An Experimental Comparison of Source Location Privacy Methods for Power Optimization in WSNs

GUILLERMO SUAREZ-TANGIL, ESTHER PALOMAR, BENJAMIN RAMOS, ARTURO RIBAGORDA
Carlos III University of Madrid
Computer Science and Engineering Department
Av. Universidad 30, Leganes, Madrid
SPAIN
gtangil@pa.uc3m.es, {epalomar, benja1, arturo}@inf.uc3m.es

Abstract: Wireless sensor technology may be envisioned as the promising substitute of traditional collaborative distributed schemes, especially for monitoring applications. However, the widespread adoption of such technology is being slowed down because of growing security concerns, specially critical in stationary sensor topologies which create new threats to the privacy of individuals and organizations. As a countermeasure against source observability and traceability, several source–location privacy solutions have been presented so far. The goal of this paper is twofold. We present an experimental comparison of the energy consumption of existing strategies (i.e. Flooding, Phantom and Fake Routing) as well as proposing a novel energy–aware approach aimed at disguising the real source in the presence of a global eavesdropper. A major finding of this preliminary work is that source anonymity is reached with minimum consumption.

Key–Words: Wireless Sensor Network, Source–Location Privacy, Energy Consumption, Power Optimization, Experimental Comparison, Sensor Lifetime.

1 Introduction

Geographical monitoring applications are envisioned to be mostly carried out on wireless sensor networks (WSNs) in the near future. However, the widespread adoption of such technology is being slowed down because of growing public concerns about its associated security threats, which are specially critical in stationary sensor topologies. There has been an immense interest in this regard since works presented in [1, 2] incorporate privacy concerns in the design of routing algorithms, aimed at protecting the location of the messages' senders in WSNs.

Furthermore, there are also negative implications regarding the abilities of an adversary to observe the content of the transmitted messages. This fact is crucial especially when context is associated to a sensed data and the message's source and its location are transmitted in clear. For example, in a supermarket scenario, information about consumer habits is also at stake; movements may be tracked and recorded. Moreover, in other scenarios like the *panda hunter game*, assets may be easily linked to a specific location [3]. Put simply, the source is assumed to transmit a sequence of packets, all describing one event. The adversary will be able to gather the source of the transmission based on the activity change. The local adversary, who is standing next to the sink, will follow the transmissions one relay at the time. The global adversary, with unlimited resources, will hear and capture any packet sent in the network (either with a very powerful antenna or having his own sensor network deployed in the area).

To counter this challenge, the design of source–location privacy should fulfill the following security goals in order to prevent adversaries from predicting source locations within a reasonable period of time:

- Anonymous interactions: Transmitted messages do not reveal event–related information, i.e. the source identifier or the asset location coordinates.
- Untraceable routes: Uncertainty of the source–sink route is mandatory.

Most approaches deal with different types of passive attacks, e.g. a global eavesdropper, based on generating fake network traffic in order to disguise the real source. In this paper we explore the advantages and disadvantages of several privacy–enhancing solutions proposed so far, and also propose a novel energy–aware scheme. Our goal is to compare the energy consumption of the existing source–location privacy strategies with our scheme. We found in preliminary experiments that our initial design improves the energy cost for this domain.

The rest of the paper is organized as follows. In Section 2 we describe the problematic issues in defining the network and adversary models which usually represent the main building block for the design of source–location privacy schemes. Section 3 explores the existing privacy–enhancing solutions for WSNs as well as its drawbacks. In Section 4 we introduce our proposal, whereas the experimental comparison is evaluated in Section 5. Finally, Section 6 concludes the paper and outlines research directions.

2 Preliminaries

The vast majority of the formalisms presented so far to deal with privacy concerns in WSNs have been highly context–dependent, imposed by a certain network scenario and/or focused on a specific adversary model. Hence, assumptions made in establishing the network and attack model have influenced the design of security protocols in WSNs. In this section, we briefly outline these implications.

