Wireless Pers Commun (2014) 75:1115–1140
DOI 10.1007/s11277-013-1411-6

Privacy Provisioning in Wireless Sensor Networks

Manjusha Pandey · Shekhar Verma

Published online: 23 September 2013

Abstract Privacy is a necessary component of any security discussion. Privacy and security must be considered separately as well as together. Threads of privacy are visible throughout the procedure of building security solutions for any system. The notions for privacy also play an important role in the technical implications of privacy preservation within any network or its subsystems. Analyzing the mitigations and protections for privacy are considered in privacy notions. The privacy notions being theoretical systems and identified vulnerabilities in the sensor networks not being mapped to such privacy preservation notions necessities the analytical review of privacy provisioning in wireless sensor network (WSN) being formalized within a framework consisting or the vulnerabilities associated with each component of the network and mechanisms of privacy preservation along with the privacy notions. The current research paper provides an analytical review of the privacy provisioning in WSNs with the perspective of development of a proposed framework for privacy notions and quantitative as well as qualitative measures associated with the privacy preservation in sensor network mapped with the various techniques implemented for privacy preservation of different components of the network and the network as a whole. This analytical review has been done on the basis of set of beliefs that an adversary has while launching an attack on the network. Further the existing techniques for privacy preservation of receiver and source node location, location of sink node, traffic analysis prevention and preserving temporal privacy in WSN have been analyses on the basis of adversary's set of beliefs' mitigated by them and the notion of privacy implemented by these techniques like k-anonymity, l-diversity or t-closeness. The present effort aims to provide the researchers with an insight of the new concept of belief mitigation for privacy provisioning in WSN.

Keywords WSN · Privacy · Change of beliefs · Belief mitigation · k-Anonymity · *l*-Diversity · t-Closeness

M. Pandey (✉) · S. Verma
Indian Institute of Information Technology, Allahabad, India
e-mail: manjushapandey82@gmail.com; rs58@iiita.ac.in

1 Introduction

Evolved adhoc networks with self configuring capacity are called as wireless sensor network (WSN) [1].The constituent of these networks are tiny equipments called sensor nodes or motes that use radio signals to communicate among themselves. As they are tiny equipments senor nodes are to be deployed in large quantity to monitor the physical world and communicate the sensed data to various real life applications. Also we can say that WSN acts as an interface between the real physical world and virtual world of networks.

Being an interface they help the applications to observe physical spaces at a finer resolution over the spatial–temporal scales. Due to its unique characteristics the WSN are favored for a wide variety of applications ranging from industry to education to operational. With the advances in nano technology the tiny sensor nodes now a day are much more efficient than their predecessors. The functioning of a node includes sensing of data, processing of data within the network as well as communication of data, fusion and correlation tasks. The devices used in sensor network include seismic, acoustic, magnetic, thermal, infrared devices. Nodes in WSNs perform a distributed algorithm to route sensed data through its neighboring nodes, as per the requirements of the application. The physical configuration of node placement is important for relaying of data though there are various other parameters effecting the transmission like obstructions, interference, environmental factors, antenna orientation and mobility of sensor nodes [2,3]. Ultimate Security and privacy requirements in sensor networks are preventive measures that based on preventive measures that ensure the prevention of attacks. Making attacks difficult along with authentication, confidentiality, and integrity of sensed data. The detective measures that include detecting the attack along with distinction between attack and network failure. Reactive measures are taken to subsidize the attacks by modifying level of security or to either neutralize of counter the attack. Sensor nodes are susceptible to physical capture, but because of their targeted low cost, tamper-resistant hardware are unlikely to prevail. Sensor nodes use wireless communication, which is particularly easy to eavesdrop on [4].

Similarly, an attacker can easily inject malicious messages into the wireless network [5]. Advanced Shear design complexity and high rate of energy consumption for physical tamper proofing makes it less feasible in the case of wireless senor networks. The WSN is more susceptible to denial of service attacks as they use wireless radio transmission with additional constraint on size, cost and limited energy [6]. Erratic topology of WSN facilitates the adversary for various types of link attacks like passive eavesdropping and active interfering. A number of attacks from various perspectives are possible for a sensor node revealing the secret information, message interfering and node impersonation etc. with the requirement of robustness in the implemented security schemes for very large scale deployments. Hence a promising approach is to use efficient cryptographic approaches; most of these schemes use symmetric key cryptography [7]. Managing key distribution is also unique in WSN because of the constraint like small memory capacity that makes centralized keying not possible to use. Almost all the recently used standard security protocols are designed and developed for two party setting and do not scale to large number of nodes in the network [8,9].

Maximization of security level and minimization of resource consumption is the conflicting goal to be achieved particularly for the case of WSN. Any of the efficient solutions need to give a good optimization between these two. Asymmetric cryptography is not used as it is too expensive for the case of energy WSN [10,11].

2 Privacy in WSN: Issues and Challenges

Designing security schemes for WSN pose significant challenges because of the adhoc nature of the network. WSN specifically poses many constraints in contrast to the traditional computer networks. Some of characteristics of WSNs majorly affecting security schemes in WSN include

(a) *Wireless Communication Medium*
The broadcast nature of wireless medium makes it inherently much more vulnerable as it eases to eavesdrop. The communications on wireless medium could easily be intercepted, altered or replayed by the attackers. Also the wireless medium makes it easier for the adversary to intercept valid packets and inject malicious packets instead. This problem of malicious packets injection is not unique to WSN but also existed in the traditional networks been though WSN are more susceptible to it.

(b) *Ad-Hoc Nature of Node Deployment*
Sensor networks do not have any statically defined structure as the network topology is erratic because of node failure, addition of nodes, and mobility in the nodes. As the sensor nodes may fail the network need to have self configuration property. And the privacy schemes developed must be in accordance with this dynamic environment.

(c) *Hostile Environment*
The sensor nodes function in a hostile environment which is also a challenging factor for privacy provisioning in WSN. Being in hostile environment the motes face possibility of being captured or destructed by the attacker. The attacker can extract useful information from a node by capturing it and disassembling it. Thus highly hostile nature of environment in which sensor nodes are deployed is a major challenge of privacy provisioning in WSN.

(d) *Resource Constraints*
Resource consuming security and privacy mechanisms are considerable challenge to be implemented in a resource constrained WSNs. Hardware constraints like restricted bandwidth, computational complexity and limited memory space are major threats to the efficiency of privacy provisioning schemes in WSN. Energy being the most precious resource for sensor nodes with energy expensive communication extends a need for special effort to make privacy provisioning energy efficient.

(e) *Immense Scale*
The traditional networks do not tend to have hundreds to thousands of nodes as is the case with WSN. This proposed scale of sensor networks prove to be a significant challenge for security mechanisms. Hence providing security and privacy on such networks is equally challenging as the mechanisms need to be scalable with the large number of nodes in network along with higher computation and communication efficiency.

(f) *Communication Unreliability*
Most of the security and privacy provisioning protocols rely heavily on some predefined protocol that in turn is dependent on the data communication patterns in WSN. Unreliability of communication patterns is due to the unreliable data transfer. Conflicts in data transfer and latency associated with it. The packet transmission by sensor nodes is based on routing that is connectionless and thus inherently unreliable. The broadcast nature of transmission in WSN may lead to communication conflicts. While the multi-hop routing, network congestion and node processing lead to latency in data communication. Hence achieving synchronization among sensor nodes is difficult.

(g) *Unattended Operation*
Sensor nodes are left unattended for long period of time depending on the function of sensor network. The unattended sensor nodes are vulnerable to three main cautions: *Exposure to Physical Attacks* The open environment in which sensor nodes are deployed are also open to adversary, bad weather and many unexpected natural calamities. Hence the sensor network suffer with higher probability of physical attacks then the traditional networks. *Managed Remotely* Remotely managed sensor networks make the detection of physical tampering difficult with many physical maintenance issues. *No Central Management Point* Lack of central management point increases the vitality o sensor networks. The network management becomes difficult, inefficient and fragile if designed incorrectly.

