R. Rios, and J. Lopez, "Analysis of Location Privacy Solutions in Wireless Sensor Networks", *IET Communications*, vol. 5, pp. 2518 - 2532, 2011.
http://doi.org/10.1049/iet-com.2010.0825
NICS Lab. Publications: https://www.nics.uma.es/publications

Analysis of Location Privacy Solutions in Wireless Sensor Networks

Ruben Rios
Computer Science Department
University of Malaga
29071, Spain
ruben@lcc.uma.es

Javier Lopez
Computer Science Department
University of Malaga,
29071, Spain
jlm@lcc.uma.es

Abstract

Extensive work has been done on the protection of Wireless Sensor Networks (WSNs) from the hardware to the application layer. However, only recently, the privacy preservation problem has drawn the attention of the research community because of its challenging nature. This problem is exacerbated in the domain of WSNs due to the extreme resource limitation of sensor nodes. In this paper we focus on the location privacy problem in WSNs, which allows an adversary to determine the location of nodes of interest to him. We provide a taxonomy of solutions based on the power of the adversary and the main techniques proposed by the various solutions. In addition, we describe and analyse the advantages and disadvantages of different approaches. Finally, we discuss some open challenges and future directions of research.

1 Introduction

In the near future the world is expected to be scattered with smart objects offering potentially tremendous opportunities both to industry and final users. In such foreseeable scenario, already known as the Internet of Things [4], everyday objects will be fitted with computational and sensing capabilities. However, before the success of such scenario, many outstanding challenges need to be examined. Among them we deem critical the ability to integrate different supporting technologies in such a way that they can provide an adequate level of security to the services being offered and to the data traversing the network. Well-known security and privacy problems already exist in current networks, hence, the creation of a trustworthy and resilient network as a conglomerate of networks and services might be unlikely if both security and privacy are not taken into consideration in advance.

Under the vision of a world cluttered with smart interacting objects, we focus on one of its supporting technologies, Wireless Sensor Networks (WSNs), and its privacy implications. WSNs are ad hoc networks comprised of tiny devices (sensor nodes) which are able to monitor the physical phenomena occurring in their vicinity. These devices can communicate the sensed information in a hop-by-hop fashion to a special node called the base station or the sink, which collects and analyses the received data. Sensor nodes are extremely constrained in terms of processing, storing and communicating capabilities as well as in terms of battery supply, which in most cases is irreplaceable. Such resource limitation makes sensor nodes extremely vulnerable to different types of threats and attacks [39]. In particular, privacy is one of such threats.

The concept of privacy is broad in the sense that it encompasses several interpretations depending on the context as well as the cultural aspects [3]. Traditionally, privacy was considered "the right (of people) to be let alone" [41], however, as the technology evolves, the concept

has acquired new nuances that affect both individuals and corporations. From a technological viewpoint there are mainly two approaches to the protection of privacy, the privacy-by-design and the privacy-by-policy approach [38]. While the former focuses on the application of privacy enhancing technologies in the system design to prevent the leakage of sensitive information, the latter emphasizes the need to inform users about data collection, to minimize the collection and misuse of personal data, and to adequately protect these data (i.e. Fair Information Practices). However, a privacy-by-design approach is usually desirable over a privacy-by-policy approach since it reduces the risk of user data misuse.

In general, the design of privacy enhancing technologies has traditionally been focused on preventing the behaviour of individuals from being tracked without their explicit consent. This perspective can be referred to as social privacy. However, in the field of WSNs there are also network privacy considerations that might help reveal not only information about the network itself but also about the events being monitored, which in turn might be related to individuals. It is interesting to note that in such scenarios traditional security mechanisms on their own cannot provide an adequate protection level. While message content confidentiality can be protected through encryption, an adversary can still gain sensitive information about individuals or their context by observing the performance of the network. Specifically, privacy preservation is further complicated in WSNs due to the nature of these networks. Firstly, the computational limitations of sensor nodes require the adaptation of cryptographic techniques [30] and, in some cases, it might render the use of heavy cryptographic algorithms infeasible. Secondly, sensor nodes communicate wirelessly and usually with a single entity, the base station, which is placed several hops away. Thus, the messages must travel several hops from the source to the destination. Moreover, WSNs tend to be deployed in hostile and unattended environments, which allows adversaries to directly interact with the nodes, monitor the communications and move in the field without any restrictions. Consequently, the adversary might be able to retrieve sensitive information about the network itself and the events being monitored by simply observing the communications.

The contribution of this work is an analysis of the existing solutions to a specific privacy problem in WSNs, namely the location privacy problem. The goal of these solutions is to protect both the source nodes and the base station from being localized by an unauthorised party. The location of source nodes is important since it allows an adversary to estimate the location of the events being monitored. These events might be related either to individuals or to important assets which must be carefully protected. In contrast, an adversary able to find and compromise the base station acquires full control of the network. The criticality of the location privacy problem is reflected in current WSNs scenarios, which range from military to healthcare applications, but also in the scenarios that are to appear with the development of the Internet of Things. Furthermore, from the analysis of solutions we elaborate a complete taxonomy that is expected to guide in the design of new improved solutions by taking into consideration the advantages and disadvantages of the solutions to date.

The structure of this paper is as follows. In section 2 we overview previous anonymity solutions that were proposed in other types of networks, since they resemble some of the ideas proposed in the field of WSNs. Section 3 will introduce the problem of privacy preservation due to the open nature of the communications in WSNs, showing potential vectors of attack to the privacy of the network. Subsequently, in section 4 we will present and analyse the set of solutions that have been proposed so far to protect both source and receiver location privacy from different types of adversaries. Finally, section 5 will conclude the paper outlining some of the open challenges which, in our opinion, will steer privacy research in the future.

2 Previous Work

In this section we will review some of the most significant techniques that have been proposed in order to provide anonymity to the communications in different types of networks. This section is divided into solutions for wired and for wireless networks, respectively. Thus, it aims to provide a clearer view of the needs for the development of privacy-preserving next generation networks, consisting of a congeries of technologies, both wired and wireless. Moreover, some of the techniques presented in this section influenced the development of new solutions in the field of WSNs.

2.1 Anonymous Communications in Wired Networks

The work presented by Chaum [7] is the starting point in the development of solutions for anonymous communications. The main focus in that work is to provide users with an anonymous e-mail system based on a special type of device called the mix. The main functionality of a mix is to receive a cryptographically encrypted message and transform it into a new message indistinguishable from the originally input one. In order to send a message, the source creates several layers of encryption over the message using the public keys of the different mixes that the message will traverse. Furthermore, messages are temporarily stored and released after a period of time following a different pattern so that an attacker observing the mix cannot correlate inputs and outputs. The main drawback of this approach is that the induced delays might be considerably high in order to provide an adequate time decorrelation.

New approaches based on Chaum's work were later presented to protect general Internet traffic. Onion routing [28] and Tor [11] provide application independent anonymous connections in near real time by creating connections through a set of machines called the onion routers. Whenever an application establishes a connection, it first connects to an onion proxy, which is the entrance point to the anonymous network. The onion proxy is in charge of determining a series of onion routers that will define the bidirectional path that the packets of that specific connection will traverse. The path is constructed by using the cryptographic material of each of the onion routers, which is included in a data structure called the onion. Once the path has been established, the application data is sent through the onion network by adding a layer of encryption for each of the hops in the anonymous path. Each of the onion routers peels off its corresponding layer, changing the appearance of the data, and forwards it to the next onion router. One of the downsides of these approaches is that they are based on a network core which the users must fully trust.

Also decentralized approaches such as Crowds [29] and Hordes [18] were proposed. Both schemes are based on the idea of making individuals "disappear" into a group of peers. Upon receiving a message from a peer, the recipient will randomly choose whether to forward it to another peer or to finally submit it to the real destination. Each member of the path must remember its predecessor and successor so that subsequent messages coming from the same source follow the same path through the anonymous network. Note that any member of the path has only a local view of the route that a message traverses so that no peer can determine who is the actual origin of a message. Furthermore, since all communications are re-encrypted at every hop, a local eavesdropper cannot easily determine the destination of a message unless the originator decides to send the message directly to the destination. The main difference between Crowds and Hordes is in the way responses are sent back to the origin. In Hordes it is done by multicasting messages, which provides a better performance.

2.2 Anonymous Communications in Wireless Networks

Some work has also been done in the realm of ad hoc networks. In particular, most of the results in this domain focus on the protection of on-demand routing protocols for mobile ad hoc networks (MANETs), which require a route discovery phase before message transmission.