2.1 Network Models

In general, the network model assumes that sensor nodes are static and organized into grids where sensors at neighboring cells are directly connected. However, assumptions on the number of sinks (or base stations) and their locations generally lead to different privacy solutions. Regarding mobility, sinks may be static or mobile, whereas in terms of location they may be situated in centrally or off–center (in other regions of the sensing field). Any combination of the above may determine different constrained scenarios and even result in such limited, specific solutions. This fact is specially critical when determining the adversary model. On the other hand, the procedure for event report and/or data dissemination also determines the required security solution. Finally, the event generation rate and the amount of data generated by each event are also important factors when designing privacy–aware protocols.

2.2 Adversary Models

Assuming a passive adversary scenario, attackers attempt to be as "invisible" as possible until locating the asset/s. Within this scenario, we have identified two types of adversary which define the strategy to discover the source location, as follows:

1. A simple adversary with limited resource focuses his strategy on being located close to the sink and tracing back the incoming communications in a hop–by–hop fashion. This type of adversary should be able to physically move from place to place.
2. An unlimited–resource adversary, however, possesses more powerful devices and/or multiple cooperative attackers dispersed over the WSN [4] in order to collect *all* the messages transmitted in the network.

This way, specific solutions to specific attack scenarios are generally invalid in other cases. In fact, it is essential to determine whether the adversary can access all the information about the whole network topology.

3 Preserving Source–Location Privacy

In both wired and wireless domains there has been intense research activity to provide the communicating parties with anonymity e.g. preventing traffic analysis [5, 6]. However, the majority of the presented approaches are not appropriate for WSNs, due to the resource–constrained environment and its inherent characteristic of routing (interested readers will find a complete survey on routing protocols for WSNs in [7]). In this section, we overview two promising approaches to source–location privacy, i.e. flooding–based and fake messaging; pointing out main advantages and disadvantages.

3.1 Flooding–based approach

Flooding-based routing mechanisms where everyone participates in data forwarding has been employed so far in preserving privacy for message sources [8]. This techniques are usually easy to implement since simplifies the routing protocol. However, the communication cost of message flooding might be prohibitively expensive in WSNs (see Fig. 1–(a)). Several works deal with this cost using limited and/or probabilistic flooding strategies but still remain inefficient when providing source location privacy [9]. In fact, flooding has been proved not to work with local adversary because the "front" of flooding propagation moves as fast as the shortest route, and the source location can be inferred by the front's gradient. For instance, adversaries can use temporal dependency between transmissions to trace messages' forwarding paths [10]. These attacks are easily launched by mobile adversaries, especially when assets are static and sources send multiple packets over a period of time.

To counter these problems, some works study the application of probabilistic flooding, but having a poor message delivery ratio (less than 70%) [11]. On the other hand, in scenarios where assets are mobile and the eavesdropper attack is limited, *accumulative broadcast* (i.e. a minimum–energy broadcasting [12]) provides sufficient uncertainty within a reasonable period of time to prevent the attacker from reaching the source.

Additionally, *gossip* and *rumor*-based routing consist of a limited flooding–based strategy where relaying nodes forward received messages according to a given retransmission probability. Generally, both models work well in dense networks but present several differences between each other. For instance, gossip protocols perform reliable network broadcasts, probabilistically. On the contrary, in rumor routing, sources maintain a unique path between a given destination [7]. The latter also presents some differences with regard to *direct diffusion* baseline, which is not a good choice for environmental monitoring, and offers poor source location privacy [13]. Due to space limitations, we could not include further details of these mechanisms which are not indeed appropriate to achieve source–location privacy in current WSN deployments.

3.2 Phantom Routing

Since forwarding decisions through the random walk are made locally and independently of the location context, this technique seems promising for protecting source location from limited eavesdroppers. Adversaries must then apply backtracking strategies, which usually require attackers to increase theirs resources (e.g. the radio range and/or the number of observation points) several times over. However, it has been proved that without geographical directrices the random walk loops around the source with high probability [14].