3 Wireless Sensor Networks Components Insight

This section provides a detail about various entities within a sensor network that play an important role in the privacy provision of these networks. The section further details about the role of adversaries and their type along with the privacy vulnerabilities associated with the different entities of sensor network.

3.1 Sensor Node

Sensor node [29] is the basic component a sensor network is made up of. Sensor node a low cost tiny devices that are spread in a number of hundreds over the application area randomly. Because of the ad hoc nature of sensor network there is no predefined topology sensor nodes are randomly deployed and follow a self configuration policy with the addition of new nodes or die out of older nodes as the sensor nodes are battery powered resource limited devices that die up in due course of time. Following Fig. 1 describes various components of a sensor node and processing associated with these components of the node also called as sensor motes.

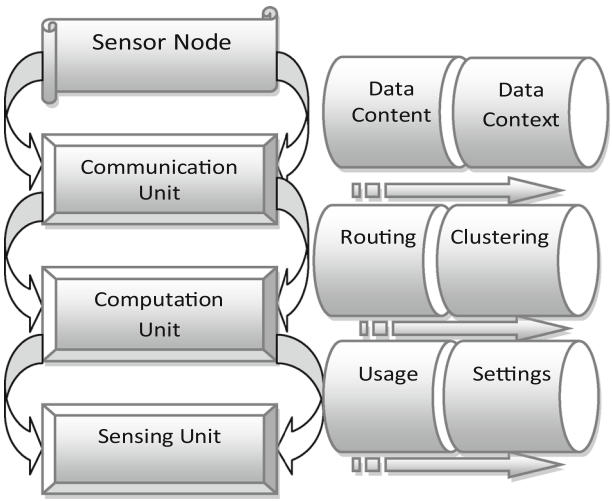

Fig. 1 Components of a sensor node (mote)

The self configuration nature of sensor networks makes them more useful to be deployed in regions having difficult terrains where the legacy networks could not be deployed. In a way sensor networks because of its basic unit sensor have made the mankind able to reach to otherwise very difficult geographical parts of the globe. A sensor node is composed of three constituents the communication unit, the computation unit and the sensing unit. The sensing unit is responsible for sensing the application specific data for which the network has been deployed and this has to be programmed within the node. The computation unit is responsible for all in network computations being ad hoc network having restricted or no infrastructure available the nodes themselves act as routers and other components like cluster heads. All these in network processing are done by the computation unit of the sensor node. The communication unit of sensor node is responsible for communication of data within the network. Sensor networks are a multi hop network which means data is communicated hop by hop from the source node to the receiver node.

3.2 Base Station

Base station acts as the central data repository for the WSN where all sensors communicate their data to. The base station then propagates the data collected and communicated by the sensors to the specific applications for further processing and analysis of the sensed data. In a way the base station acts as a gateway interface between the sensor network and the application oriented service provider. As the gateway the base station needs to be either user specific or application specific depending upon the services being provided by the service provider to which the base station is propagating data to. The base station is also responsible for some in network processing that may be internetwork or intra-network depending on the application (Fig. 2).

4 Privacy Vulnerabilities Associated with the Network Components

Each and every entities of the WSN possesses different set of attributes and contribute to the working of network as a whole in entirely different manner hence the privacy related

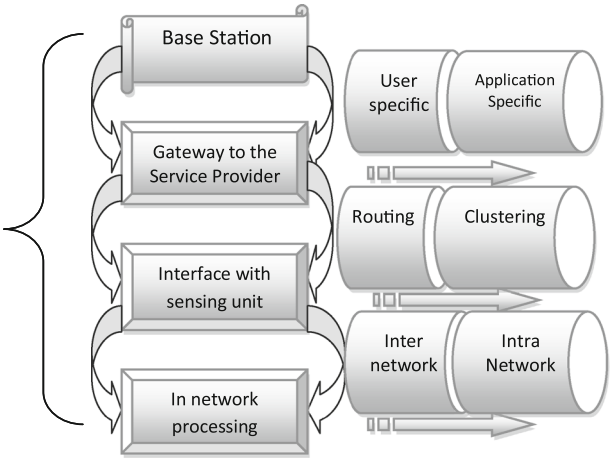

Fig. 2 Base station

Fig. 3 Adversary roles

vulnerabilities related to these entities also differ a lot, this section provides a detail about the type of adversaries and vulnerabilities associated with each and every entity in the WSNs.

4.1 Adversary Roles

The adversary is defined as any foreign unauthorized entity in the network that is trying to violate the proper functioning and protocols of the network. The adversary are said to have different roles of active attackers, passive attackers and global adversary. The active attackers are set of adversaries that modify or fabricate the sensed data by compromising the sensor node or faking as a part of the node themselves. Passive attackers are the set of attackers that do not participate in the network processing instead perform the passive analysis of functioning of network and design the attacks accordingly. While the global adversaries are more powerful role of attackers often having enhanced capabilities through which they can launch both active as well as passive set of attacks on the network (Fig. 3).

Identifying vulnerabilities is one of the difficult processes that refer to trying and anticipating how the attacker may try to exploit the network and its functioning for its own advantages. The analysis of vulnerabilities is dual process of examining the network loopholes along with the potentials of adversary to start with the examination of vulnerabilities of the in the system we must start with the top of the target list that is the network as whole. When we talk of the network itself we need to work on the building trust within the network privacy preservation in the network. For the trust development within the network as a whole the developers need to design and implement preventive measures that may be more of theoretical and implemented afterwards. After creating a theoretical model of the attacks and vulnerabilities possible within the network the particular component of resource is identified with the associated vulnerabilities. The two basic components of the sensor networks include sensor node and base station. The security and privacy vulnerabilities associated with base station may be network specific, user specific and application specific that need o be dealt with the detective measures. While the sensor node vulnerabilities include vulnerability associated with the device itself as the network is an unattended network leading to the vulnerability of sensor nodes being physically tampered (Fig. 4).

5 Privacy Provisioning

Intuitively, privacy breach is defined as the information gain by an unwanted method, procedure or entity. The basic principles of privacy include: *Accountability* all the entities and procedures involved in data exchange maintain the responsibility of purpose identification for the information in exchange. *Purpose Identification* The information in exchange must have limiting use, disclosure and retention. There must be no usage retention or disclo-

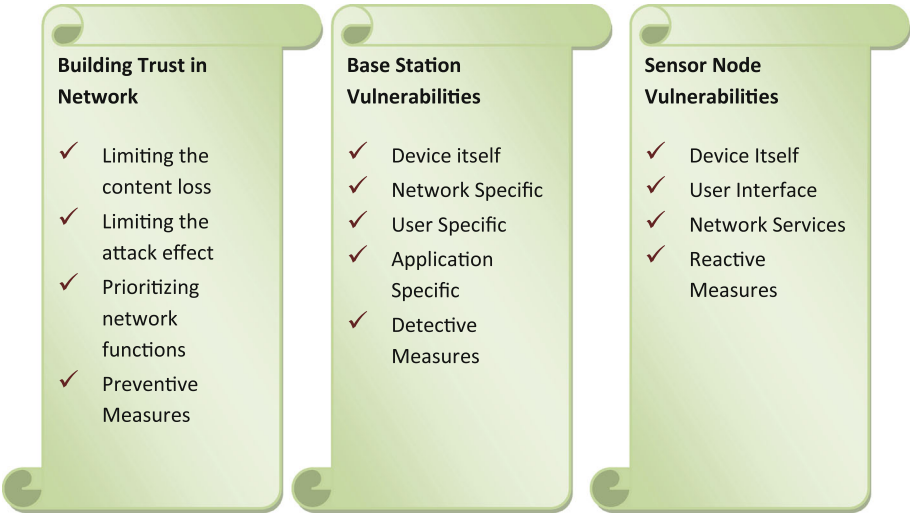

Fig. 4 Vulnerabilities associated with various entities of wireless sensor network

sure other than identified purposes of the information in exchange. *Challenging Compliance* Compliance of the above mentioned privacy principles are challenging.

