MAINTAINING TRAJECTORY PRIVACY IN MOBILE WIRELESS SENSOR NETWORKS

by
OSMAN KİRAZ

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University

August 2012

MAINTAINING TRAJECTORY PRIVACY IN MOBILE WIRELESS SENSOR NETWORKS

APPROVED BY

Assoc. Prof. Dr. Albert Levi ....

(Thesis Supervisor)

Assoc. Prof. Dr. Erkay Savaş ....

Asst. Prof. Dr. Hüsnü Yenigün ....

Assoc. Prof. Dr. Özgür Erçetin ....

Assoc. Prof. Dr. Özgür Gürbüz ....

DATE OF APPROVAL ....

MAINTAINING TRAJECTORY PRIVACY IN MOBILE WIRELESS SENSOR NETWORKS

Osman Kiraz

Computer Science and Engineering, MS Thesis, 2012

Thesis Supervisor: Assoc. Prof. Albert Levi

Keywords: Trajectory Privacy, Security, Mobile Wireless Sensor Networks

Abstract

Sensors are tiny, resource-limited devices that are deployed in different areas to gather information for specific purposes. Wireless sensor networks consist of sensors with limited communication range and one or more sink nodes that are responsible for collecting the produced data by the sensors. Mobile wireless sensor networks is a subdomain of wireless sensor networks in which sensors and/or sinks are mobile. Trajectory privacy of the sink node is one of the security issues that are emerged with mobile wireless sensor networks. In this thesis, we propose a scheme for the trajectory privacy of mobile sink nodes. The proposed scheme is based on random distribution of data packets. In this scheme, sensor nodes do not use and need location information of the mobile sink or its trajectory. We performed simulation based and analytical performance evaluations for the proposed scheme. The results show that a network with up to 99% data delivery rate can be obtained by appropriate configuration of the scheme parameters while maintaining the trajectory privacy of the mobile sink node. In addition to that, the proposed scheme has economical resource usage since it does not involve any kind of cryptographic mechanism.

HAREKETLİ KABLOSUZ DUYARGA AĞLARINDA YÖRÜNGE GİZLİLİĞİNİ SAĞLAMA

Osman Kiraz
Bilgisayar Bilimi ve Mühendisliği, Yüksek Lisans Tezi, 2012
Tez Danışmanı: Doç. Dr. Albert Levi

Anahtar Kelimeler: Yörünge Gizliliği, Güvenlik, Hareketli Kablosuz Duyarga Ağları

Özet

Duyargalar küçük, amacına göre çeşitli alanlara dağıtılmış, sınırlı kaynağa sahip belirli amaçlar için bilgi toplayan cihazlardır. Telsiz duyarga ağları; sınırlı iletişim alanına sahip duyargalar ve duyargaların ürettiği bilgileri toplamakla sorumlu alıcı düğümden oluşur. Hareketli telsiz duyarga ağları ise hareket kabiliyetine sahip bileşenlerinden dolayı telsiz duyarga ağlarının alt alanıdır. Alıcı düğümün yörünge güvenliği hareketli telsiz duyarga ağları için ortaya çıkan güvenlik sorunlarından biridir. Bu tezde, telsiz duyarga ağlarında alıcı düğümün yörünge güvenliği için şema önerilmektedir. Önerilen şemanın temeli veri paketlerinin rastgele dağıtımına dayanmaktadır. Bu şemada, duyarga düğümleri hareketli alıcı düğümün yeri veya yörüngesi bilgisine ihtiyaç duymazlar. Önerilen şema için, simulasyona dayalı ve çözümlemeler içeren başarım değerlendirmesi gerçekleştirdik. Sonuçlar göstermiştir ki hareketli alıcı düğümün yörünge güvenliği sağlanırken uygun iletişim kuralları değişkenleri seçildiği takdirde %99'a varan veri iletimi başarı yüzdesine sahip ağ elde edilebilir. Buna ek olarak, önerilen iletişim kuralları her hangi bir şifreleme mekanizması içermediği için hesaplı kaynak kullanımına sahiptir.

Dedicated to the joy of life…

Acknowledgements

I would like to thank my thesis advisor Dr. Albert Levi for his guidance and especially for his psychological support.

Special thanks are due to Dr. Erkay Savaş, Dr. Hüsnü Yenigün, Dr. Özgür Gürbüz and Dr. Özgür Erçetin for their kindness to join my jury.

Also, many thanks to Herr Ergin Gündüz for his support and valuable comments.

I also would like to thank my beloved family for their endless support.

Last but not least; I am grateful to life itself for letting me having the joy of struggling.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Wireless Sensor Networks (WSNs) [1] have emerged as a new generation of distributed embedded systems that provide observations on the physical world at low cost and with high accuracy. A wireless sensor network consists of a large number of tiny, low-powered, energy-constrained smart sensor nodes with sensing, data processing and wireless communication components. Sensor nodes in WSNs are small battery powered devices with limited energy resources, and their batteries cannot be recharged once the sensor nodes are deployed. WSNs have become an exciting research and development area [2] in the last decade and can be used in many various applications, including battlefield surveillance, harbor monitoring, healthcare, etc.

In spite of serving solutions such as monitoring wide areas with easy deployment, WSNs suffer from the following drawbacks [3]:

- Near-sink sensors drain energy faster than the other sensors in the network since near-sink sensors does not only need to deliver their own data packets, but also should forward data packets originated from the other sensors. As a result, the near-sink sensor rapidly falls out of function and this disables the functionality of the entire network.

- Due to the abovementioned reason, near-sink sensors produce high network traffic. This permits attackers, as mentioned in [8], benefit from network traffic analysis for exposing location of sink nodes.

- It may not be feasible to deploy a fixed sink in areas such as battle fields, volcanic areas, underwater zones, etc.

- Deployment in the abovementioned areas creates coverage uncertainty.

In a relative study, Di Pietro et. al. [4] states that Mobile Wireless Sensor Networks (MWSNs) is an alternative to traditional WSNs. MWSNs may be used for overcoming some handicaps of WSNs such as coverage uncertainty. If, for instance, sensors are mobile, they can move toward uncovered area of the network after deployment. The advances in robotics and wireless communication technologies have enabled the development of new architectures for MWSNs which have drawn considerable attention from the research community in the last decade [3].

The network architectures of MWSNs are classified into three categories.

- Static sensor nodes with mobile sink: Sensors are static and one or more mobile collectors periodically visit the deployed area for collection. An example for this kind of network architecture would be sensors that are deployed in a volcanic area and a helicopter as the mobile collector responsible for periodically collecting the data.

- Static sink with mobile sensor nodes: Sensors are mobile and one or more static collectors collect the sensed data when the mobile sensor node falls into the transmission range. Animal with the attached sensors and the sink nodes at the places where animals frequently visit is an example of this kind of network architecture.

- Mobile sink with mobile sensor nodes: Both sensors and sink(s) are mobile. Sensors, with capability of controlling depth of their position, deployed underwater and a few unmanned submarines periodically visit the deployed area for collection of the data is an example for mobile sink with mobile sensor nodes network architecture [5].

MWSNs have their own unique properties such as having dynamic mobile network topology. Since sensor and sink nodes are not always in direct communication, sensor nodes should have the data storage capability. These unique properties have brought many new security challenges. As mentioned in [1], [2], [3], [4], [5] [8] and [9], approaches for general network security issues cannot be applied to the WSNs due to the special characteristics of WSNs. Ren et. al [25] states that the unique properties of MWSNs also prevents the implementation of traditional computer security approaches which are applicable to security issues of static WSNs.

As having mobile sink is part of some network architectures of MWSNs, it is also a key player for the applications that are built on these architectures. For some applications, the owner and the user of the network would be different. For instance, a set of sensors can be deployed on oceanic area in order to collect data about the geographical properties. The users of this network would be oil companies with their own mobile collectors. Since these companies are competitors, they would be interested in each other's data collection region. Therefore, the location privacy of the collectors of mobile companies is a security concern. Drastically, the network could be a military one and the mobile collector could be a soldier. The interest of the attacker would be not only the current location of mobile sink, but also the patrolling trajectory. Thus, the trajectory of the mobile sink is a new security challenge emerged with MWSNs.

To the best of our knowledge, there is only one work[19] addressing the topic of protecting location privacy of a mobile sink. Again to the best of our knowledge, there is no work in the literature addressing the problem of protecting trajectory privacy of a mobile sink.

1.1. Contribution of the Thesis

In this thesis, we propose a scheme to maintain trajectory privacy of mobile sink(s) for mobile wireless sensor networks with mobile sink and mobile sensor nodes network architecture. Our literature search suggests that our work is the first one in the literature addressing the concern of trajectory privacy of mobile sinks. Our scheme relies on homogeneously distributing the sensed data through the network. The proposed scheme does not change the actions of sensor nodes in the infinite unattendance of the mobile sink or in the constant attendance of the mobile sink. Therefore, traffic analysis does not give any information about the mobile sink's location and trajectory.

Since our scheme excludes the location of the mobile sink in the header of packets, it does not require any cryptographic functionality for maintaining trajectory privacy of the mobile sink. This makes our proposal lightweight in terms of memory and computational power. Our performance evaluation shows that our scheme supplies high data delivery rate (up to 99% for certain configurations).

1.2. Organization of the Thesis

The rest of the thesis is organized as follows. Section 2 gives general background information on location privacy approaches in wireless sensor networks and presents existing solutions in the literature. In Section 3, details of the proposed scheme are explained. Section 4 presents the performance evaluation of the proposed scheme. Finally, Section 5 concludes the thesis.

2. BACKGROUND ON LOCATION PRIVACY IN WIRELESS SENSOR NETWORKS

WSNs are deployed in unattended areas and due to the motivation of applications of WSNs such as battlefield surveillance, location privacy of sensor nodes and sink node(s) are important security concerns. In this section, due to the lack of research on the trajectory privacy of sink nodes in MWSNs, we will present general background on location privacy in WSNs. Location privacy concern in MWSNs is classified into two categories: (i) location privacy of sensor nodes, (ii) location privacy of sink node(s).

2.1. Location Privacy of Sensor Nodes

In [6], “*Panda Hunter Game*” is proposed for modeling the location privacy concern of sensor nodes. In the *Panda-Hunter Game,* panda-detection sensor nodes have been deployed by the Save-The-Panda Organization to monitor a vast habitat for pandas [7]. As soon as a panda is observed, the corresponding sensor node makes observations, and sends this message towards the base station via multi-hop routing techniques. Meanwhile, due to the open nature of WSNs, an armed panda hunter may overhear the message. The hunter, by back-tracing the routing path, can find out the location of the sensor that generates the message of panda location.

In [6], *Random-Walk Routing scheme* is proposed for protecting the location privacy of the sensor nodes where the sensors have the mobility capability. The idea is that sensors randomly move for a certain amount of time or distance and then forward the message. If an attacker back traces the forwarded packet, she will only be able to find out an intermediate node’s location. Due to the energy limitations of the sensor nodes, it is not feasible to let source node to make a long distance random walk. Thus, if the attacker is not interested in exact location of the source node, but the region of it, the random-walk scheme does not succeed. In addition to that, this kind of approach is still vulnerable to the location privacy concern of sink node(s).

Dummy data injection is another technique proposed for protection of the location of sensor node [15]. The idea is letting the sensor nodes to distribute dummy data packets in predetermined time intervals or with a predetermined probability. This technique also relies on the perturbation of network traffic which increases the communication overhead and it still does not prevent the high traffic rate at near-sink sensor nodes.

The proposed technique in [30], *Fake Data Source,* is similar to the dummy data injection. Here instead, predetermined nodes behave as the data source and distribute fake data packets at the same time interval of distribution of real data packets. This method also enforces the attacker to make more analysis and computation but still does not provide an appropriate privacy for the location of the sensor nodes. In addition to that, the high energy consumption and communication overhead are also handicaps of this technique.