Several proposals address this challenge by introducing bias into the walking process [15]. However, *Phantom routing* excels other flexible approaches. This is a two-stage routing scheme that first consists of a “directed” walk along a random direction and a subsequent flooding path phase towards the sink (see Fig. 1–(b)). The objective of the former is to direct the message to a phantom source h_{walk}–hops away from the real source [11]. Limitations of both stages in the presence of a global eavesdropper show the inefficacy of the approach.

To solve these limitations, recent works propose similar two-stage hybrid approaches, most in the way of establishing a flooding backbone. For example, in the scheme presented in [16], the message source randomly selects an intermediate node in the sensor domain, i.e. one of her adjacent neighboring nodes is selected as a phantom source, who then transmits the data packet to a ring–shaped backbone node. In this case, authors assume that the adversaries are unable to monitor the entire network. We further elaborate on the use of backbone nodes in the following section.

3.3 Fake/Dummy messaging

The key idea of fake messaging is to protect sources’ location by introducing more sources which inject fake packets into the network, as depicted in Fig. 1–(c). In order to decrease the communication cost and the number of network collisions caused by dummy baseline strategies, recent approaches deploy a special layout over the network to carry out special tasks such as packet filtering and forwarding towards the sink. This underlying network infrastructure, also called backbone, consists of a set of selected sensor nodes which are located strategically and play a special role for the whole sensor community. For example, in work [17], these selected nodes, called *proxies*, store and relay (re–encrypted) real data packets from neighboring sensors around them, whereas dummy data remains filtered by discarding.

In this context, introducing appropriate delays and buffering at proxies is essential, i.e. proxies must emit messages (either bogus or real) following a certain outgoing traffic rate [18]. However, the construction of an efficient backbone (e.g. an induced graph of a minimum connected dominating set), which is a NP–hard problem, penalizes these approaches, and should usually be approximated by heuristics. Our approach agrees that the only way to deceive a global adversary requires a consistently data packet (dummy or valid) transmission.

Figure 1–(c) depicts a brief description of the operation for the above proposals.

4 Our Approach

In this section we introduce our novel two–phase routing approach to protect an observed object (e.g. an important person who is being followed, or endangered species) from the global adversary. Hence, the scheme maintains source–location privacy based on the following assumptions:

1. In an **initialization phase**, two special sensor roles are established: the *squad header* and the *dummy source*.

2. In a **Wake-up phase**, squad header will trigger wake-up calls towards a dummy source.

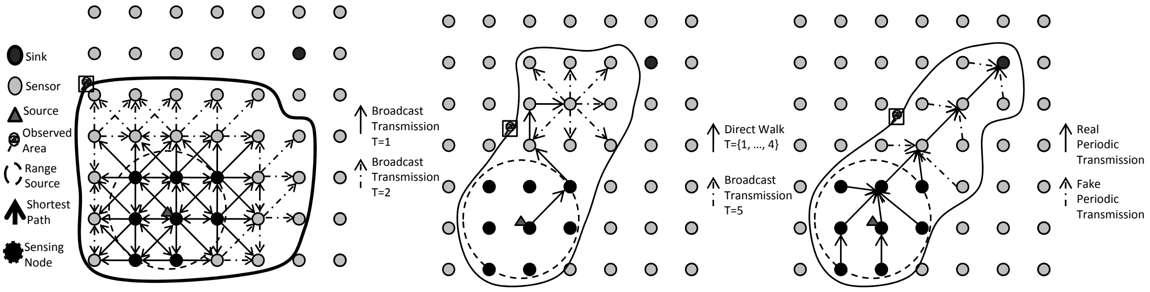
Sink
Sensor
Source
Observed Area
Range Source
Shortest Path
Sensing Node
Broadcast Transmission T=1
Broadcast Transmission T=2
Direct Walk T={1, ..., 4}
Broadcast Transmission T=5
Real Periodic Transmission
Fake Periodic Transmission

Figure 1: (a) Flooding-based approach, (b) Phantom Routing, and (c) fake messaging.

3. Direct **transmission to sink**. Messages from dummy sources are propagated by identifying the shortest path between the dummy source and the sink.