5.1 Privacy Provisioning: Quantitative Metrics

Privacy preservation quantification [31] is the quantity used to measure privacy that indicates numerically how closely the original attribute values could be deduced form the released set of information otherwise within the network. For example if an adversary knows that the set of neighbors to a sensor node always contains three nodes then the disclosure of neighboring node attribute is reduced by a conditional differential entropy value. The key privacy measure could be based on the differential entropy of any random variable given as:

$$h(N) = -\int_{DN} f_N(b) \log_2 f_N(b)\, db \tag{1}$$

where:

$$h(N) \text{ is the differential entropy of } N\, DN \text{ is the domain of } N$$

For a random variable U distributed uniformly between 0 and b

$$h(U) = \log_2(b) \tag{2}$$

And for $b = 1, h(U) = 0$

That represents the uniform distribution. A measure for privacy inherent in a random variable N is given by $2^{h(N)}$.

So a random variable U distributed uniformly over 0 to b will have privacy denoted by $\prod(U)$

Given as:

$$\prod(U) = 2^{h(N)} = 2^{log_{2(b)}} = b \tag{3}$$

Given one random variable M the conditional entropy of N becomes

$$h(N|M) = -\int_{D(N,M)}^{\cdot} f_{m,n}(b,a)\, log_2 f_{N|M=a}(b) \tag{4}$$

Hence the average conditional privacy for H given M is denoted by $\prod(N|M)$ and is given by

$$\prod(N|M) = 2^{h(N|M)} \tag{5}$$

Conditional Privacy loss for N give M denoted by ρ (N | M) can be given as

$$\begin{aligned} \rho(N|M) &= 1 - \prod(N|M) \prod(N) \\ &= 1 - 2^{\mathrm{h(N|M)}}/2^{\mathrm{h(N)}} \\ &= 1 - 2^{-I(N:M)} \\ \text{Here I (N : M)} &= \mathrm{h(N)} - \mathrm{h(N|M)} \\ &= \mathrm{h(M)} - \mathrm{h(N|M)} \end{aligned} \tag{6}$$

Is called the mutual information [32] between the two random variables N and M hence ρ(N|M) is defined as a fraction of privacy lost because of the disclosure of M.

The quantization of privacy alone is not sufficient without the study and analysis of quantity as well as quality of the privacy preservation offered by various privacy preservation mechanisms that could be calculated quantitatively based on the above metrics and qualitative analysis is discussed in the next section.

Other related privacy metrics include

Average privacy: this metric measure the reconstruction probability of any random data value within the information in exchange within the network.

Worst case privacy: This metric measure the maximum reconstruction probability across all the values within the information in exchange within the network.

Re-interrogated Privacy: A common system in which the adversary does not have access to original data along with the additional situations where in the adversary could not re-interrogate for the original data adds up to the privacy preservation.

Amplification Privacy: It's a particularly strong notion of privacy that guarantees strict limits on privacy breaches of individual entities information, independent of the distribution of true data.

5.2 Privacy Provisioning: Qualitative Metrics

5.2.1 Privacy Versus Utility

Privacy Preservation techniques are evaluated with two basic metrics: level of privacy preservation guaranteed and the level of application specific data utility preserved that could be measured as the loss of accuracy of the data classification and its application oriented utility preserved. Generally to preserve the content privacy the original data is perturbed as newer forms to enhance the difficulty of estimation of original data to unauthorized users. Another

way to preserve the context privacy of the data is K-anonymity through which the data is disguised among the similar set of data to preserve its context privacy. K-anonymity is a popular way of measuring the level of privacy by enabling the measurement of effective estimation of the original data record to a k-grouped data record. Data utility typically refers to the amount of efforts required to rebuild the original data and its application orientation after perturbation of anonymity used for privacy preservation. Hence to have an effective and efficient privacy preservation mechanism data privacy along with the data utility must be considered for the design of such mechanisms. The utility measures for privacy preservation include: *Query answering accuracy, classification accuracy, distribution similarity, generalization heights and discern ability.*

5.2.2 Personalized Versus Utility Based Privacy

In general data privacy is the measure of degree of uncertainty according to which the original private data can be inferred. As not all the entities of a network need to have similar levels of concerns for privacy thus there must be provisions for variations with personalized privacy for different entities of the network. This practically means that the value of K-anonymization is not fixed but varies with varying records of different entities. From the technical point of view we must have the value of K for anonymization that varies with different set of records. Another way to model personalized anonymity may be to specify different level of privacy for sensitive attributes. This technique assumes that an individual can specify a node of the domain generalization hierarchy in order to decide the level of anonymity that the system can work with. But the major disadvantage of personalized privacy is that the privacy preservation algorithm complexity increases with the increased levels of personalized privacy preservation implemented and it also affects the utility of data and its application orientation very much.

5.2.3 Centralized Versus distributed privacy

The key goal of distributed privacy provisioning is to allow computations of useful aggregates over the entire data set without compromising the privacy of individual data sets within the different entities of the network. Thus different entities of the network must collaborate to obtain the original data result. The major disadvantage of distributed privacy preservation is that it overlaps closely with the cryptography for determining secure multi-party computations. These mechanisms are repeated many times over the given function evaluations hence the computational effectiveness of the approach is very important. While the centralized privacy preservation though is much simpler than the distributed approaches they are not very much feasible for resource limited networks like sensor networks. A related problem with the centralized privacy provisioning is that it provides a central point of failure within the network adding to the vulnerability of the entire network processing being dependent on single central privacy providing entity.

5.2.4 Change of Beliefs

The adversary has some prior belief about the sensitive attribute value of an individual. And to obtain an effective privacy these prior believes need to be transformed to some distracting posterior beliefs. Information gain by an adversary can be represented as the difference between the posterior belief and the prior belief. The novelty of our approach

is that we separate the information gain into two parts: that about the whole population in the released data and that about specific individuals. To motivate the approach, following thought of experiment was performed: First an observer has some prior belief $B0$ about an individual's sensitive attribute. Then, in a hypothetical step, the observer was given a completely generalized version of the data table where all attributes in a quasi-identifier were removed (or, equivalently, generalized to the most general values). The observer's belief was influenced by **Q**, the distribution of the sensitive attribute value in the whole table, and changed to $B1$. Finally, the observer was given the released table. By knowing the quasi-identifier values of the individual, the observer was able to identify the equivalence class that the individual's record was in, and learned the distribution **P** of sensitive attribute values in this class. The observer's belief changed to $B2$. The l-diversity requirement was motivated by limiting the difference between $B0$ and $B2$ (although it did so only indirectly, by requiring that **P** had a level of diversity). It was chosen to limit the difference between $B1$ and $B2$. In other words, it was assumed that **Q**, the distribution of the sensitive attribute in the overall population in the table, was public information. It did not limit the observer's information gained about the population as a whole, but limited the extent to which the observer could learn the additional information about the specific individuals. To justify the assumption that **Q** should be treated as public information, it was observed that with generalizations, the most one could do was to generalize all quasi-identifier attributes to the most general value. It showed that as long as a version of the data was to be released, a distribution **Q** would be released. It was also interpreted that if one wanted to release the table at all, one intended to release the distribution **Q** and the distribution was what made the data in the table useful. In other words, one wanted **Q** to be public information. A large change from $B0$ to $B1$ meant that the data table contained a lot of new information, e.g., the new data table corrected some widely held belief that was wrong. In some sense, the larger the difference between $B0$ and $B1$, the more valuable was the data. Since the knowledge gained between $B0$ and B1 was about the whole population, it did not limit the gain. The researcher limited the gain from $B1$ to $B2$ by limiting the distance between **P** and **Q** intuitively, if $\mathbf{P}=\mathbf{Q}$, then $B1$ and $B2$ should be the same. If **P** and **Q** were closer, then $B1$ and $B2$ should be closer as well, even if $B0$ might be very different from both $B1$ and $B2$.

6 Related Work

Privacy preservation in WSNs has been an important point of concern because of the ever increasing adaptability and applicability of sensor networks in day to day life. The advent of concept of smart home, office automation and smart military applications based on WSNs has made privacy issues in WSN of utmost importance. Table 1 below gives a comprehensive survey of some of the privacy provisioning mechanisms proposed and implemented by various researchers.

