Journal of Computing and Information Technology - CIT 22, 2014, 1, 1–20
doi:10.2498/cit.1002318

A Survey on Privacy Preserving Data Aggregation Protocols for Wireless Sensor Networks

Joyce Jose, Josna Jose and M. Princy

Dept. Information Technology, Karunya University, Karunya Nagar, Coimbatore, India

The data aggregation is a widely used mechanism in Wireless Sensor Networks (WSNs) to increase lifetime of a sensor node, send robust information by avoiding redundant data transmission to the base station. The privacy preserving data aggregation is a challenge in wireless communication medium as it could be eavesdropped; however it enhances the security without compromising energy efficiency. Thus the privacy protecting data aggregation protocols aims to prevent the disclosure of individual data though an adversary intercept a link or compromise a node's data. We present a study of different privacy preserving data aggregation techniques used in WSNs to enhance energy and security based on the types of nodes in the network, topology and encryptions used for data aggregation.

Keywords: data integrity, homogeneous networks, heterogeneous networks, privacy preserving data aggregation, wireless sensor networks, data privacy, end to end data aggregation, hop by hop aggregation

1. Introduction

The Wireless Sensor Networks [1] is an ad hoc network consisting of a large number of distributed autonomous wireless devices called sensors, which is densely deployed in remote areas to detect the environmental conditions such as temperature, pressure, humidity, sound, vibration, motion, pollutants etc. The sensor nodes are resource constrained in terms of energy, memory and computation capabilities. There are three types of nodes: normal sensor nodes (leaf nodes), intermediate nodes (aggregator), base station (BS or sink or query server). A sensor node has the capability of sensing, processing and communicating the data collected from the environment in which it is deployed and report it to the base station located at remote places. An aggregator aggregates sensed data with other data received from multiple sensor nodes based on some preferred aggregation functions (SUM, AVG, Count, MAX, MIN etc.) and forward aggregation results to another aggregator or BS. Finally, the BS processes the received data and derives the significant information reflecting the events in the target field.

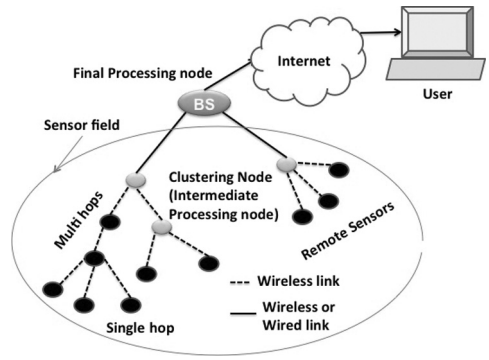


Figure 1. Typical wireless sensor network arrangement.

The dense deployment of resource constraint sensor nodes in terms of energy, memory, bandwidth, communication and computational capabilities in close proximity sense the same data, which in turn increases the redundant data in the network. The transmission of these redundant data incurs the sensor node energy. A data aggregation [2] mechanism avoids the redundant data transmission employed in WSNs at aggregator level, which reduces the energy consumption of a node and thereby increases the network lifetime. Figure 2 shows the impact of data aggregation in WSNs. The extension

of this approach is in-network data aggregation [3] which aggregates data progressively as the data passes through the network. The data aggregation can be done in tree, cluster or hybrid topology.

The remote/hostile environment of wireless sensor nodes makes the security a challenging problem during data aggregation. However the use of WSNs in security critical application increases the need of data privacy during aggregation. The eavesdropping attack in the wireless link and the compromise attacks in the sensor nodes loose data privacy during data aggregation and help the adversaries to know more about the behavior of the individual nodes. The privacy preserving data aggregation protocols achieve data privacy by protecting the transmission of a node's data from its neighbors, because the neighbors can know the aggregated sum and encryption key by compromising of nodes. There are two types of privacy concerns in WSNs: internal privacy and external privacy. The internal privacy is to maintain the privacy of a sensor node from other trusted participating sensor nodes in the WSNs, whereas the external privacy is about to protect the data from outsiders (adversaries). Thus the privacy protecting data aggregation aims to protect individual nodes' privacy by different transmission trends, in such a way that the adversaries can't get the sensitive information of a particular node even if the adversaries can overhear and decrypt the data. In addition, the verification of data integrity in privacy preserving protocols maintains the consistency and correctness of messages and it provides a guarantee to the BS that the data is original with no alternation during transmission. This increases the accuracy of aggregation which helps the BS to take critical decisions based on the aggregated result reached at the BS.

