I. J. Computer Network and Information Security, **2013, 7, 66-74**
Published Online June 2013 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2013.07.08

Integrity Protecting and Privacy Preserving Data Aggregation Protocols in Wireless Sensor Networks: A Survey

Joyce Jose
Post Graduate Scholar,
Dept. Information Technology, Karunya University, Coimbatore, India
joycejose1990@gmail.com

M. Princy
Lecturer,
Dept. Information Technology, Karunya University, Coimbatore, India
princym@karunya.edu

Josna Jose
Post Graduate Scholar,
Dept. Information Technology, Karunya University, Coimbatore, India
josnajose1990@gmail.com

Abstract—The data aggregation is a widely used energy-efficient mechanism in wireless sensor Networks (WSNs), by avoiding the redundant data transmitting to base station. The deployment of wireless communicating sensor nodes in the hostile or unattended environment causes attack more easily and the resource limited characteristics make the conventional security algorithms infeasible, hence protecting privacy and integrity during data aggregation is a challenging task. The privacy of a sensor data ensures, it is known only to itself and the integrity guarantees sensor data has not tampered during data aggregation. The Integrity Protecting Privacy preserving Data Aggregation (IPPDA) protocols ensures a robust and accurate results at the base station. This paper summarises on such IPPDA protocols during data aggregation.

Index Terms **—**Wireless Sensor Networks, data privacy, data integrity, energy efficiency, data aggregation

I. INTRODUCTION

The Wireless Sensor Networks (WSNs) [1] consist of a large number of small, low cost and resource constrained sensor nodes to cooperatively monitor the environmental conditions at different locations. Sensor nodes have the capability of sensing, processing and communicating the data, also is resource constrained in terms of power, memory and computation capabilities. Each sensor node collects the data from the environment in which it is deployed and report it to the base station located at the remote place by two ways, either single hop or multi hop. In single hop communication, each node directly sends the data to the sink, it requires considerable power consumption. Comparatively in multi hop communication, each node in the path to the sink act as a repeater, thereby reducing the long range transmission.

Three types of nodes are present in a WSN; these are Base Station (BS or sink or Query Server), intermediate node (aggregator) and leaf node (normal sensor node). The BS is a node where the aggregation results are destined, responsible for processing the received data from the sensor network derives meaningful information reflecting the events in the target field. The intermediate node performs sensing, aggregation and forward data from the leaf node to upper aggregator or sink. The normal sensor node sense, aggregate and forward data. The sensor networks can be deployed to monitor environment, habitat, military and surveillance applications.

The data collected from the sensor nodes is correlated in terms of time and space, transmit partially processed data to the sink node, which requires data aggregation [2]. It is the process of gathering data from the sensor nodes and aggregate these data using aggregation functions such as MAX, MIN, SUM, AVERAGE, HISTOGRAM, etc. The data aggregation avoids redundant data and limits number of transmissions by minimizing communication overhead to extend network lifetime. Extension of this approach is in-network data aggregation [3], where aggregated data are progressively passed through the network.

Protecting the privacy of data collected from the sensor node is a challenge in the data aggregation. Data privacy can be defined as the process in which the adversaries can overhear and decrypt the data. But still it can provide a mechanism to prevent them from getting the private information. i.e, control disclosure of any details about the data. To achieve privacy, it is required to protect the transmission trend of a node's secret data from neighbors, because the neighbors know the aggregated sum and the encryption key. It looses the end to end confidentiality. Two types of privacy concerns in WSNs: internal privacy and external privacy. Internal privacy is about to maintain the privacy of a sensor node from other trusted participating sensor nodes of the WSN, whereas the external privacy is about to protect the data from the outsiders (adversaries).

Since the aggregation result is used for making the critical decisions, the accuracy of data received at BS is crucial, so the aggregation result must be verified before accepting it. Thus, ensuring integrity of the data is important in WSNs. It maintains the consistency and correctness of the message. In this paper, we provide an overview and comparison of different existing IPPDA protocols. To the best of our knowledge, this is the first survey about the integrity protecting and privacy preserving data aggregation protocols in WSNs.

The rest of the paper is organized as follows. Section II classifies the existing IPPDA protocol. Section III presents major application areas of IPPDA protocols. In section IV, we compare the different IPPDA protocols based on some metrics. Section V concludes the work.

II. CLASSIFICATION OF IPPDA PROTOCOLS

Based on the type of sensor nodes, protocols for integrity protecting and privacy preserving data aggregation is of two types, heterogeneous and homogeneous protocols. In homogeneous protocols, all the nodes in the network have the same resources, an aggregator performs sense, aggregate and forward the result to sink. Any sensor nodes can play the role of an aggregator. In heterogeneous protocols, more than one type of sensor nodes exists, an aggregator only aggregates the data, transmits result to sink, but does not sense. The IPPDA protocols are further classified based on the network topology, cluster and tree structure.

