EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Anti-traffic analysis attack for location privacy in WSNs

Bi Di Ying[1*], Dimitrios Makrakis[2] and Hussein T Mouftah[2]

Abstract

Traditional encryption and authentication methods are not effective in preserving a sink's location privacy from a global adversary that is monitoring the network traffic. In this paper, we first propose a novel anti-traffic analysis (ATA) method to preserve the sink's location privacy. In order to confuse a local or global adversary, each node generates dummy messages, the number of which is dependent on the number of the node's children. Hence, ATA is able to prevent the adversary from acquiring valuable information on the sink's location through the traffic analysis attack. However, a larger number of dummy messages lead to consumption of extra energy. Then, we design our improved ATA (IATA) in such a way that we select some sensors to act as fake sinks, to ensure that sensors around fake sinks generate dummy messages and discard received dummy messages. Since the problem of the optimal fake sinks' placement is nondeterministic polynomial time (NP)-hard, we employ local search heuristics based on network traffic and security entropy. Performance analysis of the ATA scheme can protect the sink's location privacy, and IATA scheme can reduce energy consumption.

Keywords: Sensor network; Traffic analysis attack; Privacy

1. Introduction

Wireless sensor networks (WSNs) are deployed to support the sensing and communication needs of the deploying entity. Due to the broadcasting nature of wireless communication medium, adversaries can eavesdrop on network traffic to obtain valuable information. Existing security technologies cannot always protect the cyber-security needs of users and the run applications, in terms of data confidentially and integrity and user privacy and anonymity. Network traffic analysis can be used by an adversary to extra important information related to the node location, functionality, and identity. Traffic patterns of WSNs can reveal a great deal of contextual information, which can disclose the location of critical nodes. For example, sensing data are transmitted along relatively fixed paths connecting source nodes to a sink. This produces quite easily identifiable traffic patterns that reveal a sink's location. In addition, the sensing nodes having one-hop distance from the sink have to forward a significantly greater volume of packets, since they have to route all the traffic generated by all those nodes that are farther than nodes having one-hop away from the sink. An adversary having a global view of WSN's traffic activity can deduce the location of the sink by observing and analyzing the traffic volume distribution within WSN's coverage area for an adequately long time interval. Discovery of a sink's location may allow the adversary to launch precise physical and cyber attacks against the sink and thereby disable the network.

Evidently, location privacy is very important, especially for unattended WSN deployments in harsh or hostile environments. Recently, a number of location privacy protection methods have been developed for sensor networks, to resist the various types of traffic analysis attacks (e.g., those based on monitoring traffic patterns, traffic rates, and traffic volumes). Most of them are designed to protect source location privacy against an adversary that is only capable of eavesdropping on a limited portion of the network at a time [1-4]. However, the contributions in the current literatures related to a sink's location privacy are limited [5-7]. These methods involve multipath routing, fake message injection; however, those techniques become ineffective in the presence of a global adversary.

* Correspondence: yingbidi@mail.zjgsu.edu.cn
[1]School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China
Full list of author information is available at the end of the article

In order to prevent traffic analysis attacks by the global adversary, we propose an anti-traffic analysis (ATA) approach to protect the sink's location privacy by artificially homogenizing traffic intensity. In order to confuse a local or global adversary, each node generates dummy messages. The number of dummy messages is dependent on the number of the node's children (A node 'X' whose messages to the sink have to pass through node 'Y' is considered as 'child or kid' of 'Y'.). This approach is able to hide the sink's location. Performance analysis shows that our ATA can protect the sink's location under the global adversary by launching traffic analysis attacks.

Due to the bandwidth and energy consumed by the transmission of the dummy messages, the stronger the protection for the actual sink is, the higher the network load and energy consumption will be. Therefore, we then design an improved ATA (IATA) in such a way that one can easily tune the trade-off between the protection strength and the overhead. In this IATA, we select some sensors to act as fake sinks and to imitate the actual sink's behavior by generating dummy messages. Since the problem of optimal fake sinks' placement is nondeterministic polynomial time (NP)-hard, we employ local search heuristics that make use of traffic volume and a sink's location privacy level quantifying parameter to decide the fake sinks' placement positions. Simulation results show that IATA can reduce energy consumption while maintaining the sink's location privacy. The contributions in this paper are the following:

(1) The design of a new ATA scheme that protects sink's location privacy is provided. To do so, ATA homogenizes artificially the traffic intensity distribution over the coverage area of the WSN. It consists of a new topology discovery protocol and a novel technique that evens out WSN's traffic volume distribution over its coverage area.
(2) An IATA is proposed to reduce extra energy induced by dummy messages. This IATA chooses some sensors instead of all sensors to imitate sink's behaviors; thus, it can provide a trade-off between the protection strength and the communication overhead.

Performance evaluation conducted through computer simulations confirms the worthiness of the proposed technology. Our two approaches for protecting the sink's location privacy have distinct properties that make them suitable for different applications. The remainder of the paper is organized as follows. In Section 2, the related work is surveyed. Section 3 presents goals. Section 4 provides our ATA approach, and then performance analysis is given. Section 5 further proposes an IATA approach. Section 6 gives simulation performances of the ATA scheme and the IATA scheme. Finally, Section 7 concludes the paper.

2. Related work

Privacy has been an active area of research in recent years [8]. For example, Zhou et al. [9] designed a new multimedia traffic classification and analysis method for handling the heterogeneity of diverse applications. Li et al. [10] proposed an efficient intrusion detection protocol based on energy prediction in cluster-based WSNs. The aim of this protocol is to resist denial-of-service attacks due to the broadcast nature of wireless communication. In location-based services, a user may want to protect his/her location and can try doing so by using various existing techniques, such as k-anonymity [11-13]. However, those technologies cannot resist passive traffic analysis attacks where an adversary monitors packet transmissions to infer locations of critical - to the infrastructure - elements. Deng et al. in [1] identified two classes of passive traffic analysis attacks that can be applied to WSNs, namely, rate monitoring attack and time correlation attack. In the rate monitoring attack, an attacker monitors the packet transmission rate of nodes close to the attacker and moves gradually closer to the nodes that have a higher packet sending rate, eventually reaching the sink. In a time correlation attack, an attacker determines the correlation in frame sending time between a node and its neighbors. The node that is forwarding the same frame is expected to transmit it right after or at least very soon after receiving it. Thus, by identifying the forwarding nodes of the packet (as it propagates towards the sink), the attacker can deduce the routing path that leads to the sink node. It can also reverse the direction it follows and identify the source as well. Existing literatures against passive traffic analysis attacks are generally proposed to protect source nodes' location privacy protection:

(1) Source nodes' location privacy protection
Several schemes dealing with the source nodes' location privacy can be found in [2,4,14-22]. For example, Kamat et al. [4] proposed the phantom flooding protocol transmission of a packet to defend against an external adversary. It forwards packets from the source node to the sink using a random-walk-based approach. Yang et al. [14] used a proxy-based filtering method. Some sensors are selected as source proxies to collect and filter dummy messages. This method reduces the communication cost by dropping many dummy messages while providing source event unobservability. Xi et al. [15] proposes the sink region routing method. In this case, the source node selects an intermediate node within a designated area close to the sink node. The area should be large enough to make it impossible for a local attacker to monitor the entire region. The

techniques described in [14] and [15] cannot protect against the global attacker.

(2) Sink's location privacy protection

A variety of approaches have been used for this purpose, such as fake message injection, randomization of forwarding delay, and use of fake sinks in order to hide the real sinks' positions [3,5,6,23-33]. For example, Nezhad et al. [24] proposed an anonymous topology discovery protocol where all nodes were allowed to forward route discovery messages and incoming/outgoing labels assigned to nodes. This method hides the location of a sink. However, a route discovery message may fail to discover all sensors since only one copy of this message is forwarded by each node; in other words, this protocol may lead to some sensors becoming isolated or separated from the network. Compared to [24], in the topology discovery phase of our proposed ATA scheme, the sink broadcasts a message which requests for establishing a routing tree; thus, our ATA can avoid some areas which have isolated nodes.

Li et al. [31] proposed an intelligent fake packet injection scheme based on the random walk. This scheme provides a balance between the packet delivery latency and the sink's location privacy. Yao et al. [32] and Chen et al. [33] studied further the random-walk approach of [31]. However, Li et al., Yao, and Chen and Lou [31-33] cannot resist passive traffic analysis attacks under a global attacker. Compared to [31-33], our ATA scheme uses fake messages instead of the random-walk method and makes each node have the same traffic volume. Thus, our scheme ATA can resist traffic analysis attacks launched by a global attacker. Ebrahimi et al. [28] attempted to protect the sink's location privacy by having the sensors located in low-traffic-activity areas to send fake packets, in order to distract the attention of the local adversary. Compared to [28], our ATA scheme does not only make sensors in lower-traffic-activity areas generate fake messages, but also let nodes close to the sink generate dummy messages. Thus, our ATA scheme can prevent traffic analysis attacks under a global attacker.