To address the privacy of sensor data during aggregation, many data aggregation protocols exist in WSNs. In this paper, we provide a comprehensive summary of different privacy preservation protocols in the WSNs and some of the application areas.

The rest of the paper is organized as follows. Section 2 includes the application areas of PPDA protocols. Section 3 classifies the existing privacy preserving data aggregation protocol. In Section 4, we compare the different PPDA protocols based on some metrics. Section 5 concludes the work.

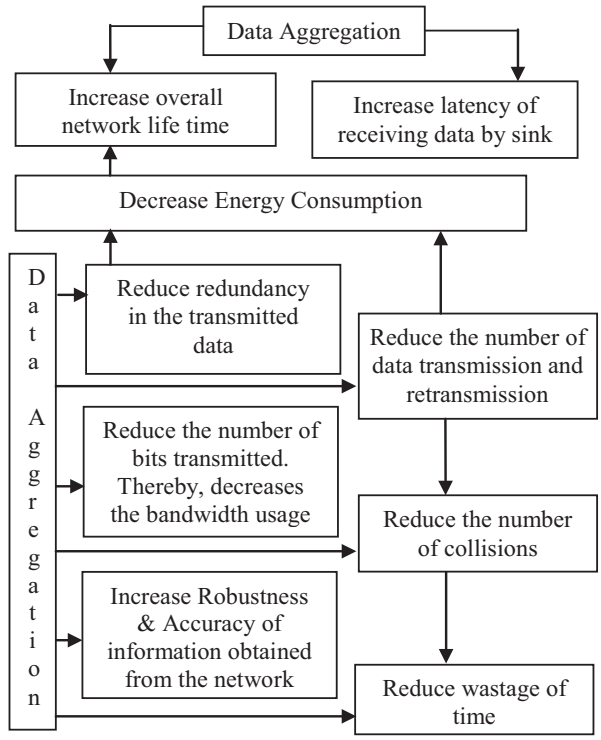


Figure 2. Impact of data aggregation in WSNs.

2. Application Areas of PPDA Protocols

2.1. Health Monitoring

There are two major health monitoring applications for WSNs [5]. First, to monitor the performance of an athlete such as tracking respiration and pulse rate using wearable sensors. Secondly to monitor the health of patients, e.g. personal weight, blood sugar level, blood pressure level, etc. The sensor measurements should be kept secret from other people during transmission to sink node.

2.2. Military Surveillance

The WSN can be used to replace the guards and sentries around defensive perimeters, keeping soldiers out of harm's way, to locate and identify troops, vehicles and targets for potential attack. So privacy of the sensor data is always critical and it should be preserved during data aggregation.