4.1 Network Initialization

A static grid of N nodes is divided into d sensor squads called *dummy populations*. Additionally, a leader, tagged *squad header*, is periodically selected within each dummy population. Each squad header is responsible for appointing a *dummy source*. A distribution function governs the interval between two consecutive calls. The creation of these neighboring groups and leader selection can be deployed by recent works proposed in the literature such as [19].

4.2 Wake-up Phase

Therefore, sensors remain asleep until the corresponding *squad header* n_h sends a wake-up call m_{w-up} to a dummy source n_d among its dummy population, which happens periodically and randomly within a certain rate:

$$n_h \to n_d : m_{w-up} = enc_K(ID_h, ID_d, wakeup)$$

where $enc_K(x)$ is the symmetric/asymmetric encryption of message x using K as key.

Though the use of a distribution function does imply some determinism, the strength lies in the pseudo–randomness of the dummy source selection. Moreover, implicit energy costs are saved since sensors are not transmitting all the time.

In this context, the geographic disposition of dummy sources around the header becomes a key factor, i.e. one event is generally captured by a subset of nodes adjacent to the source of the event. This way dummy nodes will capture the event and according to their position, the message will be sent towards the sink through a different path.

4.3 Direct Transmission to Sink

Upon receiving a wake-up call, dummy sources create a valid/fake response message m_{res} and send it to the corresponding relaying node towards the sink by the shortest path:

$$n_d \to S : m_{relay} = m_{res} = E_K(data)$$

Finally, the routing backbone is responsible for an efficient data dissemination, m_{relay}. Contributions in this area focus on filtering and the encapsulation of data/fake packets, while proving robustness to lost connections and nodes.

Remarks. Note that on average, each sensor transmits (direct to the sink) at a certain instant. From the eavesdropper's viewpoint, nodes who were first to transmit are dummy, and hence a source of confusion. In sections below we conducted several experiments to prove such a strategy can be designed as energy efficient compared to other approaches described in [3, 17].

5 Experimental Evaluation

We have conducted simulation experiments to evaluate the performance of our proposal using Castalia WSN Simulator [20]. Tests conducted show that our proposal presents less or equal energy consumption when comparing it with traditional proposals, i.e. (i) flooding, (ii) fake messaging, (iii) dummy wake-up, (iv) persistent sending using a round robin scheduling, and (v) phantom routing.

5.1 Simulation Scenario

In our simulation, a Panda-Hunter-Game scenario was built. We considered a static sensor network consisting of $n = \{i^2 \ \forall i = 2...19\}$ sensor nodes. The nodes communicate with each other using a CC2420 radio. Additionally the following protocols, already implemented in Castalia, are used: TMAC for the medium

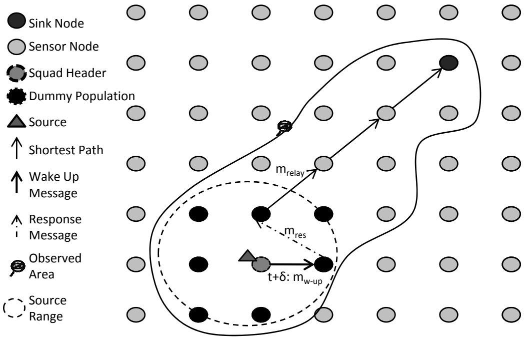

Figure 2: Our Dummy Wake-Up scheme.

access control and Multipath for the network routing (when flooding is not used). Sink is located in the central area of the grid, whereas node distance is 5 meters. The asset moves through the deployment area at $1m/s$. Thus, each node senses at a sampling rate of 5s. Therefore, we use an exponential distribution with $\mu = 5$ for fake messaging, slotted round robin, and wake-up calls generation. We obtain the average power consumption per node after 10 runs of 600s each.

5.2 Power Comparison

The experimental results, shown in Figure 3, allow us to conclude the following:

- Small deployments present similar costs.
- Large networks (i.e. from 36 nodes upwards) show significant differences in terms of power costs, specially when comparing flooding and fake Messaging in scenarios with very dynamic assets.
- A simple fixed scheduling, like Round Robin, results in very low power consumption but lacking of scalability as traffic rate increases.
- Phantom routing presents several improvements as it has already been shown in the literature, however, power consumption is expected to get worse as the number of assets increases.
- Our dummy wake-up proposal presents similar estimations to phantom routing.