2.2. Location Privacy of Sink Node(s)

The location privacy of the sink node(s) can be motivated with such an example: movement-detection sensor nodes are deployed in an area to analyze activities of enemies and movement of troops. One or more sinks which are attached to a soldier are used to access the sensed data by sensor nodes. The exposition of the location of sink (and soldier) puts the life of soldier in danger, and also may reveal the entire network’s secrets since the sink node may hold the authentication keys and pairwise keys of the network.

The traffic-analysis attack for tracing the location of sink node is introduced and studied in [8]. Based on the basic observation, near-sink nodes forward more packets than the sensors further away from the sink. An adversary can analyze network traffic intensity at various locations. This analysis may help adversary to estimate the direction of the sink because denser network traffic may mean the location is closer to the sink. The packet-tracing attack for tracing the location of sensor nodes is addressed in [6]. The attack is performed by eavesdropping on the traffic. The adversary is able to perform a hop-by-hop trace toward the original data source.

Flooding-Based Routing scheme is studied in [9, 10, 11, 12, and 13] as a counter measure for the traffic-analysis attack. Each intermediate node broadcasts the received message to its neighbors. As a result, the entire network participates in forwarding one single message to the sink node(s). This approach hardens the traffic analysis for an adversary to trace transmission route back to the sink node. In [14, 15], a minor modification of flooding-based routing scheme (called as *Probabilistic flooding)* is proposed for overcoming the extreme energy consumption of flooding-based routing schemes. In probabilistic flooding, broadcasting the received message to its neighbors is limited with a probability. An intermediate node forwards with a predetermined probability (here, if the predetermined probability equals to 1, it is actually the implementation of flooding-based routing scheme). Despite that all the proposed schemes based on flooding perturb the expected network traffic analysis, they still suffer from not preventing the observable high traffic rate at near-sink sensor nodes and cause extreme communication overhead.

In [6], *phantom routing* is proposed as a more powerful scheme than the abovementioned techniques. They study the variations of flooding-based and single-path routing techniques and claim that none of these schemes provide location privacy of sink node. In phantom routing, the delivery of each message experiences two phases: (1) the random walk phase, which may be a pure random walk or a directed walk, meant to direct the message to a phantom source, and (2) a subsequent flooding/single-path routing stage, meant to deliver the message to the sink. When the source sensor node generates a message, the message is unicasted in a random fashion for a predetermined number of hops. After the hops, in phantom flooding phase the message is flooded using baseline (probabilistic) flooding. With the technique, various routes are produced along one single sensor node to the sink node, which hardens the analysis of an attacker. Although the simulation results have yielded better results according to the previous approaches, *phantom routing* also suffers from not preventing the observable high traffic rate at near-sink nodes and it increases the communication overhead.

In [17], *Location-Privacy Routing* protocol is proposed for the location privacy of the static sink node(s). The scheme allows sensor nodes to select routing paths randomly based on a predetermined probability. Each sensor node's neighbors are divided into two lists: (i) the ones with longer route to the sink node, and (ii) the ones with shorter route to the sink node. When a sensor generates a data packet, it forwards the packet through longer route neighbors with a predetermined probability. Otherwise, it forwards the packet through shorter route neighbors. Although this approach generates various routes along to the sink node, each route will end up around the near-sink sensor nodes. Thus, both network traffic analysis and trace routing would be successful attack methods for exposing the location of sink node.

In [18], *Controlling Transmission Rate* technique is proposed for keeping the same transmission rate among all sensors by controlling delay of actual data packets. Since the asymmetric traffic flow enables an attacker to observe higher network traffic at near sink sensor nodes, with this scheme the amount of traffic per unit time is aimed to be controlled. However, a global attacker may still have the capability of observing the number of packets that are received and forwarded. Thus, even though the transmission rate of near-sink sensor nodes stays at normal values, the volume of packets that they deal with is still important information for an attacker to find out the location of sink node(s).

In [19], a randomized routing scheme is proposed in order to maintain location privacy of sink node for MWSNs with mobile sinks. Packets are forwarded for a predetermined number of hops along a random path and the destination field is not included in the header of the packets. Each intermediate sensor node stores the received packet in its buffer and forwards it if the predetermined hop count is not reached. Since there is no information about the sink nodes in the forwarded packets, location privacy is maintained. However, to be able to have high delivery rate, predetermined hop count should be selected large, which in turn causes higher network traffic.

3. THE PROPOSED SCHEME FOR MAINTAINING TRAJECTORY PRIVACY OF MOBILE SINK

In this section we propose a scheme for preserving the trajectory privacy of sink nodes in mobile wireless sensor networks with mobile sink node(s) and mobile sensor nodes. The proposed scheme relies on the random distribution of packets and storing the packets in intermediate nodes with a predefined probability. Our scheme does not release any address information about the mobile sink node. In addition to these, the scheme does not contain any cryptographic mechanism. Since we do not have any extra cryptographic mechanism, our scheme is computationally lightweight.

The rest of this section is organized as follows. The network assumptions and threat model is explained in Section 3.1. Our proposed approach is detailed in Section 3.2

The notations that are used to describe and analyze the proposed scheme are given in Table 1.

Table 3.1: List of notations used in Section 3

A	Size of the network area
N	Number of nodes in the network
B	Buffer size of a sensor node.
L	Number of different nodes desired to keep copy of data.
P_S	Probability of storing a received data.
$time_i$	i^{th} second of simulation.
D_{DR}	Data delivery rate of the network.
$total_{D_R}$	Number of distinct data packets received by the mobile sink
$total_{D_{AR}}$	Number of data packets received by the mobile sink
$total_{D_G}$	The total number of generated data packets by the mobile sensor nodes
$total_{D_F}$	The total number of forwarded data packets by the mobile sensor nodes
L_R	Remaining number of different nodes desired to keep copy of data.
L_M	Number of different nodes desired by active attacker to keep copy of data
D_G	Data packet generated by a mobile sensor node.
S_F	The mobile sensor node that forwarded data.
S_S	Selected mobile sensor node among neighbor nodes to forward data.
S_G	The mobile sensor node that generates the data.
S_R	The mobile sensor that received data packet.
DGR	Ratio of mobile sensor nodes that generates data at same time interval
NL_S	Neighbor list of a mobile sensor node.
P_F	Probability of sending fake beacon.
T_B	Predetermined time for broadcasting beacon by mobile sink node.
T_F	Predetermined time for broadcasting fake beacon by sensor nodes.
X←Y	Assignment of X to Y.
$S_G \xrightarrow{D_G} S_S$	S_G sends D_G to S_S

3.1. Network Assumptions and Threat Model

In this section, the assumptions of the networks and the abilities of an attacker are given. In Section 3.1.1 the general assumptions of the network are explained. Section 3.1.2 presents the assumptions about the mobile sink node. In Section 3.1.3, the assumptions of the mobile sensor nodes are given explained. Finally, Section 3.1.4 gives the assumption on the abilities of an attacker.

3.1.1. General Assumptions of the Network

The network consists of mobile sensor nodes and a mobile sink node. The sensor nodes are deployed randomly with uniform distribution. There is a risk of non-delivery of a packet in the case the transmission range of holders of the packet does not coincide with the trajectory of the mobile sink. Corollary, the time between the generation and delivery of a packet may lengthen.

Since our main focus is on the trajectory privacy of mobile sink node, other security issues that can be preserved with cryptography are not taken into consideration. Thus, neither private nor public key cryptography is implemented for the data forwarding process.

3.1.2. Assumptions on the Mobile Sink Node

Mobile sink has a predetermined set of trajectories and travels on one of the randomly selected trajectories for one data collection phase. Mobile sink occasionally broadcasts beacon through nearer sensor nodes. Mobile sink has the capability of filtering duplicate data packets.

3.1.3. Assumptions on the Mobile Sensor Nodes

Each sensor node has the same capability in terms of transmission range, battery power, storage and computational power. Each sensor node has a limited transmission range for wireless communication and can exchange packets directly with its neighbor nodes. Each sensor node has a limited buffer and releases the oldest packet if a new packet received or generated and the buffer is full. Even if the packet is delivered to the mobile sink, it is not released from the buffer if there is still space in the buffer. The sensor nodes that their transmission range falls into location of the mobile sink transfer the packets that are stored in their buffer. Each sensor node chooses a random destination within its transmission range and moves towards it with a fixed predetermined velocity. Each node repeats this process immediately when it reaches the destination.

3.1.4. Assumptions on the Abilities of an Attacker

An attacker cannot hear the direct communication between the mobile sink and the mobile sensor node. This assumption is fair enough since otherwise analytically no defense system can maintain the privacy of mobile sink node. With this assumption, attacks containing trace routing technique will not be sufficient since the route of a packet does not change with the existence of a mobile sink. To strengthen the attacker, it is assumed that the attacker would know about the packets with their context that are collected by the mobile sink, as the collection of the data is published in public. With this assumption, attacker would also trace route of her own packets and would learn about if they are collected and know about which sensor nodes have received her packets. An attacker may deploy malicious sensor nodes into the network. Hence, she may at least be aware of the time and location of the direct communication of the mobile sink with her own malicious sensor nodes. An attacker can capture packets and read the contexts of them. Also packet capturing is not an ideal attack technique for an attacker since there is no information about mobile sink in the header of packets. Precisely, the sensor nodes of the network ignore the location or trajectory of the mobile sink.

3.2. The Proposed Approach

In this section, the details of the proposed scheme are given. Section 3.2.1 states the motivation behind this approach. The general overview of the proposed scheme is presented in Section 3.2.2. In Section 3.2.3 storage management is detailed. In Section 3.2.4 the initial phase of the packet distribution is described. In Section 3.2.5 the intermediate phase of the packet distribution is given. Finally, in Section 3.2.6 data collection mechanism is explained.

3.2.1. Motivation

Although wireless sensor networks promise a wide spectrum of applications that cannot be or not easy to be applied by general network schemes, they also bring a wide spectrum of new security concerns. Mobile wireless sensor networks is a subdomain of wireless sensor networks and due to the mobile architecture of these networks, even more new security issues have emerged that cannot be solved by the approaches developed for traditional wireless sensor networks.

Location privacy of a mobile sink is one of the unique security concerns of mobile wireless sensor networks because the sink node is generally assumed to be static in terms of physical location in traditional wireless sensor networks. Moreover, the privacy techniques [20, 21, 22, 23] related with location privacy in general networks are far away from the derivation of the security concern into the architecture of mobile wireless sensor networks. Thus, these approaches cannot be applied to MWSNs. In some applications such as the owner of mobile sinks are in competition with each other, an attacker may be interested in the previous trajectories followed by the mobile sink or the prediction of the future trajectories of the mobile sink nodes. In the literature, only a few works exist on location privacy of the mobile sink nodes. To the best of our knowledge, no work so far published on the topic of the trajectory privacy of mobile sink node.

Our aim with this thesis is to highlight the problem of trajectory privacy of the mobile sink in mobile wireless sensor networks and propose a scheme that maintains the trajectory privacy while preserving desirable network property such as high data delivery rate.

3.2.2. Overview of the Scheme

The proposed scheme is based on homogenous distribution of the data packets by random forwarding and random movement of the mobile collector node. Our scheme aims to preserve the trajectory privacy of the mobile sink while keeping the data delivery rate and communication overhead at acceptable values. A depiction of the network is given in Figure 3.1.