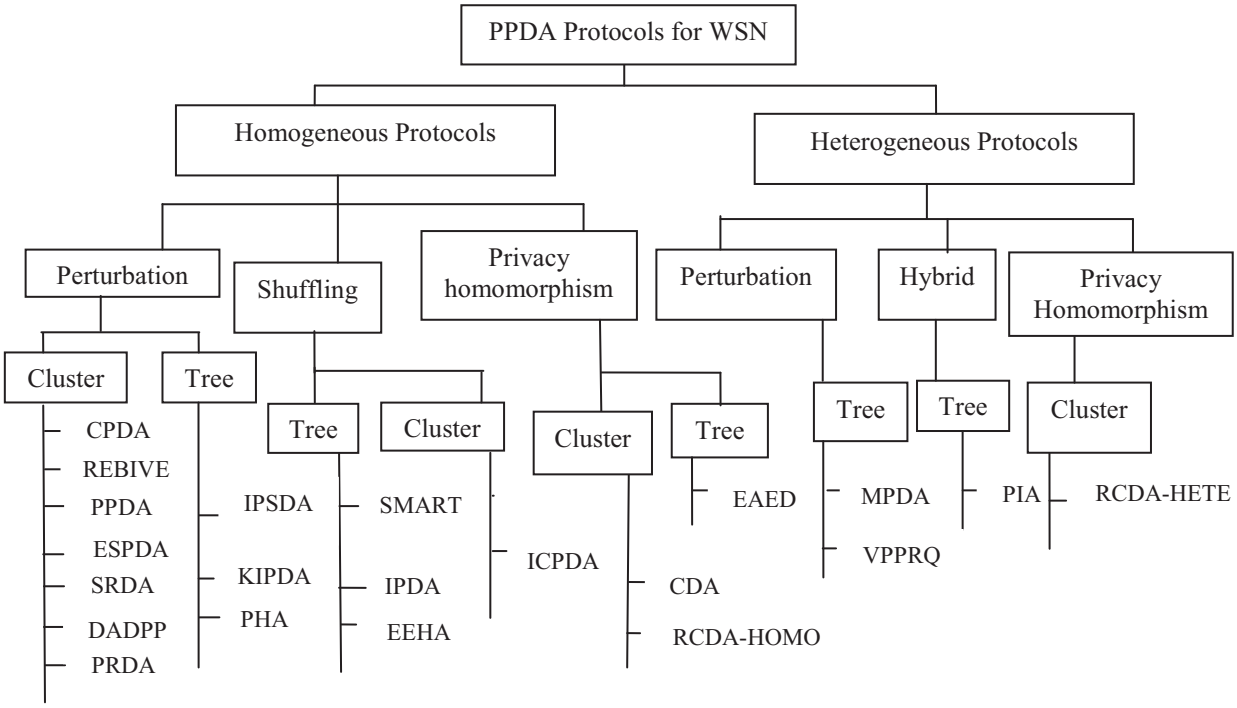


Figure 3. Classification of existing PPDA protocols for Wireless Sensor Networks.

2.3. Private Households

Wireless sensors could be placed in houses to collect statistics about water, gas and electricity consumption within a large neighborhood. The aggregated population statistics is helpful for individuals, businessmen and government agencies to plan the resources. However, individual readings reveal the daily activities of a household i.e., if the water consumed is nil, it means that there was no one in the house.

3. Classification of PPDA Protocols

Figure 3 shows the classification of the existing PPDA protocols in WSNs. Based on the type of nodes in the sensor networks, privacy preserving techniques are divided into two types: homogeneous protocols and heterogeneous protocols. In homogeneous protocols, all the nodes in the network have the same resources and the aggregator can perform sensing, aggregating and forward the aggregated result to the sink. All sensor nodes can play the role of aggregators. In heterogeneous protocols, more than one type of sensor node exist and the aggregator is considered as a special node. i.e., the aggregator plays the role for aggregation and forwarding the aggregated value to the sink, but is not used for sensing. The PPDA protocols are further divided by the type of network topology used; protocols are grouped into three classes such as cluster, tree and hybrid. The secure aggregation protocols are divided into two types; end to end aggregation and hop by hop aggregation.

In an end to end aggregation, a secure channel is established between a sensor node and the BS. It achieves end to end confidentiality. Here the aggregation is performed on the cipher text In hop by hop aggregation, the aggregator decrypts the data it receives from leaf nodes with the shared key between them, and then aggregates it, again encrypts the data and send to the upper aggregator or sink. This increases the communication overhead and doesn't guarantee the privacy at the aggregator.

3.1. Homogeneous Protocols

The homogeneous protocols are divided into different types; these are the perturbation, shuffling and privacy homomorphism. The homogeneous protocols are summarized below.

3.1.1. Perturbation

In this perturbation, each node customizes its sensed data using the encryption key and public or private seed generated by randomization technique [6, 7] to conceal the data from other nodes as well as from adversaries.