5.3 Sensor Lifetime

We evaluate the expected lifetime of the deployed scenario regarding the schemes compared above. Table 1 outlines the sensors' lifespan for a number of network sizes, assuming each sensor has two AAA-batteries[1]. For example, estimations for large networks with 100 nodes using our scheme reach a non-negligible result of 2 weeks, while flooding will deplete sensor energy supplies in only 4 days. Hence, dummy wake-up protocol enhances network lifetime up to 350% when comparing with a flooding base approach.

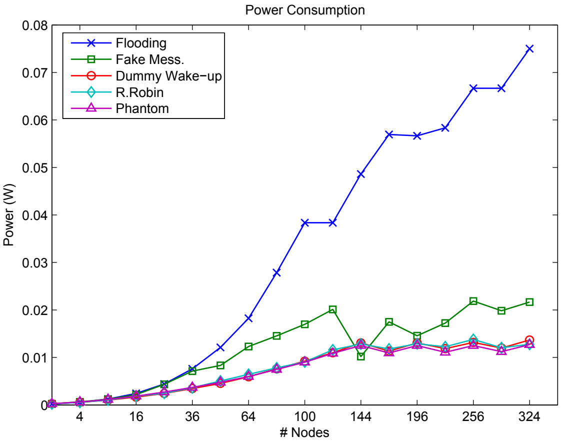

Figure 3: Average power consumption.

No. Nodes	Flooding	Fake	R. Robin	Dummy	Phantom
4	544,4	472,0	500,0	411,6	459,2
9	202,7	177,2	203,9	208,3	182,4
16	91,7	93,9	109,2	111,8	97,0
25	46,3	51,8	61,1	68,6	60,5
36	25,6	25,9	45,4	46,3	41,4
49	14,8	15,8	31,1	32,6	30,6
64	9,3	13,5	22,3	24,8	23,6
81	6,2	9,2	17,5	19,0	18,8
100	4,0	7,7	14,4	14,8	15,1
121	2,9	6,6	12,4	12,2	12,5
144	2,9	5,6	9,7	10,2	10,4
169	2,3	11,0	8,8	8,6	9,0
196	2,0	6,4	9,5	9,9	10,3
225	2,0	7,7	8,7	8,6	9,0
256	1,9	6,5	9,2	9,5	10,2
289	1,7	5,1	8,2	8,5	9,0
324	1,7	5,7	9,3	9,4	10,0
361	1,5	5,2	8,8	8,2	8,9

Table 1: Expected Lifetime (days).

6 Conclusion and Future Work

When achieving source–location privacy in WSN, persistent fake messaging has been revealed as the most promising technique to mislead the passive adversary. However, it is still a challenge to efficiently create fake sources as well as defining the appropriate traffic generation rate –specially in relatively dense

[1] For our estimations, we assume two 900 $mA \cdot h$ batteries of $1.5V$.

sensor networks. In this paper, we propose a novel energy–aware scheme based on the group formation of dummy sources. Conducted experiments comparing our proposal with other approaches presented so far conclude encouraging results in terms of power consumption and sensor lifetime.

In future works, we will further elaborate on the performance measurements particularly evaluating collisions to prove expected benefits of our proposal. Finally, we will extend our energy comparisons to include solutions such as gossip and rumor-based flooding.

References:

[1] M. Gruteser, D. Grunwald, Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking, in *Procs. of the 1st Int. Conf. on Mobile Systems, Applications and Services*, ACM, May 2003, pp. 31–42.

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

[3] N. Li, N. Zhang, S.K. Das, B. Thuraisingham, Privacy preservation in wireless sensor networks: A state-of-the-art survey, *Ad Hoc Networks*, vol. 7, no. 8, pp. 1501–1514, 2009.