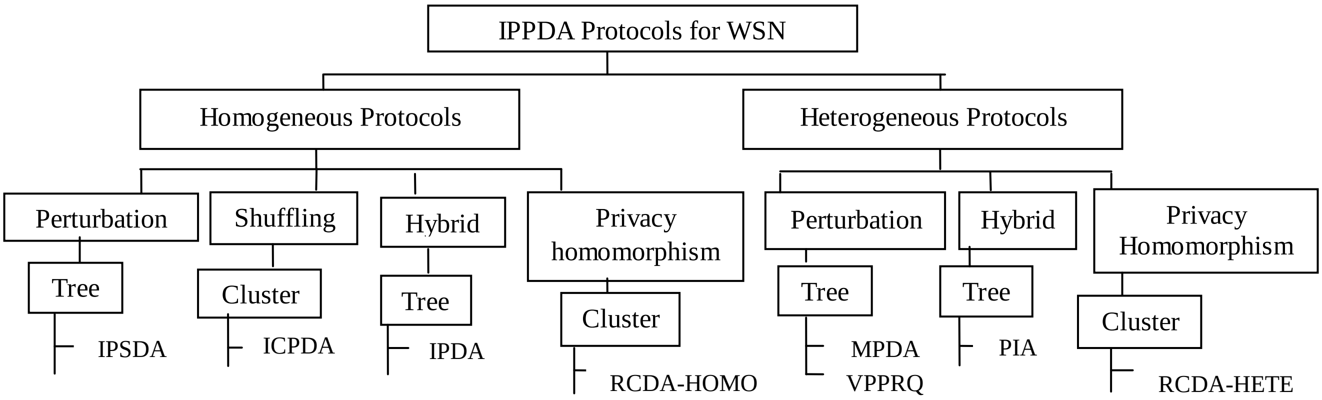


Figure 1. Classification of existing IPPDA protocols for wireless sensor networks

A. *Homogeneous Protocols*

The homogeneous protocols are divided into different types; such as perturbation, shuffling and privacy homomorphism.

1) *Perturbation*

In this technique, each sensed data is customized using encryption key and public or private seed, which is generated by randomization technique [4,5] to hide a sensitive data.

a) *IPSDA*

The IPSDA (Integrity Protecting Sensitive Data Aggregation) [6] scheme for WSNs overcomes the high communication and computational overhead of iPDA [7] protocol, but requires a considerable memory at each node. It uses an additive property of complex numbers to check integrity in data aggregation and achieve privacy from other trusted participating node as well as from adversaries. In the two parts of a complex number, the real part is used for privacy preservation and imaginary part is used for integrity checking. Every node share two keys, one key is shared with master device and other is shared with those sensor nodes lying on the aggregation tree.

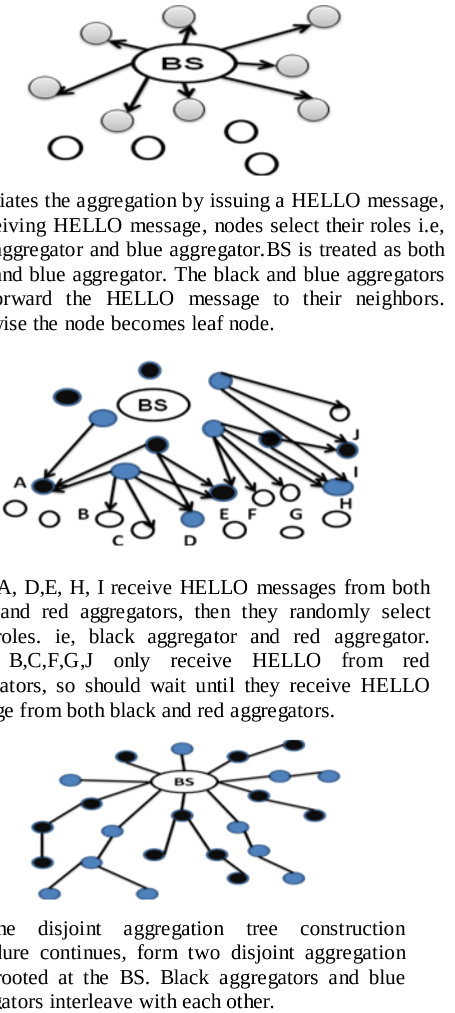
After receiving a query from the BS, each sensor node customizes its data into a complex number by combining sensitive data with a private real number and adjoins an imaginary number to it. The private real number and imaginary numbers are stored in sensor node memory and in sink node. Each node encrypts and sends the customized data to its parent node using the shared key between them. After receiving the customized value, the parent node decrypts the received value and waits for some time to guarantee that all slices are received. Then aggregate the customized data by using additive properties of complex number and send to the sink after the encryption. After receiving all the aggregated result from the intermediate nodes, sink node sum up the aggregated data. In order to get the actual data and to check the integrity, first separate the real part and imaginary part of the sum. Subtracting the sum of private real seeds from the real part of the aggregate at the BS will give the actual sum. For checking the integrity, compare the imaginary part of the sum with the sum of imaginary seeds of all sensor nodes.