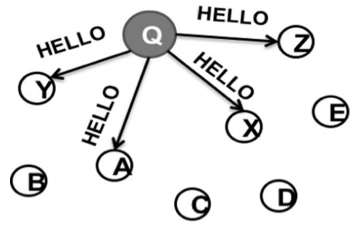
1. CPDA
   The CPDA (Cluster Based Private Data Aggregation) [8] ensures privacy of data in the cluster topology. The CPDA consists of three steps: cluster formation, calculate aggregate within the clusters and cluster data aggregation. At first, the nodes in the sensor network are grouped to form clusters. Figure 4 shows the procedure for cluster formation.

   During the aggregation in a cluster, every sensor node in a cluster customizes its private data into a polynomial of order $k-1$ using shared seeds (non private) and random numbers (private). Here k *no*. of nodes in a cluster encrypt the data and send it to all other nodes inside the cluster using shared key between them. After receiving them, each node sums up the received polynomials using the additive property of polynomials and sends its result to their CH. The aggregated data received from the cluster members are reduced by computing the inverse of a $k * k$ matrix in CH. BS receives the aggregated result from CH through the aggregation tree constructed using the TAG [9] Protocol.

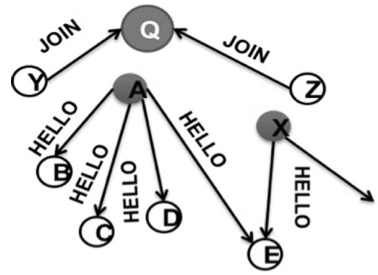
   Figure 5 shows the message exchange within a cluster. The cluster consists of three sensor nodes A, B, C and A is the cluster leader.

   The CPDA suffers from high complex computation and communication overhead, especially when the network grows. The CPDA can tolerate the collision up to a certain threshold. i.e., number of nodes in a cluster – 2. If the number of colliding nodes exceeds the threshold value, then CPDA will reveal the private data of nodes.

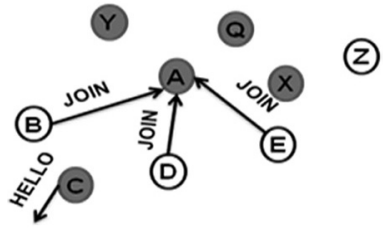
2. REBIVE
   The REBIVE (REliaBle prIVate data aggrEgation scheme) [10] is a cluster based additive data aggregation scheme to achieve privacy and perfect accuracy during the data aggregation. It has three steps: query launch



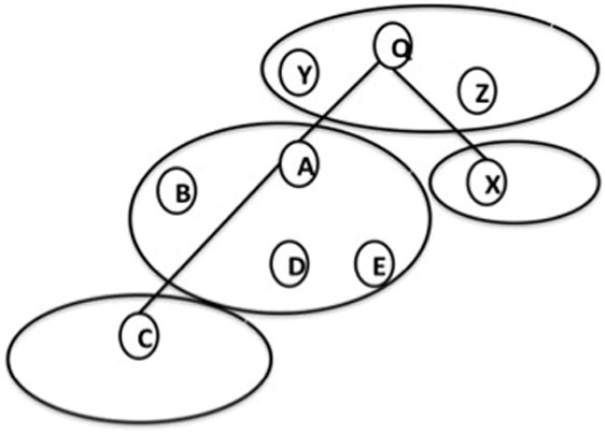
(a) Query server Q triggers a query by sending a HELLO message. A recipient of HELLO message elects itself as a cluster leader with probability pc randomly.



(b) A and X become cluster leaders, so they broadcast the HELLO message to their neighbors.

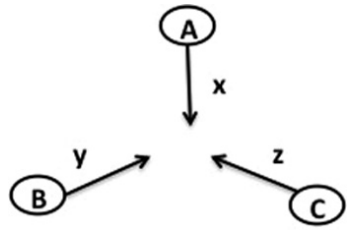


(c) Node E receives multiple HELLO messages. Then E randomly selects one to join.

