Asian Journal of Information Technology 15 (14): 2447-2456, 2016
ISSN: 1682-3915

Privacy-Preserving Data Transmission Using Geometric Data Perturbation in Wireless Sensor Networks

[1]K. Sreekumar and [2]E. Baburaj
[1]Anna University, Chennai, India
[2]Department of CSE, Narayanaguru College of Engineering , Manjalumoodu, India

Abstract: Data privacy is an important aspect to be dealt with during data transmission, data query and data storage in Wireless Sensor Networks (WSNs). Issue of privacy of the data collected and transmitted from the sensor nodes in wireless sensor networks is also of major concern. The salient features like uncontrollable environment, sensor node resource constraints and topological constraints have to be taken into consideration to have a better tradeoff among privacy, accuracy and power consumption. In this study, a new scheme, Hierarchical based Multidimensional Data Perturbation (HMDP) is proposed and Geometric Data Perturbation (GDP) is suggested for perturbing the data collected by the sensor nodes during data transmission. GDP perturbs the data randomly while broadcasting the data collected from the sensor nodes to the central processing server or another sensor node. A comparative study of the proposed method with the existing methods is presented. The proposed technique gives a better tradeoff in terms of the metrics accuracy, privacy and power consumption.

Key words: Data transmission, data query, geometric data perturbation, privacy, wireless sensor network

INTRODUCTION

Wireless Sensor Networks (WSNs) are developed and designed for a range of applications, for instance environment monitoring, habitat monitoring, health monitoring, battle-field, home automation, traffic control, etc. A wireless sensor network comprises sensor nodes widely spread to observe environmental or physical circumstances in a co-operative manner. These sensor nodes have limited battery power.

The battery power is expected to last over several years or months. So, one of the concerns is to improve the lifetime of the network which will improve the WSNs energy efficiency. It is required that the algorithm running on sensor devices has a low computational cost due to the limited memory space and processing speed of sensor nodes. In sensor network critical information is transmitted by sensor nodes. So, another challenging issue that has to be addressed in Wireless Sensor network is the privacy preservation of the transmitted data between sensors.

As per the taxonomy mentioned by Li *et al.* (2009), the privacy preservation in WSNs is broadly classified into data privacy and context privacy. Figure 1 shows the taxonomy for privacy in WSNs.

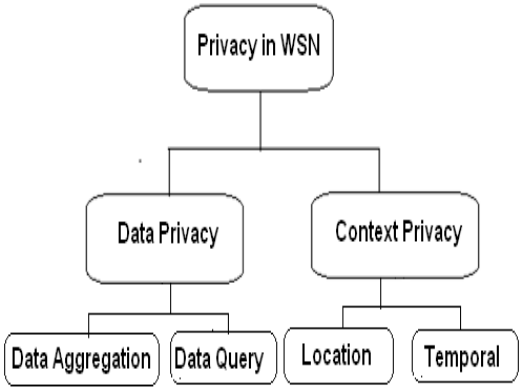

Fig. 1: Taxonomy of WSN privacy

Two major types of concerns of privacy are context oriented and data-oriented (Li *et al.*, 2009). Data-oriented issues mainly focus on the collected data privacy or posted query to WSNs. Context oriented issues mainly focus on contextual data like traffic flow timings and location in WSN. Context and data oriented privacy issues may be desecrated by traffic analysis and data attacks in that order. Figure 2 shows the above two types of attacks.

Corresponding Author: K. Sreekumar, Research Scholar, Anna University, Chennai, India

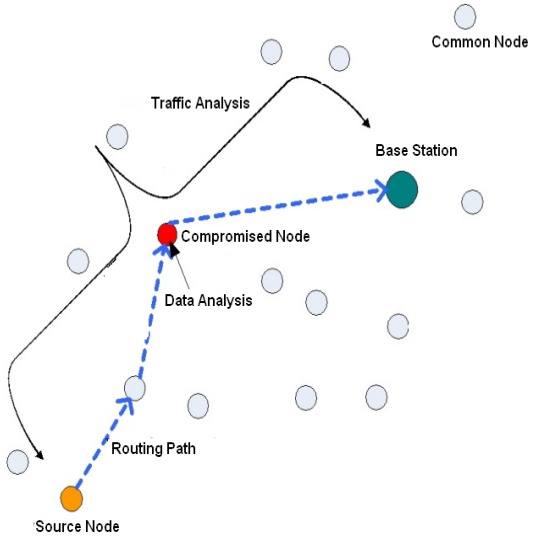

Fig. 2: Two privacy attack scenarios in WSN

In the attack of data analysis, a WSN nasty node misuses its data decryption ability to compromise the transmitted payload. Data content privacy protection is the main focus of Data-oriented privacy protection. In the case of nuclear power plant WSN, private information may include Bearing temperature, Winding temperature, Motor current, etc. There exist two kinds of adversaries which may negotiate data-oriented privacy. The external adversary that eavesdrops the data communication among sensor nodes in a WSN is one kind. This kind of adversaries can be successfully shielded against using the conventional methods of cryptographic encryption and authentication. The internal adversary which is a participating node of the WSN but captured and manipulated by unknown entities to negotiate private information is the second type. Since, it is permitted for a participating node to legally decipher data, the conventional encryption and authentication methods are not successful. So, the challenge for data-oriented privacy protection is not allow an in-house opponent from capturing the personal information, while the regular WSN process is being maintained.