2) *Shuffling*

In shuffling each node slices its data into *k* number. One piece is kept in the node itself, and the remaining *k-1* slices are encrypted and send to the *k-1* neighbors within the *h* (*h=1*, for a dense sensor networks) hops.

a) iPDA

In iPDA (integrity protecting Private Data Aggregation) [7], data privacy is achieved through slicing and assembling technique, integrity is achieved through redundancy by constructing two disjoint aggregation tree. Since each node belongs to a seperate aggregation tree, malicious nodes can only pollute the aggregation result of the aggregation tree it belongs. Hence by comparing the aggregation results from the two aggregation tree, the BS can verify the integrity of the aggregation result. Figure 2 shows the steps in the disjoint aggregation tree construction.



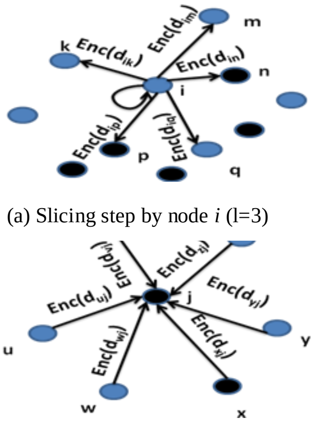
(a) BS initiates the aggregation by issuing a HELLO message, on receiving HELLO message, nodes select their roles i.e, black aggregator and blue aggregator.BS is treated as both black and blue aggregator. The black and blue aggregators will forward the HELLO message to their neighbors. Otherwise the node becomes leaf node.

(b) Node A, D,E, H, I receive HELLO messages from both black and red aggregators, then they randomly select their roles. ie, black aggregator and red aggregator. Node B,C,F,G,J only receive HELLO from red aggregators, so should wait until they receive HELLO message from both black and red aggregators.

(c) As the disjoint aggregation tree construction procedure continues, form two disjoint aggregation trees rooted at the BS. Black aggregators and blue aggregators interleave with each other.

Figure 2. Procedure for constructing a disjoint aggregation tree

To achieve privacy, each node slices its data into *l* number pieces, one is kept in the node itself and remaining *l-1* pieces are encrypted, sent to *l-1* neighbor nodes in the aggregation tree it belongs and *l* neighbor nodes in the other aggregation tree. Each node takes *2l-1* transmission in slicing step. Each node sum up received slices with one of the piece in the node and send to its parent within the same aggregation tree. The aggregation results from two different aggregation tree are compared with each other, if too much difference is obtained due to pollution attack, then the BS will reject the aggregated data. If not too much, the BS will accept. Figure 3 shows the slicing and assembling procedure in iPDA.



(a) Slicing step by node *i* (l=3)

(b) Assembling at node *j*

Figure 3. Privacy preserving data report

The advantages of iPDA are; it can build on any key management scheme. It has the low computation overhead. However, it has high communication overhead, because each node has to send the data to both the aggregation trees. It can tolerate the collusion up to a certain threshold by expanding the number of slices, so the communication overhead also raised. The accuracy of the aggregated result is reduced due to the increase in the number of messages and collision on the network.

3) *Hybrid*

It can use more than one technique to achieve privacy preservation.

a) ICPDA

The ICPDA (Integrity-protecting CPDA) [8] protocol is the extension of CPDA [9] for checking the integrity in cluster topology. In ICPDA, all nodes can participate in the calculation of an intermediate aggregation result. Hence all peer nodes in a cluster can monitor the behavior of the Cluster Head (CH), and it can report the malicious behavior of CH to the BS.

In ICPDA, privacy preserved data aggregation can be done in two ways: based on algebric properties of polynomials and Secure Multiparty Calculation (SMC) [10]. By using algebric properties of polynomials, every sensor node in each cluster customizes its private data to a polynomial of order *k-1*, using shared seeds (non private) and random numbers (private), where *k* is the total number of nodes in a cluster. Then every sensor node encrypts its customized data and sends to all nodes using a unique shared key. In SMC, slicing and assembling technique is used for privacy preservation. In both data aggregation methods, every node broadcast its assembled data to all nodes in the cluster, on receiving every node can calculate the partially aggregated result, subsequently CH sends an intermediate aggregated result to BS with signature. After receiving the aggregate, the sink checks the result, if any data found to be tampered it will discard the aggregate.