In [5], dummy sinks are introduced to confuse an adversary from tracking a packet as it moves towards a sink node. Although the inclusion of dummy sinks can protect a sensor network from local adversaries, it is not effective in the case of a global adversary, since global traffic analysis will allow the identification of all fake and real sinks, and the adversary can neutralize all of them. In [26], Mehta et al. proposed to create multiple candidate traffic traces going to the established fake sinks in order to hide the traffic aggregating around real sinks. Similarly to [5], whenever a fake sink receives a packet and broadcasts it locally, it will make the attacker believe that a real sink could be in the range of the fake sink. This method cannot protect the real sink adequately. Besides, this scheme cannot prevent the time correlation attack and rate monitoring attack. Compared to [5,26], our proposed ATA scheme makes each node have the same traffic volume. Thus, our scheme can hide the sink location completely.

Bicakci et al. [25] made all nodes including the sink to equalize the values of their total incoming and outgoing flows. Data generated by each node is destined not only to the sink but also to every other node in the network. This scheme consumes the significant amount of energy and has quite high needs in terms of processing and memory. Compared to [25], our ATA performances are much better; it is scalable and also protects privacy. The reason for having superior performances will be understood after the protocol becomes described. Ying et al. [34] designed a concealing sink location (CSL) protocol that made a node generate the same traffic volume with the sink's neighbors by transmitting a number of fake messages. This feature enables CSL to prevent the traffic analysis attack launched by a global adversary. However, the design of CSL protocol is based on the following assumptions: (i) Sensors are deployed within a circular area. (ii) The deployment is done according to a uniform distribution. (iii) The sink is located at the center of the sensor deployment area. Such conditions are restrictive and do not apply to many cases of WSN deployments. Compared to [34], our protocol can remove some of CSL's drawbacks. In this case, a node generates fake messages according to the total number of nodes whose routing path to the sink passes the node (we have named them as 'kids' of the node). This removes the above assumptions from CSL.

Table 1 summarizes the capabilities of the discussed schemes about the sink's location privacy.

3. Design goals

Protecting the sink's location privacy under the global attack model is challenging. We can encrypt and authenticate all packets during their forwarding to prevent content privacy [4]; however, this cannot solve the traffic analysis attack threat [1,35]. For example, traffic patterns of WSNs can disclose valuable statistical information that exposes the location of sink(s), thus jeopardizing their location privacy. Current literatures describe techniques that employ fake sink(s) [5,26], dummy messages [25], dummy trajectories [36], random message forwarding delay [1], multipath [24] routing, and false distances between nodes to the sink(s) [37]. However, all existing methods have one or more of the following problems: (1) Some of them only can resist traffic analysis attacks launched by a local attacker, not a global one. (2) Even those capable of defending against rate monitoring attack lunched by a global attacker cannot prevent the disclosure of statistical information that can be explored by

Table 1 Differences among literatures

References	Type of attacker it can defend against	Time correlation attack [1]	Rate monitoring attack [1]
[3]	Local	No	Yes
[5]	Local	No	No
[6]	Local	No	Yes
[24]	Local	No	No
[25]	Local and global	No	Yes
[26]	Local and global	No	No
[27]	Local	No	No
[28]	Local	No	No
[29]	Local	No	No
[30]	Local and global	No	No
[31]	Local	No	No
[32]	Local	No	No
[33]	Local	No	No
Proposed ATA scheme	Local and global	Yes	Yes

other forms of traffic analysis attacks such time correlation or traffic volume attack. (3) There is a trade-off between communication/computation/consumption cost and offered security/privacy level. Use of dummy traffic [38] increases significantly the volume of network traffic, thus increasing the communication, computation, and energy consumption costs. The goal we set for this work is to come up with a technology capable of defending sink(s) location privacy, even when the global attacker applies all the abovementioned kinds of traffic analysis attack simultaneously. In our design, we take into consideration the importance of minimizing the network traffic to allow use of lightweight processing hardware/software by the sensors and operation under high energy efficiency.

Without loss of generality and for making the understanding of the proposed technique easier to the reader, we consider that the WSN has a single sink. The traffic analysis attack model has the following properties: (1) The attacker is passive, external, and global. This is realistic [14,22,25,26,30], and previous works investigate the problem of location privacy under the global and passive attacker [14,22,25,26,30]. The global attacker is capable of monitoring all the network traffic by deploying traffic-monitoring devices (e.g., BlueRadios SMT Module, BlueRadios, Inc., Englewood, CO, USA) within the area the WSN cbnh99overs. Note that, at the current price for a BlueRadios SMT Module at $25, the attacker needs only $25,000 to build a network of 1,000 nodes [39-41]. What is more, the number of nodes can typically be smaller than the number of nodes in the target network as they monitor wireless radio signals instead of directly sensing the environment. Thus, for even moderately valuable location information, this can be worth the cost. (2) The attacker cannot distinguish between actual information carrying messages and those carrying fake information or other types of data (e.g., routing-tree formation messages). This is a valid assumption when all messages are encrypted, e.g., by using pair-wise secret keys [42].

4. Anti-traffic analysis protocol

For the reader's convenience, we provide in Table 2 the definitions of notations appearing in the remaining of this work.

4.1. Functionality description of ATA

Execution of ATA includes two main tasks: topology discovery task and data transmission task. The topology discovery is performed periodically in order to track topology changes occurring due to the energy depletion of sensor nodes. Data transmission runs after the WSN is formed and is responsible for the transfer of data (generated by the sensors) to the sink. These tasks are described below.

4.1.1. Topology discovery task

Recently, several routing-tree formation protocols for WSNs were proposed, such as the directed diffusion protocol [43], probabilistic flooding protocol [44], and controlled flooding protocol [45]; however, none of them

Table 2 Definition of notations

Parameter	Meanings
RDM	Real data message
FDM	Fake data message
$H(i)$	Number of kids node i has
$m(i)$	Number of fake messages generated by node i
ρ	The (average) generation rate of RDM messages per unit of time each node generates
$X \times Y$	2 dimensional deployment area (m)
Real sink	Rs
Fake sink	Fs
r_1	Sensors' communication range
$x(i), y(i)$	Location of a node i or a sink i
N	Total number of nodes in the network
TPN_1	Largest amount of traffic volume among the traffic volumes generated by the real sink's neighbors
$TPN_{\varepsilon+1}$	Largest amount of traffic volume in the traffic volumes generated by nodes from $(\varepsilon + 1)$ hops from the real sink
hop (i)	Number of hops forming the routing path from node i to the real sink
h_{max}	Size of the longest routing path from a source to real sink that is formed over the WSN
$\Phi(i, j)$	Number of hops from node i to the fake sink j
$\Omega(k, j)$	Number of hops from a sink k to sink j

can support the sink's location privacy when WSN is subjected to passive traffic analysis attacks. The proposed topology discovery protocol enables the sink to discover the relative positions (but not necessarily the geographic coordinates) of all sensors without compromising WSN's location privacy if the WSN is under the surveillance of a global passive attacker.

Topology discovery task is performed periodically and consists of two sequential phases: *Phase1:* build each node's route path. *Phase2:* determine the number of each node's children, and inform each node the number of the children of sink's neighbors.

(1) Phase 1

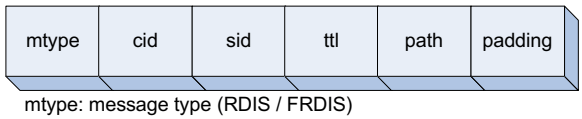
This phase has a goal to discover (for each sensor node) a route connecting the node to the sink. This phase requires the formation and transmission of two different types of messages: route discovery (RDIS) message and fake RDIS (FRIDS) message. The RDIS message's format is shown in Figure 1. '*mtype*' contains the code identifying the message as of RDIS type. The content of '*cid*' field identifies the routing-tree formation refreshing cycle and is the same for all RDIS messages generated during the same cycle. It is set by the sink and the sink increases its value by 1 when it starts a new refreshing cycle. The first value is set to '1'. The '*sid*' field contains the identifier of the sender node that is broadcasting this message. The value of '*ttl*' field indicates the time to live for the particular packet. The value K set by the sink in the *ttl* field is integer positive and should be large enough to allow the broadcasted RDIS message to reach all nodes, including those located at the edges of the WSN with a very high probability. The '*path*' field records identifiers of nodes that this message has passed through. The first value of *path* field is the sink's identifier. In order to deny the attacker to acquire information by analyzing the size of transmitted packets, we make all transmitted messages and data-carrying packets of equal size. The '*padding*' field is used to add to the RDIS structure bytes in order to reach the specific size set for all messages. The size of *mtype* field can be as small as 3 bits. The size of {*cid*/*sid*/*ttl*/*path*} segment is dependent on the size of the network how frequently the routing-tree topology is refreshed, whether truncated or full node ID address is used; however, assigning a maximum of 4 bytes to each of those fields is deemed sufficient for WSN' s applications [46].

After sensors are deployed in the area, formation of the routing tree (to the sink) is required. The process is triggered by the sink and is done so by broadcasting the first RDIS message. Any intermediate node receiving the RDIS message of the present cycle for the first time records information included in the RDIS message and rebroadcasts it after making the following modifications: it places its own identifier in the *sid*, decreases the value of the *ttl* field by 1, records the values of *path* field, and adds its own identifier in the *path* field. Any RDIS messages of the same cycle received by the node are dropped. This policy ensures that in each topology discovery execution, every node records the first RDIS message it receives and generates only one RDIS of its own during the current topology discovery task.

