Understanding Source Location Privacy Protocols in Sensor Networks via Perturbation of Time Series

Matthew Bradbury and Arshad Jhumka
Department of Computer Science
University of Warwick, Coventry,
United Kingdom, CV4 7AL
{M.Bradbury, H.A.Jhumka}@warwick.ac.uk

***Abstract*—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. Much of the previous work on SLP has focused on the development of various protocols to enhance the level of SLP imparted to the network, under various attacker models and other conditions. Other work has focused on analysing the level of SLP being imparted by a specific protocol. In this paper, we adopt a different approach where we model the attacker movement as a time series and use information theoretic concepts to infer the properties of a routing protocol that imparts high levels of SLP. We propose the notion of a *properly competing path* that causes an attacker to "stall" when moving towards the source. This concept provides the basis for developing a *perturbation model*, similar to those in privacy-preserving data mining. We then show how to use properly competing paths to develop properties of an SLP-aware routing protocol. Further, we show how different SLP-aware routing protocols can be obtained through different instantiations of the framework. Those instantiations are obtained based on a notion of information loss achieved through the use of the perturbation model proposed.**

***Index Terms*—Source Location Privacy; Wireless Sensor Networks; Entropy; Mutual Information; Time Series.**

I. Introduction

Wireless sensor networks present a difficult challenge in creating secure and private applications due to their potential to expose important information about the environment that they exist in due to the broadcast nature of wireless communications. As messages are broadcasted, they can be eavesdropped by a malicious attacker. Even if encryption is used to protect the *content* of a message the *context* of the broadcast is still exposed for a malicious eavesdropper to take advantage of. One such context problem is that of Source Location Privacy (SLP), where attackers can monitor the pattern of broadcasts to gain knowledge about the location of the source of these messages.

Source Location Privacy was initially introduced in terms of the panda-hunter game [1] where a WSN has been deployed across a large area to monitor pandas in their natural habitat. Using a directional antenna, it was shown that an attacker could identify the location of the immediate source of a message and, using this information, trace messages back through the system and find the ultimate source of the messages and thus the panda (or other asset). SLP protection schemes aim to protect against this scenario though various techniques. Many techniques centre around increasing the time an attacker would take to capture the source by changing the routing protocol.

There has been much work on SLP [2, 3], with many new techniques having been developed, and the performance typically being evaluated through large-scale simulations. There are also several works that have developed models to analyse the privacy provided by their technique or protocol. However, these models tend to be specific to the type or nature of the technique. In this paper, we adopt a novel approach based on information theoretic arguments [4]. We assume a local eavesdropper attacker and model his movement as a (clear data) time series. Then, similar to the case for privacy-preserving data mining, we develop an approach for perturbing the clear data time-series to obtain a noisy time-series. However, unlike in privacy-preserving data mining where there is trade-off between information loss and privacy loss, no such trade-off is required here, meaning that information loss can be maximized, if possible, to minimize privacy loss.

We use a suitable definition of information loss and, together with the mutual information metric, use them to determine the properties of an SLP-aware routing protocol. Specifically, we develop a perturbation model such that the normal routing protocol, which is not SLP-aware, is transformed into an SLP-aware routing protocol whereby (i) a source can still do convergecast communication and (ii) the attacker cannot reach the source within a prespecified time limit. In essence, the SLP-aware routing protocol applies the perturbation technique to perturb the clear time series to generates a noisy time series, from which the attacker can learn little in identifying the source. Such an approach is beneficial as (i) it does not make any assumptions about the nature of the attacks and (ii) it does not make any assumptions about any particular protocol implementation.

We propose a novel concept called *proper competing paths* that captures the problem of whether the attacker can be "stalled" when moving towards the source. Proper competing paths are central to the perturbation model we propose in that (i) wherever proper competing paths exist, there is an increased entropy at that point, (ii) noisy time series made up of proper competing paths are more likely to have very small mutual information with the corresponding clear data time series. We will later explain how the technique can be adapted to generate state-of-

the-art SLP-aware routing protocols. Specifically, we make the following contributions:

- We formalize the design of an SLP-aware problem as a transformation problem.
- Using information theoretic concepts, we derive the requirements necessary to minimize the amount of information leaked by a noisy time series.
- We develop the concept of proper competing paths that underpin the perturbation model proposed.
- We propose two heuristics to (i) compute the set of proper perturbation paths and to (ii) transform a normal routing protocol into an SLP-aware routing protocol.
- We explain how the heuristic can be optimized to obtain state-of-the-art SLP-aware routing protocol.

The remainder of this paper is as follows: in Section II we cover the related work, including other model-based approaches to developing SLP solutions. Section III contains a description of the network, privacy and attacker models. Section IV outlines the problem statement and details the analysis used to guide SLP protocol development. The perturbation model is described in Section V and an example case study is presented in Section VI. In Section VII we discuss some of the issues raised with our work and finally we conclude in Section VIII.

II. Related Work

In the seminal work [1] the authors proposed a solution called *phantom routing*, where a directed random walk away from or towards a landmark node is first performed to a phantom node before that message is flooded in the network. An optimised version used a single path route from the phantom node to optimise the energy usage [5]. There has been much work developing improved version of the directed random walk phase of phantom routing, such as GROW [6] which uses a bloom filter to prevent the walk doubling back on itself and angle-based techniques [7, 8] which calculate angles between certain nodes to influence the direction of the walk. Other techniques have adapted the random walk such that it forms a ring around the source and messages are routed through the ring before being forwarded onwards to the source [9, 10].

Another technique has been to use *fake sources*, which are nodes in the network that broadcast messages encrypted and padded to be indistinguishable from normal messages from the source. There have been several implementations [11, 12] with the latest focusing on implementing a protocol that can dynamically determine good parameters to use online [13]. A criticism of fake sources is that they tend to use more energy compared to phantom routing.

Other techniques consist of a hybrid between generating fake messages and having messages modify their routing path. One example is tree-based diversionary routing [14] which imposes a tree structure on the network and then routes fake messages through the tree. The idea of fogs or clouds [15, 16] is also similar where a normal message is routed round a group of nodes called a fog and then onwards to other fogs. Fake messages are used to provide additional privacy.