The first anonymous on-demand routing protocol in the literature is ANODR [17]. The route discovery process in ANODR is based on embedding cryptographic trapdoor information and an onion of pseudonyms into route request (RREQ) packets. At destination, upon the reception of the RREQ, the validity of the onion is checked and inserted, together with a proof of trapdoor opening, in a route replay (RREP) packet which is sent back to the source. If the onions are created by using trapdoor information, the participants do not need to include their identity pseudonyms in the onion, thus providing a better privacy level. Once the route is defined, data packets are locally broadcast without indicating either the sender or the next hop towards the destination. Only the nodes in the pre-established path are able to acknowledge the validity of received messages and they are the only nodes to forward the messages.

ASR [47] is basically an enhanced version of ANODR with regard to the cryptographic schemes used to create the RREQ and RREP packets during the route discovery process. The main difference between these two protocols is that ASR provides what the authors call strong location privacy. This concept refers to the inability of intermediate nodes to determine the distance, in terms of number of hops, to either the source or the destination by analysing the onion. ANODR is unable to provide this property even when it pads the onion to a fixed size in order to hide its actual size from external eavesdroppers. However, the proposed countermeasure is inefficient when the adversaries are en-route nodes because they need to make transformations on the onion, thus making them aware of its actual size. In any case, both ANODR and ASR are considered rather computationally intensive and sensitive to node mobility.

MASK [46] is founded on the concept of pairing, which is used in an authentication protocol to allow neighbouring nodes to mutually authenticate without the need to reveal their identities. Moreover, the anonymous authentication process allows the establishment of a pairwise secret key generator. The obtained keys are used by the nodes to determine if a received packet belongs to a specific connection and to perform cryptographic operations on the packets. The main limitation of this scheme is that it relies on a tight synchronization of the keys between neighbouring nodes. Moreover, during the route discovery phase, the identifier of the final destination node is contained in every RREQ packet in plaintext, thus exposing the recipient of the data.

The authors in [31] propose ARM to solve some of the limitations of previous approaches. With regard to efficiency, nodes receiving a message do not need to perform cryptographic operations to determine whether the message is directed to them because it is assumed that every node in the network shares a set of secret one-time pseudonyms with other nodes in the network. This is a strong assumption. Besides, this protocol implies some extra overheads in order to counter the threat of more powerful adversaries. Upon the reception of a RREQ packet at the destination node, it continues to forward the received packet after decreasing the time-to-live value contained in it, thus mimicking the behaviour of en-route nodes. Additionally, the destination node will generate the RREP after forwarding one or more RREQ messages. Although, this can be easily detected by a global adversary, he will not be able to determine the relationship between RREQs and RREPs. Moreover, during the data transmission process, nodes receiving a message with an unknown identifier will process the message as if they belong to the routing path.

3 The Privacy Problem in WSNs

Privacy in smart environments has traditionally been related to what is known as social privacy, which refers to the ability of collecting and analysing user data without explicit consent. However, the privacy problem in WSNs has been broadened to embrace network privacy aspects. In this scenario, an attacker might analyse the network operation in order to retrieve information about the network itself and the data being collected. In fact, there is no clear distinction between these two aspects of privacy because the events being monitored by the network might also be related to individuals. The main difference lies in the entity who aims to violate privacy. In the case of social privacy, the owner of the network is usually the privacy perpetrator because he might collect user data when the user interacts with the environment. In the network privacy case, the adversary is an outsider who takes advantage of a sensor network deployed for legitimate purposes in order to obtain information which was not intended for him.

Pai et al. [26] show how much information can be obtained from the network and the environment being monitored by simple observation of the network traffic.

(i) The frequency range might reveal the sensor platform being used. Recent technologies (e.g. micaz, IRIS, Imote2) can be easily distinguished from older ones (e.g. cricket, mica2) since the former perform in the 2.4-Gigahertz spectrum while the latter perform at sub-Gigahertz frequencies. Although this information might apparently be innocuous, it can be used to exploit platform specific vulnerabilities. In addition, carrier frequency can help to determine the owner of the network, since different frequency bands are assigned for different purposes and organizations.

(ii) The transmission rate at which messages are being delivered is a good indicator of the quantity and the nature of the events being monitored. The occurrence of events triggers the delivery of messages to the base station. Also, the non-occurrence of events might be an indicator of sensitive information. For example, in the case of a sensor network monitoring the vital signs of an individual, the lack of message transmissions may indicate that the patient suffered a cardiac arrest. On the other hand, a high transmission rate can also be used by an adversary to determine the location of the sender by using time of arrival techniques [6].

(iii) The size of the packets provides information about the type and precision of the data being collected. In particular, the use of some data aggregation protocols [27] might produce privacy breaches because the nodes receiving a message incorporate their own sensed data into the packet payload, thus increasing the size of the packets. This feature can help an adversary to infer the proximity to the base station.

(iv) The communication pattern might reveal the network topology. In order to extend battery lifetime, messages are usually transmitted in the shortest path between the source and the destination. Adversaries can take advantage of this knowledge to find out the location of important nodes in the network such as the base station or the sources of messages.

Another consideration about privacy in WSNs is made by Kamat et al. in [15]. They suggest that not only the occurrence of an event is important but that also the time at which this event takes place. This concept is named as temporal privacy. In mobile asset monitoring scenarios if an adversary is able to make an association between the time and position of the events being monitored, then he will be able to predict future behaviours. For example, in military scenarios, being in possession of such information can be a tremendous advantage in developing more effective plans of attack.

Consequently, a large amount of contextual information can be gathered by simply observing the messages being exchanged by the nodes during the network operation. This paper focuses on presenting the techniques so far developed to protect the location of nodes or the events being monitored by them.

4 Location Privacy in WSNs

In general, WSNs present well-defined communication patterns. The location privacy problem appears because immediately after event detection, the node starts transmitting messages to the base station using a fixed path. The path is usually chosen in such a way that the energy consumption is minimized. As a matter of fact, the path followed by messages exposes both the source node and the base station regardless of the use of cryptographic techniques to conceal the contents of messages. This gives rise to two different problems and the goal is to prevent an attacker from determining the location of these nodes.

- Source location privacy (SLP): aims to protect the location of the nodes reporting physical phenomena taking place in their vicinity. Note that the adversary is not usually interested in finding the device to tamper with it but to be able to determine the location of the events occurring nearby.
- Receiver location privacy (RLP): aims to protect the location of the final destination of messages, which is usually a single base station. The attacker may want to find the base station because it is the most important asset in the network; if the base station is compromised the whole network is under the control of the adversary.

The adversarial model which has been mostly considered in the literature is a passive, external attacker with either local or global eavesdropping capabilities. A passive attacker simply observers the messages being exchanged while an active attacker can also create or modify packets, disturb the routing protocol, destroy nodes, etc. Moreover, an external attacker is not part of the network, that is, no nodes are under the control of the adversary. In contrast, an internal attacker is in possession of some nodes which surreptitiously participate as legitimate network members. Therefore, internal attackers have access to more detailed information about the messages traversing the network and the protocols in use. Finally, an attacker is said to be either local or global according to his eavesdropping capabilities. Attackers with a local hearing range have similar capabilities to sensor nodes (mote-class adversaries) while global (laptop-class) adversaries have much stronger capabilities, with a monitoring radius equal or close to the entire network size.

In general, the proposed techniques to counter adversaries depend on their hearing range and recently their ability to compromise nodes. However, it is also necessary to protect the identity of the nodes sending messages. Packet headers contain relevant information about the communicating parties in order to route packets through the network. Therefore, the first step to provide location privacy is to conceal real node identities. In Figure 1 we provide a taxonomy of solutions that will guide the rest of this paper.