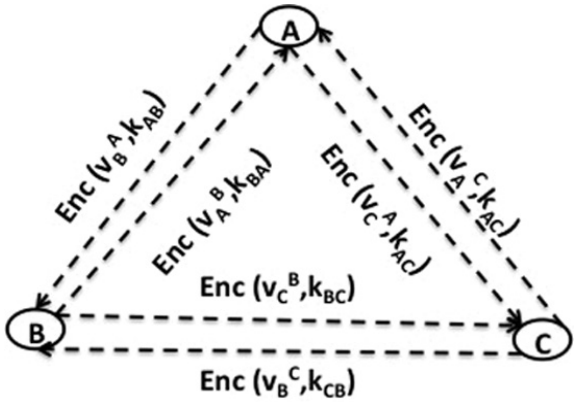


(d) Several clusters have been constructed and the aggregation tree of cluster leader is formed.

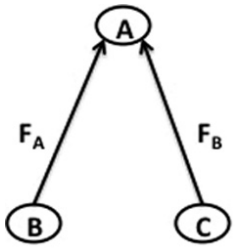
Figure 4. Procedure for cluster formation.



(a) Public seed broadcasting.



(b) Customized data encryption and sending.

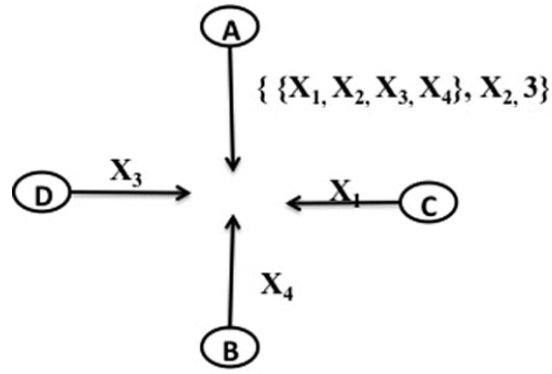


(c) Assembled information broadcasting.

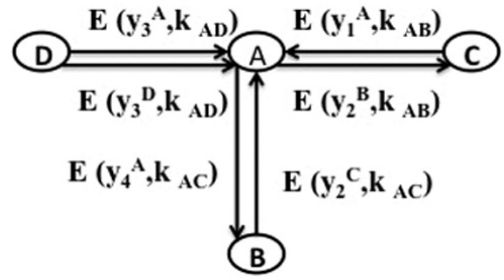
Figure 5. Message exchange within a cluster in the CPDA.

& cluster formation, data aggregation within clusters and post aggregation at query server. First the nodes in the network form clusters based on the procedure in CPDA technique. All nodes in a cluster share their sensitive data into n share using the (k, n) threshold scheme [11]. During data aggregation, the CH broadcast a message $\langle \{x_1, x_2 \ldots x_n\}, x_i, k\rangle$ into the cluster. Where n is the cluster size and k is the total number of data shares needed for reconstruction $(3 \leq k \leq n)$, x_i specifies the values at which data shares will be calculated. After receiving the message, every node in a cluster announces x value from the received message and forms a polynomial of degree $k-1$, then encrypts the customized data and sends it to all other cluster nodes using the shared key between them. During the reporting phase, each node sums up the shares received from other nodes at x_i and sends to BS through a CH. After receiving the partially aggregated data from the CH's, BS aggregates the data Because of the transmission of shares the communication overhead is increased.

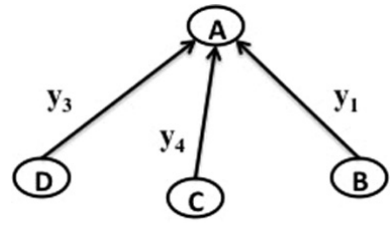
Figure 6 shows the example of data aggregation in REBIVE. The node A forms a polynomial equation as $Y_A = f_A(X) = a_0 + a_1 x + a_2 x^2$, where a_0 represents the sensing data of node A and a_1, a_2 are random coefficient known only to A.



(a) Nodes broadcasting their x_i values.



(b) Messages in and out of node A at "sharing y_i's".