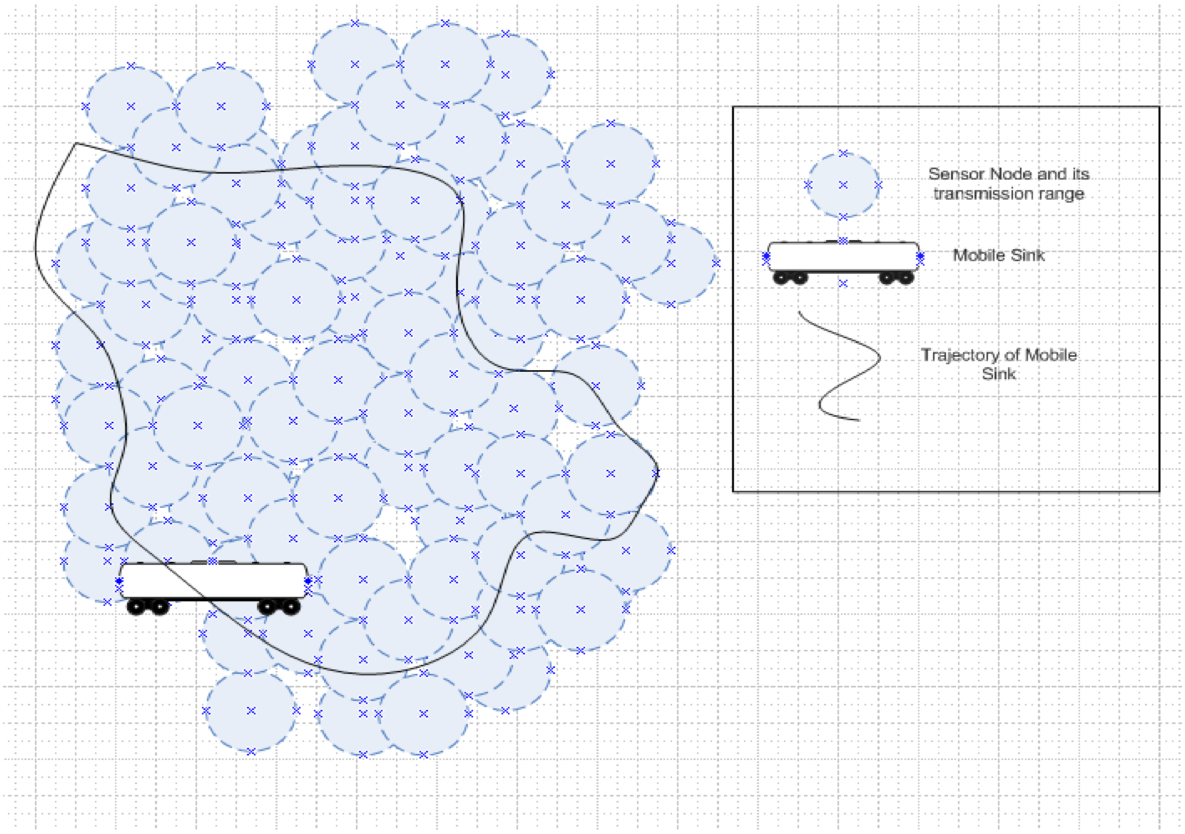
Sensor Node and its transmission range
Mobile Sink
Trajectory of Mobile Sink

Figure 3.1: MWSN with mobile sink and mobile sensor nodes

The mobile sink has a predefined set of trajectories. For each collection phase, it randomly selects one of them and travels on the selected trajectory with a preset constant speed. It broadcasts beacon for every T_B, predetermined time for broadcasting beacon, to let the sensor nodes be aware of its existence. Also each sensor broadcasts fake beacons for every T_F, predetermined time for broadcasting fake beacon, with the probability of P_F, probability of sending fake beacon. The mobile sink has the capability of filtering out duplicate data packets. The detailed information about data collection mechanism is given in Section 3.2.6.

Each sensor node has a storage, which is limited with a buffer size, B. Whenever a sensor node receives the broadcast message of the mobile sink and if its transmission range covers the location of the mobile sink, it forwards all the data packets in its buffer. The detailed information for the buffer size of a sensor node, B, and storage management are given in Section 3.2.3.

When a sensor node generates a data packet, it stores the packet in its buffer and distributes the data packet to other L sensor nodes to have them keep a copy of it. If L, number of different nodes desired to keep a copy of data, is initialized to zero, the mobile sensor nodes in the network will not forward or receive a data packet and will only interact with the mobile sink node. If, for instance, L is set to 10, then the number of copies stored for this packet by other sensor nodes in the network will be 10. The detailed information for L and the initial phase of the packet distribution are explained in Section 3.2.4.

If a sensor node receives a packet, it keeps the packet in its buffer with the probability P_S and decrements L_R, the remaining number of different nodes desired to keep a copy of data. With the probability $1 - P_S$, the packet is not stored and L_R is not decremented. The received packet is forwarded if L_R is higher than 0. The detailed information about the intermediate phase of the packet distribution is given in Section 3.2.5.

3.2.3. Storage Management

If a sensor node interacts with the mobile sink node and delivers all the data packets in its buffer, it does not necessarily clean up the entire buffer. The reason behind this is preventing an attacker to perform a successful attack, which is constructed on combination of traffic analysis and node capturing. If all of the storage of a sensor node is cleared with the interaction and a high traffic rate is observed on this node lately, the attacker would observe the empty storage by capturing this node and can conclude that the mobile sink has just passed near to this sensor node and interacted with it. In other words, cleaning up the buffer after delivering all of the stored packets helps the attacker to obtain information about a part of the trajectory of the mobile sink node. The mobile sink is assumed to have the capability of filtering out the duplicated data packets and the next interaction with the mobile sink should take some time. These facts encourage this kind of storage management approach to be applied by considering the concern of the trajectory privacy of the mobile sink node.

When a node desires to store a new data packet (either because of generation of the data packet or receiving a forwarded data packet) into its buffer, it checks the volume of the occupation of its storage and if it is equal to the buffer size of a sensor node, B, it drops the oldest packet. The data packet, which stayed longer in the buffer, has a higher probability to be already collected by the sink.

The pseudo-code of storage management of a sensor node is given in Figure 3.2

```
if $newPacketDesiredToBeStored$
  $\mathrm{usedSpace} = \mathrm{getNumberOfPacketsInBuffer}$
    if $\mathrm{usedSpace} \equiv B$
      drop $Package_{oldest}$
 $\mathrm{addPacket}(newPacket)$
```

Figure 3.2: Pseudo-code of storage management

3.2.4. Initial Phase of the Packet Distribution

When a sensor node generates data D_G, it inserts D_G into its buffer with the storage management approach that is mentioned in Section 3.2.3. If L, the predetermined different number of sensor nodes desired to keep D_G in its storage, is higher than 0, then the number of different nodes to a keep copy of data, L_R, is set to the number of different nodes desired to keep a copy of data, L. The information of L_R is attached into the header of data packet D_G. Finally, D_G is forwarded to S_S, the selected mobile sensor node among the neighbor nodes to forward data.

If L is zero, the mobile sensor node stores the generated data packet but does not forward it. In other words, the mobile sensor nodes do not interact with each other but communication takes place only between the mobile sink and the mobile sensor nodes. The pseudo-code of initial phase of the packet distribution is given in Figure 3.3

```
Case: Data $D_G$ is generated by a mobile sensor node, $S_G$
        addPacket($D_G$ )
         if $L > 0$
              $L_R \leftarrow L$
              $D_G \leftarrow D_G \,||\, L_R$
            Select $S_S$ among $NL_S$
            $S_G \xrightarrow{D_G} S_S$
        else
            Do Not Forward
```

Figure 3.3: Pseudo-code of Initial Phase of the Packet Distribution

3.2.5. Intermediate Phase of the Packet Distribution

When an intermediate node receives D_G from a mobile sensor node S_F, it stores D_G with the predetermined probability value of P_S in its buffer by applying the storage management approach mentioned in Section 3.2.3 and decrements L_R. L_R is not decremented if the data packet is not stored with the probability $1 - P_S$.

If P_S favors for storing D_G, and L_R is higher than 0, the mobile sensor node selects one neighbor node S_S among NL_S except S_F and forwards D_G with possible decremented

L_R attached to the header of D_G. The pseudo-code of intermediate phase of the packet distribution is given in Figure 3.4

```
Case: Data $D_G (D_G \| L_R)$ is received from $S_F$
        if $RNG(0,1) \le P_S$
            addPacket($D_G$)
            $L_R \leftarrow L_R - 1$
            $D_G \leftarrow D_G \| L_R$
        if $L_R > 0$
            Select $S_S$ among $NL_S - S_F$
            $S_R \xrightarrow{D_G} S_S$
```

Figure 3.4: Pseudo-code of Intermediate Phase of the Packet Distribution

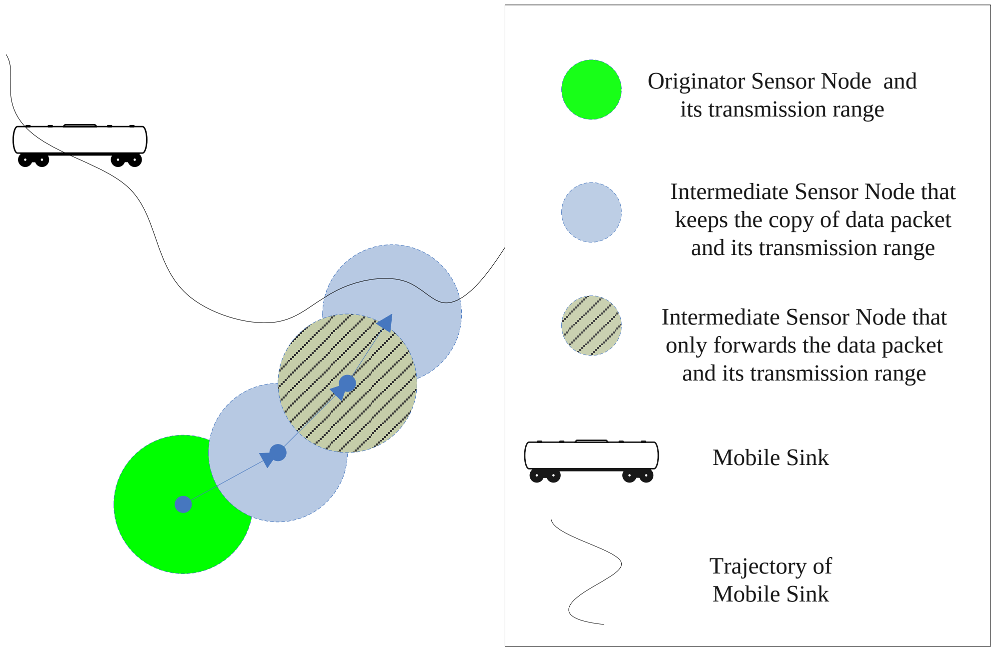
Originator Sensor Node and its transmission range
Intermediate Sensor Node that keeps the copy of data packet and its transmission range
Intermediate Sensor Node that only forwards the data packet and its transmission range
Mobile Sink
Trajectory of Mobile Sink

Figure 3.5: A local view of data distribution with $L = 2$ and $P_S = 0.5$

The reason behind not decrementing L_R when D_G is not stored, is to maintain a homogenous distribution of D_G in the entire network. By doing so, the delivery probability of D_G increases because if a mobile sensor node does not have chance to interact with the mobile sink node, the closer neighbor nodes may also have no chance to interact. The probability of having an interaction with the mobile sink node and at least one of the sensor nodes at far and different locations is higher. This situation is illustrated in Figure 3.5 with the scenario of $L = 2$ and $P_s = 0.5$. If the intermediate node had decremented L_R, it was not going to forward the data packet anymore and data was not going to be delivered because the trajectory of the mobile sink node is not in the transmission range of the originator node and the intermediate sensor nodes that stored the data packet.

3.2.6. Data Collection Mechanism

Mobile sink broadcasts beacon through nearer sensor nodes for every T_B, predetermined time for broadcasting beacon. In order to hide the existence of the mobile sink, each sensor broadcasts fake beacons for every T_F, predetermined time for broadcasting fake beacon, with the probability of P_F, probability of sending fake beacon. Thus, a sensor node cannot differentiate a beacon if it is generated by the mobile sink or by any other mobile sensor node. Sensor nodes that received a beacon broadcast packets in the buffer without dropping them as mentioned in Section 3.2.3. Mobile sink has the capability of filtering out duplicate packets and drops the packets that have been already received

4. PERFORMANCE EVALUATIONS

In this section, a detailed performance evaluation of our scheme is provided using both simulation and analytically. Section 4.1 explains the performance metrics and analyzed issues. In Section 4.2, simulation environment and setup is explained. Section 4.3 discusses the simulation and analytical results.

