarial capabilities that are listed in several threat models considered in the literature. In particular, this section reviews the threat models from four aspects of an adversary: (i) adversarial behaviour, (ii) view of the network, (iii) resources strength and (iv) network knowledge.

2.2.1 Adversarial Behaviour

An adversarial behaviour could be either active or passive. An active adversary could use positive behaviour to interfere with traffic flow or communication behaviours by injecting, modifying or blocking messages [60, 63, 64, 120, 124]. For instance, Shaikh et al. [124] describe how an active adversary uses traffic analysis attacks to track an asset. Hong et al. [64] describe another active adversary that is capable of compromising a node to block traffic and to monitor traffic flow around nodes. He et al. [60] mention data pollution attack where an adversary tampers with intermediate aggregation results to make the sink receive the wrong aggregation results.

However in most cases, adversaries are considered as passive in the litera-

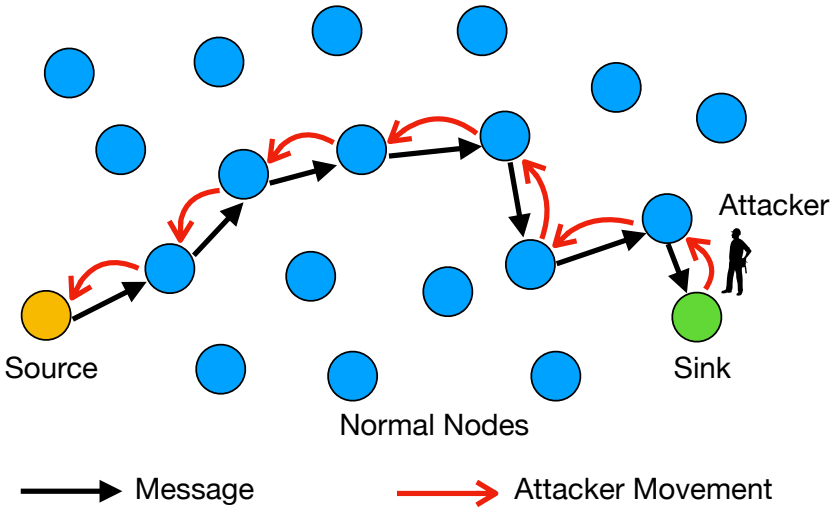


Figure 2.2: The procedures of the hop-by-hop traceback attack

ture [3, 29, 40, 45, 65, 72, 109, 116, 127, 130, 151]. A passive adversary does not actively influence the nodes or the traffic between nodes. Ozturk et al. [116] and other authors [140, 148] describe a typical behaviour of a passive adversary as follows. The adversary starts at the location of the sink and only eavesdrops on the traffic flow between nodes instead of manually changing it. When discovering an event of a message transmitted in its monitoring area, the adversary uses an attack strategy that the adversary follows the traffic between the nodes and traces back in reverse order until reaching the source. This trace strategy is called a hop-by-hop traceback attack and described in Figure 2.2 and Algorithm 1.

The hop-by-hop traceback attack only monitors traffic flow in the network, while some other attacks focus on other aspects of the network. Some more advanced attacks include rate monitoring attack [148], time correlation attack [148] and timing analysis attack [102]. In the rate monitoring attack, the adversary monitors nodes with a higher transmission rate, as intuitively these nodes are probably close to the source or the sink. In the time correlation attack, the adversary observes the correlation in transmission time between a node and its neighbour to find the route that a message travels to the sink. Finally, the timing analysis attack is used to monitor transmission patterns to discern sensitive

Algorithm 1 A Passive Adversary Strategy: Hop-by-Hop Traceback Attack

```
1: procedure HOP-BY-HOP TRACEBACK ATTACK(sink, source, msg)
2:     next_location ← sink
3:     while next_location ≠ source do
4:         LISTEN(next_location)
5:         msg ← RECEIVEMESSAGE()
6:         if ISNEWMESSAGE(msg) then
7:             next_location ← CALCULATEIMMEDIATESENDER(msg)
8:             MOVETO(next_location)
9:         end if
10:    end while
11: end procedure
```

information, such as the structure of the network and traffic flow [64]. These advanced attacks may require an adversary having the capability of monitoring a larger part of the network [37].

2.2.2 View of the Network

There are two types of adversaries when it comes to their network perspectives: the local and the global adversary. A local adversary has a local view of the network [32, 124]. Eavesdropping can be achieved using signal detection devices or other sniffers [65]. For simplicity, authors assume the adversary has a hearing radius equal to the sensor transmission radius [72, 76, 151]. A local adversary is sometimes not alone and might collaborate with others. Jhumka et al. [71] describe a threat model that involves multiple local adversaries collaborating by sharing information on the configuration and traffic in the WSN. Together these collaborating adversaries are also regarded as having a multi-local view of the network. Li and Ren [87] assume that there are some adversaries in the target area.

On the other hand, a global adversary has a full view of the network [3, 37, 45, 105, 108, 111, 114, 135, 147]. The adversary often uses its own network with sniffers to eavesdrop on all communications happening in the network. A global adversary is more powerful than a local adversary due to more knowledge of network configurations and traffic flow. Normally, SLP solutions that defend against a local adversary cannot cope with a global adversary, whereas solutions

designed against a global adversary can often deal with a local adversary [9].

2.2.3 Resources Strength

A threat model often describes the amount of resources an adversary has in terms of energy source, memory, move speed and computational capability [36, 87, 124, 127, 132]. The energy source determines whether adversaries can travel freely in the network. The adversary records data from messages tracking, so they need memory for data storage. The move speed is often considered with a passive adversary because passive adversaries often act when hearing message transmitted. Computational power for an adversary is used to track messages by calculating the directions of incoming messages or decrypting messages. In the literature, an adversary is considered as mobile with an unlimited amount of power [116]. Kamat et al. [74] define the adversary as device rich and resource rich. Device rich adversaries have the ability to assess the strength of the signal and determine the angle of arrival of a signal, for example by measuring the difference in the receiving phase of each element of an antenna array [103]. Meanwhile, resource-rich adversaries can move at any rate and has an unlimited amount of power. Besides, they also have a large of memory to store information such as messages that have been received before and nodes they have previously travelled. Jian et al. [72] also mention that an adversary has memory to remember his path and performs backtracking.

It is often assumed that strong adversaries have the ability to decrypt the contents of a message. However, in terms of context privacy, they do not have the keys to decipher the messages they overhear, so an adversary cannot obtain the contents of the message [65, 70, 86, 100, 134]. Therefore, some attack strategies (e.g., clone attack [154]) related to cryptology will not be discussed further.

2.2.4 Network Knowledge

The network knowledge of the adversary varies in the literature. Kamat et al. [74, 75] define an informed adversary who knows the location of the

sink and algorithms used in the network to protect the panda. Wang et al. and other authors [76, 110, 131, 140] also assume that the adversary knows the location of the sink and starts tracking from it. However, Deng et al. [32] assume that an adversary cannot see the sink visually in a large network. Jhumka et al. [70] assume that the adversary knows (i) the location of the sink, (ii) the network configuration and (iii) the routing algorithm. However, the attacker does not know the number of assets being monitored, and the possible location of the assets. An attacker also learns about the 1-hop neighbourhood of different nodes, depending on its location within the network. Besides, the adversary knows, not only the routing algorithm, but the protection strategy being used in the network [72, 141].

2.3 SLP-Aware Routing Protocols in Wireless Sensor Networks

The concept of the SLP problem was first introduced around 2004 [116] which proposed the panda-hunter game where the poachers only used network traffic flow to track a panda. Kamat et al. [74] formalise the SLP issue based on the panda-hunter game. Since then, many techniques have been proposed to address SLP. The solution spectrum ranges from simple solutions such as pure random walk [116] to more sophisticated techniques, such as fake sources [17, 71] and message delay [15, 64]. This section discusses two main categories of solutions: random-walk based techniques and fake-source based techniques. Other types of solutions that provide SLP in the literature are also reviewed. For each algorithm, its theory, strength and weakness will be discussed as well.

2.3.1 Random-Walk Based Techniques

Ozturk et al. [115] use the random walk as a technique to provide SLP. In the phantom routing scheme (PRS), there are two phases: (i) the random walk phase which is a pure or directed random walk, meant to deliver the message to a

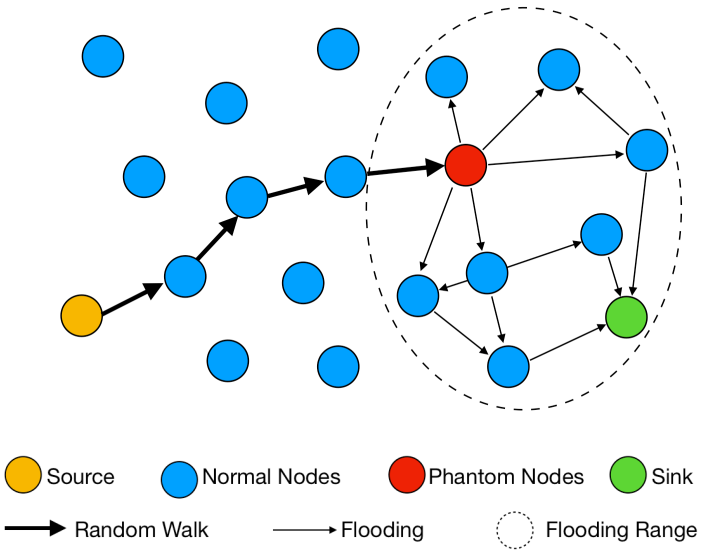


Figure 2.3: Illustration of phantom routing scheme

phantom source after travelling h_{walk} hops, and (ii) a subsequent flooding meant to deliver messages to the sink (see Figure 2.3). Ozturk et al. discuss the pure random walk in the phantom routing in detail and claim that the phantom node is within 20% of h_{walk} from the real source after h_{walk} hops (see Figure 2.4a). Then Ozturk et al. propose the directed random walk that avoids random walks cancelling each other out (see Figure 2.4b). Both sector-based directed random walk and hop-based directed random walk could guarantee phantom sources far away from the true source. Instead of using flooding for the second phase, Ozturk et al. also use single path routing algorithms, such as shortest path routing. The combination of the random walk together with single path routing is often referred to as the phantom single-path routing scheme (PSRS). Both PRS and PSRS has received a lot of attention in the literature. On the other hand, this class of solutions is known to have weaknesses [89, 124, 138], ascribing poor SLP performance to the directed random walk reusing the routing path and exposing direction information. Zhang [156] introduces an improved algorithm of a sector-based directed random walk called self-adjusting directed random walk (SADRW). Instead of dividing neighbours into two sets in the phantom routing,

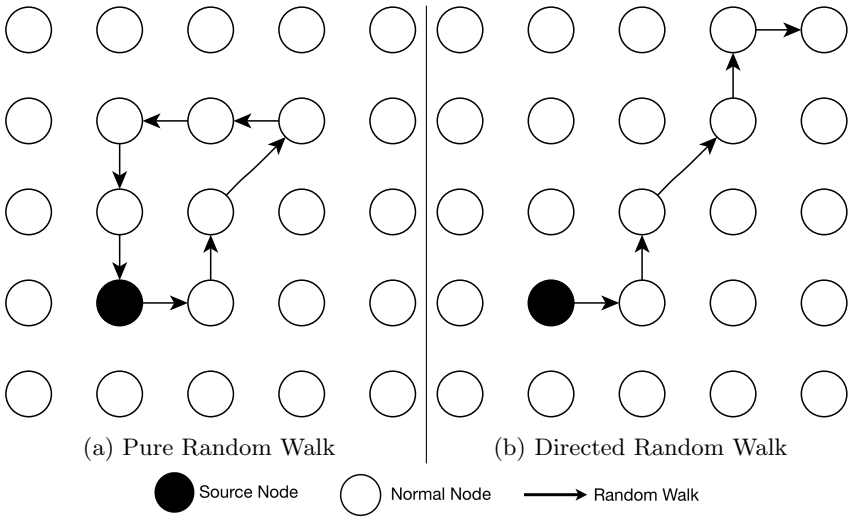


Figure 2.4: Illustration of pure random walk and directed random walk

in SADRW neighbours are divided into four different sets. Nodes randomly pick a neighbour out of one of the four directional sets and send messages to it. If an intermediate node receives the message and cannot forward it to the same direction, then it chooses a new direction to forward the message to until the message travels a total of h_{walk} hops. SADRW solves the weakness of a phantom walk that may unexpectedly terminate before h_{walk} hops, hence increasing the SLP level.

Wang et al. [140] introduce phantom routing with a locational angle (PRLA). In PRLA, the random walk is based on the inclination angle between a node and its neighbours towards the sink. PRLA works as follows. In the deployment stage, every node calculates the inclination angle between itself and its neighbours. Then, every node uses the inclination angle to calculate the forward probability of each of its neighbours. The higher the inclination angle of a neighbour, the higher the forward probability of that neighbour will be. The source sends messages to neighbours by using the forward probability of the neighbours. After the message travels h_{walk} hops or the last node is not able to forward the message with the same inclination angle, the message is forwarded to the sink using a single path routing strategy.

Yao and Wen [151] provide another improvement by introducing the directed random walk (DROW). In DROW, every node has the knowledge of its own hop-distance to the sink and the hop-distance of its neighbours to the sink. Each node chooses the neighbours with a lower hop-distance towards the sink than its parent's. When sending messages, the node randomly chooses one of its parents as the next destination. Authors claim several advantages applying to DROW such as routing diversity, long safety period [1] and energy efficiency. However, Deng et al. [32] show that DROW does not defend against a time correlation attack. Wang and Hsiang [139] mention that the direction information retrieved from the packet headers helps the adversary to find the source of messages.

Xi et al. [144] introduce the greedy random walk (GROW). In the GROW, one random walk starts from the sink and goes to a randomly chosen receptor-node. The other random walk starts from the source and meets the first random walk at the receptor-node. Then the receptor-node uses the path established by the random walk from the sink to the receptor-node to route the packet from the source to the sink. In addition, the authors use a different approach, by recording neighbours in a bloom filter which informs the choice of the next node to be used in the random walk. However, there is still scope to improve nodes that are allocated to take part in the directed random walk. Yao and Wen [151] point out that the random walk used in the GROW is inefficient at creating a safe distance between the receptor-node and the source. Wang et al. [140] state that the latency is unstable due to the usage of two random walks. Other weakness can be found from [89, 122, 124].

An algorithm called randomly selected intermediary node (RRIN) is introduced by Li et al. [88] as an improvement over PRS. Unlike PRS, RRIN does not leak any directional information via its messages. A source node sends a message to a chosen intermediate node, and the intermediate node sends the message to the sink. The choice of the intermediate node must meet the following criteria: the location of the intermediary node must be at least a minimum distance away

[1] The notion of *safety period* will be introduced in Section 2.6.

from the source and be normally distributed within the rest of the network. The authors claim that RRIN has the same latency and power consumption as PRS, but a higher safety period. Then Li et al. propose a second version of choosing intermediate nodes. Each node in the WSN has an equal probability to be the intermediate node of any given source node. The second version of RRIN consumes much more energy [89] and has a higher delay than PRS, but it does provide an even better safety period.

There are other algorithms using random walk techniques to address the SLP issue such as the random routing scheme (RRS) [96], location privacy support scheme (LPSS) [76] and network mixing ring (NMR) [87]. As random walk is one of the early techniques used to provide SLP, they could only defend against a local adversary. In fact, some solutions in the literature have discussed weaknesses, which shows that the random walk is not always effective. Therefore, algorithms need to be developed to guard against a powerful adversary with a global view.

2.3.2 Fake-Source Based Techniques

Algorithms utilise dummy messages sent by a *fake source* to provide SLP. Some nodes are chosen as fake sources and periodically send dummy messages to obfuscate the real traffic. In the early stages of this technique, Ozturk et al. [115] introduce the concept of fake sources and propose a theoretical algorithm called short-lived fake source routing (SLFSR). The solution works as follows. If a node receives a real message, it generates a probability p to decide whether to send a dummy message. If p is below a threshold P, then the node broadcasts a dummy message to all its neighbours. SLFSR consumes more energy but it could improve the safety period. However, both Kamat et al. [74] and Ozturk et al. [115] state that only one fake source at the time for only one dummy packet is not enough to distract an adversary.

Chen and Lou [23] provide two solutions called dynamic bidirectional tree (DBT) and zigzag bidirectional tree (ZBT). They both use the fake-source

technique to confuse attackers, in such ways that the attackers are not sure if they are tracking real traffic from the source, or following dummy traffic. In DBT, each node knows its distance to the sink and of its neighbours to the sink. The source randomly sends messages to neighbours with a shorter or equal hop-distance, which works similar to the first stage of phantom routing. Then intermediary nodes use a probability p to randomly select a neighbour to create a branch and forward dummy traffic for h hops. The second solution (ZBT) makes messages walk zigzags in the network. Firstly the sink generates one proxy sink with each of its sides. Then the source randomly selects a node as a proxy source which is i hops away from itself. The real traffic is following the route that messages are from the source, to the proxy source, to the proxy sink, and finally to the sink.

Jhumka et al. [71] propose another algorithm. Jhumka et al. first prove the fake sources selection problem to be NP-complete and the algorithm works as follows. The source node sends a normal message to the sink. When the sink receives it, it waits a short period and broadcasts $\langle away \rangle$ messages that floods the network. When a 1-hop neighbour of the sink receives the $\langle away \rangle$ message it becomes a temporary fake source (TFS) and broadcasts $\langle fake \rangle$ messages for a period. Before the TFS becomes a normal node they broadcast $\langle choose \rangle$ messages. When a normal node receives the $\langle choose \rangle$ message it becomes a permanent fake source (PFS) if the node believes itself to be the furthest node in the network from the sink, otherwise it will become a TFS.

Bradbury et al. [17] improve the algorithm in [71] through the online estimation of its parameters. As the consequence, the improved algorithm provides a better SLP level than [71] without requiring prior network knowledge. Then Bradbury et al. propose DynamicSPR which is an extended version of the dynamic fake source technique [17]. DynamicSPR optimises the way fake sources are allocated, in such a way that fake sources perform a directed random walk away from the sink. This algorithm reduces the number of fake sources present in the network and also the number of messages the technique sends (thus reducing

energy usage).

Many other algorithms have been proposed with state-of-the-art fake-source techniques [69, 70, 102]. However, these algorithms based on the fake-source techniques mentioned so far can only provide SLP against a local attacker. For the scope of the global attacker, a global protection scheme called Periodic is developed in which every node sends a message after a fixed period [105]. This provides perfect protection against an attacker with a global view of the network. The authors [105] create a model involving traces of source detection, which is used to measure the privacy of those traces as well as the energy cost of providing SLP. In addition, a different approach where statistical techniques are used to show that their global protection scheme provides high level of SLP [126]. Nodes use slotted transmission and send a packet at each interval. If a node does not have a real message to send at a slot, it sends a dummy message. This approach does not provide perfect global SLP as [105] does, but instead provides statistically strong SLP. Their model and solution aim to make the distribution of message broadcasts from nodes indistinguishable from a certain statistical distribution.

Other techniques consist of a hybrid between generating fake messages and having messages modify their routing path. Tree-based diversionary routing [94] imposes a tree structure on the network using fake sources at leaf nodes, with a focus on using the minimal energy possible at nodes 1-hop from the sink node to lengthen the networks lifetime. Similarly, fog or cloud techniques [36, 100] have been proposed to provide SLP where a normal message is routed through a group of nodes called a fog and then onwards to other fogs.

Perhaps the most significant disadvantage of the described fake-source based techniques is the volume of messages broadcast to provide SLP. This leads to increased energy consumption and an increased number of collisions, both of which result in a decreased message receive ratio. Another issue is that fake-source based techniques can perform poorly with multiple sources due to collisions between fake messages [82]. This means that a trade-off between energy

expenditure and privacy must be made [69], making dummy message schemes challenging for many large-scale networks.

2.3.3 Other Techniques

Apart from random-walk based and fake-source based algorithms described above, some algorithms are reviewed with other techniques.

Geographic Routing

Geographic routing algorithms use the physical position of the nodes to route messages from the source to the sink through WSNs. Shaikh et al. [124] propose identity, route and location privacy (IRL) to provide SLP. In IRL, every node classifies its neighbours as trustworthy, uncertain or untrustworthy neighbours. Messages are sent with priority to the trustworthy neighbours until they reach the sink. The weakness of this algorithm is that it can introduce cycles leading to high latency and a low receive ratio. Lightfoot et al. [90] introduce sink toroidal region routing (STaR) that adopts the notion of a random intermediary node to generate a path from the source to the sink. STaR is an improvement over RRIN, but it has limitations such as the assumption of networks consisting of small grids and message loss [90]. Both IRL and STaR provide SLP against a local adversary.

Message Delay

Generally, the message delay changes the traffic flows or patterns by holding incoming messages for a random time before forwarding them. As a consequence, nodes alter the chronological order of the messages, making it hard for a local adversary to track the traffic to the actual source. Hong et al. [64] introduce probabilistic reshaping (PRESH) to counter timing analysis attacks. The source sends messages directly towards the first intermediary node, and then randomly delays the transmission for a small amount of time following an exponential distribution. Meanwhile, there can be many other transmissions in a node's

neighbourhood during the delay period. Therefore, the adversary cannot tell where the message came from, as there are too many other transmissions that an adversary could hear. However, PRESH lacks of a routing algorithm [108]. Integer linear programming (ILP) routing [15] uses a delay strategy where nodes group received messages together and delay message transmitted by the distance of the sink to itself. As a consequence, the attacker cannot make as much forward progress towards the source. The ILP routing algorithm can provide near optimal SLP but causes high message latency [15].

Cyclic Entrapment

Algorithms in this category aim to confuse the adversary by shaping the traffic between nodes in cyclic patterns. The adversary which tracks traffic between the nodes will travel in circles without finding the actual source. The works in [113] and [133] contribute to the notion of cyclic entrapment method (CEM) and path extension method (PEM) respectively. CEM aims to trap the attacker in a cycle instead of letting them find the source node whereas PEM draws the attacker away using extended paths that broadcast fake messages. Another solution information hiding in distributing environments (iHIDE) is propose by Kazatzopoulos et al. [78]. In iHIDE, only bus nodes can activate the virtual cyclic loop node and nodes do not use probabilistic forwarding, but use time to live (TTL) counters and probability to regulate the ring traffic. Both of these two solutions use loops, where messages are sent to confuse a local adversary.

Cross-Layer Routing

Nodes normally use the network layer to exchange messages that contain information on sensed events. The solutions discussed thus far mostly concentrate on activities at the network layer. Different from the solutions discussed so far, underlying layers of the communication protocol stack could be used to provide SLP in a different way. Rather than utilising the network layer to exchange messages about events, authors [127] use control messages from the medium

access control (MAC) layer as well. In the cross-layer solution (CLS), nodes use the payload field of the beacon frames to exchange data [127]. The beacon frames might go unnoticed if the adversary does not listen to them and only checks the network layer for traffic. The weakness is that the beacon frame interval causes too much latency. Kirton et al. [80] develop a TDMA MAC schedule that can provide SLP. The 3-stage protocol works as follows. The algorithm first generates a normal data aggregation schedule (DAS). Then, it searches for a suitable location in the network where the attacker can be tricked for some time. Finally, the trick is to reassign slots to some nodes to ensure that the attacker takes a longer route towards the source, thus delaying it. Kirton et al. claim that the simulation results show great improvement at the SLP level when SLP-aware DAS is used with little messages overhead.

2.3.4 Summary of SLP-Aware Routing Protocols

So far previous sections have investigated many SLP-aware routing protocols with different techniques. However, there has been no universal solution proposed to deal with all types of adversaries. In general, the performance of SLP algorithms depends on the assumed system model and threat model. table 2.1 lists all the algorithms discussed above with specific techniques and threat models.

2.4 Other Context Privacy Issues

In the location privacy domain, some authors have investigated location privacy problems except SLP [13, 22, 32]. For instance, in order to protect sink location privacy from a powerful adversary with a global view, Chai et al. [22] propose the k-anonymity algorithm so that at least k entities in the network are indistinguishable to the nodes around the sink with regard to communication statistics. Then Chai et al. design a generic-algorithm-based quasi-optimal (GAQO) method and an artificial potential-based quasi-optimal (APQO) method to obtain optimal solutions. Chen and Lou [24] claim the location privacy of both the source

Solution	Technique	Adversarial Behaviour	View of Network	Knowledge	Weakness
PRS PSRS [115]	Random Walk	Eavesdropping & hop-by-hop traceback	Local	Location of sink	Random walk stops [89, 124, 138]
SADRW [156]	Random Walk	Eavesdropping & hop-by-hop traceback	Local	NA	NA
PRLA [140]	Random Walk	Eavesdropping & hop-by-hop traceback	Local	Location of sink	NA
DROW [151]	Random Walk	Eavesdropping & hop-by-hop traceback	Local	Location of sink	Time correlation attack [32, 139]
GROW [144]	Random Walk	Eavesdropping & hop-by-hop traceback	Local	NA	Path repetition [89, 122, 124, 140, 151]
RRIN [88]	Random Walk	Eavesdropping & hop-by-hop traceback	Multi-local	NA	Energy consumption [89]
SLFSR [115]	Fake Sources	Eavesdropping & hop-by-hop traceback	Local	Location of sink	Weak safety period [74, 115]
DBT ZBT [23]	Fake Sources	Eavesdropping & hop-by-hop traceback	Local	The protocols	NA
[71]	Fake Sources	Eavesdropping & hop-by-hop traceback	Local	Location of sink	Fixed parameters [17]
[17]	Fake Sources	Eavesdropping & hop-by-hop traceback	Local	Location of sink	Energy consumption [18]
DynamicSPR [18]	Fake Sources	Eavesdropping & hop-by-hop traceback	Local	Location of sink	No low power listening [18]
Periodic [105]	Fake Sources	Eavesdropping & Traffic analysis	Global	The topology	NA
IRL [124]	Geographic Routing	Eavesdropping & hop-by-hop traceback	Local	Location of sink	High latency [124]
STaR [90]	Geographic Routing	Compromise node Traffic analysis	Local	NA	High packet drop rate [90]
PRESH [64]	Delay	Eavesdropping & hop-by-hop traceback	Local	NA	No routing algorithm [108]
ILP [15]	Delay	Eavesdropping & hop-by-hop traceback	Local	Location of sink	Very high latency [15]
CEM [113] PEM [133]	Cyclic Entrapment	Eavesdropping & hop-by-hop traceback	Local	NA	NA
iHIDE [78]	Cyclic Entrapment	Eavesdropping & hop-by-hop traceback	Local	The topology	NA
CLS [127]	Cross-layer	Eavesdropping & hop-by-hop traceback	Local	Location of sink	High latency [127]
[80]	Cross-layer	Eavesdropping & hop-by-hop traceback	Local	Location of sink	NA

Table 2.1: Summary of SLP-aware routing protocols

and the sink becomes a critical issue in WSNs. Chen and Lou solve this issue by proposing four schemes: forward random walk (FRW), bidirectional tree (BT), dynamic bidirectional tree (DBT) and zigzag bidirectional tree (ZBT) to protect end-to-end location privacy against a local eavesdropper by using fake routes. Another solution proposed by [118] use a mechanism for preserving the anonymity of sources and sinks against global eavesdroppers.

SLP is mostly defined as a part of context privacy [100]. In general, context privacy often includes identity privacy, location privacy, timing privacy, and route privacy. Identity privacy ensures that the identity of nodes remains hidden. Location privacy focuses on the hidden location of a node. Timing privacy ensures that the temporal relation is hidden between incoming and outgoing traffic. Route privacy hides route flow in the network. Li et al. [85] provide a slightly different taxonomy to describe context privacy. Li et al. mention SLP as data source location privacy and divide context privacy between location privacy and temporal privacy. Location privacy is then further divided into data source location privacy and base station location privacy. Kamat et al. [75] discuss temporal privacy, which hides timing information related to packet creation. Shaikh et al. [124] define three elements of privacy required to achieve full network level privacy: (i) node identity privacy, (ii) route privacy and (iii) data privacy. More works are mentioned in [26, 31, 117].

Finally, the data-centric sensor network (DCS) is a specific solution in the literature. There is a demand for efficient data dissemination techniques to find relevant data, leading to the development of DCS. Data-centric WSNs normally are not designed to provide SLP and use different hardware than normal WSNs. However, saving data inside a network also creates security problems. Shao et al. [125] propose a privacy-enhanced DCS network solution called *p*DCS which offers different levels of data privacy based on different cryptographic keys for data-centric sensor networks.

2.5 Simulators and Testbeds

A variety of WSN simulators exist, many with different features. This section covers a number of the simulators considered for the simulations performed in this work.

2.5.1 TOSSIM

TinyOS is a flexible, embedded, component-based operating system and platform for low-power wireless devices, such as those used in wireless sensor networks [84]. The design motivations of the TinyOS are to deal with: (i) limited resources, (ii) reactive concurrency, (iii) flexibility and (iv) low power. These features are supported because the OS is written in a C dialect called nesC [49] optimised for the memory limits of sensor networks. TinyOS supports an event-driven concurrency model based on split-phase interfaces, asynchronous events, and deferred computation called tasks.

Codes written for TinyOS can be run in the TOSSIM simulator and also on a variety of hardware. TOSSIM is a discrete event simulator capable of accurately modelling sensor nodes and the modes of communications between them. TOSSIM provides two noise traces to model the environment: (i) The *meyer-heavy* noise sample collected from the Meyer library from Stanford and (ii) the *casino-lab* noise sample collected from Colorado schools of mines. The main difference is that meyer-heavy trace creates a very noisy environment whereas casino-lab trace generates a more quiet environment [83]. Instead of compiling a TinyOS application for a node, users can compile it into the TOSSIM framework, which runs on a laptop. Thus, this allows users to debug, test and analyse algorithms in a controlled and repeatable environment. However, TOSSIM has several disadvantages: (i) TOSSIM only supports codes written for TinyOS, and (ii) Only one hardware platform (MICAz) is supported.

2.5.2 COOJA

COOJA is a novel simulator for the Contiki operating system [39] that enables cross-level simulation: simultaneous simulation at many levels of the system. COOJA combines low-level simulation of sensor node hardware and simulation of high-level behaviour in a single simulation. COOJA is flexible and extensible in that all levels of the system can be changed or replaced: sensor node platforms, operating system software, radio transceivers, and radio transmission models [112].

The simulator is implemented in Java, making the simulator easy to extend for users, but allows sensor node software to be written in C by using the Java native interface. Furthermore, the sensor node software can be run either as compiled native code for the platform on which the simulator is run, or in a sensor node emulator that emulates an actual sensor node at the hardware level. It supports the MSP430 and ATmega CPUs in the form of support for the TelosB (Sky) and MICAz motes.

Whilst COOJA has proved to be very useful there are some issues with it that do not seem to have been resolved. There are issues with the compiler for the TelosB motes2[2] (MSP-GCC 4.6.3) where instructions are generated for the MSP430X CPU instead of the MSP430 CPU present. This is the default compiler provided by TinyOS and also the Debian operating system. Updating to a more recent version of the compiler introduced its own set of issues, so code needs to be written to avoid the generating MSP430X instructions.

2.5.3 RIOT

RIOT is a small operating system for networked, memory-constrained systems with a focus on low-power wireless Internet of Things (IoT) devices [11]. RIOT is based on a microkernel architecture inherited from FireKernel [143]. In contrast to other operating systems with similarly low memory usage (e.g., TinyOS

[2] `https://github.com/contiki-os/mspsim/blob/47ae45/se/sics/mspsim/core/MSP430Core.java#L455`

or Contiki), RIOT allows application programming with the programming languages C++ and provides multiple network stacks such as IPv6, 6LoWPAN and transmission control protocol (TCP). Advantages of the RIOT architecture thus include: (i) high reliability and (ii) a developer-friendly API.

However, RIOT has no provided simulator. This means that cycle accurate simulators like COOJA would need to be used to simulate code. This comes with the same downsides as using Contiki with COOJA, the main one being the low scalability and low simulation speed.

2.5.4 FlockLab

FlockLab is a wireless sensor network testbed developed and run by the Computer Engineering and Networks Laboratory at the Swiss Federal Institute of Technology Zurich in Switzerland [91]. FlockLab combines the capability of a logic analyser, power analyser, serial data logger, and programmable power supply with network synchronisation and deep local storage adjacent to each targetdistributed across the entire testbed.

FlockLab consists of several distributed target-observer pairs and a set of servers. *Observers* are powerful platforms that can host up to four devices under test, the targets, connected through relatively simple interface boards. They connect to several backend *servers* responsible for coordinating their distributed and synchronised operation, for processing and storing collected results, and interacting with FlockLab users. FlockLab users can program the nodes using different OS such as TinyOS and Contiki. However, FlockLab does not have enough nodes, so many SLP-routing protocols requiring hundreds of nodes cannot be tested on it. Besides, queued jobs in FlockLab cannot exceed 1 hour of queued testbed time.

2.6 Performance Attributes in Wireless Sensor Networks

The *Attribute* (or *metric*) is the standard of measurement, and it varies with the measured environment. Several attributes close to WSNs characteristics are used to evaluate network performance.

Network Lifetime

As the energy source is generally limited, protocols in WSNs must be energy efficient to maximise system lifetime. Network lifetime strongly depends on the lifetimes of the single nodes that constitute the network, thus it can be measured by generic parameters such as the time until half of the nodes die. It is also calculated as the time until message loss rate exceeds a given threshold [8].

Energy Consumption

The energy consumption is the sum of used energy of all the nodes in the network, where the used energy of a node is the sum of the energy used for communication, including transmitting, receiving, and idling. Assuming each transmission consumes an energy unit, the total energy consumption is equivalent to the total number of packets transmitted in the network [24, 106].

Message Latency

The message latency is defined as the average amount of time between sending a packet from the source, and the time for successfully receiving the message at the destination. Measurement takes into account the queuing and the propagation delay of the packets. Therefore, the latency measures time cost for the individual message [48].

Fault Tolerance

Sensors may fail due to surrounding physical conditions or when their energy runs out. It may be impractical to replace existing sensors. In response, the WSN must be fault-tolerant such that non-serious failures are hidden from the application in a way that does not hinder it. Fault-tolerance may be achieved through data replication, as in the SPIN protocol [145]. However data replication itself requires energy, thus there is a trade-off between data replication and energy efficiency.

Scalability

Scalability of a network allows more sensor nodes to be involved during network design. WSN scalability needs to consider an integrated view of the hardware and software. For hardware, scalability involves sensitivity and range of sensors, communication bandwidth of the radio, and power consumption. The software parts include the reliability of data transfer, data management and suitable algorithms for analysing the data. The combined hardware and software issues include trade-offs between on-board computations and wireless communication between nodes [8].

Receive Ratio

It is a common attribute and defined as the average percentage of messages sent by the source that arrive at the sink across multiple repeats. The sink may receive the same messages multiple times due to the messages flooding. However, the redundant messages will not be calculated by checking the ID of received messages.

Message Latency

It is the average amount of time it takes a message to travel from the source to the sink. The message latency is only calculated when a message is sent from a

source and successfully delivered to the sink. If a message cannot be successfully forwarded to the sink, the latency will not be counted instead of given infinite.

Messages Sent

It is the average number of total messages transmitted through all nodes per second in the network. A very simple energy consumption model is used where each transmission of a message by a node costs one unit of energy. This model omit other energy consumption such as the energy consumption for nodes sleep or wake-up schedules as nodes are all active in the simulations [17]. This attribute approximates the energy costs of sending and receiving which are expensive activities in WSNs [101].

Coverage

It is always advantageous to have the ability to deploy a network over a larger physical area. Multi-hop communication techniques can extend the coverage of the network, but increase the power consumption of the nodes, which may decrease the network lifetime [77].

For the scope of SLP, there are extra attributes used to evaluate the level of SLP.

Safety Period

The overall objective of any WSN-based SLP solution is to ensure that the asset (at a given location) is never captured through information leaked by the WSN. There are two observations [16, 68]: If the asset is static, then the attacker can perform an exhaustive search of the network to find the asset. In this case, the SLP problem becomes irrelevant. Specifically, if there exists no time bound on the capture time, then an exhaustive search is a trivial solution, yet effective solution. On the other hand, if the asset is mobile, then performing an exhaustive search of the network is unsuitable, as the attacker may zoom in on a given

location only to find out that the asset has moved. Thus, the SLP problem can only be considered when it is time-bounded, capturing the maximum amount of time there mobile asset will spend at a given location.

This notion of time bound has been termed as an attribute called the *safety period* in the literature [74]. The higher the safety period is, the higher the source location privacy level. However, using the safety period attribute means that simulation runtime is unbounded and potentially very large.

Thus an alternative, but analogous, definition for the safety period was used for each network size and network configuration: the safety period is obtained when protectionless flooding is used as the routing protocol. The protectionless flooding, as described by Ozturk et al. [115] works as follows: a node broadcasts a packet to its neighbours, its neighbours then broadcast the packet to their neighbours. This process continues until all nodes within the WSN have received the packet. Flooding is used as it has been argued to provide the *least* SLP level, hence any SLP improvement is due to the SLP-aware technique [74]. The safety period is then obtained by increasing this value to account for the attacker potentially making bad moves. This definition is commensurate with [17, 69, 71, 134], but uses a different multiplicative factor due to the difference in the type of SLP technique being used.

Capture Ratio

Capture ratio is defined as the number of experiments ending in the capture of an attacker in the safety period divided by the total number of experiment repeats for a specific parameter combination [69]. The lower the capture ratio is, the higher the source location privacy level.

$$\text{Capture Ratio} = \frac{\text{number of experiments ending in a capture}}{\text{total number of experiments}} \tag{2.1}$$

Attacker Distance

It is the average attacker distance from the source recorded at the end of a run of simulation [15, 18]. For example, in [15], the object of the SLP-aware routing protocol is to maximise the attacker's distance from the source, hence keeping the source safe.

CHAPTER 3

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

3.1 Introduction

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. In seminal work on SLP, *phantom routing* [115] was proposed as an approach to address the issue. This protocol combines random walk technique and messages flooding to attract the adversary to the wrong location, hence keeping the source location safe. However, authors [89, 124, 138] claim the weaknesses of phantom routing such as poor SLP and unexpected termination of random walk. Apart from those shortcomings mentioned above, the range of experiments conducted in phantom routing is restrictive in the literature, such that little is known about the ability of phantom routing with multiple sources and various network configurations. Therefore, results presented in support of phantom routing have not included considerations for practical scenarios.

To better understand the SLP issue in WSNs, this chapter implements phantom routing and investigates shortcomings by conducting an in-depth investigation of phantom routing under various scenarios. Specifically, this chapter considers up to three sources and various network configurations. Besides, four parameters are identified, namely (i) length of random walk, (ii) source period, (iii) network size and (iv) number of sources, that impact the performance of phantom routing. These parameters are varied to assess their impact as well as a range of experiments are conducted to validate these findings, both individually

and in combination, on phantom routing as a viable approach to the problem of SLP. The results demonstrate that previous work in phantom routing does not generalise well to multiple sources and different network configurations.

The rest of this chapter is organised as follows. Section 3.2 presents the reasons why the phantom routing is worth for assessment and the implementation of phantom routing. Section 3.3 shows the problem statement and Section 3.4 presents the models assumed. Details of the experiments conducted are provided in Section 3.5. Section 3.6 shows the procedure of phantom routing execution in a safety period. Section 3.7 presents the results of these experiments. Section 3.8 concludes this chapter with a summary of contributions.

3.2 Phantom Routing

This section first explains the reasons why phantom routing is chosen for assessment and details the implementation of phantom routing.

3.2.1 Why Phantom Routing?

The phantom routing is a popular and simple protocol to provide SLP. Many similar routing protocols are proposed based on the idea of phantom routing [140, 156]. The reasons why the phantom routing is selected and assessed rather than other random-walk based routing protocols are:

- Results presented in support of phantom routing are often *single* source, thus have not included considerations for practical network configurations and multiple sources, omitting simulations and analyses with larger network sizes.