Many of the techniques described thus far demonstrated their performance by simulation. There have only been a few approaches where information theoretic, statistical models or analysis have been used to either assist in designing a SLP protocol or evaluating how well it performs. For example, [17] developed a global protection scheme called Periodic in which every node send a message after a fixed period. This allowed perfect protection against an attacker with a global view of the network. In designing this solution the authors created a model involving traces of source detections, which was used to measure the privacy of those traces as well as the energy cost of providing SLP.

Following the global attacker theme, [18] took a different approach where statistical techniques were used to show that their global protection scheme provided high levels of SLP. This approach did not provide perfect global SLP as [17] did, but instead provided *statistically strong* SLP. Their model and solution aimed to make the distribution of message broadcasts from nodes indistinguishable from a certain statistical distribution (using the Anderson-Darling Test).

Other global protection schemes such as [19, 20] have also performed an analysis of their algorithms to justify their effectiveness. Although global protection schemes tend to be easier to formalise and reason about compared to providing SLP against a local attacker, there have been several different analysis approaches. In [21] the amount of source-location information that an individual message can leak to attackers was measured. They go on to analyse the effect of multiple routing paths originating from the source node, showing that more paths of longer lengths increases the SLP provided.

In [22] the authors performed an analysis to determine the safety period for their technique (the higher the safety period the greater the SLP provided). Their analysis focused on a tree-based scheme and how an attacker would navigate it. Their analysis requires a bounded time for message forwarding and that the number of messages repeated follows the Poisson probability distribution.

[23] takes a different approach and uses a thorough information theoretic analysis. The location of the asset is modelled as a random variable with the value in the set of all network nodes. Using a matrix of message forwarding probabilities the quantity of location information leaked is calculated. This solution is the most general as it does not rely on individual paths. However, the formalisation of the routing matrix is unlikely to be applicable to more complicated local routing schemes.

These analyses succeeded in assisting with the techniques they accompanied, but, there are deficiencies in these local routing analyses. For example a network may contain many separate routing paths, but if an attacker never encounters them then they will not increase the location privacy as [23] says they should. Also many analyses focus on paths that solely originate from the source, none have taken the perspective of fake sources or multiple real sources and how they affect the information leakage. Another issue is that the aim of these analyses have been to evaluate a single protocol, none that we

are aware of is designed to be generic enough to evaluate a wide range of protocols.

III. System Model

In this section, we present the models we assume in this paper.

A. Network Model

We consider a sensor network to be a graph $G = (V, E)$ where V represents the set of nodes and E the set of links between the nodes. When a link exists between two nodes m, n, then m and n can directly communicate with each other and are called neighbours. There is a dedicated node in the network called a sink, which is responsible for data collection and linking the network to the outside world. A *path* is a finite sequence $\langle n_1 \cdot n_2 \dots n_j \rangle$ of nodes. A *source-converging* path is a path with $n_j = s$ (In this paper, unless specified otherwise, a path means a source-converging path). A finite path can be converted into an infinite path through the introduction of loops. For example, when a (finite) path terminates at a source, we can augment it through the infinite repetition of the final node, i.e., self-loop at the source $\langle n_1 \cdot n_2 \dots s \cdot s \cdot s \dots \rangle$. For a given (finite) path p, we denote by $|p^n|$ (resp. $|p^t|$) the number of nodes (resp. the number of transitions) in p. We also denote by ${}^l\bar{p}$ (resp. $\bar{p}^l$) the prefix of p of length l (resp. the suffix of p of length l).

The nodes sense the environment, and when a node detects an event of interest (i.e., the presence of an asset) the node broadcasts information about the event to the network and is called a *source* node. In this paper, we focus on the case where there is only a *single* source node in the network (i.e., there is only a single asset to protect). We denote the source by s. We also assume that the source transmits a single message in every time unit. When a node transmits a new message, its neighbours forward the message in the following time unit. This process is repeated until the message is eventually collected at the sink. We thus assume that all nodes take exactly one time unit to receive, process and forward a new message. The message is routed towards the sink using a multi-hop routing protocol $\mathcal{R}$. We model the routing protocol $\mathcal{R}$ as a matrix, with $\mathcal{R}[i, j]$ representing the probability that node j receives a message from node i *first*, after the message is transmitted by the source. This means that a node can only potentially receive a message first from a neighbour that is closer to the source node. In other words, a node *cannot* receive a message first along a route which is not the shortest path from the source to that node.

B. Privacy Model

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. However, we make two observations:

1) 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.
2) 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 *safety period* in the literature. There are two alternative definitions of safety period: The first, used primarily by routing-based techniques, e.g. [5], is where the safety period is defined as the time required to capture the asset. The aim of these techniques is to maximise the safety period, i.e., the higher the time to capture, the higher the SLP level provided.

The second notion of safety period is used where it is desirable to bound the amount of time SLP is being considered for, i.e., if an attacker fails to capture a source within the specified safety period, then we say SLP has been provided. That notion of safety period intuitively captures the maximum time an asset will be at a given location before its next movement. Often, this can be obtained from previous data gathering to know more about such mobile assets.

This second notion of safety period is more generic that the first one in that, rather than attempting to maximise the amount of time an asset isn't captured (as under the first definition), the second definition captures the fact that the asset can't be reached before a certain time limit, i.e., setting the time limit to be ∞ in the second instance results in the first definition.

In this paper, we thus use the time bound model of safety period.

C. Attacker Model

It was shown in [24] that the strength of a WSN attacker can be factored along two dimensions, namely (i) presence and (ii) actions. Presence may, for example, be local while actions can be eavesdropping, crash or reprogramming among others. In this paper, we assume a *distributed eavesdropper* attacker. We chose a distributed attacker as the attacker can move around the network, gathering further information and the only action that he can perform is eavesdropping. Though being a weak attacker model, an attacker with a stronger set of actions will likely interfere with its stated objective of capturing the asset. For example, if the attacker attempts to jam signals at a given location, then the attacker cannot progress within the network to reach the asset within the specified safety period.