4.1. Performance Evaluation Metrics & Analyzed Issues

We are going to evaluate the performance of our scheme using the following metrics and issues.

Data Delivery Rate ($\boldsymbol{D_{DR}}$): Since the proposed scheme does not establish a route toward the mobile sink nodes, delivery of a data packet is not guaranteed. Thus, delivery rate of the generated data packets is one of the main metrics of our performance evaluation in order to measure the success of our proposed scheme. The ratio of the number of distinct data packets received by the mobile sink over the total number of generated data packets by the mobile sensor nodes gives D_{DR}:

$$D_{DR} = \frac{total_{D_R}}{total_{D_G}} \tag{1}$$

Hiding Ratio: Our threat model proposes that an attacker can deploy her own nodes into the network. Thus, hearing a beacon by a malicious node gives information about the location of mobile sink node. In order to avoid this situation, our scheme lets mobile sensor node to broadcast fake beacons for every T_F, predetermined time for broadcasting fake beacon, with the probability of P_F, probability of sending fake beacon. In this way, a mobile sensor node that receives a beacon cannot differentiate if the beacon is generated by the mobile sink or by any other mobile sensor node. We compute the ratio of the number of fake beacons heard generated by mobile sensor nodes and total number of heard beacons. The average of this ratio yields hiding ratio:

$$Hiding\ Ratio\ = Mean(\frac{number\ of\ heard\ fake\ beacons}{total\ number\ of\ heard\ beacons}) \qquad (2)$$

Communication Overhead: One of the most important mechanisms of our scheme to be successful in terms of D_{DR} is distributing the generated data packets to the different locations of the network. As a side effect of this mechanism, high network traffic is expected. Number of copies in the network may increase the probability of deliverance but higher number of packet forwarding is required to have more number of copies of a packet. Thus, we evaluate the communication overhead in terms of amount of transmissions among the mobile sensor nodes and the amount of generated data packets.

Resilience against Traffic Analysis Attacks: Since the traffic analysis attack is one of the most studied attacks for location privacy in WSNs, we analyze different traffic rate of the different regions of the network and compare them with each other to measure the resilience of our scheme against traffic analysis attacks.

Resilience against Node Fabrication Attacks: We evaluate the effect of node fabrication attacks (An attacker deploy her own sensor node into the network and make them participate in the network scheme) by modeling two types of attacks: (i) pure passive attack, (ii) active attacks. Details of these attack models are given in Section 4.3.5 and in Section 4.3.6 relatively.

4.2. Simulation Environment and Setup

Simulation is implemented using Omnet++ Network Simulation Framework [24] in Solaris 10 (SunOS 5.10) using Intel Xeon X5675 3.06 Ghz CPU. In our simulations, 100 nodes (N) are uniformly distributed over a field of $\mathrm{A} = 100\mathrm{m} \times 100\mathrm{m}$. We run the simulations for $400\,seconds$. A mobile sink enters into the sensor area at $time_{50}$ and follows a predetermined trajectory which falls out of the simulation area after $time_{400}$. Speed of the mobile sink is with $1\,m/s$. Mobile sink broadcasts beacon for every T_B, where $T_B = 1\mathrm{s}$. Sensor nodes and sink node have a communication range of $10\,meters$. Each sensor node selects a random destination within its transmission range and moves towards it with $1m/s$ speed and repeats this process immediately after reaching its destination. From $time_0$ to $time_{350}$, at every 5 seconds, a randomly selected predetermined portion of the sensor nodes (DGR) generate data packets. From $time_0$ to $time_{350}$, sensor nodes broadcasts beacon for every T_F, where $T_F = 1\mathrm{s}$. Each set of simulation scenarios is performed 10 times and average values are reported to converge the randomization.

4.3. Simulation and Analytical Results

We perform three basic simulation scenarios with various set of parameter values:

- **Benign network**: We have simulated the network without any attack to observe the performance of the network under normal circumstances with various scheme parameter values.
- **Network under pure passive attack**: We have simulated the proposed network scheme with malicious nodes which generate a data packet and do not forward it.

- **Network under active attack**: We have simulated the network with malicious nodes that are also actively participating in the data distribution process.

In Section 4.3.1 we give the results for benign network scenario in terms of data delivery rate, and discuss about the effects of number of different nodes desired to keep copy of data, L, buffer size of a sensor node, B and probability of storing a received data, P_S. Hiding ratio and effects of P_F, probability of sending fake beacon, on hiding ratio is discussed in Section 4.3.2. Communication overhead is analyzed in Section 4.3.3. In Section 4.3.4, traffic analysis attack is discussed by observing the traffic rates of the network for its different subregions. In Section 4.3.5 the network under pure passive attack is analyzed. In Section 4.3.6 the network under active attack is discussed and analyzed. Finally, in Section 4.3.7 performance difference of our scheme with the approach studied by Ngai et al. in [19] is presented.

4.3.1. Data Delivery Rate

Figure 4.1 shows the data delivery rate for various values of L while keeping the sensor node's buffer size B fixed to 10 packets and DGR fixed to 0.15 (i.e. for each 5 simulation time, a randomly selected 15% of the mobile sensor nodes generate data packet). The scenario is processed for benign networks.

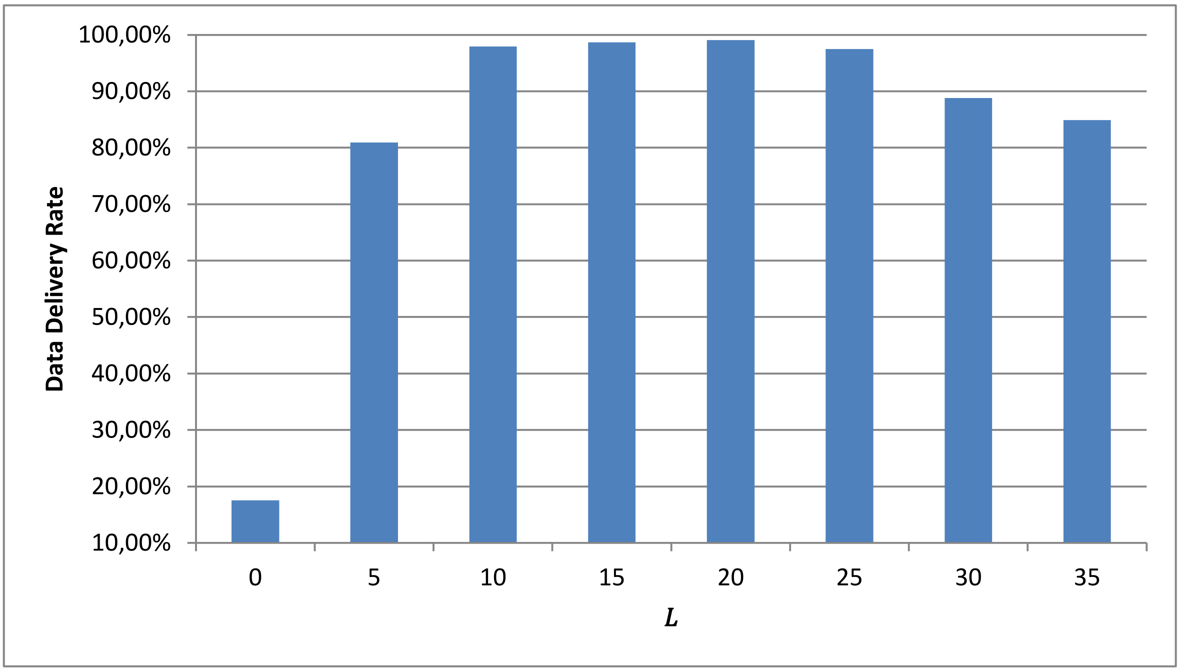

Figure 4.1: Data Delivery Rate vs. L for benign networks ($B = 10$, $P_S = 0.5$ and $DGR = 0.15$)

It is observed that, with the increase of L, D_{DR} increases and comes to a saturation point between L = 10 and 20. In this setup of simulation, the actual number of generated data packets is 1065 (15% of the network generated data packets for each 5 seconds from $time_0$ to $time_{350}$). In Table 4.1, the actual number of delivered data packets for various L is given. D_{DR} starts to decrease after L = 20. For this setup, L = 10 is D_{DR} is high for L = 10 as much as L between 10 and 20, but communication overhead increases with the increase of L. Thus, L = 10 is the optimum for this simulation configuration. These results conclude that L affects the distribution of data packets among the network but after some certain point delivered packet amount decreases.

Table 4.1: Actual number of delivered data packets

L	0	5	10	15	20	25	30	35
#Delivered	187	862	1043	1051	1055	1038	946	904

Packets may not be delivered either because there have been no interaction between the mobile sink and the sensor nodes that have a copy (undelivered packets), or the packets are dropped from the buffer due to buffer overflow (buffer overflowed packets). Figure 4.2 shows the correlation between undelivered data packets and buffer overflowed data packets for the same simulation (Note: values of zeros are depicted as 1 to be able to scale the graph logarithmically.)

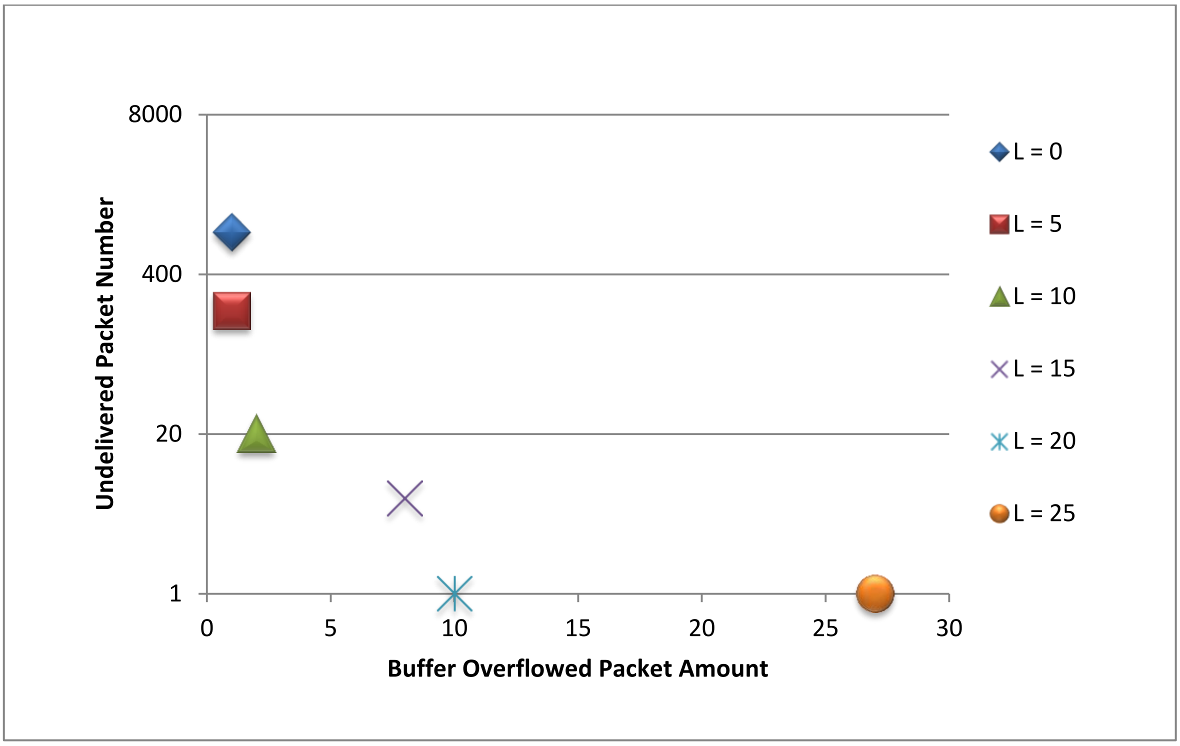

Figure 4.2: Undelivered vs. Buffer Overflowed Packets for benign networks ($B = 10$, $P_S = 0.5$ and $DGR = 0.15$)

It is observed that the amount of undelivered data packets decreases significantly and converges to 0 for $L \geq 20$. In other words, for higher values of L, there is not any data packet distributed to a set of mobile sensor nodes that are not interacting with the mobile sink node. However, the amount of buffer overflowed data packets are increasing with the increase of L.