To overcome the above data analysis issue, this paper presents a new privacy-preserving scheme, Hierarchical-based Multidimensional Data Perturbation (HMDP), using geometric data perturbation approach. The objective of this research is to secure the privacy of the data transmitted between sensors. The proposed methodology enhances the data transmission that provides security to the data. It is suggested that the HMDP when applied to the sensor nodes will perturb the data that has to be transmitted and prevent eavesdropping from external and internal adversaries.

This study has the following organization. Presentation of related work about key distribution strategies and data privacy preservation strategies for data transmission in WSNs is carried out in the 2nd section. Presentation of the novel privacy-preserving data aggregation scheme, HMDP, in WSNs is carried out in the 3rd section. Presentation about the evaluation of performance of the new strategy, HMDP, is carried out in the 4th section. Suggestion for the future work and the conclusions are finally carried out in the 5th section.

Related work: In wireless networks containing sensors one active research area is the problem of key management. Because of the high computation and storage cost, conventional methods in ad hoc networks with asymmetric keys are very costly. To bootstrap the primary trust among the sensor nodes a probabilistic key predistribution scheme was proposed by Eschenauer and Gilgor (Eschenauer and Gligor, 2002). The idea was before deployment to permit every sensor node select randomly a set of keys from a key pool so that the likelihood of sharing no less than one common key among any two sensor nodes is certain. This idea was extended further by Chan *et al.* (2003) which resulted in the development of two key predistribution strategies: the random pair-wise keys strategy and the q-composite key predistribution strategy. A key pool is also used by the q-composite predistribution strategy of keys. But, it requires two sensor nodes to calculate a pair-wise key from no less than q predistributed shared keys. In the random pair-wise keys strategy duos of sensor nodes are picked randomly and a sole random key is assigned to each pair. The security above the plan of basic probabilistic predistribution of key is improved by both the strategies. But one issue is the authentication problem and establishment of pair-wise key in wireless sensor networks involving mobile sinks.

The existing privacy-preserving techniques used the common idea of perturbing the transmitted WSN data. He *et al.* (2007) proposed two techniques for privacy preservation. One is Cluster-based Private Data Aggregation (CPDA). Here with actual data random seeds are added. Other one is Slice-Mixed Agg Rega Tion (SMART). This method first cuts data into pieces. After exchanging randomly those pieces it rebuilds data package. To hide individual raw data the above two approaches rely on the neighbors cooperation. They also concentrate on the operation of SUM aggregation. Another technique for data-transmission by preserving privacy is to generalize the data transmitted for various aggregation functions support like Median, Max/Min, etc.

Adding noise to the raw data sensed from a WSN is the basic idea of CPDA. In this way the aggregator get precise cumulative data other than individual points of data. Analogous data perturbation approaches are widely used in privacy-preserving data mining. Data mining with privacy preservation heads to inaccurate combined results since noises are independently generated at random. In CPDA the noises are designed carefully to control the cooperation between various sensor nodes. This helps the aggregator to get the precise aggregated values. CPDA categorizes sensor nodes as cluster members and cluster heads. Between them they have one-to-many mapping. The responsibility of the cluster head is in direct data aggregation from cluster members. For secured communication each pair of communication nodes use a different shared key. The noise can be used as random data in CPDA. This noise hides the raw items of data from cluster head. The noise negative effect is eliminated by the cluster nodes cooperation which helps to calculate the original sum value.

Another method used for individual protection of data in the aggregation of SUM is SMART. The technique used is initially the original data is sliced into pieces and later randomly recombine them. This scheme involves three steps. In step one slicing, within h hopes j neighbor nodes are selected by each sensor nodes to build a set S. Then the data is sliced into J pieces. After keeping one piece for itself it transmits other (J-1) encrypted pieces to (J-1) sensors selected at random from set S. In step two mixing, each sensor upon receipt of data from other sensors decrypts the data via the shared key. Every sensor waits to make certain that the process of round aggregation data slicing and receipt is finished. In step three aggregation, data is aggregated by intermediate sensor and transmit to the base station.

Another method is Generic Privacy-preservation solutions for approximate aggregation (GP^2S) (Zhang *et al*., 2008). To generalize the values of transmitted data in a WSN (Zhang *et al*., 2008) is the basic idea. The particularized content of data cannot be decoded. Still the aggregator can able to get the correct estimate of the data distribution histogram. Thus aggregates are approximated. An integer range is used by each sensor node before transmission to replace the raw data. The histogram for data collected is plotted by the aggregator with certain granularity. Then the aggregator estimates aggregates such as Median and MIN/ MAX.

Like the protection of collected data in WSN, a query posted to the WSN (Carbunar *et al*., 2010) to get the data collected too is also very crucial issues of privacy. In medical WSN when the queries have been often posted to a WSN part which covers a house of a patient then an intruder can understand that the patient's health is receiving extra notice because of problems of health. In the WSN design which is resource-constrained private data query pose challenges significantly. Restrict query processing to small range to preserve energy. Reducing the query limit in addition strengthen the likelihood for the intruders to infer the query.