- As many other routing protocols are derived from phantom routing [140, 156], they all have very similar phases: the random walk phase and the flooding phase. If the phantom routing has been proved that it can only deal with one source, the deduction can be made that those protocols

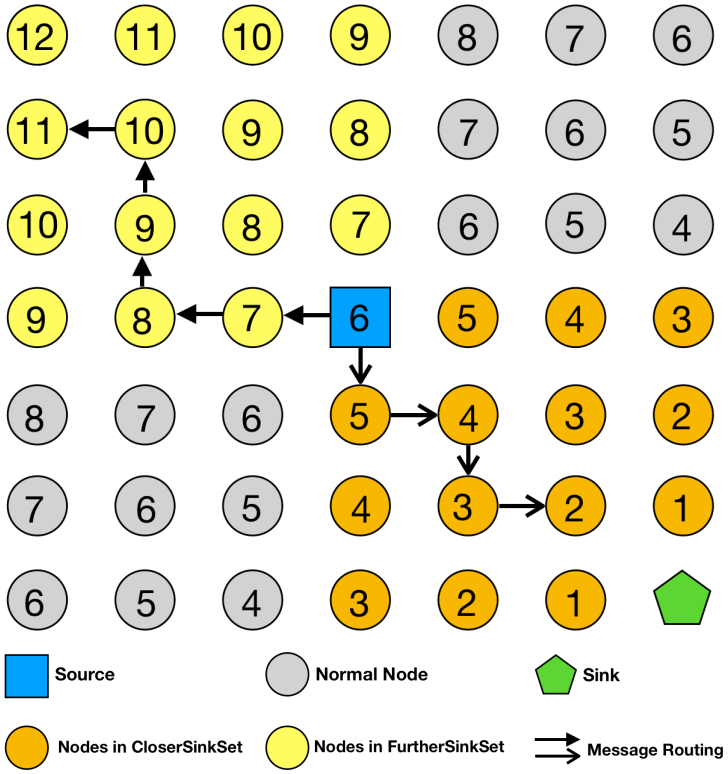


Figure 3.1: Illustration of neighbours division in phantom routing.
In a network size of 7×7, the number in a node indicates the distance between the sink and the node.

similar to phantom routing also cannot deal with multiple sources, hence avoiding investigate all random-walk based routing protocols.

3.2.2 Phantom Routing Implementation

Since phantom routing was first proposed in 2004, it attracted much attention [74, 115, 156]. In this chapter, the phantom routing is implemented with a hop-based directed random walk [115]. Phantom routing uses Algorithm 2, Algorithm 3 and is described below:

Deployment Phase

At the initial stage, before the source sends normal messages, the sink acting as a landmark node floods *beacon* messages through the network. The purpose is to

Algorithm 2 Random Walk Phase in Phantom Routing

```
1: procedure RANDOM WALK PHASE(msg, s)
2:     msg.S_dir ← ⊥
3:     msg.M_dir ← ⊥
4:     msg.h_walk ← s
5:     msg.S_dir ← CHOOSEONESET(msg)
6:     while msg.h_walk ≠ 0 do
7:         msg.M_dir ← CHOOSEONENEIGHBOUR(msg.S_dir)
8:         if ISREACHSINK(msg) = True ∨ msg.M_dir = ⊥ then
9:             msg.h_walk ← 0
10:            break
11:        end if
12:        msg.h_walk ← msg.h_walk − 1
13:        FORWARDMESSAGE(msg.M_dir)
14:    end while
15: end procedure
```

make a node knows its neighbouring nodes and obtains the distance to the sink. Based on the distance of a node to the sink, each node maintains two sets for all its neighbours: *CloserSinkSet* contains all the neighbours whose hop counts to the sink are smaller than or equal to the node's hop count to the sink, and *FurtherSinkSet* includes neighbours with a larger hop count to the sink. After neighbour nodes are partitioned, the source randomly picks one of these two sets and sends normal messages to one neighbour in the chosen set. An example is shown in Figure 3.1. If *CloserSinkSet* is chosen, intermediate nodes always forward messages to the nodes whose distance to the sink is smaller than itself. In addition, Figure 3.1 shows approximately half nodes in the network can be chosen into either *CloserSinkSet* or *FurtherSinkSet*. Furthermore, the rest of nodes do not belong to either *CloserSinkSet* or *FurtherSinkSet* cannot be selected as phantom nodes.

Random Walk Phase

During random walk phase, messages are always sent to the neighbour in the chosen set. If a message is blocked (e.g., there is no neighbour in the chosen set so messages cannot be forwarded) the random walk phase stops. In other cases, when a message travels s hops (assuming random walk length is s), it has

Algorithm 3 Flooding Phase in Phantom Routing

```
1: procedure FLOODING PHASE(msg)
2:     if msg.h_walk = 0 then
3:         if ISREACHSINK(msg) = False then
4:             FLOODING(msg)
5:         end if
6:     end if
7: end procedure
```

finished the random walk phase.

Flooding Phase

When the random walk phase ends, if the message does not reach the sink, the message then floods the network until it reaches the sink.

3.3 Problem Statement

The problem to be addressed in this chapter is the following: In a WSN, the phantom routing protocol is used to deliver messages from the source to the sink. When an attacker is initially located at the sink and starts receiving messages sent by the source to the sink, an important problem is to assess the performance attributes (i.e., capture ratio, receive ratio, message latency and messages sent) of phantom routing under multiple sources and various network configurations. Formally, the problem specification is shown in Figure 3.2.

3.4 Models

The system model and threat model used in the experiments are presented following on from description in Chapter 2.

3.4.1 System Model

The system model is described from three aspects: (i) the sensor nodes presenting the information and knowledge of a node, (ii) the network containing the network

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 phantom routing protocol $\mathcal{P}$,
- A distributed eavesdropper $\mathcal{A}$ that is located at the sink initially,
- A set of source locations $\mathbb{L}$,
- A set of network configurations $\mathbb{C}$ and
- A safety period P_{safety},

Objective:

- Assess the attributes performance of phantom routing over P_{safety} in the presence of $\mathbb{L}$ sources using $\mathcal{P}$ in G, $\mathbb{C}$ and $\mathcal{A}$.

Figure 3.2: Problem statement: Assessment of phantom routing under multiple sources and various network configurations

deployment, and (iii) the message structure demonstrating the information preserved in a message.

Sensor Nodes Modelling

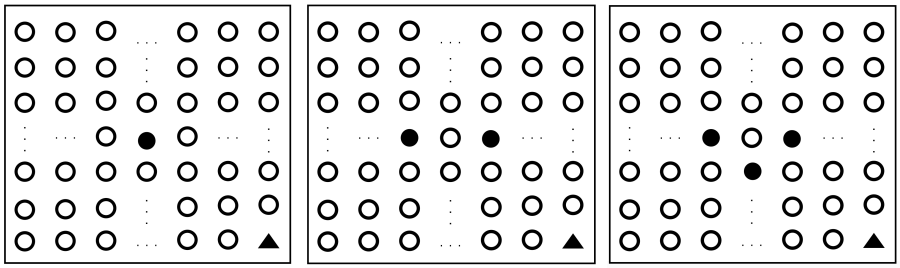
The wireless sensor node is a small computing device with communication and computation capabilities. There are multiple sources collecting data and only one sink receiving data in the wireless sensor network. When a source detects an event, it will route, in collaboration with other nodes, the message to the sink. All the nodes are stationary, i.e., they do not move in the network. It is assumed that all nodes have the same communication range. A node m that can directly receive a message from a node n is called a *neighbour* of n and a node will have knowledge of all of its neighbouring nodes. Each node has a unique node ID. The nodes do not have GPS capability.

Network Modelling

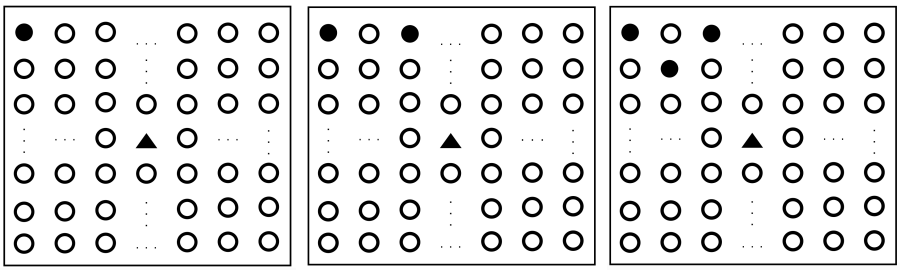
The sensor network is assumed divided into cells where each pair of nodes can communicate directly with each other. The link between node pairs may be

unidirectional or bidirectional. A collection of nodes consist of the network regarded as undirected graph $G = (V, E)$, where V represents the set of N wireless sensor nodes, and E is a set of links connecting two distinct nodes. 1-hop neighbours are defined as two nodes $m \in V$, $n \in V$ and $m, n \in E$, i.e., m and n are in each other's communication range. The graph $G = (V, E)$ defines the configuration of the network. This thesis focuses on the grid-like network configuration, i.e., network of size $n \times n = N$.

In order to test the performance of phantom routing in practical scenarios, up to three sources and different network configurations are considered. For the network configurations, the *SourceCorner* configuration and the *SinkCorner* configuration are chosen. The three sources positions and the network configurations are shown in Figure 3.3. The reason why the sources are clustered together is that the thesis envisions providing SLP for several assets grouping together. Nodes are located 4.5 meters apart in the network.



(a) SinkCorner configuration with 1, 2, 3 sources



(b) SourceCorner configuration with 1, 2, 3 sources

Figure 3.3: Network configurations with multiple sources

Message Structure

The message content is encrypted, thus it can only be read by the correct node (e.g., the sink) and not by an attacker. A normal message contains such information when routed in the network:

$$< id, \mathcal{S}_{dir}, \mathcal{M}_{dir}, h_{walk}, broadcast >$$

where id is the unique sequencing number of messages sent from the source. $\mathcal{S}_{dir}$ is the random walk set (i.e., $CloserSinkSet$ or $FurtherSinkSet$). $\mathcal{M}_{dir}$ is the random walk direction. h_{walk} is the length of random walk. The value of $broadcast$ is a boolean type indicating the status of a message whether it is used for broadcasting or not.

3.4.2 Threat Model

This section assumes a *distributed eavesdropper* which 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. The adversary capabilities can be categorised into the following four domains:

Adversary Behaviour

There is a passive adversary that initially starts at the sink since it is guaranteed to detect the arrival of a message at that location. Wherever the attacker is located, upon receiving (i.e., overhearing) the *first* new message at that location, the attacker moves to the neighbour who relayed the message. The reason to focus on the first new message is that the message has, with high probability, travelled along the shortest path from the source to the sink. Thus, when the attacker hears a new message, it makes a step towards the source. This process can be repeated a number of times until the attacker reaches the source location, whereby it captures the asset.

View of the Network

The adversary only has a local view where it can eavesdrop a 1-hop neighbourhood comprised of different nodes. This can be achieved with signal detection devices or other sniffers. The reason why this thesis focuses on a distributed attacker with small visibility of the network is that an attacker physically presenting in the network is more likely to have a limited power. For attackers to gain global visibility of the network they will need to expend significant resources. For instance, attackers need to deploy their own WSN to monitor the WSN that is monitoring the assets, or has long range directional antennas. However, either of which costs too much. Therefore, this work focuses on a single attacker with a local view of the network.

Resources Strength

The adversary has a large energy source, i.e., there is no need to assume an infinite energy source but rather than the amount of energy required for the task is much less than the amount of energy available. The attacker does not keep track of historical information, i.e., it may revisit a node that it has previously visited. Besides, the attacker has the ability to determine the source of a message that it overhears through the use of a directional antenna and obtain the strength of the signal using spectrum analysers. However, the attacker does not read the messages it overhears, so cannot obtain the contents of a message.

Network Knowledge

The attacker does not know the locations of nodes. Specifically, the attacker may know the configuration of the network but not the specific locations of nodes. For example, the attacker may know that the configuration is a grid, but not the placement of the nodes in the grid. The attacker also knows the location of the sink (similar to the assumption made in seminal work by [74]).

3.5 Experimental Setup

The TOSSIM (V2.1.2) simulation environment was used in all experiments [83]. An experiment is made from a single execution of the simulation environment using a specified protocol configuration, network nodes and a safety period. An experiment terminated when *any* source node had been captured by an attacker during the safety period or the safety period had expired.

A square grid network layout of size $n \times n$ was used in all experiments, with $n \in \{11, 15, 21, 25\}$, i.e., networks with 121, 225, 441 and 625 nodes respectively. Node neighbourhoods were generated using *ideal* communication model , which is a unit disk graph radio model (UDGM) where a perfectly reliable network link exists between the edges of a node's neighbours. The noise model was created using the *meyer-heavy* noise sample file provided with TOSSIM[1].

The time interval between two messages sent from the real sources is called the *source period*. For example, the source period is 0.5 second when 2 messages are sent from the source per second. The source period was normalised with respect to the number of sources so that any configuration will have the same overall source period. The source period was set to be either 0.25, 0.5, 1.0, or 2.0 second(s) (i.e., 4, 2, 1, 0.5 messages are sent from sources per second correspondingly). The random walk length is regarded as a few hops and normally does not exceed the sink-source distance in phantom routing [74]. Therefore, the experiments set the fixed short random walk to 2, 5 and 8 hops. At least 2000 repeats were performed for each combination of source location and parameters.

The different safety periods P_{safety} were calculated as the following, where $\mathcal{TT}$ is the safety period for protectionless flooding and ψ is the *safety factor*. A large safety factor results in a long safety period, meaning that the simulation has excessive time to run. Therefore, a large safety factor is not reasonably given for the simulation runtime, but normally set more than 1.0 (i.e., the safety period is longer than the time taken of flooding).

The safety period was calculated with the safety factor 1.3 from Equation 3.1.

[1]The first 2500 lines of meyer-heavy.txt were used.

The reason to choose the value 1.3 is because the safety period is longer than the time taken of protectionless flooding. In fact other safety factor values were also applied. The time taken $\mathcal{TT}$ for each network size when source period was 1 second, for protectionless flooding is shown in Table 3.1, Table 3.2 and Table 3.3[2]. The tables show that for a larger network size an adversary needs longer time to reach the source location due to the longer sink-source distance[3].

$$P_{safety} = \psi \times \mathcal{TT} \tag{3.1}$$

To evaluate the performance of routing protocols, four performance attributes were used: (i) capture ratio, (ii) receive ratio, (iii) message latency and (iv) messages sent, which were discussed in Section 2.6.

Network Size	**SinkCorner (1 source)**	**SourceCorner (1 source)**
11×11	12.479 ± 2.440	12.826 ± 2.660
15×15	17.574 ± 3.041	18.059 ± 3.401
21×21	25.617 ± 4.102	25.813 ± 4.169
25×25	30.855 ± 4.452	31.071 ± 4.634

Table 3.1: Time taken (seconds) of flooding for each network size with one source

Network Size	**SinkCorner (2 sources)**	**SourceCorner (2 sources)**
11×11	16.319 ± 5.789	15.286 ± 5.794
15×15	21.794 ± 6.421	20.609 ± 6.111
21×21	29.846 ± 7.077	28.557 ± 6.615
25×25	35.255 ± 7.354	33.894 ± 6.971

Table 3.2: Time taken (seconds) of flooding for each network size with two sources

3.6 Demonstration of Simulation Procedure

Before exploring the impact of various parameters, this section explains the mechanism of simulations in terms of message timings. An example shown

[2]The results are generated from 10000 repeats of protectionless flooding with meyer-heavy noise model and ideal communication model.

[3]The results of sink-source distance are presented in the Appendix B.

Network Size	**SinkCorner (3 sources)**	**SourceCorner (3 sources)**
11×11	14.892 ± 3.951	14.478 ± 4.448
15×15	20.527 ± 4.650	19.898 ± 4.960
21×21	28.281 ± 5.267	28.140 ± 5.745
25×25	33.815 ± 5.806	33.699 ± 6.459

Table 3.3: Time taken (seconds) of flooding for each network size with three sources

in Figure 3.4 is a run of simulation with configurations of network size of 11×11, the SourceCorner configuration, a safety period of 50 seconds, a source period of 1 second and a random walk length of 5 hops. Every node is associated with the unique node ID starting with 0. There are three types of messages: *away*, *beacon* and *normal* message. There are also three types of nodes in the network: source, sink and normal nodes. The procedures are described as follows:

1. Initially nodes in the network randomly boot up in 1 second. Then the sink floods *away* messages to inform the network of distance from each node to the sink. *Beacon* messages are flooded by the sink a little bit after *away* messages to help ensure the information spreads correctly.

2. As the source period is 1 second, messages are only sent during this initial timing period of a second. In most cases, all the nodes are involved in sending messages transmitted. However, In an abnormal case only a few nodes are involved in transmission (e.g., at time 14.0 seconds). This is because the message is lost due to unreliable links during the random walk phase, thus no flooding messages are sent through the network.

Although the safety period is 50 seconds, the experiment terminates at 17.5 seconds, suggesting the adversary reaches the source location and catches the source at 17.5 seconds.

3.7 Simulation Results

This section conducts experiments to examine the impact of varied random walk lengths in Subsection 3.7.1, source periods in Subsection 3.7.2, network sizes

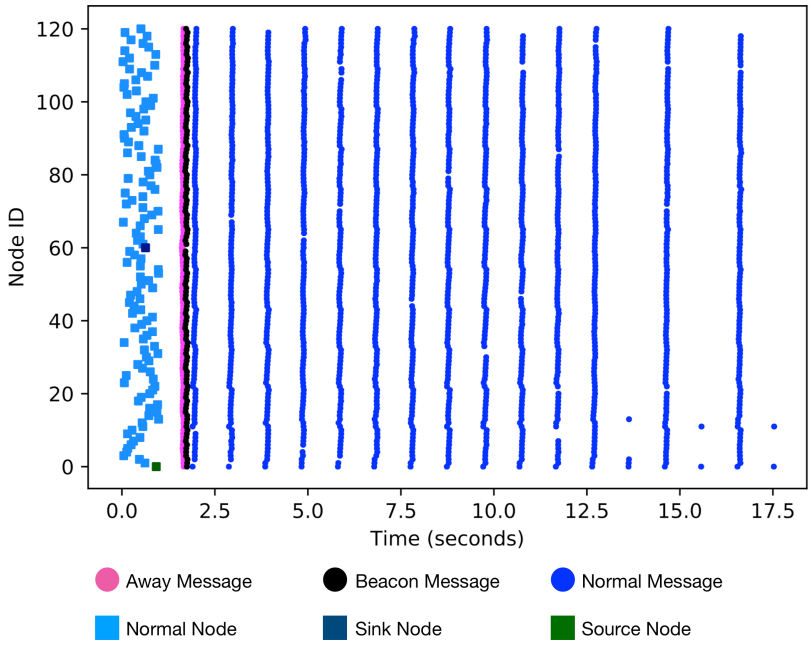


Figure 3.4: Demonstration of messages sent in the safety period

in Subsection 3.7.3, and numbers of sources in Subsection 3.7.4. In addition, results of other attributes are discussed in Subsection 3.7.5.

3.7.1 Results: Impact of Random Walk Length on SLP

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

Intuition: The intuition behind this investigation is the following: Denote the length of the random walk by h and denote the source period by r. If the number of paths from a source to nodes in $\mathcal{P}^h$ is similar, then the expected number of times nodes will hear a message is per unit time if $\frac{1}{r|\mathcal{P}^h|}$. Thus, the mean time between two successive messages heard by an attacker is approximately $r|\mathcal{P}^h|$. 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:

$$\bar{T_c} \approx \sum_{i=1}^{h} r|\mathcal{P}^i| \tag{3.2}$$

Therefore, the conjecture is: 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. **Results:** To address this conjecture, experiments were conducted for each network size and network configuration. The length of the random walk was varied to be either 2, 5, or 8 hops and the network size is 11×11 (i.e., 121 nodes). From Figure 3.5, the following observations can be made:

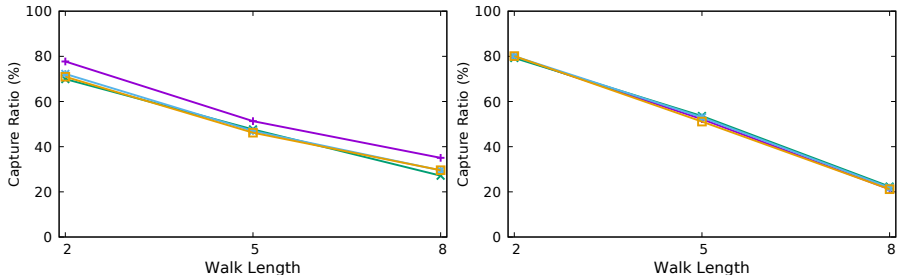
1. In both network configurations, an increase in the length of the random walk leads to a corresponding decrease in the capture ratio, thereby confirming the conjecture.

2. The SourceCorner configuration yields better SLP level than the SinkCorner configuration especially with long random walk length and multiple sources. In both configurations the capture ratio is around 80% with random walk length of 2 hops. However, in the SourceCorner configuration the capture ratio is below 20% with random walk length of 8 hops.

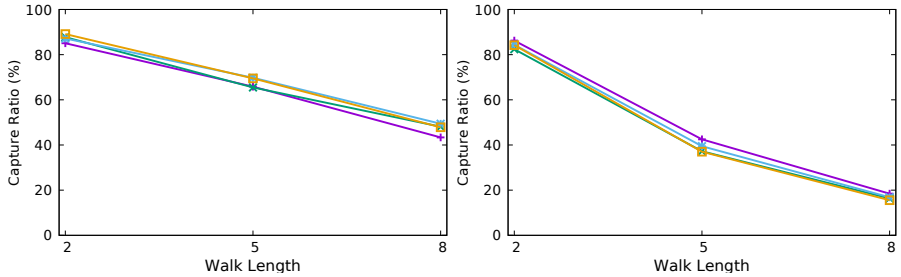
To have a better insight into the impact of phantom walk length, the section also investigates the receive ratio in such situations. The results are shown in Figure 3.6, and the following observations are made:

1. With an increase of random walk length, the receive ratio in the SinkCorner configuration performs better than the SourceCorner configuration. For instance, in the SinkCorner configuration with three sources, the receive ratio is 70%. However, in the SourceCorner configuration with three sources, the receive ratio decreases to 50% (see Figure 3.6c).

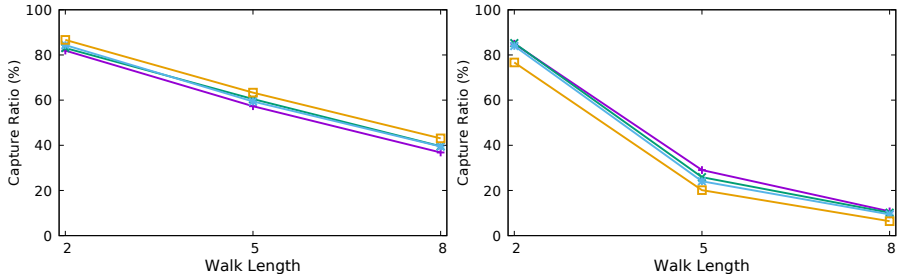
2. The receive ratio is at the same level within different source periods. Besides, receive ratio decreases with an increasing number of sources. This can be explained by multiple sources causing collisions when more than



(a) One Source: SinkCorner and SourceCorner configuration respectively



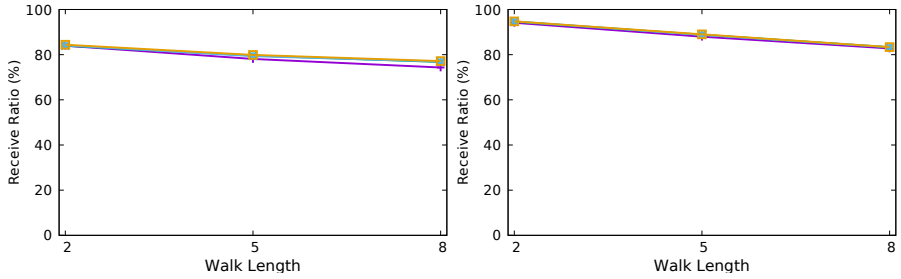
(b) Two Sources: SinkCorner and SourceCorner configuration respectively



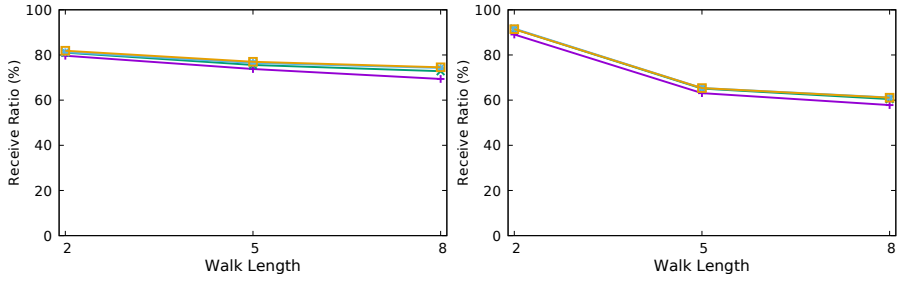
(c) Three Sources: SinkCorner and SourceCorner configuration respectively

Figure 3.5: Impact of random walk length: Capture ratio with multiple sources and network configurations

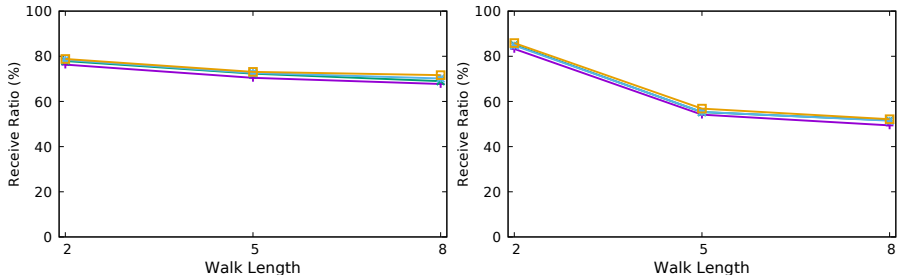
one message arrives simultaneously at a node. This message collisions can also explain why multiple sources yield a better SLP level than a single source.



(a) One Source: SinkCorner and SourceCorner configuration respectively



(b) Two Sources: SinkCorner and SourceCorner configuration respectively



(c) Three Sources: SinkCorner and SourceCorner configuration respectively

Figure 3.6: Impact of random walk length: Receive ratio with multiple sources and network configurations

3.7.2 Results: Impact of Source Period on SLP

In real-world scenarios, it is to be 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 source periods have on the provision of SLP.

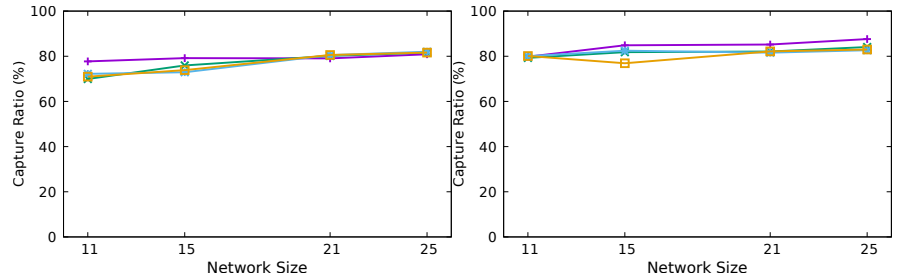
Intuition: The intuition behind this investigation is the following: Denote the set of nodes h hops away from the source by S^h, where h is the length of the random walk, and denote the source period by r. If all the nodes in S^h can be reached independently by a similar number of paths, then the expected number of messages received by any node $n \in S^h$ per unit time is approximately $\frac{1}{r|S^h|}$. Thus, if an attacker has reached the node n, the lower 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 lower r can cause the attacker to capture the source before the safety period elapses. Thus, the conjecture is: As the value of r increases it will result in a lower capture ratio.

Results: To address this conjecture, experiments were conducted for each network size and network configuration. The source period of the application was varied such that the time cost of messages transmitted per message was 0.25, 0.5, 1.0 or 2.0 sent from a source in the network.

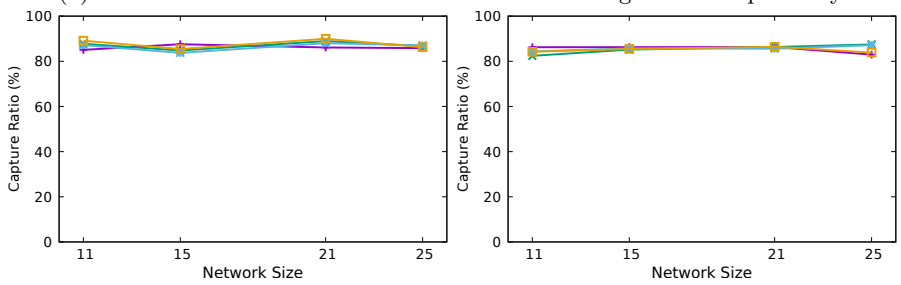
From Figure 3.5, it can be observed that, across all network sizes and configurations, a decrease in the source period leads to a corresponding increase in the capture ratio for one source, thereby confirming the conjecture. As more messages are being sent in the same period of time, the attacker has a greater number of chances to move towards the source in response to a message. There are two observations regarding these results:

1. In the network configuration of one source, the lowest source period yields the highest capture ratio.

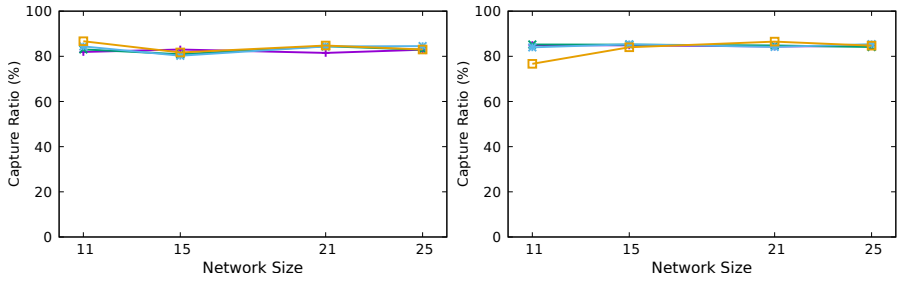
2. In the case of three sources, the low capture ratio is achieved with low source period in the SinkCorner configuration, while the high capture ratio is achieved with high source period in the SourceCorner configuration. These are due to the different network configurations and message collision.



(a) One Source: SinkCorner and SourceCorner configuration respectively

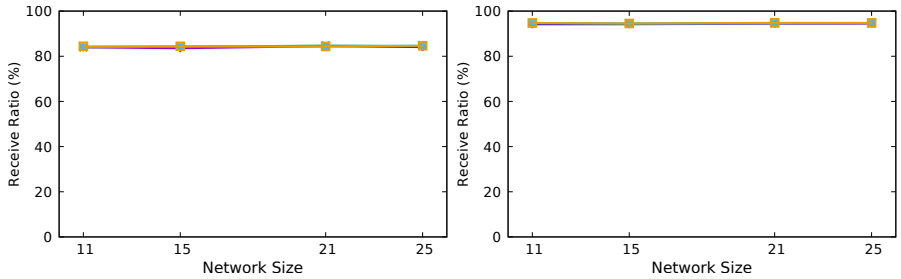


(b) Two Sources: SinkCorner and SourceCorner configuration respectively

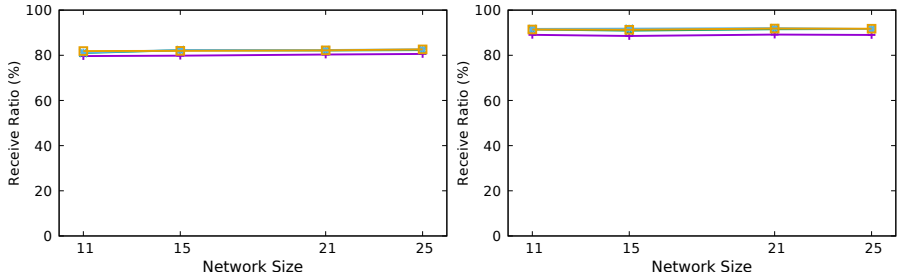


(c) Three Sources: SinkCorner and SourceCorner configuration respectively

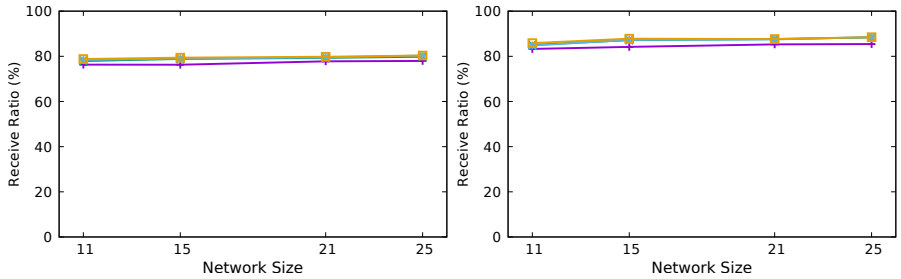
Figure 3.7: Impact of network sizes: Capture ratio with multiple sources and network configurations when random walk length is 2 hops



(a) One Source: SinkCorner and SourceCorner configuration respectively



(b) Two Sources: SinkCorner and SourceCorner configuration respectively



(c) Three Sources: SinkCorner and SourceCorner configuration respectively

Figure 3.8: Impact of network sizes: Receive ratio with multiple sources and network configurations when random walk length is 2 hops

3.7.3 Results: Impact of Network Size on SLP

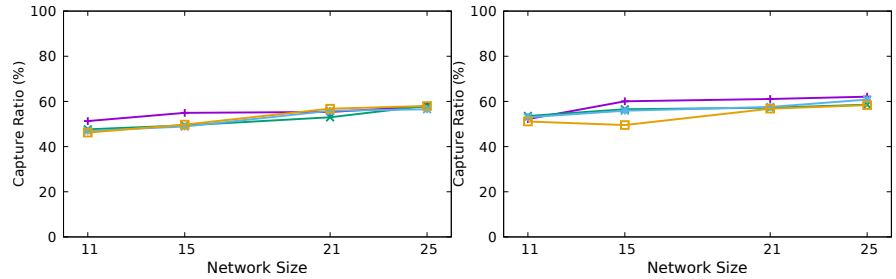
This section explores the impact of network size to see whether phantom routing works efficiently with a larger network.

Intuition: The intuition behind this investigation is the following: Denote the set of nodes h hops away from the source by S^h, where h is the length of the random walk. Given a network size N and fulfil $N^2 > S^h$, an attacker can be attracted to S^h phantom nodes. Therefore, the conjecture is: As the fixed value of h it will not result in any change of capture ratio.

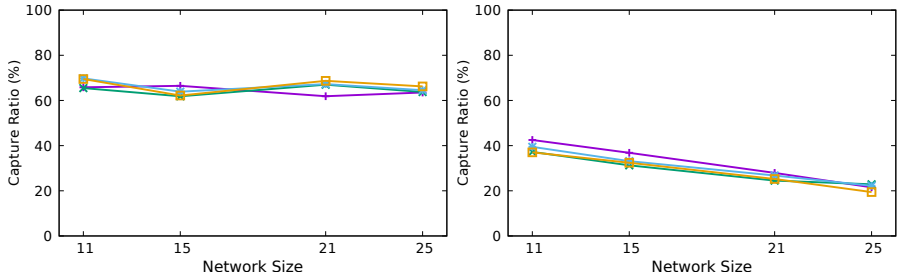
Results: The network sizes were varied as 11, 15, 21, 25 and random walk length h 2, 5 and 8 hops. Figure 3.7, Figure 3.9 and Figure 3.11 contain the capture ratio results and Figure 3.8, Figure 3.10 and Figure 3.12 contain the receive ratio results. The following observations can be made:

1. In the SinkCorner configuration, the SLP level does not change with the network sizes with multiple sources (see Figure 3.7, Figure 3.9 and Figure 3.11). For instance, the capture ratio remains at the level of 80% with two sources in Figure 3.7b and 60% with three sources in Figure 3.9c.

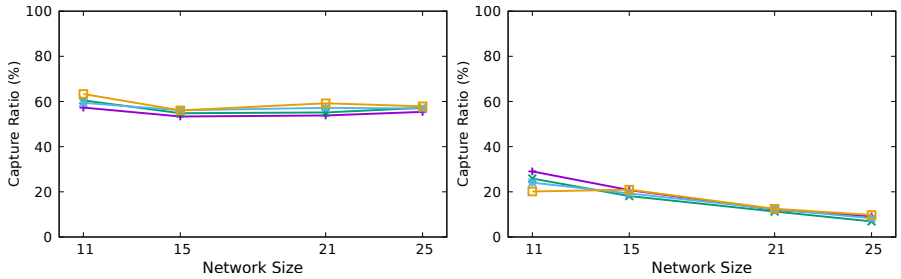
2. The situation in the SourceCorner configuration is opposite to the SinkCorner configuration, where the capture ratio decreases with the increase of network sizes. The reason is due to low receive ratio in the SourceCorner configuration. In the SourceCorner configuration, receive ratio is lower than the SinkCorner configuration with multiple sources (e.g., see Figure 3.10b and Figure 3.10c). The reason behind this case can be explained as follows: In the SourceCorner configuration, source locates in the corner of the network. In this case, messages are always forwarding towards the sink (i.e., the $FurtherSinkSet$ is void and only the $CloserSinkSet$ is chosen). Assuming there are S^h phantom nodes in SinkCorner configuration, only half of S^h phantom nodes exist in SourceCorner configuration, thus causing more message collisions in the SourceCorner configuration.