For some MWSNs, the volume of sensed data by sensor nodes would be very high and for some others it would be very small. We expect less buffer overflow for networks with lower data generation rate, DGR. DGR is not a part of the proposed scheme definition, but it is a simulation parameter for us to model networks with different rate of data generation. For this reason, we have processed simulations with various values of DGR while keeping the other network factors fixed ($B = 10$, $P_S = 0.5$ and $L = 10$).

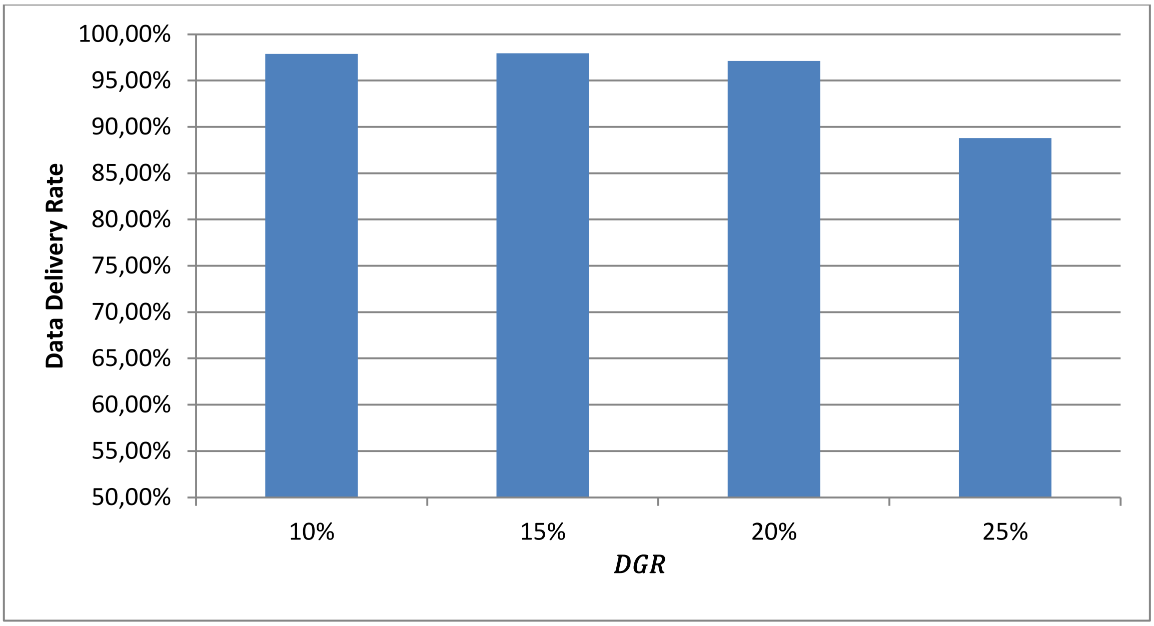

Figure 4.3: Data Delivery Rate vs. DGR for benign networks ($B = 10$, $P_S = 0.5$ and $L = 10$).

Figure 4.3 shows the effect of different values of DGR on D_{DR}, data delivery rate. Up to the value DGR = 20%, we have D_{DR} over 97%. However, at the point DGR = 25%, a tremendous decrease on D_{DR} is observed, which is almost 10 points. The main reason of this decrease is the fluctuation of the amount of buffer overflowed packets. Figure 4.4 is depicted for the same scenario of Figure 4.3 and it shows the relation between data generation rate, DGR, and the number of packets buffer overflowed. For a limited B, we observe that there exists a threshold of DGR and after passing this threshold, the packets start to be dropped due to buffer overflow. For our simulation setup, this threshold happens to be around 20%-25%.

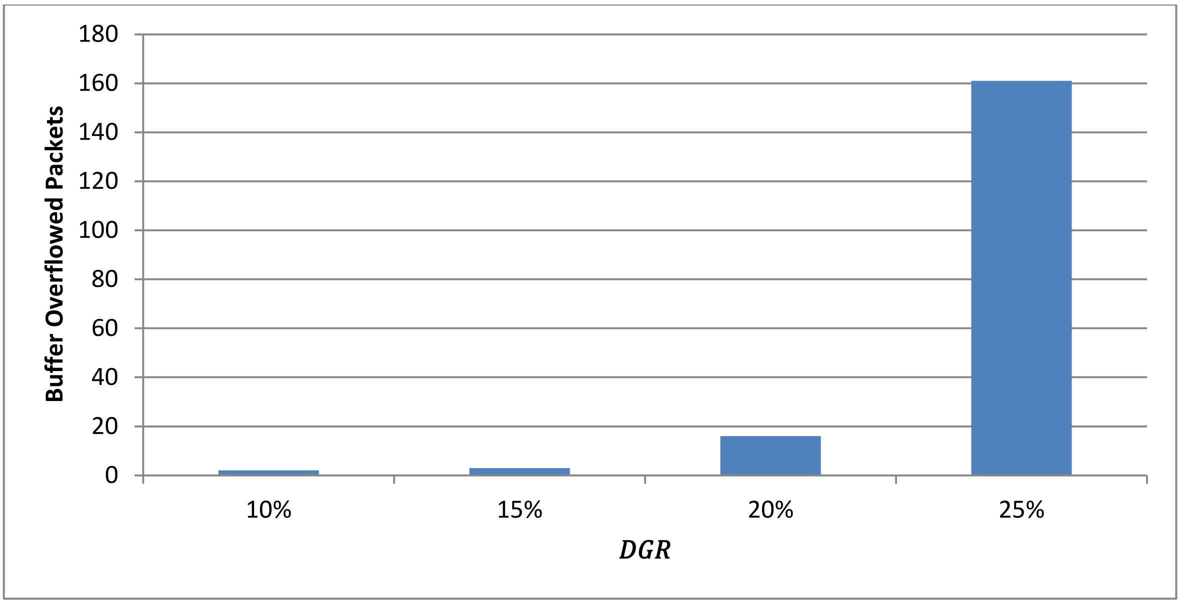

Figure 4.4: Buffer Overflowed Packets vs. DGR for benign networks ($B = 10$, $P_S = 0.5$ and $L = 10$).

In Figure 4.5, effects of different P_S values on D_{DR} is depicted with the fixed values $B = 10$, $L = 10$ and $DGR = 0.15$. It is observed that D_{DR} is decreased with increase in probability of storage. The observation shows that distribution of the packets through the network increases with the decrease of P_S. However, the acceleration of this decrease is low until the value of P_S = 0.5. Thus, P_S = 0.5 is optimum for this simulation setup.

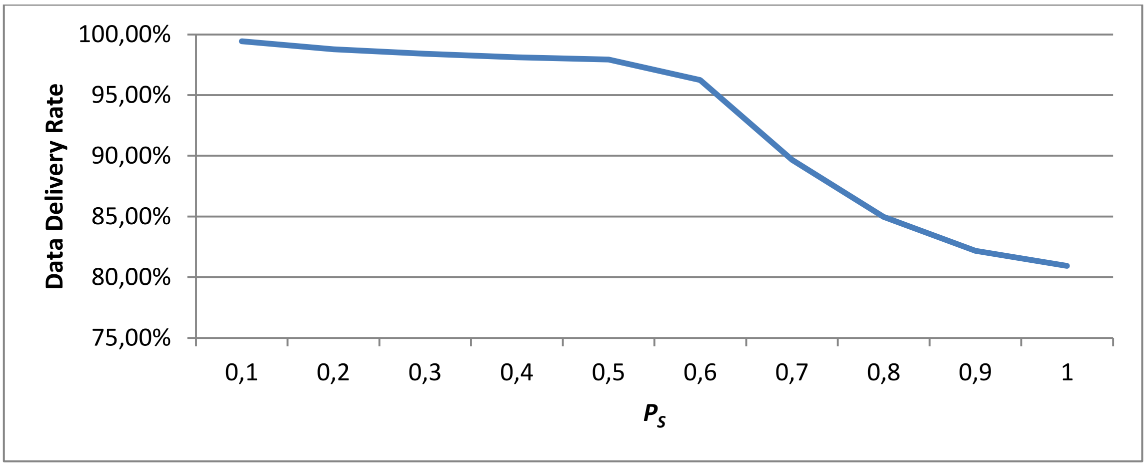

Figure 4.5: Data Delivery Rate vs. P_S for benign network ($B = 10$, $L = 10$ and $DGR =$ 0.15).

To sum up, we have three basic parameters (L, B and P_S) for our proposed scheme. D_{DR} increases linearly with respect to the buffer size of a sensor node, B. As it is observed in Figure 4.1, D_{DR} will be less than 20% if there is only space for the self-generated packets in the storage (L = 0). On the other hand, if B was infinite, there would not be any buffer overflowed packets. According to Figure 4.2, D_{DR} would have been 100% in case of infinite B. For a certain amount of increase in L, we observe fast convergence of D_{DR} to 100%. However, due to the high network traffic and limited B, data delivery rate starts to decrease for higher values of L. In another aspect, high network traffic and limited B increase the number of buffer overflowed packets, which in turn decrease D_{DR}.

All these simulations demonstrate that with a fine tuning of the parameters of our scheme, it is easy to maintain a high D_{DR} but these parameter values would differ from network to network because every network may have different limitations such as low buffer size. Based on the application area, the data generation rate of the networks may differ. It may be less for networks to observe geographical properties of an area but it may be high for a network that senses radioactivity in a nuclear station.

4.3.2. Hiding Ratio

In a scenario where mobile sensor nodes do not broadcast fake beacons, an attacker is able to get information about trajectory of mobile sink node via deployed malicious nodes. In our scheme, with the integration of fake beacons, the attacker does not know if the beacon is generated by the mobile sink node or by any other node. However, if the probability of receiving a beacon from the mobile sink node is higher, the attacker can use this statistics to reveal the trajectory.

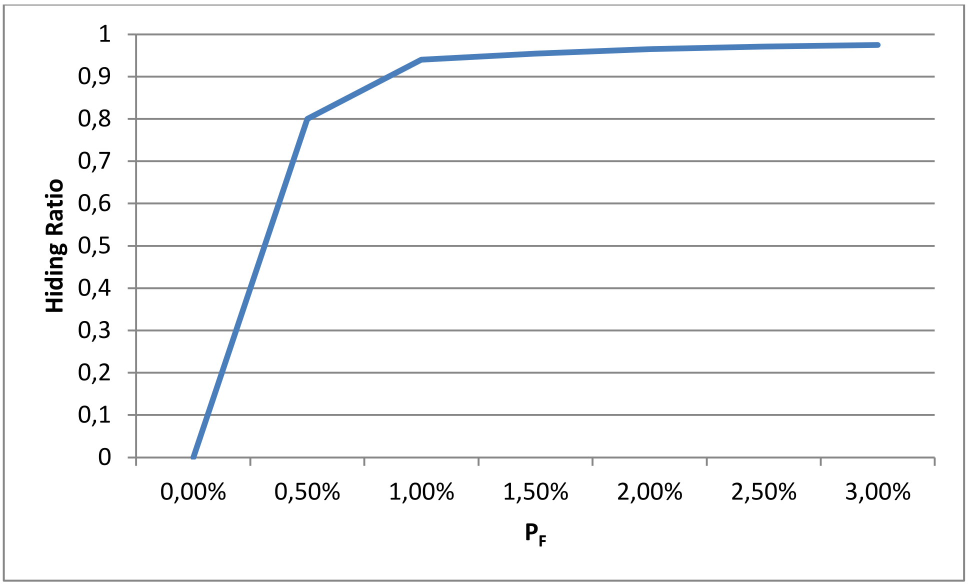

Figure 4.6: Hiding Ratio vs. P_F for benign network ($B = 10$, $P_S = 0.5$, $L = 10$ and $DGR = 0.15$).