The protection of privacy benefit relies on the cost of other metrics. Data broadcasting techniques like SMART, GP^2S and CPDA gives good security for the collected data privacy from distinct sensors. But on other measures they have limitations. For SMART and CPDA, the cost primarily depends on the consumption of power because of the sliced data exchange. For GP^2S, since only approximate values, e.g. MIN, MEDIAN and SUM are acquired the cost arise from the aggregated data accuracy. So, the prime challenge in the design of a new data privacy-preserving transmission method is to find a good tradeoff among three metrics: accuracy, privacy and power consumption considering the salient features like uncontrollable environment, sensor-node resource constraints and topological constraints (Akyildiz *et al*., 2002; Niu *et al*., 2014). The proposed method, Hierarchical-based Multidimensional Data Perturbation (HMDP), achieves a better tradeoff between the three metrics, i.e. privacy, accuracy and power consumption.

MATERIALS AND METHODS

System model

Design consideration: In various WSN applications data privacy protection is a key concern. Privacy-preserving strategy requirements to sustain data privacy are presented under this section. The subsequent criterion sum up the necessary characteristics of the proposed data transmission scheme, HMDP.

Data privacy: One prime concern toward many resident applications, involving wireless networks containing sensors, is privacy. By eavesdropping on the network intruders try to collect more information. So, there is a significance to build up data transmission strategies against eavesdropping to guarantee privacy of the data.

Efficiency: For privacy protection additional overhead is introduced in private data aggregation schemes. But, a good private scheme of data aggregation should maintain low overhead. Minimizing the total messages transmitted inside the sensor network is the goal of data transmission scheme which will eventually reduce the usage of power and resource. In the proposed scheme to reduce the overhead, perturbation technique is mainly employed

instead of encryption. The perturbed data does not require encryption as the perturbed data itself is having very high privacy. The bandwidth efficiency is achieved by data aggregation by using in-network processing.

Accuracy: An accurate sensor data aggregation is expected without revealing the precise value of individual sensor to other sensors as the constraint. For private data aggregation scheme's performance estimation, accuracy should be a criterion.

Hierarchical-based Multidimensional Data Perturbation (HMDP) scheme for data transmission: We provide a new scheme for preserving privacy, HMDP, in wireless networks containing sensors by using Geometric Data Perturbation technique. The new scheme for preserving privacy is accomplished by using four phases; phase involving construction of the network, key distribution phase, data perturbation phase and finally, the phase involving transmission of the data. In the phase of construction of the network, every sensor finds their child nodes, sibling nodes and parent node by transferring messages through broadcasting. In the phase of key distribution, key discovery between sensor nodes is carried out for encryption of any perturbation parameter if needed. In the phase of data perturbation, each node perturbs the collected data using Geometric Data Perturbation. In the final phase of data transmission, the sensor node transmits the perturbed data along with the encrypted perturbation parameters like rotation, translation and noise addition to the destination. Only the perturbation parameters are encrypted. All the data owned by sensors are aggregated by the sink. Detailed steps are explained below.

Network construction phase: The HMDP strategy uses the topology which is based on tree. The communication range between cluster heads affects the clustering-based topology. Also in the clustering-based topology for network construction large amount of messages are required. Initially by transferring the HELLO message, one query is triggered by a sink node. After the receipt of the message, HELLO, the node checks the source of the received message. If within the communication range a sink node is located then the sensor node receives the HELLO message and puts the sink as a parent node. It also sets the current node as an access point. The HELLO message will be forwarded to the neighbors by a node if it becomes an access point. Otherwise within its communication range if an access point is located then the node accepts the HELLO message from the access point and puts the access point as the parent. Or else, the

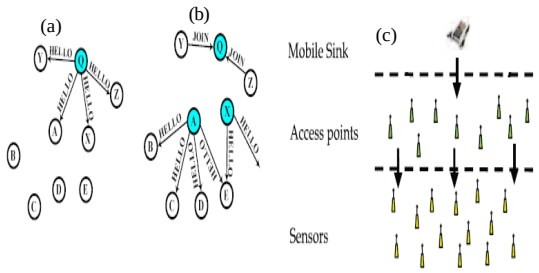

Fig. 3: Construction of the network

sensor node stays for some interval of time to accept the HELLO messages from the neighbors and chooses any of the nodes within its communication range with minimum hop count to the nearest access point as the parent node using the broadcast of a JOIN message and sets the access point same as the access point of the parent. This is illustrated in Fig. 3a and b. As this process goes on, a network is created as shown in Fig. 3c. To avoid network imbalance the maximum number of child nodes are specified. If there is imbalance in the network, more energy is consumed by all the sensor nodes in the imbalanced area than other areas.

This shows the algorithm for the construction of the network. Initially, a HELLO message is flooded by a sink node to a node which is nearest inside its transmission limit (steps 01 and 02). The node receiving the HELLO message from the sink puts its level and becomes access point and transmits the HELLO message to other nodes (steps 03-13). When a JOIN message is received by a node, then the node sending the message JOIN is set as the child. Then it sets the access point of the child same as the access point of the parent. Parent node having maximum child nodes send out the RESET message to child nodes indicating them to choose a new parent node (steps 14-19).