The ICPDA can detect data pollution attack by enabling the peer monitoring mechanism, which requires higher bandwidth, thereby increases the communication overhead than CPDA. It is incapable of detecting the data collusion attack. Accuracy level is same as that of CPDA.

4) *Privacy Homomorphism*

In privacy homomorphism, the arithmetic operations are done on the encrypted data without decryption, which reduces the energy consumed for the decryption at aggregators.

a) RCDA

The RCDA (Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks) [11][21] uses elliptic curve based additive privacy homomorphism (EC-ElGamal) technique, which allow aggregation to be carried out in cipher text directly, no need to decrypt at aggregator node for data aggregation. Thus, RCDA achieves end to end data privacy with reduced energy. i.e, the sensor data encrypted with BS public key is decrypted at BS with its private key. The RCDA provides a mechanism to recuperate all sensed data from the aggregated result at BS, to verify the integrity and authenticity of all sensed data, after which BS can perform any type of aggregation, later to the recoverd data. RCDA uses aggregated digital signature scheme to provide authenticity and integrity of all sensing data. RCDA provides a mechanism for homogenous (RCDA-HOMO) and heterogeneous network (RCDA-HETE) respectively.

- *RCDA-HOMO*

The sensor node generates the cipher text using base station's public key and digital signature using its private key to CH. Data and digital signature from different sensor nodes are aggregated separately to produce aggregated cipher text and aggregated digital signature respectively at every CH. Upon receiving in the BS, it decrypts the aggregated cipher text using BS's private key and recovers all sensing data. Thus, BS can overcome the limitation of aggregating and verifying the integrity of all sensed data using aggregated digital signature. This scheme for heterogeneous network is called native RCDA-HETE scheme.

- *RCDA-HETE*

It consists of two types of sensors. H sensors and L sensors. H sensors are tamper resistant with strong computing capability. To reduce the computation overhead of L sensors it uses symmetric key shared to the H sensor for encryption and send the encrypted data to H sensor, it decrypts the data coming from different sensor nodes and generates an aggregated result. CH encrypts it with the BS public key, generate a digital signature of the aggregated data and pass the aggregated cipher text and signature pair to the higher level H sensor. Here digital signature and aggregated cipher text are aggregated separately with other H sensor's aggregated cipher text and digital signature. Upon receiving at BS, integrity of each aggregated data is verified by recovering it from the aggregated data and the integrity of each sensed data is verified using the MAC received from the sensors during Intercluster Encrypt phase.

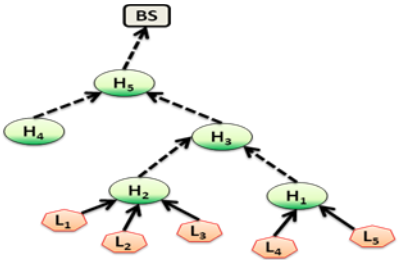


Figure 4. Heterogeneous WSN

B. *Heterogeneous Protocols*

In this, the sensor nodes in the network are different in terms of resources.

1) *Perturbation*

This technique is same as perturbation technique in homogeneous networks.

a) MPDA

The MPDA (Multidimensional Privacy–preserving Data Aggregation scheme for Wireless Sensor Networks) [12] can combine more than one diverse sample data. It uses a super increasing sequence and a perturbation technique for multidimensional aggregation, the aggregated result arrives at sink. The MPDA consists of four steps: system initialization, sensor and aggregation node initialization, deployment and neighbor key discovery phase and multidimensional privacy preserving data aggregation. Figure 5 shows the multidimensional privacy preserving scheme.

For the secure transmission of data, each node and aggregator compute *k* neighbor key, and each neighbor key is symmetrically shared by the aggregator and sensor node (pairing). The multidimensional privacy preservation consists of two parts, private data aggregation and recovery of aggregated data. First part focuses on additive aggregation, when the sensor node receives a query from the sink, the node calculates the customized value (c_i) by using the private key (S_i) and the sequence a= (a_1…. a_n) produced by the sink.

$$c_i = \sum_{j=1}^{n} a_j.(m_{ij} + b_i) \bmod p \qquad (1)$$

Where b_i=h (E||S_i). E is the unique identifier of the query. It then calculates the hash, h_i=h(c_i||key_{ij}||T) and sends the message to the aggregator node in the format (N_i||T||c_i||h_i). After receiving the message from the neighbor sensor node, the aggregator checks the timestamp for transmission delay. If it is within the period, then aggregator node verifies the message by calculating the hash using the neighbor key (key_{ij}). If it is correct, accept the encrypted message c_i, otherwise reject it. After receiving the k valid encrypted data from the sensor nodes, the aggregator node calculates the aggregated data c= $\sum_{i=1}^{k} c_i \bmod p$. Then encrypts the aggregated result using self-encryption technique [13] and send to the sink. The self encryption technique can hide the node information with the public key, then no one knows the information except the sink. After receiving the message at the sink, it can recover the message by using aggregated data c, from aggregator. Calculate the difference in time, if it is not in time, the message will be rejected. Then compute hash and compare it with the hash received, if it is not equal, reject. The message (M) is recovered using the equation (2).