4.1 Node Identity Concealment

Despite of the use of cryptographic mechanisms to protect the contents of the packets, there is much relevant information contained in the packet headers but it is not usually protected. Sender and destination identifiers (IDs) are sent in clear in packet headers for routing purposes. An attacker eavesdropping network communications, after a certain amount of time, might

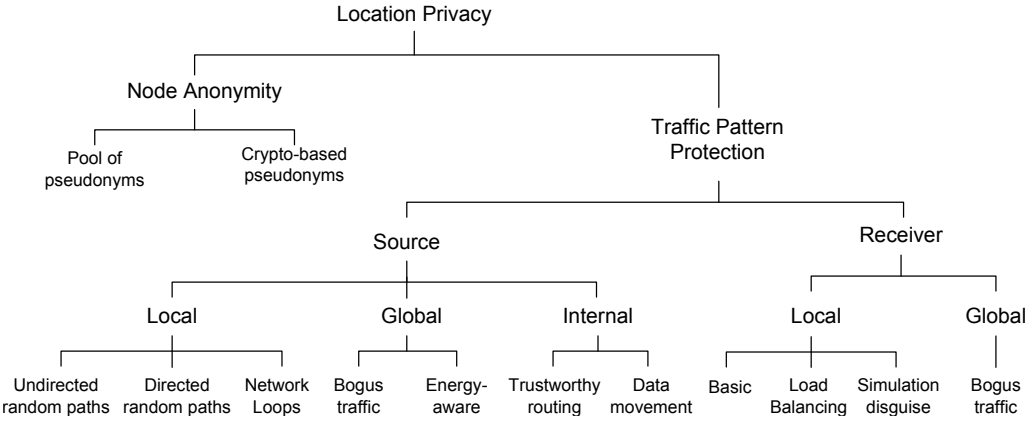


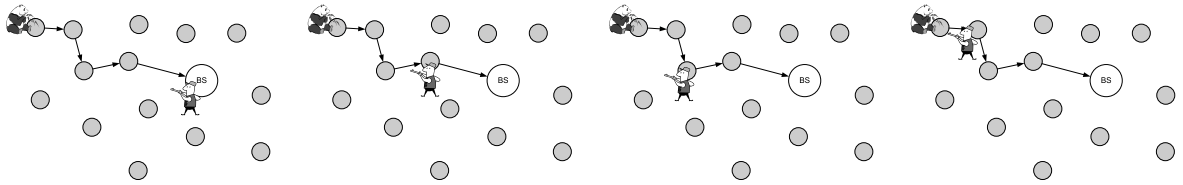
Figure 1: Taxonomy of location privacy techniques in WSNs

capture a sufficient number of packets to produce a network map. In such case, he simply needs to wait next to the base station for packets to arrive in order to retrieve the source ID and by using the network map translate it into a physical location, where the event has occurred.

Several techniques based on the use of pseudonyms have been proposed to provide node identity protection. A pseudonym is a name or identifier that can be used instead of a real name. Using fixed pseudonyms eventually provides no protection because the attacker is able to relate a pseudonym to a node. Therefore, these techniques are based on the use of dynamic pseudonyms, which are periodically changed. Misra et al. [22] developed two anonymous schemes. In the Simple Anonymity Scheme (SAS) each node is provided with a randomly distributed set of pseudonyms ranges from a network-wide pool of pseudonyms. The node will select a different pseudonym for each transaction. The major limitation of SAS is the amount of memory needed to store the a sufficiently large pseudonym space. The second approach, Cryptographic Anonymity Scheme (CAS), uses a keyed hash function to generate the pseudonyms. Neighbouring nodes share a fixed random number and a sequence number which are used as an input to the keyed hash function. This reduces the amount of storage at the expense of a greater computational cost.

Previous schemes assume an attacker cannot compromise the secrets shared by sensor nodes. Ouyang et al. [24] propose two methods based on keyed hash chains to reduce the impact of shared secrets being compromised. In Hashing-based ID Randomization (HIR), every sensor node shares a pairwise key with its neighbours for the computation of hash chains, which are used as pseudonyms. Messages are of the form $M = H_j||H_s||t||Data$, where H_j is a keyed hash chain used by the recipients to determine if M is directed to them, H_s is a keyed hash chain of depth t intended for the base station to find the real sender of the message, and $Data$ is the encrypted payload of the message. In the other method, Reverse HIR (RHIR), the hash chains are created during deployment and then used in reverse order. This provides better security because if the key is compromised in HIR at transaction i the attacker is able to compute any identifier ID_{i+k}. The downside of RHIR is greater memory requirements since the hash values are stored and not discarded until they are used. Other schemes based on similar ideas have been proposed in the literature. For example, in [19] hash chains are also used in reverse order as a complement to their privacy preservation scheme. The main difference with the previous approach in that the network is divided into cells and they propose to use only part of the hash function output as the ID to preserve energy.

Although node identity protection is necessary to preserve location privacy, a more skilled attacker can still perform traffic analysis attacks to determine the location of nodes of interest



(a) Initial state (b) Jump towards received packet (c) Distance is reduced (d) One step away from target

Figure 2: The Panda Hunter Game

to him. In the following sections we review some of the attacks and countermeasures which have been developed against these adversaries depending on their power.

4.2 Source Location Privacy

Source location privacy refers to the ability of protecting the location of the events being reported by sensor nodes. This problem was first analysed by Ozturk et al. [25] where the Panda Hunter Game was proposed. This game considers a large sensor network deployed to enable biologists to monitor the behaviour of pandas in their environment. In such scenario, a hunter tries to take advantage of the messages being reported by the network to hunt pandas.

Different techniques have been developed in the literature to counter such adversaries. These techniques vary depending on the power of the adversary.

4.2.1 Local Adversaries

Local adversaries are capable of monitoring only a small portion of the network, typically the equivalent to an ordinary node. Moreover, they are equipped with a device capable of determining the direction of the received packets. Thus, when an adversary overhears a packet he moves towards the node which sent it. This process is repeated at the next node until the adversary reaches the true source of event data (see Figure 2). Therefore, the goal is to mislead the adversary in order to increase the safety period, that is, the number of packets sent by the source before the attacker reaches the source node.

Clearly, a local adversary is able to find the source because the packets always follow the same path. Therefore, current proposals are mainly based on traffic pattern randomization. However, since data packets no longer follow the shortest path, the probability of packet loss is increased as well as the delivery time and, more importantly, more energy is consumed. The following techniques make a trade-off between network performance and privacy protection level.

Undirected Random Paths Phantom Flooding [25] resulted from the analysis of two widely used families of routing protocols in WSNs, flooding-based and single-path routing protocols. Surprisingly, both provide the same privacy protection level because the shortest path is always contained in a flooding, however, baseline flooding implies an energy consumption significantly higher. A probabilistic flooding forwards packets following a probability distribution, which reduces the energy consumption but also reduces the probability of an attacker reaching the source. Therefore, the authors propose to make every packet experience two phases, a walking phase and a flooding phase. In the walking phase, the packet travels h hops in a random walk until a phantom source. Then, in the next phase, a baseline or probabilistic flooding is initiated by the phantom source, which finally delivers the packet to the base station. Thus, every

(a) One random path

(b) Next random path

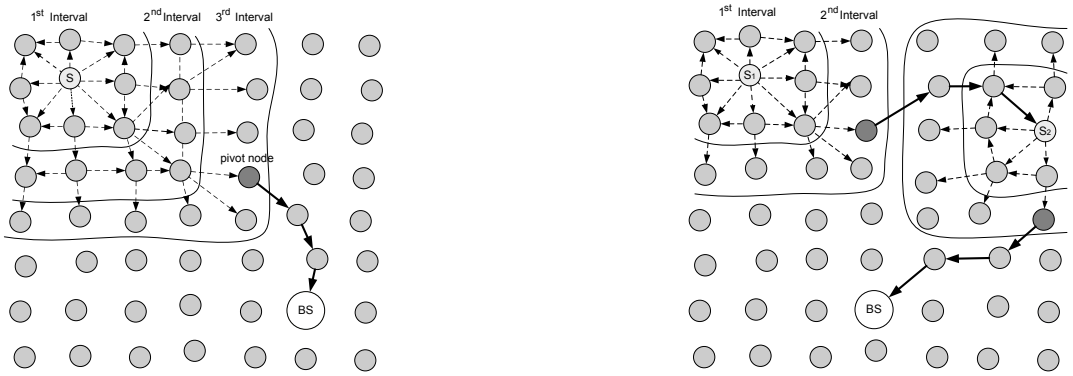
Figure 3: Phantom Routing Single-Path with $h = 4$

packet will traverse a new random route to the destination thus reducing the privacy risk. Also a Phantom Routing protocol, where the phantom source initiates a single-path routing phase, is presented (see Figure 3). The main limitations of this approach is the great amount energy consumed in the flooding phase. Also, the walking phase, which is dependent on h, introduces significant overhead. Furthermore, the walking phase must be carefully designed in order to avoid having phantom sources close to each other or to the real source.

To reduce the concerns about random walks staying close to each other or the real source, a two-way random walk is presented in [43]. Greedy Random Walk (GROW) first creates a static random walk (path of receptors) from the base station that is in charge of intercepting the packets originated from the source nodes. Subsequently, data packets are sent on a greedy random walk that will eventually arrive at a receptor node, from which the data will be forwarded to the base station following the established path. It is a greedy algorithm because the packets attempt to expand as far as possible by choosing the next hop from the set of neighbours who have not yet participated in the random walk. Also, the reason for using two intersecting random walks is due to the fact that the time to reach the destination with a single random walk goes to infinite while the probability of two random walks not intersecting decreases exponentially with time [33]. Despite these properties, the main drawback of this work is the significant increase in the delivery time of the packets.