Pseudo code for the construction of the network

Network construction algorithm 1:

```
Command construction of network (message msg,
msgtype msgtype){
    Step 01: If (Node Q is a Sink) {
    Step 02: Flood(InitialLevel, BaseStationID);  Exit;}
    Step 03: Wait until the receipt of the msg HELLO ;
    Step 04: When (a node receiving msg from
                                        another node) {
    Step 05: If (msgType = HELLO) {
    Step 06: Locate ParentID, RectLevel,
                                RectHopCont from msg ;
    Step 07: If (NodeParent is Sink node) {
    Step 08: Set currentNode as AccessPoint; }
    Step 09: Network.CurrEntry++ ;
    Step 10: If(RectHopCont + 1 < CurHopCont) then
                        CurHopCont = RectHopCont + 1;
```

Step 11: else Break;
Step 12: If (TOS_LOC_ADD <>LeafNode)
Step 13: Flood(CurrentLevel, CurrentNodeID) ; }
Step 14: If(msgType = JOIN){
Step 15: If(Child nodes of parent < Max allowed Child nodes)
Step 16: Network.Parent = ParentID ;
Step 17: If NodeParent is not an access point {
Step 18: Set child node access point = parent node access point; }
Step 19: else transmit (RESET) msg to node}
Step 20: }
End algorithm

Key distribution phase: After the construction of sensor network, every sensor node establishes pair-wise keys. To build our approach we use the Blundo scheme. The Blundo scheme gives better security guarantee. In the proposed concept, to improve authentication and setup key between stationary access and sensor nodes a static polynomial pool is used. The authentication points of access for the network are the stationary access nodes. The transmission of the aggregated data to the sink from the sensor nodes is initiated by the stationary access node upon receiving the request for data message from sink. From the pool of static polynomial a subset of polynomials are selected randomly by all sensors and the stationary access nodes. Our plan is separated into two stages: predistribution of static and discovery of key among a sensor node and a stationary access node. Likewise, a mobile polynomial is shared by every stationary access node with the mobile sink.

Data perturbation phase: In the phase of data perturbation, every node perturbs the sensed data using the proposed method, i.e. Geometric Data Perturbation. GDP is a mixture of noise addition, random translation perturbation and random rotation perturbation. The representation of GDP is as follows:

$$G(X) = RX + \Psi + \Delta \quad (1)$$

Where Ψ is a random translation matrix, R is a random orthogonal matrix and X indicates the original normalized dataset with N rows and d columns. Ψ = t x 1^1 and t are generated randomly using the uniform distribution over [-1, 1]. Δ is the noise matrix with independent identically distributed (i.i.d.) elements which is used in distance perturbation. To symbolize the privacy assurance of DPi, denoted by ñi use the minimum privacy guarantee by default. ρ is bounded by some value, say ci which may be different for different data sets and it is equal to or greater than zero. Definition of a random translation matrix is as given below (Zhang *et al*., 2005).

Definition: If $\Psi = t \times 11$, $t = [t1, t2, \ldots, td]^1$ ($0 <= ti <= 1$, $1 <= I <= d$) and $1 = [1, 1, \ldots, 1]^1$ then a matrix Ψ can be called as a translation matrix. Based on the uniform distribution over [-1, 1] t is generated randomly. Δ is the noise matrix with independent identically distributed (i.i.d) with zero mean and small variance elements which is employed in distance perturbation such that for certain kind of attacks the perturbation is resilient.

It is easy to verify that $\Psi_1 \pm \Psi_2$ and $R\Psi_i$ are also translation matrices if R is an orthogonal transformation and Ψ_1 and Ψ_2 are translation matrices. Both noise component of G(X) and t can be generated independently even though it is difficult to locate a suitable R in terms of the resistance to attacks. In the initial study Δ is with some common setting such as Gaussian $N(0, \sigma^2)$ and σ (Chen *et al*., 2005).

One of the challenges for applying GDP is to unify securely the different perturbed data of sensor nodes while privacy of each node is preserved and the pooled data utility is also preserved well. The quality of unified perturbation is affected by three factors and they are the pooled data utility, each data sets privacy guarantee and perturbation unification protocol efficiency. This shows the geometric data perturbation algorithm.

Pseudo code for data perturbation

Data perturbation algorithm 2:

Algorithm (Perturbing the nuclear power plant data at Sensor Node)
Input: X the original data set, priority of sensitive data, no of iterations
Output: R_t, Selected rotation matrix, the random translation, the noise level, the no of iterations
Step 01: Calculate the covariance matrix
Step 02: Generate randomly the translation Matrix
Step 03: For every iteration perform the following
Randomly generate rotation matrix
Swap the rows of selected rotation matrix
If the privacy guarantee of swapped rotation matrix
Then generate $\hat{X}$ else Rt = R'
Step 04: Perform perturbation G(X) = RX+Ψ Δ
End algorithm

The parameters rotation and translation are preserving the distance while Δ is slightly perturbing the distance as shown in the example in Fig. 4. The advantages of using GDP in WSN are as follows:

- GDP gives better accuracy than any other perturbation method by keeping the distances well
- Works well with multidimensional data too
- Provides better privacy guarantee
- GDP does not involve slicing or dicing of data but only involves rotation, translation and noise addition thus better accuracy can be achieved

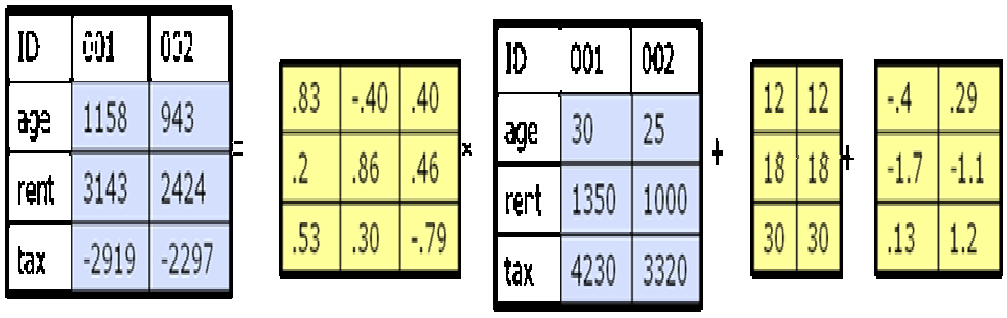

Fig. 4: Example calculation of GDP

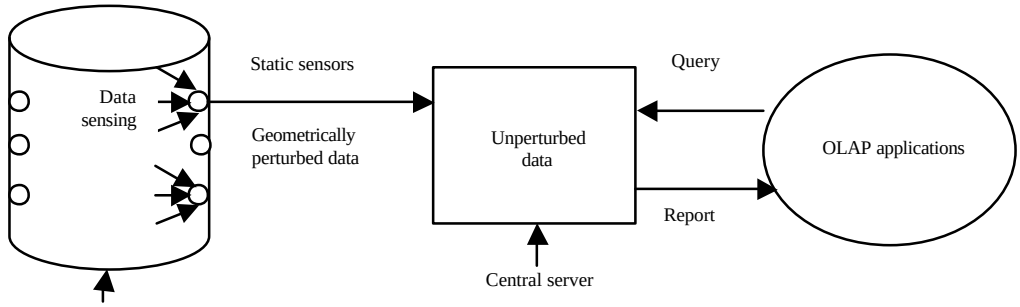

Fig. 5: System architecture for HMDP

Data transmission phase: In the phase of transmission of data, every node transmits the perturbed data along with the encrypted parameters to the parent node. Then, the received data is analyzed by the parent. All the perturbed data from the nodes hierarchy are aggregated by a sink node. For data transmission Time Division Multiple Access (TDMA) process (Madden *et al*., 2005) is used which will also be used to avoid the wireless sensor network's communication loss. Every child node transmits the perturbed data during its specific time of transmission. Figure 5 shows the algorithm for data aggregation. Aggregation of data starts from leaf node (step 01 and 02). For the purpose of aggregation, child node data can be received by an intermediate node and re-perturb the data with its private data (steps 03-13). By this way, every sensor node's perturbed data along with the encrypted parameters arrive at a sink which further transmits the aggregated data to the client (steps 14-18).

Pseudo code for data aggregation

Data aggregation algorithm 3:

Command Data Aggregation (Message msg, msg Type msg Type)
Step 01: If (node = Leaf)
Step 02: Transmit Msg (enc Par Data, per Data) to Node Parent
Step 03: Else {
Step 04: If (node receive msg(enc Par Data, per Data) from sensornode {
Step 05: If (node = InternalNode) {
Step 06: Store encParData,perData from msg
Step 07: DecryptedParData = decrypt(encParData, PairwiseKey)
Step 08: Unperturbed Data = unperturb (per Data, decrypted Par Data)
Step 09: Aggregatedper Data += unperturbed Data
Step 10: If(every data from childNode is received)
Step 11: New Per Data = GDP (aggregatedpe rData, Parameter (R,T,N))
Step 12: New EncPar Data = Hash Function (Parameter (R,T,N), PairwiseKey)
Step 13: SendMessage(encParData,perData) to NodeParent}
Step 14: If (node = SinkNode){
Step 15: Stores encPar Data, per Data from msg
Step 16: Decrypted Par Data = decrypt (encPar Data, PairwiseKey)
Step 17: Uunperturbed Data = unperturb (per Data, decrypted Par Data)
Step 18: Send Message (decrypted Par Data, unperturbed Data) to User}}}
End Algorithm

During data transmission phase for encryption, a collision-resistant hash function such as MD5 (Rivest, 1992) is used. To protect this scheme of security against replication attack on stationary access node, one-way hash chains algorithm (Lamport, 1981) in combination with static polynomial pool-based strategy (Liu and Ning, 2003) is used.

System architecture: This shows the system architecture of our proposed system, HMDP. The data collected by the sensor nodes are perturbed using GDP and are aggregated in a sink which further transmits the aggregated data to the central server. The data is unperturbed before being stored in the server. The stored data in the server is in turn used by various applications.