Right after WSN's deployment, there is a concern that if the sink is the first one to generate transmission, there is a small chance to compromise its position. This is possible if the attacker has location identification capability. To eliminate even this possibility, each sensor performs the following process: each sensor decides with a pre-set probability μ to generate FRDIS message (FRDIS message format is shown in Figure 1). If a receiving node receives the FRDIS for the first time, it records the value of *cid*, modifies *sid* and *ttl* by following the same method used for RDIS messages, and rebroadcasts it. All subsequently received copies of the FRDIS message are dropped by the node.

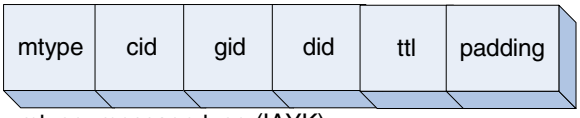
(2) Phase 2

The objective of this phase is to make each node aware of how many nodes will be using it as a relay when they are sending messages to the sink. This is done as follows. Each node generates and sends an 'I am your kid' (IAYK) message towards the sink, which is routed on the established route path connecting the node to the sink. IAYK's structure is shown in Figure 2. The value of *cid* field identifies the route-update cycle. All IAYK messages generated and transmitted during a certain topology discovery cycle contains the same *cid* value. The '*gid*' field contains the identifier of the node that generates the specific IAYK message.



mtype: message type (RDIS / FRDIS)
cid: sequence number of routing-refreshing cycle
sid: identifier of the node sending the message
ttl: Time to Live
path: set of identifiers of nodes that the message visited
padding: is variable in message size

Figure 1 RDIS/FRDIS message format.



mtype: message type (IAYK)
cid: sequence number of route-update cycle
gid: identifier of the node generating the message
did: identifier of the node receiving the message
ttl: Time to Live
padding:is variable in message size

Figure 2 IAYK message format.

Each node can determine the number of children it has by calculating the number of IAYK-type messages it receives, each having different *gid* field values from the others. Each node, when receiving the IAYK message for the first time, records the next hop node's identifier (as determined by the routing path) into the '*did*' field, decreases the value of the *ttl* field by 1, and transmits the modified IAYK message. The size of *gid* and *did* is depending on the network size; however, size of 4 bytes is sufficient for most WSNs [46].

After each IAYK message reaching the sink, the sink's neighbors generate a total kids (TK) message, whose format is shown in Figure 3, and transmit this message. The *gid* field indicates the identifier of the node generating the TK message; the '*kid*' field contains the total number of this node's kids. Upon receiving this message, each node checks if the node identifier value recorded in *gid* field is in its route path or not. If it is, the receiving node records the value contained in the *kid* field, decreases the value of the *ttl* field by 1, and then transmits the message. If the node receives TK more than once, it ignores and discards all the follow-up receptions. The length of the *kid* field depends on the network size. A size of 4 bytes is considered sufficient for most WSN deployment [46].

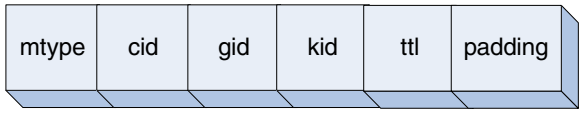
4.1.2. Data transmission task

By the completion of the routing-tree formation task, a packet routing topology that has a tree-like structure with the sink being its root, has been generated. As explained earlier, the closer a node is to the sink (in terms of the number of hops), the larger the number of messages it has to transmit becomes, since it tends to be the traffic forwarder (to the sink) of a larger number of sensor nodes. The traffic-monitoring attacker can easily identify this trend of traffic volumes and from that deduce a well-confined region within which the sink is expected to be located. In order to solve this problem, we introduce a mechanism that has an objective to have all nodes generate equal volume of traffic. This prevents the global attacker from being able to acquire valuable statistical knowledge through traffic analysis.

To achieve this, a node generates two different types of messages: real data message (RDM) and fake data message (FDM). RDMs carry useful information collected by sensors and destined for the sink, while FDMs are generated for confusing the attacker.

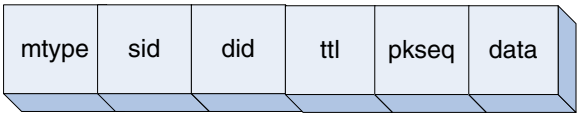
Figure 4 provides the structure of RDM and FDM packets. The *mtype* field contains the code identifying this packet as a 'data' carrying message. For RDM or FDM packets, *mtype* = DATA. *sid* and *did* fields contain the identifiers of the sender and its one-hop forwarder node on the sender's routing path to the sink. The *ttl* field contains the time to live of the particular packet. As the packet passes through a node, its value is decremented by one. When $ttl = 0$, the node drops the packet. The source node sets a value of *ttl* equal to the number of hops to the sink of its route path. The source node sets $ttl = 0$ for FDM packets. '*pkseq*' contains the sequence number that uniquely identifies the specific packet from any other packets sent by the specific source to the sink. Size of 4 bytes is sufficient for the *pkseq* field [46]. The '*data*' field contains the actual data payload if the packet is RDM or a meaningless pattern if it is FDM.

After receiving a packet, node i decrypts it and moves its payload content upwards. If it is a DATA-type message, this node checks *ttl* to find out if it should be forwarded or dropped. If $ttl > 0$, the node places its own identifier in *sid* and its one-hop forwarder node's identifier in *did*, and decreases the value of *ttl* by 1. In order to prevent traffic analysis attack, the receiving node does not forward the RDM immediately. It places this RDM in a buffer containing forwarded RDMs (includes RDMs generated by the node itself and those for which it acts as a relay). In the mean time, this node has to generate an average of $m(i)$ FDMs (in reference to the unit of time). That brings the total average traffic of node i to $m(i) + g_i + f(i)$, where g_i is the average number of RDMs generated by node i and $f(i)$ is the average number of



mtype: message type (TK)
cid: sequence number of route-update cycle
gid: identifier of the node generating the message
kid: number of the node (gid)'s children
ttl: Time to Live
padding: is variable in message size

Figure 3 TK message format.



mtype: message type (DATA)
sid: identifier of the node sending the message
did: identifier of the node receiving the message
ttl:Time to Live
pkseq: sequence of data message
data: data payload

Figure 4 Format of RDM and FDM.

RDMs forwarded by node i, $f(i) = \sum_{\text{node } j \in I_{1-hop}}^{H(i)} g_j$. $H(i)$ represents the number of children node i has, and node j belongs to the set I_{1-hop}, which contains all children of node i, one-hop away from it.

To increase the statistical uncertainty for the attacker, the node can randomize (in accordance to certain distribution or distributions) the inter-departure times of RDM and FDM packets. Running a random generator, the node i selects the message at the head of the RDM queue with probability $p_{\text{RDM}}(i) = \frac{g_i + f(i)}{g_i + f(i) + m(i)}$. If sending of the RDM message is not selected, it generates and transmits a FDM message. If the RDM is selected but the RDM queue is empty, it transmits nothing.

4.2. Energy consumption

Energy consumption for communication is very important for sensors due to limited resources in sensors. Energy consumption for packet transporting in the WSN is in proportion to the distance. The distance to neighbors can increase or decrease the energy consumption of radio channel to transmit a data bit. Besides, the factors like the number of packets and the size of packets are also important in determining the amount of energy consumption of sensors. Therefore, we use a radio energy consumption model [47] which derived the energy consumption of transmit and receive, an L-bit message from different sensors. In this model, during radio operation transmit and receive, circuitry dissipates (each) E_{elec} =50 nJ/bit and the transmit amplifier E_{amp} =100 pJ/bit/m^2. Thus, to transmit an L-bit-long message to a receiving node located at distance d from the transmitter, we have the following amount of energy consumption:

$$E_{T_x}(L, d) = E_{\text{elec}} \times L + E_{\text{amp}} \times L \times d^2 \quad (1)$$

The energy consumed to receive this message equals

$$E_{R_x}(L, d) = E_{\text{elec}} \times L \quad (2)$$

We represent with S_{1-hop} the set containing as members all nodes having one-hop distance from the sink and L_{DATA} the size of RDMs and FDMs (in bits). The energy consumption occurring during the data transmission phase is

$$\text{cost}_{\text{Data}} = (E_{T_x}(L_{\text{DATA}}, d) + E_{R_x}(L_{\text{DATA}}, d)) \times \rho \times \sum_{\text{node } k \in S_{1-hop}} \left[g_k + f(k)\right]^2 \quad (3)$$

In phase 1, each node that receives an RDIS or FRDIS message for the first time should re-broadcast this (RDIS or FRDIS) message. We define as L_{RDIS} (bits) the size of the RDIS and FRDIS messages. We assume the probability of the selected source nodes to initiate transmission of FRDIS messages is μ; thus,

$$\text{cost}_{\text{Topo1}} = (E_{T_x}(L_{\text{RDIS}}, d) + E_{R_x}(L_{\text{RDIS}}, d)) \times \left(\frac{N\pi r_1^2}{S_{\text{whole}}} - 1\right) \times (N + N\mu) \quad (4)$$

where S_{whole} is the size of the deployment area, r_1 is the communication radius of each sensor, and N is the total number of sensors deployed into the area.