We assume the distributed eavesdropper to be initially located at the sink, since he 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. To achieve this, we assume that the attacker has sufficient capabilities to determine

the direction in which he receives the message, although the range of its detection is assumed to be limited and does not extend to the entire network. 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 node, whereby it captures the asset. Such a routing protocol that provides little protection to the source location is called a *normal* routing protocol $\mathcal{N}$ or protectionless.

The distributed eavesdropper attacker does not keep track of history information, i.e., it may revisit a node that it has previously visited. Thus, a path an attacker takes to capture an asset may contain loops. Also, given a safety period λ and since only one message is sent in a single time unit, we focus on path of length at least λ, i.e., if a path ending at the source s is of length less than λ, then we extend the path with sufficient number of repetitions of s until the path has length of λ.

IV. Problem Statement and Analysis

In this section, we first state the problem we address and subsequently analyse the problem to obtain relevant SLP-aware routing protocols.

A. Problem Statement

The problem we address is the following: 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 determine the (maximum) probability that the attacker will be able to reach the source and capture the asset within a specified maximum time bound. To ensure that the attacker does not follow any path he so desires (hence performing random backtracking), the attacker uses the routing protocol to achieve his objective of reaching the source node.

Formally, the problem specification is shown in Figure 1.

Given:

- A network $G = (V, E)$,
- A distributed eavesdropper attacker $\mathcal{A}$ that is located at the sink initially,
- A source location $s \in V$,
- A safety period λ,
- A maximum capture threshold δ that determines the SLP level required and
- A normal routing protocol $\mathcal{R}_{\mathcal{N}}$,

Objective:

Transform $\mathcal{R}_{\mathcal{N}}$ into $\mathcal{R}_{\mathcal{S}}$ such that:

- There exists a path from s to the sink using $\mathcal{R}_{\mathcal{S}}$.
- $\mathcal{A}$ reaches s with probability of at most δ within λ using $\mathcal{R}_{\mathcal{S}}$.

Fig. 1: Problem Definition: Transformation to a SLP-Aware Routing Protocol.

We call $\mathcal{R}_{\mathcal{S}}$ a λ-SLP routing protocol (or simply an SLP-aware routing protocol). Specifically, the objective is to understand the steps required to transform a normal (non SLP-aware) routing protocol into an SLP-aware routing protocol, i.e., to determine the properties that underpin a λ-SLP routing $\mathcal{R}_{\mathcal{S}}$. Observe that $\mathcal{R}_{\mathcal{N}}$ and $\mathcal{R}_{\mathcal{S}}$ do not need to have any relationship, with the exception that a source still needs to be able to send messages to the sink.

One way towards solving the above problem is to develop a protocol and then perform a performance analysis of the protocol to determine its efficiency, for example [23], thereby understanding the possible bottlenecks. Such an analysis then allows the protocol under analysis to be possibly refined. However, this technique constrains the search for a protocol that can optimally solve the above problem.

Given that an attacker takes a step along a single hop within a given time unit, we model the attacker movement as a time series. In doing this, we analyse the problem from the perspective of privacy protection of time-series data. This enables us to consider routing protocols abstractly to be able to determine the necessary properties. We thus consider the problem of quantifying the protection of time-series data which have been perturbed by some arbitrary perturbation model. As the perturbation model is closely related to the transformation of $\mathcal{R}_{\mathcal{N}}$ into $\mathcal{R}_{\mathcal{S}}$, this then provides the ability to determine the development of a SLP routing protocol that potentially minimises privacy loss, i.e., a δ-SLP routing protocol.

B. Problem Analysis

We first present the notations used in the rest of the paper before subsequently presenting an analysis of the transformation required.

1) Notation: We use the following notations in the rest of the paper, where $\mathcal{X}$ and $\mathcal{Y}$ are two random variables:

$$\Pr_{\mathcal{X}}(x) = \Pr(\mathcal{X} = x)$$
$$\Pr_{\mathcal{X}\mathcal{Y}}(x, y) = \Pr(\mathcal{X} = x, \mathcal{Y} = y)$$

2) Definitions:

- $\mathcal{N}$ is a random variable of attacker transitions under a protectionless routing protocol $\mathcal{R}_{\mathcal{N}}$. In essence, we let a trace of clear time-series data of an attacker movement under a protectionless routing $\mathcal{R}_{\mathcal{N}}$ to be a stochastic process $\mathcal{N} = \{\mathcal{N}_i\}$, where i is the time index and the $\mathcal{N}_i$'s form a sequence of random variables.
- $\mathcal{S}$ is a random variable of attacker transitions under a SLP routing protocol $\mathcal{R}_{\mathcal{S}}$. Thus, we let the trace of noisy time-series data generated by an SLP routing protocol $\mathcal{R}_{\mathcal{S}}$ to be a stochastic process $\mathcal{S} = \{\mathcal{S}_i\}$, where i is the time index and the $\mathcal{S}_i$'s form a sequence of random variables.
- $\mathcal{R}_{\mathcal{N}}$ is a routing matrix for normal routing, where $\mathcal{R}_{\mathcal{N}}[i, j]$ represents the probability j receives a message from i first under $\mathcal{N}$
- $\mathcal{R}_{\mathcal{S}}$ is a routing matrix for SLP aware routing, where $\mathcal{R}_{\mathcal{S}}[i, j]$ represents the probability j receives a message from i first under $\mathcal{S}$
- There is a safety period λ which is the amount of time units or steps it takes the attacker to reach a source (on average) from the time it first receives a normal message.
- The time domain is denoted by $\mathcal{T}$.

Note that our routing matrices are different to those in [23] which contain the probability that the routing algorithm chooses the next node. Our routing matrices contain the probability a message is received from a node.

Thus, in this setting, the objective can be recast as follows: develop a transformation $\mathcal{P}$ that transforms $\mathcal{R}_{\mathcal{N}}$ into $\mathcal{R}_{\mathcal{S}}$ such that, using some information theoretic measure of a *dissimilarity* metric, the dissimilarity between $\mathcal{N}$ and $\mathcal{S}$ captures the SLP level imparted by $\mathcal{P}$, which implements the SLP-aware protocol.