RESULTS AND DISCUSSION

Evaluation: The privacy-preserving scheme of data aggregation presented in this paper is evaluated in this section. We evaluate the performance of our method in terms of efficiency, aggregation accuracy and privacy preservation.

Environment: We use NS2 simulator. In the implementation of our simulation, 50 sensor nodes are placed in a 200×200 m^2 area. We describe the node density as the number of deployed nodes in the field. The data packet generated from the source node is being forwarded over multiple hops toward the sink.

Evaluation metrics: The main evaluation metrics to evaluate the usefulness of the proposed scheme are accuracy, power consumption and privacy:

- Privacy means the privacy protection degree given via the observed methods
- The degree of accuracy comprises two aspects: obtained data precision of the sink and the ease of use of the (expected) data to sink (i.e., the data to the sink is deliverable or not)
- The measure of power consumption focuses on the extra spends energy in WSN for the transmission of additional messages

Measuring privacy: Privacy is commonly measured by the difficulty level in estimating the original data from the perturbed data. Given the data perturbation technique, the more difficult the original values can be estimated from the perturbed data, the higher level of data privacy this technique provides:

$$P = A + V \tag{2}$$

Where, A is the actual data, V is the random variable and P is the perturbed data, i.e., the difficulty level to convert P to A:

$$\text{Privacy} = \text{stddev}(P - V) \tag{3}$$

Measuring accuracy: The accuracy is intimately related to information loss resulting from the hiding strategy. Less information loss increases the data quality. The dissimilarity between the original data D and its frequency fD(i), fD'(i) is the frequency of the perturbed data D' is measured as follows:

$$\text{Diss}(D, D') = \frac{\sum_{i=1}^{n} |fD(i) - fD'(i)}{\sum_{i=1}^{n} fD(i)} \tag{4}$$

The loss of information is described as the proportion among the sum of absolute faults made in frequency computation of their items from the perturbed data and sum of all the item frequencies in the initial data set.

Overhead: The power consumption is calculated in terms of the execution time taken by the algorithm. The reduction in execution time significantly reduces the power consumption in WSNs. The measure of consumption of power focuses on the extra required messages for storage, processing and broadcasting.

Result: Packet Delivery Ratio (PDR) is the proportion of the total number of packets successfully delivered to the destination and the number of packets generated by the source. If the amount of malicious node increases, PDR decreases. Also the higher mobility of nodes causes PDR to decrease. Higher PDR ensures secure data transmission with lesser or no packet loss. PDR is high in our proposed concept using GDP method as shown in Fig. 6:

$$\text{PDR}(\%) = \frac{\text{No. of packets successfully delivered to the destination}}{\text{No. of packets generated by the source node}} \tag{5}$$

Figure 7 shows the privacy of the proposed scheme, HMDP, against the existing scheme CPDA. It is shown that as the epoch increases, the privacy also increases. With increase in the epoch, the proposed scheme gives a privacy which is almost same like the existing scheme CPDA. Figure 8 shows the accuracy of the proposed scheme, HMDP, against the existing schemes CPDA and SMART. It is shown that as the epoch increases, the accuracy also increases. With increase in the epoch, the proposed scheme gives a higher accuracy. For example for the epoch, say 50, the accuracy of the proposed scheme is 0.9122. It is clear that the proposed scheme outperforms the existing methods.

The average increase in the accuracy of the proposed scheme against the existing scheme, SMART is 0.24823. The average increase in the accuracy of the proposed scheme against the existing scheme, CPDA is 0.22305.

Figure 9 shows the consumed energy in Jules of the proposed scheme, HMDP, against the existing schemes CPDA and SMART. As the time increases, the proposed scheme consumes less energy than the existing schemes. For example, for the epoch, say 50, the consumption of energy of the proposed scheme is 2.243.

The proposed scheme outperforms the existing methods. The average decrease in the consumption of