(a) One Source: SinkCorner and SourceCorner configuration respectively

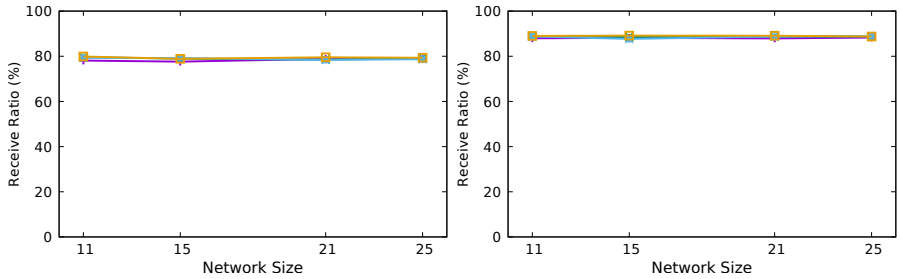


(b) Two Sources: SinkCorner and SourceCorner configuration respectively

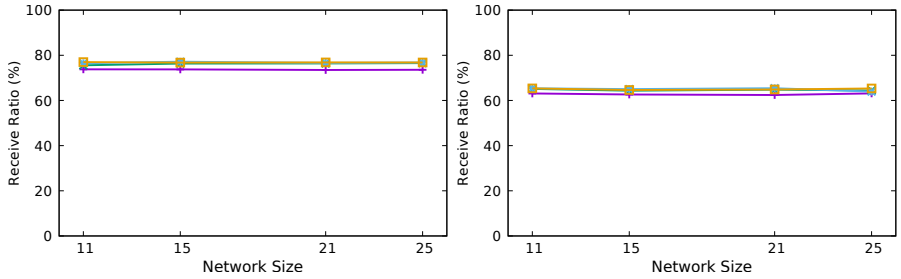


(c) Three Sources: SinkCorner and SourceCorner configuration respectively

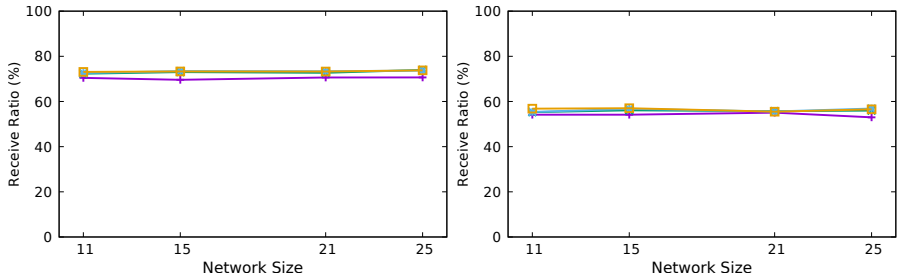
Figure 3.9: Impact of network sizes: Capture ratio with multiple sources and network configurations when random walk length is 5 hops



(a) One Source: SinkCorner and SourceCorner configuration respectively

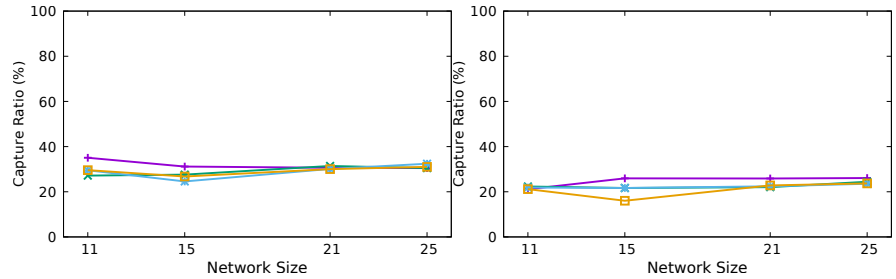


(b) Two Sources: SinkCorner and SourceCorner configuration respectively

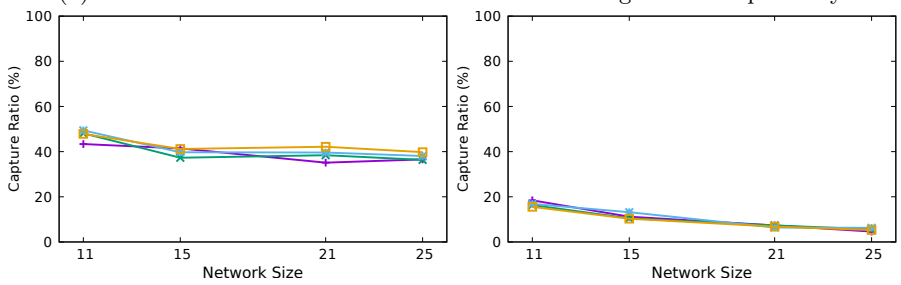


(c) Three Sources: SinkCorner and SourceCorner configuration respectively

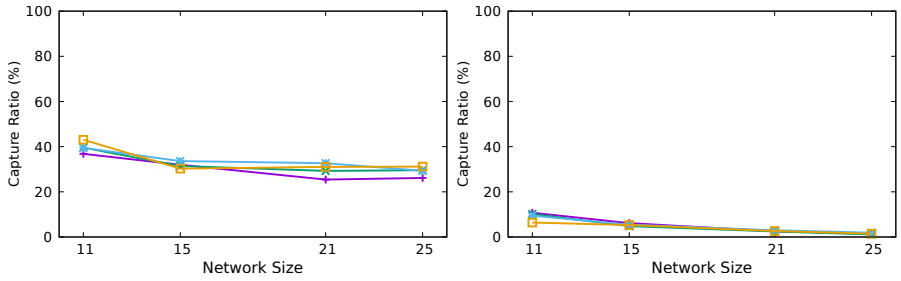
Figure 3.10: Impact of network sizes: Receive ratio with multiple sources and network configurations when random walk length is 5 hops



(a) One Source: SinkCorner and SourceCorner configuration respectively

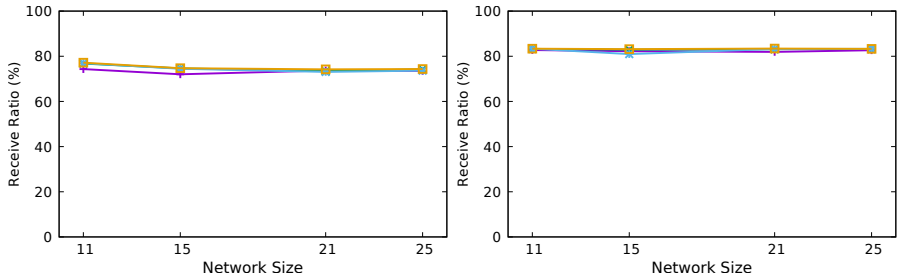


(b) Two Sources: SinkCorner and SourceCorner configuration respectively

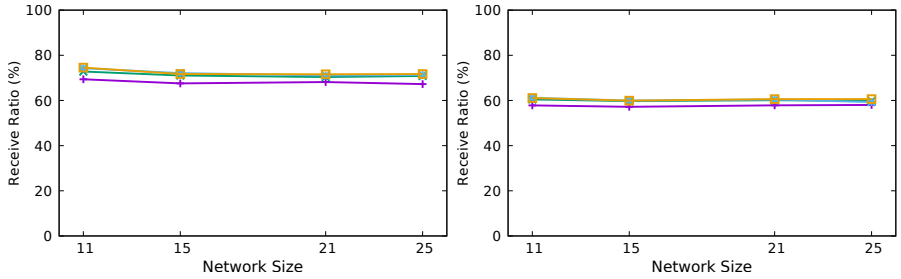


(c) Three Sources: SinkCorner and SourceCorner configuration respectively

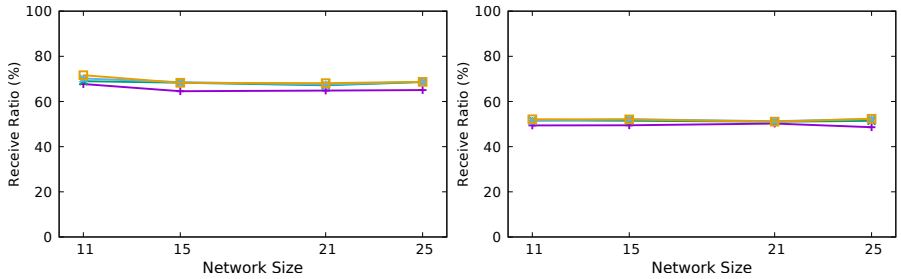
Figure 3.11: Impact of network sizes: Capture ratio with multiple sources and network configurations when random walk length is 8 hops



(a) One Source: SinkCorner and SourceCorner configuration respectively



(b) Two Sources: SinkCorner and SourceCorner configuration respectively



(c) Three Sources: SinkCorner and SourceCorner configuration respectively

Figure 3.12: Impact of network sizes: Receive ratio with multiple sources and network configurations when random walk length is 8 hops

3.7.4 Results: Impact of Number of Sources on SLP

In real-world scenarios, for instance, in an animal monitoring scenario, pandas equipped with sensors as sources are often grouping together. Therefore, there is a need to analyse the impact of multiple sources in the network. This section will present results under the different number of sources and analyse the effect different network sizes have on the provision of SLP.

Intuition: The length of the random walk is denoted by h and the number of phantom nodes with source number i by S_i^h. For multiple sources, three sources have more phantom nodes than two sources and one source situations (i.e., $S_3^h > S_2^h > S_1^h$). It indicates that an adversary may be attracted to more different phantom sources with multiple sources. On the other hand, if sources are scattered over the network, the attacker may easier to capture the source since it is trying to capture any one source rather than capturing all sources. Therefore, the simulations need to be conducted for validation.

Results: The length of the random walk h was 5 hops and the source periods were 0.25, 0.5, 1.0 and 2.0 seconds. There are following observations from Figure 3.13 and Figure 3.14:

1. For all source periods, the capture ratio slightly increases with multiple sources in the SinkCorner configuration, whereas capture ratio decreases in the SourceCorner configuration. Besides, the results also show that the capture ratio varies for different network sizes.

2. In all configurations, in a small network size one source performs best SLP (e.g., low capture ratio), but performs the worst SLP in a large network size. On the other hand, for three sources, the SLP level is very poor in a small network size but best in a large network size.

3. The receive ratio performs the same level with different network sizes in the SinkCorner configuration. However, the receive ratio in the SourceCorner configuration is worse than SinkCorner configuration and decrease with

the increasing number of sources. The reasons have been explained in Subsection 3.7.3.

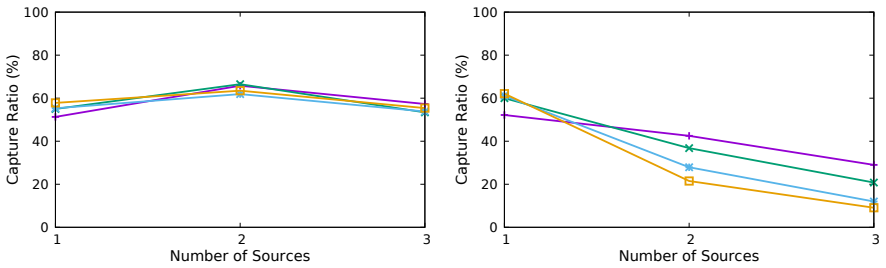
3.7.5 Results: Other Attributes Discussion

So far results presented in this chapter have observed the impact of (i) random walk length, (ii) source period, (iii) network size and (iv) number of sources. In this section, in order to have a comprehensive understanding of the results, other two attributes are investigated. *Message latency* defines the time cost of a message sent from the source to the sink. This attribute does not impact the SLP level of phantom routing, because the latency is only calculated when a message is successfully delivered to the sink from the source. On the other hand, *messages sent* defines the average number of messages transmitted by all nodes in the network per second. It impacts the SLP level due to the source period and the receive ratio. For example, a high source period will imply a low capture ratio because the attacker will overhear only a few messages and would not be able to track the asset down. In a similar way, if a node does not receive a normal message, it is not going to forward it, reducing the number of messages sent[4]. The results in Figure 3.15 and Figure 3.16 were generated with random walk length of 5 hops.

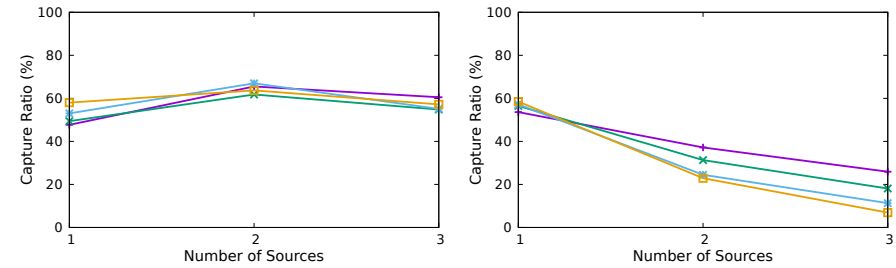
Message Latency

Intuitively, the random walk hop is denoted by h, sink-source distance by Δ_{ss} and transmission time cost between two nodes by α. When the $CloserSinkSet$ is chosen, messages are always forwarded to the sink, so the *minimum* message latency is $\alpha \times \Delta_{ss}$. On the other hand, when the $FurtherSinkSet$ is selected, messages are always forwarded far away from the sink in the random walk phase and then flooded to the sink. The *maximum* message latency is $\alpha \times (2h + \Delta_{ss})$.

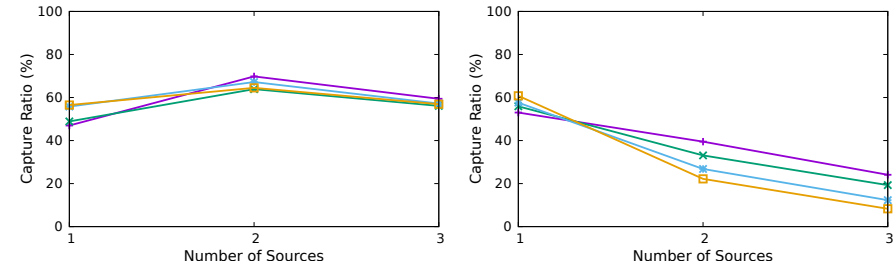
[4]The attribute *message sent* is not the *mutually preferentially independent* attribute, which will be further discussed in Section 5.4.



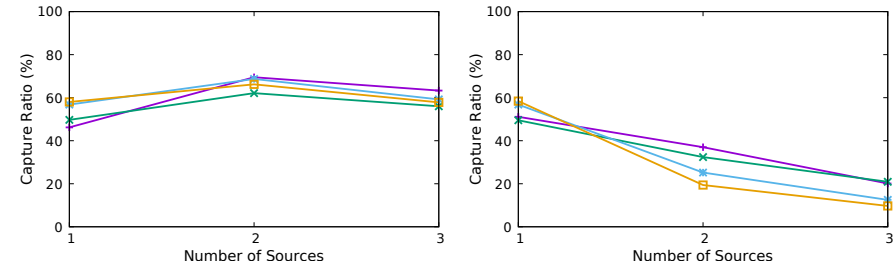
(a) Results of SinkCorner and SourceCorner configurations when source period is 0.25 second



(b) Results of SinkCorner and SourceCorner configurations when source period is 0.5 second

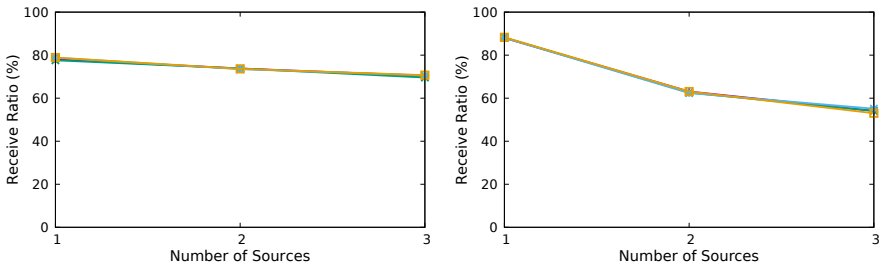


(c) Results of SinkCorner and SourceCorner configurations when source period is 1.0 second

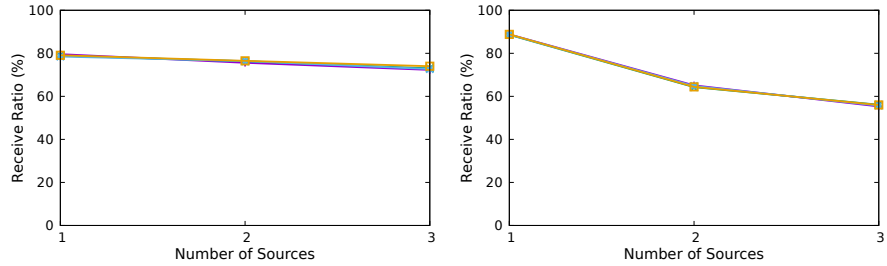


(d) Results of SinkCorner and SourceCorner configurations when source period is 2.0 seconds

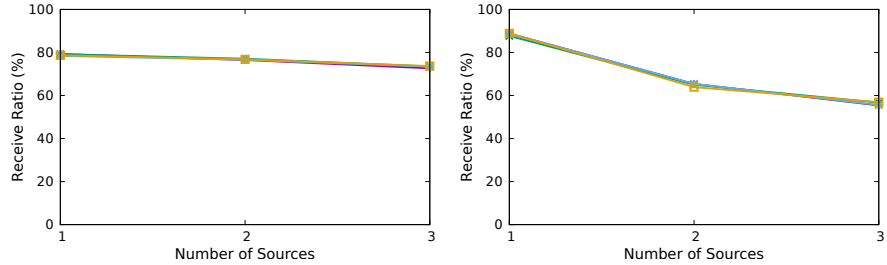
Figure 3.13: Impact of source numbers: Capture ratio with multiple network sizes and source periods



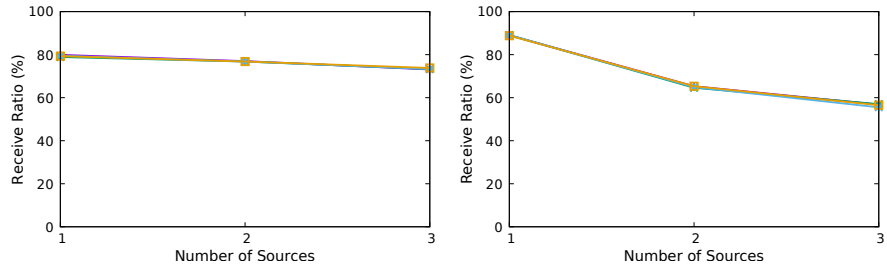
(a) Results of SinkCorner and SourceCorner configurations when source period is 0.25 second



(b) Results of SinkCorner and SourceCorner configurations when source period is 0.5 second



(c) Results of SinkCorner and SourceCorner configurations when source period is 1.0 second



(d) Results of SinkCorner and SourceCorner configurations when source period is 2.0 seconds

Figure 3.14: Impact of source numbers: Receive ratio with multiple network sizes and source periods

Therefore, the message latency lat is in between:

$$lat \in [\alpha \times \Delta_{ss}, \alpha \times (2h + \Delta_{ss})] \tag{3.3}$$

As messages in the SourceCorner configuration are always sent towards the sink, the conjecture is: The latency in the SourceCorner configuration is less than the SinkCorner configuration. The observations from Figure 3.15 are:

- Message latency in the SinkCorner configuration is between 75 and 150 milliseconds while in the SourceCorner configuration the latency yields 50 to 125 milliseconds, which proves the conjecture.
- As phantom routing does not apply any delay technique, message latency is not affected by different source periods and number of sources.
- Message latency increases with larger network sizes as the sink-source distance increases.

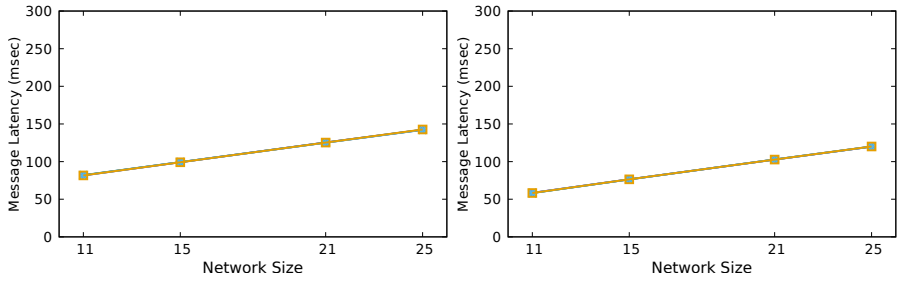
Messages Sent

Many routing schemes are constrained by energy consumption, therefore it is necessary to investigate this attribute. The conjecture is: Low source periods and large network sizes yield high messages sent.

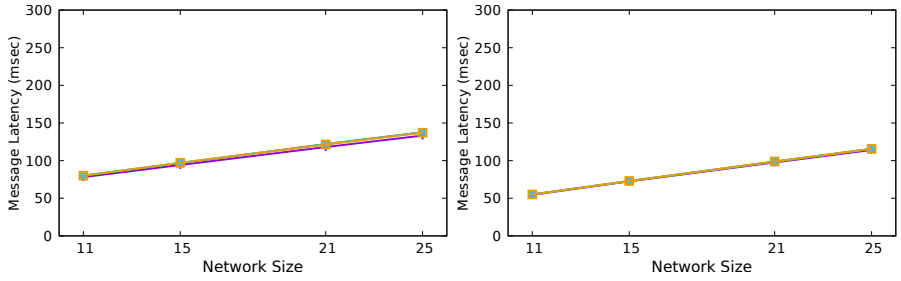
From Figure 3.16 such conclusions can be made: First, for a given network size, nodes transmit more messages in a smaller source period, hence more energy cost. Second, the energy cost increases with larger network sizes.

3.8 Summary

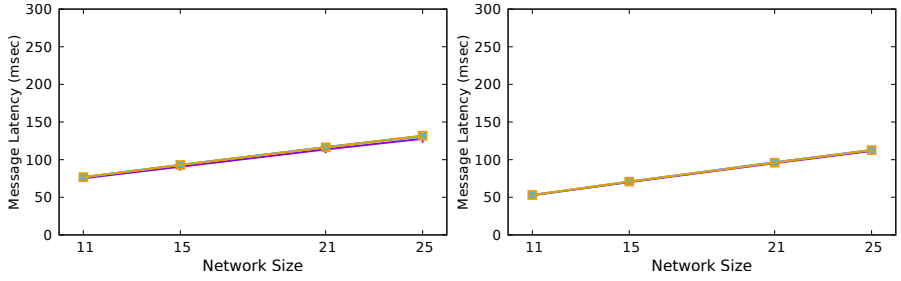
This chapter has investigated the performance of phantom routing, a well-known algorithm that provides SLP in WSNs, under various network scenarios. The chapter has considered multiple sources, various network configurations and four application parameters: (i) the length of random walk, (ii) source period, (iii)



(a) One Source: SinkCorner and SourceCorner configuration respectively

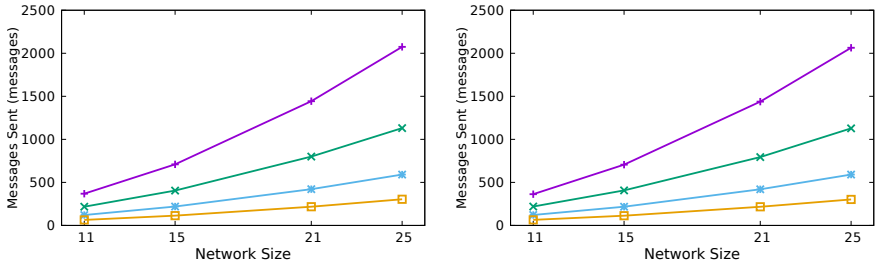


(b) Two Sources: SinkCorner and SourceCorner configuration respectively

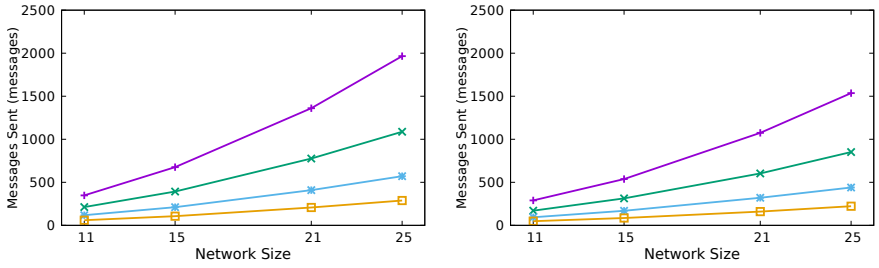


(c) Three Sources: SinkCorner and SourceCorner configuration respectively

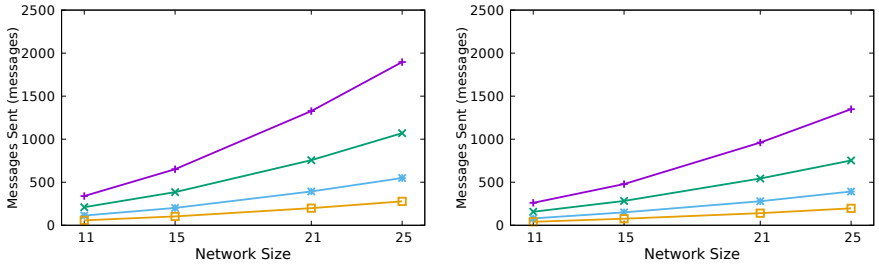
Figure 3.15: Message latency with multiple sources and network configurations



(a) One Source: SinkCorner and SourceCorner configuration respectively



(b) Two Sources: SinkCorner and SourceCorner configuration respectively



(c) Three Sources: SinkCorner and SourceCorner configuration respectively

Figure 3.16: Messages sent with multiple sources and network configurations

network size and (iv) number of sources. These parameters are varied to assess their impact, both individually and in combination on phantom routing. The results show that:

- An increase in the length of the random walk leads to a corresponding increase in the SLP level.
- An decrease of source periods causes a decrease in the SLP level provided, i.e., the capture ratio increases.
- The SLP level does not change subject to network sizes.
- An increasing number of sources also causes an increase of capture ratio.
- Receive ratio affects the performance of the SLP level.

The overarching results summarise the main findings in this chapter. Meanwhile, the shortcomings of phantom routing can be summarised as:

- Phantom routing preserves low SLP level with a short random walk.
- Phantom routing cannot provide a high level of SLP with larger network sizes.
- Phantom routing cannot deal with multiple sources on SLP.

Thus, the main contributions of this chapter are to:

- Define the practical scenarios: multiple sources and different network configurations.
- Identify four parameters that impact the performance of phantom routing.
- Derive expressions that capture the impact of the four identified parameters, as well as conducting a range of experiments to validate these findings.
- Demonstrate that, under varied network conditions, the phantom routing has bad performance, confirming an initial conjecture that phantom routing works well under specific conditions but requires fine-tuning in order to realise optimal performance.

Overall, the results show that the SLP level of phantom routing can drop by a high factor under some practical scenarios. The conclusion is that phantom routing is not as effective as initially claimed, as it was previously evaluated under a restricted set of circumstances and network configurations.

CHAPTER 4

Phantom Walkabouts in Wireless Sensor Networks

4.1 Introduction

Phantom routing is a typical routing protocol that deals with the SLP issue in wireless sensor networks (WSNs). The benefit of phantom routing is that it provides better SLP than the protectionless flooding protocol and causes low messages overhead. However, Chapter 3 has shown the shortcomings of phantom routing by conducting an in-depth investigation under various network configurations. The results demonstrate that phantom routing is not effective under practical scenarios. Therefore, there is a need to develop a new SLP-aware routing protocol which achieves better performance than phantom routing.

This chapter presents the concept of *phantom walkabouts*, a routing protocol that is designed to provide a better SLP with multiple sources than phantom routing. Phantom walkabouts aims to lead an adversary roaming around in the network, hence keeping the source location safe. Compared to phantom routing, phantom walkabouts has the following features: first, each node in the network has the probability of becoming a phantom node, while only a limited number of nodes could become a phantom node in the phantom routing. Second, phantom routes with variable random walk lengths are performed. Third, phantom walkabouts provides the routing with biased random walk as a special case that solves the weakness of long random walks in a certain network configuration. Multiple sources and various network configurations are also considered in the simulations. The results show that phantom walkabouts could provide high SLP level with trade-offs for other attributes.

The rest of this chapter is organised as follows. Section 4.2 presents moti-

vations which leads to the development of phantom walkabouts. This section analyses some cases to demonstrate why phantom routing cannot provide high level of SLP. Section 4.3 presents the implementation of phantom walkabouts and the difference between phantom routing and phantom walkabouts. Problem statement is presented in Section 4.4. Section 4.5 lists the parameters used in the experiments. Experimental evaluation is then presented in Section 4.6, and some discussions about the approach are in Section 4.7. Finally, Section 4.8 summarises the work of phantom walkabouts.

4.2 Motivations of Phantom Walkabouts

The section first briefly reviews phantom routing from Chapter 3 and then illustrates motivations for the development of phantom walkabouts by presenting the weakness of short random walk.

4.2.1 Phantom Routing Review

Chapter 3 considered four parameters that effect phantom routing performance: the length of random walk, source period, network size and the number of sources. By conducting extensive experiments, the results show that (i) phantom routing achieves poor SLP with small random walk hops, (ii) phantom routing cannot provide a high level of SLP with large network sizes, and (iii) phantom routing cannot deal with multiple sources.

Chapter 3 also considered the combined effect of some of these parameters and claimed that phantom routing is not as effective as initially claimed especially under a restricted set of circumstances and network configurations.

4.2.2 Motivations of Phantom Walkabouts

Since the focus in this thesis has been on SLP-aware routing protocols, the capture ratio is considered to be the most important attribute. Figure 4.1 shows a typical scenario during an execution of phantom routing where the source

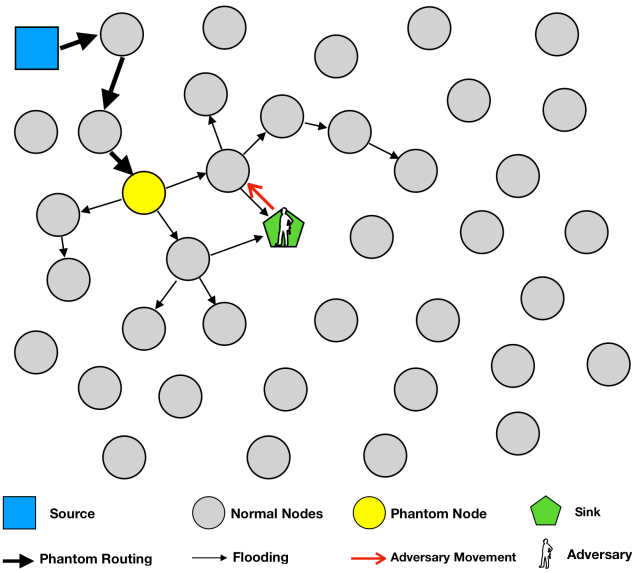


Figure 4.1: Illustration of the message routing with short random walks

sends a message to a phantom node which lies somewhere between itself and the sink. At the beginning, a source sends a message with a few random walk hops, then the message reaches a phantom node after finishing random walk. When the phantom node floods the message to the sink, the first movement of the adversary is always towards the phantom node and source as well. Therefore, routes with a short random walk do not lure the adversary far away from the source. In a case when an adversary is close to the source, the random location of the phantom node may pull an adversary away from the source after flooding (see Figure 4.2). However, in most cases the source can be easily captured since the attacker gets too close to the source.

In order to avoid the risk that the phantom node is positioned close to the real source, it would be beneficial to make the attacker's first movements away from the source, as shown in Figure 4.3. To achieve this, a longer random walk can be used, where the random walk length exceeds the sink-source distance.

As such, phantom walkabouts with a mix of short and long random walks will achieve a higher level of SLP than phantom routing. The phantom walkabouts parametrisation is denoted by $PW(m, n)$, where m and n denote the number

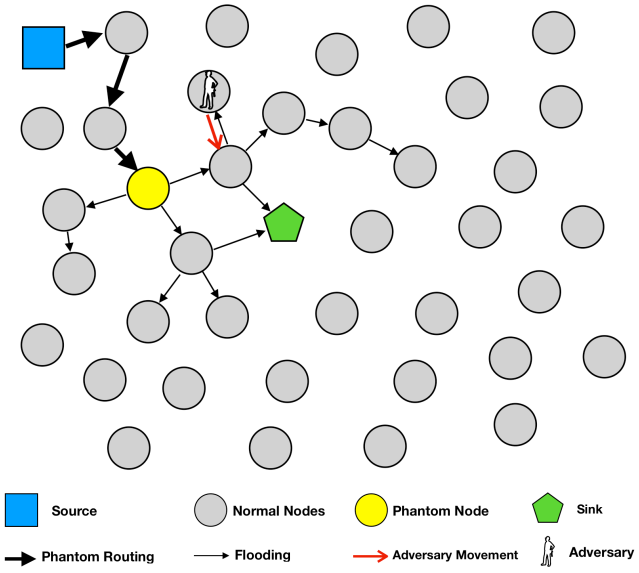


Figure 4.2: Illustration of the message routing with short random walks when an adversary is close to the source

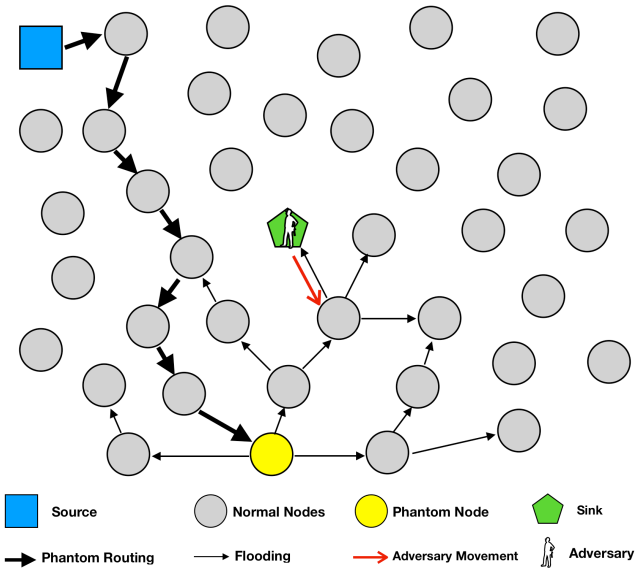


Figure 4.3: Illustration of the message routing with long random walks

of short and long random walk respectively to be performed in a cycle. For example, $PW(1,1)$ denotes a repeating sequence of a short random walk followed by a long random walk. Later, this chapter will investigate the performance of phantom walkabouts with various parameterisations (i.e., $PW(1,0)$, $PW(1,1)$, $PW(1,2)$ and $PW(0,1)$).

4.3 Implementation of Phantom Walkabouts

This section proposes a novel SLP routing protocol, termed *phantom walkabouts*, which is a more generic version of the phantom routing strategy that addresses the impact caused by small random walks in phantom routing. Phantom walkabouts is basically phantom routing with variable random walk lengths. The results will show better SLP than phantom routing. This section explains the rationale behind the protocol and algorithms for forming the random walk, the biased random walk and the overall phantom walkabouts algorithm. The differences between the phantom routing and phantom walkabouts are also summarised in Table 4.2. Table 4.1 lists the most commonly used notations in the implementation of phantom walkabouts.

Table 4.1: Commonly used notations

Notation	**Description**
msg	The normal message
$\mathcal{S}_{dir}$	The random walk set of a message
$\mathcal{M}_{dir}$	The random walk direction of a message
$\mathcal{B}_{dir}$	The biased random walk direction of a message
$\mathcal{P}_{biased}$	The probability of biased random walk
$\mathcal{TT}$	The time taken (seconds) of protectionless flooding
P_{safety}	The safety period (seconds)
M_s	The message with the short random walk
M_l	The message with the long random walk
Δ_{ss}	The distance in hops between the sink and the source
h_{walk}	The remaining hops of the random walk

4.3.1 Random Walk in Phantom Walkabouts

In the random walk phase of phantom routing some exceptional situations have never been considered. For example, not every node could become a phantom node; badly choice of phantom nodes may reduce the SLP; the random walk stops when next location of a message is void. Therefore, a new random walk algorithm is introduced to deal with neighbour nodes division and exceptional termination of random walks. The new random walk algorithm works for both short and long random walks.

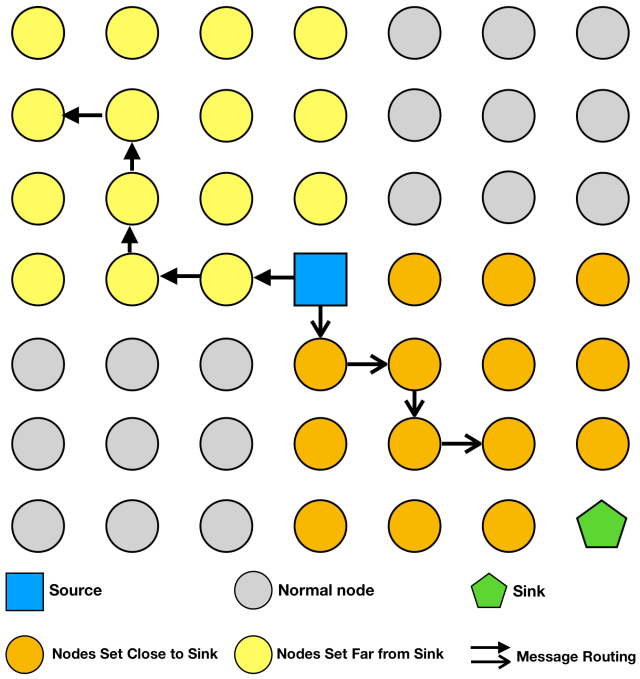


Figure 4.4: Illustration of neighbours division in phantom routing

Neighbour division in random walk phase

Different from phantom routing that each node's neighbours are divided into two sets, each node's neighbours are now divided into four sets in different directions. This division can be done as follows: A node in the network acts as another landmark node. As shown that in phantom routing the sink divides a node's neighbours into two sets by flooding beacon messages, the landmark

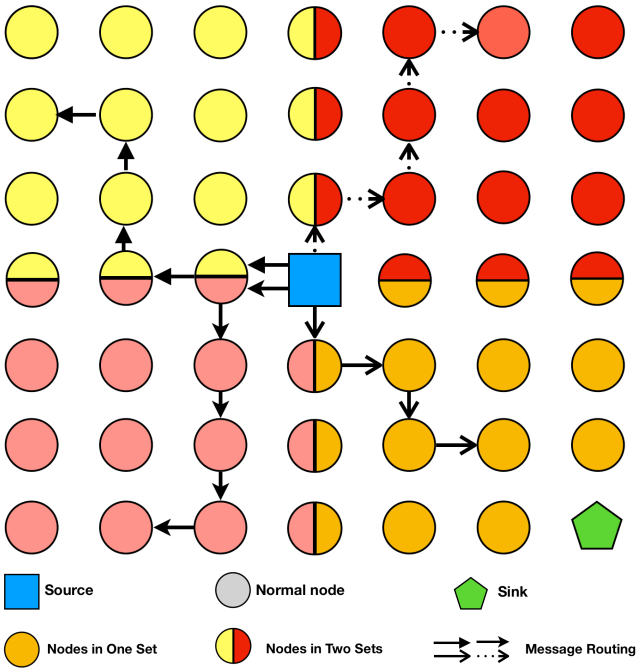


Figure 4.5: Illustration of neighbours division in phantom walkabouts

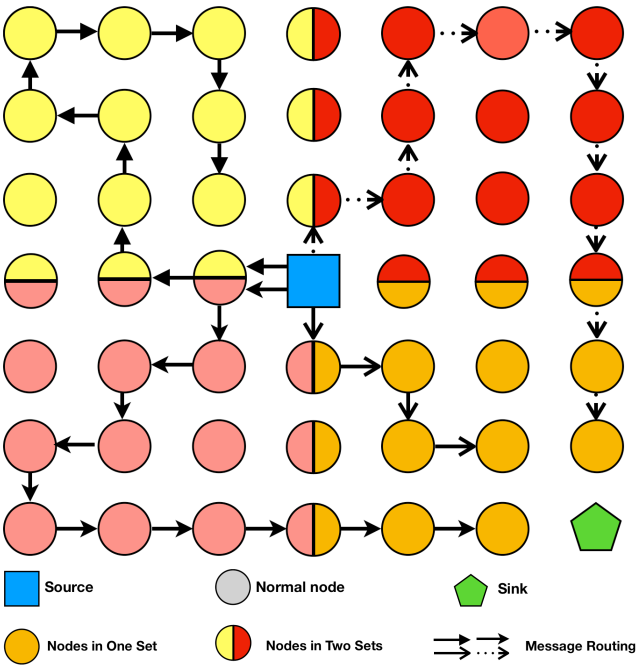


Figure 4.6: Illustration of the routing with random walk in phantom walkabouts

Algorithm 4 Random Walk Phase in Phantom Walkabouts

```
1: procedure RANDOM WALK PHASE(msg, l)
2:     msg.S_dir ← ⊥
3:     msg.M_dir ← ⊥
4:     msg.h_walk ← l
5:     msg.S_dir ← CHOOSEONESET(msg)
6:     while msg.h_walk ≠ 0 do
7:         msg.M_dir ← CHOOSEONENEIGHBOUR(msg.S_dir)
8:         if ISREACHSINK(msg) = True then
9:             msg.h_walk ← 0
10:            break
11:        end if
12:        if msg.M_dir = ⊥ then
13:            msg.M_dir ← CHOOSEONENEIGHBOUR(CloserSinkSet)
14:        end if
15:        msg.h_walk ← msg.h_walk − 1
16:        FORWARDMESSAGE(msg.M_dir)
17:    end while
18: end procedure
```

node in phantom walkabouts also floods beacon messages to divide a node's neighbours into another two sets, thus achieving the neighbour division into four sets[1]. In other words, totally two waves of beacon messages are flooded by two selected landmark nodes. In general, neighbours can be divided more than four sets depending on requirements [2]. More sets indicate that messages could be forwarded to different directions, hence luring the adversary to different locations in the network. Of course, more work in the deployment phase will be carried out in terms of more neighbour-division sets. In the implementation of the new random walk algorithm, the corner node is chosen as a landmark node (e.g., node located in the corner). Figure 4.4 and Figure 4.5 demonstrate the difference between neighbours division. The benefit of the division is: Given a SinkCorner configuration, in the random walk phase of phantom routing half of the nodes in the network could be used as phantom nodes whereas almost all the nodes could become phantom nodes in the new random walk algorithm.

[1]Observe that this does not restrict the network configuration to be a grid, but the nodes can be partitioned into these four sets.

[2]Due to the case that nodes are not equipped with GPS, more landmark nodes need to be selected to divide neighbours into multiple sets.

Algorithm 5 Flooding Phase in Phantom Walkabouts

```
1: procedure FLOODING PHASE(msg)
2:     if msg.h_walk = 0 then
3:         if ISREACHSINK(msg) = False then
4:             FLOODING(msg)
5:         end if
6:     end if
7: end procedure
```

Random walk termination handling in random walk phase

If a message is blocked during the random walk phase (e.g., messages reach the borderline of the network), the random walk stops in the phantom routing. However, in the new algorithm, a new direction will be chosen and assigned to the *CloserSinkSet*. This procedure is demonstrated in Figure 4.6. In certain extreme situations when *CloserSinkSet* is void, the random walk terminates and the node becomes the phantom node. Because it is believed that the phantom node is farthest from the real source and ensures the safety of the source node that its location will be hard to track. When a message travels l hops (assuming random walk length is l), it has finished the random walk phase.

Similar to phantom routing, if a message does not reach the sink, the flooding phase will start once the random walk phase ends. The algorithms of these two are shown in Algorithm 4 and Algorithm 5.

4.3.2 Biased Random Walk Routing in Phantom Walkabouts

The routing with random walk, especially a long random walk ensures that phantom nodes are far away from the real source. However, there is a weakness that needs to be addressed for certain configurations. Specifically, consider the *SourceCorner* configuration, messages will always be transmitted towards the sink. Owing to the random nature of the walk, the random walk may "go through" the sink. In this case, the attacker will notice the message and will move towards the source, increasing the chance of a source capture. Figure 4.7

shows the issue.