(c) Nodes reporting y_i's to cluster head A.

Figure 6. Illustration of data aggregation in REBIVE.

3. PPDA
The PPDA (Privacy Preserving Data Aggregation) [12] is an additive aggregation technique that preserves the privacy of the data based on a key management scheme and randomized data perturbation technique. The computation is based on Secure Multi party Computation (SMC) [13] and recovery of the aggregated data through modular arithmetic. PPDA is applicable only for the SUM aggregation function. The PPDA has two parts: secure key management and privacy preservation.

Key mapping is stored in the server before distribution. This could be done in two ways, aggregator to source and source to source key exchange. Every sensor node in a network is assigned with K number of keys, in that k keys are used for source to source communication and $K-k$ keys are used for an aggregator to source communication. The source node generates a random number (R_n) between 1 and $K-k$ and sends it to server. Each node encrypts its message by $(R_n)^{\text{th}}$ key from the key bank stored in a node. Since randomization of R_n^{th} key is different for every sensor node, this avoids the guessing attack. In a source to source key establishment, each node separately permutes the key bank of k keys, records it and passes the permute function to every other node through server using the same procedure as in source to aggregator key establishment.

During the data transmission, the server initiates a source node, hides its data by adding a random number to its own data and does the modular operation with the sum of private data. i.e., $R_1 = (r_1 + x_1) \mod X$, where r_1 is the random number added by node 1 and X represents the sum of private data. The node finds its neighbor to which it is connected and passes this information to the server, it keeps the details of already participated nodes. If the neighbor node is a non participating node, the server will inform the source node and pass the customized data to the neighbor node. Upon receiving, the neighbor node adds its sensed data to the customized data and performs a modular operation with the sum of private data. i.e., $R_N = (R_{N-1} + X_{N-1}) \mod X$. This procedure goes on, the server finds that all the nodes have participated, sends the data it has received from the last node to the first source node to compute the aggregated data. The first node subtracts the random number it has added to the sum and performs a modular operation with the sum of private data. i.e., $X = (R_N + R_1) \mod X$, and sends it back to the server.

In PPDA, source to source communication is via the server, there is no direct link between one source node to the other source node, which increases the communication overhead. It achieves privacy against the BS, each time the initiator node calculates the SUM and compares it with private data of the initiator. If not equal, the initiator sends a message to the server.

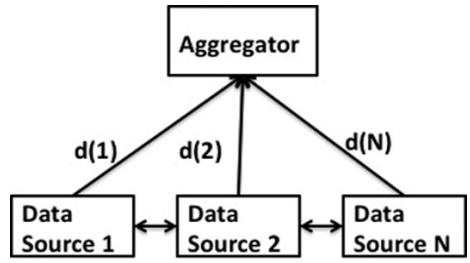


Figure 7. Privacy preserved data aggregation system model.

4. ESPDA
The ESPDA (Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks) [14] improves the energy efficiency and privacy preservation in WSNs. In ESPDA, each node hides its data by converting the actual data into pattern code and sends the pattern code instead of actual data. The pattern codes are generated using the pattern seeds broadcasted by the sink using the common session key. Whenever the sensor node receives the pattern seed broadcasted from the CH, it creates a lookup table containing critical value, interval and threshold. In a multidimensional sensor network, the data sensed from the environment is represented as parameters. Whenever the node senses data from the environment, its parameters are compared with the interval in the lookup table and the corresponding critical value is assigned to each parameter. Then

concatenate all the critical value, it will generate the pattern code.

For the secure data aggregation, the ESPDA uses an energy efficient key distribution scheme. Each node in a network is assigned with a session key, unique ID and a node specific key during the manufacturing phase and the BS is assigned with common secret key, session key, secret key and ID pairs of all nodes in the network. If a node wants to send data to the sink, first request the CH to get the session key transmitted by the sink. Upon receiving, every node calculates encryption key by XOR ing the session key with the node specific key and encrypts its data and sends the data, MAC, ID, time stamp to the CH using the NOVSF-BH technique [15,16]. Whenever the CH receives the data from cluster members, it compares the pattern code with one another, if redundancy exists, CH requests to send original data of the node having unique pattern code and send acknowledgement to other nodes for discarding the data. After which the sensor node switches to sleep mode to conserve power.