Wang et. al [40] propose Random Parallel routing, which assigns every sensor node n parallel routing paths to the base station. Messages are evenly distributed to different paths so that the adversary traceback time is the same at any path. Also, the paths are sufficiently separated so that when the adversary is placed in one of the paths he is unable to overhear messages coming from other paths. Consequently, the adversary is forced to stay in a single path, which significantly increases the safety period. However, this approach presents several disadvantages in terms of computation, storage and privacy protection. In practice, calculating and using the n parallel paths is a complex task in large sensor networks, where the topology is quite dynamic due to instability of wireless links. Also, sensor nodes must locally store a large number of routing paths which requires significant memory consumption. Finally, since the paths are parallel to each other, retrieving several packets on any path provides a good estimate of the direction to the source. This feature might help the adversary to infer the location of the source.

A cross-layer approach is presented by Shao et al. in [35]. This approach is similar to Phantom Routing but it hides the walking phase in the MAC layer to prevent the adversary from overhearing messages in the vicinity of the source node. Event data is cryptographically hidden within beacon frames, which are control messages widely used in WSNs for network configuration purposes. Since beacons are periodically broadcast regardless of the occurrence of real events, the attacker is unable to distinguish legitimate beacons from those containing



(a) Single cross-layer approach

(b) Dual cross-layer approach

Figure 4: Cross-layer schemes for source location privacy

event data. At the end of the walking phase, data reaches a pivot node where the information is extracted and normally routed to the base station. This process is shown in Figure 4a, where the dotted arrows are beacon frames and solid arrows represent messages forwarded in the routing layer. To further mislead the adversary event data can move between layers several times on its way to the base station (see Figure 4b). The use of a cross-layer approach improves the privacy protection level because the adversary cannot capture packets near the source. The main objection is that beacons are periodically broadcast at intervals ranging from milliseconds to several hundreds of seconds which introduces a large delay in the delivery of messages.

Directed Random Paths Instead of simply sending packets at random, some authors proposed the use of directed walking phases. By having a walking phase guided by certain parameters, either the packet delivery time is reduced or the privacy protection level is increased, or both. The first solution to include a directed walking phase is Phantom Routing [25, 16]. In Phantom Routing every node separates its neighbours into two groups depending on whether they are in the direction or in the opposite direction to the base station. Thus, during the walking phase, the next hop in the path is randomly selected only from the set of nodes in the direction to the base station. In this way, they countered the problem of phantom sources staying in the vicinity of the source. Next, in the routing phase, the phantom source forwards the packet to the base station using a single-path routing protocol. A single-path phantom routing besides significantly reducing the energy consumption, it provides a higher privacy protection level since the resulting single paths originated from the phantom sources will evade the hearing range of the adversary more easily than a flooding-based strategy.

Later, the authors in [42] observed that originating larger random walks does not necessarily increase the protection level of the sources because phantom sources are not always placed in a secure location to initiate the routing phase. In general, an adversary placed in the shortest path from the base station to the source is more likely to find the source if the angle of arrival of the messages is usually less pronounced (see Figure 5). Therefore, the authors proposed the Phantom Routing with Location Angle (PRLA), which introduces inclination angles to direct random walks, which prioritizes the selection of phantom sources leading to a larger angle of arrival. In terms of efficiency PRLA improves the safety period with respect to Phantom Routing by simply directing random walks based on the angle of arrival. Therefore, it would be possible to reduce the number of hops performed during the walking phase in PRLA while keeping an adequate privacy level, which is translated into energy efficiency.

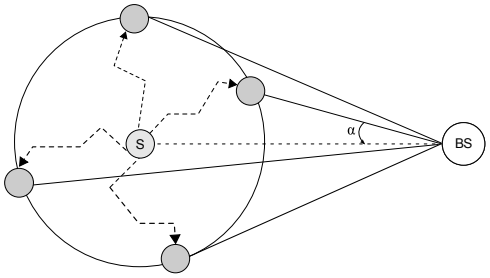
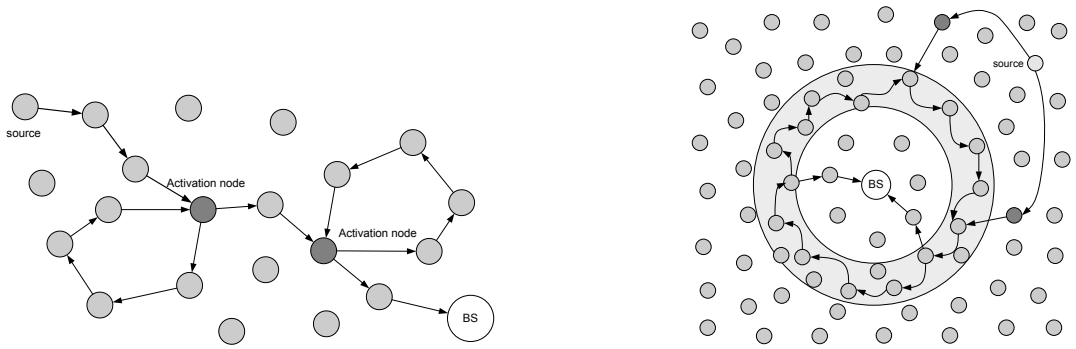


Figure 5: Path inclination w.r.t shortest path

In addition to proposing the Random Parallel routing (Section 4.2.1), Wang et. al [40] developed the Weighted Random Stride (WRS) routing. In this scheme every sensor is allowed to make its own routing decisions based on its limited network topology knowledge. Whenever a node transmits a message to the base station it does it based on two parameters, a forwarding angle and a stride. The node selects an angle and chooses a neighbour that matches that angle. The message is forwarded to the chosen node, which subsequently forwards the message to one of its neighbours that matches the predefined angle. This process is repeated for the number of hops defined by the stride. Once the stride is finished, the receiving node starts a new stride. In practice, instead of using forwarding angles, the coverage radio of the nodes are divided into sectors and the next hop is randomly selected from one of these sectors. In order to produce longer routing paths, and thus reducing the chances of the adversary, larger forwarding angles are also prioritized in this scheme. However, this also increments energy needs.

The authors in [20] propose the Random Intermediate Node (RRIN) scheme, where data packets are sent to a random node that will eventually send the packet to the base station. Intermediate nodes are chosen based on their relative location and taking into consideration that they are placed at least at a distance d_{min} from the source node. Since the RRIN scheme assumes that nodes only have knowledge of immediate neighbours, the use of relative locations guarantee that the packets will reach the closest node to the location calculated by source and from there the packets will be forwarded to the base station. The main advantage of this approach is that intermediate nodes are randomly chosen from those at least at a distance d_{min} from the source, which is one of the limitations of traditional random walks. Besides that the length of the paths tend to be too long, RRIN has other drawbacks. The size of the network clearly influences the probability of a node becoming an intermediate node, hence it provides the adversary with additional information about the location of the source. Also, no mechanism prevents intermediate nodes from being selected from the proximities of the source-destination shortest path, which was one of the problems addressed by PRLA and WRS.

Network Loop Methods Some approaches based on the creation of loops of messages have been developed to mislead local adversaries. In the Cyclic Entrapment Method (CEM) [23] several traps are set to the adversary on his way to the source. These traps are in the form of decoy messages which are sent in a circular fashion to keep the adversary away from the source. After deployment nodes decide whether to generate a network loop based on some probability. A loop is triggered when a real data packet is received at any of the loop activation nodes. When tracing back the path to the source node, local adversaries will at some time reach the loop, where the path forks in several directions (see Figure 6a). At this point, the adversary must decide in which direction to move. In case of making a wrong decision, he is trapped in the loop until he discovers. The safety period is increased several orders of magnitude with the number of active loops. However, a more skilled adversary might avoid loops depending on the



(a) Cyclic Entrapment Method (b) Network Mixing Ring

Figure 6: Network Loops Methods

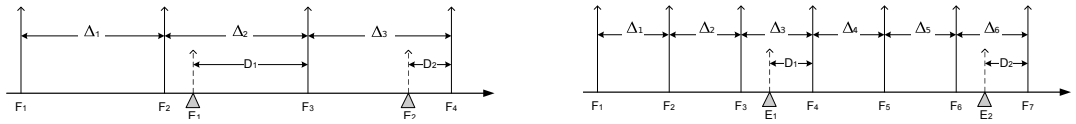
angle of arrival of the different messages reaching the activation node since messages with a larger inclination angle are more likely to lead to a loop.