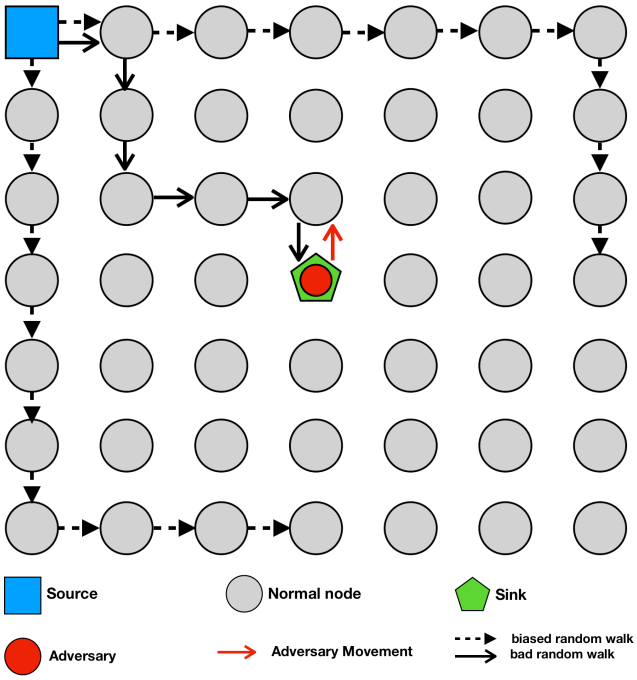


Figure 4.7: Illustration of bad random walks and biased random walks in the SourceCorner configuration

To address this issue, a *biased random walk* is developed, for the specific configuration so as to avoid the risk of a random walk close to the attacker. In the biased random walk messages are not forwarded in the direction of the sink. Instead, messages are transmitted by following borderline nodes to avoid being captured in the random walk phase. The routing with biased random walk is described as follows and shown in Algorithm 6 and Algorithm 5.

- The source first chooses a set out of four neighbour sets, and then assigns a direction from the chosen set for a message. The chosen direction is called the biased direction ($\mathcal{B}_{dir}$). The message direction $\mathcal{M}_{dir}$ is always following the $\mathcal{B}_{dir}$.
- When a node receives a message, the random value $r \in [0, 1]$ is generated. The fixed parameter $\mathcal{P}_{biased}$ is set in the experiments to make sure a message has a high probability of walking along the previous biased

direction. Normally the value of $\mathcal{P}_{biased}$ is set larger than 0.5 but less than 1. For instance, if $\mathcal{P}_{biased}$ is set to 0.7, it indicates the message has a 70% probability of being transmitted along the previous biased direction. The node decides the message direction $\mathcal{M}_{dir}$ by the following equation:

$$\mathcal{M}_{dir}(r, \mathcal{P}_{biased}, \mathcal{S}_{dir}, \mathcal{B}_{dir}) = \begin{cases} \mathcal{B}_{dir} & \text{if } r \in [0, \mathcal{P}_{biased}] \\ \mathcal{S}_{dir} \setminus \{\mathcal{B}_{dir}\} & \text{otherwise} \end{cases} \tag{4.1}$$

- When the message direction is blocked, it indicates that the message reaches the end of this direction. The message will choose a new biased direction to continue the random walk until random walk finishes. If the new biased random walk direction is empty again, the random walk phase stops. Then the flooding phase starts.

There is one remaining issue: The source cannot recognise where the sink is. In other words, the source does not know the network configuration thus it does not know whether the biased random walk should be applied. This issue is addressed as follows: In a grid network, three corner landmark nodes are chosen to flood beacon messages in the deployment stage[3]. After three waves of flooding, the distance between the sink to each landmark node is calculated and sent to the source through flooding. In reality, the sink is not exactly located in the centre of the network, so biased random walk is used for the network configuration where the sink is located in the small centre area. Assuming that three landmark nodes are n_1, n_2, n_3 and distances between the sink and each corresponding landmark node are Δ_{n1}, Δ_{n2}, Δ_{n3}, the biased random walk is used when the distance fulfils Equation 4.2, where Δ is a small threshold for defining the centre area in the network. The solution of how to determine the network configuration is also explained graphically in Figure 4.8.

[3]The location and number of landmark nodes varies for different network configurations.

Algorithm 6 Biased Random Walk in Phantom Walkabouts

```
1: procedure BIASED RANDOM WALK(msg, l, P_biased)
2:     msg.S_dir ← ⊥
3:     msg.M_dir ← ⊥
4:     msg.B_dir ← ⊥
5:     msg.h_walk ← l
6:     msg.S_dir ← CHOOSEONESET(msg)
7:     while msg.h_walk ≠ 0 do
8:         msg.M_dir ← CHOOSEONENEIGHBOUR(msg.S_dir)
9:         if msg.B_dir = ⊥ then
10:            msg.B_dir ← msg.M_dir
11:        end if
12:        r ← GENERATERANDOMNUMBER(0, 1)
13:        if ISREACHSINK(msg) = True then
14:            msg.h_walk ← 0
15:            break
16:        end if
17:        if r ≥ P_biased then
18:            msg.M_dir ← CHOOSEONENEIGHBOUR(msg.S_dir \ msg.B_dir)
19:        end if
20:        if msg.M_dir = ⊥ then
21:            msg.M_dir ← CHOOSEONENEIGHBOUR(msg.S_dir)
22:            msg.B_dir ← msg.M_dir
23:        end if
24:        msg.h_walk ← msg.h_walk − 1
25:        FORWARDMESSAGE(msg.M_dir)
26:    end while
27: end procedure
```

$$|\Delta_{n1} - \Delta_{n2}| + |\Delta_{n1} - \Delta_{n3}| + |\Delta_{n2} - \Delta_{n3}| \leq \Delta (\Delta \in \mathbb{Z}^{+}) \tag{4.2}$$

4.3.3 Phantom Walkabouts

This section formalises the phantom walkabouts technique, which extends the phantom routing protocol by adopting variable lengths of phantom routing. When the source routes a message M using phantom walkabouts, a decision is needed regarding whether M goes on a short (M_s) or long (M_l) random walk route. The sequencing of messages looks like as follows:

$$\underbrace{M_s, \cdots, M_s}_{m}, \underbrace{M_l, \cdots, M_l}_{n}, \underbrace{M_s, \cdots, M_s}_{m}, \underbrace{M_l, \cdots, M_l}_{n}, \cdots$$

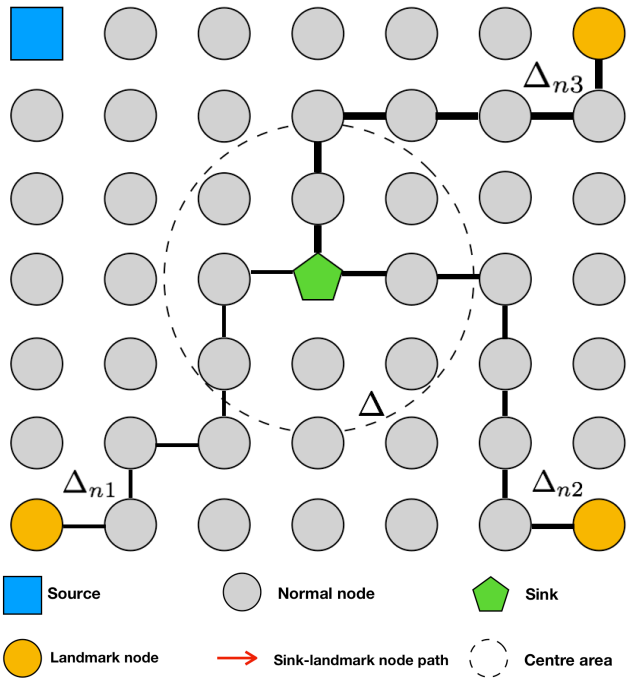


Figure 4.8: Illustration of how the source determines the network configuration using landmark nodes. Landmark nodes notify Δ_{n1}, Δ_{n2}, Δ_{n3} to the source by flooding. Then the source knows the network configuration through Equation 4.2.

Therefore, phantom walkabouts consists of m messages on short random walks and n messages on long random walks, before the cycle is repeated. For example, the messages sequences in $PW(1,2)$ are as follows:

$$\underbrace{M_s}_{1}, \underbrace{M_l, M_l}_{2}, \underbrace{M_s}_{1}, \underbrace{M_l, M_l}_{2}, \underbrace{M_s}_{1}, \cdots$$

Phantom walkabouts adopts all the techniques described in Subsection 4.3.1, Subsection 4.3.2 and Subsection 4.3.3. The phantom walkabouts algorithm is shown in Algorithm 7.

Algorithm 7 Phantom Walkabouts

```
1: procedure PHANTOM WALKABOUTS(m, n, PW(m,n))
2:     m', n' ← m, n
3:     while True do
4:         if m' > 0 then
5:             msg ← GENERATEMESSAGE()        ▷ The message contains short random walk
6:             ROUTING(msg)
7:             m' ← m' − 1
8:         else if m' = 0 ∧ n' > 0 then
9:             msg ← GENERATEMESSAGE()        ▷ The message contains long random walk
10:            ROUTING(msg)
11:            n' ← n' − 1
12:        else
13:            m', n' ← m, n
14:        end if
15:    end while
16: end procedure
```

4.3.4 Summary: Difference between Phantom Routing and Phantom Walkabouts

Before introducing the theory of phantom walkabouts, the section briefly summarises the difference between the phantom routing and phantom walkabouts shown in Table 4.2.

Difference	Phantom Routing	Phantom Walkabouts
The length of random walk	The length of random walk is fixed to a few hops.	1. The messages contain long random walk which exceeds sink-source distance. 2. The short and long random walks repeat in the phantom walkabouts
The neighbour sets of a node	The neighbours of a node are classified into two sets: the *CloserSinkSet* and the *FurtherSinkSet*	More than two neighbour sets of a node are classified depending on the choice of the landmark nodes in the network (e.g., four sets in this chapter)
The number of phantom nodes	Half the nodes in the network can be chosen as the phantom nodes (except for the source and the sink)	All the nodes in the network can be chosen as the phantom nodes (except for the source and the sink)
Random walk termination	Random walk stops in some exceptional cases (e.g., a message reaches the border node and cannot be forwarded)	Random walk continues when facing exceptional cases
Random walk techniques in different network configurations	The fixed random walk technique is used for all network configurations	Biased random walk technique is used for special network configuration (e.g., *SourceCorner* configuration)

Table 4.2: The Difference between phantom routing and phantom walkabouts

4.4 Problem Statement

The problem to be addressed is: In a WSN, phantom walkabouts is used to deliver messages from the source(s) to the sink. When an attacker is initially located at the sink and starts receiving messages sent by the source(s) to the sink, an important problem is to analyse the impact on SLP and associated performance attributes (i.e., capture ratio, receive ratio, latency and messages sent) of phantom walkabouts under various parameterisations. Formally, the problem specification is shown in Figure 4.9.

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 phantom walkabouts protocol $PW(m, n)$ where a pair (m, n) for short and long random walk lengths,
- A distributed eavesdropper $\mathcal{A}$ that is located at the sink initially,
- A set of source locations $\mathbb{L}$,
- A set of network configurations $\mathbb{C}$ and
- A safety period P_{safety},

Objective:

- Evaluate the attributes performance of $PW(m, n)$ with various parameterisations of (m, n) over P_{safety} in the presence of $\mathbb{L}$ sources in G, $\mathbb{C}$ and $\mathcal{A}$.

Figure 4.9: Problem statement: Evaluation of phantom walkabouts with various parameterisations

4.5 Experimental Setup

A square grid network layout of size $n \times n$ was used in all experiments, with $n \in \{11, 15, 21, 25\}$, i.e., networks with 121, 225, 441 and 625 nodes respectively. The node neighbourhoods were generated using an ideal radio model and the noise model was created using the meyer-heavy noise sample file provided with TOSSIM.

The source period was set to be either 0.25, 0.5, 1.0, or 2.0 second(s) (i.e., 4, 2, 1, 0.5 messages were sent from sources per second correspondingly). In Subsection 4.3.2, the parameter $\mathcal{P}_{biased}$ and Δ are introduced to implement biased random walks. The larger the value of $\mathcal{P}_{biased}$ is, the greater chance that the random walk follows the biased direction. The simulation set the value of $\mathcal{P}_{biased}$ to 0.9 and Δ to 5 hops.

When choosing the length of the short and long random walks for phantom walkabouts, a variety of parameter combinations were considered. The experiments set the short random walk series $S = \{2, 3, \ldots, 0.5 \times \Delta_{ss}\}$, and long random walk series $L = \{2 + \Delta_{ss}, \ldots, 1.5 \times \Delta_{ss}\}$, where Δ_{ss} is the sink-source distance. In the phantom walkabouts, short and long random walks were randomly generated from S and L during simulation runtime. For the phantom routing, the random walk length was fixed to 5 hops. The safety period was calculated with the safety factor 1.3 from Equation 3.1.

4.6 Simulation Results

As explained earlier, a message with short random walk will initially direct the attacker towards the source while a long random walk will direct the attacker away from the source, thereby possibly increasing the SLP. This section seeks to determine whether the hypotheses hold. In each graph from Figure 4.10 to Figure 4.17, the first row shows the results of phantom routing, followed by results of phantom walkabouts $PW(m, n)$ by varying m and n. The results of phantom routing are used as a baseline for comparison.

4.6.1 Baseline: Phantom Routing with Multiple Sources

The section first establishes the base case against which subsequent improvements will be evaluated. In the previous work on phantom routing, the length of the random walk has typically been small, less than the sink-source distance. Figure 4.10a to Figure 4.17a contain the results of phantom routing.

Capture Ratio: The results of SLP are shown in Figure 4.10a and Figure 4.11a. Two important observations are made:

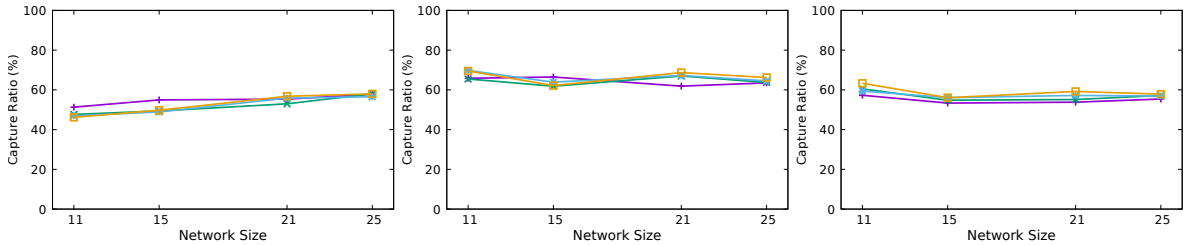
- In the SinkCorner configuration, multiple sources have the same SLP. The situation is opposite to the SourceCorner configuration, where the capture ratio decreases with an increasing number of sources.
- For both SinkCorner and SourceCorner configuration, the SLP is not subject to different source periods.

Receive Ratio: Figure 4.12a and Figure 4.13a contain the results.

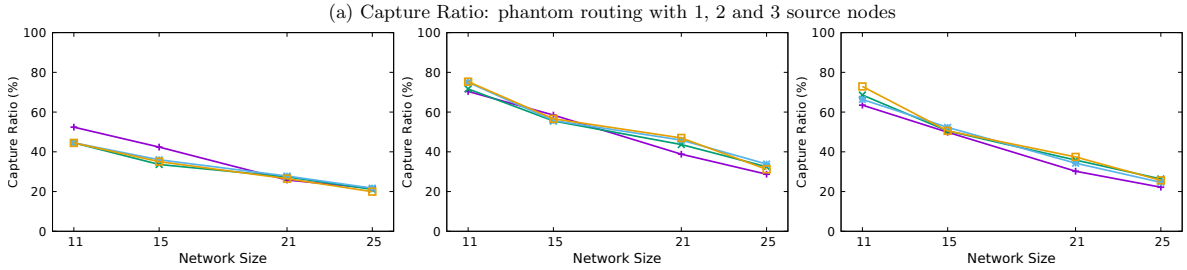
- In the SinkCorner configuration, the receive ratio is at a high level (approximately 80%) and a low source period results in the lowest receive ratio. With an increase in the number of sources, the receive ratio slightly decreases (e.g., to 70% with 3 sources). This outcome is due to messages collision.
- In the SourceCorner configuration, one source yields the highest receive ratio. However, the receive ratio decreases to 60% with 3 sources, which is worse than the SinkCorner configuration.

Messages Sent: From Figure 4.14a and Figure 4.15a, the number of messages increases with increasing network sizes. It can also be observed that the number of messages transmitted is different at various source periods. However, the number of messages transmitted gets smaller with an increasing number of sources. This is due to the fact that a lower receive ratio limits the number of messages that can be transmitted.

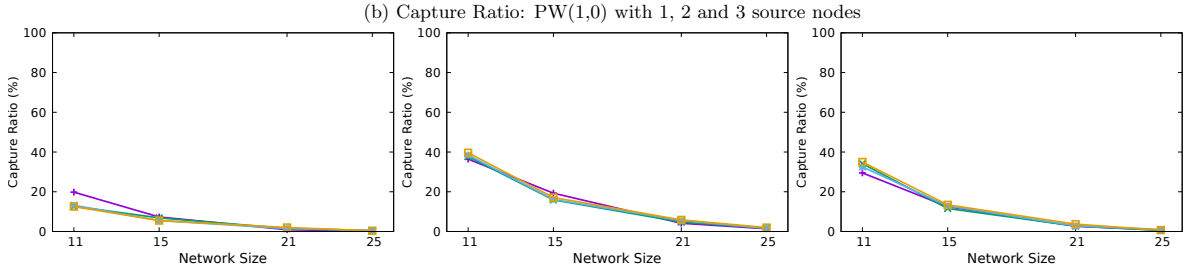
Message Latency: In the SinkCorner configuration, the latency is between 75 and 150 milliseconds for 1, 2 and 3 sources. In addition, a low source period causes the relatively low latency in the network. However, in the SourceCorner configuration, the latency is between 50 to 125 milliseconds, which is lower than the SinkCorner configuration. All the source periods generate the same level of message latency.



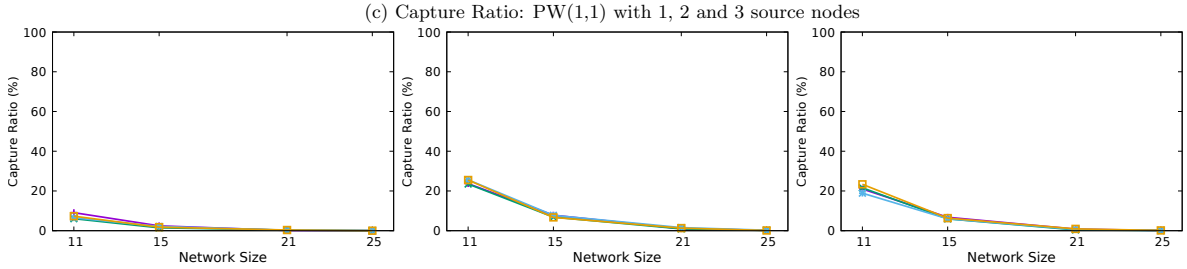
(a) Capture Ratio: phantom routing with 1, 2 and 3 source nodes



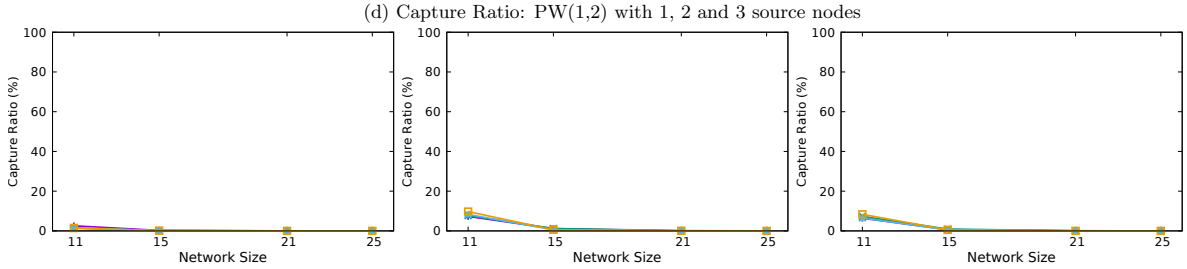
(b) Capture Ratio: PW(1,0) with 1, 2 and 3 source nodes



(c) Capture Ratio: PW(1,1) with 1, 2 and 3 source nodes

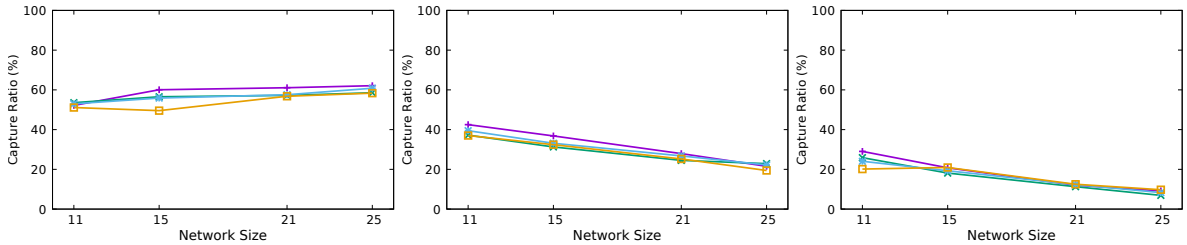


(d) Capture Ratio: PW(1,2) with 1, 2 and 3 source nodes

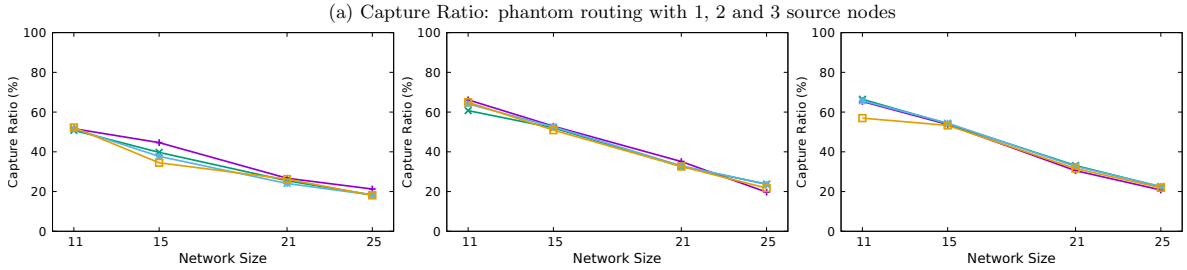


(e) Capture Ratio: PW(0,1) with 1, 2 and 3 source nodes

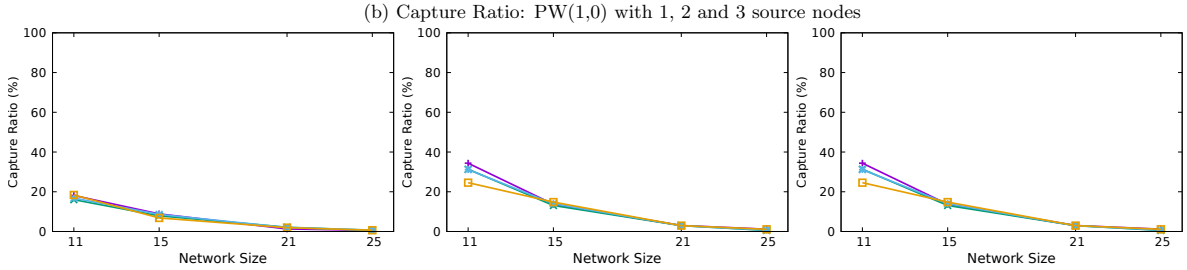
Figure 4.10: SLP level of protocols for 1, 2 and 3 sources respectively in SinkCorner configuration



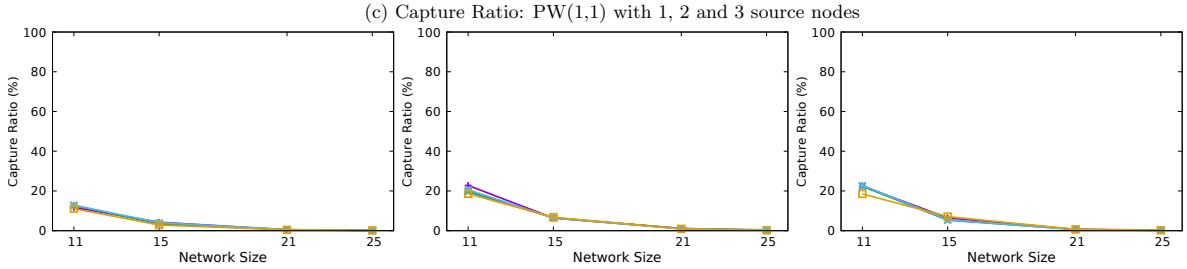
(a) Capture Ratio: phantom routing with 1, 2 and 3 source nodes



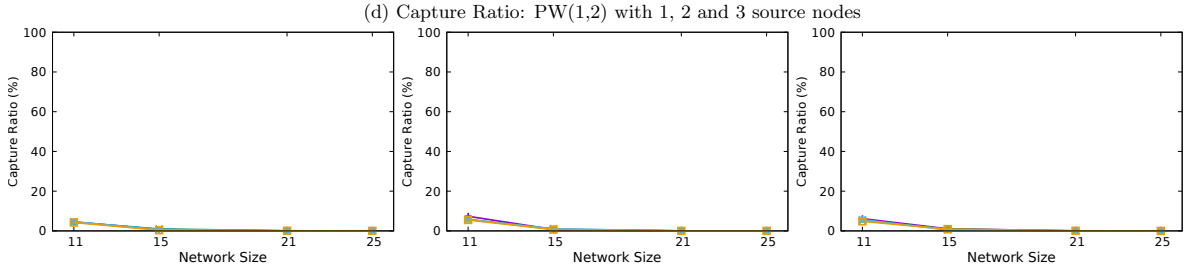
(b) Capture Ratio: PW(1,0) with 1, 2 and 3 source nodes



(c) Capture Ratio: PW(1,1) with 1, 2 and 3 source nodes



(d) Capture Ratio: PW(1,2) with 1, 2 and 3 source nodes



(e) Capture Ratio: PW(0,1) with 1, 2 and 3 source nodes

Figure 4.11: SLP level of protocols for 1, 2 and 3 sources respectively in SourceCorner configuration

4.6.2 PW(1,0): SLP with Multiple Sources using Short Random Walks

It has been shown that phantom routing does not provide a high SLP level with small random walks. $PW(1,0)$ means only short random walks are adopted in the network. The main differences between phantom routing and $PW(1,0)$ are (i) neighbour nodes divisions and (ii) the randomised length of short random walks in the random walk phase.

Capture Ratio: Figure 4.10b and Figure 4.11b contain the results of phantom routing. Two important observations are made:

- In the SinkCorner configuration with 1 source, the SLP increases (i.e., capture ratio decreases) with decreasing source periods. This is intuitive in the sense that the capture ratio can be expected to be higher as more messages are sent by the nodes and can be captured by the attacker. However, the SLP increases with an increasing source period with multiple sources. The conjuncture is that a higher number of messages leads to a much lower safety period, meaning that it is difficult for an attacker to capture the source within the safety period.

- $PW(1,0)$ provides a better SLP than phantom routing in both configurations with large network sizes. For instance, in the SinkCorner configuration with 25×25 network size and 3 sources, the capture ratio is 20% in $PW(1,0)$, while the capture ratio remains 60% in phantom routing. Therefore, $PW(1,0)$ yields high SLP in a large network size rather than a small network size. Besides, generally $PW(1,0)$ performs better in the SourceCorner configuration than SinkCorner configuration.

Receive Ratio: As the different algorithm is used in the $PW(1,0)$, receive ratio will also need to be investigated. From Figure 4.12b and Figure 4.13b, the following results can be observed:

- $PW(1,0)$ and phantom routing have the same level of receive ratio in the SinkCorner configuration. Another observation is that in the SourceCorner

configuration the receive ratio of $PW(1,0)$ is better than the SinkCorner configuration. For instance, the receive ratio is 60% in phantom routing with 3 sources, while the receive ratio is over 80%. The reason is due to the scatter of phantom nodes in $PW(1,0)$.

- For both configurations, the receive ratio decreases with an increase of sources.

Messages Sent: From Figure 4.14b and Figure 4.15b, messages sent of $PW(1,0)$ in both configurations are lower than phantom routing, especially in a low source period. For instance, from Figure 4.14b, when the source period is 0.25 second in the 25×25 network size, nodes transmit over 2000 messages in phantom routing, whereas less than 1500 messages are transmitted in $PW(1,0)$. This happens due to the unreliability of network links [12], causing a proportion of messages to never reach the phantom nodes during the random walk phase. The messages are thus not flooded, thereby reducing the expected number of messages being sent.

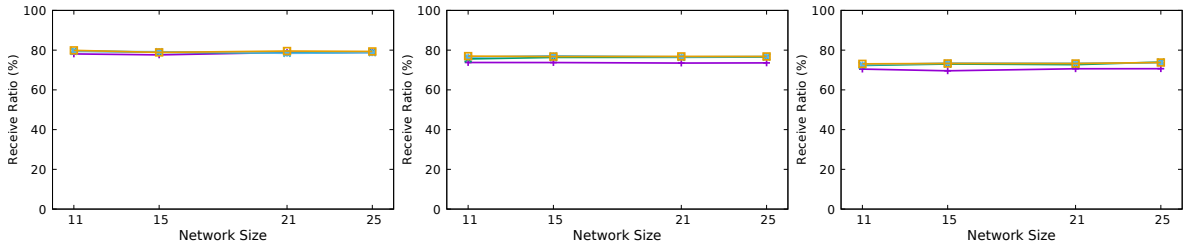
Message Latency: The latency is very similar to phantom routing for both configurations and different source periods, which indicates that $PW(1,0)$ will not cause a high message latency.

4.6.3 PW(1,1): SLP with Multiple Sources using Alternating Short and Long Random Walks

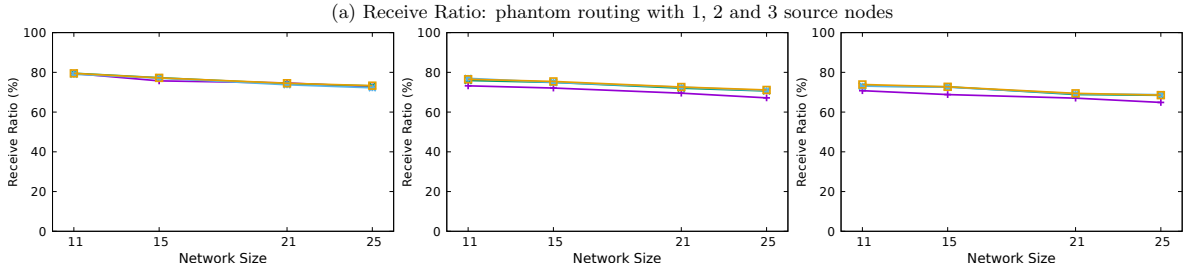
As has been shown, phantom routing and $PW(1,0)$ with random walk yields lower SLP. To try and achieve a better SLP level, this section considers the case where a short and a long random walk are chosen alternately, termed as *phantom walkabouts*. The results are shown from Figure 4.10c to Figure 4.17c.

Capture Ratio: Alternating between a short and a long random walk in phantom walkabouts, $PW(1,1)$ yields better SLP than phantom routing. Several observations can be made from Figure 4.10c and Figure 4.11c:

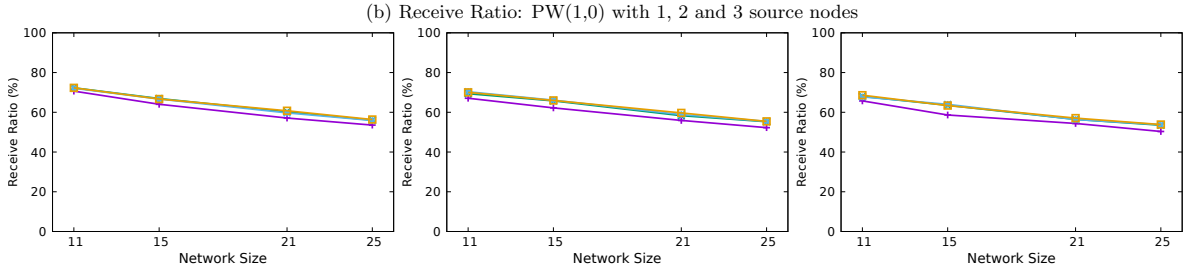
- $PW(1,1)$ yields, in general, a much higher level of SLP than phantom



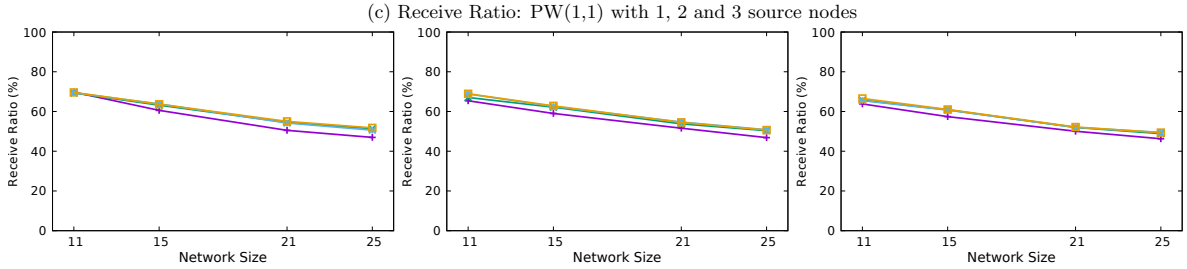
(a) Receive Ratio: phantom routing with 1, 2 and 3 source nodes



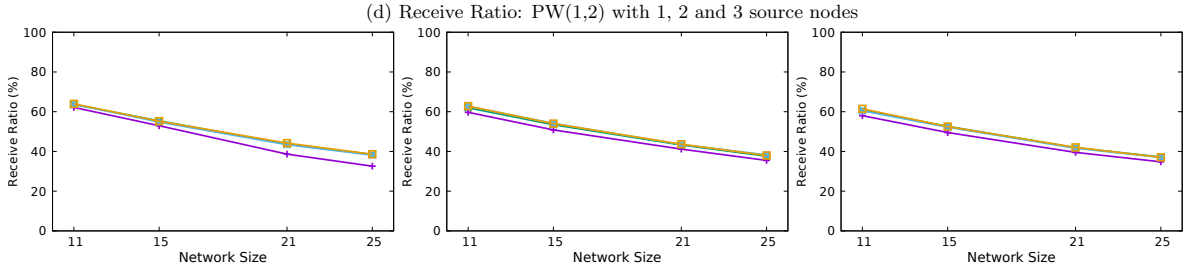
(b) Receive Ratio: PW(1,0) with 1, 2 and 3 source nodes



(c) Receive Ratio: PW(1,1) with 1, 2 and 3 source nodes

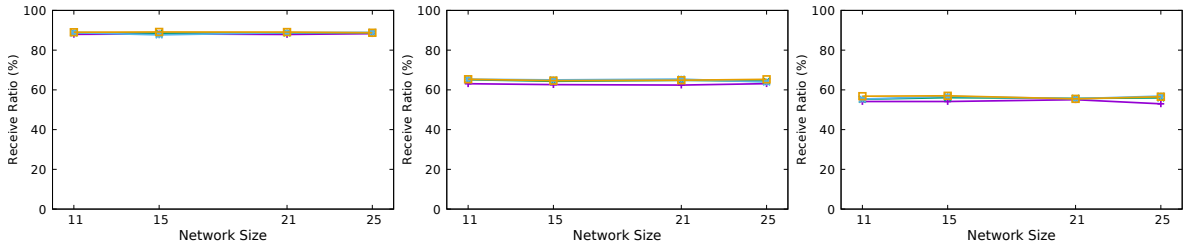


(d) Receive Ratio: PW(1,2) with 1, 2 and 3 source nodes

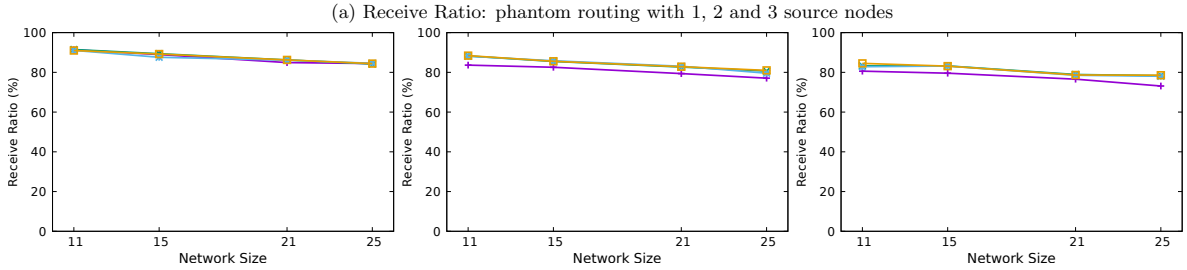


(e) Receive Ratio: PW(0,1) with 1, 2 and 3 source nodes

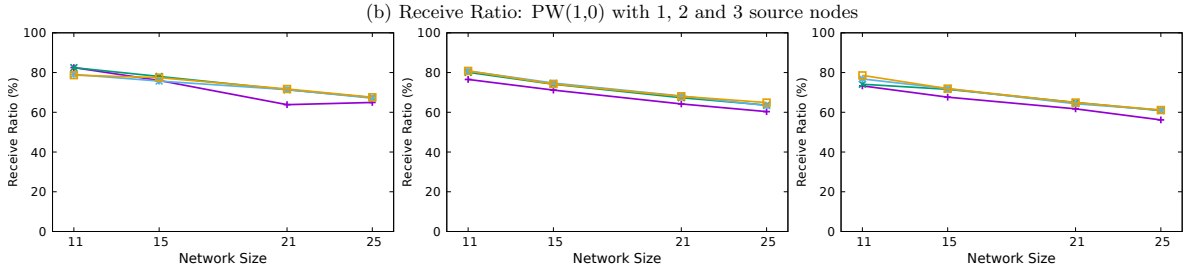
Figure 4.12: Receive ratio of protocols for 1, 2 and 3 sources respectively in SinkCorner configuration



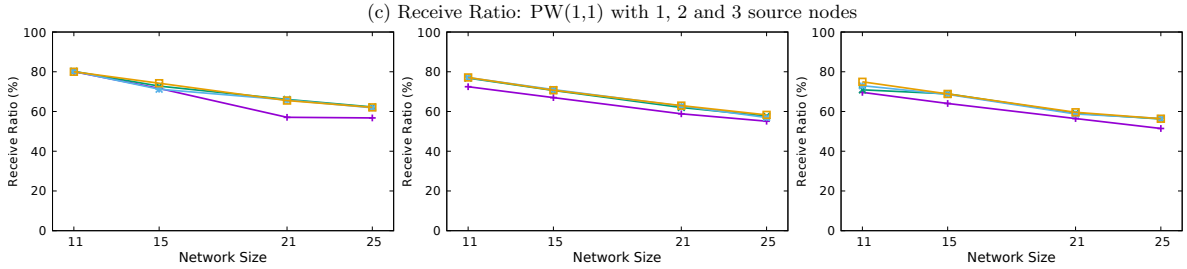
(a) Receive Ratio: phantom routing with 1, 2 and 3 source nodes



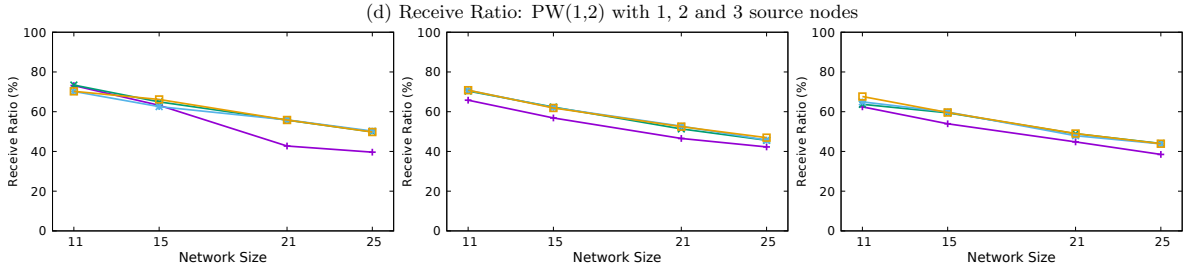
(b) Receive Ratio: PW(1,0) with 1, 2 and 3 source nodes



(c) Receive Ratio: PW(1,1) with 1, 2 and 3 source nodes



(d) Receive Ratio: PW(1,2) with 1, 2 and 3 source nodes



(e) Receive Ratio: PW(0,1) with 1, 2 and 3 source nodes

Figure 4.13: Receive ratio of protocols for 1, 2 and 3 sources respectively in SourceCorner configuration

routing, especially for the larger-sized networks. The capture ratio reduces at least 20% compared to phantom routing.