In phase 2, each node sends an IAYK message according to the route table. Representing by L_{IAYK} (bits) the size of the IAYK message, we have the following:

$$\text{cost}_{\text{Topo2}} = (E_{T_x}(L_{\text{IAYK}}, d) + E_{R_x}(L_{\text{IAYK}}, d)) \times \sum_{\text{node } j \in S_{1-hop}} [H(j) + 1] \quad (5)$$

Each one-hop node of the sink generates a TK message with the size of L_{TK} bits; thus,

$$\text{cost}_{\text{Topo3}} = (E_{T_x}(L_{\text{TK}}, d) + E_{R_x}(L_{\text{TK}}, d)) \times \left(\frac{N\pi r_1^2}{S_{\text{whole}}} - 1\right) \times (N - 1) \quad (6)$$

Therefore, $\cos t_{\text{Topo1}}$, $\cos t_{\text{Topo2}}$, and $\cos t_{\text{Topo3}}$ are occurring during the topology discovery, which is performed infrequently; thus, energy consumption $\cos t \cong \cos t_{\text{Data}}$.

4.3. Performance simulation analysis

We evaluated the performance of our approach, the conventional scheme, and the CSL scheme [34] through simulation using OPNET [48]. Note that by the term 'conventional scheme' we mean a WSN that is not transmitting FDMs; thus, it is highly vulnerable to traffic analysis attacks. All simulation parameters are listed in Table 3. We performed 100 simulation runs with different seeds and calculated average values.

4.3.1. Energy consumption and network life

Figure 5 shows how the total energy consumed by sensors is impacted by the simulation time (there is no dead node during the simulation time). The CSL scheme has the highest energy consumption compared to ATA scheme and conventional scheme.

As shown in Figure 5, the conventional scheme has lower energy consumption. It is thus self-evident that the average node lifetime of conventional WSN deployments will be longer compared to ATA scheme. However, nodes at one-hop distance from the sink have to act as traffic forwarders of nodes that are not in the sink's hearing distance; thus, they end up dying considerably earlier compared to the rest of the nodes. When all one-hop nodes die, WSN is dead as well despite the fact many nodes remaining alive, since the collected information cannot be passed to the

Table 3 Simulation configurations

Parameter	Value
Number of nodes	50
Deployment area	250 m × 250 m
Channel bandwidth (kHz)	22.000
Channel data rate	11 Mbps
Short retry limit in MAC layer	7
Long retry limit in MAC layer	4
Wireless protocol	IEEE 802.11
MAC layer buffer size (in bits)	256,000
Data payload (in bytes)	1,024
Packet generation at start time (s.)	5.0
Inter arrival time of packet (s.)	Constant (1.0)
Communication range	40 m
E_{elec}	50 nJ/bit
E_{amp}	100 pJ/bit/m^2

sink. A better assessment of WSN's life duration is to look at the times the sink's neighbor dies (in which case information routed through that node does not reach the sink and becomes lost). In the following results, the initial energy stored in each node is 1.5 J; the simulation ends up until all nodes are dead. We start presenting results for the time it takes to have the last 'dead' neighbor of the sink (to be referred as *last-to-die-sink-neighbor (LtoDSN)*) for the conventional scheme, ATA scheme, and CSL scheme. They are shown in Figure 6. *X* axis corresponds to the time interval between the starting times of two successive topology discovery cycles. We see a gradual increase of the LtoDSN life duration as the frequency of refreshing the topology discovery is increased. For ATA and CSL, this change is modest whereas for the conventional scheme is quite significant. This dependency is due to the following

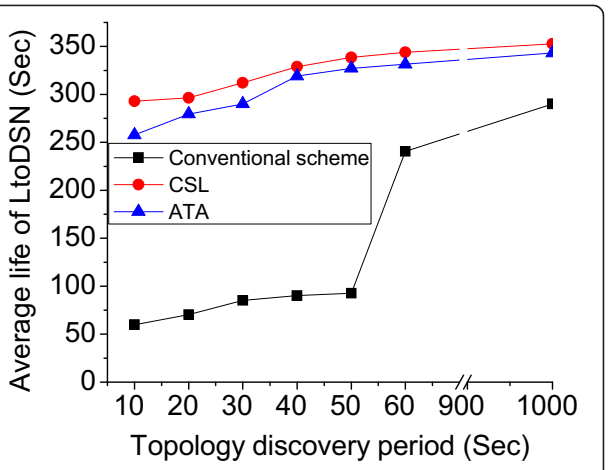


Figure 6 Average life of LtoDSN.

reason. In-between two consecutive topology discovery refreshing events, some nodes with distance longer than one-hop from the sink become drained and die. Packets sent by nodes whose routing path to the sink passes through such node will be terminated and will not reach the sink's neighbor; thus, the sink's neighbors have the lower traffic volume to forward to the sink, resulting in conservation of energy. However, when a new cycle for the topology discovery is running, new routing paths will be formed, enabling again those nodes that have been cut off from the sink to reconnect, placing again the one-hop from sink nodes at a position of having to forward a higher traffic volume of packets, thus draining themselves faster.

Figure 7 shows the progression of node depletion versus time. We define 'active node' that this node is alive itself and its parents still are alive. We do not refresh the routing tree in the simulation. Compared to CSL, ATA loses its first node approximately 27 s later than the CSL scheme (about 15 s) and this trend continues while the

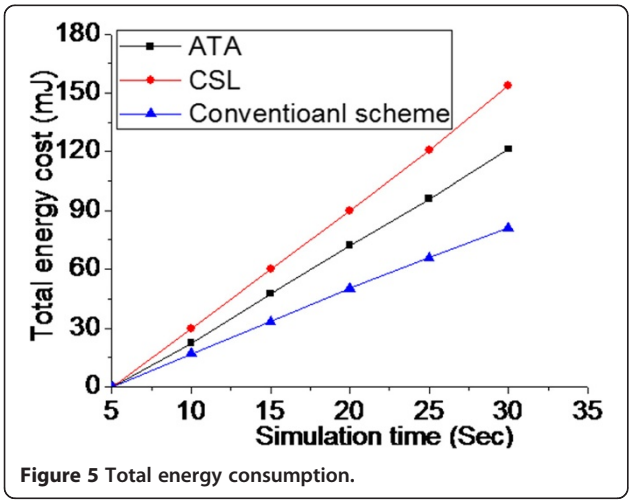


Figure 5 Total energy consumption.

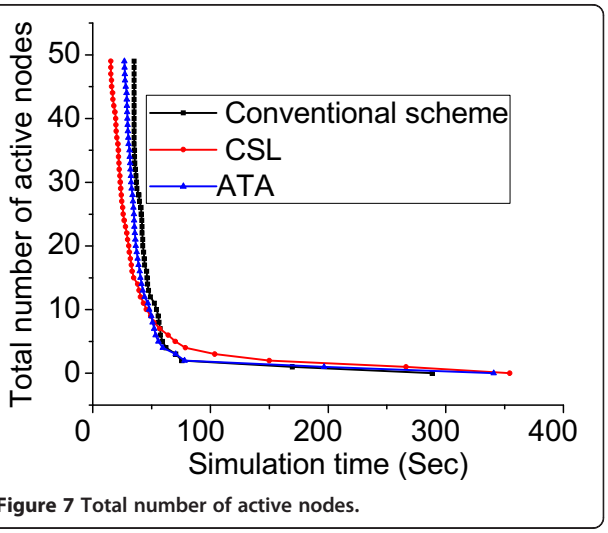


Figure 7 Total number of active nodes.

active nodes are in range of 49 to 10. This advantage comes with the very important additional benefit of having a secure scheme (ATA) capable of neutralizing the danger of traffic volume monitoring attack. Compared to the conventional scheme, the time duration when using ATA or CSL has close values, and in any case, the closeness between the two curves remains close for the whole range of time display. That means fake messages in ATA scheme have no huge influence on the average life of active node.

A way to understand the coverage ability WSN has and how it changes in time is to determine the number of packets the sink receives within a set time window, as it moves in time. Figure 8 is doing just that for three schemes using a window of 8 s. The values shown at 16 s indicate the number of RDM packets the sink received within the time interval [(16-8) = 8 s; 16 s], and goes on. The conclusions are the following: the CSL scheme has lowest RDM delivery up until 64 s, and afterwards, the trend reverses. The gap of RDM delivery between ATA and conventional scheme is very small, especially after 64 s.

4.3.2. Security

An area with a dimension of 240 m × 240 m is assigned as the area monitored by the global attacker. A grid is formed within this area, segmenting it into smaller squares of 20-m length each. Traffic-monitoring devices (passing the information to the attacker) are placed at the locations [$x = 20$ m × j; $y = 20$ m × k] where $1 \le j \le 12$, $1 \le k \le 12$. The sink is at location [$x = 120.0$ m; $y = 120.0$ m]. The simulation time is 1,000 s, the topology discovery is performed at the beginning and remains unchanged, and enough energy is stored in each node to remain alive for the duration of the simulation. Let P_i be the average traffic volume measured by the monitoring device i divided by the summation of traffic volume measured and reported by all monitoring devices. An attacker is trying to identify the location (or at least the region) where the largest volume of traffic volume is occurring and will conclude the sink should be located within that area. In an approximate sense, it can be associated with the probability that the sink is located within the area the monitoring device i operates and can provide a measure for the degree of certainty of this been the case.