$$M=c-\sum_{j=1}^{n} a_j \cdot \sum_{i=1}^{k} b_i \bmod p \qquad (2)$$

MPDA scheme preserves privacy against passive attack and sensor node compromise attack. The compromising of a sensor node does not disclose the data sensed by other nodes or aggregator nodes and its private key.

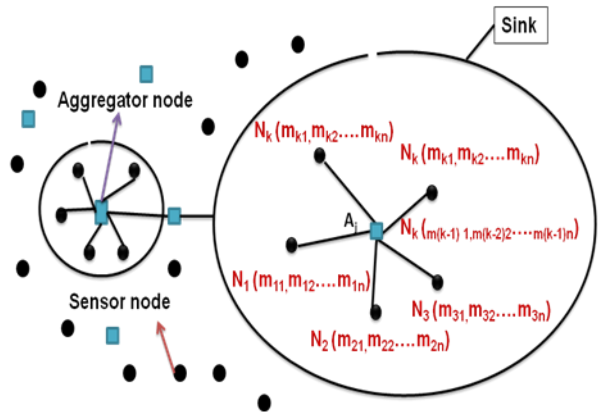

Sink
Aggregator node
Sensor node
$N_k(m_{k1,}m_{k2}....m_{kn})$
$N_k(m_{k1,}m_{k2}....m_{kn})$
A_j
$N_k(m_{(k-1)\ 1,}m_{(k-2)2}....m_{(k-1)n})$
$N_1(m_{11,}m_{12}....m_{1n})$
$N_3(m_{31,}m_{32}....m_{3n})$
$N_2(m_{21,}m_{22}....m_{2n})$

Figure 5. Multidimensional Privacy-preserving data aggregation in WSNs

b) VPPRQ

In VPPRQ (Verifiable Privacy-Preserving Range Query in two tired sensor networks) [14] provide a privacy preserving storage scheme based on the bucketing scheme [15,16]. In this scheme, every sensor node generates environmental data values in a fixed rate and periodically submits the collected data to the closest storage node for each epoch (time interval between two submissions). Every sensor node has a unique ID, and it shares a distinct secret key with sink, so that compromising one sensor node does not affect the security of other sensor data. Before sending the data to the storage node, encrypt the data using the shared key between sensor node and sink, then attach a tag to the encrypted data based on which bucket the data falls into. The data values with same tag can be encrypted as a block. Upon receiving the encrypted data from source node, the storage node divides the data into multiple buckets, and each bucket is assigned a tag. The sensors and sink should agree on the same bucket partition. When a sink node wants to process a range query over the data stored on the storage node, it obtains the result based on the tag corresponding to the query range, instead of decrypting the data and return the result. This reduces the energy wasted for the decryption. At the sink, it decrypts the received data based on the key shared with the sensor node, obtain the real sensor data. The main problem is the transmission delay, i.e, the time required to send data from source to sink. It is unsuitable for applications like event detection and tracking objects because this application requires an immediate response. This scheme focus on the problem when the storage node and sensor node is compromised. The compromising of storage node leads to reveal sensed data of nodes and data fidelity attack (the adversary tries to reply wrong information for the user query and make them to accept it). To prevent the storage node from false reply, an encoding number is generated on each sensor if no data in a query range are collected from the sensor.

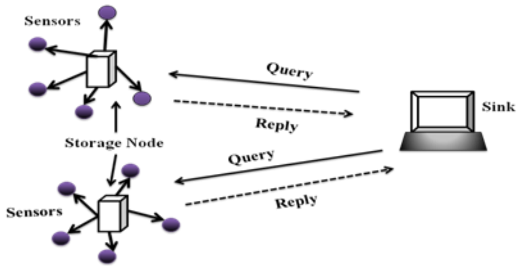


Figure 6. Two-tier model

2) Hybrid

It is same as in homogeneous networks. It uses more than one type of privacy preserving technique in data aggregation.

a) PIA

The PIA (Privacy Preserving Integrity Assured data aggregation) [17] addresses the integrity assured data aggregation with efficiency and privacy as joint objective. Two types of Integrity Verification (IV): centralized and distributed. In both types, the integrity of aggregate is verified by re-computing the aggregation function of the raw data. In centralized IV, after receiving data from sensor node, the server checks the integrity. In distributed IV, the sensor node recompute the aggregation function using the data from other sensors. If all the results obtained from sensor node is same, then aggregation is considered as secure. It distributes the communication throughput in the network. The PIA proposed four symmetric key solutions for the single aggregator model.