3) Assumptions:

- An attacker can distinguish between a message that has been seen before and a message that has never been seen.

C. Analysis

In this section, we identify the characteristics of a routing protocol that can provide a high level of SLP. To do this, we need a measure of privacy for evaluating the level on SLP enhancement provided by a given solution. There are several potential definitions for a privacy metrics and this paper provides a survey of some of these metrics [25].

In this paper, we focus on the *mutual information* metric between two random variables $\mathcal{N}$ and $\mathcal{S}$, denoted by $I(\mathcal{N};\mathcal{S})$, which is defined as follows:

$$I(\mathcal{N};\mathcal{S}) = H(\mathcal{N}) - H(\mathcal{N}|\mathcal{S}) \tag{1}$$

$H(\mathcal{N})$ denotes the entropy on H whereas $H(\mathcal{N}|\mathcal{S})$ denotes the entropy on H given S. If the uncertainty on $\mathcal{N}$ is the same when (an SLP protection scheme) $\mathcal{S}$ is in use, then it implies that no SLP protection is lost, i.e., the attacker does not learn anything new about $\mathcal{N}$ through $\mathcal{S}$. Then, it means that $\mathcal{N}$ and $\mathcal{S}$ are independent.

$$\text{if } H(\mathcal{N}) = H(\mathcal{N}|\mathcal{S}) \implies I(\mathcal{N};\mathcal{S}) = 0 \tag{2}$$

Now, to obtain $I(\mathcal{N};\mathcal{S})$, we need to calculate $H(\mathcal{N}|\mathcal{S})$:

$$\begin{aligned} H(\mathcal{N}|\mathcal{S}) &= \sum_{f\in\mathcal{S}} \Pr_{\mathcal{S}}(f)\, H(\mathcal{N}|\mathcal{S}=f) \\ &= -\sum_{n\in\mathcal{N}}\sum_{f\in\mathcal{S}} \Pr_{\mathcal{S}}(f)\Pr_{\mathcal{N}}(n|f)\log\Pr_{\mathcal{N}}(n|f) \\ &= -\sum_{n\in\mathcal{N}}\sum_{f\in\mathcal{S}} \Pr_{\mathcal{S}}(f)\frac{\Pr_{\mathcal{NS}}(n,f)}{\Pr_{\mathcal{S}}(f)}\log\Pr_{\mathcal{N}}(n|f) \\ &= \sum_{n\in\mathcal{N}}\sum_{f\in\mathcal{S}} \Pr_{\mathcal{NS}}(n,f)\log\frac{\Pr_{\mathcal{S}}(f)}{\Pr_{\mathcal{NS}}(n,f)} \end{aligned} \tag{3}$$

Substituting Equation 3 into Equation 1, mutual information is then given by:

$$I(\mathcal{N};\mathcal{S}) = \sum_{n\in\mathcal{N}}\sum_{f\in\mathcal{S}} \Pr_{\mathcal{NS}}(n,f)\log\frac{\Pr_{\mathcal{NS}}(n,f)}{\Pr_{\mathcal{N}}(n)\Pr_{\mathcal{S}}(f)} \tag{4}$$

To evaluate $I(\mathcal{N};\mathcal{S})$, we need to evaluate $\Pr_{\mathcal{NS}}(n,f)$ in Equation 4. Using a time period $\lambda' \propto \lambda$, we have:

$$\Pr_{\mathcal{NS}}(n,f) = \Pr_{\mathcal{N}}(n|f)\Pr_{\mathcal{S}}(f) \tag{5}$$

$$= \sum_{\tau=0}^{\lambda'} \Pr_{\mathcal{NT}}(n,\tau|f)\Pr_{\mathcal{S}}(f) \tag{6}$$

$\Pr_{\mathcal{NS}}(n,f)$ captures the probability that the attacker will take transition n within λ' steps, if transition f is his next transition. Here, with a safety period of λ, the value λ' represents the amount of time units the attacker needs to be "stalled" (e.g., either through diversion or bypassing it). Typically, the safety period $\lambda \leq \lambda' + \alpha$, where α represents the minimum number of time units for an attacker to reach the source, starting from the sink. Thus, in the rest of the paper, since $\lambda' \propto \lambda$, λ' can be easily evaluated from λ.

$$\Pr_{\mathcal{N}}(n|\mathcal{S}=f) = \sum_{n\in\mathcal{N}}\sum_{\tau=0}^{\lambda'} \Pr_{\mathcal{NT}}(n,\tau|f) \tag{7}$$

$$= \sum_{n\in\mathcal{N}} \left(\omega^{f} \cdot \sum_{\tau=0}^{\lambda'} (\mathcal{R}'_{\mathcal{S}})^{\tau} \cdot \omega^{n\top} \right) \tag{8}$$

ω^x is a vector whose length is equal to the number of nodes in the network and is defined as follows:

$$\omega^x = \begin{cases} 1 & \text{if } x\text{th entry} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

On the other hand, $\omega^{n\top}$ represents the transpose of ω^n. The matrix $\mathcal{R}'_{\mathcal{S}}$ is obtained from $\mathcal{R}_{\mathcal{S}}$ as follows:

$$\mathcal{R}'_{\mathcal{S}}[i,j] = \begin{cases} \mathcal{R}_{\mathcal{S}}[i,j] & \text{if } (i,j) \neq n \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

For $I(\mathcal{N};\mathcal{S})$ (Equation 4) to be minimised, then $\Pr_{\mathcal{NS}}(n,f)$ needs to be 0 for all relevant (n,f) combinations, i.e., the objective is to define $\mathcal{R}_{\mathcal{S}}$ such that $\Pr_{\mathcal{N}}(n|f)$ is minimised.

Now, to compute $\Pr_{\mathcal{N}}(n|f)$, we need $\mathcal{R}'_{\mathcal{S}}$ which, in turn, requires $\mathcal{R}_{\mathcal{S}}$ to be defined[1]. The way $\mathcal{R}_{\mathcal{S}}$ is defined gives rise to different SLP protocols. As mentioned earlier, we seek a notion of dissimilarity between $\mathcal{R}_{\mathcal{S}}$ and $\mathcal{S}$ such that the dissimilarity level is indicative of the SLP level provided by $\mathcal{S}$. It was shown in [26] that the notion of information loss varies inversely with privacy loss (see Figure 2), i.e., the higher the information that is lost or the more perturbed the clear data time-series is, the less privacy is lost.