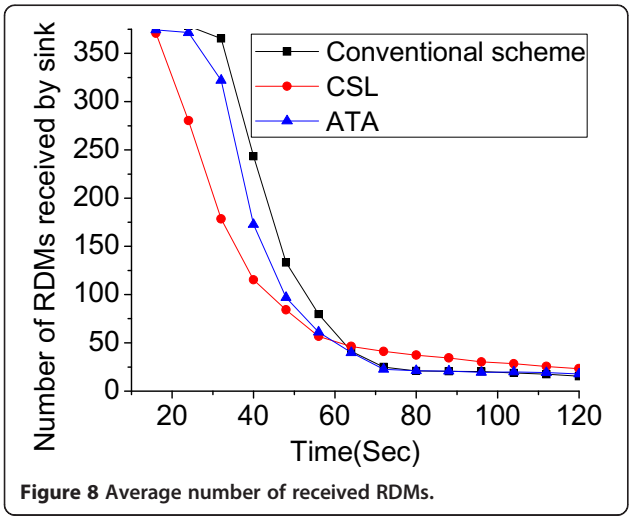


Figure 8 Average number of received RDMs.

Figure 9a, b, c provides the values of P_i at different locations of the monitored area for the conventional scheme, ATA scheme, and CSL scheme. X axis and Y axis define the plane on which sensor nodes and traffic-monitoring devices are deployed. The perpendicular - to X and Y - plane axis displays the value of P_i. It is evident that in the case of the conventional system, P_i peaks at the sink's location. Also, the strength increases (almost) consistently, as we move from the boundaries of the monitored area, where the sink is located. On the contrary, in the case of ATA, the distribution of P_i appears to be close to uniform, not allowing the attacker to develop confidence in terms of the sink's location. In the case of CSL, P_i's distribution is considerably less spread out compared to the ATA scheme, thus reducing significantly uncertainty.

5. Preserving sink's location: an improved ATA scheme

There is a trade-off between information delivery performances, energy efficiency, and location privacy by using dummy traffic to hide the real sink's location. If all packets are real event packets, the communication/computation cost will be lower; however, it will be very easy for a global attacker to trace the packets. If we make all nodes having the traffic volume using dummy traffic, it will significantly increase the network. Our goal is to minimize the network traffic while to guarantee the real sink (RS)'s location privacy.

To address this problem, we propose an improved ATA (IATA) scheme. In this IATA, we select several nodes to act as 'fake sinks' (FS) and emulate traffic patterns similar to the RS, in order to confuse the global attacker. We take into account the case that nodes deployed into an area $X \times Y$ according to a grid network. We define as 'Mixnode' each node that satisfies the following conditions: (i) it is located within any circle having as center a FS or the RS and radius $r_1 \times \varepsilon$, where r_1 is the sensor's communication range and ε is a positive integer; (ii) it produces traffic volume equal to TPN_1 (TPN_1 is the largest amount of traffic volume among the traffic volumes generated by the real sink's neighbors). We define as UnMixnode each node that: (i) it is not located within any circle having as center a FS or the RS and radius $r_1 \times \varepsilon$; (ii) it produces traffic volume equal to $\mathrm{TPN}_{\varepsilon+1}$ ($\mathrm{TPN}_{\varepsilon+1}$ is the largest amount of

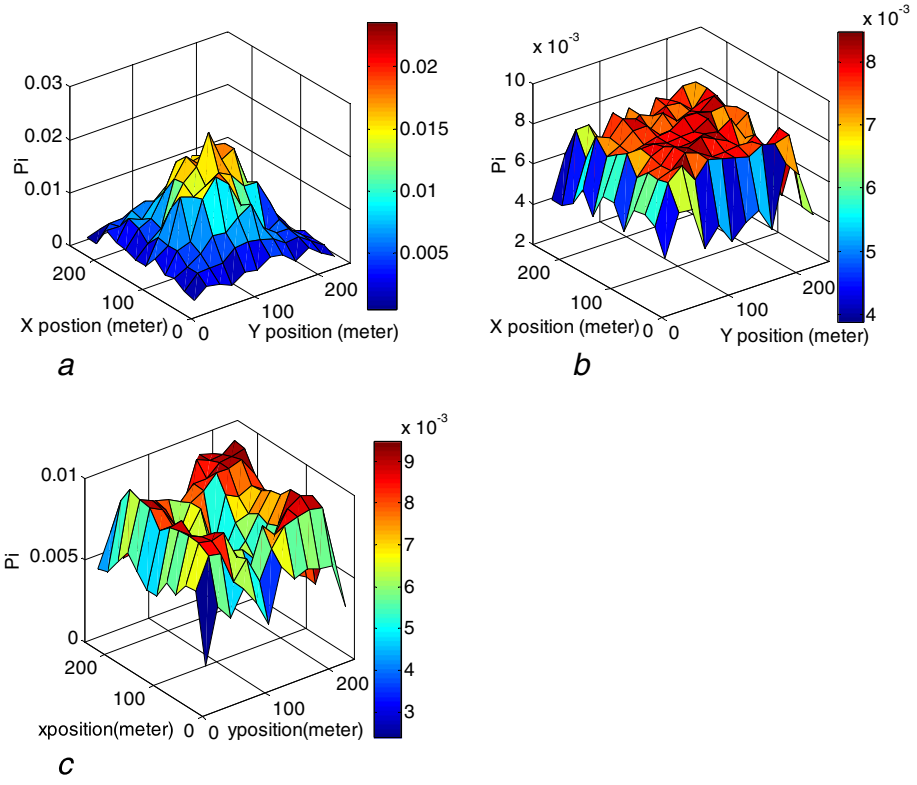


Figure 9 Conventional scheme (a), ATA scheme (b), CSL scheme (c).

traffic volume among the traffic volumes generated by nodes from ($\varepsilon + 1$) hops from the real sink).

5.1. Fake sink's placement

Deploying fake sinks at the right locations is crucial to the network's performance in terms of RDM end-to-end delay, RDM delivery ratio, and sink's location privacy. For example, if all fake sinks are the real sink's neighbors as well, or they are deployed close to each other, or the radius $r_1 \times \varepsilon$ is small, the scheme would not work and the global attacker will be able to determine the sink's location. This is because it has the fewer number of nodes which generate fake messages when the value of $r_1 \times \varepsilon$ is small, and the attacker can guess the sink's location with a high probability under having a knowledge of the area where the real sink is. If the circle having center RS and radius $r_1 \times \varepsilon$ is large enough to cover the whole sensor deployment area, it has the largest traffic volume of the network since all nodes located in the circle produce traffic volume equal to TPN_1. Since heavy traffic volume increases RDM end-to-end delay and RDM delivery ratio, our objective is to minimize the network's traffic along with maintaining the real sink's location privacy. The optimization criterion used in the selection of the FS locations is formed to materialize the abovementioned objective.

We express with N_{in} the number of Mixnodes and with N the total number of nodes in the network. The total traffic volume of the network, TV, can be

$$\mathrm{TV} = N_{in}\mathrm{TPN}_1 + (N - N_{in})\mathrm{TPN}_{\varepsilon+1} \tag{7}$$

The goal of the attacker is to identify a set C_{TV} of devices that represent the set of possible locations (or regions) of the sink. Hence, the goal of the attacker is to discover the sink's location (or region). This set indicates that the attacker believes that the objects being observed are close to some of the devices in C_{TV}. Since Mixnodes have the same (and maximum) traffic volume, thus, by observing the traffic volume, the attacker cannot distinguish if a Mixnode is a sink or not. Assume that the total number of these regions is N_{rn}, then the probability of any region member of C_{TV} including the location of the real sink is $1/N_{rn}$, which reflects the level of uncertainty the attacker has in terms of identifying correctly a region of the sink. We define and use as a privacy measure of the SPP scheme the entropy

$$\mathrm{ent} = \sum_{C_{\mathrm{TV}}} -\frac{1}{N_{rn}} \log_2 \frac{1}{N_{rn}} = \log_2 N_{rn} \tag{8}$$

Define $\Omega(k, j)$ be the number of hops from the fake sink k to the fake sink j and $h_{\max}$ be the maximum

number of hops between all nodes and the real sink. ε should be satisfied: (i) To avoid the case without any fake sink, ε should be less than (equal to) $h_{\max}/2$. (ii) To make sure that there has been no overlapped area among the circles with center FS or RS and radius $r_1 \times \varepsilon$, $\Omega(k,j)$ should be larger than (equal to) 2 ε, where $0 < k, j \le N_{rn} - 1$. (iii) To make sure FSs should be located in this deployment area, we have to satisfy this condition of $\min(X - x(j), x(j), Y - y(j), y(j))/r_1 \ge \varepsilon$, where $(x(j), y(j))$ is the sink j's position. Our goal is to provide a method that enables to minimize traffic volume while to maintain the location privacy larger than (equal to) a threshold δ. Therefore, we use a metric of

$$\eta = \min(\mathrm{TV}) \tag{9}$$

$$s.t. \begin{cases} \mathrm{TV} > 0 \\ \mathrm{ent} > \delta \\ \min(X - x(j), x(j), Y - y(j), y(j))/r_1 \ge \varepsilon \\ \dfrac{\Omega(k,j)}{2} \le \varepsilon \le \dfrac{h_{\max}}{2} \\ 0 < k, j \le N_{rn} - 1 \end{cases}$$

From Equation 9, the total number of fake sinks and their fake sinks' positions depends on the values of ent and TV, conditions of $\min(X - x(j), x(j), Y - y(j), y(j))/r_1 \ge \varepsilon$, and $\frac{\Omega(k,j)}{2} \le \varepsilon \le \frac{h_{\max}}{2}$. However, the condition of $\frac{\Omega(k,j)}{2} \le \varepsilon \le \frac{h_{\max}}{2}$ is linked with the fake sinks' positions, and the total traffic volume of the network depends on the number of fake sinks directly. Thus, the fake sink's placement problem is NP-hard.