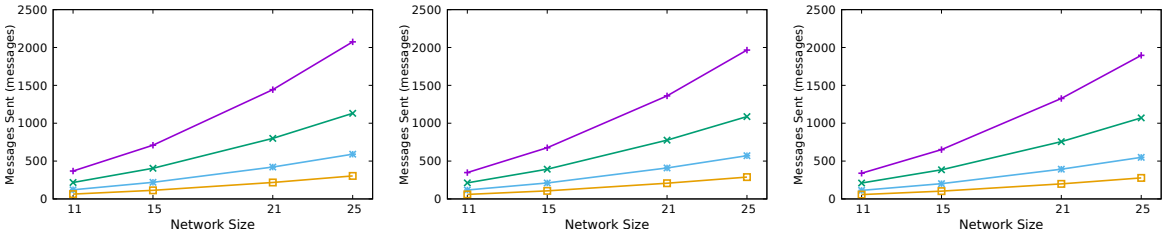
- $PW(1,0)$ in the SourceCorner configuration yields better SLP than the SinkCorner configuration.
- In general, the network with 2 sources yields the worst SLP.

Receive Ratio: It has been shown that $PW(1,1)$ provides better SLP level than phantom routing. In addition, there is a need to investigate the high level of $PW(1,1)$ is due to the efficiency of phantom walkabouts or due to the unreliability of the network.

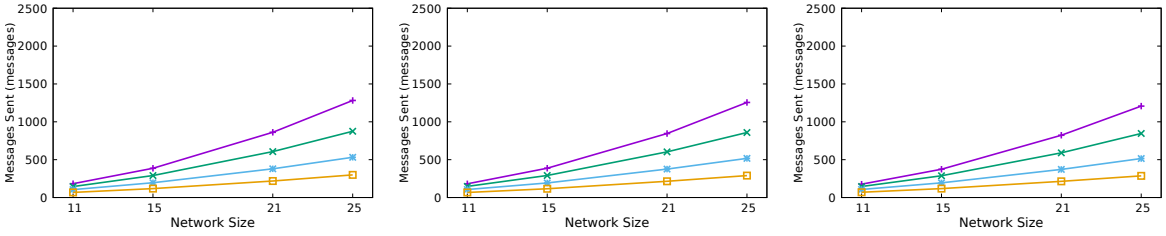
- In the SinkCorner configuration, the receive ratio in a small network size (e.g., the 11×11 network size) is 10% lower than in phantom routing. In addition, receive ratio decreases with an increase in network size and the number of sources. For instance, in the configuration of 25×25 network size and 3 sources, the receive ratio is 50% compared to 60% in the $PW(1,0)$ and 70% in the phantom routing.
- In the SourceCorner configuration with 1 source, the receive ratio is lower than phantom routing. However, multiple sources yield better receive ratio than phantom routing.

Messages Sent: The messages sent in $PW(1,1)$ is lower than both phantom routing and $PW(1,0)$. The conjecture is that the low receive ratio causes messages sent decreases.

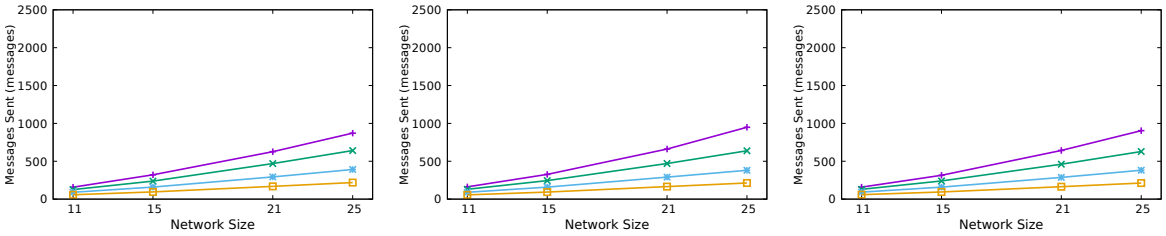
Message Latency: The result that latency is higher than phantom routing and $PW(1,0)$ is due to long random walks used in $PW(1,1)$, hence causing a longer time cost from the source to the sink.



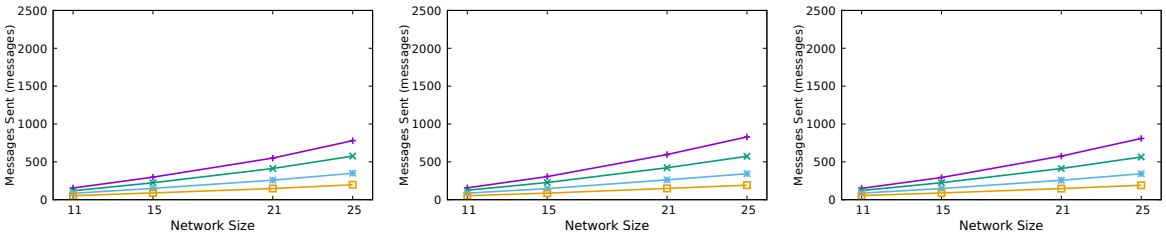
(a) Messages sent: phantom routing with 1, 2 and 3 source nodes



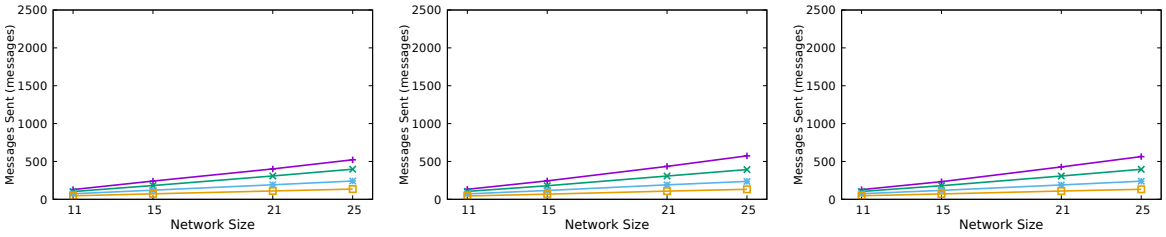
(b) Messages sent: PW(1,0) with 1, 2 and 3 source nodes



(c) Messages sent: PW(1,1) with 1, 2 and 3 source nodes

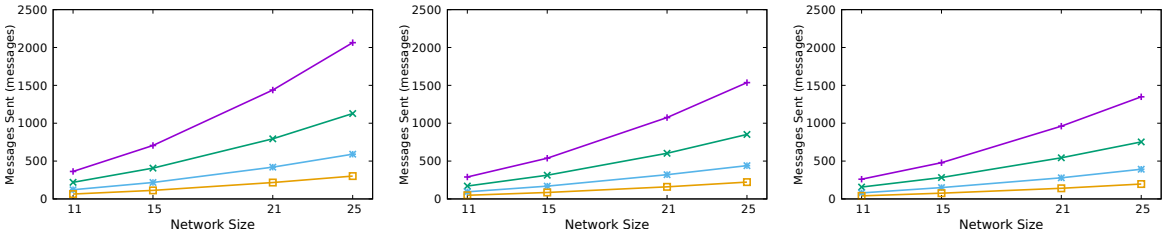


(d) Messages sent: PW(1,2) with 1, 2 and 3 source nodes

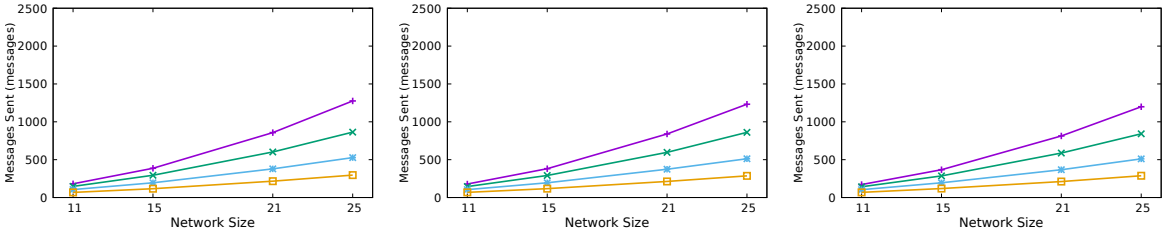


(e) Messages sent: PW(0,1) with 1, 2 and 3 source nodes

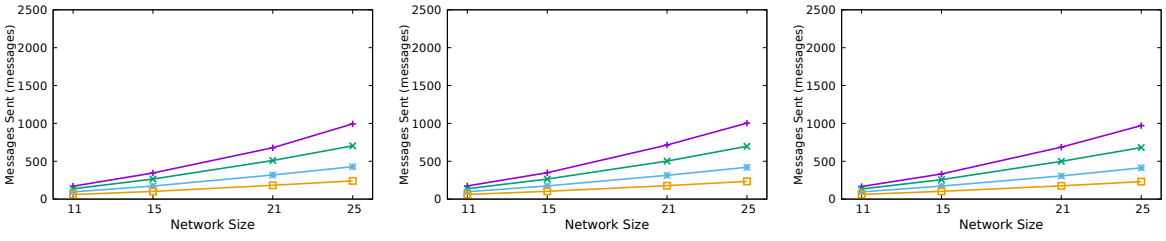
Figure 4.14: Messages sent of protocols for 1, 2 and 3 sources respectively in SinkCorner configuration



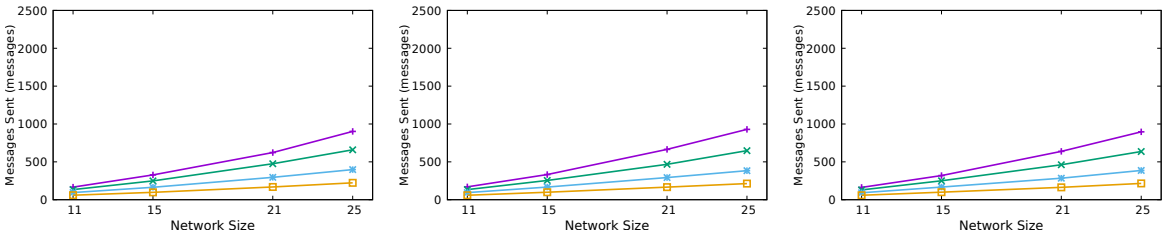
(a) Messages sent: phantom routing with 1, 2 and 3 source nodes



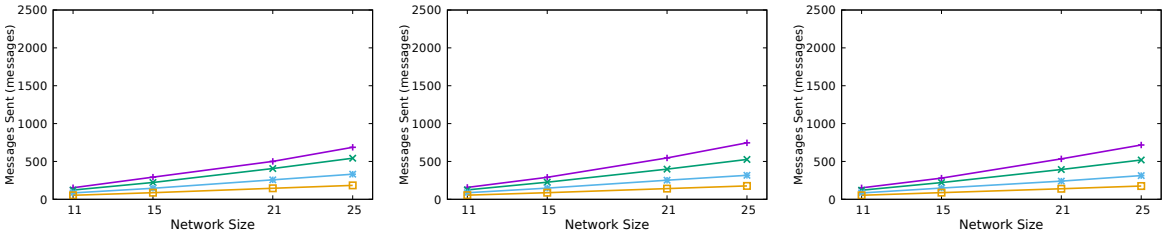
(b) Messages sent: PW(1,0) with 1, 2 and 3 source nodes



(c) Messages sent: PW(1,1) with 1, 2 and 3 source nodes



(d) Messages sent: PW(1,2) with 1, 2 and 3 source nodes



(e) Messages sent: PW(0,1) with 1, 2 and 3 source nodes

Figure 4.15: Messages sent of protocols for 1, 2 and 3 sources respectively in SourceCorner configuration

4.6.4 PW(1,2): SLP with Multiple Sources using Alternating One Short and Two Long Random Walks

To further investigate the efficacy provided by phantom walkabouts, the new routing mechanism consisted of messages being routed through a short random walk followed by two long random walks. Subsequent routing then follows this cycle. Owing to the two long random walks in phantom walkabouts, $PW(1, 2)$ produces better SLP.

Capture Ratio: It can be observed that this routing mechanism provided a big improvement in capture ratio over the base case (Figure 4.10d and Figure 4.11d). The capture ratio reduces at least 40% compared to phantom routing in the SinkCorner configuration.

Receive Ratio: From Figure 4.12d and Figure 4.13d, although there is huge improvement in the SLP over phantom routing, there is at least 10% reduction in the receive ratio. Furthermore, the receive ratio decreases to 50% with 3 sources.

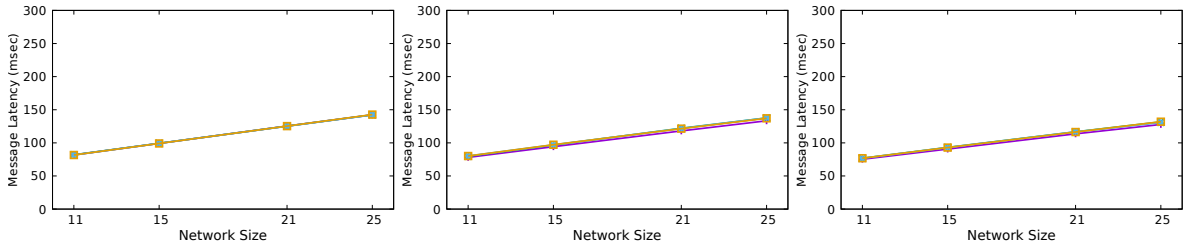
Messages Sent: In both configurations, the messages sent in $PW(1, 2)$ is lower than phantom routing. This reason is also due to the low receive ratio.

Message Latency: The latency increases from 20 to 30 milliseconds compared to $PW(1, 1)$ as extra one long random walk is used in the each repeat of phantom walkabouts.

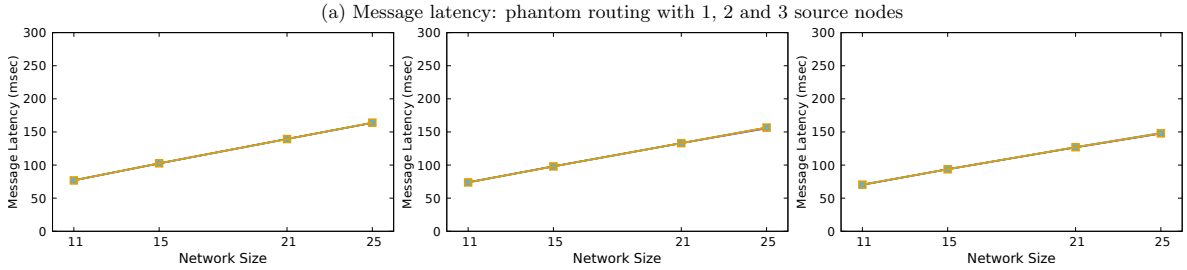
4.6.5 PW(0,1): SLP with Multiple Sources using Long Random Walks

Finally, the section investigates the case that only long random walks exist in phantom walkabouts (i.e., $PW(0, 1)$). From previous knowledge, $PW(0, 1)$ produces the best SLP but yields a poor receive ratio.

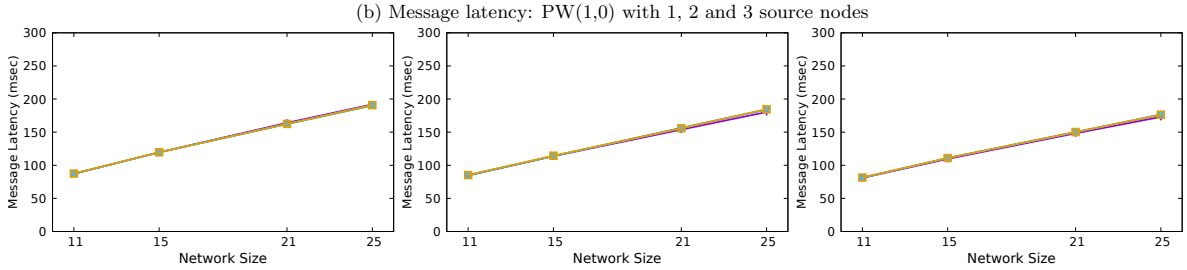
Capture Ratio: As can be observed from Figure 4.10e and Figure 4.11e, the level of SLP provided with a longer random walk is much higher than phantom routing, thereby corroborating the hypothesis. In both configurations, the



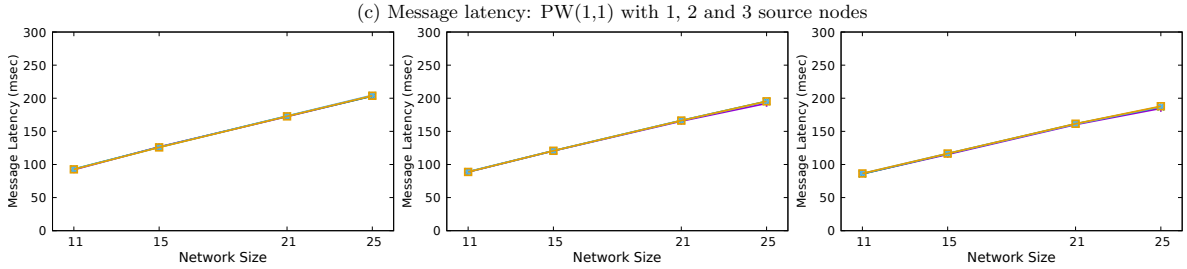
(a) Message latency: phantom routing with 1, 2 and 3 source nodes



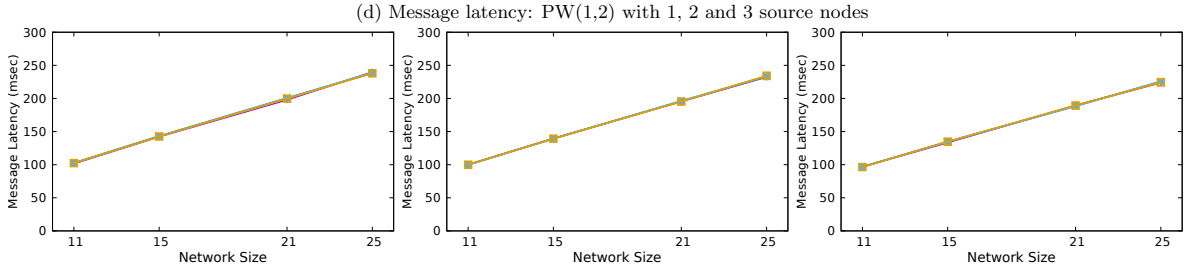
(b) Message latency: PW(1,0) with 1, 2 and 3 source nodes



(c) Message latency: PW(1,1) with 1, 2 and 3 source nodes

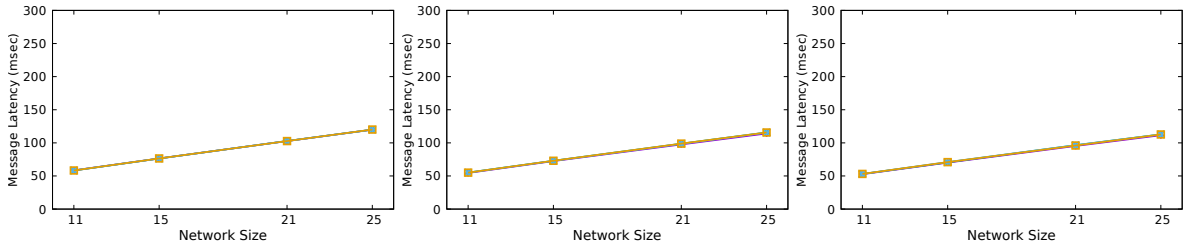


(d) Message latency: PW(1,2) with 1, 2 and 3 source nodes

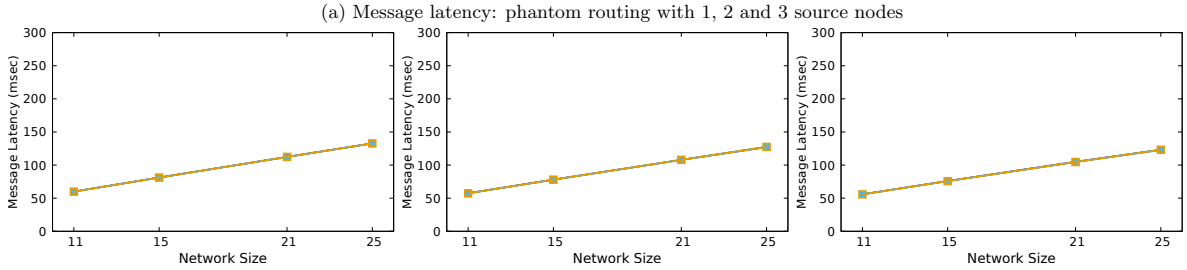


(e) Message latency: PW(0,1) with 1, 2 and 3 source nodes

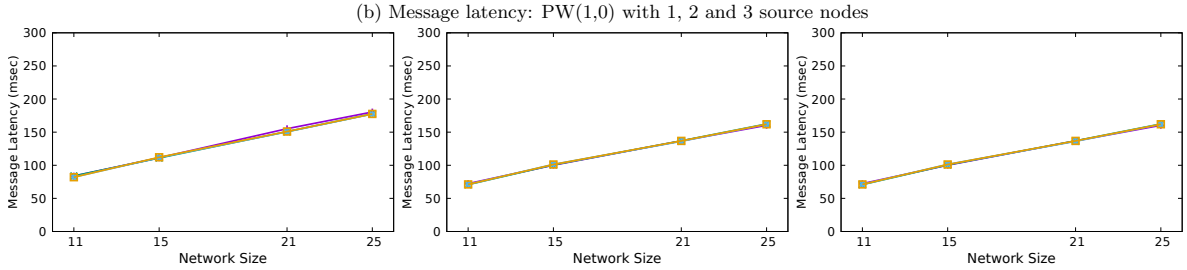
Figure 4.16: Message latency of protocols for 1, 2 and 3 sources respectively in SinkCorner configuration



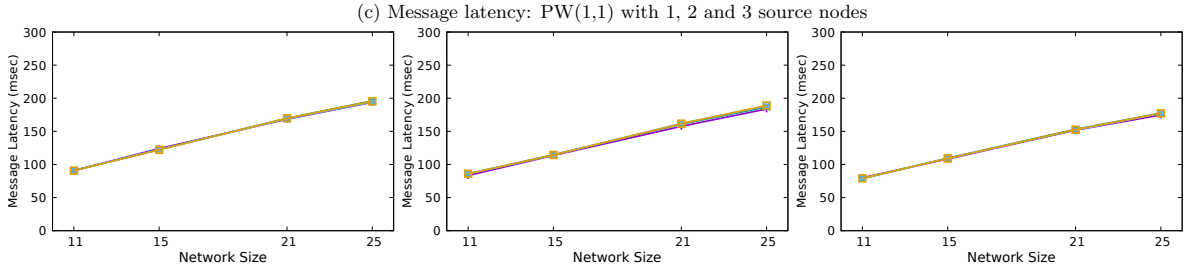
(a) Message latency: phantom routing with 1, 2 and 3 source nodes



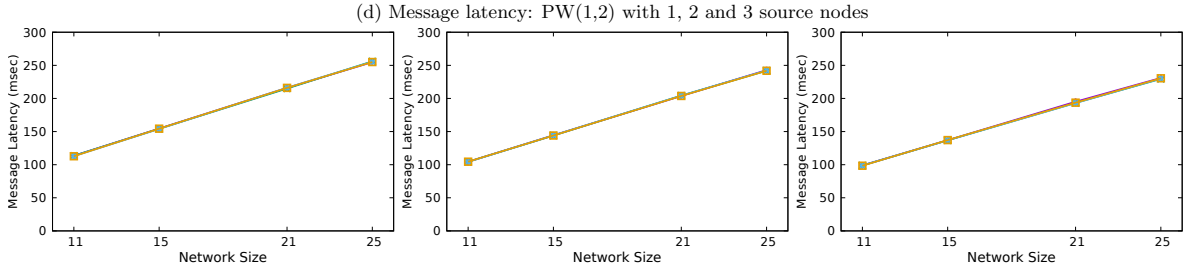
(b) Message latency: PW(1,0) with 1, 2 and 3 source nodes



(c) Message latency: PW(1,1) with 1, 2 and 3 source nodes



(d) Message latency: PW(1,2) with 1, 2 and 3 source nodes



(e) Message latency: PW(0,1) with 1, 2 and 3 source nodes

Figure 4.17: Message latency of protocols for 1, 2 and 3 sources respectively in SourceCorner configuration

capture ratio is no more than 10%.

Receive Ratio: On the other hand, though the capture ratio with long random walks is much lower than with short random walks, the decrease in receive ratio is around 20%, which is not nominal (see Figure 4.12e and Figure 4.13e). This shows that phantom routing, with a long random walk, offers a much higher level of SLP at the expense of decrease in the receive ratio.

Messages Sent: $PW(0,1)$ yields the lowest messages sent across all cases. Again, the reason is due to the lowest receive ratio in $PW(0,1)$.

Message Latency: The latency in $PW(0,1)$ is between 100 and 250 milliseconds, which doubles in phantom routing as all routes are with the long random walk.

4.7 Discussion

This section contains discussions about some issues and observations that arise as a result of this work.

Flooding phase in phantom walkabouts

In the phantom walkabouts, flooding is used to deliver messages from the phantom node to the sink. The reason to adopt flooding is simplicity. It requires no costly topology maintenance or complex route discovery. However, it has several shortcomings including implosion, overlap and energy consumption [62]. Therefore, the flooding protocol can be replaced by other enhanced protocols such as gossiping [14] and sensor protocols for information via negotiation (SPIN) [119]. However, the receive ratio in phantom walkabouts may decrease. Besides, the safety period needs to be recalculated based on the corresponding baseline protocol instead of baseline flooding.

Performance of receive ratio

With long random walk in phantom walkabouts, the receive ratio is at a low level due to the unreliability of network links [12], causing a proportion of messages to never reach the phantom nodes during long random walks. Therefore, there is a need to create a mechanism to provide high message delivery. Retransmission is a solution to ensure reliability along the route. It works as follows: When a message is transmitted from node $\mathcal{J}$ to a neighbour node $\mathcal{K}$, $\mathcal{K}$ may send an acknowledgement(ACK) message when it has received the message from $\mathcal{J}$. If $\mathcal{J}$ does not receive any ACK message from $\mathcal{K}$, it means that the message may not have been successfully delivered to $\mathcal{K}$. In this case, $\mathcal{J}$ will resend the same message to $\mathcal{K}$. Retransmission will stop when an ACK message is received or the maximum number of retransmission have been sent. The implementation of messages retransmitted is the next step for future research.

Trade-offs between SLP and other attributes

Phantom walkabouts provides a high level of SLP, but by reducing some attributes performance such as receive ratio. Table 4.3 compares the performance of four attributes between phantom routing and phantom walkabouts with the SourceCorner configuration and network size 11×11. These results indicate that there is trade-offs needed to be made between the capture ratio and other attributes. To obtain a better capture ratio using this technique a decrease in other attributes will be required. For many applications a decrease of this magnitude will be acceptable. For example, protecting pandas will not be adversely affected. But in some scenario where message latency or receive ratio is important, such as on a battlefield or in healthcare, the selection of which protocol must be thoroughly considered. Thus, the trade-offs should be made under the consideration of such practical scenarios.

	Capture Ratio	Receive Ratio	Messages	Latency
Phantom Routing	53.0%(-)	88.8%(-)	1854(-)	58(-)
$PW(1,0)$	51.7%(-2.5%)	91.1%(+2.6%)	2478(+33.7%)	60(+3.4%)
$PW(1,1)$	16.8%(-68.3%)	79.1%(-10.9%)	2342(+26.3%)	82(+41.3%)
$PW(1,2)$	12.8%(-75.8%)	80.2%(-9.7%)	2338(+26.1%)	91(+56.9%)
$PW(0,1)$	4.6%(-91.3%)	70.2%(-20.9%)	2113(+14.0%)	113(+94.8%)

Table 4.3: Comparison of attributes results under the given network configuration. Phantom routing is used as a baseline.

4.8 Summary

This chapter first briefly reviews some shortcomings of phantom routing which have been detailed in Chapter 3. For example, phantom routing cannot handle multiple sources in SLP. Then a novel technique called phantom walkabouts is proposed, which extends the phantom routing, to provide a better level of SLP. Phantom walkabouts utilises sequences of short random walks and long random walks to attempt to make the attacker move in the wrong direction, as opposed to phantom routing (with small random walks) where an attacker moves towards the source. The various parameterisations in phantom walkabouts provide different levels of SLP. The main contributions of this chapter are to:

- Analyse the weakness of short random walk in phantom routing and conjecture interleave sequences of short random walks and long random walks could achieve a high level of SLP.
- Propose phantom walkabouts, a novel and more general technique than phantom routing. phantom walkabouts solves several weaknesses of phantom routing and helps achieve a better SLP with multiple sources and different network configurations.
- Show the viability of phantom walkabouts through simulations. For example, under certain parameterisations, phantom walkabouts achieves extremely high SLP with acceptable decrease in other attributes.

The future plan is to investigate phantom walkabouts with dynamic short and long random walks, i.e., in the current experiments, the values are fixed for

short random walks and long random walks in one phantom walkabouts repeat. However, this needs not be the case. In the dynamic phantom walkabouts, the values of short random walks and long random walks in one phantom walkabouts repeat can be decided by the routing protocol itself in the runtime. Besides, the work has mostly focused on the SinkCorner and SourceCorner configuration. Of course, the results in this chapter are only applicable to these network configurations. Different phantom walkabouts may be necessary for other network configurations. For example, a non-grid network configuration may show very different SLP levels with $PW(m, n)$.

CHAPTER 5

A Decision Theoretic Framework for Selecting SLP-Aware Routing Protocols

5.1 Introduction

A number of techniques have been proposed to provide SLP, such as phantom routing using random walks [115, 156], message delay [15, 64], fake sources [17, 71] and many others [78, 80, 124]. In general, the objective of SLP protection can be informally stated as the provision of a high level of source location privacy while expending as little energy as possible. Thus, the various techniques above navigate this trade-off solution space. However, when several conflicting objectives are involved, navigating this space becomes more challenging. For example, very low level of message delivery yields high level of SLP (i.e., low capture ratio), but such low level of receive ratio is intolerant in the WSN as the sink cannot receive data from the source. Thus, this chapter proposes a methodology where routing protocols are first profiled to capture their performance according to a desired set of attributes. Then, a decision theoretic procedure is adopted for selecting the most appropriate SLP routing algorithm for the type of network and application under study. The results demonstrate the viability of the approach through various case studies, and show how the suitability of different SLP protocols vary according to the application under study.

In the Section 5.2, some SLP-aware routing protocols are reviewed, as well as several protocols with *minimal* SLP by contrast. All these protocols will be used to create a performance library. Section 5.3 and Section 5.4 present a decision theoretic procedure for selecting the most appropriate SLP-aware routing algorithm. Problem statement is presented in Section 5.5. In Section 5.6

an example is given to demonstrate the execution of the decision theoretic procedure. The adopted system and simulation approach are outlined in Section 5.7. Section 5.8 presents three case studies to showcase the viability of the approach taken. Section 5.9 concludes this chapter with a summary of contributions.

5.2 Routing Protocols Review

The section reviews the SLP-aware routing protocols that will be analysed in this chapter. However, the framework can be extended to handle any other SLP-aware routing protocol.

5.2.1 Protectionless Flooding and Protectionless CTP

Two routing algorithms that provide *minimal* SLP will be evaluated in this work to compare against the SLP techniques. The first is flooding, in which a source floods a message through the network by having each node that receives it forward the message. Flooding is included as seminal work demonstrated that the protocol provided minimal SLP [74]. The second is CTP [50] (the Collection Tree Protocol) which uses the expected number of transmissions to gauge the reliability of a link to form a routing tree from every node in the network to the sink. CTP is included as it is a state-of-the-art reliable routing protocol for WSNs. No work thus far has analysed its ability to provide SLP.

5.2.2 Phantom Routing

In seminal work [116], the authors proposed a solution called phantom routing, which combines a random walk phase and a baseline flooding phase to provide SLP level and reliable messages sent. The protocol works as follows: Each node maintains two sets for all its neighbours where *CloserSinkSet* contains all the neighbours whose hop counts are smaller than or equal to the nodes hop counts and *FurtherSinkSet* includes neighbours with larger hop counts. After neighbour nodes partition, source node(s) randomly pick either of two sets and send normal

messages to one neighbour in the chosen set. During the random walk phase, if a message is blocked the random walk phase stops (e.g., there is no neighbour in the chosen set, so messages cannot be forwarded). In other case, when a message travels s hops (assuming random walk length is s), it has finished the random walk phase. When the random walk phase ends, if the message does not reach the sink node, the message is flooded throughout the network so that it reaches the sink node.

5.2.3 Phantom Walkabouts

Phantom walkabouts is an algorithm using a random walk technique to provide SLP [55]. This new technique, which uses a mix of short and long random walks, achieves a higher level of SLP than phantom routing with a bounded message overhead. The phantom walkabouts parameterisation is denoted by $PW(m_s, m_l)$, where m_s and m_l denote the number of short and long random walks respectively to be performed in a cycle. When a source node routes a message M using phantom walkabouts, a decision is needed regarding whether M goes on a short or long random route. The sequencing of messages is as follows:

$$\underbrace{M_s, \cdots, M_s}_{m_s}, \underbrace{M_l, \cdots, M_l}_{m_l}, \underbrace{M_s, \cdots, M_s}_{m_s}, \underbrace{M_l, \cdots, M_l}_{m_l}, \cdots$$

For instance, $PW(1, 1)$ denotes a repeating sequence of 1 short random walk followed by 1 long random walk. Therefore, phantom walkabouts consists of m_s messages on the short random walk M_s and m_l messages on the long random walk M_l, before the cycle is repeated.

5.2.4 DynamicSPR

DynamicSPR is an extended version of the dynamic fake source technique [17] in which fake sources are allocated away from the real source and sink in order to provide a *pull* in that direction. This technique dynamically determines

parameters online to be able to adjust to a changing network environment. DynamicSPR [18] optimises the way fake sources are allocated, in such a way that fake sources perform a directed random walk away from the sink. This reduces the number of fake sources present in the network and also the number of messages the technique sends (thus reducing energy usage).

5.2.5 ILP Routing

In ILP Routing [15], the problem of the SLP-aware routing of messages from a source to a sink was modelled as an Integer Linear Programming (ILP) optimisation problem. Using an ILP solver an optimal solution was obtained when trying to maximise the attacker's distance from the source. As the optimal solution requires global knowledge, the authors implemented a distributed version that had a message take a directed walk around the sink to approach it from a direction other than the one the source was in. Messages were delayed by different amounts so that they reached a similar point at a certain distance. By doing this the attacker makes less progress, due to messages being grouped at a similar location and because messages would be missed that take a different path.

5.3 Decision Theoretic Procedure Overview

Given the number of SLP-aware routing protocols, each one optimising one or more attributes, it becomes challenging to select a protocol for a given application. For example, if an application requires a high level of privacy and is supposed to run for a short time, selecting a protocol that trades-off privacy for lower energy consumption will not be suitable. Thus, there is a need to develop a framework that can guide a network or application designer in selecting the appropriate SLP-aware routing protocol.

The section concerns about structuring the preferences to simplify the trade-off analysis. As multi-attribute optimisation is concerned, this section provides

Table 5.1: Commonly used symbols

Symbol	Description
$\mathcal{NC}$	The network configuration
$\mathcal{P}$	The name of a given routing protocol
$r_{\omega}^{\mathcal{NC},\mathcal{P}}$	The result of a attribute under $\mathcal{NC}$ and $\mathcal{P}$
$R_{\omega}^{\mathcal{NC},\mathcal{P}}$	The normalised result of a attribute under $\mathcal{NC}$ and $\mathcal{P}$
$r^{\mathcal{NC},\mathcal{P}}$	The result vector of all attributes under $\mathcal{NC}$ and $\mathcal{P}$
$R^{\mathcal{NC},\mathcal{P}}$	The performance vector of all attributes under $\mathcal{NC}$ and $\mathcal{P}$
$U_{\omega}^{\mathcal{NC},\mathcal{P}}$	The utility of a single attribute under $\mathcal{NC}$ and $\mathcal{P}$
$U^{\mathcal{NC},\mathcal{P}}$	The utility of performance vector under $\mathcal{NC}$ and $\mathcal{P}$
u_a	The aspiration vector
λ_{ω}	The weight of a single attribute
Δ_{ss}	The distance in hops between the sink and the source
$\mathcal{TT}$	The time taken (seconds) of protectionless flooding
P_{safety}	The safety period (seconds)

a brief overview of the theory underpinning the generation of multi-attribute utility functions. Table 5.1 summarises the most commonly used symbols in the chapter.

5.3.1 Introduction to Decision Theory (DT)

Very often, real-world cases deal with multiple attributes. This chapter assumes that the real-world cases have n evaluators, $E_1, E_2, \ldots E_n$, evaluating attributes $a_1, a_2, \ldots a_n$ respectively, such that $(E_1(a_1), E_2(a_2), \ldots E_n(a_n)) = (q_1, q_2, \ldots, q_n)$, where each q_i captures the "performance" of the protocol for a particular attribute and the vector $(q_1, q_2, \ldots, q_n)$ is the "performance" vector of a protocol. An overall relevance function, G, may be expressed in additive form:

$$G(q_1, q_2, \ldots, q_n) = \sum_{i=1}^{n} (\lambda_i * G_i(q_i)) \tag{5.1}$$

where G_i's are single-attribute or individual relevance functions [51, 79], and $\sum_{i=1}^{n} \lambda_i = 1$ iff the attributes are *mutually preferentially independent*, i.e., trade-off between pairs of attributes is independent of the values of other attributes. Such a property is important to keep the selection "local", i.e., the trade-offs

between a pair of attributes need not consider the values of *all* other attributes. A higher value of λ_i is indicative of the higher importance of a corresponding attribute. Thus, to generate the overall relevance function, each λ_i needs to be determined, subject to the constraint $\sum_{i=1}^{n} \lambda_i = 1$. Also, each individual relevance function G_i needs to be generated by the system administrator or application developer. Determining an accurate G_i is a challenging process. An interested reader is directed to [51, 79] for more information about generating such functions, which is beyond the objectives of this thesis. *arbitrary* functions are used to showcase the decision theoretic methodology proposed.

5.3.2 Decision Theory-Based Heuristic

The section presents a novel two-step decision procedure (or heuristic) that helps to choose the most suitable SLP-aware routing algorithm from a set of contenders for a given application:

Step 1 - Profiling and Filtering

1. For various network configurations (size, safety factor, noise models etc.), run all the protocols to obtain their respective performance profiles, i.e., to generate their performance vectors. This step can be done once and the profiles are stored in a database or library.

2. Determine all decision attributes for the application (e.g., capture ratio, receive ratio or message latency). If the attributes are not mutually preferentially independent, at this point they are either transformed so that they can satisfy this property or more sophisticated techniques are required. Also, determine the network configuration which the application will be running under. This is called the *input network configuration* (or input configuration). If there is no profile associated with the input configuration, then either the protocol has to be run under this new configuration (and added to the library) or a profile exists in the library for a configuration that is close enough to the input configuration.

3. For a given input network configuration, determine a performance vector that best represents the application's requirements, i.e., determine a performance vector that captures the acceptable value boundary for each attribute. The boundary is the maximal or minimal acceptable value, depending on the attribute type. This vector is referred as the *aspiration* vector.

4. For the given input network configuration, remove all vectors that are either *dominated* by the aspiration vector since they fall short of the application's requirements for input configuration (i.e., all entries in the aspiration vector are better than the corresponding ones in the vector under consideration).

5. If there are no candidates left, go to 3. Else, for each attribute, determine the minimum and maximum values from the remaining alternatives. This is done to help in determining normalised single-attribute functions (range from 0 to 1).

Step 2 - Characterisation and Selection

1. Determine the (i) individual weights (or importance) of each attribute, (ii) individual relevance function and (iii) the overall relevance function.

2. For each algorithm (i.e., performance vector) in the set of remaining contenders, insert their attribute values in the overall relevance function to obtain their respective relevance or utility values.

3. Select the alternative with the highest relevance value.

5.4 Decision Theoretic Procedure for Selecting SLP-Aware Routing Protocols

This section imposes the decision theoretic framework on the protocols and now explain the steps in more detail.