To this end, we adapt the following definition of information loss [25] (Equation 11), which is used in privacy-preserving data mining, to suit the SLP problem (see Equation 12):

$$IL(D_{\mathcal{N}}, D_{\mathcal{S}}) = \frac{\sum_{i=1}^{n} |f_{D_{\mathcal{N}}}(i) - f_{D_{\mathcal{S}}}(i)|}{\sum_{i=1}^{n} f_{D_{\mathcal{N}}}(i)} \tag{11}$$

[1]Since $\mathcal{R}'_{\mathcal{S}} = \mathcal{R}_{\mathcal{S}}$ for all entries except for one entry n, we will use $\mathcal{R}_{\mathcal{S}}$ to mean $\mathcal{R}'_{\mathcal{S}}$ in the rest of the paper.

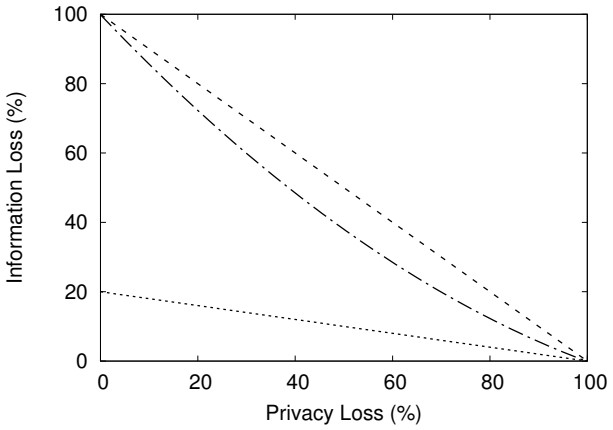

Fig. 2: Information Loss vs Privacy Loss representative of [26, Fig. 8] depending on the perturbation model used.

where $D_{\mathcal{N}}$ and $D_{\mathcal{S}}$ represent the clear and noisy domains respectively, and $f_D(i)$ represents the frequency of the data item i in domain D.

Since we assume normal transitions can be used, especially after the safety period has elapsed, we adapt the information loss definition as follows:

$$IL(D_{\mathcal{N}}, D_{\mathcal{S}}) = \frac{\sum_{i=1}^{n} \left| f_{D_{\mathcal{N}}}(i) - f_{D_{\mathcal{S}}}(i^{\lambda}) \right|}{\sum_{i=1}^{n} f_{D_{\mathcal{N}}}(i)} \tag{12}$$

Where $\mathcal{F}_{\mathcal{D}_{\mathcal{N}}}(i)$ and $\mathcal{F}_{\mathcal{D}_{\mathcal{S}}}(i^{\lambda})$ are defined as:

$$\mathcal{F}_{\mathcal{D}_{\mathcal{N}}}(i) = \begin{cases} 1 & \text{if transition } i \text{ is used in } \mathcal{N} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

$$\mathcal{F}_{\mathcal{D}_{\mathcal{S}}}(i^{\lambda}) = \begin{cases} 1 & \text{if } i \text{ is not taken within } \lambda \text{ steps in } \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

Basically, Equation 12 states that the more dissimilar the set of transitions taken within λ time units are, the greater is the information loss, hence the lesser the privacy loss.

If $IL(D_{\mathcal{N}}, D_{\mathcal{S}}) = 1$ (i.e., is maximum), then it implies that $D_{\mathcal{N}} \cap D_{\mathcal{S}} = \emptyset$. In other words, to minimise privacy loss, $\mathcal{R}_{\mathcal{N}}$ and $\mathcal{R}_{\mathcal{S}}$ cannot share any transition. More specifically, it means that, though $\mathcal{R}_{\mathcal{N}}$ and $\mathcal{R}_{\mathcal{S}}$ can share transitions, an attacker cannot take some transition in $\mathcal{R}_{\mathcal{N}}$ under $\mathcal{R}_{\mathcal{S}}$ within λ time units. This then means that $\mathcal{R}_{\mathcal{N}}$ has to be transformed in such a way that *for a certain duration*, for any transition (i, j) unique in $\mathcal{R}_{\mathcal{S}}$, an attacker at location j needs to receive a message from i first.

V. Perturbation Model: Proper Competing Paths

To understand the transformation of $\mathcal{R}_{\mathcal{N}}$ into $\mathcal{R}_{\mathcal{S}}$, i.e., to understand how $\mathcal{R}_{\mathcal{N}}$ can perturbed into $\mathcal{R}_{\mathcal{S}}$, we introduce the concept of *competing paths*.

Definition 1 (Competing Paths): Given a network $G = (V, E)$ and a protectionless routing protocol $\mathcal{R}_{\mathcal{N}}$, two distinct paths p_1 and p_2 under $\mathcal{R}_{\mathcal{N}}$ compete at a node $n \in V$ iff the following are satisfied:

- p_1 and p_2 are source-converging paths.
- $\exists (i, j), (i, j') \in E : (i, j) \in p_1 \wedge (i, j') \in p_2 : i = n$
- $\mathcal{R}_{\mathcal{N}}[j, n] > 0 \wedge \mathcal{R}_{\mathcal{N}}[j', n] \geq 0, j \neq j'$

The idea of competing paths is that if one path is part of the clear data time-series, then the other can be used in the noisy data time-series. Specifically, it means that if the attacker is more likely to follow a given path p_1 under $\mathcal{R}_{\mathcal{N}}$, then the attacker can be made to follow path p_2 under $\mathcal{R}_{\mathcal{S}}$. We call node n a *junction* node. To make this concrete, consider Figure 3. Since $\mathcal{R}_{\mathcal{N}}[2, 5] = 0.5$ and $\mathcal{R}_{\mathcal{N}}[4, 5] = 0.5$, then paths $\langle (5, 2) \cdot (2, 1) \rangle$ and $\langle (5, 4) \cdot (4, 1) \rangle$ compete at node 5. Observe that the notion of competing paths increases the entropy at the node they are competing at. We call p_1 a normal path and p_2 a perturbed path.