7 Analytical Review

The privacy measures have been proposed for hiding the data content and its context being propagated. It is done through different privacy preservation techniques. The aim is to develop framework for privacy provision and making the data and its context undistinguishable. These measures of privacy preservation may be mapped to various privacy provisioning techniques

Table 1 Related work

Author/year	Title	Proposed mechanism	Implementation	Limitations
Ouyang et al. [12]	Providing anonymity in wireless sensor networks	Baseline and probabilistic flooding mechanisms	Each sensor to broadcast the data it receives from one neighbor to all of its other neighbors	Needs a cache at every sensor node to store the packet that has already been received so that it can compare duplicate packets and discard them
Xi et al. [13]	Preserving source location privacy in monitoring-based wireless sensor networks	Random walk mechanisms	Data first performs a few steps of random walk from data source, and then, by employing probabilistic flooding scheme, it is transmitted towards base station	The amount of overhead incurred to simulate a source or to redirect traffic randomly and these schemes also introduce a delay in delivering the packets which may not be useful in real time applications
Deng et al. [14]	Intrusion tolerance and anti-traffic analysis strategies for wireless sensor networks	Dummy data mechanism	Fake data packets are introduced to perturb the traffic patterns observed by the adversary	Consumes a lot of bandwidth and hence a higher communication cost
Ozturk et al. [15]	Source-location privacy in energy-constrained sensor network routing	Fake data sources mechanism	One or more sensor node simulates the behavior of a real data source in order to confuse the adversaries	Major challenge for the design of this technique is how to simulate the behavior of data sources without being detected
Deng et al. [16]	Countermeasures against traffic analysis attacks in wireless sensor networks	Routing with multiple parent	Multiple-parent scheme was introduced to balance the traffic load between parents and children, such that an adversary cannot easily identify which node is nearer to the base station	A malicious node can claim a low level value to attract traffic from other nodes, or it can use unfair media access control mechanisms to occupy the wireless channel
Xi et al. [17]	Preserving source location privacy in monitoring-based wireless sensor networks	Routing with random walk	Simple random walk divides the neighbors of a sensor into two lists—closer and further lists according to the hop count from the base station. When sensor forwards data, it randomly selects a next hop neighbor from one of those two lists	Large amount of overhead incurred to simulate a source or to redirect traffic randomly and these schemes also introduce a delay in delivering the packets which may not be useful in real time applications

Table 1 continued

Author/year	Title	Proposed mechanism	Implementation	Limitations
Syverson et al. [18]	Anonymous connections and onion routing	Deco-relating parent–child relationship by randomly selecting sending time	Period of time T is divided into m slots when there are one parent and (m−1) children for a sensor	A malicious node can claim a low level value to attract traffic from other nodes, or it can use unfair media access control mechanisms to occupy the wireless channel
Xi et al. [19]	Preserving source location privacy in monitoring-based wireless sensor networks	Hiding traffic pattern by controlling transmission rate	Sensor close to the base station needs to not only send its own data but also relay data from sensors further away from the base station, and therefore features a high transmission rate	Rate needs to be controlled at every sensor node but to implement this realistically every sensor node must have a buffer so that it can delay the packet and there is a uniform rate at every node
Hong et al. [20]	Elective probabilistic approach protecting sensor traffic	Propagating dummy data	Fake-packet injection is used to prevent an adversary from identifying the real data transmission pattern	Assumption that an adversary cannot distinguish real data from fake data
Przydatek et al. [21]	SIA: Secure Information Aggregation in Sensor Networks	Aggregator nodes, aggregating information requested by a query	Efficient random sampling is used with interactive proofs, protocols for secure computation are used	Efficiently storage of past data and authenticator, Verifier needs to compute many one-way functions to derive current key
He et al. [22]	PDA: Privacy-Preserving Data Aggregation in Wireless Sensor Networks	Cluster-based Private Data Aggregation (CPDA)	CPDA Based on clustering protocol and algebraic properties of polynomials	Accuracy of CPDA dependent on P_c (probability of any node to become cluster head)
		Slice-Mix-AggRegaTe (SMART)	SMART uses slicing techniques and the associative property for mixing	Not efficient for malicious active attackers

Table 1 continued

Author/year	Title	Proposed mechanism	Implementation	Limitations
Kamat et al. [23]	Temporal privacy in wireless sensor networks	Uses adaptive buffering at intermediate nodes to obfuscate temporal information	Privacy enhancement strategies of buffer preemption has been implemented	Is applicable to delay tolerant networks only Implemented for non-intrusive adversary model
Li et al. [24]	Providing source-location privacy in wireless sensor networks	Two-step routing strategy with message routed by, randomly selected intermediate node(s)	Selection of multiple random intermediate nodes is done based on angle and quadrant	Efficient selection of intermediate nodes required
Edith et al. [25]	On providing sink anonymity for sensor networks	Randomized Routing with Hidden Address (RRHA)	Prevents attackers from obtaining the receiver address destination field or predict the location of the sinks by flow of network traffic	Energy consumption is fairer
Shaikh et al. [26]	Network Level Privacy for Wireless Sensor Networks	Identity, Route and Location (IRL) privacy algorithm	Full network level privacy is provided against privacy disclosure attacks	Assumed secure encryption and decryption with no knowledge of network topology to every node
Zhang et al. [27]	GP2S: Generic Privacy-Preservation Solutions for Approximate Aggregation of Sensor Data (concise contribution)	Data aggregation is provided for variety of queries privacy for individual data and aggregate data	Implemented for sensor in-network data aggregation	Not considering the active attackers and in-network intruders

proposed and implemented in WSNs. The three most popular privacy notions as reported may be of k-anonymity [22], l-diversity [23] and t-closeness [24].

k-anonymity

The k-anonymity notion in any data set has been based on the set of quasi-identifier attributes. A Quasi-identifier Attribute is defined as an attribute of the data set that does not identify the subject by itself, instead a combination of similar attributes are required to identify the subject e.g., the parent–child relationship between various sensor nodes. A set of data may have k-anonymity notion of privacy if every record in the data set is made indistinguishable from other k−1 with respect to every set of quasi-identifiers. Therefore, for every k-anonymous set of data there would be combination of values of the quasi-identifiers for which there exist at least k records that share those values. Hence the privacy notion of k-anonymity has been said to be in a data set T if it has satisfied the condition that it is indistinguishable from other $k-1$ records in the set T. k-anonymity notion of privacy provisioning normally creates groups. These groups thus formed become responsible for leaking some vital information due to lack of diversity in the sensitive attribute and it does not protect against the attacks based on the background knowledge. Therefore an improved notion of l-diversity was proposed.

l-diversity

In the equivalence classes formed if, there are at least l "well-represented" values of the sensitive attribute then the data set may said to have l-diversity. The whole data set would be said to have l-diversity if each and every equivalence class within the data set bears l-diversity. A sensitive attribute is one which is kept secret for any individual entity of the data set. It is done not to provide revelation of sensitive attribute to people who have no direct access to the original data like the sensor node id in case of sensor networks. The disadvantages associated with l-diversity is that it is limited in its assumptions with the adversarial knowledge and in principle it effectively assumes all the attributes to be categorical. To improvise these disadvantages of l-diversity a new notion of t-closeness was introduced.

t-closeness

The t-closeness privacy notion is the phenomenon in which an equivalence class is said to have t-closeness when the distance between distribution of the attributes in the whole data and the distribution of sensitive attribute in any equivalence class is no more than a threshold value t. If all the equivalence classes within the dataset have t-closeness then the whole data set is said to have t-closeness.

All the above three privacy notions form the theoretical basis for the privacy measures of any privacy preservation algorithm. The basic characteristics of the three privacy preservation notions considered for the analytical study of privacy preservation methods could be summarized as in the Table 2.

7.1 Privacy Preservation Through Data Aggregation

Data aggregation algorithms aim to gather and aggregate data in an energy efficient manner. This helps in the enhancement of the network lifetime. The WSNs consists of numerous tiny sensor nodes having less power and require lesser power consumption for processing as compared to transmitting data. So it is preferable to do in network processing inside network. This has dual advantage one it and reduces the packet size and is energy efficient thus improving the network lifetime. The following threat model1 presents a threat of being identified to aggregator node.