5.4.1 Step 1: Profiling and Filtering SLP-Aware Routing Algorithms

1. Profiling the Protocols

In the first phase of Step 1, the network administrator runs every protocol under consideration under various network configurations. This step need not be repeated for every application, but is a one-time activity. These profiles (or protocol performance vectors) can then be saved or stored in a protocol library that can be used whenever a new application is developed. If a new protocol is developed, then the process is repeated for the new protocol, and its (normalised) performance profile is added to the library.

2. Determining Decision Attributes

There are four decision attributes: (i) capture ratio (cr), (ii) receive ratio (rr), (iii) message latency (lat) and (iv) messages sent ($mSent$). These four attributes could be classified into *gain* type (high value is better, e.g., receive ratio) and *cost* type (high value is worse, e.g., capture ratio, message latency, messages sent). Decision Attributes can differ depending on the applications.

For a given network configuration $\mathcal{NC}$ and given protocol $\mathcal{P}$, the result vector $r^{c,\mathcal{P}}$ can be determined experimentally (e.g., through simulations). The vector contains the recorded (raw) values of all decision attributes.

$$r^{c,\mathcal{P}} = (r^{c,\mathcal{P}}_{cr}, r^{c,\mathcal{P}}_{rr}, r^{c,\mathcal{P}}_{lat}, r^{c,\mathcal{P}}_{mSent}) \tag{5.2}$$

Please note that, since the focus in this chapter has been for SLP-awareness, the attributes of interest capture both the SLP level and WSNs performance (e.g., capture ratio, receive ratio) and these are used in the vector. However, the conjecture is that a similar heuristic can be used but for different objectives, requiring a different set of attributes. An example of a result vector for an arbitrary SLP-aware protocol $\mathcal{P}^{'}$ for an input configuration $\mathcal{NC}^{'}$, using the above

attributes, could be:

$$r^{c',P'} = (10\%, 90\%, 2500, 12800) \tag{5.3}$$

However, these attributes do not satisfy the mutually preferentially independent property. This is apparent as, for example, the capture ratio attribute is dependent on the receive ratio attribute. For example, a low receive ratio will imply a low capture ratio because the attacker will have overheard only a few messages and would not have been able to track the asset down. In a similar way, messages sent attribute is related to the receive ratio in that, if a node does not receive a normal (data) message, then it is not going to forward it, reducing the number of messages sent.

To address this issue, some of these attributes are opted to transform to attempt to introduce the mutual preferential independence property. Since receive ratio is the one attribute that seems to affect both capture ratio and messages sent, these two attributes are normalised with respect to receive ratio, i.e., these attributes are penalised with respect to the receive ratio. On the other hand, message latency is independent of receive ratio. Thus, the attributes are redefined as follows:

$$R_{\omega}^{c,\mathcal{P}} = \begin{cases} r_{\omega}^{c,\mathcal{P}} / r_{rr}^{c,\mathcal{P}} & \text{if } \omega \in \{cr, mSent\}, \\ r_{\omega}^{c,\mathcal{P}} & \text{otherwise.} \end{cases} \tag{5.4}$$

3. Determining Aspiration Vector

After a consideration of different scenarios, the aspiration value is chosen for each attribute to remove any results not able to meet the scenario requirements. The aspiration value defines the minimal (or maximal in the case of cost criterion) acceptable value for each attribute. The aspiration vector is denoted by μ_a.

4. Filtering the Protocols

Based on the performances of the various protocols, it is obvious that those protocols that are worse (in all attributes) than all other protocols can be removed from the list as it implies that such protocols will never get selected when there is always another protocol that can deliver a better result. In this case, based on the selection of protocols none of them is dominated and thus none is removed from the list. For example, *ILP Routing-Max* has a very low capture ratio but a very high latency as its mechanism is based on using time redundancy to achieve privacy.

5. Checking Remaining Values

Each attribute needs to be checked whether any results are left in the scope of the aspiration value. If not, the procedure returns to phase 3 and reselect the aspiration value as all remaining results do not satisfy the aspiration value.

5.4.2 Step 2: Characterisation and Selection of SLP-Aware Routing Algorithms

1. Determining Weights and Utility Functions

In the first phase of Step 2, individual weight λ_ω ($\omega \in \{cr, rr, lat, mSent\}$) is chosen which represents the importance of attributes. From Step 1, the aspiration vector is used to generate the utility function for each attribute. The total utility obtained by a algorithm is shown below, where $U_\omega^{c,\mathcal{P}}$ is single attribute utility and $R_\omega^{c,\mathcal{P}}$ is a result value in $R^{c,\mathcal{P}}$.

$$U^{c,\mathcal{P}}(\lambda_\omega, R_\omega^{c,\mathcal{P}}) = \sum \lambda_\omega \cdot U_\omega^{c,\mathcal{P}}(R_\omega^{c,\mathcal{P}}) \tag{5.5}$$

2. Inserting Attribute Values

In the second phase of Step 2, each attribute uses the utility function and the remaining normalised result vector (i.e., performance vector) in the library as

input to calculate the utility value. Then the final utility of algorithm under $\mathcal{NC}$ can be calculated by Equation 5.5.

3. Selecting Utility Value

After proceeding through the steps described above, the performance library has obtained utility values for all algorithms and have chosen the best algorithm in terms of the highest utility value. In this case, under network configuration $\mathcal{NC}$ and the given scenario, the algorithm with the highest utility value has the best performance.

5.5 Problem Statement

The problem in this chapter is the following: In a WSN, there is a bunch of SLP-aware routing protocols can be used to deliver messages from the source to the sink. An attacker exists initially located at the sink and starts receiving messages sent by the source to the sink. Given a practical application, an important problem is to choose the best performing routing protocol to meet the requirements of the given application in terms of trade-offs. Formally, the problem specification is shown in Figure 5.1.

5.6 An Example: Execution of Decision Theoretic Procedure

This section provides a brief example of the execution of the decision theoretic procedures during the *profiling and filtering* phase (Subsection 5.4.1) and the *selection* phase (Subsection 5.4.2).

Figure 5.2 is an example containing the results of the capture ratio, receive ratio, message latency and messages sent for the various protocols under consideration, with network configuration $\mathcal{NC}$ = (grid, 121, SourceCorner, CasinoLab, LowAsymmetry, 1.2) which specifies that the network is a grid network of 121

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 set of SLP-aware routing protocols $\mathbb{P}$,
- A distributed eavesdropper $\mathcal{A}$ that is located at the sink initially,
- A given practical application $\mathcal{S}$,
- A set of network configurations $\mathbb{C}$ and
- A set of safety periods λ,

Objective:

- Choose the highest utility value of SLP-aware routing protocol from $\mathbb{P}$ over λ in G, $\mathbb{C}$, $\mathcal{S}$ and $\mathcal{A}$.

Figure 5.1: Problem statement: Selection of the best performing SLP-aware routing protocol under a practical application

nodes (i.e., 11×11), *SourceCorner* configuration (the source is at one corner of the grid), casino-lab noise model, low asymmetry communication model and 1.2×$\mathcal{TT}$ safety period respectively. From Table 5.3, the safety period is 11.9 seconds. For simplicity, the overall utility of phantom routing is evaluated and $PW(1, 1)$ is for comparison under the such network configuration.

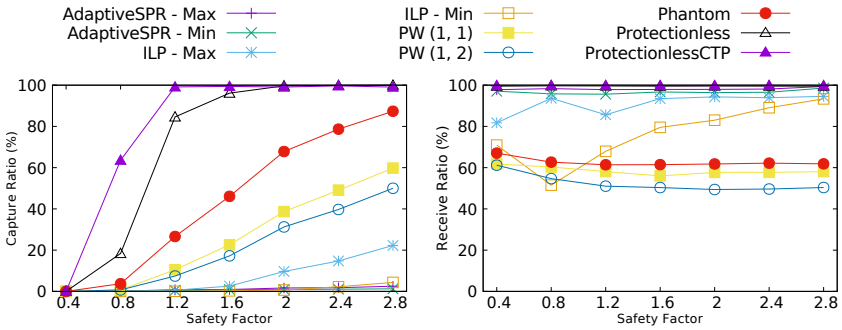
5.6.1 Step 1: Profiling and Filtering SLP-Aware Routing Algorithms

This section now explains Step 1 of the decision theoretic procedure.

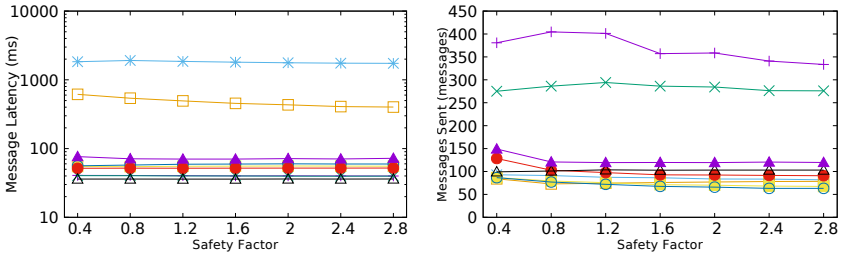
1. Profiling the Protocols

A number of routing protocols are executed, as explained in Section 5.2, some SLP-aware and others not. Figure 5.2 shows the graphs of capture ratio, receive ratio, message latency and messages sent.

Under configuration $\mathcal{NC}$ (as above), the result vector of phantom routing is (0.25, 0.65, 0.05, 100), with 25% capture ratio, 65% receive ratio, 50 ms for latency and 100 messages for messages sent. Similarly, the result vector of

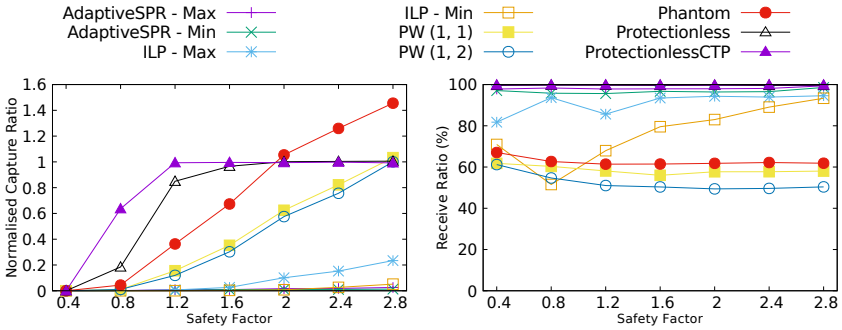


(a) Capture Ratio (b) Receive Ratio

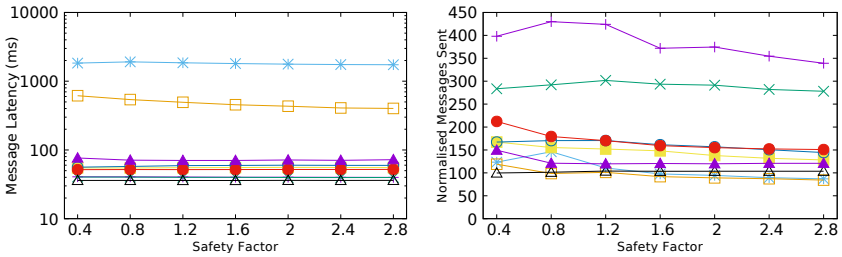


(c) Message Latency (d) Messages Sent

Figure 5.2: An Example: Protocols results of multiple attributes



(a) Normalised Capture Ratio (b) Receive Ratio



(c) Message Latency (d) Normalised Messages Sent

Figure 5.3: An Example: Multiple protocols results of normalised capture ratio and messages sent

$PW(1,1)$ is (0.1, 0.6, 0.06, 75). As discussed in section 5.4.1, results of attributes in the vector are not mutually preferentially independent, so they need to be normalised. The normalised results are shown in Figure 5.3. The normalised result vector of phantom routing is calculated, which is $(0.38, 0.65, 0.05, 153.85)$, while the normalised result vector of $PW(1,1)$ is $(0.17, 0.6, 0.06, 125)$. The normalised result vectors are used, or performance vectors, to form the protocol performance library, from which the most suitable SLP-aware routing protocol is to be selected. Since the objective here is to show the execution of the decision theoretic procedure, The example focuses on phantom routing and $PW(1,1)$.

2. Determining Relevant Attributes

The four attributes are: (i) capture ratio, (ii) receive ratio, (iii) message latency and (iv) messages sent. In general, for SLP, capture ratio and message or energy overhead are the two most important overheads. However, there are applications when other parameters such as receive ratio is important. In this chapter, for the applications, the above four attributes are considered relevant.

3. Determining Aspiration Vector

The aspiration value for each attribute needs to be determined. For instance, if 0.5 is given as an aspiration value for capture ratio and 60% for receive ratio, the utility of capture ratio and receive ratio will be set to 0 at these values. The system designer may consider that SLP cannot be provided if the capture ratio is greater than 0.5 and that the routing protocol cannot work properly if the receive ratio is lower than 60%. In this example, the aspiration vector is set to $(1.0, 1.0, 1.0, 1000)$. This means that the normalised capture ratio is set to 1.0, the receive ratio is set to 100%, latency is set to 1000 ms and the normalised messages sent are set to 1000.

4. Filtering the Protocols

Having the normalised results of both phantom routing and $PW(1,1)$, the results first need to be determined whether these two results are dominated by each other. If one protocol dominates the other, then there is a clear winner and the dominating protocol is selected as the best one. From the performance library, only the capture ratio and messages sent of $PW(1,1)$ performs better than phantom routing. Hence, $PW(1,1)$ does not dominate phantom routing and vice-versa. Therefore, in this case, both results are retained for further consideration.

5. Checking Remaining Values

Since neither $PW(1,1)$ nor phantom routing dominate each other, there is a need to determine whether they are dominated by the aspiration vector. The aspiration vector is $(1.0, 1.0, 0.1, 1000)$ while the performance profiles of phantom routing and $PW(1,1)$ are $(0.38, 0.65, 0.05, 153.85)$ and $(0.17, 0.6, 0.06, 125)$ respectively. As can be observed, the aspiration vector does not dominate either of the protocols, hence both protocols are still under consideration.

5.6.2 Step 2: Characterisation and Selection of SLP-Aware Routing Algorithms

Since there is more than a single protocol still in contention (i.e., there is no clear winner), step 2 is detailed in this section to select the better protocol.

1. Determining Weights and Utility Functions

Based on the aspiration vector, the sigmoid functions are used to build the utility functions for attributes[1]. Assuming in a scenario, all attributes are sensitive. For example, the utility keeps in a high value when capture ratio is below 0.2. However, the utility decrease quickly when capture ratio is larger than 0.2. In a

[1]Other utility functions could be also applied and they will be shown in Section 5.8.

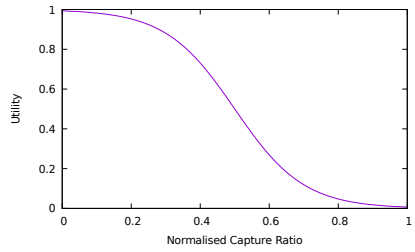
similar way, the message latency is intolerant when it is longer than 1 second. Therefore, the parameters are chosen and generate the utility functions for the four attributes, as shown in Equation 5.6 to Equation 5.9. These four sigmoid functions are shown in Figure 5.4 respectively. For simplicity, a weight vector is adopted with equal values $\lambda = (0.25, 0.25, 0.25, 0.25)$, meaning that all four attributes are equally as important.

$$U_{cr}^{\mathcal{NC},\mathcal{P}}(R_{cr}^{\mathcal{NC},\mathcal{P}}) = \frac{1}{1 + e^{10(R_{cr}^{\mathcal{NC},\mathcal{P}} - 0.5)}} \tag{5.6}$$

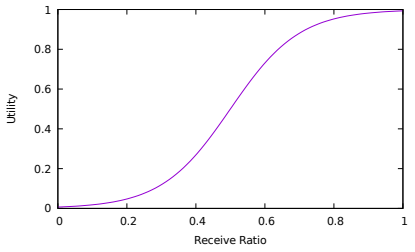
$$U_{rr}^{\mathcal{NC},\mathcal{P}}(R_{rr}^{\mathcal{NC},\mathcal{P}}) = \frac{1}{1 + e^{10(-R_{rr}^{\mathcal{NC},\mathcal{P}} + 0.5)}} \tag{5.7}$$

$$U_{lat}^{\mathcal{NC},\mathcal{P}}(R_{lat}^{\mathcal{NC},\mathcal{P}}) = \frac{1}{1 + e^{10(R_{lat}^{\mathcal{NC},\mathcal{P}} - 0.5)}} \tag{5.8}$$

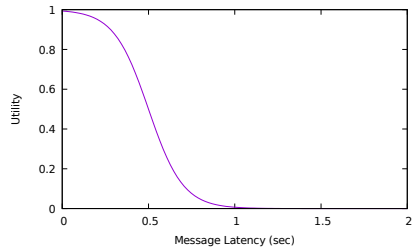
$$U_{mSent}^{\mathcal{NC},\mathcal{P}}(R_{mSent}^{\mathcal{NC},\mathcal{P}}) = \frac{1}{1 + e^{0.01(R_{mSent}^{\mathcal{NC},\mathcal{P}} - 500)}} \tag{5.9}$$



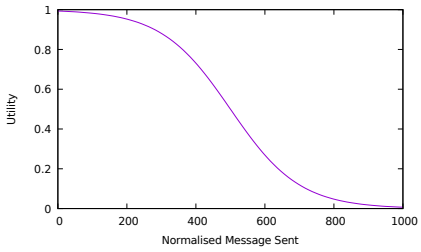
(a) The Utility of Normalised Capture Ratio



(b) The Utility of Receive Ratio



(c) The Utility of Message Latency



(d) The Utility of Normalised Messages Sent

Figure 5.4: Illustration of the utility functions in the example

2. Inserting Attribute Values

Using the utility functions previously identified (Equation 5.6 to Equation 5.9), the utility value of each attribute can be calculated in phantom routing: $U_{cr}^{\mathcal{NC},\mathcal{P}}(0.38) = 0.77$, $U_{rr}^{\mathcal{NC},\mathcal{P}}(0.65) = 0.82$, $U_{lat}^{\mathcal{NC},\mathcal{P}}(0.05) = 0.99$ and $U_{mSent}^{\mathcal{NC},\mathcal{P}}(153.85) = 0.97$. Finally, using the identified weight vector λ, the final utility of protocol $\mathcal{P}$ phantom routing under $\mathcal{NC}$ is:

$$\begin{aligned} U^{\mathcal{NC},\mathcal{P}}(\lambda_\omega, R_\omega^{\mathcal{NC},\mathcal{P}}) &= \sum \lambda_\omega \cdot U_\omega^{\mathcal{NC},\mathcal{P}}(R_\omega^{\mathcal{NC},\mathcal{P}}) \\ &= 0.25 \times 0.77 + 0.25 \times 0.82 + 0.25 \times 0.99 + 0.25 \times 0.97 \\ &= 0.89 \end{aligned} \tag{5.10}$$

Similarly, the utility value of $PW(1,1)$ also can be calculated: $U_{cr}^{\mathcal{NC},\mathcal{P}}(0.17) = 0.96$, $U_{rr}^{\mathcal{NC},\mathcal{P}}(0.6) = 0.73$, $U_{lat}^{\mathcal{NC},\mathcal{P}}(0.06) = 0.99$ and $U_{mSent}^{\mathcal{NC},\mathcal{P}}(125) = 0.98$. The final utility is 0.92.

3. Selecting Utility Value

Comparing the final utility of phantom routing and $PW(1,1)$, $PW(1,1)$ is selected as the better algorithm to provide SLP under a network configuration $\mathcal{NC}$ and with weight λ, as it is the one with the highest utility value.

5.7 Experimental Setup

A square grid network layout of size $n \times n$ was used in all experiments, with $n \in \{11, 21\}$, i.e., networks with 121 and 441 nodes respectively. The source period was set to 1 second. Node neighbourhoods were generated using *low-asymmetry*, which uses the LinkLayerModel tool provided with TOSSIM to generate link strengths between nodes using the parameters shown in Table 5.4. Links generated with low-asymmetry have a small probability of becoming asynchronous. Studies have shown that unreliable and asymmetric links can have a major impact on the performance of protocols [153].

The safety periods are calculated with different safety factors varying from 0.4, 0.8, 1.2 until 2.8 from Equation 3.1. The time taken for each network size and network configuration, for protectionless flooding is shown in table Table 5.2 and Table 5.3.

Network Size	casino-lab ideal	casino-lab low-asymmetry	meyer-heavy ideal	meyer-heavy low-asymmetry
11 × 11	12.355 ± 2.333	9.613 ± 2.231	12.479 ± 2.440	16.703 ± 6.578
21 × 21	25.059 ± 3.743	19.838 ± 3.597	25.617 ± 4.102	36.222 ± 10.994

Table 5.2: Time taken (seconds) of flooding for SinkCorner configuration with various models

Network Size	casino-lab ideal	casino-lab low-asymmetry	meyer-heavy ideal	meyer-heavy low-asymmetry
11 × 11	12.732 ± 2.736	9.928 ± 3.424	12.826 ± 2.660	22.890 ± 12.792
21 × 21	25.727 ± 4.111	20.405 ± 5.018	25.813 ± 4.169	46.836 ± 17.429

Table 5.3: Time taken (seconds) of flooding for SourceCorner configuration with various models

Phantom Routing and Phantom Walkabouts

These two algorithms both rely on the random walk technique. When choosing the length of the short and long random walks, a variety of parameter combinations were considered. The experiments set the short random walk series $S = \{2, 3, \ldots, 0.5 \times \Delta_{ss}\}$, and long random walk series $L = \{2 + \Delta_{ss}, 3 + \Delta_{ss}, \ldots, 1.5 \times \Delta_{ss}\}$, where Δ_{ss} is the sink-source distance. In phantom walkabouts, short and long random walk lengths are randomly generated from S and L respectively. For phantom routing, the random walk length is $0.5 \times \Delta_{ss}$ hops.

DynamicSPR

For this technique, as it aims to dynamically determine the parameters to use online, there are few parameters to specify. Other than the previously mentioned parameters, only the approach used needs to be specified. The approach determines how many fake messages are sent over the lifetime of

Name	Value
PATH_LOSS_EXPONENT	4.7
SHADOWING_STANDARD_DEVIATION	3.2
D0	1.0
PL_D0	55.4
NOISE_FLOOR	-105
S	[0.9 -0.7; -0.7 1.2]
WHITE_GAUSSIAN_NOISE	4

Table 5.4: LinkLayer model parameters for the low-asymmetry radio model

a temporary fake source. There are three options: `Fixed1`, `Fixed2` and `Rnd`. `Fixed1` sends a single fake message over the duration, `Fixed2` sends two fake messages over the duration and `Rnd` sends either 1 or 2 messages randomly chosen.

ILP Routing

This algorithm has four parameters: maximum walk length, buffer size, the number of messages to group and the probability that the message is sent directly to the sink. The same parameters are used in [15]. As the maximum walk length is simply to provide a finite bound in large networks, it was set to 100 hops. The number of messages to group was varied between {1, 2, 3, 4}. The buffer size was set to 10 messages as no more than 10 concurrent messages are expected being sent in the network at one time. Finally, the probability of sending a message directly to the sink was set to 20% as it was identified as a good setting in [15].

5.8 Case Studies: Routing Protocol Selection for Different Application Scenarios

This section will develop three case studies to showcase both the applicability of, and the generality allowed by, the methodology. The three case studies are: (i) an animal protection scenario, (ii) a non-critical asset monitoring scenario and (iii) a security-critical military scenario.

The first phase reuses the library that has already been built, consisting of a number of protocols that have been profiled. The library is denoted by $\mathbb{L}$ and the protocols used are listed in Table 5.5. Next, the set of decision attributes contain (i) capture ratio, (ii) receive ratio, (iii) message latency and (iv) messages sent.

In the next step, rather than deciding on the input network configuration, the section will eschew this step so as to keep the discussion as general as possible. This section also has the following aspiration performance vector:

$$\mu_a = (\min\{R_{cr}^{c,\mathcal{P}} \mid \mathcal{P} \in \mathbb{L}\}, \max\{R_{rr}^{c,\mathcal{P}} \mid \mathcal{P} \in \mathbb{L}\}, \min\{R_{lat}^{c,\mathcal{P}} \mid \mathcal{P} \in \mathbb{L}\}, \min\{R_{mSent}^{c,\mathcal{P}} \mid \mathcal{P} \in \mathbb{L}\}) \tag{5.11}$$

Thus, this means that all protocols are in contention and will be under consideration, i.e., there is no protocol filtering at this time.

The first phase of Step 2 is to decide the importance of each of the attributes (from capture ratio, receive ratio, message latency and messages sent) and to create a vector of weights quantifying the preference of each metric. Utilising the data and methods presented in Subsection 5.4.1, it is possible to calculate the utility of each protocol-parameter combination and generate plots to show which combination provides the highest utility for the given scenario. Specifically, using input network configuration $\mathcal{NC}$, it is possible to then select the most appropriate protocol.

For attributes, both non-linear[2] and linear functions[3] are used to model the utility of each parameter, as shown in Table 5.6. The reason why these functions is: The choices would like to show that the framework works with the combination of linear and non-linear functions. For those important attributes, non-linear functions are used to satisfy the quick change rate of the utility while linear functions are used for a smooth change rate. Parameters for the different attribute utility functions in different scenarios are shown in Table 5.7.

For all case studies, the network is assumed to be a grid with SourceCorner

[2]The sigmoid function $f(x) = \frac{1}{1+e^{k(-x-x_0)}}$ is used as the utility function for receive ratio, and $g(x) = \frac{1}{1+e^{k(x-x_0)}}$ for the rest attributes.

[3]The utility function is $f(x) = kx + x_0$

and SinkCorner configurations. Note that these assumptions need not be the case and there is no constraint imposed by the approach that precludes certain types of networks.

Algorithm Name	Technique	SLP-aware?
Protectionless	Flooding	No
Protectionless CTP	Collection Tree Protocol	No
DynamicSPR	Fake Messages	Yes
ILP Routing	Directed Walk	Yes
Phantom Routing	Directed Random Walk	Yes
Phantom Walkabouts	Directed Random Walk	Yes

Table 5.5: Protocols library ($\mathbb{L}$)

Attribute	Function Model Types		
	Animal Protection	Asset Monitor	Military
Normalised Capture Ratio	Non-Linear	Linear	Non-Linear
Receive Ratio (%)	Linear	Non-Linear	Non-Linear
Message Latency (seconds)	Linear	Linear	Non-Linear
Normalised Messages Sent	Linear	Non-Linear	Linear

Table 5.6: Model types of attribute utility funtions

Attribute	Scenario								
	Animal Protection			Asset Monitoring			Military		
	k	x_0	weight	k	x_0	weight	k	x_0	weight
Normalised Capture Ratio	50.0	0.1	0.4	-1.0	1.0	0.2	50.0	0.1	0.4
Receive Ratio (%)	1.0	0.0	0.2	20.0	-0.8	0.4	20.0	-0.8	0.25
Message Latency (seconds)	-0.5	1.0	0.2	-0.5	1.0	0.1	10.0	0.5	0.25
Normalised Messages Sent	-0.0005	1.0	0.2	0.01	400	0.3	-0.0005	1.0	0.1

Table 5.7: Parameters for attribute utility functions in different scenarios

5.8.1 Animal Protection Scenario

In this scenario, to prevent the rare animal from being captured by a poacher, SLP is crucial. Badger protection [41] and the WWF's Wildlife Crime Technology Report [1] are real world examples of animal protection. Therefore, capture ratio is the most important attribute animal protection. To maximise network lifetime, the message overhead needs to be reduced while receive ratio needs to be high as well to better understand the animal's behaviour. Thus, the weighting vector

is set to be $\lambda = (0.4, 0.2, 0.2, 0.2)$. These weights and respective utility values (see Tables 5.6 and 5.7 for each attribute) are used to produce overall utility plots, allowing us to deliberate which protocol and parameter combination would be most suited for the application. Figure 5.5 shows the overall utility values generated using the weights and utility functions previously specified[4].

Protocol Selection

An input network configuration is required from the utility values. For example, if the animal is expected to trigger a node that is at the top-left corner of a grid network of 11×11 that has been deployed and that the animal is expected to be constantly on the move (i.e., spending only a short time at a given location), and the environment is expected to be lightly noisy (similar to the casino-lab noise model) and the links are expected to be often unidirectional, then the input configuration can be settled with LowAsymmetry communication model. For instance, the configuration may be as follows: $\mathcal{NC}$ = (grid, 121, SourceCorner, CasinoLab, LowAsymmetry, 1.2). Thus, the protocol performance is discussed under the SourceCorner and SinkCorner configuration.

- **SourceCorner Configuration:** The configuration $\mathcal{NC}$ = (grid, 121, SourceCorner, CasinoLab, LowAsymmetry, 1.2) will correspond to Figure 5.5c and the best protocol is AdaptiveSPR because AdaptiveSPR-Min and AdaptiveSPR-max provide near-comparable performance. On the other hand, for example, if the input configuration is $\mathcal{NC}$ = (grid, 441, SourceCorner, CasinoLab, Ideal, 1.2), then the protocols that achieve the best trade-offs are $PW(1, 1)$ and $PW(1, 2)$ (see Figure 5.5b).

- **SinkCorner Configuration:** Differing from results in SourceCorner configuration, the best protocol is ILP Routing with configuration $\mathcal{NC}$ = (grid, 121, SourceCorner, CasinoLab, LowAsymmetry, 1.2) (see Figure 5.6c). Similarly, ILP Routing also dominates other routing protocols with the

[4]For simplicity, each configuration was described under corresponding graph with such format: number of nodes/noise model/communication model.

input configuration $\mathcal{NC}$ = (grid, 441, SourceCorner, CasinoLab, Ideal, 1.2)(see Figure 5.6b).

5.8.2 Asset Monitoring Scenario

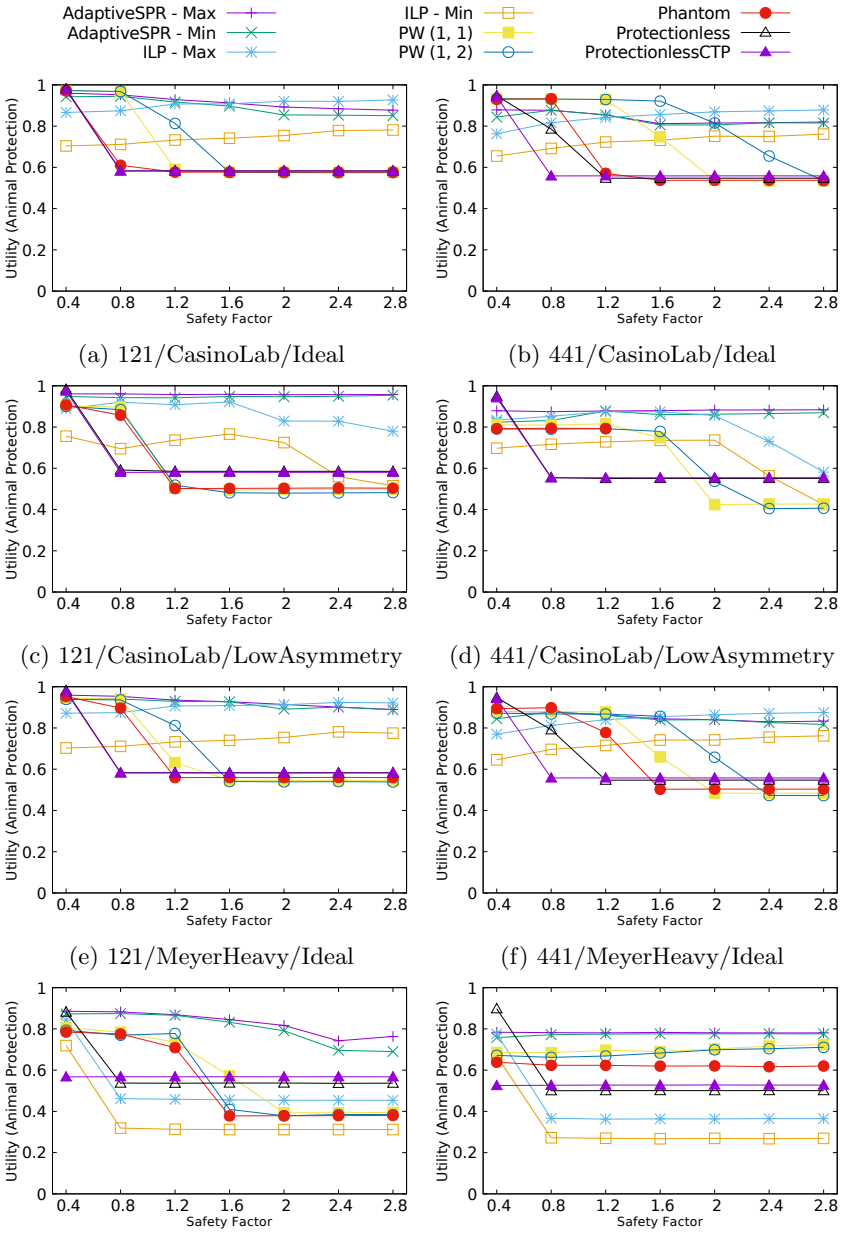
Sensors are often deployed in the body of bridges or in a building to monitor product quality [5]. They can also be deployed to monitor and understand animal behaviour (e.g., Great Duck Island [101]), differently from animal protection, as explained in the previous section. For this type of application, it could be assumed that receive ratio is the most important factor. This would leave capture ratio, latency and messages sent to be less important. Assume the weighting vector $\lambda = (0.2, 0.4, 0.1, 0.3)$ respectively representing capture ratio, receive ratio, latency and messages sent.

Protocol Selection

From the utility values, the input network configuration is required. For example, if the animal is expected to trigger a node that is at the top-left corner of a grid network of 11×11 that has been deployed. Since the animal is not expected to be very mobile and the environment can be expected to be noisy (i.e., similar to the meyer-heavy noise model) with unidirectional links due to a lack of line-of-sight transmission, then the input configuration can be as follows: $\mathcal{NC}$ = (grid, 121, SourceCorner, MeyerHeavy, LowAsymmetry, 2).

On the other hand, if the environment is not very noisy, i.e., similar to the casino-lab noise model, the network configuration would be $\mathcal{NC}$ = (grid, 121, SourceCorner, CasinoLab, LowAsymmetry, 2).

- **SourceCorner Configuration:** When $\mathcal{NC}$ = (grid, 121, SourceCorner, MeyerHeavy, LowAsymmetry, 2), this configuration will correspond to the Figure 5.7g and the best protocol is Protectionless CTP. The reasons why a non-SLP routing protocol is selected are (i) the Protectionless CTP has the highest utility and (ii) the SLP is not the most important attribute



(a) 121/CasinoLab/Ideal

(b) 441/CasinoLab/Ideal

(c) 121/CasinoLab/LowAsymmetry

(d) 441/CasinoLab/LowAsymmetry

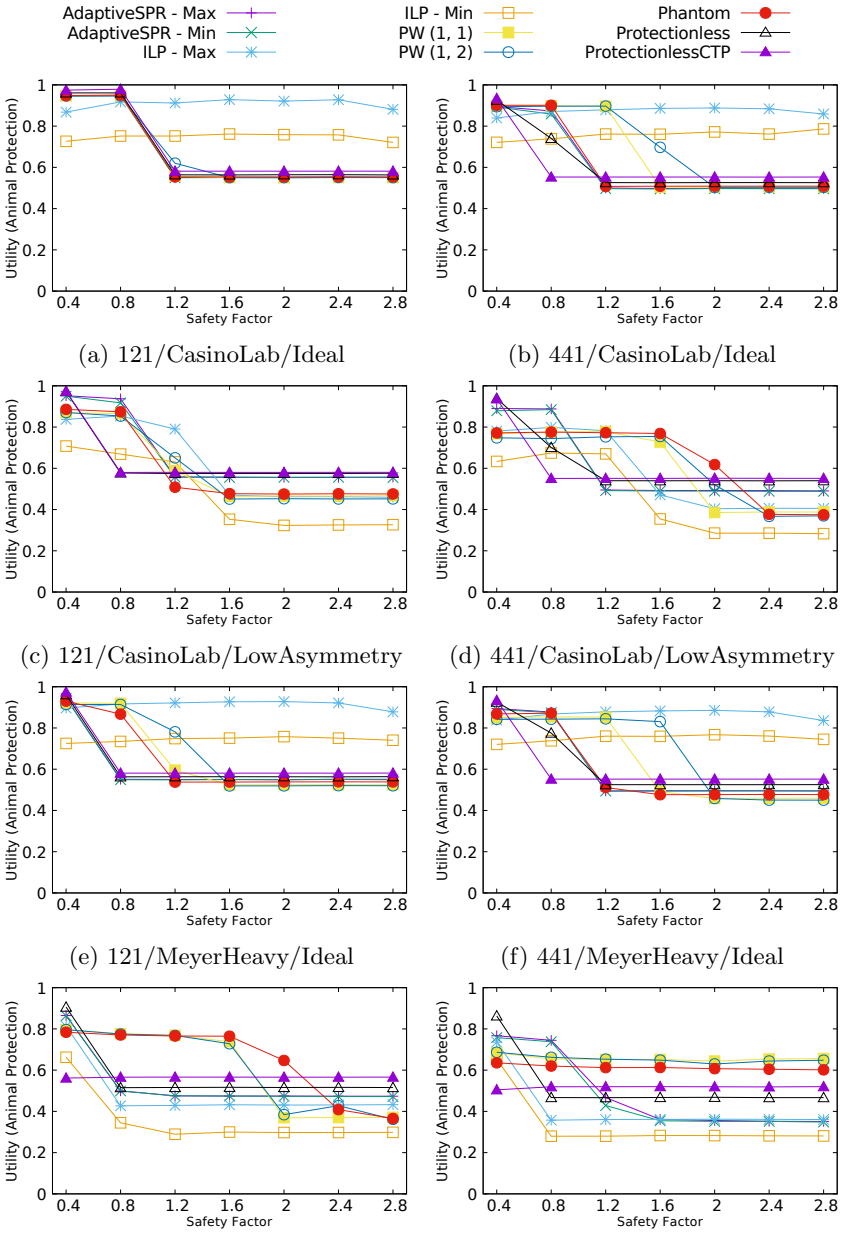
(e) 121/MeyerHeavy/Ideal

(f) 441/MeyerHeavy/Ideal

(g) 121/MeyerHeavy/LowAsymmetry

(h) 441/MeyerHeavy/LowAsymmetry

Figure 5.5: Utility of animal protection scenario in SourceCorner configuration



(a) 121/CasinoLab/Ideal

(b) 441/CasinoLab/Ideal

(c) 121/CasinoLab/LowAsymmetry

(d) 441/CasinoLab/LowAsymmetry

(e) 121/MeyerHeavy/Ideal

(f) 441/MeyerHeavy/Ideal

(g) 121/MeyerHeavy/LowAsymmetry

(h) 441/MeyerHeavy/LowAsymmetry

Figure 5.6: Utility of animal protection scenario in SinkCorner configuration

in this scenario. If the environment is not noisy, with $\mathcal{NC}$ = (grid, 121, SourceCorner, CasinoLab, LowAsymmetry, 2), ILP Routing is the protocol that achieves the best trade-off (Figure 5.7c).