The authors in [19] propose a two-phase routing scheme. In the first phase, for every new message, the source selects a RRIN that will initiate the next phase. The strategy used for the selection of the RRIN takes into consideration a minimum distance to the source node as presented in Section 4.2.1. In the second phase, the intermediate node sends the packet to the closest node in the Network Mixing Ring (NMR). The NMR is a virtual ring of nodes surrounding the base station whose aim is not to trap the adversary but to mix up real messages making them undistinguishable to the adversary. Packets in the mixing ring are relayed in a clockwise direction (see Figure 6b) and change their appearance at every hop to thwart traffic analysis. After a random number of hops in the ring, the packets are finally forwarded to the base station. Since the attacker is unable to figure out the initiator of a message from the nodes forwarding it within the ring, he is also unable to determine the location of the real source of the messages. The main drawback of this approach is that energy consumption is quite unbalanced depending on the functionality of the nodes. Ring nodes are more likely to deplete their batteries than other nodes and since ring nodes are surrounding the base station it will eventually become isolated from nodes outside the ring.

4.2.2 Global Adversaries

Previous techniques are not effective against adversaries with a larger hearing range. More powerful adversaries are able to monitor larger areas and therefore obtain a better picture of the path followed by the messages. In particular, global adversaries are capable of monitoring all the traffic generated by the WSN. This type of adversary is not necessarily a single attacker equipped with a powerful antenna. In fact, several colluding attackers wisely deployed in the field might achieve an equally effective monitoring range. Such adversaries can easily detect the source of event messages among mere intermediaries because sensor nodes are programmed to immediately report event data as soon as it is detected.

In order to deal with such powerful adversaries, the general solution is to hide event messages within fake message transmissions. Note that real and fake messages must be indistinguishable from the point of view of the adversary. Consequently, both types of messages must have, on average, the same length and be encrypted with a shared secret key, which allows the next hop to authenticate the message transmission while message injection is avoided. Moreover, the use of dummy traffic implies a significant waste of energy, which reduces the operational life of the



(a) Perfect event unboservability

(b) Reducing message latency

Figure 7: Perfect event source unobservability

sensor nodes. The aforementioned issues are considered in the solutions analysed next.

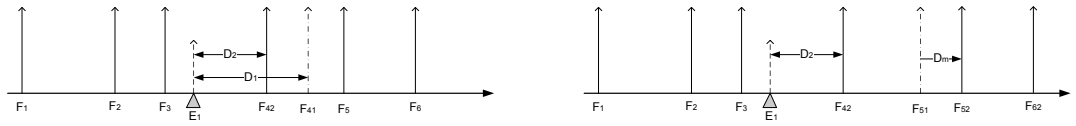
Dummy Traffic Injection The idea of introducing fake messages was first used in [25] with the Short-lived and Persistent Fake Source strategies. In both strategies some nodes simulate the presence of real events in their vicinity to mislead the adversary. In the former, upon the reception of a message the node decides to produce a fake packet and flood the network with it based on a constant probability. In the latter, every node decides whether to become a fake source based on a certain probability. However, these strategies are still inefficient against a global adversary. The Short-lived scheme cannot mislead global adversaries because fake packets are only produced after the reception of a real one. Therefore the adversary can easily detect the real sources. In the Persistent scheme, the attacker can detect real sources because fake sources are not constantly sending messages. In any case, these approaches were devised to deal with local adversaries only.

The threat of global adversaries is first considered in [21]. The authors propose the Periodic Collection scheme, where every sensor node sends messages at a fixed rate, Δ, regardless of the existence of event data to report. Upon the occurrence of a real event, the message is temporarily stored until the next scheduled transmission time. Since real and bogus traffic are indistinguishable and sent at a fixed rate, this method provides perfect event source unobservability (see Figure 7a). However, there is an important trade-off between energy consumption and packet delivery time. The inter-transmission times must be carefully adjusted since a small Δ value implies a large amount of extra traffic (see Figure 7b), while a large Δ introduces significant delays to the data packets sent to the base station.

Energy-Aware Approaches The Periodic Collection approach presented in [21] was too energy consuming for sensor nodes. Thus, several solutions were devised to deal with this problem. The goal was to provide an adequate privacy protection level without incurring an excessive delay in nodes transmissions and also without exhausting the battery of the nodes.

Mehta et al. [21] proposed a second technique which aimed to reduce energy consumption by decreasing the number of nodes transmitting fake messages. This scheme is named Source Simulation. The idea is to simulate the presence of the asset being monitored in different areas of the WSN. The problem lies in the difficulty of accurately modelling the movement of an object to appear as real to the adversary. In the case of moving objects, such as animals or vehicles, having a static subset of sensors constantly sending messages can be easily interpreted by a global eavesdropper as a mechanism to deceive him. Therefore, sensor nodes must be carefully programmed to transmit fake messages following a coherent pattern that resembles a real object. Moreover, this process should be carefully tailored for any type of asset being monitored.

In order to reduce network traffic while maintaining source unobservability, Yang et al. [44] proposed a filtering scheme based on the use of proxy nodes. In this approach, named the Proxy-based Filtering Scheme (PFS), sensor nodes produce real or dummy traffic, at a probabilistic



(a) Statistically strong unobservability

(b) Message delay for mean adjustment

Figure 8: Statistically strong unobservability

rate, depending on the presence of events. Upon receiving a dummy packet at one of the proxy nodes, it simply drops the packet in order to reduce network traffic. However, upon receiving a real packet, it will be re-encrypted and temporarily stored for later forwarding. Note that if the proxy has no real traffic to forward when its timer expires it must transmit bogus messages, otherwise the message distribution would change and the attacker could easily detect the areas of the network reporting real events. To further reduce dummy traffic a multi-layered proxy architecture is proposed. In the Tree-based Filtering Scheme (TFS), real or bogus packets might traverse several proxy nodes. Thus network traffic is reduced at the expense of an increased network delay due to buffering at each layer. Both PFS and TFS provide a good level of privacy to the source nodes, however, the attacker can still use monitoring rate techniques to breach privacy in a different way, by identifying the proxy nodes, which are important for the operation of the network.

In addition to the cross-layer scheme described in Section 4.2.1, Shao et al. [35] proposed a simpler version which can be used to protect from global adversaries while preserving nodes batteries. In this version, instead of using a cross-layer approach, the data are conveyed to the base station using only beacon frames in the MAC layer. Since beacons are periodically broadcast regardless of the occurrence of real events, this approach provides perfect event source unobservability at no additional cost. However, since the time between consecutive beacons is relatively large, the solution is only practical for some applications where no tight time restrictions exist. Furthermore, the delivery time is highly dependent on the distance from the source to the base station.

Also, some statistical approaches were devised in order to reduce message delivery time while preserving batteries lifetime. Shao et al. [36] introduce the concept of Statistically Strong Source Unobservability. The idea is depicted in Figure 8a. Nodes transmit fake messages according to a probability distribution (F_i) in such a way that upon the occurrence of a real event (E_i) it is transmitted in the shortest time possible (F_{42}), before the next scheduled fake transmission (F_{41}) and without altering the overall probability message distribution. To achieve this, every node keeps a sliding window of previous inter-transmission times $\{\Delta_1, \Delta_2, ..., \Delta_{n-1}\}$. When a real event occurs, Δ_n is set to a value very close to 0 and gradually incremented until the whole sliding window passes a statistical (goodness of fit) test. Thus, the real event transmission can be sent ahead of the scheduled time without alerting the adversary even if he performs statistical tests on inter-transmission times. Besides, as real messages are re-scheduled, the presence of bursts of events might skew the mean of the distribution. To overcome this problem, a mean recovery mechanism is introduced, which delays subsequent transmissions (see Figure 8b).

Recently, Alomair et al. [2] showed that a global adversary can perform in a more efficient way. Instead of focusing on a single sliding window, the attacker might attempt to spot differences between any two sliding windows in order to detect the presence of real events. Therefore, it is necessary to provide the property of Interval Indistinguishability. The authors note that in the presence of intervals with real events, short inter-transmission times followed by long inter-transmission times are more likely to happen due to the mean recovery mechanism proposed by

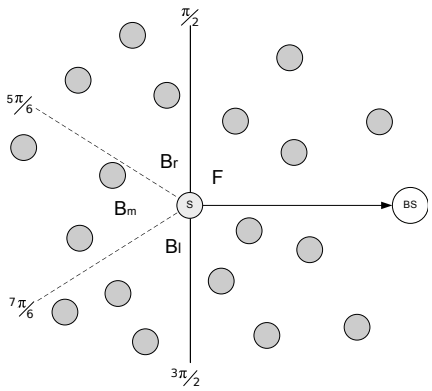


Figure 9: Neighbouring nodes division in IRL

Shao et al. Consequently, by counting the number of short-long inter-delays an attacker might be able to distinguish intervals containing real events. In order to reduce the chances of the attacker, the proposed solution is to make fake intervals resemble intervals with real events by introducing some statistical interdependence between fake inter-delays.