[4] K. Mehta, L. Donggang, M. Wright, Location Privacy in Sensor Networks Against a Global Eavesdropper, in *Procs. of the IEEE Int. Conf. on Network Protocols*, 2007, pp. 314–323.

[5] A. Wadaa, S. Olariu, L. Wilson, M. Eltoweissy, K. Jones, Providing Anonymity in Wireless Sensor Networks, in *Procs. of the 10th Int. Conf. on Paral. and Distrib. Systems*, IEEE Computer Society, 2004, pp. 1521–9097.

[6] R. Dingledine, N. Mathewson, P. Syverson, Tor: the second-generation onion router, in *Procs. of the 13th conf. on USENIX Security Symposium*, USENIX Association, 2004, pp. 21–21.

[7] K. Akkaya, M. Younis, A Survey on Routing Protocols for Wireless Sensor Networks, *Ad Hoc Networks*, vol. 3, pp. 325–349, May 2005.

[8] Y. Xiang, D. Chen, X. Cheng, K. Xing, M. Song, Localized Flooding Backbone Construction for Location Privacy in Sensor Networks, in *Procs. of the ACIS Int. Conf. on Soft. Engineering, AI, Net., and Paral./Distrib. Comp.*, IEEE Computer Society, 2007, pp. 167–171.

[9] P.Th. Eugster, R. Guerraoui, S.B. Handurukande, P. Kouznetsov, A.M. Kermarrec, Lightweight probabilistic broadcast, *ACM Trans. Comput. Syst*, vol. 21, no. 4, pp. 341–374, 2003.

[10] J. Kong and X. Hong, ANODR: anonymous on demand routing with untraceable routes for mobile ad-hoc networks, in *Procs. of the 4th Int. symposium on Mobile ad hoc net. & computing*, ACM, June 2003, pp. 291–302.

[11] P. Kamat, Y. Zhang, W. Trappe, C. Ozturk. , Enhancing Source–Location Privacy in Sensor Network Routing, in *Procs. of the 25th Int. Conf. on Distrib. Comp. Syst.*, IEEE Computer Society, 2005, pp. 599–608.

[12] I. Maric, R. Yates, Performance of Repetition Codes and Punctured Codes for Accumulative Broadcast, in *Procs. of the Modeling and Optimization in Mobile, Ad Hoc and Wireless networks Workshop*, March 2003.

[13] R. Govindan, D. Estrin, Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks, in *Procs. of the 6th ACM Int. Conf. on Mobile comp. and net.*, ACM, 2000.

[14] Y.L. Yun, J. Ren, Providing Source-Location Privacy in Wireless Sensor Networks, in *International conference on Wireless Algorithms, Systems and Applications*, Springer Berlin/Heidelberg, 2009, pp. 338–347.

[15] Y. Xi, L. Schwiebert, W. Shi, Preserving source location privacy in monitoring–based wireless sensor networks, in *Procs. of the 20th Int. Paral. and Distrib. Processing Symposium*, IEEE Computer Society, 2006.

[16] Y.L. Yun, J. Ren, Preserving Source-Location Privacy in Wireless Sensor Networks, in *Sixth Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, IEEE, 2009, pp. 338–347.

[17] Y. Yang, M. Shao, S. Zhu, B. Urgaonkar, G. Cao, Towards event source unobservability with minimum network traffic in sensor networks, in *Procs. of the 1st ACM conf. on Wireless network security*, ACM, 2008, pp. 77–88.

[18] A. Abbasi, A. Khonsari, S.M. Talebi, Source location anonymity for sensor networks, in *6th IEEE conf. on Consumer Communications and Netw. Conf.*, IEEE, 2009, pp. 588–592.

[19] O. Younis, S. Fahmy, HEED: A Hybrid, Energy-Efficient, Distrib. Clustering Approach for Ad Hoc Sensor Networks, *IEEE Transactions on Mobile Computing*, vol. 3, pp. 366–379, 2004.

[20] A. Boulis, Castalia: revealing pitfalls in designing distributed algorithms in WSN, in *Proc. of the 5th int. conf. on Embedded networked sensor systems*, ACM, 2007, pp. 407–408.