However, as can be observed, not all competing paths can prevent the attacker from reaching the source within the required safety period. We thus strengthen the notion of competing paths to that of *proper* competing paths.

Corollary 1: All paths compete at the sink.

Definition 2 (Proper Competing Paths): Given a network $G = (V, E)$ and a protectionless routing protocol $\mathcal{R}_{\mathcal{N}}$, two distinct paths p_1 and p_2 under $\mathcal{R}_{\mathcal{N}}$ compete properly at a node $n \in V$ iff the following are satisfied:

- p_1 and p_2 are source-converging paths.
- $\exists (i, j), (i, j') \in E : (i, j) \in p_1 \wedge (i, j') \in p_2 : i = n$
- $\mathcal{R}_{\mathcal{N}}[j, n] > 0 \wedge \mathcal{R}_{\mathcal{N}}[j', n] = 0$

Here, for two proper competing paths, the attacker cannot receive the message first along one of these paths and we call node n a *proper junction* node. Thus, path p_1 should be perturbed into path p_2 in the noisy data time-series. The intuition is that the attacker, at a proper junction node, has two distinct choices and one of those choices is one he would unlikely have made under normal circumstances. As before, we call p_1 a normal path and p_2 a properly perturbed path.

Lemma 1: Given a network $G = (V, E)$, a protectionless routing protocol $\mathcal{R}_{\mathcal{N}}$, an attacker $\mathcal{A}$ that starts at the sink, safety period λ, and a path p_1 under $\mathcal{R}_{\mathcal{N}}$ with $|p_1^t| \leq \lambda$. Then there exists a path p_2 with $|p_2^t| > \lambda$ such that $\exists n \in p_1, p_2$ and p_1 and p_2 properly compete at n.

Proof. We have to prove that, along p_1, there exists a node n such that p_1 and p_2 properly compete at n and that $|p_2^t| > \lambda$.

We prove this by contradiction: We assume that there exists no such path p_2 and subsequently show a contradiction. Consider a node $i \in p_1$ and one of its neighbours j which is further from the source. According to our assumption, we have that $0 < \mathcal{R}_{\mathcal{N}}[i, j] \leq 1$. Since we assume that no p_2 exists that can properly compete with p_1, i.e., no proper junction node exists, we have that $\forall j, (i, j) \in E : 0 < \mathcal{R}_{\mathcal{N}}[i, j] \leq 1$. However, $\exists k : (i, k) \in E$ is further from the source than i, then we have that $\mathcal{R}_{\mathcal{N}}[k, i] = 0$, which is a contradiction. Thus, such a p_2 exists and the proper junction node is node j. Path p_2 is also of length $|p_2^t| > \lambda$, by repeating the transition (i, j) λ times. □

The intuition is that p_1 is a path that an attacker may follow under the protectionless protocol $\mathcal{R}_{\mathcal{N}}$ to capture the asset,

while p_2 provides a "diversion" via a path that the attacker will not normally follow. This path should be long enough to adequately delay the attacker. Further, p_2 also captures the fact that p_1 can be perturbed into p_2 at the identified proper junction node. Such a path p_2 will need to be guaranteed to occur under a routing protocol.

Using Figure 3, for example, we have the following:

- $p_1 := \langle(5,4)\cdot(4,1)\rangle$. Setting $\lambda = 4$, we have $p_1 := \langle(5,4)\cdot(4,1)\cdot(1,1)\cdot(1,1)\cdot(1,1)\cdot(1,1)\cdot(1,1)\rangle$.
- Then, p_2 as follows: $p_2 := \langle(5,4)\cdot(4,5)\cdot(5,4)\cdot(4,5)\cdot(5,4)\cdot(4,1)\cdot(1,1)\rangle$. Notice that the attacker reaches the asset at node 1 in p_1 within λ steps while it doesn't in p_2.

At this point, a proper junction point needs to be selected. Since the attacker starts at the sink, the sink is the first candidate to be considered as a proper junction point. If the sink is not a proper junction point, we iteratively select a normal path under $\mathcal{R}_\mathcal{N}$ at random and identify proper junction points. Once a proper junction node is selected, a properly perturbed path needs to be selected to obtain the routing matrix $\mathcal{R}_\mathcal{S}$, i.e., choosing a set of properly perturbed paths.

Hence, we propose the following heuristic to generate the set of properly perturbed paths at a given proper junction node.

Algorithm 1 Generating set of properly perturbed paths

```
NormalPaths ← set of all normal paths from sink
while NormalPaths ≠ ∅ do
    NextNode ← sink
    while NextNode ≠ source do
        if ∃ i : R_N[i, NextNode] = 0 then
            ppp ← {p | generate path p through i
                such that ^λp̄ ∉ R_N}
        else
            NextNode ← CHOOSE({i|(i, NextNode) ∈
E, R_N[i, NextNode] ≠ 0})
```

Then, we generate the routing matrix $\mathcal{R}_\mathcal{S}$ for the perturbed time series as follows:

Algorithm 2 Generating SLP-aware routing matrix

```
R_S ← R_N
select one path p_2 ∈ ppp from above.
for (i, j) ∈ p_2 do
    R_S[, i] ← 0          ▷ Deleting normal paths
    R_S[j, i] ← 1         ▷ Ensuring attacker follows chosen
                            properly perturbed path
```

Theorem 1 (Privacy Loss): A network $G = (V, E)$, A distributed eavesdropper attacker $\mathcal{A}$ that is located at the sink initially, a source location $s \in V$, a safety period λ, a maximum capture threshold δ that determines the SLP level required and a normal routing protocol $\mathcal{R}_\mathcal{N}$, then $\mathcal{R}_\mathcal{S}$ generated as above results in $\Pr_\mathcal{N}(n|\mathcal{S} = f) = 0$ (Equation 8), with $\delta = 0$.