Table 2 Basics of privacy notions

Privacy notation	K-anonymity [33]	L-diversity [34]	t-closeness [35]
Core concept	At least (K−1) packets with same attributes with respect to a common attribute called quasi identifier	Has L-diverse well represented values for the sensitive attributes in an equivalence class	Concept is to make the difference between the prior believe and posterior believe of adversary higher
Implementation	Provides identity privacy but is unable to limit attribute discloser	Entropy L-diversity: where low entropy of the values must be maintained Recursive L-diversity: ensures most frequent values does not appear frequently	An equivalence class is said to have t-closeness if the distribution of sensitive attribute is no more than a threshold value Key idea is to limit the correlation between quasi-identifier and sensitive attribute
Limitation	Problems are Homogeneity and background knowledge attack	Problem is with the probabilistic inference attack and similarity attack	One major limitation is the concept of utility of data should not be affected by the notation

Threat Model1

Traffic Load imbalance at sink node and cluster heads (B_3)

1. **In_Traffic** = Current node (received (msg));
2. **Out_ traffic** = Current node (sent (msg));
3. **If** (Out_traffic > In_traffic)
4. **If** (current node = = sink node)
5. **Get** (sink node)
6. **Else**
7. **Get** (cluster node or aggregator node)

CPDA: adds randomly generated noise to the raw data for privacy preservation the scheme Depicts t-closeness and K-anonymity as the scheme classifies the senor nodes into categories having specific distributed values based on the function that could be used to generated t-closeness in the set of attributes. Also the scheme depicts K-anonymity if we consider many to one relationship between the cluster head and nodes within the cluster. This co-relation could be used for generating equivalence class. While CPDA does not exhibits l-Diversity as there may not be l-Ddiverse data values to be transmitted in the data payload. *SMART: data is sliced mixed and then aggregated to preserve the content privacy of data* SMART depicts l-Diversity because the data is sliced mixed and aggregated at each node this provides l well represented values to the data being propagated to the base station. Also the scheme depicts K-Anonymity if we consider the nodes slices to be quasi identifiers.

The Table 3 above gives a detail about the set of beliefs that an adversary has if it is going to attack a sensor network having privacy provisioning by data aggregation and the considered set of techniques of privacy preservation through data aggregation. Analysis of the considered set of techniques on the basis of the three privacy notions reveals that SMART does not have t-Closeness as there are variations that could not be distributed with specific threshold value due to slicing and mixing of data.*GPˆ2S: Integer range of the data is transmitted by the node and appropriate aggregation is done at the aggregator node* The scheme has t-Closeness as the scheme follows generalized distribution of integers being transmitted instead of the data providing a range of distribution of values. Also the mechanism has K-anonymity if we

Table 3 Data aggregation based privacy provisioning in WSN

Considered set of beliefs	
$\mathbf{B}_1$	Packet receiving rate of the aggregator node is higher than the ordinary nodes
$\mathbf{B}_2$	Packet sending time of the aggregator node would be higher than the ordinary node
$\mathbf{B}_3$	Packet length would be reduced at the aggregator node
$\mathbf{B}_4$	Pronounced traffic patterns would be there at the aggregator node almost comparable to the sink node for higher node density in the network
Considered techniques of privacy preservation through data aggregation in WSN	
$\mathbf{T}_1$	CPDA: adds randomly generated noise to the raw data for privacy preservation
$\mathbf{T}_2$	SMART: data is sliced mixed and then aggregated to preserve the content privacy of data
$\mathbf{T}_3$	GPˆ2S : Integer range of the data is transmitted by the node and appropriate aggregation is done at the aggregator node
$\mathbf{T}_4$	Bucketing scheme: bucket size attached with the encrypted data is based on generalization

consider the integer range to be quasi identifier. The mechanism does not has l-Diversity as the data transmitted could not be represented to have l-Diverse values. *Bucketing Scheme : bucket size attached with the encrypted data is based on generalization* The bucketing scheme depicts t-Closeness defined as per the distribution of bucket size. Also the scheme represents to be having K-anonymity with bucket size to be the quasi–identifier but the scheme does not have l-Diversity.

7.2 Privacy Preservation Through Anti-Traffic Analysis Schemes

The communication traffic in WSNs is very much dependent on the application for which the network has been established. Types of communication traffic that exist in a sensor network include data traffic, routing discovery traffic, link layer feedback and hello message, etc. Following threat model 2 provides a traffic monitoring attack at the sink node.

Threat Model 2

Traffic rate monitoring at sink node (B_3)

1. **In_Traffic** = Current node (received(msg));
2. **Out_ traffic** = Current node (sent (msg));
3. **If** (Out_traffic > In_traffic)
4. (current node = = sink node);
5. **Get** (sink node)
6. **Else**
7. Current node = Random (neighbor node)

Many other types of anti-traffic analysis threats are also there, Table 4 below gives a set of beliefs' that an adversary has while launching traffic analysis attacks in WSNs. Also the table provides detail of considered set of anti-traffic analysis privacy provisioning techniques for analytical review based on the three privacy notions.

Anti-traffic analysis privacy provisioning in sensor network has a major impact on the content privacy preservation. *Multi parent routing* mitigates the B_3 and B_4 and depicts

Table 4 Anti-traffic analysis privacy provisioning in WSN

Considered set of beliefs	
$\mathbf{B}_1$	Maximum traffic near the base station
$\mathbf{B}_2$	Rate monitoring (i.e. packet receiving rate maximum at the Base Station then sending rate)
$\mathbf{B}_3$	Time correlation could be used to deduce the source and destination
$\mathbf{B}_4$	Backtracking could also help deduce source and destination
Considered techniques for privacy preservation through anti-traffic analysis schemes in WSN	
$\mathbf{T}_1$	Multi parent routing
$\mathbf{T}_2$	Controlled random walk
$\mathbf{T}_3$	Random fake path
$\mathbf{T}_4$	Multiple hotspot generation
$\mathbf{T}_5$	Hiding traffic patterns
$\mathbf{T}_6$	Propagating dummy data

K-anonymity to a larger extent if we consider the selected parent node to be given to the quasi identifier for the equivalence class. Also MPR could be used to provide l-Diversity to some extend as there are multiple parents providing l-Diversity values to the parent attribute of a node. *Controlled Random Walk* The scheme mitigates B_1, B_2, B_3 and B_4 and depicts all the three forms of privacy notions i.e. K-Anonymity l-Diversity t-Closeness. The scheme depicts K-anonymity as some nodes could randomly be selected multiple times for data forwarding making it to be eligible for quasi–identifier of an equivalence class to be defined. l-Diversity is defined to be depicted by the scheme as there are l-Diverse nodes that could be selected for random propagation of data it provides l-Diversity to the defined equivalence class of data forwarding. The scheme has t-Closeness because the rate of random forwarding of data is controlled by a probability hence providing the attribute a probabilistic distribution to be used as a measure of t-Closeness. *Random Fake Path* The generation of random fake paths along with the original path provides l-Diversity followed by transmitted data. While the scheme does not have K-anonymity as it is not necessary that there could be $K-1$ similar fake paths generated for transmitted data. Also the scheme does not follow any specified distribution on the basis of which the fake data paths could be generated thus there is a very little possibility of having t-Closeness in the privacy preservation. *Multiple Hotspot Generation* The scheme mitigates B_1, B_2, and B_3 and B_4. This scheme depicts only K-anonymity with respect to the pronounced traffic generated at $K-1$ different locations within the sensor network. While it does not provide l-Diversity or t-Closeness as we could not predict any probable distribution of hotspots generated in the network or could not give any well represented values to the different hotspots generated. *Hiding Traffic Patterns Through transmission rate control* The scheme has t-Closeness privacy notion for privacy preservation as the traffic patterns are stabilized by controlling the traffic rate. The control of traffic rate is done on the basis of some predetermined distribution that could be utilized to generate t-Closeness in the data propagated. *Propagating Dummy Data* Propagating dummy data packets provide the network with privacy notions of K-anonymity as there may be $(K-1)$ similar dummy packets in the network with similar set of attributes for generation of an equivalence class.