The CH performs aggregation on the received data, appends its ID and transmits to the BS. After receiving the aggregated data from the cluster heads, it determines the secret key of the sender node by using the sender and cluster head ID. To decrypt, the aggregated data is the session key is XOR'd with the nodes' specific secret key. The BS uses the time stamp and the session key to check the data freshness, MAC is used to verify data integrity. If the data is altered/replayed, the BS discards the data and requests the sender node to transmit the data again. The MAC is generated by using the memory efficient CBC-MAC protocol.

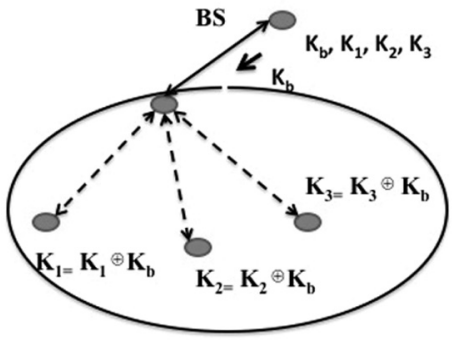


Figure 8. A cluster of sensors and BS computing encryption keys, $k_{i,b}$.

The ESPDA avoids the redundant data transmission to the CH and uses the sleep active protocol to improve the energy efficiency. It provides better communication security by using NOVSF block hopping technique. It does not transmit the symmetric key between CH and sensor node, hence increases the security.

5. SRDA
In SRDA (Secure Reference-Based Data Aggregation) [18], sensor nodes send differential data instead of raw data by comparing raw data with reference data, thus reducing the number of bits transmitted from sensor node to CH, which improves the energy. SRDA uses an algorithm with security margin as adjustable parameter; it is calculated based on the number of hops from the BS. First the leaf node sends its raw data to CH for a session; CH creates a reference entry for that node. A sensor node sends differential data to CH for subsequent readings. Finally when the session ends for a sensor node, CH removes the reference entry for the node from the CH. This method is independent of clustering scheme which can be applied at any level. If the reference value is greater than differential value, the efficiency of the scheme will increase.

6. DADPP
The DADPP (Data Aggregation Different Privacy levels Protection) [19] provides different level of data aggregation privacy based on different node numbers for pretreating the data. The different level of privacy is achieved based on Shao *et al* [20] scheme and achieving privacy based on CPDA protocol. In DADPP, a hierarchical wireless sensor network is constructed in such way that the sensor nodes form several clusters, each of which has a CH below the BS. According to the level of privacy, all nodes within a same cluster are partitioned into different groups with the same privacy level. Data are pretreated only in the same group and the size defines privacy levels of a group. The lowest privacy levels consist of partitioned groups with three nodes. The upper privacy levels contain two partitioned groups with

four sensor nodes. If all sensor nodes of a cluster belong to a single group, it becomes the highest privacy level. The data aggregation is same as that of the CPDA [8]. First the data are aggregated on the group and sent to CH; it aggregates and sends to the BS. The DADPP reduces the traffic by partitioning a cluster with n sensor nodes into multiple in-networks with pretreatment of groups according to the desired privacy levels. It has high communication and computation overhead, it is increased based on privacy levels.

7. PRDA
The PRDA (Privacy-preserving Robust Data Aggregation) [21] is used to provide privacy preservation and robust data aggregation in WSN's. It consists of three steps: local cluster formation, twin key establishment and data aggregation. Clusters are formed in the network using the clustering method in CPDA [8]. Each cluster forms a different Hamiltonian circuit [22] and every node in a cluster shares a twin key with other nodes in a cluster. Each node contains pre-deployed key ring of k symmetric keys chosen from a larger key pool of size P stored in node's memory. Each node anonymously checks which of its k keys are shared with other nodes, CH computes the aggregate of the cluster together with a twin key announcement protocol-each node should be aware of the twin key it posses are used by another node in the cluster. During data aggregation, the aggregate is routed twice along the Hamiltonian circuit and each node adds its sensed value and the shadow value corresponding to twin key for encryption, after which the aggregate is sent to the BS for further processing. PRDA has limited applications because the key pool is only known to the node manufacturer. It has high communication overhead because the aggregate is routed twice the Hamiltonian circuit and all nodes participated during the routing of individual cluster aggregate to the BS. It also decreases the lifetime of sensor nodes because the nodes are idle, until BS receives the aggregated value from CH.