In the first solution, homomorphic encryption scheme hides data during transmission in a centralized IV . It combines the homomorphism and MAC [18] to construct an authenticated encryption scheme for the aggregator node. The MAC is used to check the integrity of the aggregated result at the sink. It only supports additive aggregation functions such as SUM, average and standard deviation.

The second solution uses the Order Preserving Encryption Scheme (OPES) [19] to preserve the privacy of distribution of data. OPES uses to verify the integrity of comparison based aggregation function. Sensors encrypt the data by using OPES with master secret key shared by all sensors then aggregator calculates and sends to the user. The user decrypts the aggregated data in centralized IV.

The third solution uses a secure hierarchical in networking aggregation (SHIA) [20] scheme for adapting distributed IV. This scheme supports any aggregation function because the sensor nodes have the access to all raw data. The communication overhead of each node is *O(N)*. Where *N* is the total number of sensor node in a network. This scheme preserves privacy without using any additional privacy preservation mechanism such as encryption. A commitment is constructed during the aggregation process. After calculating the aggregate, each sensor node independently reconstructs the commit tree and ensure that the data is not modified or discard the contribution of the node by any adversary.

The fourth solution used to improve the privacy and integrity of the third solution by using a logical aggregation tree within the aggregator node. Each sensor node has a communication overhead of *O (log N)*. It only supports decomposable functions such as mean, standard deviation, count, MIN/MAX. It uses distributed IV.

III. APPLICATION AREAS OF IPPDA PROTOCOLS

A. *Health monitoring*

There are two major health monitoring application for WSNs. First, to monitor the performance of an athlete such as tracking respiration and pulse rate using wearable sensors. Secondly, to monitor the health conditions of patients such as personal weight, blood sugar level, blood pressure level, etc. The sensor measurements should be kept secret from other people during the transmission to sink node, so the integrity and privacy of the data are needed.

B. *Military Surveillance*

The WSNs replaces the guards and sentries around the defensive areas, so it helps to keep the soldiers out of harm's way. It also helps to locate and identify troops, vehicles and targets for potential attack. So the privacy and integrity of the sensor data are always critical and it should be preserved during data aggregation.

C. *Private Households*

Sensors could be placed in houses to collect the details of water, gas and electricity consumption within a large neighborhood. The aggregated population statistics are helpful for individuals, business and government agencies to plan the resources. However, individual readings reveal the daily activities of households such as, which time all family members are absent in home.

IV. COMPARATIVE STUDY OF DIFFERENT IPPDA PROTOCOLS

A. *Communication Cost (CMC)*

Communication overhead is the major problem in WSNs. It is calculated by counting the number of messages generated by each node in the WSNs. The communication cost of protocols belongs to three classes: Low, Medium, High, when the number message (m) generated per node $m \geq 3$, $3>m>1$, $m=1$ respectively.

B. *Computational Cost (CPC)*

This is the processing overhead of processor to achieve privacy preserving data aggregation. The values are High, Medium and Low. The CPC is high: if a sensor node performs many encryption/decryption, arithmetic operation and other operations. Medium: If a node performs a couple of encryption/decryption, some arithmetic operation. Low: if a sensor node performs few arithmetic operations, one encryption or decryption.

C. *Energy Consumption (EC)*

This is the total energy spent by the WSN to collect data from the source node. It is calculated based on the size of the payload, and the number of messages generated in the network during data aggregation. The values are High, Medium, Low.

D. *Accuracy (AUC)*

The final decision at the BS is based on the aggregation accuracy. The aggregation accuracy is defined as a ratio between the aggregated data received at the sink and the sum of actual data. The values are Low, Medium, High.

E. *Encryption Type (ET)*

It determines the type of encryption. It is either end to end encrypted data aggregation or hop by hop encrypted data aggregation. In an end to end encrypted data aggregation, only one decryption at sink node. Other intermediate nodes, aggregate the encrypted data using a privacy homomorphism technique. In hop by hop encrypted data aggregation, each node performs the encryption and decryption. Here the aggregation is done on the plain text, so the chance of modification of data is more.

F. *P vs OUT, P vs IN, P vs BS*

It determines whether the protocol protects the privacy of data from outsiders, insiders and BS. This is specified as privacy vs outsiders (P vs OUT), privacy vs insiders (P vs IN) and privacy vs BS (P vs BS).

G. Data Loss Resiliency (DLR)

To determine, the protocol can compute the correct aggregate even though there is some data loss or some nodes fail to participate. The values are Yes or No.