Theorem 1: The fake sinks' placement problem is NP-hard.

Proof: we prove the NP-hardness of the fake sinks' placement problem by reducing the well-know knapsack problem defined as follows:

The knapsack problem [49]**:** Given a set of items z_1, z_2, ..., z_n, each item z_i with a weight w_i and a value v_i, the maximum weight that we can carry in the bag is W. Thus, the objective is

$$\text{maxmizes} \quad \sum_{i=1}^{n} v_i, \quad \text{subject to} \quad \sum_{i=1}^{n} w_i \le W \tag{10}$$

We create a fake sink i as an item z_i, and traffic volume generated by Mixnodes located in the range of the fake sink i as a v_i. The total entropy should not be less than δ. Position changes of fake sinks may cause recalculations for traffic volume, numbers of fake sinks, and so on. Thus, we have to minimize the total traffic volume given the condition of the entropy that is $\ge \delta$. It is easy to see that the fake sinks' placement problem is in NP class as the objective function associated with a given solution can be evaluated in a polynomial time. Thus, we conclude that this fake sinks' placement problem is NP-hard [14]. We give the following algorithm of placing fake sinks, where hop(i) is the number of hops forming the routing path from node i to the real sink, $\Phi(i,j)$ is the number of hops from node i to the fake sink j, node i' s position is $(x(i), y(i))$. Lines 4 to 11 try to find the first fake sink which satisfies the condition of hop(i) $\ge d$ and $\min(X - x(i), x(i), Y - y(i), y(i))/r_1 \ge \varepsilon$. Lines 12 to 27 try to find the other fake sinks in the deployment area. Line 28 computes the multiobjective optimization value. Line 31 gives the best placements of fake sinks.

Algorithm 1. Fake sinks' placement

1: For $\varepsilon = 1 : h_{\max}/2$
2: Compute TPN_1 & $TPN_{\varepsilon+1}$
3: For $d = 2\varepsilon : h_{\max}/2 - \varepsilon$
4: For i=1:N
5: NumFS=0;
6: If $hop(i) \ge d$ & $\min(X - x(i), x(i), Y - y(i), y(i))/r_1 \ge \varepsilon$
7: Store node i as a fake sink;
8: NumFS++;
9: Break;
10: End
11: End
12: While(flag)
13: For i=1:N
14: If $hop(i) \ge d$ & all $\Phi(i,j) \ge d$, $j \in [1, NumFS]$
15: and $\min(X - x(i), x(i), Y - y(i), y(i))/r_1 \ge \varepsilon$
16: If node i is not in the storage
17: NumFS++;
18: flag=1;
19: break;
20: End
21: End
22: else if(i==N)
23: flag=0;
24: End
25: End
26 End
27 End while
28: Computer: ent and TV
29: End
30: End
31: $\eta = \min(TV)$ given the conditions of $TV > 0$, $ent > \delta$, $\min(X - x(j), x(j), Y - y(j), y(j))/r_1 \ge \varepsilon$ and $\frac{\Omega(k,j)}{2} \le \varepsilon \le \frac{h_{\max}}{2}$
32: record " ε and d".

Figure 10 gives an example of optimal fake sinks' placement. Two hundred fifty-six nodes are uniformly deployed into an area with 500 m × 500 m, and a real sink is located at position (250 m, 250 m). The communication range is 30 m. We can see that there are four fake sinks whose positions are approximately (105 m, 105 m), (345 m, 405 m), (105 m, 375 m), and (375 m, 105 m).

5.2. Details of IATA scheme

The IATA protocol includes two phases: topology discovery phase and data transmission phase. In the topology discovery phase, a routing path of each node which connects the node to the sink is setup by the ATA method. After the topology discovery phase, a topology map has been generated which has a tree-like structure with the sink being the root. Besides, each node knows the route path, the number of its children, the total number of

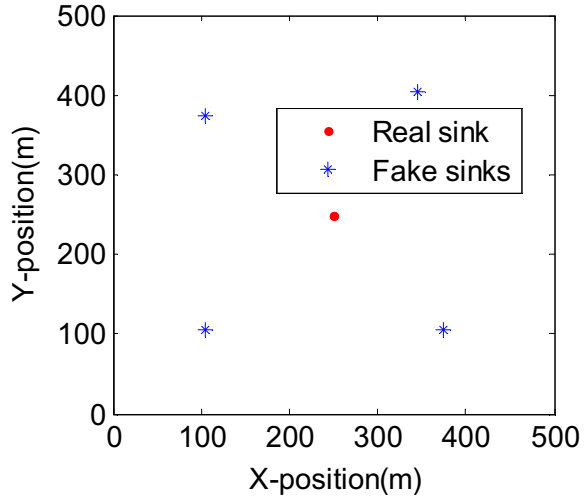


Figure 10 Optimal fake sinks' placement.

children of sink's neighbors, the fake sinks' positions, and ε. In addition, each node can check if it is Mixnode or UnMixnode by knowing the fake sinks' positions and ε.

The process of transmitting data messages is as follows.

(1) If one Mixnode MN_i receives RDMs and FDMs, it will forward RDMs (including RDMs generated by itself and forwarded RDMs) according to the route, and discard FDMs. In order to prevent the attacker from drawing conclusions by identifying and tracing successive transmissions, which might eventually lead the attacker to the sink, the Mixnode does not forward the RDMs immediately. It places RDM in a buffer containing nonprocessed yet RDM (includes RDM generated by the node itself and those for which it acts as a relay).

(2) Then, the Mixnode MN_i generates $m(MN_i)$ FDMs and sends them according to the route path. Let $H(MN(i))$ represent the number of kids Mixnode MN_i has; we have $m(MN_i) = TPN_1 - g_{MN_i} - f(MN_i)$, where g_{MN_i} is an average number of RDMs generated by Mixnode MN_i, $f(MN_i)$ is an average number of RDMs forwarded by Mixnode MN_i, $f(MN_i) = \sum_{\text{node } j \in MNI_{1-hop}}^{H(MN(i))} g_j$, and node j belongs to a set MNI_{1-hop} of children of Mixnode.

(3) If one UnMixnode UMN_i receives RDM and FDM, it will forward RDMs (including RDM generated by himself and forwarded RDM) according to the route, and discard FDMs. Then, UMN_i generates $m(UMN_i)$ FDMs and sends them according to the route path. Note that $m(UMN_i) = TPN_{\varepsilon+1} - g_{UMN_i} - f(UMN_i)$, where g_{UMN_i} is an average number of RDMs generated by UnMixnode UMN_i, $f(UMN_i)$ is an average number of RDMs forwarded by UnMixnode UMN_i, $H(UMN(i))$ represent the number of kids UnMixnode UMN_i has $f(UMN_i) = \sum_{\text{node } j \in UMNI_{1-hop}}^{H(UMN(i))} g_j$, and node j belongs to a set $UMNI_{1-hop}$ of children of UnMixnode UMN_i.

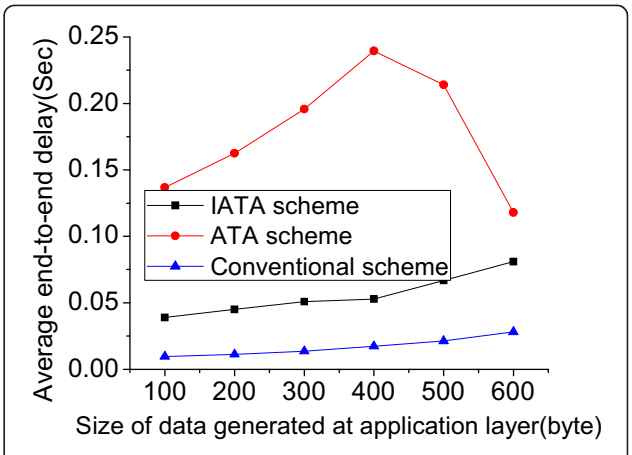


Figure 11 End-to-end delay.

6. Performance evaluation

We evaluate the performance of our ATA, IATA, and the conventional scheme, through simulation using OPNET. The reported results correspond to Table 3, with one sink and 256 nodes which are deployed into vertexes into a grid of 30-m length each, and the radius of each node is 30 m.