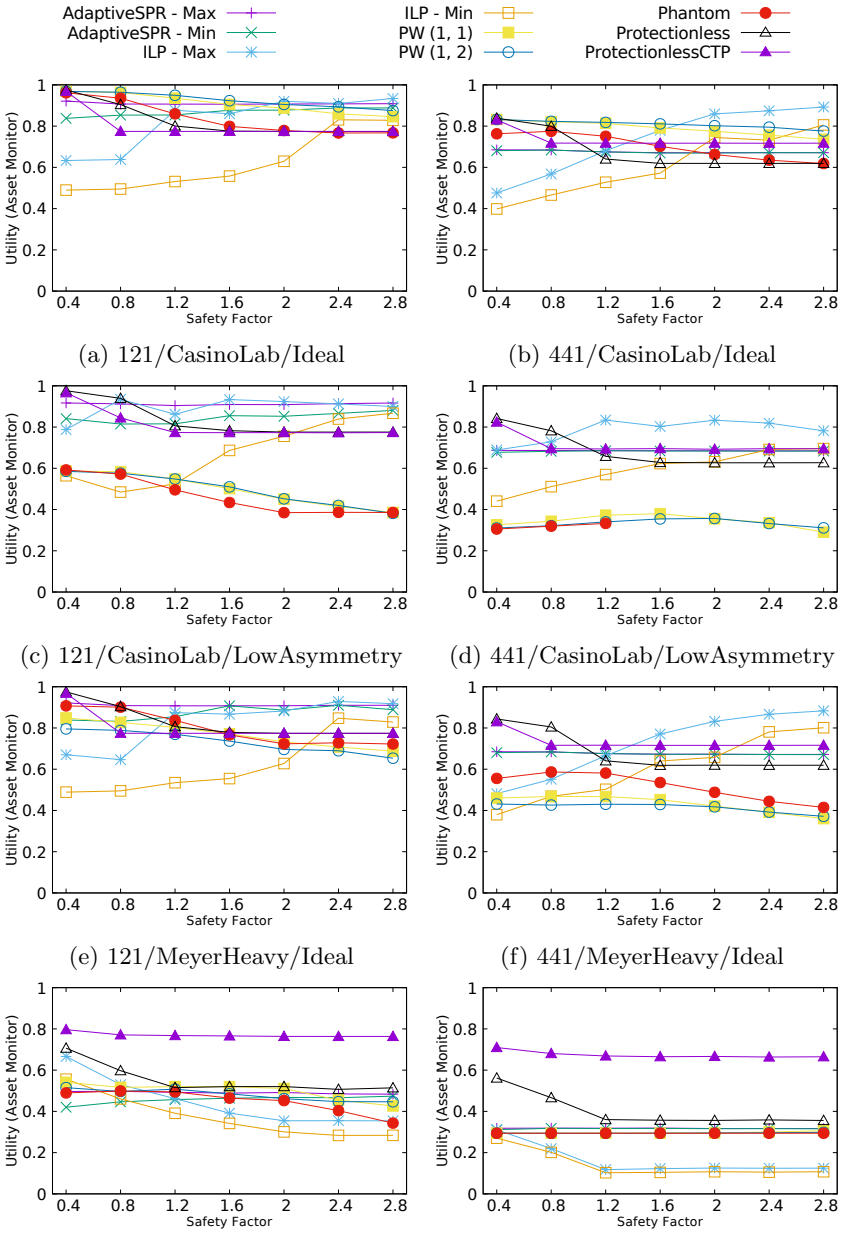
- **SinkCorner Configuration:** When $\mathcal{NC}$ = (grid, 121, SinkCorner, MeyerHeavy, LowAsymmetry, 2), as the same as the results in SourceCorner configuration, protectionless CTP is the best protocol (Figure 5.8g). AdaptiveSPR has the best performance under $\mathcal{NC}$ = (grid, 121, SinkCorner, CasinoLab, LowAsymmetry, 2)(Figure 5.8c).

5.8.3 Military Scenario

The use of sensor networks in military situations includes communication, battlefield surveillance and battle damage assessment among many others [5]. When these activities are carried out by military personnel, then SLP is extremely important. Furthermore, these networks may be short-lived, thus message overhead is not very important. On the other hand, latency and receive ratio are important, though less than SLP, to ensure soldiers can communicate in near real-time. Thus, the weight vector is $\lambda = (0.4, 0.25, 0.25, 0.1)$.

Protocol Selection

From the utility values, the input network configuration is required. For example, the surveillance activity to be carried out is expected to trigger a node that is at the top-left corner of a grid network of 11×11 that has been deployed. Since the personnel are expected to be very mobile and the environment can be expected to be noisy (i.e., similar to the MeyerHeavy noise model) with unidirectional links due to a lack of line-of-sight transmission, then the input configuration can be as follows: $\mathcal{NC}$ = (grid, 121, SourceCorner, MeyerHeavy, LowAsymmetry, 1.2). On the other hand, if the environment is not very noisy, i.e., similar to the casino-lab noise model, the network configuration would be $\mathcal{NC}$ = (grid,441, SourceCorner, CasinoLab, LowAsymmetry, 1.2).



(a) 121/CasinoLab/Ideal

(b) 441/CasinoLab/Ideal

(c) 121/CasinoLab/LowAsymmetry

(d) 441/CasinoLab/LowAsymmetry

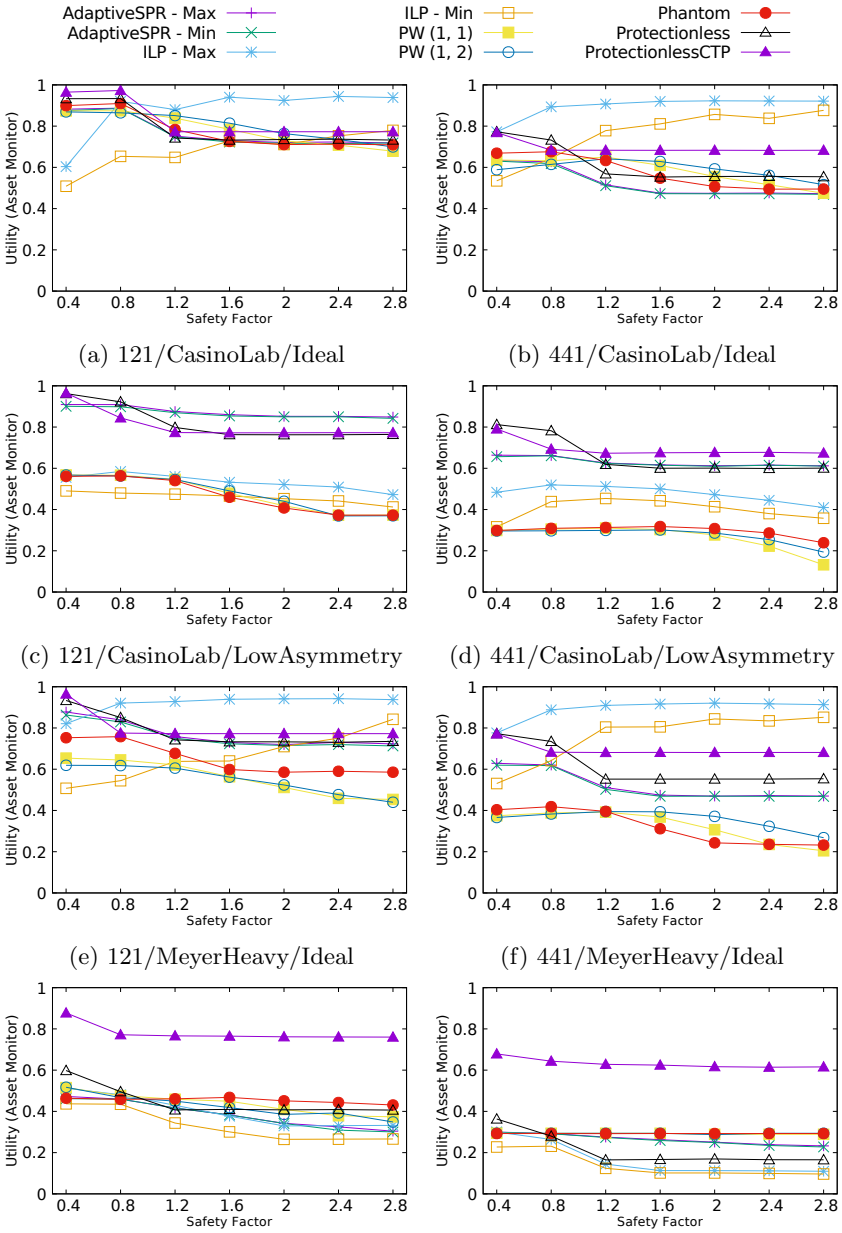
(e) 121/MeyerHeavy/Ideal

(f) 441/MeyerHeavy/Ideal

(g) 121/MeyerHeavy/LowAsymmetry

(h) 441/MeyerHeavy/LowAsymmetry

Figure 5.7: Utility of asset monitoring scenario in SourceCorner configuration



(a) 121/CasinoLab/Ideal

(b) 441/CasinoLab/Ideal

(c) 121/CasinoLab/LowAsymmetry

(d) 441/CasinoLab/LowAsymmetry

(e) 121/MeyerHeavy/Ideal

(f) 441/MeyerHeavy/Ideal

(g) 121/MeyerHeavy/LowAsymmetry

(h) 441/MeyerHeavy/LowAsymmetry

Figure 5.8: Utility of asset monitoring scenario in SinkCorner configuration

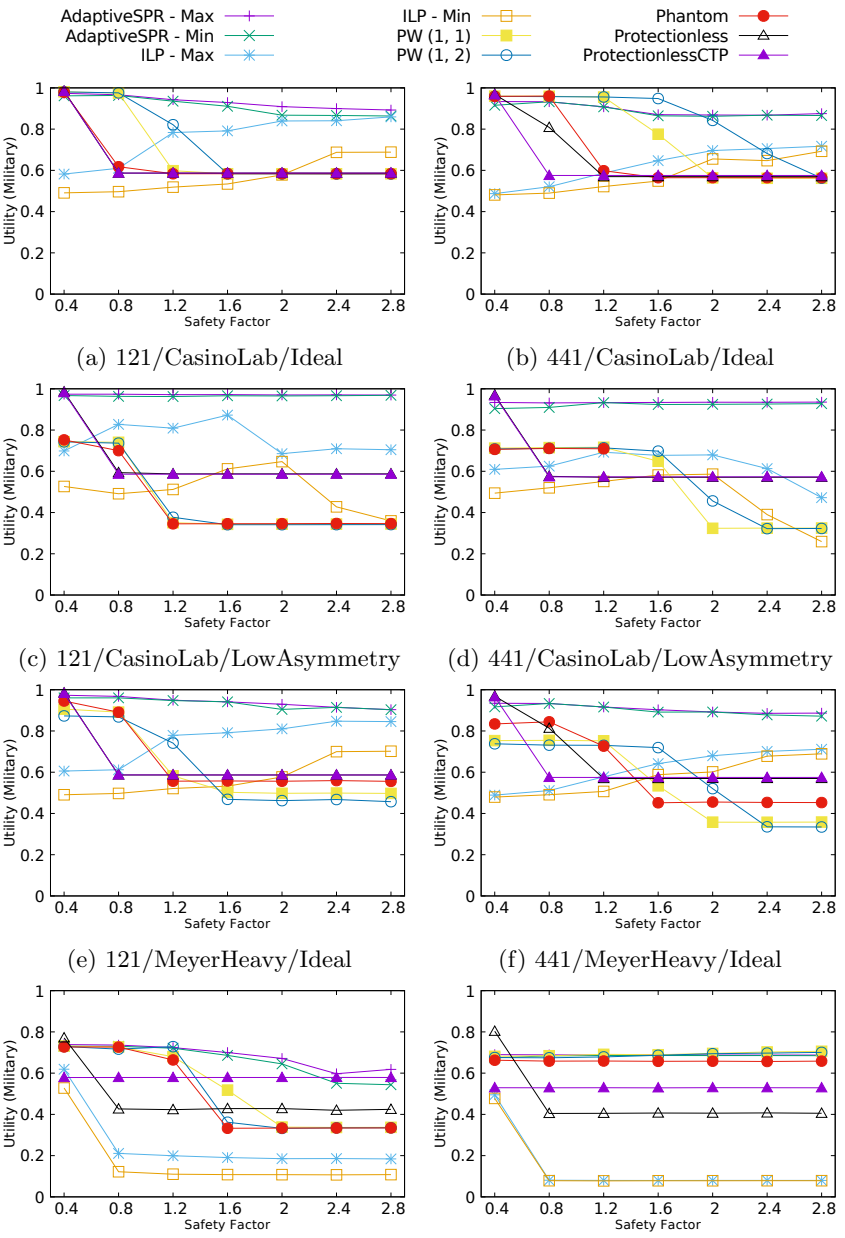
- **SourceCorner Configuration:** When $\mathcal{NC}$ = (grid, 121, SourceCorner, MeyerHeavy, LowAsymmetry, 1.2), this configuration will correspond to Figure 5.9g and the best protocol is AdaptiveSPR. AdaptiveSPR is also the protocol that achieves the best trade-off in Figure 5.9d with a network configuration $\mathcal{NC}$ = (grid,441, SourceCorner, CasinoLab, LowAsymmetry, 1.2). In the ideal environment, apart from ILP Routing, other protocols are good choices. However, when the environment gets worse (e.g., links are not reliable between two nodes), Protectionless and Protectionless CTP generally become the protocols of choice.

- **SinkCorner Configuration:** There are several protocols that achieve same best performance when $\mathcal{NC}$ = (grid, 121, SourceCorner, MeyerHeavy, LowAsymmetry, 1.2): phantom routing, PW(1,1) and PW(1,2)(see Figure 5.10g). These two protocols also have the best performance with $\mathcal{NC}$ = (grid,441, SourceCorner, CasinoLab, LowAsymmetry, 1.2)(Figure 5.10d).

5.9 Summary

Source location privacy (SLP) is becoming an important property for a large class of security-critical wireless sensor network applications such as monitoring and tracking. Many routing protocols have been proposed that provide SLP, all of which provide a trade-off between SLP and energy. Experiments have been conducted to gauge the performance of the proposed protocols under different network parameters such as noise levels. As there exists a plethora of protocols which contain a set of possibly conflicting performance attributes, it is difficult to select the SLP protocol that will provide the best trade-offs across them for a given application with specific requirements.

This chapter investigates the performance of SLP-aware routing protocols in wireless sensor networks (WSNs). The chapter especially focuses on the selection between SLP-aware routing protocols that achieves the best trade-offs among a set of attributes in the different scenarios. The methodology is based on the



(a) 121/CasinoLab/Ideal

(b) 441/CasinoLab/Ideal

(c) 121/CasinoLab/LowAsymmetry

(d) 441/CasinoLab/LowAsymmetry

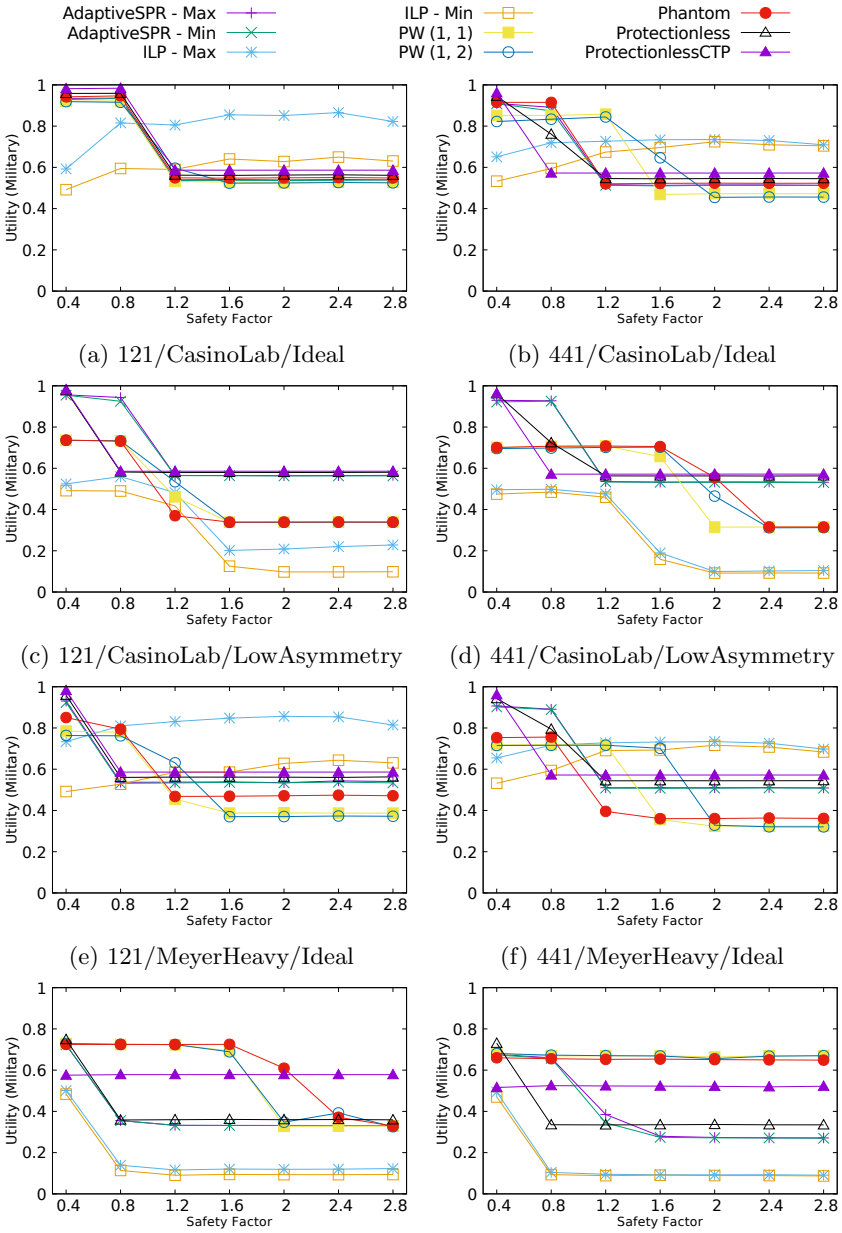
(e) 121/MeyerHeavy/Ideal

(f) 441/MeyerHeavy/Ideal

(g) 121/MeyerHeavy/LowAsymmetry

(h) 441/MeyerHeavy/LowAsymmetry

Figure 5.9: Utility of military scenario in SourceCorner configuration



(a) 121/CasinoLab/Ideal

(b) 441/CasinoLab/Ideal

(c) 121/CasinoLab/LowAsymmetry

(d) 441/CasinoLab/LowAsymmetry

(e) 121/MeyerHeavy/Ideal

(f) 441/MeyerHeavy/Ideal

(g) 121/MeyerHeavy/LowAsymmetry

(h) 441/MeyerHeavy/LowAsymmetry

Figure 5.10: Utility of military scenario in SinkCorner configuration

existence of a library of performance profiles of various routing algorithms and the decision theoretic procedure allows trade-offs to be assessed. This can be achieved when the attributes are mutually preferentially independent. Utility functions, weights of the attributes and network configurations are inputs that have to be provided by network administrators.

The main contribution is developing a decision theoretical framework that enables the the selection of the best performing SLP-aware protocols for the task. The framework has two steps: protocols profiling and protocol selection. The protocol selection step is based on a decision theoretic procedure that first removes dominated protocols and then formalises the notion of relevance using suitable utility functions. Furthermore, three different case studies are proposed to showcase the viability of the approach. The framework is not only applicable to the protocols in this chapter, but can be extended to handle any other SLP-aware routing protocols.

In future work, there are a few additions need to be done. First, the research will focus on generating profiles for other network configuration where both the source and sink are located in the opposite corners (called *FurtherSinkCorner*) in the network. Second, other protocols will be considered, such as OLSR [66], for protectionless routing to provide a baseline profile. And finally, further research plans on investigating the suitability of preference learning [47] for selecting appropriate values for aspiration levels.

CHAPTER 6

Conclusions and Further Work

The work described in this thesis has been concerned with SLP-aware routing protocols under practical scenarios in wireless sensor networks (WSNs). As wireless sensor networks have been applied across a spectrum of application domains such as asset monitoring, there is a need to consider security and privacy issues. One such issue is that of source location privacy (SLP) where the location of the source in the network needs to be kept secret from a malicious adversary. Several techniques have been proposed to provide SLP against an eavesdropping attacker, whereas results presented in support of these protocols have not included considerations for practical scenarios, omitting simulations and analyses with various network configurations. This thesis investigates SLP-aware routing protocols under practical scenarios in terms of routing protocols, parameterisations and trade-offs in WSNs.

The thesis first evaluates phantom routing by conducting an in-depth investigation of multiple attributes under various network configurations. The results demonstrate that previous work in phantom routing does not generalise well to different network configurations such as multiple sources. In order to address some weaknesses of phantom routing, phantom walkabouts is proposed, a novel and more general version of phantom routing, which performs phantom routes of variable lengths. Results from extensive simulations indicate a high level of SLP is achievable as a trade-off for a decrease in other attributes (e.g., the receive ratio). Finally the thesis proposes a decision theoretic procedure used to select SLP-aware routing protocols that achieve the best trade-offs among a set of attributes, which addresses the existing difficulty of selecting a SLP protocol that will provide the best trade-offs for a given application with specific

requirements.

The key contributions of this thesis are summarised in the first three sections of this chapter. Further works are then presented in Section 6.4.

6.1 Assessing the Performance of Phantom Routing on Source Location Privacy under Practical Scenarios

As WSNs have been applied across a spectrum of application domains, 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 address this problem. However, results presented in support of phantom routing have not included considerations for practical network configurations, and have omitted simulations and analyses with multiple sources.

Chapter 3 investigates the performance of phantom routing, a well-known algorithm that provides SLP in the WSN, under various network scenarios. Four application parameters are considered: (i) length of the random walk, (ii) source period, (iii) network size and (iv) number of sources. Meanwhile, various attributes related to the simulations are comprehensively discussed. The results show that (i) an increase in the length of the random walk leads to a corresponding increase in SLP level, (ii) a decrease of source period causes a decrease in the SLP level provided (i.e., the capture ratio increases), (iii) the SLP level does not change in response to changing network sizes, (iv) the increasing number of sources also causes an increase of capture ratio and (v) the receive ratio affects the performance of SLP level. The conclusion is that phantom routing is not as effective as initially claimed, as it was previously evaluated under a restricted set of circumstances and network configurations, e.g., one single source.

6.2 Developing Phantom Walkabouts to Achieve High Level of SLP

In seminal work on SLP, phantom routing was proposed as a viable approach to address SLP. However, the work has shown the limitations of phantom routing such as poor performance with multiple sources. Therefore, there is a need to develop a new routing protocol to perform better SLP than phantom routing can achieve.

In Chapter 4, the research starts with review of phantom routing, and then propose a novel technique called phantom walkabouts, which extends phantom routing to provide a better SLP level. Phantom walkabouts interleaves sequences of short random walks and long random walks to make the attacker move in the wrong direction, as opposed to phantom routing (with small random walks) where the attacker moves towards the source, hence improving the SLP level. A large-scale experiments are conducted to evaluate the efficiency of phantom walkabouts. The results have shown that phantom walkabouts provides much better level of SLP with multiple sources at certain parametrisation at the expense of energy and/or data yield. For some applications receive ratio and latency of this magnitude will be acceptable. For example, in a panda protection scenario the receive ratio will not be adversely affected.

6.3 A Decision Theoretic Framework for Selecting SLP-Aware Routing Protocols

Many routing protocols have been proposed that provide SLP, all of which provide a trade-off between SLP and energy. Experiments have been conducted to gauge the performance of the proposed protocols under different network parameters such as noise levels. As there exists a plethora of protocols which contain a set of possibly conflicting performance attributes, it is difficult to select the SLP protocol that will provide the best trade-offs across them for a given

application with specific requirements.

A decision theoretic procedure is proposed in Chapter 5 for selecting the SLP-aware routing algorithm that achieves the best trade-offs among a set of attributes. A methodology is proposed where routing protocols are first profiled to capture their performance under various protocol configurations. Then, a novel decision theoretic procedure is presented for selecting the most appropriate SLP routing algorithm for the application and network under investigation. The methodology is based on the existence of a library of performance profiles of various routing algorithms. The decision theoretic procedure allows trade-offs to be assessed. This can be achieved when attributes are mutually preferentially independent. Utility functions, weights of the attributes and network configurations are inputs that have to be provided by network administrators. The results have presented three case studies to showcase the viability of the approach.

6.4 Directions for Further Work

Following on from the work presented in this thesis, some more general directions are outlined for the work in the future.

Generic network configurations in wireless sensor networks

This thesis adopts a grid network configuration where nodes are evenly deployed in the network, but irregularly shaped networks should be considered. The network should be modelled under more realistic conditions where nodes are randomly deployed in the network, so that there are not always four neighbours for each node. In addition, results from simulations using TOSSIM do not work on real sensor devices, which limits practical analysis. Instead, it would be beneficial working on the SLP problem through testbeds that are deployed with real sensor network devices in the network. For instance, MoteLab [142] addresses this need allowing users to access and schedule experiments on real hardware via a web interface.

Developing dynamic protocols to address SLP

The phantom walkabouts technique presented in Chapter 4 uses fixed values to solve the SLP issue. For instance, the experiments used a given value for short random walks and a different value for long random walks (i.e., m and n are fixed in $PW(m, n)$). However, this need not be the case. The main weakness of phantom walkabouts is that it is not generically applicable to any network configuration. One further plan is to investigate phantom walkabouts with variable short and long random walks where parameters are decided in the runtime rather than in the compile time of the firmware. The conjecture would be that better trade-offs can be achieved by varying the length of short and long random walks. Besides, another research plan is to investigate, as mentioned in Chapter 4, the influence of other network configurations (e.g., the non-grid network configuration).

Developing a dynamic decision theoretic framework for SLP-aware routing protocol selection

The one drawback of the framework introduced in Chapter 5 is that most of the parameters must be provided manually by the network administrator. Namely, they must supply the network configuration, definitions of aspiration levels for the criteria, and the criteria as well. In practice, it may be hard for an administrator to determine if, for example, 20 milliseconds of latency is acceptable while 19 milliseconds is not. It would be worth considering an application of preference learning methodology [47] to help the administrator in selecting the appropriate values of aspiration levels. Other protocols will be also considered, such as OLSR [66], for protectionless routing to provide a baseline profile.

Mobile nodes in the wireless sensor networks

So far, solutions of location privacy have considered the static network, i.e., nodes are stationary in the network. However, considering the mobile nodes such as the mobile source or mobile sink would be a challenging task from the

perspective of location privacy. Undoubtedly, the new system model, threat model and theory need to be redefined to handle the emerging issue. Phantom walkabouts also needs a redesign. For example, when a node senses the change of environment, it can randomly move towards the sink instead of relying on multiple-hop communication. Another possible improvement is the sink can move to the direction of the incoming message with the long random walk. Because in some network configuration (e.g., SourceCorner configuration) the last node in the long random walk routing path is far away from the source. Furthermore, introducing mobile nodes is consistent with the scenarios of the Internet of Things (IoT), where embedded devices with sensors are attached to everyday objects even people. In this case, solving the privacy in this situation may be regarded as solutions of location privacy in IoT.

Bibliography

[1] Wildlife Crime Technology Project, 2012. URL https://www.worldwildlife.org/projects/wildlife-crime-technology-project.

[2] Tinyos CC2420.h, 2014. URL https://github.com/tinyos/tinyos-main/blob/master/tos/chips/cc2420/sim/CC2420.h.

[3] A. Abbasi, A. Khonsari, and M. S. Talebi. Source location anonymity for sensor networks. In *2009 6th IEEE Consumer Communications and Networking Conference, CCNC 2009*, pages 1–5, 2009. doi: 10.1109/CCNC.2009.4784915. URL https://doi.org/10.1109/CCNC.2009.4784915.

[4] I. F. Akyildiz and M. C. Vuran. *Wireless Sensor Networks*. John Wiley and Sons Ltd, 2010. doi: 10.1002/9780470515181. URL https://doi.org/10.1002/9780470515181.

[5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Communications Magazine*, 40(8):102–105, 2002. doi: 10.1109/MCOM.2002.1024422. URL https://doi.org/10.1109/MCOM.2002.1024422.

[6] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38(4):393–422, 2002. ISSN 13891286. doi: 10.1016/S1389-1286(01)00302-4. URL https://doi.org/10.1016/S1389-1286(01)00302-4.

[7] J. Al-Karaki and A. Kamal. Routing Techniques in Wireless Sensor Networks: A Survey. *IEEE Wireless Communications*, 11(6):6–28, 2004. doi: 10.1109/MWC.2004.1368893. URL https://doi.org/10.1109/MWC.2004.1368893.

[8] L. Alazzawi and A. Elkateeb. Performance Evaluation of the WSN Routing Protocols Scalability. *Journal of Computer Systems, Networks, and Communications*, 2008:1–9, 2008. ISSN 1687-7381. doi: 10.1155/2008/481046. URL `https://doi.org/10.1155/2008/481046`.

[9] B. Alomair, A. Clark, J. Cuellar, and R. Poovendran. Statistical framework for source anonymity in sensor networks. In *IEEE Global Telecommunications Conference*, pages 1–6, 2010. ISBN 9781479968183. doi: 10.1109/TMC.2011.267. URL `https://doi.org/10.1109/TMC.2011.267`.

[10] T. Arampatzis, J. Lygeros, and S. Manesis. A Survey of Applications of Wireless Sensors and Wireless Sensor Networks. In *Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, 2005.*, pages 719–724, 2005. ISBN 0-7803-8937-9. doi: 10.1109/.2005.1467103. URL `https://doi.org/10.1109/.2005.1467103`.

[11] E. Baccelli, O. Hahm, M. Gunes, M. Wahlisch, and T. Schmidt. RIOT OS: Towards an OS for the Internet of Things. In *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 79–80, 2013. ISBN 978-1-4799-0056-5. doi: 10.1109/INFCOMW.2013.6970748. URL `https://doi.org/10.1109/INFCOMW.2013.6970748`.

[12] N. Baccour, A. Koubaa, C. Noda, H. Fotouhi, M. Alves, H. Youssef, M. A. Zuniga, C. A. Boano, K. Roemer, and D. Puccinelli. Radio Link Quality Estimation in Wireless Sensor Networks: A Survey. *ACM Transactions on Sensor Networks (TOSN)*, 8(4):34:1–34:33, 2012. doi: 10.1145/2240116.2240123. URL `https://doi.org/10.1145/2240116.2240123`.

[13] N. Baroutis and M. Younis. Using fake sinks and deceptive relays to boost base-station anonymity in Wireless Sensor Network. In *2015 IEEE 40th Conference on Local Computer Networks (LCN)*, pages 109–116, 2015. doi:

10.1109/LCN.2015.7366289. URL https://doi.org/10.1109/LCN.2015.7366289.

[14] S. Boyd, A. Ghosh, B. Prabbakar, and D. Shah. Gossip algorithms: design, analysis and applications. *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, 3:1653–1664, 2005. ISSN 0743-166X. doi: 10.1109/INFCOM.2005.1498447. URL http://doi.org/10.1109/INFCOM.2005.1498447.

[15] M. Bradbury and A. Jhumka. A Near-Optimal Source Location Privacy Scheme for Wireless Sensor Networks. In *2017 IEEE Trustcom/BigDataSE/ICESS*, pages 409–416, 2017. doi: 10.1109/Trustcom/BigDataSE/ICESS.2017.265. URL https://doi.org/10.1109/Trustcom/BigDataSE/ICESS.2017.265.

[16] M. Bradbury and A. Jhumka. Understanding Source Location Privacy Protocols in Sensor Networks via Perturbation of Time Series. In *IEEE Conference on Computer Communications*, pages 1–9, 2017. doi: 10.1109/INFOCOM.2017.8057122. URL https://doi.org/10.1109/INFOCOM.2017.8057122.

[17] M. Bradbury, M. Leeke, and A. Jhumka. A dynamic fake source algorithm for source location privacy in wireless sensor networks. In *14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, pages 531–538, 2015. ISBN 9781467379519. doi: 10.1109/Trustcom.2015.416. URL https://doi.org/10.1109/Trustcom.2015.416.

[18] M. Bradbury, A. Jhumka, and M. Leeke. Hybrid online protocols for source location privacy in wireless sensor networks. *Journal of Parallel and Distributed Computing*, 115:67–81, 2018. ISSN 07437315. doi: 10.1016/j.jpdc.2018.01.006. URL https://doi.org/10.1016/j.jpdc.2018.01.006.

[19] E. Callaway, P. Gorday, L. Hester, J. Gutierrez, M. Naeve, B. Heile, and V. Bahl. Home networking with IEEE 802.15.4: a developing standard for low-rate wireless personal area networks. *IEEE Communications Magazine*, 40(8):70–77, 2002. doi: 10.1109/MCOM.2002.1024418. URL `https://doi.org/10.1109/MCOM.2002.1024418`.

[20] B. Carbunar, Y. Yu, L. Shi, M. Pearce, and V. Vasudevan. Query privacy in wireless sensor networks. In *2007 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pages 203–212, 2007. doi: 10.1109/SAHCN.2007.4292832. URL `https://doi.org/10.1109/SAHCN.2007.4292832`.

[21] C. Castelluccia, A. C.-F. Chan, E. Mykletun, and G. Tsudik. Efficient and Provably Secure Aggregation of Encrypted Data in Wireless Sensor Networks. *ACM Trans. Sen. Netw.*, 5(3):20:1–20:36, 2009. doi: 10.1145/1525856.1525858. URL `https://doi.org/10.1145/1525856.1525858`.

[22] 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), 2012. doi: 10.1155/2012/648058. URL `https://doi.org/10.1155/2012/648058`.

[23] H. Chen and W. Lou. From nowhere to somewhere: Protecting end-to-end location privacy in wireless sensor networks. In *Conference Proceedings of the IEEE International Performance, Computing, and Communications Conference*, pages 1–8, 2010. ISBN 9781424493302. doi: 10.1109/PCCC.2010.5682341. URL `https://doi.org/10.1109/PCCC.2010.5682341`.

[24] H. Chen and W. Lou. On protecting end-to-end location privacy against local eavesdropper in wireless sensor networks. *Pervasive and Mobile Computing*, 16:36–50, 2015. ISSN 15741192. doi: 10.1016/j.pmcj.2014.01.006. URL `https://doi.org/10.1016/j.pmcj.2014.01.006`.

[25] J. Chen, X. Du, and B. Fang. An efficient anonymous communication

protocol for wireless sensor networks. *Wireless Communications and Mobile Computing*, 12(14):1302–1312, 2011. doi: 10.1002/wcm.1205. URL https://doi.org/10.1002/wcm.1205.

[26] X. Chen, K. Makki, K. Yen, and N. Pissinou. Sensor network security: a survey. *IEEE Communications Surveys and Tutorials*, 11(2):52–73, 2009. ISSN 1553-877X. doi: 10.1109/SURV.2009.090205. URL https://doi.org/10.1109/SURV.2009.090205.

[27] O. Chipara, C. Lu, T. C. Bailey, and G. C. Roman. Reliable clinical monitoring using wireless sensor networks. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, pages 155–168, 2010. ISBN 9781450303446. doi: 10.1145/1869983.1869999. URL https://doi.org/10.1145/1869983.1869999.

[28] C. Y. Chong and S. P. Kumar. Sensor networks: evolution, opportunities, and challenges. *Proceedings of the IEEE*, 91(8):1247–1256, 2003. doi: 10.1109/JPROC.2003.814918. URL https://doi.org/10.1109/JPROC.2003.814918.

[29] M. Conti, R. Di Pietro, and L. V. Mancini. ECCE: Enhanced cooperative channel establishment for secure pair-wise communication in wireless sensor networks. *Ad Hoc Networks*, 5(1):49–62, 2007. ISSN 15708705. doi: 10.1016/j.adhoc.2006.05.013. URL https://doi.org/10.1016/j.adhoc.2006.05.013.

[30] M. Conti, L. Zhang, S. Roy, R. Di Pietro, S. Jajodia, and L. V. Mancini. Privacy-preserving robust data aggregation in wirelesssensor networks. *Security and Communication Networks*, 2(2):195–213, 2009. doi: 10.1002/sec.95. URL https://doi.org/10.1002/sec.95.

[31] M. Conti, J. Willemsen, and B. Crispo. Providing Source Location Privacy in Wireless Sensor Networks: A Survey. *IEEE Communications Surveys*

and Tutorials, 15(3):1238–1280, 2013. doi: 10.1109/SURV.2013.011413.00118. URL https://doi.org/10.1109/SURV.2013.011413.00118.

[32] J. Deng, R. Han, and S. Mishra. Countermeasures Against Traffic Analysis Attacks in Wireless Sensor Networks. In *First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05)*, pages 113–126, 2005. doi: 10.1109/SECURECOMM.2005.16. URL https://doi.org/10.1109/SECURECOMM.2005.16.

[33] J. Deng, R. Han, and S. Mishra. Decorrelating wireless sensor network traffic to inhibit traffic analysis attacks. *Pervasive and Mobile Computing*, 2(2):159–186, 2006. doi: 10.1016/j.pmcj.2005.12.003. URL https://doi.org/10.1016/j.pmcj.2005.12.003.

[34] R. Di Pietro and A. Viejo. Location privacy and resilience in wireless sensor networks querying. *Computer Communications*, 34(3):515–523, 2011. ISSN 01403664. doi: 10.1016/j.comcom.2010.05.014. URL https://doi.org/10.1016/j.comcom.2010.05.014.

[35] R. Di Pietro, P. Michiardi, and R. Molva. Confidentiality and integrity for data aggregation in wsn using peer monitoring. *Security and Communication Networks*, 2(2):181–194, 2009. doi: 10.1002/sec.93. URL https://doi.org/10.1002/sec.93.

[36] M. Dong, K. Ota, and A. Liu. Preserving source-location privacy through redundant fog loop for wireless sensor networks. In *IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM)*, pages 1835–1842, 2015. doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.274. URL https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.274.

[37] R. Doomun, T. Hayajneh, P. Krishnamurthy, and D. Tipper. SECLOUD: Source and destination seclusion using clouds for wireless ad hoc networks.

In *Proceedings - IEEE Symposium on Computers and Communications*, pages 361–367, 2009. ISBN 9781424446711. doi: 10.1109/ISCC.2009.5202367. URL https://doi.org/10.1109/ISCC.2009.5202367.

[38] W. Du, J. Deng, Y. S. Han, S. Chen, and P. K. Varshney. A key management scheme for wireless sensor networks using deployment knowledge. In *IEEE INFOCOM 2004*, page 597, 2004. doi: 10.1109/INFCOM.2004.1354530. URL https://doi.org/10.1109/INFCOM.2004.1354530.

[39] A. Dunkels, B. Gronvall, and T. Voigt. Contiki - a Lightweight and Flexible Operating System for Tiny Networked Sensors. In *29th Annual IEEE International Conference on Local Computer Networks*, pages 455–462, 2004. doi: 10.1109/LCN.2004.38. URL https://doi.org/10.1109/LCN.2004.38.

[40] N. Dutta, A. Saxena, and S. Chellappan. Defending wireless sensor networks against adversarial localization. In *Proceedings - IEEE International Conference on Mobile Data Management*, pages 336–341, 2010. ISBN 9780769540481. doi: 10.1109/MDM.2010.75. URL https://doi.org/10.1109/MDM.2010.75.

[41] V. Dyo, K. Yousef, S. A. Ellwood, D. W. Macdonald, A. Markham, N. Trigoni, R. Wohlers, C. Mascolo, B. Pásztor, and S. Scellato. WILDSENSING: Design and deployment of a sustainable sensor network for wildlife monitoring. *ACM Transactions on Sensor Networks*, 8(4): 1–33, 2012. ISSN 15504859. doi: 10.1145/2240116.2240118. URL https://doi.org/10.1145/2240116.2240118.

[42] E. Ekici, S. Vural, J. McNair, and D. Al-Abri. Secure probabilistic location verification in randomly deployed wireless sensor networks. *Ad Hoc Networks*, 6(2):195–209, 2008. ISSN 15708705. doi: 10.1016/j.adhoc.2006.11.006. URL https://doi.org/10.1016/j.adhoc.2006.11.006.

[43] R. El-Badry, A. Sultan, and M. Youssef. HyberLoc: Providing physical

layer location privacy in hybrid sensor networks. In *IEEE International Conference on Communications*, pages 1–5, 2010. ISBN 9781424464043. doi: 10.1109/ICC.2010.5502104. URL https://doi.org/10.1109/ICC.2010.5502104.

[44] L. Eschenauer and V. D. Gligor. A Key-management Scheme for Distributed Sensor Networks. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pages 41–47, 2002. doi: 10.1145/586110.586117. URL https://doi.org/10.1145/586110.586117.

[45] Y. Fan, Y. Jiang, H. Zhu, and X. Shen. An efficient privacy-preserving scheme against traffic analysis attacks in network coding. In *Proceedings - IEEE INFOCOM*, pages 2213–2221, 2009. ISBN 9781424435135. doi: 10.1109/INFCOM.2009.5062146. URL https://doi.org/10.1109/INFCOM.2009.5062146.

[46] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi. In-Network Aggregation Techniques for Wireless Sensor Networks: a Survey. *IEEE Wireless Communication*, 14(2):70–87, 2007. doi: 10.1109/MWC.2007.358967. URL https://doi.org/10.1109/MWC.2007.358967.

[47] J. Frnkranz and E. Hllermeier. *Preference Learning*. Springer-Verlag New York, Inc., 1st edition, 2010. doi: 10.1007/978-3-642-14125-6. URL https://doi.org/10.1007/978-3-642-14125-6.