7.3 Privacy Preserving of Location Privacy of Source and Receiver

Confidentiality of the source sensor's location and receiver senor location must be preserved for making the network reliable. This helps the suitable sensor routing protocols to make it difficult for an adversary to backtrack to the origin of the sensor communication. The threat model 3 provides a description of backtracking implemented by an adversary for revelation of source sensor location.

Threat Model 3

Backtracking for source location revelation (B_6)

1. **Attack=** Sink add;
2. **While** (does not capture Source node)
3. **Listen** (next_msg) **and**
4. **If** (receive msg)
5. **If** (Is new msg(receive msg))
6. **Then** Next location = Getimmediate Sender (msg);
7. **Move to** (next location);

Table 5 Source and receiver location privacy provisioning in WSN

Considered set of beliefs	
$\mathbf{B}_1$	Traffic patterns could be used to reveal the source and destination location privacy
$\mathbf{B}_2$	Traffic patterns could be beneficial to deduce the physical configuration of the network
$\mathbf{B}_3$	Largest density of communication activities leads towards the discloser of identity of and location of base station
$\mathbf{B}_4$	Pronounced traffic patterns correspond to highest activity region in the network
$\mathbf{B}_5$	Communication could be initiated by the source node upon the occurrence of any event thus depicting the event generation in the network
$\mathbf{B}_6$	Backtracking the hop by hop data transmission from base station towards the node may reveal the source node location
$\mathbf{B}_7$	Statistical traffic analysis over a period of time may reveal the patterns of occurrence of event in the network
$\mathbf{B}_8$	There is a parent child relationship between two senor nodes if there is short time interval between sending time at a sensor node and receiving data at its neighbor node
Considered techniques of preserving location privacy of the source and receiver nodes	
$\mathbf{T}_1$	Flooding techniques
$\mathbf{T}_2$	Probabilistic flooding
$\mathbf{T}_3$	Random walk
$\mathbf{T}_4$	Dummy data mechanism
$\mathbf{T}_5$	PFS (Probabilistic Filtering Scheme)
$\mathbf{T}_6$	Fake data source mechanism
$\mathbf{T}_7$	PRLA (Phantom Routing with Locational Angle)

Table 5 below presents the set of beliefs' an adversary prevails while attacking the source and receiver location privacy in WSN along with the considered set of techniques proposed and implemented by various researchers for source and receiver location privacy prevention.

Most of the privacy prevention mechanisms of source and receiver location privacy are based on the hiding of traffic patterns generated and received by the sensor node. Like *Flooding Techniques* The flooding technique for location privacy preservation in WSN mitigates B_3, B_4, B_5, B_6 and B_7 and provides K-anonymity notion of privacy preservation as there are $(K-1)$ similar data packets flooded within the network to generate an equivalence class with node is as the quasi identifier. *Probabilistic Flooding* The probabilistic flooding techniques for location privacy mitigates B_3, B_4, B_5, B_6 and B_7 and provides t-Closeness along with the K-anonymity for the same reason as the flooding technique. To achieve t-Closeness we may utilize the probabilistic distribution followed for implementation of probabilistic flooding. This probabilistic distribution could be used to reduce the distribution to a specific value. *Random Walk* Random walk mitigates B6 and poses K anonymity as there would be $K-1$ number of times a node would be selected for random walk. Also the scheme has l-Diversity as there would be l different values of random paths followed by the forwarded packet. *Dummy Data Mechanism* The dummy data mechanism proposed for generation and forwarding of dummy data along with the original packet for privacy preservation of the data. Also the route followed by the data could be disguised under the wrap of traffic generated by the dummy data packets in the network. The scheme provides K-anonymity as there would be $K-1$ similar data packets propagating in the network thus providing sufficient quasi identifier values for the equivalence class of the transmitted packets. Also the traffic patterns would be having $K-1$ quasi identifier values and equivalence class defined on them. *PFS (Packet Filtering Scheme)* As the major disadvantage of dummy data packet generation

scheme is that the a lot of energy is wasted. Sensor networks being an energy constrained network needs to improve the scheme along with the preservation of advantages it provides for privacy protection. Hence packet filtering was introduced before those filter the packets at intermediate nodes before sending to the sink node. The scheme has K-anonymity at the level of dummy data packet generation and l-Diversity and t-Closeness at the proxy packet filtering scheme level. *Fake Data Source Mechanisms* The fake data source mechanism provides l-Diversity as there would be l-Different data sources for the data packets generated and transmitted to be represented in the equivalence class. *PRLA (Phantom Routing with Locational Angle)* Locational angle introduces the inclination angles to direct random walks that is used to priorities the selection of phantom sources leading to a larger angle of arrival. There are set of angle arrivals that may be useful for making the equivalence class of $K-1$ similar angles of arrival to provide K-anonymity. Also based on the different well represented values of angle of arrival we may have l-Diversity and since there would be a proper distribution of range of angles of arrival which could be used for the implementation of t-Closeness in the mechanisms data.

7.4 Privacy Preservation of Base Station Location

Base station is the central point of failure in any wireless WSN, hence if the base station is compromised the hole network functioning can be compromised. There are many techniques proposed and implemented in literature for location privacy preservation of base station. Threat model 4 describes the time correlation attack that can be launched at the base station by an adversary.

Threat Model 4

Time correlation at sink node (B_3)

1. **In_Traffic** = Current node (received (msg));
2. **In_traffic_time** = Current (received (msg)) _time
3. **Out_ traffic** = Current node (sent (msg));
4. **Out_traffic_time** = Current (sent(msg))_time
5. **If** (Out_traffic_time > In_traffic_time +general_out_traffic_time)
6. (Current node = = sink node);
7. **Get** (sink node)
8. **Else** Current node = Random (neighbor node);

Table 6 gives the considered set of beliefs of an adversary that wish to launch an attack against the sink location privacy in sensor network along with the considered set of mechanisms for sink location privacy being analyzed as per the three privacy notions of k-anonymity, l-diversity and t-closeness.

One of the techniques for sink location privacy preservation is *by Changing data appearance by re encryption* The privacy preservation mechanism based on changing data by re-encryption mitigates B_2. Change of data appearance by re-encryption provides us with the different set of well represented values for data encryption that could be used for implementation of l-Diversity. Also the re-encryption would be following a probabilistic distribution. This probabilistic distribution would be useful for the implementation of t-Closeness in case of changing data appearance by re-encryption mechanism. *Flooding Techniques* The flooding techniques for location privacy provide K-anonymity notion to the privacy preservation in

Table 6 Sink location privacy provisioning in WSN

Considered set of beliefs	
$\mathbf{B}_1$	Destroying or isolating the base station may lead to malfunction of the whole network
$\mathbf{B}_2$	Location information o the base station may be included in the payload
$\mathbf{B}_3$	There will be a parent child relationship between the base station and nodes nearer to it
$\mathbf{B}_4$	There will be difference in time interval of sending and receiving data by base station as base station will be receiving message and transmitting beacons only
Considered techniques for privacy preservation of location privacy of base station	
$\mathbf{T}_1$	Changing data appearance by re encryption
$\mathbf{T}_2$	Routing with multiple parents
$\mathbf{T}_3$	Routing with random walk
$\mathbf{T}_4$	Deco-relating parent child relationship
$\mathbf{T}_5$	Hiding traffic patterns
$\mathbf{T}_6$	Propagating dummy data