H. Delay (DLY)

This is the time taken to get the sensed data from the source to the sink. The delay is less in the end to end encrypted data aggregation and more in hop by hop encrypted data aggregation due to the decryption at aggregator. The increased delay causes an increase in communication overhead. It thereby increases the energy consumption. The values are High, Medium, Low.

I. Aggregation Function (AF)

To determine how many aggregation functions can support by the aggregation scheme, among Sum, Average, Count, Standard deviation, Max, Min, Variance, Histogram and Median. The two classes: Numerous and Few.

J. Pay Load Size (PLS)

It determines the real size of the message after applying privacy preserving operation. The values are Large, Medium, Small.

K. Memory Consumption (MC)

The WSNs are limited in memory. Memory consumption determines the amount of memory needed to store the keys, integer ranges and variables.

L. Network Topology (NT)

It determines the topology used in data aggregation, either tree or cluster.

Table 1. Comparison of different Integrity Protecting and Privacy Preserving Data Aggregation (IPPDA) protocols

Protocols	CMC	CPC	EC	AUC	ET	P vs OUT	P vs IN	P vs BS	DLR	DLY	AF	PLS	MC	NT
IPSDA	L	L	L	H	Hop by hop	Y	Y	Y	N	M	F	G	H	Tree
IPDA	H	M	H	H	Hop by hop	Y	Y	Y	N	H	F	S	M	Tree
ICPDA	H	L	M	H	Hop by hop	Y	Y	Y	N	M	F	M	H	Cluster
RCDA	M	M	M	H	End to end	Y	Y	N	N	L	U	G	H	Cluster
MDPA	L	H	M	H	Hop by hop	Y	Y	Y	N	M	F	G	H	Tree
VPPRQ	L	M	M	H	End to end	Y	N	N	Y	H	-	G	H	Tree
PIA	L	M	L	M	Hop by hop	Y	N	N	Y	M	U	S	M	Tree

Legend: Y=Yes, N=No, "-"=not mentioned, H=High, L=Low, M=Medium, F=Few, G=Large, S=Small, U=Numerous

VI. CONCLUSIONS

Privacy preserving and integrity protecting data aggregation is crucial for security critical applications. This paper presented seven different IPPDA protocols in WSNs, that achieves privacy preservation and integrity protection as joint objective. The IPPDA protocols are classified based on the type of aggregator, technique used to achieve privacy and topology used for data aggregation. This paper also discussed some of the application areas of IPPDA protocols and finally, we compare the existing IPPDA protocol based on some performance metrics. The resource limited sensor nodes need new approaches for secure data aggregation. The energy-efficient and high accuracy secure data aggregations are key areas of research in WSNs. We hope this paper will be a guide to know the working of varied protocols in nutshell.

REFERENCES

[1] Vaibhav Pandey, Amarjeet, Narottam Chand, A review on data aggregation techniques in wireless sensor network, *Journal of Electronics & Electrical Engineering*, ISSN: 0976-8106 & E-ISSN: 0976-8114, Vol.1, Issue 2, 2010, pp-01-08.

[2] Nadini.S.Patil, Prof. P. R. Patil, Data Aggregation in wireless sensor network, *IEEE International Conference on Computational Intelligence and Computing Research*, 2010, ISBN 97881 8371 3627.

[3] K. Akkaya and I. Ari. In-network Data Aggregation in Wireless Sensor Networks, Handbook of Computer Networks, Ed. H. Bidgoli, John Wiley & Sons, Vol. 2, pp. 1131-1146, 2008.

[4] Agrawal, R; Srikant, R, Privacy-preserving data mining, In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, Dallas, TX, USA, May 15–18, 2000; pp. 439–450.

[5] Kargupta, H.; Datta, Q.W.S.; Sivakumar, K, On the privacy preserving properties of random data perturbation techniques, In *Proceedings of the IEEE International Conference on Data Mining*,

Melbourne, FL, USA, November 19–22, 2003; pp. 99–106.

[6] Rabindra Bista, Hye-Kyeom Yoo, Jae-Woo Chang, A New Sensitive Aggregation Scheme for Protecting Integrity in Wireless Sensor Networks, proceedings of *the 10th IEEE International Conference on Computer and Information Technology (CIT 2010)*.

[7] W.He, X. Liu, H. Nguyen, K. Nahrstedt, T. Abdelzaher, iPDA : An Integrity –Protecting Private Data Aggregation Scheme for Wireless Sensor Networks, *IEEE MILCOM*, November 2008, pp 1-7.

[8] Wenbo He, Xue Liu, Hoang Nguyen, Klara Nahrstedt, "A Cluster-Based Protocol to Enforce Integrity and Preserve Privacy in Data Aggregation", Proceedings of the *29th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW '09)*, 2009 pp 14-19.