[48] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-Delay Trade-off in Wireless Networks. In *IEEE INFOCOM 2004*, pages 464–475, 2004. doi: 10.1109/INFCOM.2004.1354518. URL https://doi.org/10.1109/INFCOM.2004.1354518.

[49] D. Gay, P. Levis, R. von Behren, M. Welsh, E. Brewer, and D. Culler. The nesC Language: A Holistic Approach to Networked Embedded Systems. In *Proceedings of the ACM SIGPLAN 2003 Conference on Programming*

Language Design and Implementation, pages 1–11, 2003. doi: 10.1145/781131.781133. URL https://doi.org/10.1145/781131.781133.

[50] O. Gnawali, R. Fonseca, K. Jamieson, M. Kazandjieva, D. Moss, and P. Levis. CTP: An efficient, robust, and reliable collection tree protocol for wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 10(1):16:1–16:49, 2013. doi: 10.1145/2529988. URL https://doi.org/10.1145/2529988.

[51] S. Greco, M. Ehrgott, and J. R. Figueira. *Multiple Criteria Decision Analysis: State of the Art Surveys*. Springer New York, 2016. doi: 10.1007/978-1-4939-3094-4. URL https://doi.org/10.1007/978-1-4939-3094-4.

[52] C. Gu. SLP Algorithms (TinyOS), 2014. URL https://bitbucket.org/Chen_Gu/slp-algorithms-tinyos.

[53] C. Gu. Simulation results for PhD thesis, 2018. URL https://doi.org/10.5281/zenodo.1323732.

[54] C. Gu, M. Bradbury, A. Jhumka, and M. Leeke. Assessing the Performance of Phantom Routing on Source Location Privacy in Wireless Sensor Networks. In *IEEE 21st Pacific Rim International Symposium on Dependable Computing*, pages 99–108, 2015. ISBN 9781467393768. doi: 10.1109/PRDC.2015.9. URL https://doi.org/10.1109/PRDC.2015.9.

[55] C. Gu, M. Bradbury, and A. Jhumka. Phantom walkabouts in wireless sensor networks. In *Proceedings of the ACM Symposium on Applied Computing*, pages 609–616, 2017. doi: 10.1145/3019612.3019732. URL https://doi.org/10.1145/3019612.3019732.

[56] C. Gu, M. Bradbury, J. Kirton, and A. Jhumka. A decision theoretic framework for selecting source location privacy aware routing protocols in wireless sensor networks. *Future Generation Computer Systems*, 2018.

ISSN 0167739X. doi: 10.1016/j.future.2018.01.046. URL https://doi.org/10.1016/j.future.2018.01.046.

[57] V. C. Gungor and G. P. Hancke. Industrial Wireless Sensor Networks: Challenges, Design Principles, and Technical Approaches. *IEEE Transactions on Industrial Electronics*, 56(10):4258–4265, 2009. doi: 10.1109/TIE.2009.2015754. URL https://doi.org/10.1109/TIE.2009.2015754.

[58] J. Gutierrez, M. Naeve, E. Callaway, M. Bourgeois, V. Mitter, and B. Heile. IEEE 802.15.4: a developing standard for low-power low-cost wireless personal area networks. *IEEE Network*, 15(5):12–19, 2001. doi: 10.1109/65.953229. URL https://doi.org/10.1109/65.953229.

[59] W. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. Abdelzaher. PDA: Privacy-Preserving Data Aggregation in Wireless Sensor Networks. In *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, pages 2045–2053, 2007. doi: 10.1109/INFCOM.2007.237. URL https://doi.org/10.1109/INFCOM.2007.237.

[60] W. He, H. Nguyen, X. Liuy, K. Nahrstedt, and T. Abdelzaher. iPDA: An integrity-protecting private data aggregation scheme for wireless sensor networks. In *MILCOM 2008 - 2008 IEEE Military Communications Conference*, pages 1–7, 2008. doi: 10.1109/MILCOM.2008.4753645. URL https://doi.org/10.1109/MILCOM.2008.4753645.

[61] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishna. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4):660–670, 2002. doi: 10.1109/TWC.2002.804190. URL https://doi.org/10.1109/TWC.2002.804190.

[62] W. R. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive protocols for information dissemination in wireless sensor networks. In *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing*

and networking - MobiCom '99, pages 174–185, 1999. doi: 10.1145/313451.313529. URL https://doi.org/10.1145/313451.313529.

[63] A. Hessler, T. Kakumaru, H. Perrey, and D. Westhoff. Data obfuscation with network coding. *Computer Communications*, 35(1):48–61, 2012. doi: 10.1016/j.comcom.2010.11.004. URL https://doi.org/10.1016/j.comcom.2010.11.004.

[64] X. Hong, P. Wang, J. Kong, Q. Zheng, and J. Liu. Effective probabilistic approach protecting sensor traffics. In *IEEE Military Communications Conference (MILCOM)*, volume 1, pages 169–175, 2005. ISBN 0780393937. doi: 10.1109/MILCOM.2005.1605681. URL https://doi.org/10.1109/MILCOM.2005.1605681.

[65] D. Huang. Traffic analysis-based unlinkability measure for IEEE 802.11b-based communication systems. In *Proceedings of the 5th ACM workshop on Wireless security (WiSe '06)*, pages 65–74, 2006. doi: 10.1145/1161289.1161303. URL https://doi.org/10.1145/1161289.1161303.

[66] P. Jacquet, P. Mühlethaler, T. H. Clausen, A. Laouiti, A. Qayyum, and L. Viennot. Optimized link state routing protocol for ad hoc networks. *Proceedings of the IEEE International Multi Topic Conferences*, pages 62–68, 2001. doi: 10.1109/INMIC.2001.995315. URL https://doi.org/10.1109/INMIC.2001.995315.

[67] A. Jhumka. Crash-Tolerant Collision-Free Data Aggregation Scheduling for Wireless Sensor Networks. In *Reliable Distributed Systems, 2010 29th IEEE Symposium on*, pages 44–53, 2010. doi: 10.1109/SRDS.2010.14. URL https://doi.org/10.1109/SRDS.2010.14.

[68] A. Jhumka and M. Bradbury. Deconstructing source location privacy-aware routing protocols. In *Proceedings of the Symposium on Applied Computing*, pages 431–436, 2017. doi: 10.1145/3019612.3019655. URL https://doi.org/10.1145/3019612.3019655.

[69] A. Jhumka, M. Leeke, and S. Shrestha. On the use of fake sources for source location privacy: Trade-Offs between energy and privacy. *Computer Journal*, 54(6):860–874, 2011. doi: 10.1093/comjnl/bxr010. URL https://doi.org/10.1093/comjnl/bxr010.

[70] A. Jhumka, M. Bradbury, and M. Leeke. Towards understanding source location privacy in wireless sensor networks through fake sources. In *IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, pages 760–768, 2012. ISBN 9780769547459. doi: 10.1109/TrustCom.2012.281. URL https://doi.org/10.1109/TrustCom.2012.281.

[71] A. Jhumka, M. Bradbury, and M. Leeke. Fake source-based source location privacy in wireless sensor networks. *Concurrency Computation Practice and Experience*, 27(12):2999–3020, 2015. ISSN 15320626. doi: 10.1002/cpe.3242. URL https://doi.org/10.1002/cpe.3242.

[72] Y. Jian, S. Chen, Z. Zhang, and L. Zhang. Protecting receiver-location privacy in wireless sensor networks. In *26th IEEE International Conference on Computer Communications*, pages 1955–1963, 2007. doi: 10.1109/INFCOM.2007.227. URL https://doi.org/10.1109/INFCOM.2007.227.

[73] J. R. Jiang, J. P. Sheu, C. Tu, and J. W. Wu. An Anonymous Path Routing (APR) Protocol for Wireless Sensor Networks. *Journal of Information Science and Engineering*, 27(2):657–680, 2011.

[74] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk. Enhancing Source-Location Privacy in Sensor Network Routing. In *25th IEEE International Conference on Distributed Computing Systems (ICDCS)*, pages 599–608, 2005. ISBN 0-7695-2331-5. doi: 10.1109/ICDCS.2005.31. URL https://doi.org/10.1109/ICDCS.2005.31.

[75] P. Kamat, W. Xu, W. Trappe, and Y. Zhang. Temporal Privacy in Wireless Sensor Networks: Theory and Practice. *ACM Trans. Sen. Netw.*, 5(4):

28:1–28:24, 2009. ISSN 1550-4859. doi: 10.1145/1614379.1614380. URL https://doi.org/10.1145/1614379.1614380.

[76] L. Kang. Protecting Location Privacy in Large-Scale wireless sensor networks. In *IEEE International Conference on Communications*, pages 1–6, 2009. ISBN 9781424434350. doi: 10.1109/ICC.2009.5199372. URL https://doi.org/10.1109/ICC.2009.5199372.

[77] T. Kavitha and D. Sridharan. Security Vulnerabilities In Wireless Sensor Networks : A Survey. *Journal of Information Assurance and Security*, 5 (1):31–44, 2010.

[78] L. Kazatzopoulos, C. Delakouridis, G. Marias, and P. Georgiadis. iHIDE : Hiding Sources of Information in WSNs. In *Second International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing*, pages 8 pp.–48, 2006. ISBN 0769525490. doi: 10.1109/SECPERU.2006.11. URL https://doi.org/10.1109/SECPERU.2006.11.

[79] R. L. Keeney and H. Raiffa. *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge University Press, 1993.

[80] J. Kirton, M. Bradbury, and A. Jhumka. Source Location Privacy-Aware Data Aggregation Scheduling for Wireless Sensor Networks. In *International Conference on Distributed Computing Systems (ICDCS)*, pages 2200–2205, 2017. doi: 10.1109/ICDCS.2017.171. URL https://doi.org/10.1109/ICDCS.2017.171.

[81] K. H. Kwong, T. T. Wu, H. G. Goh, K. Sasloglou, B. Stephen, I. Glover, C. Shen, W. Du, C. Michie, and I. Andonovic. Practical considerations for wireless sensor networks in cattle monitoring applications. *Computers and Electronics in Agriculture*, 81:33–44, 2012. doi: 10.1016/j.compag.2011.10.013. URL https://doi.org/10.1016/j.compag.2011.10.013.

[82] J. Laikin, M. Bradbury, C. Gu, and M. Leeke. Towards fake sources for source location privacy in wireless sensor networks with multiple sources.

In *IEEE International Conference on Communication Systems*, pages 1–6, 2016. doi: 10.1109/ICCS.2016.7833572. URL https://doi.org/10.1109/ICCS.2016.7833572.

[83] 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*, pages 126–137, 2003. ISBN 1-58113-707-9. doi: 10.1145/958491.958506. URL https://doi.org/10.1145/958491.958506.

[84] P. Levis, S. Madden, J. Polastre, R. Szewczyk, K. Whitehouse, A. Woo, D. Gay, J. Hill, M. Welsh, E. Brewer, D. Culler, W. Weber, J. Rabaey, and E. Aarts. TinyOS: An Operating System for Wireless Sensor Networks. In *Ambient Intelligence*, pages 115–148. Springer, Berlin, Heidelberg, 2005. doi: 10.1007/3-540-27139-27. URL https://doi.org/10.1007/3-540-27139-2{_}7.

[85] N. Li, N. Zhang, S. K. Das, and B. Thuraisingham. Privacy preservation in wireless sensor networks: A state-of-the-art survey. *Ad Hoc Networks*, 7(8):1501–1514, 2009. ISSN 15708705. doi: 10.1016/j.adhoc.2009.04.009. URL https://doi.org/10.1016/j.adhoc.2009.04.009.

[86] N. Li, M. Raj, D. Liu, M. Wright, and S. K. Das. Using data mules to preserve source location privacy in Wireless Sensor Networks. In *International Conference on Distributed Computing Computing and Networking*, pages 309–324, 2012. doi: 10.1016/j.pmcj.2012.10.002. URL https://doi.org/10.1016/j.pmcj.2012.10.002.

[87] Y. Li and J. Ren. Mixing Ring-Based Source-Location Privacy in Wireless Sensor Networks. In *2009 Proceedings of 18th International Conference on Computer Communications and Networks*, pages 1–6, 2009. doi: 10.1109/ICCCN.2009.5235216. URL https://doi.org/10.1109/ICCCN.2009.5235216.

[88] Y. Li, L. Lightfoot, and R. Jian. Routing-based source-location privacy in wireless sensor networks. In *2009 IEEE International Conference on Electro/Information Technology*, pages 29–34, 2009. ISBN 9781424434350. doi: 10.1109/EIT.2009.5189579. URL https://doi.org/10.1109/EIT.2009.5189579.

[89] L. Lightfoot, Y. Li, and J. Ren. Preserving source-location privacy in wireless sensor network using STaR routing. In *IEEE Global Telecommunications Conference*, pages 1–5, 2010. ISBN 9781424456383. doi: 10.1109/GLOCOM.2010.5683603. URL https://doi.org/10.1109/GLOCOM.2010.5683603.

[90] L. Lightfoot, Y. Li, and J. Ren. STaR: design and quantitative measurement of source-location privacy for wireless sensor networks. *Security and Communication Networks*, 9(3):220–228, 2016. doi: 10.1002/sec.527. URL https://doi.org/10.1002/sec.527.

[91] R. Lim, F. Ferrari, and M. Zimmerling. FlockLab: A testbed for distributed, synchronized tracing and profiling of wireless embedded systems. In *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 153–165, 2013. doi: 10.1145/2461381.2461402. URL https://doi.org/10.1145/2461381.2461402.

[92] S. Lin, J. Liu, and Y. Fang. ZigBee Alliance. In *IEEE International Conference on Automation and Logistics*, pages 1979–1983, 2007. doi: 10.1109/ICAL.2007.4338898. URL https://doi.org/10.1109/ICAL.2007.4338898.

[93] G. Liu, R. Tan, R. Zhou, G. Xing, W.-Z. Song, and M. J. Lees. Volcanic earthquake timing using wireless sensor networks. In *ACM/IEEE International Conference on Information Processing in Sensor Networks*, pages 91–102, 2013. doi: 10.1145/2461381.2461396. URL https://doi.org/10.1145/2461381.2461396.

[94] 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. *IEEE Access*, 2:633–651, 2014. doi: 10.1109/ACCESS.2014.2332817. URL https://doi.org/10.1109/ACCESS.2014.2332817.

[95] R. Lu, X. Lin, H. Zhu, and X. S. Shen. TESP2 : Timed Efficient Source Privacy Preservation Scheme for Wireless Sensor Networks. In *IEEE International Conference on Communications*, pages 1–6, 2010. ISBN 9781424464043. doi: 10.1109/ICC.2010.5502142. URL https://doi.org/10.1109/ICC.2010.5502142.

[96] X. Luo, X. Ji, and M.-S. Park. Location Privacy against Traffic Analysis Attacks in Wireless Sensor Networks. In *2010 International Conference on Information Science and Applications*, pages 1–6, 2010. ISBN 978-1-4244-5942-1. doi: 10.1109/ICISA.2010.5480564. URL https://doi.org/10.1109/ICISA.2010.5480564.

[97] J. P. Lynch. A Summary Review of Wireless Sensors and Sensor Networks for Structural Health Monitoring. *The Shock and Vibration Digest*, 38(2): 91–128, 2006. ISSN 0583-1024. doi: 10.1016/j.comnet.2010.05.003. URL https://doi.org/10.1016/j.comnet.2010.05.003.

[98] M. Mafuta, M. Zennaro, A. Bagula, G. Ault, H. Gombachika, and T. Chadza. Successful Deployment of a Wireless Sensor Network for Precision Agriculture in MalawiWiPAM. In *IEEE International Conference on Networked Embedded Systems for Every Application*, pages 1–7, 2012. doi: 10.1109/NESEA.2012.6474009. URL https://doi.org/10.1109/NESEA.2012.6474009.

[99] H. Mahmoud and A. Fahmy. *Wireless Sensor Networks: Concepts, Applications, Experimentation and Analysis*, volume 14. Springer Singapore, 1st

edition, 2016. ISBN 978-981-10-0412-4. doi: 10.1007/978-981-10-0412-4. URL https://doi.org/10.1007/978-981-10-0412-4.

[100] M. Mahmoud and X. Shen. A cloud-based scheme for protecting source-location privacy against hotspot-locating attack in wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 23(10):1805–1818, 2012. ISSN 10459219. doi: 10.1109/TPDS.2011.302. URL https://doi.org/10.1109/TPDS.2011.302.

[101] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson. Wireless Sensor Networks for Habitat Monitoring. In *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pages 88–97, 2002. ISBN 1-58113-589-0. doi: 10.1145/570738.570751. URL https://doi.org/10.1145/570738.570751.

[102] A. Majeed, K. Liu, and N. Abu-Ghazaleh. TARP: Timing Analysis Resilient Protocol for Wireless Sensor Networks. In *2009 IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, pages 85–90, 2009. ISBN 978-0-7695-3841-9. doi: 10.1109/WiMob.2009.24. URL https://doi.org/10.1109/WiMob.2009.24.

[103] G. Mao, B. Fidan, and B. D. Anderson. Wireless sensor network localization techniques. *Computer Networks*, 51(10):2529–2553, 2007. ISSN 13891286. doi: 10.1016/j.comnet.2006.11.018. URL https://doi.org/10.1016/j.comnet.2006.11.018.

[104] K. Martinez and J. K. Hart. Glacier Monitoring: Deploying Custom Hardware in Harsh Environments. In *Wireless Sensor Networks: Deployments and Design Frameworks*, pages 245–258. 2010. ISBN 978-1-4419-5834-1. doi: 10.1007/978-1-4419-5834-1. URL https://doi.org/10.1007/978-1-4419-5834-1.

[105] K. Mehta, D. Liu, and M. Wright. Location Privacy in Sensor Networks Against a Global Eavesdropper. In *2007 IEEE International Conference on*

Network Protocols, pages 314–323, 2007. doi: 10.1109/ICNP.2007.4375862. URL https://doi.org/10.1109/ICNP.2007.4375862.

[106] L. Miao, K. Djouani, A. Kurien, and G. Noel. Network coding and competitive approach for gradient based routing in wireless sensor networks. *Ad Hoc Networks*, 10(6):990–1008, 2012. doi: 10.1016/j.adhoc.2012.01.001. URL https://doi.org/10.1016/j.adhoc.2012.01.001.

[107] G. Mulligan. The 6LoWPAN Architecture. In *Proceedings of the 4th Workshop on Embedded Networked Sensors*, pages 78–82, 2007. doi: 10.1145/1278972.1278992. URL https://doi.org/10.1145/1278972.1278992.

[108] A. A. Nezhad, A. Miri, and D. Makrakis. Location privacy and anonymity preserving routing for wireless sensor networks. *Computer Networks*, 52 (18):3433–3452, 2008. ISSN 13891286. doi: 10.1016/j.comnet.2008.09.005. URL https://doi.org/10.1016/j.comnet.2008.09.005.

[109] E. C. H. Ngai. On providing sink anonymity for sensor networks. In *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, pages 269–273, 2009. ISBN 978-1-60558-569-7. doi: 10.1145/1582379.1582439. URL https://doi.org/10.1145/1582379.1582439.

[110] E. C. H. Ngai and I. Rodhe. On providing location privacy for mobile sinks in wireless sensor networks. *Wireless Networks*, 19(1):115–130, 2013. ISSN 10220038. doi: 10.1007/s11276-012-0454-z. URL https://doi.org/10.1007/s11276-012-0454-z.

[111] S. Ortolani, M. Conti, B. Crispo, and R. Di Pietro. Event handoff unobservability in WSN. In *Proceedings of international conference on Open research problems in network security*, volume 1, pages 20–28, 2011. ISBN 9783642192272. doi: 10.1007/978-3-642-19228-9_3. URL https://doi.org/10.1007/978-3-642-19228-9_3.

[112] F. Osterlind, A. Dunkels, J. Eriksson, N. Finne, and T. Voigt. Cross-level sensor network simulation with COOJA. In *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, pages 641–648, 2006. doi: 10.1109/LCN.2006.322172. URL https://doi.org/10.1109/LCN.2006.322172.

[113] Y. Ouyang, Z. Le, G. Chen, J. Ford, and F. Makedon. Entrapping adversaries for source protection in sensor networks. In *Proceedings - WoWMoM 2006: 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks*, pages 23–34, 2006. ISBN 0769525938. doi: 10.1109/WOWMOM.2006.40. URL https://doi.org/10.1109/WOWMOM.2006.40.

[114] Y. Ouyang, Z. Le, D. Liu, J. Ford, and F. Makedon. Source Location Privacy Against Laptop-class Attacks in Sensor Networks. In *Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks*, pages 1–10, 2008. ISBN 978-1-60558-241-2. doi: 10.1145/1460877.1460884. URL https://doi.org/10.1145/1460877.1460884.

[115] C. Ozturk, Y. Zhang, and W. Trappe. Source-location privacy in energy-constrained sensor network routing. In *Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks - SASN '04*, pages 88–93, 2004. ISBN 1581139721. doi: 10.1145/1029102.1029117. URL https://doi.org/10.1145/1029102.1029117.

[116] C. Ozturk, Y. Zhang, W. Trappe, and M. Ott. Source-location privacy for networks of energy-constrained sensors. In *Proceedings - Second IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems*, pages 68–72, 2004. doi: 10.1109/WSTFES.2004.1300417. URL https://doi.org/10.1109/WSTFES.2004.1300417.

[117] S. Pai, S. Bermudez, S. B. Wicker, M. Meingast, T. Roosta, S. Sastry, and D. K. Mulligan. Transactional confidentiality in sensor networks. *IEEE*

Security and Privacy, 6(4):28–35, 2008. ISSN 15407993. doi: 10.1109/MSP.2008.107. URL https://doi.org/10.1109/MSP.2008.107.

[118] H. Park, S. Song, B. Y. Choi, and C. T. Huang. PASSAGES: Preserving Anonymity of Sources and Sinks against Global Eavesdroppers. In *2013 Proceedings IEEE INFOCOM*, pages 210–214, 2013. ISBN 9781467359467. doi: 10.1109/INFCOM.2013.6566765. URL https://doi.org/10.1109/INFCOM.2013.6566765.

[119] A. Perrig, R. Szewczyk, V. Wen, D. Culler, J. D. Tygar, A. Perrig, and R. Szewczyk. SPINS: Security Protocols for Sensor Networks. *Wireless Networks*, 8(5):521–534, 2002. doi: 10.1023/A:1016598314198. URL https://doi.org/10.1023/A:1016598314198.

[120] S. Peter, D. Westhoff, and C. Castelluccia. A survey on the encryption of convergecast traffic with in-network processing. *IEEE Transactions on Dependable and Secure Computing*, 7(1):20–34, 2010. doi: 10.1109/TDSC.2008.23. URL https://doi.org/10.1109/TDSC.2008.23.

[121] R. Rios and J. Lopez. Source Location Privacy Considerations in Wireless Sensor Networks. In *4th International Symposium of Ubiquitous Computing and Ambient Intelligence (UCAmI10)*, pages 29 – 38, 2010. ISBN 978-84-92812-61-5. URL https://www.nics.uma.es/sites/default/files/papers/Rios2010.pdf.

[122] R. Rios and J. Lopez. Exploiting context-awareness to enhance source-location privacy in wireless sensor networks. *Computer Journal*, 54(10): 1603–1615, 2011. ISSN 00104620. doi: 10.1093/comjnl/bxr055. URL https://doi.org/10.1093/comjnl/bxr055.

[123] R. Rios, J. Lopez, and J. Cuellar. *Location Privacy in Wireless Sensor Networks*. CRC Press, Inc., 1st edition, 2016. ISBN 9781498776332. URL https://www.

crcpress.com/Location-Privacy-in-Wireless-Sensor-Networks/Rios-Lopez-Cuellar/p/book/9781498776332.

[124] R. A. Shaikh, H. Jameel, B. J. D'Auriol, H. Lee, S. Lee, and Y. J. Song. Achieving network level privacy in wireless sensor networks. *Sensors*, 10 (3):1447–1472, 2010. ISSN 14248220. doi: 10.3390/s100301447. URL https://doi.org/10.3390/s100301447.

[125] M. Shao, S. Zhu, W. Zhang, G. Cao, and Y. Yang. PDCS: Security and privacy support for data-centric sensor networks. In *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, pages 1298–1306, 2007. doi: 10.1109/INFCOM.2007.154. URL https://doi.org/10.1109/INFCOM.2007.154.

[126] M. Shao, Y. Yang, S. Zhu, and G. Cao. Towards Statistically Strong Source Anonymity for Sensor Networks. In *The 27th Conference on Computer Communications. IEEE*, pages 51–55, 2008. ISBN 9781424420261. doi: 10.1109/INFOCOM.2008.19. URL https://doi.org/10.1109/INFOCOM.2008.19.

[127] M. Shao, W. Hu, S. Zhu, G. Cao, S. Krishnamurthy, and T. La Porta. Cross-layer enhanced source location privacy in sensor networks. In *6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON 2009)*, pages 1–9, 2009. ISBN 9781424429080. doi: 10.1109/SAHCN.2009.5168923. URL https://doi.org/10.1109/SAHCN.2009.5168923.

[128] B. Sheng and Q. Li. Verifiable privacy-preserving range query in two-tiered sensor networks. In *Proceedings - IEEE INFOCOM*, pages 457–465, 2008. ISBN 9781424420261. doi: 10.1109/INFOCOM.2007.18. URL https://doi.org/10.1109/INFOCOM.2007.18.

[129] J. H. Song, V. W. S. Wong, and V. C. M. Leung. Wireless Location Privacy Protection in Vehicular Ad-Hoc Networks. *Mobile Networks and*

Applications, 15(1):160–171, 2010. ISSN 1938-1883. doi: 10.1109/ICC.2009.5199575. URL https://doi.org/10.1109/ICC.2009.5199575.

[130] P. Spachos, L. Song, and D. Hatzinakos. Opportunistic routing for enhanced source-location privacy in wireless sensor networks. In *25th Biennial Symposium on Communications, QBSC 2010*, pages 315–318, 2010. ISBN 9781424457090. doi: 10.1109/BSC.2010.5472946. URL https://doi.org/10.1109/BSC.2010.5472946.

[131] P. Spachos, L. Song, F. M. Bui, and D. Hatzinakos. Improving source-location privacy through opportunistic routing in wireless sensor networks. In *2011 IEEE Symposium on Computers and Communications*, pages 815–820, 2011. ISBN 978-1-4577-0680-6. doi: 10.1109/ISCC.2011.5983942. URL https://doi.org/10.1109/ISCC.2011.5983942.

[132] G. Suarez-tangil, E. Palomar, B. Ramos, and A. Ribagorda. An Experimental Comparison of Source Location Privacy Methods for Power Optimization in WSNs. In *Proceedings of the 3rd WSEAS International Conference on Advances in Sensors, Signals and Materials*, pages 79–84, 2010. ISBN 9789604742486. URL http://dl.acm.org/citation.cfm?id=1950175.1950191.

[133] W. Tan, K. Xu, and D. Wang. An anti-tracking source-location privacy protection protocol in WSNs based on path extension. *IEEE Internet of Things Journal*, 1(5):461–471, 2014. ISSN 2327-4662. doi: 10.1109/JIOT.2014.2346813. URL https://doi.org/10.1109/JIOT.2014.2346813.

[134] A. Thomason, M. Leeke, M. Bradbury, and A. Jhumka. Evaluating the impact of broadcast rates and collisions on fake source protocols for source location privacy. In *EEE International Conference on Trust, Security and Privacy in Computing and Communications*, pages 667–674, 2013. ISBN 9780769550220. doi: 10.1109/TrustCom.2013.81. URL https://doi.org/10.1109/TrustCom.2013.81.

[135] P. Venkitasubramaniam and L. Tong. Anonymous networking with minimum latency in multihop networks. In *Proceedings - IEEE Symposium on Security and Privacy*, pages 18–32, 2008. ISBN 9780769531687. doi: 10.1109/SP.2008.18. URL https://doi.org/10.1109/SP.2008.18.

[136] A. Wadaa, S. Olariu, L. Wilson, M. Eltoweissy, and K. Jones. On Providing anonymity in wireless sensor networks. In *Proceedings of 10th International Conference on Parallel and Distributed Systems*, pages 411–418, 2004. ISBN 0-7695-2152-5. doi: 10.1109/ICPADS.2004.1316121. URL https://doi.org/10.1109/ICPADS.2004.1316121.

[137] C. Wang, G. Wang, W. Zhang, and T. Feng. Reconciling Privacy Preservation and Intrusion Detection in Sensory Data Aggregation. In *Proceedings IEEE INFOCOM*, pages 336–340, 2011. doi: 10.1109/INFCOM.2011.5935177. URL https://doi.org/10.1109/INFCOM.2011.5935177.

[138] H. Wang, B. Sheng, and Q. Li. Privacy-aware routing in sensor networks. *Computer Networks*, 53(9):1512–1529, 2009. ISSN 13891286. doi: 10.1016/j.comnet.2009.02.002. URL https://doi.org/10.1016/j.comnet.2009.02.002.

[139] H.-J. Wang and T.-R. Hsiang. Defending Traffic Analysis with Communication Cycles in Wireless Sensor Networks. In *2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks*, pages 166–171, 2009. doi: 10.1109/I-SPAN.2009.78. URL https://doi.org/10.1109/I-SPAN.2009.78.

[140] W. P. Wang, L. Chen, and J. X. Wang. A source-location privacy protocol in WSN based on locational angle. In *IEEE International Conference on Communications*, pages 1630–1634, 2008. ISBN 9781424420742. doi: 10.1109/ICC.2008.315. URL https://doi.org/10.1109/ICC.2008.315.

[141] X. Wang, X. Li, Z. Wan, and M. Gu. CLEAR : A Confidential and Lifetime-Aware Routing Protocol for Wireless Sensor Network. In *Personal*

Indoor and Mobile Radio Communications 2009 IEEE 20th International Symposium on, pages 2265–2269, 2009. doi: 10.1109/PIMRC.2009.5450077. URL https://doi.org/10.1109/PIMRC.2009.5450077.

[142] G. Werner-Allen, P. Swieskowski, and M. Welsh. MoteLab: a wireless sensor network testbed. In *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005.*, pages 483–488, 2005. doi: 10.1109/IPSN.2005.1440979. URL https://doi.org/10.1109/IPSN.2005.1440979.

[143] H. Will, K. Schleiser, and J. Schiller. A real-time kernel for wireless sensor networks employed in rescue scenarios. In *2009 IEEE 34th Conference on Local Computer Networks*, pages 834–841, 2009. doi: 10.1109/LCN.2009.5355049. URL https://doi.org/10.1109/LCN.2009.5355049.

[144] Y. Xi, L. Schwiebert, and W. Shi. Preserving source location privacy in monitoring-based wireless sensor networks. In *Proceedings 20th IEEE International Parallel and Distributed Processing Symposium*, page 8 pp., 2006. doi: 10.1109/IPDPS.2006.1639682. URL https://doi.org/10.1109/IPDPS.2006.1639682.

[145] D. Xiao, M. Wei, and Y. Zhou. Secure-SPIN: Secure Sensor Protocol for Information via Negotiation for Wireless Sensor Networks. In *2006 1ST IEEE Conference on Industrial Electronics and Applications*, pages 1–4, 2006. doi: 10.1109/ICIEA.2006.257149. URL https://doi.org/10.1109/ICIEA.2006.257149.

[146] Y. Xiao, V. K. Rayi, B. Sun, X. Du, F. Hu, and M. Galloway. A survey of key management schemes in wireless sensor networks. *Computer Communications*, 30(11):2314–2341, 2007. doi: 10.1016/j.comcom.2007.04.009. URL https://doi.org/10.1016/j.comcom.2007.04.009.

[147] W. Yang and W. T. Zhu. Protecting Source Location Privacy in Wireless Sensor Networks with Data Aggregation. In *7th international con-*

ference on Ubiquitous intelligence and computing, pages 252–266, 2010. doi: 10.1007/978-3-642-16355-5_22. URL https://doi.org/10.1007/978-3-642-16355-5_22.

[148] 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 - WiSec '08*, pages 77–88, 2008. ISBN 9781595938145. doi: 10.1145/1352533.1352547. URL https://doi.org/10.1145/1352533.1352547.

[149] Y. Yang, X. Wang, S. Zhu, and G. Cao. SDAP: A secure hop-by-hop data aggregation protocol for sensor networks. *ACM Transactions on Information and System Security*, 11(4):1–43, 2008. doi: 10.1145/1380564.1380568. URL https://doi.org/10.1145/1380564.1380568.

[150] Y. Yang, J. Zhou, R. H. Deng, and F. Bao. Better security enforcement in trusted computing enabled heterogeneous wireless sensor networks. *Security and Communication Networks*, 4(1):11–22, 2011. ISSN 09739769. doi: 10.1002/sec.179. URL https://doi.org/10.1002/sec.179.

[151] J. Yao and G. Wen. Preserving source-location privacy in energy-constrained wireless sensor networks. In *Proceedings - International Conference on Distributed Computing Systems*, volume 1, pages 412–416, 2008. ISBN 9780769531731. doi: 10.1109/ICDCS.Workshops.2008.42. URL https://doi.org/10.1109/ICDCS.Workshops.2008.42.

[152] W. Ye, J. Heidemann, and D. Estrin. An energy-efficient MAC protocol for wireless sensor networks. In *Proceedings of 21st Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 1567–1576, 2002. doi: 10.1109/INFCOM.2002.1019408. URL https://doi.org/10.1109/INFCOM.2002.1019408.

[153] M. Z. Zamalloa and B. Krishnamachari. An analysis of unreliability and asymmetry in low-power wireless links. *ACM Transactions on Sensor*

Networks, 3(2):1–34, 2007. doi: 10.1145/1240226.1240227. URL https://doi.org/10.1145/1240226.1240227.

[154] Y. Zeng, J. Cao, S. Member, S. Zhang, S. Guo, and L. Xie. Random-Walk Based Approach to Detect Clone Attacks in Wireless Sensor Networks. *IEEE Journal on Selected Areas in Communications*, 28(5):677–691, 2010. doi: 10.1109/JSAC.2010.100606. URL https://doi.org/10.1109/JSAC.2010.100606.

[155] J. Zhang and V. Varadharajan. Wireless sensor network key management survey and taxonomy. *Journal of Network and Computer Applications*, 23 (2):63–75, 2010. doi: 10.1016/j.jnca.2009.10.001. URL https://doi.org/10.1016/j.jnca.2009.10.001.

[156] L. Zhang. A Self-Adjusting Directed Random Walk Approach for Enhancing Source-Location Privacy in Sensor Network Routing. In *international conference on Wireless communications and mobile computing*, pages 33–38, 2006. ISBN 1-59593-306-9. doi: 10.1145/1143549.1143558. URL https://doi.org/10.1145/1143549.1143558.

[157] R. Zhang, Y. Zhang, and K. Ren. Distributed Privacy-Preserving Access Control in Sensor Networks. *IEEE Transactions on Parallel and Distributed Systems*, 23(8):1427–1438, 2012. doi: 10.1109/TPDS.2011.299. URL https://doi.org/10.1109/TPDS.2011.299.

[158] W. Zhang, C. Wang, and T. Feng. GP2S: Generic privacy-preservation solutions for approximate aggregation of sensor data. In *6th Annual IEEE International Conference on Pervasive Computing and Communications, PerCom 2008*, pages 179–184, 2008. doi: 10.1109/PERCOM.2008.60. URL https://doi.org/10.1109/PERCOM.2008.60.

Appendices

APPENDIX A

Result Reproduction

The results presented in this thesis were obtained with reproducibility in mind. All results obtained from simulations TOSSIM are deterministic for a given random seed. As part of this work a simulation framework used to run simulations, gather results, and analyse them was developed. The source codes for this framework can be found at (1) and the main development repository for TinyOS can be found at (2).

1. `https://bitbucket.org/Chen_Gu/slp-algorithms-tinyos`
2. `https://github.com/MBradbury/tinyos-main`

The raw results generated by TOSSIM can be found at the following location:

`https://doi.org/10.5281/zenodo.1323732`

This results include all TOSSIM results presented in Chapter 3, Chapter 4 and Chapter 5, and the directory tree of the results is shown in Appendix A.1.

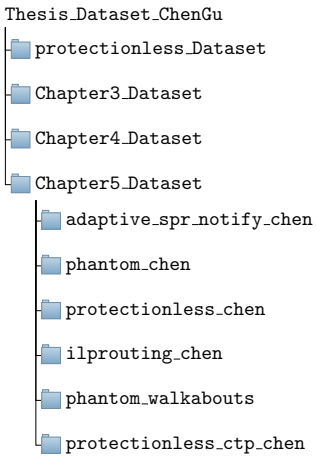


Figure A.1: The directory tree of the results file

APPENDIX B

Results of Sink-Source Distance

In the thesis, the network is modelled as a graph $G = (V, E)$ where V is the set of nodes and E is the set of links. An link is a 2-tuple (u, v) where u is the origin and v is the target. Each node in V is assigned a 2D coordinate and the distance is calculated as one hop between each node. The reason why hops are used as the distance between nodes is that an adversary traces back to the source by following these links. The hop distance is used in many places in the thesis. For example, nodes know 1-hop neighbours (i.e., the communication range of nodes is 1 hop); The attacker can sense all the communications within 1 hop distance; In the phantom routing and phantom walkabouts, the landmark node floods *away* and *beacon* messages to the normal nodes and notify them the distance between sink and themselves.

The distance between the sink and the source is important since the message is sent from the source to the sink. The sink-source distance is used to divide the source's neighbours into different sets. The sink-source distance varies depending on the network configurations and network sizes. Given a configuration and a network size, the sink-source distance presents the *minimum* hops that an adversary can travel back from the sink to the source. The sink-source distances used in Chapter 3 and Chapter 4 are shown in Table B.1, Table B.2 and Table B.3, and the results in Table B.4, Table B.5 and Table B.6 are used in Chapter 5. The results show that the sink-source distance varies under different models and network sizes[1].

[1] All the results are generated from 10000 repeats of protectionless flooding.

Network Size	SinkCorner (1 source)	SourceCorner (1 source)
11×11	10.3 ± 0.2	10.3 ± 0.2
15×15	14.1 ± 0.1	14.3 ± 0.2
21×21	20.1 ± 0.1	20.4 ± 0.2
25×25	24.2 ± 0.1	24.4 ± 0.2

Table B.1: The sink-source distance (hops) under the meyer-heavy communication model and ideal noise model with 1 source

Network Size	SinkCorner (2 sources)	SourceCorner (2 sources)
11×11	10.1 ± 0.2	9.3 ± 0.3
15×15	14.2 ± 0.2	13.4 ± 0.2
21×21	20.2 ± 0.2	19.4 ± 0.2
25×25	24.2 ± 0.2	23.4 ± 0.2

Table B.2: The sink-source distance (hops) under the meyer-heavy communication model and ideal noise model with 2 sources

Network Size	SinkCorner (3 sources)	SourceCorner (3 sources)
11×11	9.9 ± 0.4	9.0 ± 0.3
15×15	14.0 ± 0.3	13.0 ± 0.3
21×21	20.0 ± 0.3	19.1 ± 0.2
25×25	24.0 ± 0.3	23.1 ± 0.2

Table B.3: The sink-source distance (hops) under the meyer-heavy communication model and ideal noise model with 3 sources

Network Size	SinkCorner (1 source)	SourceCorner (1 source)
11×11	10.1 ± 0.1	10.3 ± 0.2
15×15	14.1 ± 0.1	14.3 ± 0.2
21×21	20.1 ± 0.1	20.3 ± 0.2
25×25	24.2 ± 0.1	24.4 ± 0.2

Table B.4: The sink-source distance (hops) under the casino-lab communication model and ideal noise model with 1 source

Network Size	SinkCorner (1 source)	SourceCorner (1 source)
11×11	7.7 ± 0.6	7.9 ± 0.7
15×15	10.8 ± 0.7	10.9 ± 0.7
21×21	15.3 ± 0.7	15.4 ± 0.7
25×25	18.3 ± 0.7	18.4 ± 0.7

Table B.5: The sink-source distance (hops) under the casino-lab communication model and low-asymmetry noise model with 1 source

Network Size	SinkCorner (1 source)	SourceCorner (1 source)
11×11	9.3 ± 0.9	9.7 ± 0.9
15×15	12.8 ± 0.8	13.2 ± 0.8
21×21	17.9 ± 0.8	18.4 ± 0.8
25×25	21.4 ± 0.8	21.8 ± 0.8

Table B.6: The sink-source distance (hops) under the meyer-heavy communication model and low-asymmetry noise model with 1 source