Proof. By construction, $\sum_{n\in\mathcal{N}}\left(\omega^f \cdot \sum_{\tau=0}^{\lambda}(\mathcal{R}_\mathcal{S}^{'})^\tau \cdot \omega^{n^\top}\right) = 0$ as $\sum_{\tau=0}^{\lambda}(\mathcal{R}_\mathcal{S}^{'})^\tau = 0$, with $(\mathcal{R}_\mathcal{S}^{'})^\tau = 0$ defined as in Equation 10, solving the problem of Figure 1 . □

The intuition here is that, by construction, the attacker is sent on a diversion long enough that it only takes a normal transition after the safety period has elapsed. Further, Theorem 1 also shows that the resulting SLP-aware routing protocol causes minimal privacy loss as the attacker does not reach a normal transition within the safety period, i.e., according to Equation 12, the information loss is maximal as $\mathcal{R}_\mathcal{S}$ reuses some transitions of $\mathcal{R}_\mathcal{N}$, but only after the safety period is over.

VI. A Case Study

In this section, we present a simple case study showing the viability of our approach. Using the network shown in Figure 3, with the sink at the centre of the network and the source on the border of the network, we define $\mathcal{R}_\mathcal{N}$ as below. For example, $\mathcal{R}_\mathcal{N}[1,2] = 1$ means that node 2 will receive a message first from node 1. On the other hand, $\mathcal{R}_\mathcal{N}[2,5] = 0.5$ means there is a 50% chance that node 5 will receive a message first from node 2 (the other possibility is from 4). The set of normal paths is given by $\{\langle(5,2)\cdot(2,1)\rangle, \langle(5,4)\cdot(4,1)\rangle\}$. Here, we set λ, the safety period to 4.

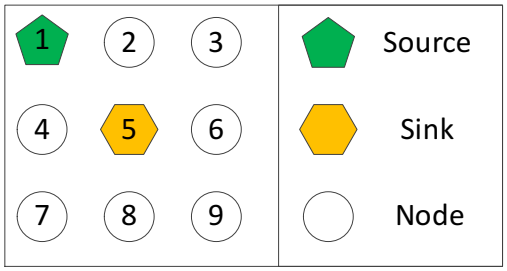

Fig. 3: Example 3×3 network grid

$\mathcal{R}_\mathcal{N}$:

Sending Nodes \ Receiving Nodes	1	2	3	4	5	6	7	8	9
1	0	1	0	1	0	0	0	0	0
2	0	0	1	0	0.5	0	0	0	0
3	0	0	0	0	0	0.5	0	0	0
4	0	0	0	0	0.5	0	1	0	0
5	0	0	0	0	0	0.5	0	0.5	0
6	0	0	0	0	0	0	0	0	0.5
7	0	0	0	0	0	0	0	0.5	0
8	0	0	0	0	0	0	0	0	0.5
9	0	0	0	0	0	0	0	0	0

(15)

Here, the first proper junction node is node 5, the sink. Any properly perturbed path will start with transition $(5,6)$ or $(5,8)$. Applying the heuristic to generate $\mathcal{R}_\mathcal{S}$ shows that (i) it is first required to reset the values of node 5 and then set the value of the relevant transition to 1. This is shown in the matrix $\mathcal{R}_\mathcal{S}$, where the old value is stricken off and replaced by the new value, shown in bold. For example, it means that $\mathcal{R}_\mathcal{S}[6,5]$ needs to be set to 1 to ensure that the attacker moves to node

6 from the sink, rather than move towards either node 2 or 4. This is performed for each transition in one of the properly perturbed paths.

As can be observed, an attacker will now take the following path: $5 \cdot 6 \cdot 9 \cdot 8 \cdot 7 \cdot 4 \cdot 1$, meaning that the attacker requires 6 transitions to reach the source, more than the safety period. Hence, it means that the attacker cannot catch the source before the safety period has expired.

$\mathcal{R}_{\mathcal{S}}$:

Sending Node \ Receiving Nodes	1	2	3	4	5	6	7	8	9
1	0	1	0	1	0	0	0	0	0
2	0	0	1	0	~~0.5~~ **0**	0	0	0	0
3	0	0	0	0	0	~~0.5~~ **0**	0	0	0
4	0	0	0	0	~~0.5~~ **0**	0	1	0	0
5	0	0	0	0	0	~~0.5~~ **0**	0	~~0.5~~ **0**	0
6	0	0	0	0	~~$\emptyset$~~ **1**	0	0	0	~~0.5~~ **0**
7	0	0	0	0	0	0	0	~~0.5~~ **1**	0
8	0	0	0	0	0	0	0	0	~~0.5~~ **1**
9	0	0	0	0	0	~~$\emptyset$~~ **1**	0	0	0

(16)

VII. Discussion

In this section, we briefly discuss some issues raised by the methodology.

- There are different ways to obtain the relevant proper perturbing paths. Specifically, a path that contains no loop may be obtained while another path with loops may be considered. Since the notion of paths captures source-converging paths, it means that Equation 12 will never be a maximum as $D_{\mathcal{S}}$ will contain some elements of $D_{\mathcal{N}}$. However, minimizing the number of common elements will result in high information loss. Specifically, it is better to introduce loops in (the non-overlapping elements of) $D_{\mathcal{S}}$ than in $D_{\mathcal{N}}$ as this will reduce the number of common elements.
- In this paper, we have assumed that links are bidirectional and lossless, such that the sum of each column, apart from the source, adds to 1. However, when links become unidirectional or lossy (for example due to message collisions), the sum may be less than 1. This also means that the domain $D_{\mathcal{S}}$ may be different to when the links are bidirectional. In this case, some of the work we propose here will have to be adapted to specifically account for unidirectional links. On the other hand, if the unidirectional nature of links is transient, i.e., is short-lived, the current framework can still work if nodes are made to perform retransmissions (at the link-layer level). However, this technique will not work if message collisions occur. Most often, sensor nodes are not equipped with collision detectors and it is entirely possible that the matrix $\mathcal{R}_{\mathcal{N}}$ is different to the one assumed, as a node j may receive a message from node i first (in practice) rather than from node k (as specified by $\mathcal{R}_{\mathcal{N}}[k, j] = 1$ and $\mathcal{R}_{\mathcal{N}}[i, j] = 0$). Dealing with these two issues are part of our future work.
- The framework we have proposed is independent of any routing protocol. However, to provide SLP, a routing protocol will need to provide guarantees that $\mathcal{R}_{\mathcal{S}}$ will be supported. In the example shown in Section VI, since node 5 needs to receive a message from node 6 first, then a possible implementation may require control messages to be sent to notify node 6 to send a "fake" messages, identical to normal messages, to node 5 to get the attacker to move to node 6 first. So, in effect, our framework provides indication of the requirements to provide high levels of SLP. On a further note, instantiating the routing matrix $\mathcal{R}_{\mathcal{S}}$ differently actually gives rise to different state-of-the-art SLP-aware routing protocols. For example, if $\mathcal{R}_{\mathcal{S}}$ is set as follows, then the routing protocol is the one that has been proposed in [13].