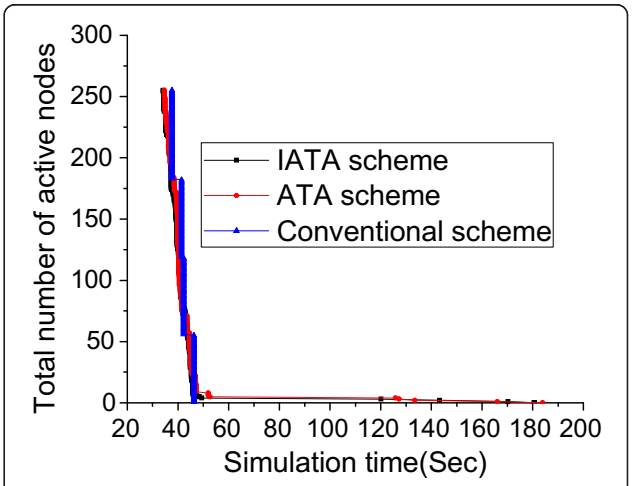


Figure 12 Total number of active nodes in three schemes.

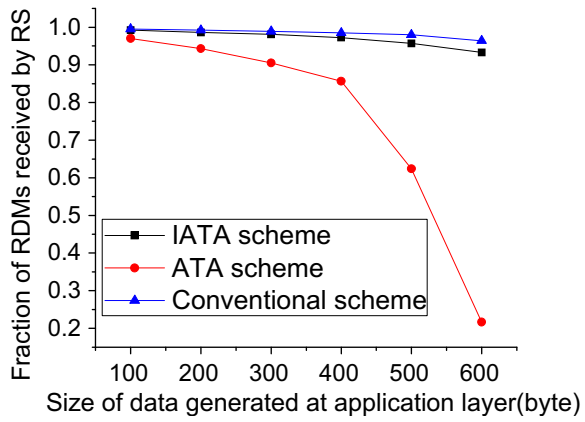


Figure 13 Fraction of RDMs received by the real sink.

Figure 11 illustrates the behavior of the average RDM delivery time (average end-to-end delay) as the size of data packet generated at the application layer. Each node generates a data packet per interval time 1.0 s. Enough energy has been provided to all nodes, to ensure all of them remain alive for the duration of the simulation. We can see that the delivery delay is increasing with the size of a data packet in IATA and conventional scheme. On the contrary, delivery delay in the ATA scheme increases until the size of a data packet is 400 bytes and then decreases quickly. The reason for this is that when the size of a data packet is larger than 400 bytes, ATA has a higher packet loss in the MAC layer, which leads to those successful delivery packets reducing waiting time in the MAC layer buffer. Compared to ATA, IATA has a significantly better performance that keeps the values of average end-to-end delay 0.08 s.

Figure 12 shows the progression of node depletion versus time. Initial energy stored in each node is 0.5 J; the simulation ends up until all nodes are dead. Each sensor node generates a data packet of size 100 bytes at its application layer every second. Curves in all the three schemes are very close, and the total number of active nodes decreases sharply during a period [35 s 42 s]. That means fake messages in ATA scheme and IATA scheme have no any influence on the average life of active node.

Figure 13 displays the fraction of RDMs received by the real sink. There is no dead node during the simulation time of 1,000.0 s. The delivery ratio in IATA and the conventional scheme keeps at least 0.95, while in ATA scheme, the fraction of RDMs received by the real sink decreases sharply after the size of data packet 400 bytes. The main reason is that ATA generates too many fake messages, which leads to poor performances such as delivery delay and delivery ratio.

An area with dimensions 500 m × 500 m is assigned as the area monitored by the global attacker. A grid is

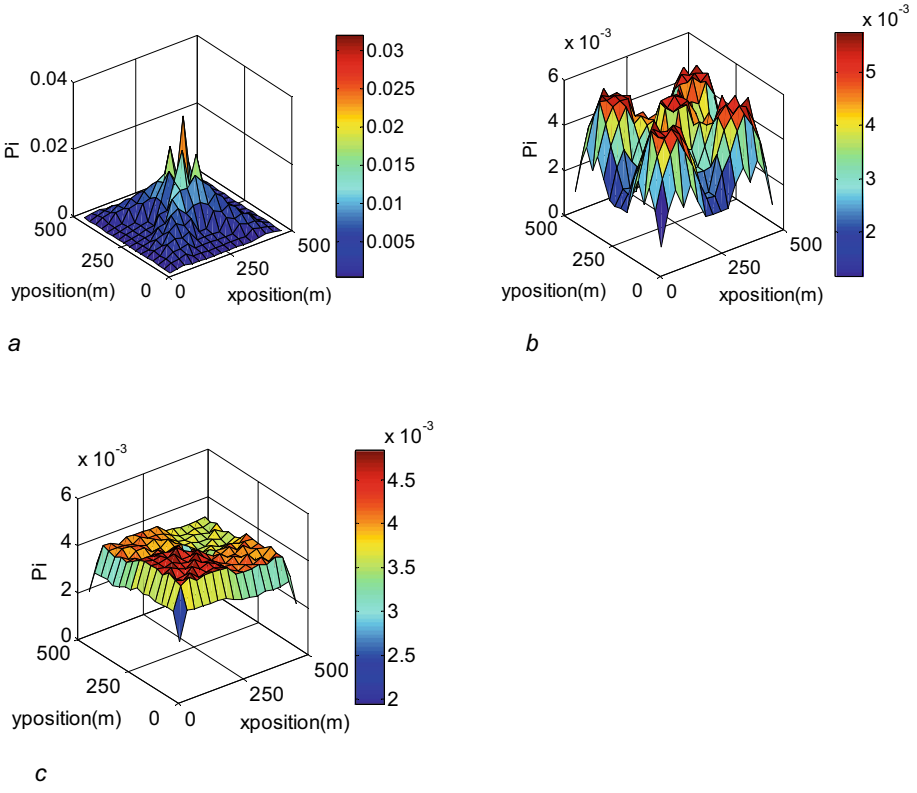


Figure 14 Conventional scheme (a), IATA (b), ATA (c).

formed within this area, segmenting it in smaller squares of 30-m length each. Traffic-monitoring devices (passing the information to the attacker) are placed at the centers of grids. Figure 14 displays the values of P_i at different locations of the monitored area. Figure 14a, b, c provides results for the conventional scheme, IATA scheme, and ATA scheme. X axis and Y axis define the plane on which sensor nodes and traffic-monitoring devices are deployed. The perpendicular - to X and Y - plane axis displays the value of P_i. It is evident that in the case of the conventional system, P_i peaks at the sink's location. On the contrary, in the case of IATA, the distribution of P_i appears to uniform within five different regions, which will enhance significantly uncertainty. In the case of ATA, the highest values of P_i in the region of 250 m × 250 m are distributed uniformly.

7. Conclusions

Sink is the connecting point of the sensor network with the entity making use of its collected results; thus, the ability of the sink to be capable to receive collected information is very crucial. In this paper, after analyzing the sink's location privacy problem, we firstly describe and analyze a new ATA scheme aiming at concealing the sink's location by using fake message injection. Then, we design an improved ATA (IATA) scheme where some nodes are selected to act the fake sinks, and sensors around fake sinks generate dummy messages and discard received dummy messages. Performance analysis of the ATA scheme can protect the sink's location privacy, and the IATA scheme can reduce traffic volume.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the Natural Science Foundation of Zhejiang Province LQ13F010001, Y201328392, National Natural Science Foundation of China 61301142, and SRF for ROCS, SEM (2013[1792]).

Author details

[1]School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China. [2]School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa K1N 6N5, Canada.