In Figure 4.6, hiding ratio is depicted for different values of P_F, probability of sending fake beacon, with the fixed values $B = 10$, $P_S = 0.5$, $L = 10$ and $DGR = 0.15$. Trivially, hiding ratio is 0 for $P_F = 0.0\%$ since any beacon heard by a mobile sensor node is generated by the mobile sink. For $P_F = 0.5\%$, hiding ratio is 0.8 which also proposes that on average 20% of the beacons heard by a mobile sensor node is generated by the mobile sink node. With the increase of P_F, hiding ratio increases and converges to 100%. The acceleration of the increase in hiding ratio reaches the saturation point at $P_F = 1.0\%$. Thus $P_F = 1.0\%$ is optimum for this simulation setup.

In Figure 4.7, the ratio of number of broadcasts of packets in the buffer due to fake beacons over number of broadcasts of packets in the buffer due to actual beacons are given for the scenario depicted in Figure 4.6. It is observed that as the P_F increases, number of extra broadcasted packets increase linearly. Hiding ratio never reaches to 1.0. That is to say, the minimum P_F that supplies desirable hiding ratio is optimum. For this simulation setup $P_F = 1.0\%$ is optimum where hiding ratio is above 0.94.

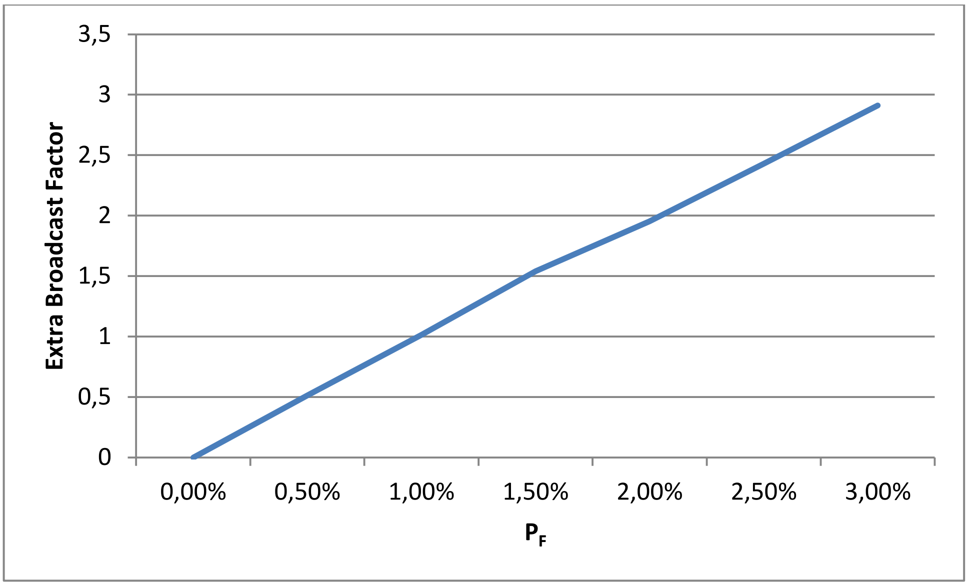

Figure 4.7: Extra Broadcast Factor vs. P_F for benign network ($B = 10$, $P_S = 0.5$, $L = 10$ and $DGR = 0.15$).

4.3.3. Communication Overhead

In our scheme, due to the fact that we are interested in distributing the data packets among the entire network as much as possible, our approach is directly affected by the number of transmissions. The expected number of transmissions for one successful packet delivery is $\frac{L}{P_S}$. The expected number of forwarded data packets is calculated as follows:

$$E\left[total_{D_F}\right] = \sum_{i=1}^{total_{D_G}} \frac{L}{P_S} = \frac{total_{D_G} \times L}{P_S} \tag{3}$$

It is expected to have 20 transmissions for having 10 nodes with the copy of the data packet according to (3). Figure 4.8 shows the linear relationship between L and the total number of forwarded data packets by the mobile sensor nodes, $total_{D_F}$. It can be observed that with the increase of L, the amount of transmissions among the mobile sensor nodes increase linearly, which is expected according to (3). In Table 4.2, $E\left[total_{D_F}\right]$ and actual values of $total_{D_F}$ are given for the simulation scenario depicted in Figure 4.8.

Table 4.2: $E\left[total_{D_F}\right]$vs. $total_{D_F}$

L	$E\left[total_{D_F}\right]$	$total_{D_F}$
0	0	0
5	10650	10256
10	21300	21285
15	31950	33674
20	42600	41445
25	53250	52186

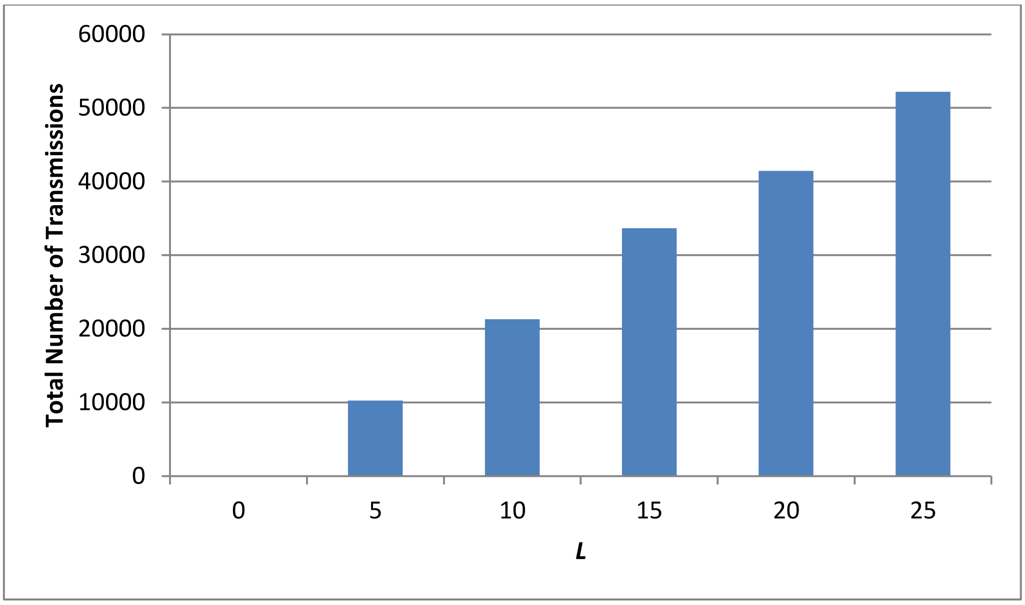

Figure 4.8: $total_{D_F}$ vs. L for benign network ($B = 10$, $P_S = 0.5$ and $DGR = 0.15$)

In Figure 4.9, extra collection factor (the ratio of the number of distinct data packets received by the mobile sink to the number of data packets received by the mobile sink) is given for various L and for the same scenario depicted in Figure 4.8. It is observed that with the increase of L, the number of same packets received for one singe data packet increases. In other words, for one single packet, the mobile sink collects extra data packets and has to filter out them. As mentioned in Section 4.3.1, L = 10 is optimum for this simulation setup since the extra effort for collection increases where D_{DR} slowly increases for L > 10.

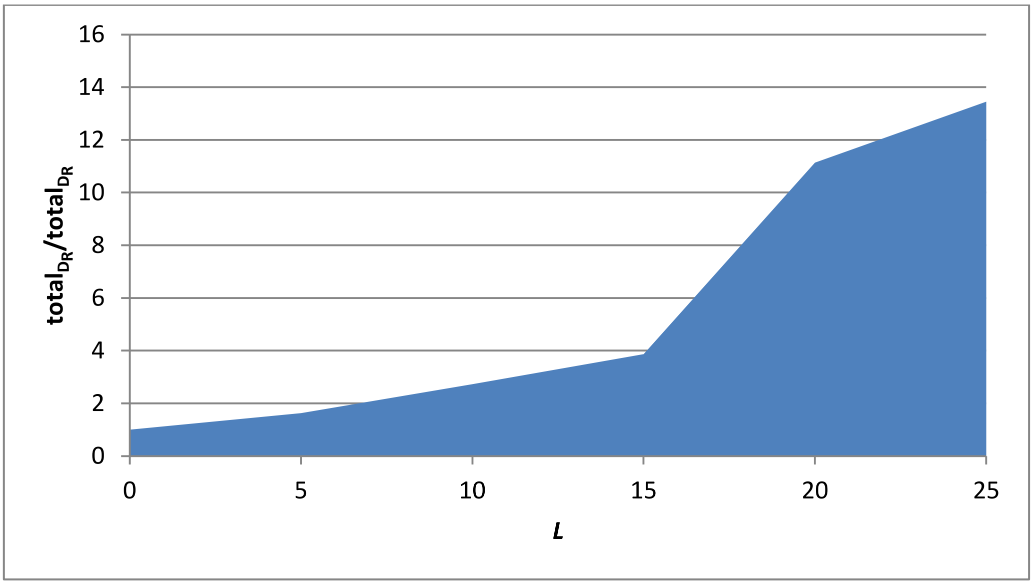

Figure 4.9: $total_{D_R}/total_{D_{AR}}$ vs. L for benign network (B = 10, P_S=0.5 and DGR = 0.15)

In conclusion, the parameter setup directly effects the communication overhead in the network. Although the simulation results show a linear increase in the traffic rate according to the parameter values, we find the increase acceptable. Under normal circumstances, there had to be a forwarding mechanism for any scheme because it is not feasible for a mobile sink to visit every mobile node in the network. Thus, a certain amount of transmissions is expected for any. In our scheme, this communication overhead is manageable and can be foreseen. Hence, with the fine tuning of scheme parameters, the communication overhead can be maintained while keeping the data transmission rate at desirable values.

4.3.4. Traffic Analysis Attack

As mentioned in [8], Traffic Analysis Attack is a powerful technique used by attackers for location privacy concerns in WSNs. Thus, most of the approaches for preserving privacy location involve a counter-measure for traffic analysis. In our proposed scheme, traffic analysis does not yield any useful information for an attacker since our scheme's routing is independent of the location of the mobile sink node. Precisely, mobile sensor nodes do not take into account the trajectory of the mobile sink while distributing generated data packets. Eventually, even if there is no mobile sink in the network, the behavior and consequently the traffic rate of the network do not change. Actually, the nature of our scheme produces a network traffic that can be predicted and due to this prediction, any abnormal traffic rate information can be used for the security systems. In other words, the traffic analysis actually can be used as a security tool for the network.

To illustrate the deterministic behavior of our scheme in terms of network traffic, we divide A, size of the network area, into 25 subregions $(20\text{m} \times 20\text{m})$ and compared $total_{D_F}$ per each subregion. See Figure 4.10 for subregions' illustration.

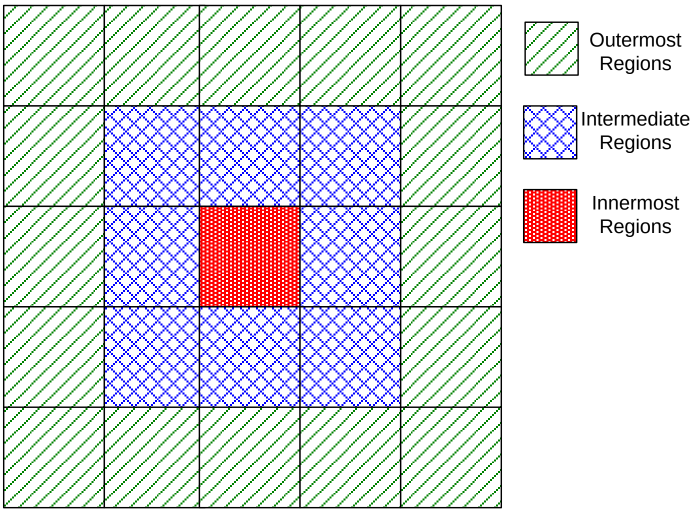

Figure 4.10: Subregions of A

Analogically, sensor nodes near to the edges of the network have less traffic rate and the sensor nodes in the middle of the network have higher rates. In Figure 4.11, surface illustration of the network according $total_{D_F}$ of subregions is depicted for the simulation scenario with values $B = 10$, $P_S = 0.5$, $L = 10$ and $DGR = 0.15$. It is observed that $total_{D_F}$, the total number of forwarded data packets by the mobile sensor nodes, increases from outermost regions to innermost regions. In addition to that, regions in the same layer have almost same $total_{D_F}$. Despite the innermost region has the highest traffic, some of the trajectories of the mobile sink do not cover the innermost region sensor nodes' transmission range. Moreover, there is no different traffic rate between the same layer subregions where some of them involve the trajectory and some do not.