Sending Node \ Receiving Nodes	1	2	3	4	5	6	7	8	9
1	0	1	0	1	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0
3	0	0	0	0	0	0.5	0	0	0
4	0	0	0	0	0	0	1	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0.5	0	0	0	0.5
7	0	0	0	0	0	0	0	1	0
8	0	0	0	0	0.5	0	0	0	0.5
9	0	0	0	0	0	0.5	0	0	0

(17)

VIII. Conclusion

In this paper, we address the source-location privacy problem in wireless sensor networks from an information theoretic viewpoint. One major advantage of using such an approach is that it allows specific attacks and protocols to be abstracted away, focusing instead on the amount of information that is leaked or lost/gained by attacker. While several other works focused on analysing specific routing protocols or privacy metrics, we focused on understanding the basis of routing transformations to minimize the mutual information metric. Our framework is novel in that it allows the SLP-aware routing matrix to be configured in different ways, to give rise to potentially different SLP-aware routing protocols.

As future work, we will focus on (i) integrating energy usage into the framework so a trade-off between SLP and energy can be analysed and (ii) implementing the relevant SLP-aware routing protocols that provide potentially optimal SLP levels to gauge their performance in more natural settings. We will also study the problem of selecting the best proper junction node(s).

References

[1] 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*, ser. SASN '04. New York, NY, USA: ACM, 2004, pp. 88–93.

[2] M. Conti, J. Willemsen, and B. Crispo, "Providing source location privacy in wireless sensor networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1238–1280, 2013.

[3] R. Rios, J. Lopez, and J. Cuellar, *Foundations of Security Analysis and Design VII: FOSAD 2012/2013 Tutorial Lectures*. Cham: Springer International Publishing, 2014, ch. Location Privacy in WSNs: Solutions, Challenges, and Future Trends, pp. 244–282.

[4] C. E. Shannon, "A mathematical theory of communication," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55, Jan. 2001.

[5] 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, 2005. ICDCS 2005. Proceedings.*, Jun. 2005, pp. 599–608.

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

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

[8] P. Spachos, D. Toumpakaris, and D. Hatzinakos, "Angle-based dynamic routing scheme for source location privacy in wireless sensor networks," in *Vehicular Technology Conference (VTC Spring), 2014 IEEE 79th*, May 2014, pp. 1–5.

[9] S. Li, Y. Xiao, Q. Lin, and Z. Qi, "A novel routing strategy to provide source location privacy in wireless sensor networks," *Wuhan University Journal of Natural Sciences*, vol. 21, no. 4, pp. 298–306, 2016.

[10] L. Yao, L. Kang, F. Deng, J. Deng, and G. Wu, "Protecting source–location privacy based on multirings in wireless sensor networks," *Concurrency and Computation: Practice and Experience*, vol. 27, no. 15, pp. 3863–3876, 2015.

[11] A. Jhumka, M. Bradbury, and M. Leeke, "Towards understanding source location privacy in wireless sensor networks through fake sources," in *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom'12)*, Jun. 2012, pp. 760–768.

[12] ——, "Fake source-based source location privacy in wireless sensor networks," *Concurrency and Computation: Practice and Experience*, vol. 27, no. 12, pp. 2999–3020, 2015.

[13] 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 (TrustCom'15)*, Aug. 2015, pp. 531–538.

[14] J. Long, M. Dong, K. Ota, and A. Liu, "Achieving source location privacy and network lifetime maximization through tree-based diversionary routing in wireless sensor networks," *Access, IEEE*, vol. 2, pp. 633–651, 2014.

[15] M. Dong, K. Ota, and A. Liu, "Preserving source-location privacy through redundant fog loop for wireless sensor networks," in *2015 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)*, Oct. 2015, pp. 1835–1842.

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

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

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

[19] H. Park, S. Song, B. Y. Choi, and C. T. Huang, "Passages: Preserving anonymity of sources and sinks against global eavesdroppers," in *INFOCOM, 2013 Proceedings IEEE*, Apr. 2013, pp. 210–214.

[20] A. Proano, L. Lazos, and M. Krunz, "Traffic decorrelation techniques for countering a global eavesdropper in WSNs," *IEEE Transactions on Mobile Computing*, vol. PP, no. 99, pp. 1–1, 2016.

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

[22] A. A. Nezhad, A. Miri, and D. Makrakis, "Location privacy and anonymity preserving routing for wireless sensor networks," *Computer Networks*, vol. 52, no. 18, pp. 3433–3452, 2008.

[23] S. Armenia, G. Morabito, and S. Palazzo, "Analysis of location privacy/energy efficiency tradeoffs in wireless sensor networks," in *Proceedings of the 6th international IFIP-TC6 conference on Ad Hoc and sensor networks, wireless networks, next generation internet*, ser. NETWORKING'07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 215–226.

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

[25] E. Bertino, D. Lin, and W. Jiang, *A Survey of Quantification of Privacy Preserving Data Mining Algorithms*. Boston, MA: Springer US, 2008, pp. 183–205.

[26] D. Agrawal and C. C. Aggarwal, "On the design and quantification of privacy preserving data mining algorithms," in *Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, ser. PODS '01. New York, NY, USA: ACM, 2001, pp. 247–255.