[9] He W, Liu X, Nguyen H, Nahrstedt K, Abdelzaher T, PDA: Privacy-preserving data aggregation in wireless sensor networks. *Proceedings of 26th IEEE International Conference on Computer Communications (Infocom 2007)*, Anchorage, Alaska, USA, May 2007:2045-2053.

[10] O. Goldreich, Secure multi-party computation, Working Draft, Version 1.3, 2011.

[11] Chien -Ming Chen, Yue-Hsun Lin, Ya-Ching Lin, Hung-Min Sun, RCDA: Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks. *IEEE TRANSACTIONS ON PARRALLEL AND DISTRIBUTED SYSTEMS*, VOL 23, NO 4, APRIL 2012.

[12] Xiaodong Lin, Rongxing Lu, Xuemin (Sherman) Shen, MPDA: multidimensional privacy-preserving aggregation scheme for wireless sensor networks, *Wireless Communications and Mobile Computing*, 2010; 10:843-856.

[13] Lin X, Lu R, Shen X, Nemoto Y, Kato N, SAGE: a strong privacy preserving scheme against global eavesdropping for e-health systems, *IEEE Journal of Selected Areas of Communications* (in press).

[14] Sheng, B.; Li, Q. Verifiable privacy-preserving range query in two-tiered sensor networks. In *Proceedings of the 27th IEEE International Conference on Computer Communications*, *INFOCOM*, Phoenix, AZ, USA, April 15–17, 2008; pp. 457–465.

[15] Hacigumus, H.; Iyer, B.R.; Li, C.; Mehrotra, S. Executing SQL over encrypted data in the database service provider model. In *Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data*, Madison, WI, USA, June 3–6, 2002; pp. 216–227.

[16] Hore, B.; Mehrotra, S.; Tsudik, G. A privacy-preserving index for range queries. In *Proceeding of the 28th Very Large Database Conference*, *VLDB*, Toronto, Canada, August 29–September 3, 2004; pp. 720–731.

[17] Taban, G.; Gligor, V.D. Privacy-preserving integrity-assured data aggregation in sensor networks. In *Proceeding of International Symposium on Secure Computing*, *SecureCom*, Vancouver, Canada, August 29–31, 2009; pp. 168–175.

[18] Bellare, M.; Canetti, R.; Krawczyk, H. Keying hash functions for message authentication. In *Proceedings of Annual International Cryptology Conference, Crypto*, Santa Barbara, CA, USA, August 18–22, 1996; pp. 1–16.

[19] Agrawal, R.; Kiernan, J.; Srikant, R.; Xu, Y. Order preserving encryption for numeric data. In *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data*, Paris, France, June 13–18, 2004, pp. 563–574.

[20] Chan, H.; Perrig, A.; Song, D. Secure hierarchical in-network aggregation in sensor networks. In *Proceedings of the 13th ACM Conference on Computer and Communications Security*, Alexandria, VA, USA, October 30–November 3, 2006; pp. 278–287.

[21] Hung-Min Sun, Chien-Ming Chen, Yue-Hsun Lin, and Ya-Ching Lin, Supplementary Material: "Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks", SUPPLEMENTAL MATERIAL FOR THE IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, available on http://doi.ieeecomputersociety.org/10.1109/TPDS.2011.29.

Joyce Jose received the B Tech degree in Computer Science & Engineering from Mahatma Gandhi University in the year 2011, Currently she is a Post Graduate Scholar in Network and Internet Engineering at Karunya University. She has published a paper titled "A survey on Secure Data Aggregation Protocols in Wireless Sensor Networks" in International Journal of Computer Applications (0975 – 8887) Volume 55– No.18, October 2012. Her research interest in the area of data aggregation in Wireless Sensor Networks.

M. Princy pursued her M Tech degree specialized in Network and Internet Engineering, in the year 2011, her research interest is in the area of Sensor Networks , has presented papers in various national and international conferences , also published a paper in International Journal of Computer Applications paper titled "*Analysis on Scheduling and Load Balancing Techniques in Wireless Mesh Networks*", March 2012. She has done project on sleep scheduling algorithm for power efficiency in Sensor Networks and guiding projects in Power efficient aggregation and routing in Sensor Networks and planned to pursue her doctorate in the same.

Josna Jose received the B Tech degree in Computer Science and Engineering from Mahatma Gandhi Univeristy in the year 2011, Currently she is a Post Graduate Scholar in Network and Internet Engineering at Karunya University. She has published a paper titled “A survey on Secure Data Aggregation Protocols in Wireless Sensor Networks” in *International Journal of Computer Applications (0975 – 8887) Volume 55– No.18, October 2012.* Her research interest in the area of data aggregation in Wireless Sensor Networks.