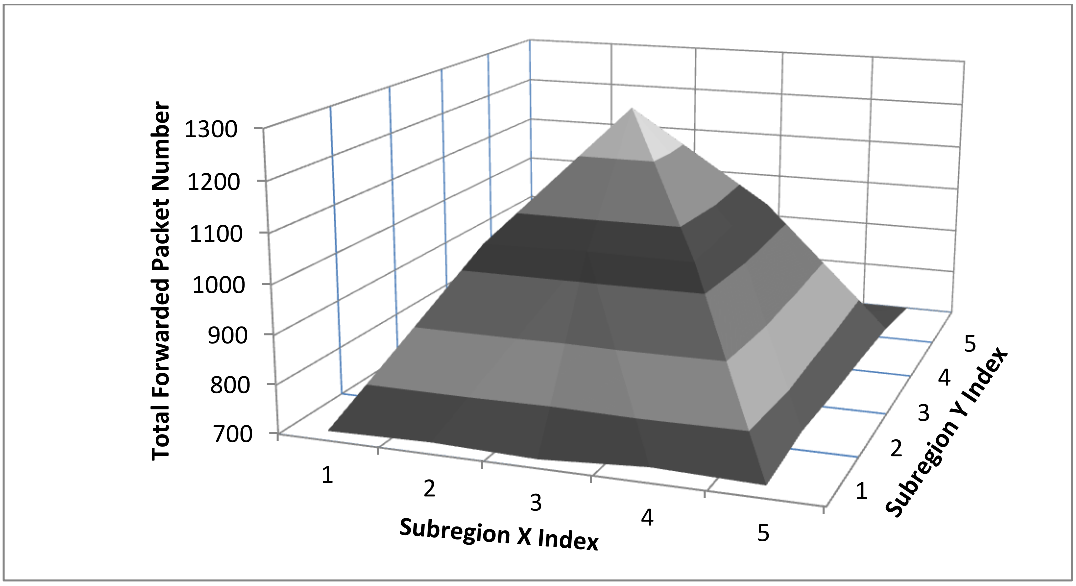

Figure 4.11: Traffic Illustration Based on Subregions ($B = 10$, $P_S = 0.5$, $L = 10$ and $DGR = 0.15$.)

Because of the deterministic behavior of network traffic for the networks having our proposed scheme, observing that two same layer subregions having significantly different traffic rates do not conclude about the trajectory of the mobile sink. Actually, this kind of abnormality is not expected and may reflect a malicious behavior in the network, such as dysfunction of sensors or deployment of malicious sensor nodes into the region.

Thus, traffic analysis can be used as a tool for intrusion detection system for our scheme, rather than a tool for attackers to expose trajectory of the mobile sink node.

4.3.5. Network under Pure Passive Attack

In *Pure Passive Attack* model, an attacker deploys her own static sensor nodes into the network area with her own generated data packets but do not distribute these packets through the network. In case of receiving a data packet from other nodes, it is processed via proposed scheme principles.

For pure passive attacks, interaction with the mobile sink gives exact information about the location of the mobile sink. In addition to that, no interaction provides the information that the location of the malicious node is not part of the trajectory.

We have processed simulations with various values of DGR while keeping the other network factors fixed ($B = 10$, $P_S = 0.5$ and $L = 10$) and 6 malicious nodes in addition.

Out of 6 malicious nodes, 2 of them have interacted with the mobile sink node and 4 of them have not interacted with the mobile sink node. In other words, the attacker have learnt 2 points of the trajectory and learnt that 4 locations do not fall into the trajectory while having a network with 5,67% (6 out of 100 + 6) of the sensor nodes are malicious

We ignore, a wise ignorance in favor of the attacker, the fact that the trajectory also contains locations in areas with absence of any mobile sensor nodes and we ignore the time dimension of a trajectory. For this analysis, the trajectory is a set of the locations where the mobile sink interacted with the mobile sensor nodes. Thus, we conclude that the number of locations constructing the trajectory is equal to the number of distinct sensor nodes interacted with the mobile sink. Under these extreme assumptions, the least number of nodes to be maliciously deployed in the network to learn entire trajectory is equal to the number of the mobile sensor nodes interacted with the mobile sink, say α. The probability of selecting a location that falls into the trajectory point is equal to $\alpha/_N$. If learning β percentage of the trajectory points is assumed to be enough for an attacker to induce the rest of the trajectory, the expected number of nodes should be deployed is calculated as follows:

$$E[N_M] = \frac{N}{\alpha} \times \beta \times \alpha = N \times \beta \tag{4}$$

So, even if β = 20% is enough to learn the rest of the trajectory, number of the nodes should be deployed is the 20% of the total number nodes in the network. In conclusion, we have served assumptions in favor of the attacker such as ignoring the time dimension of a trajectory ignoring the locations that are not interacting with any mobile sink node. Yet, we concluded that the attacker should deploy an infeasible amount of nodes in the network to learn the trajectory of the mobile sink node. Thus, our scheme is resilient against pure passive attacks.

Simulation scenario process for pure passive attack has the same configuration setup with the simulations depicted in Figure 4.3. In Figure 4.12, the correlation of $total_{D_F}$, total number of forwarded data packets by the mobile sensor nodes, and data generation rate, DGR, is given for benign networks and networks under pure passive attack. It is observed that number of transmissions is almost sam for benign networks and networks under pure passive attack. Pure passive attack does not put an abnormal behavior in terms of network traffic rate. Since the attack is passive, the traffic analysis is not successful for detecting pure passive attacks.

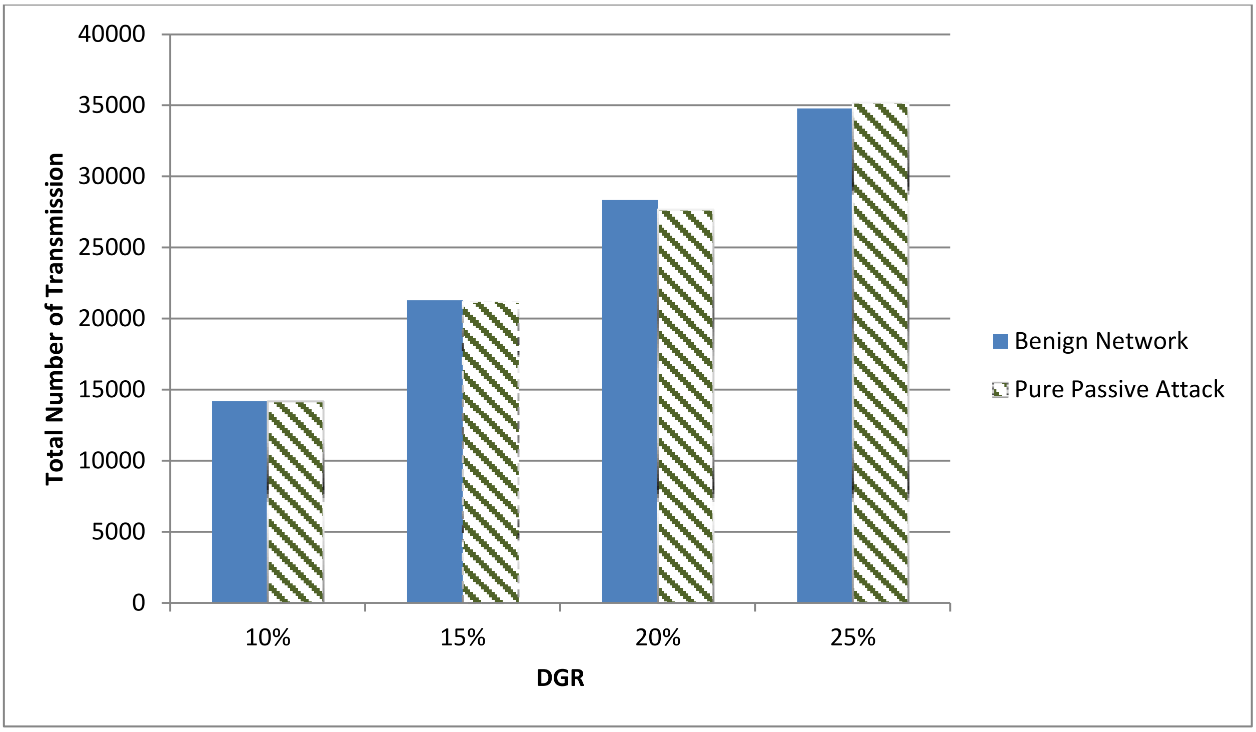

Figure 4.12: $total_{D_F}$ vs. DGR for Networks under Pure Passive Attack (6 Malicious Nodes) and Benign Networks ($B = 10$, $P_S = 0.5$ and $L = 10$)

4.3.6. Network under Active Attack

In Section 3.1, the assumption is given that the contexts of data packets collected by the mobile sink are published in public. With the existence of this assumption, we have conducted *Active Attack* and processed simulations to observe the resilience of the proposed scheme in terms of trajectory privacy. In Active Attack model, an attacker deploys her own mobile sensor nodes into the network area with her own generated data packets. Data packets of the malicious nodes are distributed through the network with L_R, equals to L_M, number of different nodes desired by active attacker to keep copy of data.

In case of malicious nodes receive data packets from other nodes, scheme rules are followed (for received packets, L is taken into consideration). The attacker can also trace her own packet after forwarding. Thus, she can check possibly modified L_R values by overhearing the forwarded packets during the intermediate phase of the packet distribution (See Section 3.2.5 for details of intermediate phase of the packet distribution). By doing so, she can learn which other mobile sensor nodes keep a copy of her own data packet. In the end, by analyzing the network report that is assumed to be published regularly, she has the information about if her packet is collected by the mobile sink. In addition to that, the attacker has location information of mobile sensor nodes that have kept a copy of her data packet. Precisely, she knows L_M number of sensor nodes stored her packet and her packet has been collected by the mobile sink.

We have processed simulations with various values of L_M while keeping the other parameters fixed ($B = 10$, $P_S = 0.5$,$L = 10$ and $DGR = 0.15$) and one malicious node is deployed. Figure 4.13, shows the relationship between the total numbers of benign sensor nodes participated in the delivery of the malicious node's data packet and L_M, the number of different nodes desired by active attacker to keep copy of data. Results show that for $L_M = 1$, the attacker finds out one benign mobile sensor node that has interacted with the mobile sink for sure. For $L_M = 10$, attacker learns that a portion of out of 10 mobile sensor nodes has interacted with the mobile sink node for sure, but she doesn't know how many and which of these nodes have interacted with the mobile sink node. The actual number of benign sensor nodes delivered the copy of malicious data to the mobile sink is 5, but the attacker does not have this information.