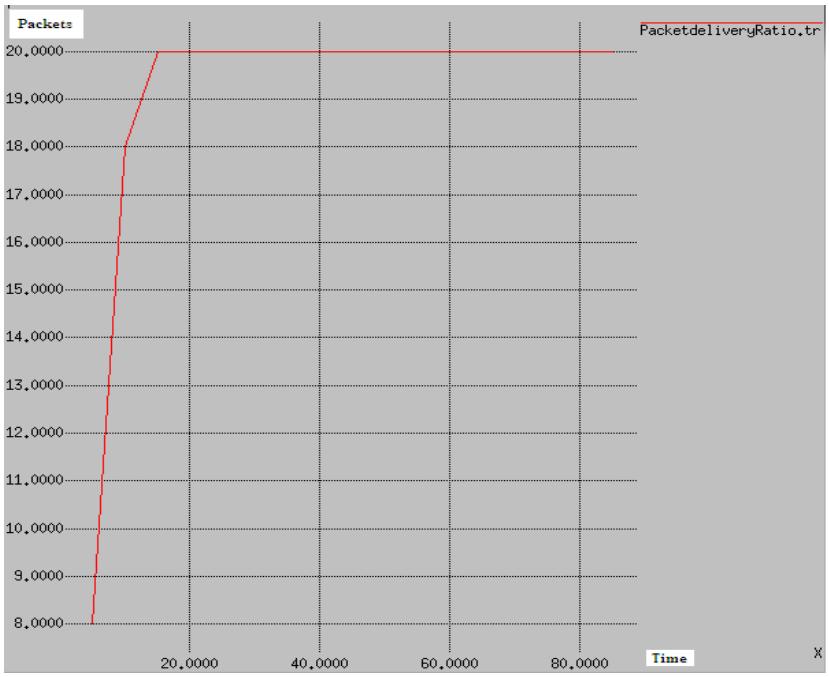

Fig. 6: Packet delivery ratio with varying seconds using GDP

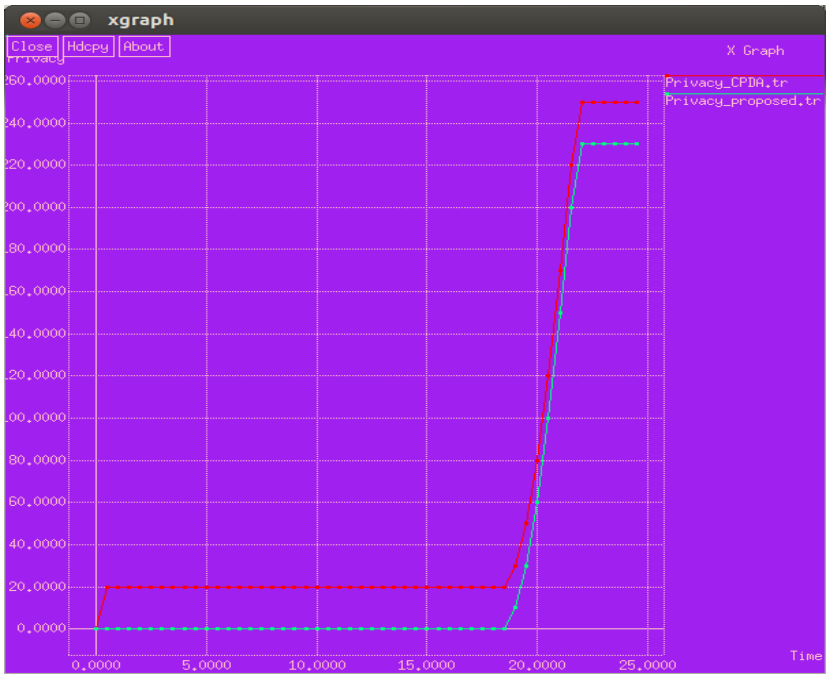

Fig. 7: Privacy with varying seconds

energy of the proposed scheme against the existing scheme, SMART is 1.02187. The average decrease in the consumption of energy of the proposed scheme against the existing scheme, CPDA is 0.70606. Figure 10 shows the communication overhead of the proposed scheme, HMDP, against the existing schemes CPDA and SMART. It is clear that the proposed scheme has less communication overhead than the existing schemes. This is because in the proposed scheme additional messages are not required like in the existing methods. Thus the proposed scheme outperforms the existing methods.

Table 1: Comparison between data privacy techniques

Techniques	Privacy	Accuracy	Power consumption
CPDA	Shared key security dependent	No data loss provides the exact sum	Requirement of huge amount of data communication for calculation of the result of aggregation
SMART	Privacy protection is 100%	If no loss of data then possible to get the exact sum	Transmit then receive data slices
GP^2S	100% privacy protection	Data histogram is plotted approximately	This involves no extra consumption of energy on additional communication and overhead
Based on K-Anonymity	Dependent on k which is a parameter	Results of query available beginning one interesting and (k-l) additional cells of uninteresting	Uninteresting cells data are queried and collected
HMDP	100% protection for privacy	No data is lost	No extra energy consumption