4.2.3 Internal Adversaries

More skilled adversaries might also compromise and control a subset of nodes in such a way that they keep on participating in the packet forwarding process and other network tasks. These internal adversaries provide the attacker information about the packets in transit and also about the protocols being used by the network. In the following we present the only works that, to the best of our knowledge, consider the threat of internal adversaries.

The authors in [32] propose the Identity, Route and Location privacy (IRL) algorithm. In fact, this work was not focused on the protection against internal adversaries but since it introduces the notion of trust and reputation in the routing process, it seems suitable for such purpose. In IRL, every node classifies its neighbours in four groups depending on their position with respect to the base station: forward (F), right backward (Br), left backward (Bl) and middle backward (Bm); as shown in Figure 9. Subsequently, neighbours are also classified into either trustworthy or untrustworthy based on the number of successfully forwarded packets. Nonetheless, the calculation of the trust values could be extended to incorporate new parameters, such as the presence of communications with external entities or with other non-neighbouring nodes in order to identify internal adversaries. Finally, when a node needs to send a message to the base station, it checks whether there are any trustworthy nodes to select in the direction to the base station. Among all trustworthy nodes one is randomly selected. If there are no trustworthy nodes, the same process is repeated for Br and Bl. As a last resort, the node tries to send the packet in the opposite direction to the base station. In the case no trustworthy nodes are found, the node simply drops the packet. Therefore, every message follows a different (random) path comprised of trustworthy nodes only.

Two additional mechanisms are proposed in IRL to further protect source nodes. In the first case, the ID of received packets are replaced at each hop with the ID of the node receiving the packet before it is forwarded. Thus, nodes are unable to determine if the previous node is the real source or a mere intermediary. The second mechanism consists of hiding the real source of the packet within the payload, which is also encrypted with a secret key shared between the source node and the base station. However, using end-to-end encryption protects the data from internal adversaries but it might also impede the use of data-aggregation mechanisms which is a common strategy in WSNs to reduce traffic and preserve energy. Some works [12, 45] have

been proposed to protect the data payload while preserving normal operation of the network.

In contrast, Shao et al. [37] present *p*DCS which aims to provide security and privacy to Data-Centric Sensor (DCS) networks with an itinerant base station. In DCS networks, event data are labelled depending on type of event and the label determines where the data are stored. Therefore, there are sensing nodes and storing nodes, which allows for a more practical and energy-efficient way of querying for a specific type of data. *p*DCS aims to prevent adversaries from compromising sensor nodes and obtaining the event data contained in them. The strategy used by *p*DCS is the following. Firstly, if the adversary compromises a storing node he will not able to decrypt the data contained in the sensor because these data are encrypted with the key of the sensing nodes which reported them. Secondly, if the compromised node is a sensing node, the attacker is unable to determine where is stored the data sensed in previous intervals because *p*DCS uses a secure mapping scheme based on keyed hash functions to select the storing nodes. Moreover, when a node is found to be compromised there is a node revocation mechanism in order to prevent the attacker from obtaining the location of future event data. Finally, the attacker could still perform traffic analysis attacks during the delivery of sensed data to find the location of the source of events. The authors do not propose new protection measures but suggest the use of some of the already existing location privacy techniques.

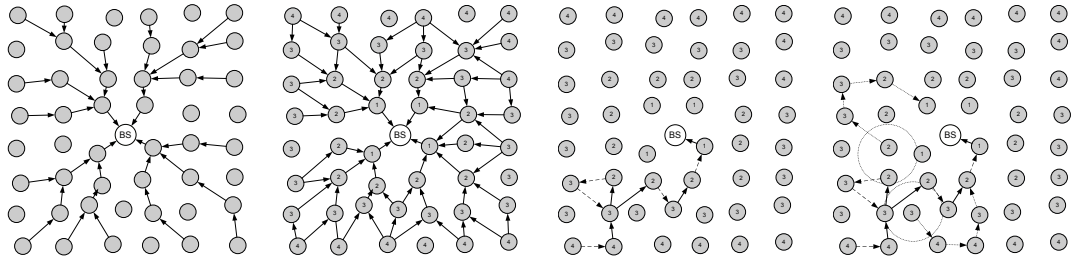
4.3 Receiver Location Privacy

Receiver location privacy is usually devoted to protecting the location of the base station. The base station is the most important asset in the network because it is responsible for processing and analysing all the information collected by the sensor nodes. Additionally, it serves as an interface between the user and the monitored field, allowing the user to access or send commands to sensor nodes. Thus, an adversary aware of the location of the base station can compromise it, or even destroy it, rendering the WSN useless.

4.3.1 Local Adversaries

In a local adversarial model, the attacker usually starts at a random position in the field and analyses the traffic in a limited area surrounding him. A local adversary can perform three basic traffic analysis attacks [8]. Firstly, in the content analysis attack, the adversary looks for valuable data that might lead him to the base station in either the packets headers or the payload. A similar attack based on the introduction of undetectable marks to data packets that allows the adversary to track them has also been proposed in [34]. Secondly, in the time correlation attack, the adversary observes the relation between the time at which a node and its neighbours send packets. Finally, in the rate monitoring attack, the strategy of the adversary is to move in the direction of the nodes with higher forwarding rates since nodes in the vicinity of the base station receive more packets than remote nodes.

Basic Countermeasures In order to prevent previous traffic analysis attacks, some basic countermeasures have been proposed. Encryption can be used to protect against content analysis attacks and packet marking attacks. However, an attacker can still differentiate between incoming and outgoing packets in a node, which allows him to determine the direction to the base station. Therefore, it is imperative to provide input-output indistinguishability. This can be achieved by encryption and padding at each hop [11]. However, even if the attacker cannot distinguish between different packets he can force the generation of packets by interacting with the nodes while he monitors the sending times between neighbouring nodes. These type of attacks can be prevented by means of buffering techniques, which introduce random delays in the packet forwarding process at every node. The rate monitoring attack is due to the nature



(a) Single-path routing (b) MPR (c) MPR + RW (d) MPR + RW + FP

Figure 10: Comparison of Load Balancing Techniques

of WSNs communications which present a specific pattern, that is, the source nodes send information to a single base station in relatively fixed paths (see Figure 10a). Intuitively, the solution to protect the base station from being localized consists of balancing the traffic load of the network. The straightforward solution is to make nodes transmit at a constant rate even if they have nothing to send [8]. Also, flooding-based protocols can provide an adequate protection because in such protocols every node that receives a packet broadcasts it, however, these techniques reduce the batteries lifetime.

Traffic Load Balancing Previous techniques can be further improved with the introduction of a set of load balancing techniques. Deng et al. [9, 10] proposed several techniques which can be used in conjunction. Firstly, in Multi-Parent Routing (MPR) every (child) node forwards data packets to a random (parent) node closer to the base station, thus balancing the amount of traffic among the various parents (see Figure 10b). Secondly, a Random Walk (RW) is added to MPR. In this approach, every node decides with a certain probability p to send a packet to a random parent or to start a random walk with probability $1 - p$. This introduces random routes in any direction, not necessarily in the direction to the base station, at the cost of a higher message delivery delay. In Figure 10c we show a trace of MPR and RW, where dashed arrows represent a RW choice and general arrows represent an MPR choice. Finally, Fractal Propagation (FP) is added to MPR and RW. In this scheme, when a node overhears a real packet, it creates and forwards a fake message with a certain probability. The performance of this scheme is depicted in Figure 10d, where dotted circles and arrows indicate the creation of a dummy random path. Since nodes close to the base station overhear more packets, the amount of fake packets generated increases significantly in that region. To address this problem, a Differential Fractal Propagation scheme (DFP) is proposed. In DFP, nodes adjust their probability of generating dummy traffic depending on the number of packets they overhear. This scheme not only reduces the amount of energy wasted by the sensors but also provides a better protection to the base station because an attacker knows that the traffic load next to the base station is always higher.

Additionally, Jian et al. [13, 14] claim that even if fake messages are introduced in different directions, these can be easily discarded because of an abrupt change in their trajectory. The authors assume that the attacker can determine the direction of both incoming and outgoing packets from a node, which requires either a large hearing range or that packet headers reveal routing information. To counter such adversary, the authors propose the Location Privacy Routing (LPR), in which every node decides with a certain probability P_f whether to forward a received packet to a parent node or to a child node. The value of P_f must be carefully tuned to balance the privacy level and the energy consumption, delivery time, and delivery ratio. In any case, the value of P_f should be biased in such a way that the traffic eventually reaches the

base station. Therefore, the traffic trend must point to the base station, however, this gives the attacker the opportunity to find it. To overcome this problem fake messages are injected in the opposite direction of the base station.