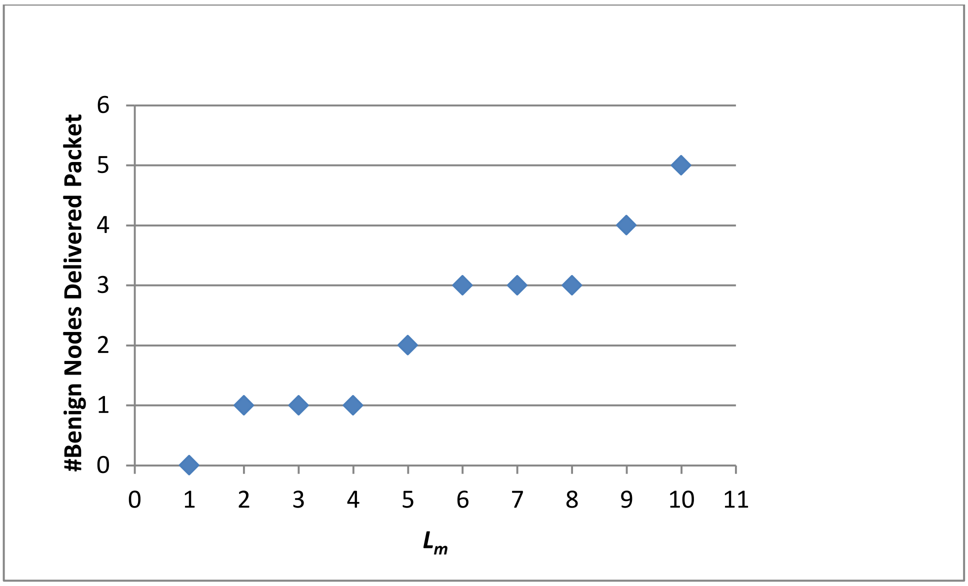

Figure 4.13: L_M vs. Total Number of Benign Nodes Participated in Delivery of Malicious Packet for network under active attack with one malicous node ($B = 10$, $P_S = 0.5$,$L = 10$ and $DGR = 0.15$)

In case of the mobile sink interacts at least with one of the sensor nodes among L_M sensor nodes, the probability of interaction with the mobile sink for a specific sensor node is calculated as follows.

$$\frac{2^{L_M-1}}{2^{L_M}-1} \tag{5}$$

The limit of (5) as L_M approaches to infinity is 0.5 as shown below.

$$\lim_{L_M\to\infty}\frac{2^{L_M-1}}{2^{L_M}-1} = 0.5 \tag{6}$$

Expected number of nodes that have interacted with the mobile sink node is given below.

$$E[N_I] = 0.5 \times L_M \quad (7)$$

If an attacker deploys a node with the value of L_M , she may assume that $L_M/2$ of L_M nodes have been interacted with the mobile sink node. If we check Figure 4.13, we can conclude that the assumption of the attacker holds. For instance, for $L_M = 10$, 5 benign mobile sensor nodes have interacted with the mobile sink. However, attacker does not know which $L_M/2$ nodes are these nodes.

In the end, all these derivations are not too much useful information for the attacker. First of all, (7) also holds for the network with 100% delivery rate. In a network with %100 data delivery rate, a benign node's packet will be delivered for sure. It is assumed that an attacker can overhear and trace route all network traffic between mobile sensor nodes. Thus, instead of deploying a malicious node, attacker can just assume that the benign node is the node she deployed and use this node for her analysis. That is to say, all benign sensor nodes are acting as malicious nodes, as the attacker has the entire network with her own nodes and setting L_R to L. She has all these observations for each packet, but cannot find out which of these nodes that have actually interacted with the mobile network. On the other hand, mobile sensor nodes do not do anything different where all packets are delivered with 100% rate and where no mobile sensor sink node traveled around the sensor area, which brings data delivery rate to 0%. Consequently, the attacker derives the problem of finding out the trajectory of mobile sink into the problem of finding out which L_M nodes is part of the trajectory. The probability of one mobile sensor node to be in the trajectory of the mobile sink is 50%. Thus, our scheme is resilient against network under active attack.

4.3.7. Performance Difference of Our Scheme and Ngai et. al.

In Figure 4.14, we have compared our proposed scheme with the approach studied by Ngai et al. in [19]. The results show that both approaches yield high D_{DR}, data delivery rate.

However, our scheme converges faster since approach of Ngai et al. ignores if a packet stored or not stored by an intermediate node and decrements L. Thus, our approach achieves desirable D_{DR} values around $L = 10$ where their approach achieves around $L = 20$.

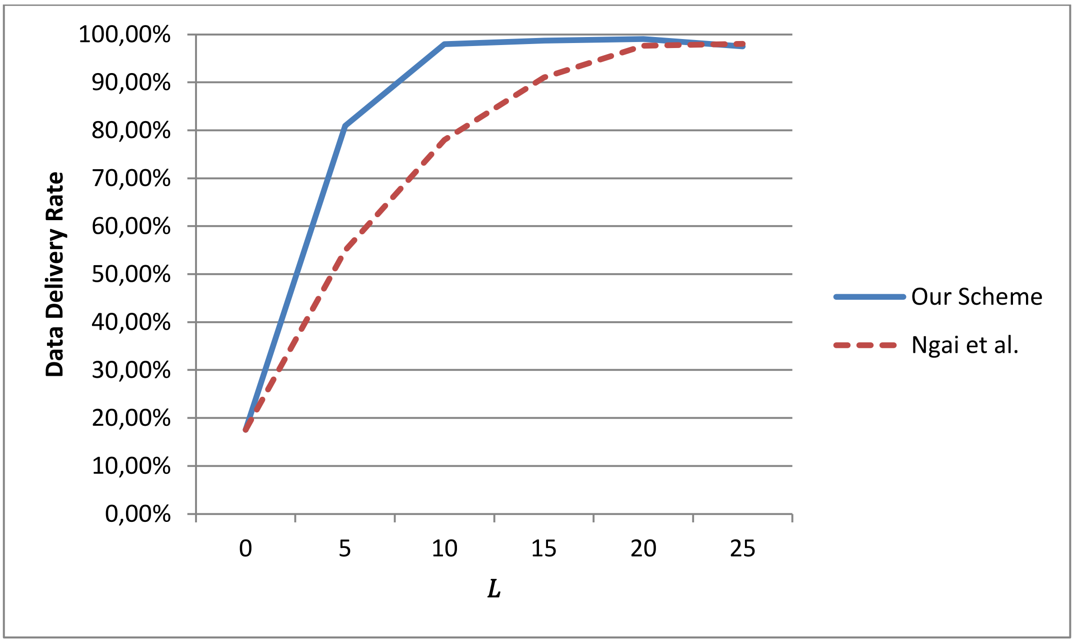

Figure 4.14: Our scheme vs. Ngai et al. in terms of D_{DR} ($B = 10$, $P_S = 0.5$ and $DGR = 0.15$)

5. CONCLUSION

In this thesis, we highlight a new type of security challenge for mobile wireless sensor networks, the trajectory privacy of mobile collector nodes. To the best of our knowledge, there has been no approach proposed in order to maintain the trajectory privacy of the mobile collector nodes. We have proposed an abstract network scheme to maintain trajectory privacy of mobile sink(s) for mobile wireless sensor networks with mobile sink with mobile sensor nodes network architecture. Our scheme is based on randomly distributing the data packets among the network without taking account into the trajectory privacy of the mobile sink node.

We have performed simulations and analysis to evaluate the proposes scheme. The results show that with fine tuning of scheme parameters, data delivery rate reaches up to 100%. The network yields a deterministic communication overhead that can be maintained at desirable ratios with the configuration of scheme parameters. We have also analyzed our scheme against traffic analysis attack and observed that our scheme is resilient to these kinds of attacks. On the contrary, it is observed that the traffic analysis can be used as an intrusion detection tool due to the deterministic behavior of the network in terms of communication overhead. We have proposed two different attack models (*pure passive attack* and *active attack*) with wise assumptions in favor of attackers and have shown that our scheme is also resilient against these types of attack models.

6. REFERENCES

[1] J. Yick, B. Mukherjee, D. Ghosal, “*Wireless sensor network survey*”, *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 52, pp. 2292-2330, 2008.

[2] Z. Yu, X. Fu, Y. Cai, and M. C. Vuran, "*Distributed Vision Graph Construction in Wireless Multimedia Sensor Networks*", *Journal of Sensors*, vol. 11, issue 3, pp. 3381-3400, 2011.

[3] L. Vieira, U. Lee, and M. Gerla, “*Phero-trail: a bio-inspired location service for mobile underwater sensor networks*”, *IEEE Journal on Selected Areas in Communications*, vol. 28, issue 4, pp. 553-563, 2010.

[4] R. Di Pietro, L. V. Mancini, C. Soriente, A. Spognardi, and G. Tsudik, “*Catch me (if you can): data survival in unattended sensor networks.*”, *Proceedings of the 2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications*, pp. 185–194, 2008.

[5] K.-S. Hung, K.-S. Lui, and Y.-K. Kwok, *“A trust-based geographical routing scheme in sensor networks.”, Proceedings of IEEE WCNC ’07*, pp. 3123–3127, 2007.

[6] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, *“Enhancing source-location privacy in sensor network routing,”*, *Proceedings of the 25th IEEE International Conference on Distributed Computing Systems (ICDCS ’05),* pp. 599–608, 2005.

[7] “*WWWF - the conservation organization*”, *http://www.panda.org/.*

[8] J. Deng, R. Han, and S. Mishra, *"Countermeasures against traffic analysis attacks in wireless sensor networks", Proceedings of the 1st IEEE International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm '05),* pp. 113 – 126, 2005.

[9] C. L. Barrett, S. J. Eidenbenz, L. Kroc, M. Marathe, and J. P. Smit, *"Parametric probabilistic sensor network routing,", in Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications,* 2003.

[10] Z. Cheng and W. Heinzelman, *"Flooding Strategy for Target* Discovery *in Wireless Networks" , in proceedings of the Sixth ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2003)*, 2003.

[11] C. Intanagonwiwat, R. Govindan, and D. Estrin, *"Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks", in Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networks (MobiCOM),* 2000.

[12] H. Lim and C. Kim, *"Flooding in Wireless Ad-hoc Networks", in IEEE Computer Communications*, vol. 24, issue 4, pp. 353-363, 2001.

[13] D. Braginsky and D. Estrin, *"Rumor routing algorthim for sensor networks", in Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications,* 2002.

[14] P. Th. Eugster, R. Guerraoui, S. B. Handurukande, P. Kouznetsov, and A.-M. Kermarrec*, " Lightweight probabilistic broadcast", ACM Transactions on Computer Systems (TOCS),* vol. 21, no. 4, pp. 341 – 374, 2003.

[15] M. Shao, Y. Yang, S. Zhu, and G. Cao, *"Towards statistically strong source anonymity for sensor networks", in Proc. IEEE INFOCOM '08,* pp. 51 –55, 2008.

[16] K. Mehta, D. Liu, and M. Wright, *"Location privacy in sensor networks against a global eavesdropper",in Proc. IEEE Int. Conf. on Network Schemes (ICNP '07)*, pp. 314 –323, 2007.

[17] Y. Jian, S. Chen, Z. Zhang, and L. Zhang, "*Protecting receiver-location privacy in wireless sensor networks", in Proc. IEEE INFOCOM '07*, pp. 1955–1963, 2007.

[18] R. Han, and S. Mishra, "*Decorrelating wireless sensor network traffic to inhibit traffic analysis attacks", Pervasive and Mobile Computing (Elsevier)*, vol. 2, no. 2, pp. 159 – 186, 2006.

[19] E. C. H. Ngai, and I. Rodhe, *"On providing location privacy for mobile sinks in wireless sensor networks", In MSWiM'09: Proceedings of the 12th ACM international conference on modeling, analysis and simulation of wireless and mobile systems*, pp. 116–123, 2009.

[20] D. Chaum, *"Untraceable electronic mail, return addresses, and digital pseudonyms", Communications of the ACM*, vol. 24, pp. 84–88, 1981.

[21] M.Reed, P. Syverson, and D. Goldschlag, *"Anonymous connections and onion routing", IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 482–494, 1998.

[22] "Mixmaster remailer", "*http://mixmaster.sourceforge.net/".*

[23] M. Gruteser and D. Grunwald, *"Anonymous Usage of Location-based Services through Spatial and Temporal Cloaking", in Proceedings of the International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2003.

[24] *"OMNeT++ Network Simulation Framework", http://www.omnetpp.org/".*

[25] Y. Ren, V. Oleshchuk, F. Y. Li and X. Ge, *"Security in Mobile Wireless Sensor Networks – A Surver", Journal of Communications*, vol. 6, pp. 128-142, 2011