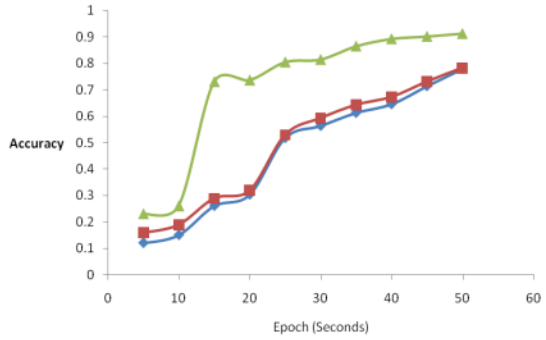

Fig. 8: Accuracy with varying second

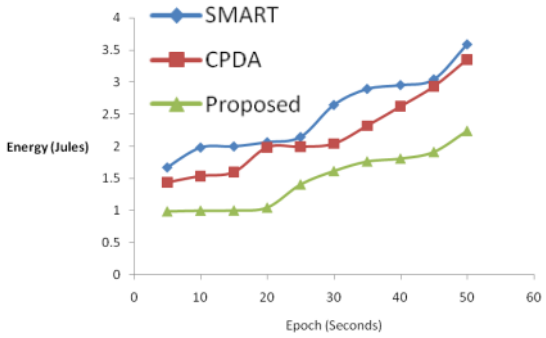

Fig. 9: Energy consumption with varying seconds

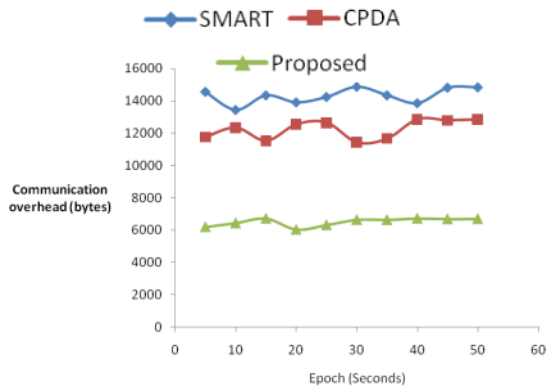

Fig. 10: Communication overhead with varying seconds

Table 1 depicts the comparisons. The proposed system is an efficient technique that gives a better tradeoff between the parameters privacy, accuracy and power consumption. The limitations of this model are the assumptions that the sensor nodes involved are of static in nature and this scheme works well with the multidimensional data.

In the proposed scheme to reduce the overhead, perturbation technique is mainly employed instead of encryption. The perturbed data does not require encryption as the geometrically perturbed data itself is having very high privacy.

Only the perturbation parameters, like rotation, translation and noise addition, need to be encrypted if necessary. This reduces the amount of work to be carried out which in turn reduces the power consumption and the overhead. Also, it reduces the cost of encryption of the entire database. The proposed scheme has reasonable computation cost.

CONCLUSION

We have proposed a new scheme, HMDP, in this paper to resolve the data privacy issue of the WSN. The main aim of the proposal is to reduce the power consumption and to have a better tradeoff between privacy and accuracy. The amount of data to be encrypted is minimized using the GDP and thus the power consumption is reduced.

We have shown that the new scheme which involves Geometric Data Perturbation technique and pair-wise key distribution outdo the existing techniques in terms of accuracy and power consumption and provides a matching privacy.

The proposed system is an efficient technique that gives a better tradeoff between the parameters privacy, accuracy and power consumption. GDP application in context-privacy area of WSNs offers itself as a potential area for research study.

REFERENCES

Akyildiz, I.F., W. Su, Y. Sankarasubramaniam and E. Cayirci, 2002. Wireless sensor networks: A survey. Comput. Networks, 38: 393-422.

Carbunar, B., Y. Yu, W. Shi, M. Pearce and V. Vasudevan, 2010. Query privacy in wireless sensor networks. ACM. Trans. Sensor Netw. TOSN., Vol. 6, 10.1145/1689239.1689244

Chan, H., A. Perrig and D. Song, 2003. Random key predistribution schemes for sensor networks. Proceedings of the IEEE Symposium on Security and Privacy, May 11-14, 2003, Berkeley, CA., USA., pp: 197-213.

Chen, K., G. Sun and L. Liu, 2007. Towards attack-resilient geometric data perturbation. Proceedings of the 2007 SIAM International Conference on Data Mining, April 26-28, 2007, SIAM International, Minneapolis, Minnesota, ISBN: 978-0-89871-630-6, pp: 78-89.

Eschenauer, L. and V. D. Gligor, 2002. A key-management scheme for distributed sensor networks. Proceedings of the ACM Conference on Computer and Communications Security, November 18-22, 2002, Washington, DC., USA., pp: 41-47.

He, W., X. Liu, H. Nguyen, K. Nahrstedt and T. Abdelzaher, 2007. PDA: Privacy-preserving data aggregation in wireless sensor networks. Proceedings of the 26th IEEE International Conference on Computer Communications, May 6-12, Anchorage, pp: 2045-2053.

Lamport, L., 1981. Password authentication with insecure communication. Commun. ACM, 24: 770-772.

Li, N., N. Zhang, S.K. Das and B. Thuraisingham, 2009. Privacy preservation in wireless sensor networks: A state-of-the-art survey. Ad Hoc Netw., 7: 1501-1514.

Liu, D. and P. Ning, 2005. Establishing pairwise keys in distributed sensor networks. ACM Trans. Inform. Sys. Security, 8: 41-77.

Madden, S.R., M.J. Franklin, J.M. Hellerstein and W. Hong, 2005. TinyDB: An acquisitional query processing system for sensor networks. ACM Trans. Database Syst., 30: 122-173.

Niu, J., L. Cheng, Y. Gu, L. Shu and S.K. Das, 2014. R3E: Reliable reactive routing enhancement for wireless sensor networks. IEEE Trans. Ind. Inform., 10: 784-794.

Zhang, N., S. Wang and W. Zhao, 2005. A new scheme on privacy-preserving data classification. Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, August 21-24, 2005, ACM, Chicago, Illinois, USA, ISBN: 1-59593-135-X, pp: 374-383.

Zhang, W., C. Wang and T. Feng, 2008. GP 2S: Generic privacy-preservation solutions for approximate aggregation of sensor data (concise contribution). Proceedings of the Sixth Annual IEEE International Conference on Pervasive Computing and Communications PerCom 2008, March 17-21, 2008, IEEE, Hong Kong, China, ISBN: 978-0-7695-3113-7, pp: 179-184.