sensor networks as there may be (K−1) similar data packets flooded within the network to generate an equivalence class with quasi–identifiers. *Probabilistic Flooding* The probabilistic flooding techniques for location privacy provide t-Closeness along with K-anonymity for the same reason as flooding technique. And t-Closeness is achieved has to be done on the basis of the probabilistic distribution followed for implementation of flooding. This probabilistic distribution could be used to have the t-Closeness. *Random Walk* Random walk for source location privacy poses K-anonymity as there would be (K−1) number of times a node would be selected for random walk. Also the scheme has l-Diversity as there would be l-Different values of random paths followed by the forwarded packets. *Two-way Random Walk* Two ways random walk is used to further diversify the random walk mechanism hence adding up to l-Diversity of random walk as well as K-anonymity is increased with increased set of values of the quasi identifier to be in the equivalence class for describing K-anonymity. *Routing with multiple parents* When routing with multiple parents each node would be having more than one parent providing an ability to implement K-anonymity at each node may be having (K−1) set of packets having similar parent node id so parent node id could be considered as quasi identifier for the implementation of K-anonymity. Also each node will be having as many number of parent nodes as well represented values of parent node id thus helpful for l-Diversity privacy principle notion implementation. *Routing with random walk* For routing with random walk again we can have l-Diversity and K-anonymity. K-anonymity would be supported by (K−1) values of the random routes taken by many sensor nodes. Also the scheme would be having l-Diversity as there would be l-different well represented nodes available for taking as the next node to propagate data. *Deco-relating parent child relationship* The privacy preservation through deco relation of parent child relationship takes into consideration measures to ensure that there are a set of values instead of sole values for both the parent as well as child attributes at each node level. Thus providing K-anonymity as there would be (K−1) similar values of parent and child attributes of each sensor node thus making the attributes eligible to be used as quasi identifier. Also the scheme would be having l-Diversity as there would be l well represented values of the aforementioned attributes for generation of l-Diverse equivalence class. *Hiding traffic patterns* Privacy preservation through hiding traffic patterns of sensor network traffic utilizes some algorithm that generalizes the flow of traffic in the network thus making the patterns of traffic non-recognizable to the adversary. These generalization algorithms implemented could be used as the basis for implementation of t-Closeness privacy preservation notion in the network. *Propagating*

Table 7 Temporal privacy provisioning in WSN

Considered set of beliefs	
$\mathbf{B}_1$	Background knowledge possessed by the adversary along with the localization may help revelation of network topology to the adversary
$\mathbf{B}_2$	Network deployment along with the routing information know to the adversary may help the adversary to deduce Spatio-temporal information related to sensed data
$\mathbf{B}_3$	Spatio-temporal information and network operations help the adversary to evolve the application criteria of the sensor network
Considered techniques for temporal privacy preservation in WSN	
$\mathbf{T}_1$	Adding delay to the sensed data
$\mathbf{T}_2$	Masking spatio-temporal information related to data
$\mathbf{T}_3$	RCAD (Rate Controlled Added Delay)

dummy data Propagation of dummy data in the network is generally done in order to perceive the uniformity in the network operations of data packet propagation. This scheme though is quiet energy consuming in many a times very useful for both the content as well as context privacy in sensor network. As the dummy data generated is somewhat similar to the original data required to be propagated hence provides a scope of implementation of K-anonymity privacy preservation notion to such privacy preservation mechanism.

7.5 Temporal Privacy Preservation in WSN

The temporal privacy provisioning in WSNs is done in order to hide the spatio-temporal evolution patterns of traffic generated within the network. Table 7 gives an insight of the considered set of beliefs an adversary may have while trying to deduce the spatio-temporal patterns of traffic generated and propagated. The table also enlists various temporal privacy prevention techniques proposed and implemented followed by an analytical review of these techniques on the basis of privacy notions.

To preserve the temporal generation patterns, *adding delay to the sensed data* has been done: Adding delay to the data sensed adds up to the temporal privacy preservation of the data. This mechanism not only helps in the prevention of revealing the exact time of occurrence of data but also helps to mitigate the time correlation attacks by the adversary. The amount of delay added though must be lesser enough not to mar the relevance of data being propagated. Though the mechanism of privacy preservation enhances the K-anonymity in sensor network as a whole as of now many data packets would be having similar delay introduced and similar buffering. Also the mechanism inbuilt t-Closeness in the network as of now after adding delay and buffering data at specific sensor nodes the overall distance between the whole distribution of data and to any equivalence class generated would be comparable. *Masking spatio-temporal information related to data* Masking the spatio temporal information related to data generated is another way of handling temporal privacy in sensor networks. While the masking adds aforementioned details would be on the basis of some cryptographic techniques the mechanism adds up to the K-anonymity of privacy preservation. *RCAD (Rate Controlled Added Delay)* The rate controlled added delay proposed to introduce delay in the data packets generated but the buffering is done on the preemption basis of data packets. Though this mechanisms provides a very good performance on preventing buffering stress at the corresponding sensor node, it also has an additional advantage of providing K-anonymity, l-Diversity as well as t-Closeness privacy notions.

Table 8 Analytical depiction of privacy provisioning in WSN

Beliefs techniques	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	K-Anonymity	*l*-Diversity	t-Closeness
(a) Privacy preservation through data aggregation in WSN											
T_1									*		*
T_2									*	*	
T_3									*		*
T_4										*	*
(b) Privacy preservation through anti-traffic analysis schemes in WSN											
T_1			*	*					*	*	
T_2	*	*	*	*					*	*	*
T_3			*	*						*	
T_4	*	*	*	*					*		
T_5	*	*	*	*							*
T_6	*	*	*	*					*	*	
(c) Privacy preserving schemes for preserving location privacy of source and receiver node											
T_1			*	*	*	*	*		*		
T_2			*	*	*	*	*		*		
T_3						*			*	*	
T_4	*	*	*	*	*				*	*	
T_5	*	*	*						*		
T_6			*	*							*
(d) Techniques for privacy preservation of location privacy of base station											
T_1		*								*	*
T_2			*		*				*	*	
T_3			*		*				*	*	
T_4			*		*				*		*
T_5	*	*	*	*							*
T_6									*		
(e) Temporal privacy preservation in WSN											
T_1		*	*						*		*
T_2	*	*	*						*		
T_3		*	*						*	*	*

8 Conclusion

This paper presents an analytical review of the current privacy preservation mechanisms designed and implemented in WSN based on the considered set of beliefs mitigated by each of them and three privacy notions of K-anonymity, *l*-Diversity and t-Closeness. Our analytical study presented many interesting facts about the current scenario of privacy provisioning in WSN. The study was conducted on five different sets of privacy provisioning sets namely privacy preservation through data aggregation, privacy preservation of location of source node, destination node, base station and temporal privacy preservation. The considered set of beliefs to be mitigated for the aforementioned set of privacy provisioning algorithms were formulated after the keen observation of preconvention any adversary would be having

about the network and its entities. These beliefs need to mitigate by the privacy provisioning methods in order to achieve an efficient privacy mechanism that could be fool the adversary about the exact information related to network and its entities. Afterwards depending on the above observations the mechanisms were mapped to three privacy notions that could be implemented on the considered mechanisms. Our observations though did not presented any steep orientation of the considered mechanism towards any specific privacy notion but still almost all of the privacy provisioning mechanism could be mapped easily with the K-anonymity notion for privacy preservation as almost all the mechanism in some or the other way created a set of K similar entities that may be similar set of fake packets or similar set of aggregated packets or similar set of some entities in the network to make the real packet or entity concealed among these similar entities. While having l-Diversity requires the need to have l-different well represented values that are achieved by some of the algorithms and t-Closeness also required a probabilistic distribution among the entities that is achieved by few of the privacy provision mechanism. A depiction of the privacy notions and set of beliefs mitigated by the various privacy provisioning techniques have been given in the Table 8. The key points of our observations out of the present study could be summarize as follows:

- The data aggregation based privacy preservation techniques are able to preserve the content privacy in WSN but are not able to mitigate any of the content privacy oriented believes considered in our study.
- Almost all of the privacy provisioning techniques implemented followed K-anonymity.
- Lesser number of the techniques adheres to t-Closeness as there must be a probabilistic distribution pattern in the generated procedures of the technique to adhere to t-Closeness.
- The number of considered set of believes mitigated by the privacy provisioning techniques is maximum for the anti-traffic analysis schemes followed by the location privacy provisioning schemes.
- l-Diversity notion of privacy provisioning is also not adhered by lesser number of the techniques the reason being the requirement of l-different well representations to be implemented by the technique.
- As a whole what we may conclude is that we need to improvise or privacy provision techniques to adhere to latest and more efficient forms of privacy notions.
- Also the techniques must have the ability to mitigate the newer set of beliefs that the adversary may deduce based on the earlier set of beliefs mitigated by privacy provisioning techniques.