Simulation and Disguise Some approaches try to simulate the presence of the base station on different locations of the network to mislead the adversary or to disguise it among ordinary nodes. In [9, 10], the Differential Enforced Fractal Propagation (DEFP) is presented to further reduce the chances of the adversary. DEFP generate network areas with a high volume of fake traffic also known as hot spots. The underlying idea is to lead the adversary to a hot spot making him believe he is being directed to the base station. Also Biswas et al. [5] present a similar approach based on the simulation of several data sinks in the network. They propose to evenly distribute multiple fake base stations and to create fake traffic directed to them. However, the creation of network areas with a high traffic load in a coordinated and energy efficient way is a very challenging task due to the nature of these networks, which are highly distributed and energy constrained. Moreover, even when fake traffic is injected in a controlled way, it introduces a significant cost in terms of energy consumption.

Also, as the traffic rate of the base station is usually null in monitoring-based sensor networks, the attacker can perform rate monitoring attacks to determine its location. In order to overcome this problem, the base station might mimic the behaviour of sensor nodes, that is, forwarding received packets several hops away. This approach could be useful if the nodes are not visible to the adversary (e.g.: underground or underwater networks).Other approaches to base station concealment are proposed in [1]. In the first approach, the base station transmits some of the packets it receives with varying intensity so that the adversary believes these packets belong to ordinary sensor nodes. The second approach relocates the base station in a different area of the network. However, relocation has to deal with at least three issues, determine the best moment in time to move, selecting the next location and finally moving the base station in an energy efficient way.

4.3.2 Global Adversaries

The above techniques are considered to be effective only in a local adversarial model. However, some of these techniques could also provide some means of protection against global adversaries. In fact, they might be useful if the global adversary has no real-time analysing capabilities, that is, he is only able to retrieve a snapshot of the amount of traffic transmitted by every node during a period of time. This is the case if the adversary deploys several devices in the field to monitor node transmissions and returns occasionally to gather the stored information in a map of the network (similar to Figure 10). However, previous techniques cannot protect from an adversary with real-time monitoring capabilities. Similarly to source node protection, it is necessary to introduce dummy traffic in order to hide the location of the base station.

Buffering and Fake Transmissions The first step is to control the transmission rate of the nodes as proposed in [9, 10]. Since nodes in the vicinity of the base station have a higher forwarding rate than remote nodes, using buffering techniques can prevent a global adversary from inferring the location of the sink. However, this must be accompanied by the transmission of fake messages at the same rate if there are no real messages to forward. Furthermore, in a scenario where the attacker has no real-time analysing capabilities, the injection of fake messages in the opposite direction of the receiver [13, 14] and the different versions of the Fractal Propagation technique [9, 10] might also provide an adequate protection level. Clearly,

	Privacy	Adversary	Main technique	References
Node Identity	Source	Any	Pool of pseudonyms	[22]
			Cryptographic schemes	[22, 24, 19]
Traffic Pattern	Source	Local	Undirected random paths	[25, 43, 40, 35]
			Directed random paths	[25, 16, 42, 40, 20]
			Network loops	[23, 19]
		Global	Bogus traffic	[21]
			Energy-aware	[21, 44, 35, 36, 2]
		Internal	Trust-based	[32]
			Data movement	[37]
	Receiver	Local	Basic	[8]
			Load balancing	[9, 10, 13, 14]
			Simulation and disguise	[9, 10, 5, 1]
		Global	Bogus traffic and buffering	[9, 10, 13, 14]

Table 1: Summary of Location Privacy Solutions in WSNs

all these techniques increase the communication overhead in terms of delivery time and energy consumption.

5 Conclusions and Future Research

In this paper we have presented and analysed a wide range of solutions for the provision of location privacy in WSNs (see a summary in Table 1). From this analysis we have proposed a complete taxonomy based on the power of the adversary as well as the main techniques used by each solution. This taxonomy considers the protection of both source nodes and the base station. Source location privacy must be properly preserved because an attacker might obtain the physical location of the events occurring in the field. This information is particularly sensitive when the events are related to military applications, homeland security and healthcare, to name a few. On the other hand, receiver location privacy is also of vital importance because the base station is the most important asset for the operation of the network. If the base station is compromised or destroyed, the whole sensor network stops functioning.

Location privacy in sensor networks poses new challenges due to the limitations of sensor nodes regarding computational power, limited communication range, and, more importantly, a modest and non-rechargeable battery supply. Therefore, privacy preservation techniques must trade-off between an appropriate privacy protection level and the cost of applying countermeasures against traffic analysis attacks. Obviously, the more powerful the attacker is, the more resources the network must spend in order to cope with the threat of a privacy breach. In general, the solutions to protect, either source or receiver location privacy, against a local adversary are based on routing protocols which randomize and balance network traffic. In contrast, to counter global adversaries, the most commonly used technique is the injection of fake message transmissions but it is very costly. We have also considered another type of attacker called the internal adversary, which are compromised sensor nodes under the control of the attacker. Currently, there is no clear approach as to which is the most effective technique to defeat them because there are very few works to consider such adversarial model. The use of reputation and trust between neighbouring nodes could be an interesting way to approach the problem. Nevertheless, we expect that over the course of the next few years new works are done in the area internal attackers.

Moreover, this paper only considers one of the many possible threats to privacy in WSNs. Note that privacy problems exist at different levels. In this paper, we investigated the privacy

problem at the communications level. However, in future scenarios it is also necessary to incorporate privacy preservation mechanisms at higher levels. Consider the case where a user moves in a smart environment, the services being provided will probably adapt or move with him. In such case, an observer could manage to infer the type of services being accessed and relate them to a user regardless of using privacy-preserving techniques at the communication level. Therefore, in order to tackle the privacy problem it must be considered in a holistic way, from the communication to the application layer.

Finally, we consider that the integration of WSNs with other technologies through the Internet, with the advent of the Internet of Things, will result in one of the most promising areas of study. In such domain, it is possible that not only the base station will have a direct connection to the outside world, but each sensor node could be directly connected to the Internet with a public IP address. To avoid the threat of adversaries, these nodes could benefit from having an Internet connection to route their packets directly. On the contrary, if not all the nodes have sufficient capacity to be directly connected to the Internet, they could route their packets through more powerful nodes, either to the base station or to another sensor with direct connection to the outside. Additionally, a user could request information from certain sensor nodes directly, without passing through the base station. This process not only involves the location privacy of the queried node, but also of the user requesting the information from a remote location via the Internet. Many new scenarios may emerge from the new network topology, which might fuse with other networks. Similarly, new types of adversaries might appear. Therefore, we believe that the integration of sensor networks with the Internet will result in a prolific area of study.

Acknowledgments

This work has been partially funded by Spanish Ministry of Science and Innovation through the ARES (CSD2007-00004), SPRINT (TIN2009-09237), and IOT-SEC (ACI2009-0949) projects. The SPRINT project is co-financed by FEDER (European Regional Development Fund) and the IoT-SEC project is under the ACI-COLABORA scheme.

References

[1] U. Acharya and M. Younis. Increasing base-station anonymity in wireless sensor networks. *Ad Hoc Networks*, 8(8):791–809, 2010.

[2] B. Alomair, A. Clark, J. Cuellar, and R. Poovendran. Statistical Framework for Source Anonymity in Sensor Network. Technical Report 003, Network Security Lab (NSL), 2009.

[3] D. Banisar and S. Davies. Privacy and Human Rights - An International Survey of Privacy Laws and Practice. online, 2010.

[4] A. Bassi and G. Horn. Internet of Things in 2020 - A roadmap for the future. Technical report, Information Society and Media Directorate-general of the European Commission (DG INFSO) and the European Technology Platform on Smart Systems Integration (EPSoSS), 2008.

[5] S. Biswas, S. Mukherjee, and K. Mukhopadhyaya. A Countermeasure against Traffic-Analysis based Base Station Detection in WSN. 2008.

[6] S. Brands and D. Chaum. Distance-bounding protocols. In *EUROCRYPT '93: Workshop on the theory and application of cryptographic techniques on Advances in cryptology*, pages 344–359, Secaucus, NJ, USA, 1994. Springer-Verlag New York, Inc.

[7] D. Chaum. Untraceable Electronic Mail, Return addresses, and Digital Pseudonyms. *Commun. ACM*, 24(2):84–88, Feb. 1981.