Received: 25 November 2013 Accepted: 21 July 2014
Published: 13 August 2014

References

1. J. Deng, R. Han, S. Mishra, Countermeasures against traffic analysis attacks in wireless sensor networks, in *Proceedings of IEEE/Create Net International Conference on Security and Privacy for Emerging Areas in Communication Networks (SecureComm)* Athens, Greece, 2005), p. 113
2. C. Ozturk, Y. Zhang, W. Trappe, Source-location privacy in energy-constrained sensor network routing, in *Proceedings of 2004 ACM Workshop on Security of Ad Hoc and Sensor Networks, Washington D.C. USA*, 2004, p. 8
3. Y. Jian, S. Chen, Z. Zhang, L. Zhang, Protecting receiver location privacy in wireless sensor networks, in *Proc. of IEEE Infocom 2007, Anchorage, Alaska, USA, 2007*, 1955
4. P. Kamat, Y. Zhang, W. Trappe, C. Ozturk, Enhancing source location privacy in sensor network routing, in *Proc. of IEEE ICDCS, Columbus, Ohio, USA, 2005* p. 599
5. X. Wu, J. Liu, X. Hong, E. Bertino, Achieving anonymity in mobile ad hoc networks using fuzzy position information: mobile Ad-hoc and sensor networks. Lect Notes Comput Sci **4325**, 461–472 (2006). doi: 10.1007/11943952_39
6. Q. Gu, X. Chen, Z. Jiang, J. Wu, *Sink-anonymity mobility control in wireless sensor networks (IEEE International Conference on Wireless and Mobile Computing* (Morocco, Networking and Communications, Marrakech, 2009). p. 36
7. R. Shokri, A. Nayyeri, N. Yazdani, *P Papadimitratos Efficient and adjustable recipient anonymity in mobile ad hoc networks (IEEE International Conference on Mobile Ad hoc and Sensor Systems* Pisa, Italy, 2007). p. 1
8. A. Perrig, R. Szewczyk, J.D. Tygar, V. Wen, D.E. Culler, Spins: security protocols for sensor networks. Wirel. Netw **8**(5), 521–534 (2002). doi: 10.1023/A:1016598314198
9. L. Zhou, H.-C. Chao, Multimedia traffic security architecture for internet of things. IEEE Netw. **25**(3), 35–40 (2011). doi: 10.1109/MNET.2011.5772059
10. G. Han, J. Jiang, W. Shen, L. Shu, J. Rodrigues, IDSEP: a novel intrusion detection scheme based on energy prediction in cluster-based wireless sensor networks. IET Information Security, Institution of Engineering and Technology (IET) **7**(2), 97–105 (2013). doi:10.1049/iet-ifs.2012.0052
11. B. Gedic, L. Liu, Location privacy in mobile system: a personalized anonymization model, in *Proc. of 25th IEEE International Conference on Distributed Computing Systems, Columbus, OH, USA*, 2005, p. 620
12. B. Bamba, L. Liu, P. Pesti, T. Wang, Supporting anonymous location queries in mobile environments with privacy grid, in *Int'l Conf. World Wide Web WWW'08* Beijing, China, 2008), p. 237
13. L. Sweeney, K-anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems **10**(5), 557–570 (2002). doi: 10.1142/S0218488502001648
14. Y. Yang, M. Shao, S. Zhu, B. Urgaonkar, G. Cao, Towards event source unobservability with minimum network traffic in sensor networks, in *Proc. of ACM WiSec, Alexandria, Virginia, USA*, 2008, p. 77
15. Y. Xi, L. Schwiebert, W. Shi, *Preserving source location privacy in monitoring-based wireless sensor networks, in 20th International Conference on Parallel and Distributed Processing Symposium* Rhodes Island Greece, 2006)
16. Y. Li, J. Ren, J. Wu, Quantitative measurement and design of source-location privacy schemes. IEEE Transaction on Parallel and Distributed Systems **3**, 7 (2012). doi: 10.1109/TPDS.2011.260
17. Y. Li, J. Li, J. Ren, J. Wu, *Providing hop-by-hop authentication and source privacy in wireless sensor networks (IEEE INFOCOM2012* (USA, Orlando, FL, 2012). p. 3071
18. M.M.E.A. Mahmoud, X. Shen, A cloud-based scheme for protecting source-location privacy against hotspot-locating attack in wireless sensor networks. IEEE Transaction on Parallel and Distributed Systems **23**(10), 1805–1818 (2012). doi: 10.1109/TPDS.2011.302
19. J. Ren, D. Tang, *Combining source-location privacy and routing efficiency in wireless sensor networks (2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)* Houston, Texas, USA, 2011). p. 1
20. A. Gurjar, A.R.B. Patil, *Cluster based anonymization for source location privacy in wireless sensor network (2013 International Conference on Communication Systems and Network Technologies (CSNT)* Gwalior, India, 2013). p. 248
21. P. Spachos, S. Liang, F.M. Bui, D. Hatznakos, *Improving source-location privacy through opportunistic routing in wireless sensor networks (2011 IEEE Symposium on Computers and Communications (ISCC)* Kerkyra, Greece, 2011). p. 815
22. S. Sivashankari, M.M. Raseen, A framework of location privacy and minimum average communication under the global eavesdropper, in *2013 International Conference on Current Trends in Engineering and Technology* (ICCTET, Coimbatore, 2013), p. 392
23. B. Ying, D. Makrakis, H.T. Mouftah, *A protocol for sink location privacy protection in wireless sensor networks (2011 IEEE Global Communications Conference* Houston, TX, USA, 2011). p. 1
24. A.A. Nezhad, D. Makrakis, A. Miri, *Anonymous topology discovery for multi-hop wireless sensor networks (3rd ACM International Workshop on QoS and Security for Wireless and Mobile Networks* Greece, Crete Island, 2007). p. 78
25. K. Bicakci, I.E. Bagci, B. Tavli, Lifetime bounds of wireless sensor networks preserving perfect sink unobservability. IEEE Commun. Lett. **5**(2), 205–207 (2011). doi: 10.1109/LCOMM.2011.010311.101885
26. K. Mehta, D. Liu, M. Wright, Protecting location privacy in sensor networks against a global eavesdropper. IEEE Transaction on Mobile Computing **11**(2), 320–336 (2012). doi: 10.1109/TMC.2011.32
27. U. Acharya, M. Younis, Increasing base-station anonymity in wireless sensor networks. Journal of Ad-hoc Network **8**(8), 791–809 (2010). doi: 10.1016/j.adhoc.2010.03.001

28. Y. Ebrahimi, M. Younis, Using deceptive packets to increase base-station anonymity in wireless sensor network, in *Proc. of 7th International Conference on Wireless Communication and Mobile Computing Conference, Istanbul, Turkey*, 2011, p. 842
29. Y. Ebrahimi, M. Younis, Increasing transmission power for higher base-station anonymity in wireless sensor network, in *The proc. of the IEEE International Conference on Communications (ICC 2011)* Kyoto, Japan, 2011), p. 1
30. H. Park, S. Song, B.Y. Choi, C.T. Huang, Passages: preserving anonymity of sources and sinks against global eavesdroppers, in *Proc. on IEEE Infocom, Turin, Italy, 2013* pp. 210–214
31. X. Li, X. Wang, N. Zheng, Z. Wan, Enhanced location privacy protection of base station in wireless sensor networks, in *Proc. of 5th International Conference on Mobile Adhoc and Sensor Networks, Fujian, China, 2009* p. 457
32. J. Yao, Preserving mobile-sink-location privacy in wireless sensor networks, in *Proc. of 2nd International Workshop on Database Technology and Applications, Wuhan, China, 2010* p. 1
33. H. Chen, W. Lou, From nowhere to somewhere: protecting end-to-end location privacy in wireless sensor networks, in *Proc. of 29th International Conference on Performance Computing and Communications Conference, Hong Kong, China, 2010*. pp. 1–8
34. B. Ying, J.R. Gallardo, D. Makrakis, H.T. Mouftah, Concealing of the sink location in WSNs by artificially homogenizing traffic intensity, in *IEEE Infocom 2011, Shanghai, China*, 2011, p. 1005
35. N. Li, N. Zhang, S.K. Das, B. Thuraisingham, Privacy preservation in wireless sensor networks: a state-of-the-art survey. Ad Hoc Netw. **7**(8), 1501–1514 (2009). doi: 10.1016/j.adhoc.2009.04.009
36. O. Kiraz, A. Levi, *Maintaining trajectory privacy in mobile wireless sensor networks (2013 IEEE Conference on Communications and Network Security (CNS)* Washington, DC, USA, 2013). p. 401
37. L. Kang, *Protecting location privacy in large-scale wireless sensor network (IEEE ICC '09* Dresden, Germany, 2009). p. 1
38. C. D'iaz, B. Preneel, Taxonomy of mixes and dummy traffic. Information Security Management, Education and Privacy, IFIP International Federation for Information Processing **148**, 217–232 (2004). doi: 10.1007/1-4020-8145-6_18
39. Blue Radios Inc, Order and price info. http://www.digikey.cn/cn/zhs/techzone/wireless/supplier/BlueRadios_Inc__822.html?WT.z_ref_page_id=21. Accessed in February 2006
40. D. Niculescu, B. Nath, *Adhoc positioning system (APS) using AoA, in Proceedings of IEEE INFOCOM* San Francisco, CA, USA, 2003). p. 1734
41. A. Savvides, C. Han, M. Srivastava, *Dynamic fine-grained localization in ad-hoc networks of sensors, in Proceedings of ACM MobiCom* Rome, Italy, 2001). p. 166
42. L. Eschenaur, V. Gligor, A key-management scheme for distributed sensor networks, in *Proc. 9th ACM Conference on Computer and Communications Security, Washington, DC, USA*, 2002
43. C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, F. Silva, Directed diffusion for wireless sensor networking. IEEE/ACM Transactions on Networking (TON) **11**(1), 2–16 (2003). doi: 10.1109/TNET.2002.808417
44. K. Oikonomou, I. Stavrakakis, Performance analysis of probabilistic flooding using random graphs, in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, WoWMoM 2007. Espoo, Finland, June 2007*, 2007, p. 1
45. N.B. Chang, M.Y. Liu, Controlled flooding search in a large network. IEEE Transaction on Networking **15**(2), 436–449 (2007). doi: 10.1109/TNET.2007.892880
46. N. Sastry, D. Wagner, Security considerations for IEEE 802.15.4 networks, in *Wise'04: Proceedings of the 3rd ACM Workshop on Wireless Security, Philadelphia, PA, USA, 2004*, p. 32
47. W.T. Heinzelman, A. Chandrakasan, H. Balakrishnam, Energy-efficient communication protocol for wireless microsensor network, in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, vol. 2, Hawaii, USA*, 2000, p. 1
48. Opnet. http://www.opnet.com
49. E.D. Kamin, A parallel algorithm for the knapsack problem. IEEE Transaction on Computers **33**(5), 404–408 (1984). doi: 10.1109/TC.1984.1676456

doi:10.1186/1687-1499-2014-131
Cite this article as: Di Ying *et al.*: **Anti-traffic analysis attack for location privacy in WSNs.** *EURASIP Journal on Wireless Communications and Networking* 2014 **2014**:131.