References

1. Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications in Magazine*, *40*(8), 102–114.
2. Li, N., Zhang, N., Das, S. K., & Thuraisingham, B. (2009). Privacy preservation in wireless sensor networks: A state-of-the-art survey. *Ad Hoc Networks*, *7*(2009), 1501–1514.
3. Deng, J., Han, R., & Mishra, S. (2005). *Security, privacy, and fault tolerance in wireless sensor networks*. Boston: Artech House.
4. Karlof, C., & Wagner, D. (2003). Secure routing in wireless sensor networks: Attacks and countermeasures. *Elsevier's AdHoc Networks Journal, Special Issue on Sensor Network Applications and Protocols*, *1*(2–3), 293–315.
5. Xu, W., Trappe, W., Zhang, Y., & Wood, T. (2005). The feasibility of launching and detecting jamming attacks in wireless networks. In *ACM MobiHoc*.
6. Xu, W., Wood, T., Trappe, W., & Zhang, Y. (2004). Channel surfing and spatial retreats: Defenses against wireless denial of service. In *ACM WiSe* (pp. 80–89).

7. Gaubatz, G., Kaps, J. P., & Sunar, B. (2004). Public key cryptography in sensor networks-revisited. In *1st European workshop on security in ad-hoc and sensor networks (ESAS 2004)*.
8. Hwang, J., & Kim, Y. (2004). Revisiting random key pre-distribution schemes for wireless sensor networks. In *Proceedings of the 2nd ACM workshop on security of ad hoc and sensor networks (SASN '04)* (pp. 43–52). New York, NY: ACM Press.
9. Zhang, R., Zhang, Y. & Ren, K. (2009). Distributed privacy preserving access control in sensor networks. In *INFOCOM 2009, IEEE* (pp. 1251–1259).
10. Perrig, A., et al. (2002). SPINS: Security protocols for sensor networks. *Wireless Networks*, *8*(5), 521–534.
11. Aysal, T. C., & Barner, K. E. (2008). Sensor data cryptography in wireless sensor networks. *IEEE Transactions on Information Forensics and Security*, *3*(2), 273–289.
12. Ouyang, Y., Le, Z., Xu, Y., Triandopoulos, N., Zhang, S., Ford, J., et al. (2007). Providing anonymity in wireless sensor networks. In *IEEE international conference on pervasive services* (pp. 145–148).
13. Xi, Y., Schwiebert, L., & Shi, W. (2006). Weisong preserving source location privacy in monitoring-based wireless sensor networks. In *IEEE international parallel and distributed processing symposium*. Los Alamitos, CA: IEEE Computer Society.
14. Deng, J., Han, R., & Mishra, S. (2004). Intrusion toleranceand anti-traffic analysis strategies for wireless sensor networks. In *Proceedings of the 2004 international conference on dependable systems and networks (DSN '04)* (pp. 637–646). Washington, DC: IEEE Computer Society.
15. Ozturk, C., Zhang, Y., & Trappe, W. (2004). Source-location privacy in energy-constrained sensor network routing. In *Proceedings of the 2nd ACM workshop on security of ad hoc and sensor networks (SASN '04)* (pp. 88–93). New York, NY: ACM.
16. Deng, J., Han, R., & Mishra, S. (2005). Countermeasures against traffic analysis attacks in wireless sensor networks. In *Proceedings of the first international conference on security and privacy for emerging areas in communications networks (SECURECOMM '05)* (pp. 113–126). Washington, DC: IEEE Computer Society.
17. Xi, Y., Schwiebert, L., & Shi, W. (2006). Preserving source location privacy in monitoring-based wireless sensor networks. In *IEEE international parallel and distributed processing symposium*. Los Alamitos, CA: IEEE Computer Society.
18. Syverson, P. F., Goldschlag, D. M., & Reed, M. G. (1997). Anonymous connections and onion routing. In *Proceedings of the IEEE symposium on security and privacy* (pp. 44–54).
19. Xi, Y., Schwiebert, L., & Shi, W. S. (2006). Preserving source location privacy in monitoring-based wireless sensor networks. In *Proceedings of the 20th international parallel and distributed processing symposium (IPDPS 2006)*.
20. Hong, X., Wang, P., Kong, J., Zheng, Q., & Liu, J. (2005). Elective probabilistic approach protecting sensor traffic. In *Military communications conference (MILCOM 2005)* (Vol. 1, pp. 169–175). IEEE.
21. Przydatek, B., Song, D., & Perrig, A. (2003). *SIA: Secure information aggregation in sensor networks*. ACM SENSYS, 255–265.
22. He, W., Liu, X., Nguyen, H., Nahrstedt, K., & Abdelzaher, T. T. (2007). PDA: Privacy-preserving data aggregation in wireless sensor networks. In *26th IEEE international conference on computer communications (INFOCOM 2007)* (pp. 2045–2053). IEEE.
23. Kamat, P., Xu, W., Trappe, W., & Zhang, Y. (2007). Temporal privacy in wireless sensor networks. In *Proceedings of the 27th international conference on distributed computing systems (ICDCS '07)* (pp. 23–30). Washington, DC: IEEE Computer Society.
24. Li, Y., & Ren, J. (2009). Providing source-location privacy in wireless sensor networks. In *Proceedings of the 4th international conference on wireless algorithms, systems, and applications (WASA '09)* (pp. 338–347). Berlin, Heidelberg: Springer.
25. Ngai Edith, C.-H. (2009). On providing sink anonymity for sensor networks. In *Proceedings of the 2009 international conference on wireless communications and mobile computing (IWCMC '09)* (pp. 269–273). New York, NY: ACM.
26. Shaikh, R. A., Jameel, H., d'Auriol, B. J., Lee, S., Song, Y.-J., & Lee, H. (2008). Network Level privacy for wireless sensor networks. In *Fourth international conference on information assurance and, security* (pp. 261–266).
27. Zhang, W., Wang, C., & Feng, T. (2008). GP2S: Generic privacy-preservation solutions for approximate aggregation of sensor data (concise contribution). In *Proceedings of the 2008 sixth annual IEEE international conference on pervasive computing and communications (PERCOM '08)* (pp. 179–184). Washington, DC: IEEE Computer Society.
28. Levis, P., Lee, N., Welsh, M., & Cullar, D. (2003). TOSSIM: Accurate and scalable simulation of entire TinyOS applications. In *Proceedings of the first ACM conference on embedded networked sensor systems (SenSys 2003)*.
29. Hill, J., et al. (2000). System architecture directions for networked sensors. In *ASPLOS, 93*.

30. Alkhatib, A. A. A., & Baicher, G. S. (2012). Wireless sensor network architecture. In *International conference on computer networks and communication systems (CNCS 2012) IPCSIT* (Vol. 35). © Singapore: IACSIT Press.
31. Yuksel, E. (2011). *Qualitative and quantitative security analyses for ZigBee wireless sensor networks*. Ph.D. thesis. Technical University of Denmark.
32. Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). ISBN: 0-471-24195-4.
33. Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems*, *10*(5), 557–570.
34. Machanavajjhala, A., Kifer, D., Gehrke, J., & Venkitasubramaniam, M. (2007). *L*-diversity: Privacy beyond *k*-anonymity. *TKDD, 1*(1).
35. Li, N., & Li, T. (2007). t-closeness: Privacy beyond k-anonymity and l-diversity. In *ICDE*.

Author Biographies

Manjusha Pandey is a Ph.D. student in the Department of Information Technology at Indian Institute of Information Technology, Allahabad, India. Her research interest are Wireless Sensor Networks, Privacy in Wireless Communication, Privacy and security in Digital & Mobile Communication, Signal Processing and Vehicular Technology.

Shekhar Verma received his Ph.D. degree from IIT, Banaras Hindu University, Varanasi, India in Computer Science and Engg. He is Associate Professor in Information Technology at Indian Institute of Information Technology, Allahabad, India. His research interest areas are Computer Networks, Wireless Sensor Networks, Vehicular Technology, Cryptography, Information and Network Security.