[8] J. Deng, R. Han, and S. Mishra. Intrusion Tolerance and Anti-Traffic Analysis Strategies For Wireless Sensor Networks. *Dependable Systems and Networks, International Conference on*, 0:637, 2004.

[9] J. Deng, R. Han, and S. Mishra. Countermeasures Against Traffic Analysis Attacks in Wireless Sensor Networks. In *SECURECOMM '05: Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks*, pages 113–126, Washington, DC, USA, 2005. IEEE Computer Society.

[10] 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.

[11] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The Second-Generation Onion Router. In *SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium*, pages 21–21, Berkeley, CA, USA, 2004. USENIX Association.

[12] W. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. Abdelzaher. PDA: Privacy-Preserving Data Aggregation in Wireless Sensor Networks. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 2045–2053, May 2007.

[13] Y. Jian, S. Chen, Z. Zhang, and L. Zhang. Protecting Receiver-Location Privacy in Wireless Sensor Networks. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications*, pages 1955–1963, May 2007.

[14] Y. Jian, S. Chen, Z. Zhang, and L. Zhang. A novel scheme for protecting receiver's location privacy in wireless sensor networks. *Wireless Communications, IEEE Transactions on*, 7(10):3769–3779, October 2008.

[15] P. Kamat, W. Xu, W. Trappe, and Y. Zhang. Temporal Privacy in Wireless Sensor Networks. In *ICDCS '07: Proceedings of the 27th International Conference on Distributed Computing Systems*, page 23, Washington, DC, USA, 2007. IEEE Computer Society.

[16] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk. Enhancing Source-Location Privacy in Sensor Network Routing. In *ICDCS 2005. 25th IEEE International Conference on Distributed Computing Systems*, pages 599–608, June 2005.

[17] J. Kong and X. Hong. ANODR: ANonymous On Demand Routing with Untraceable Routes for Mobile Ad-hoc Networks. In *MobiHoc '03: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 291–302, New York, NY, USA, 2003. ACM.

[18] B. N. Levine and C. Shields. Hordes: A Multicast Based Protocol for Anonymity. *J. Comput. Secur.*, 10(3):213–240, 2002.

[19] Y. Li and J. Ren. Preserving Source-Location Privacy in Wireless Sensor Networks. In *SECON'09: Proceedings of the 6th Annual IEEE communications society conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pages 493–501, Piscataway, NJ, USA, 2009. IEEE Press.

[20] Y. Li and J. Ren. Providing Source-Location Privacy in Wireless Sensor Networks. In *WASA '09: Proceedings of the 4th International Conference on Wireless Algorithms, Systems, and Applications*, pages 338–347, Berlin, Heidelberg, 2009. Springer-Verlag.

[21] K. Mehta, D. Liu, and M. Wright. Location Privacy in Sensor Networks Against a Global Eavesdropper. In *ICNP 2007. IEEE International Conference on Network Protocols*, pages 314–323, Oct. 2007.

[22] S. Misra and G. Xue. Efficient anonymity schemes for clustered wireless sensor networks. *International Journal of Sensor Networks*, 1(1):50–63, 2006.

[23] Y. Ouyang, Z. Le, G. Chen, J. Ford, and F. Makedon. Entrapping adversaries for source protection in sensor networks. In *WOWMOM '06: Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks*, pages 23–34, Washington, DC, USA, 2006. IEEE Computer Society.

[24] Y. Ouyang, Z. Le, Y. Xu, N. Triandopoulos, S. Zhang, J. Ford, and F. Makedon. Providing Anonymity in Wireless Sensor Networks. In *Pervasive Services, IEEE International Conference on*, pages 145–148, July 2007.

[25] C. Ozturk, Y. Zhang, and W. Trappe. Source-Location Privacy in Energy-Constrained Sensor Network Routing. In *SASN '04: Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks*, pages 88–93, New York, NY, USA, 2004. ACM.

[26] S. Pai, S. Bermudez, S. Wicker, M. Meingast, T. Roosta, S. Sastry, and D. Mulligan. Transactional confidentiality in sensor networks. *IEEE Security & Privacy*, 6(4):28–35, July-Aug. 2008.

[27] R. Rajagopalan and P. Varshney. Data-Aggregation Techniques in Sensor Networks: A Survey. *Communications Surveys & Tutorials, IEEE*, 8(4):48–63, 2006.

[28] M. Reed, P. Syverson, and D. Goldschlag. Anonymous Connections and Onion Routing. *Selected Areas in Communications, IEEE Journal on*, 16(4):482–494, May 1998.

[29] M. Reiter and A. Rubin. Crowds: Anonymity for Web Transactions. *ACM transactions on information and system security*, 1(1):66–92, 1998.

[30] V. Rijmen and N. Pramstaller. *Wireless Security & Cryptography: Specifications and Implementations*, chapter Cryptographic Algorithms in Constrained Environments, pages 177–212. CRC-Press, 2007.

[31] S. Seys and B. Preneel. ARM: Anonymous Routing Protocol for Mobile Ad hoc Networks. In *Advanced Information Networking and Applications, 2006. AINA 2006. 20th International Conference on*, volume 2, pages 133–137, April 2006.

[32] R. Shaikh, H. Jameel, B. d'Auriol, S. Lee, Y.-J. Song, and H. Lee. Network Level Privacy for Wireless Sensor Networks. In *ISIAS '08. Fourth International Conference on Information Assurance and Security.*, pages 261–266, Sept. 2008.

[33] S. Shakkottai. Asymptotics of query strategies over a sensor network. In *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 1, 2004.

[34] E. Shakshuki, T. Sheltami, N. Kang, and X. Xing. Tracking Anonymous Sinks in Wireless Sensor Networks. In *Advanced Information Networking and Applications, 2009. AINA '09. International Conference on*, pages 510–516, May 2009.

[35] M. Shao, W. Hu, S. Zhu, G. Cao, S. Krishnamurthy, and T. La Porta. Cross-layer Enhanced Source Location Privacy in Sensor Networks. In *IEEE Conference on Sensor, Mesh, and Ad Hoc Communications and Networks (SECON '09)*, pages 1–9. IEEE Communications Society, June 2009.

[36] M. Shao, Y. Yang, S. Zhu, and G. Cao. Towards Statistically Strong Source Anonymity for Sensor Networks. pages 466–474, April 2008.

[37] M. Shao, S. Zhu, W. Zhang, G. Cao, and Y. Yang. pDCS: Security and Privacy Support for Data-Centric Sensor Networks. *Mobile Computing, IEEE Transactions on*, 8(8):1023–1038, Aug. 2009.

[38] S. Spiekermann and L. Cranor. Engineering Privacy. *Software Engineering, IEEE Transactions on*, 35(1):67–82, jan. 2009.

[39] J. Walters, Z. Liang, W. Shi, and V. Chaudhary. *Security in Distributed, Grid, Mobile, and Pervasive Computing*, chapter Wireless Sensor Network Security: A Survey, pages 367–411. Auerbach Publications, 2007.

[40] H. Wang, B. Sheng, and Q. Li. Privacy-aware routing in sensor networks. *Comput. Netw.*, 53(9):1512–1529, 2009.

[41] S. Warren and L. Brandeis. The Right to Privacy. *Harvard Law Review*, IV(5), December 1890.

[42] W. Wei-ping, C. Liang, and W. Jian-xin. A Source-Location Privacy Protocol in WSN Based on Locational Angle. In *ICC '08. IEEE International Conference on Communications*, pages 1630–1634, May 2008.

[43] 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*, page 8 pp., April 2006.

[44] Y. Yang, M. Shao, S. Zhu, B. Urgaonkar, and G. Cao. Towards event source unobservability with minimum network traffic in sensor networks. In *WiSec '08: Proceedings of the first ACM conference on Wireless network security*, pages 77–88, New York, NY, USA, 2008. ACM.

[45] W. Zhang, C. Wang, and T. Feng. GP^2S: Generic Privacy-Preservation Solutions for Approximate Aggregation of Sensor Data (concise contribution). In *Pervasive Computing and Communications, 2008. PerCom 2008. Sixth Annual IEEE International Conference on*, pages 179–184, March 2008.

[46] Y. Zhang, W. Liu, and W. Lou. Anonymous communications in mobile ad hoc networks. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 3, pages 1940–1951, March 2005.

[47] B. Zhu, Z. Wan, M. S. Kankanhalli, F. Bao, and R. H. Deng. Anonymous Secure Routing in Mobile Ad-Hoc Networks. In *LCN '04: Proceedings of the 29th Annual IEEE International Conference on Local Computer Networks*, pages 102–108, Washington, DC, USA, 2004. IEEE Computer Society.