8. IPSDA
The IPSDA (Integrity Protecting Sensitive Data Aggregation) [23] overcomes the high communication and computational overhead of iPDA [27]. It uses the additive property of complex numbers to achieve privacy and integrity protecting data aggregation from adversaries or trusted participating nodes in the network. In the two parts of a complex number, real number is used for privacy preservation and the imaginary part is used for integrity checking. Every node in the network shares two keys: a pair wise secret key with a master device and a symmetric pair wise key with those sensor nodes lying on the aggregation tree. It consists of four steps: create customized data from the data of the source nodes, apply additive properties of complex numbers to get the sum value of customized data, compute the aggregation result at the sink and extract actual SUM of sensor data to check data integrity at the sink.

   Each node is assigned with a private real number and an imaginary number by the node manufacture. After receiving a query from the BS, every node conceals its data by adding a private real number to it and then adjoins an imaginary unit to it. Afterwards, the leaf node encrypts its customized data and sends it to its parent using the symmetric key between them. After receiving customized data from the child nodes, the intermediate node decrypts and aggregates with the customized data using the additive property of complex numbers, encrypts it and sends to upper aggregator or sink. After receiving the aggregated data at the BS, it first decrypts the value and separates the real part and imaginary part of it. The BS keeps real seed and the imaginary seed of all the nodes in the network, it subtracts the sum of real seed from the real part separated from the aggregated data, gives the actual sum and the integrity is verified by comparing the sum of imaginary seed and the imaginary part separated from the aggregated data. Due to this complex process IPSDA suffers from high memory consumption.

9. KIPDA
The KIPDA (K-Indistinguishable Privacy-preserving Data Aggregation) [24] is a non cryptographic privacy preserving technique that protects the privacy of data in a tree topology by adding a set of camouflaged values to the message set. It is well suited for nonlinear functions such as MAX/MIN. A message set consists of single real value and

camouflaged values (both restricted and unrestricted). This is the form of k-indistinguishability from $k-1$ other values. The restricted camouflage values should be less than, or equal to the sensing data, if the MAX aggregation function is used and greater than, or equal to the sensing data, if the MIN aggregation function is used. The unrestricted camouflage values are either greater than, or less than, or equal to the sensing data. The BS keeps a Global Secret Set (GSS), it contains the possible locations for the final aggregate, it also assigns a Node Secret Set (NSS^i) for all nodes in the network, which represents the secret information about GSS, it contains the union of index set of the restricted camouflage values and real values. The NSS_T^i keeps the index set of the real values of all nodes in the message set. The aggregation trees are constructed based on the TAG data aggregation protocol. There are four phases: predistribution, reporting, aggregating and BS processing.

The BS chooses a GSS and determines the NSS_T^i and NSS^i for every node. Every node determines the values for a message set. During data aggregation, each leaf node sends its message set to its parent and the intermediate node computes the aggregate using MIN/MAX aggregation function among its child node value and its own value; pass the aggregate to the next hop. The BS will compute the aggregated data from the values in the index given by GSS in the message set. The bandwidth consumption per node in KIPDA is based on the number of values in a message set. Its energy consumption is more than end to end encryption and less than a hop by hop encryption. It does not provide privacy against BS.

10. PHA
The PHA (Perturbed Histogram based Aggregation) [25] scheme provides privacy for the queries targeted at the special node or sensor data distribution for both individual data and aggregated data. It is highly efficient and resilient to any number of node collusions, supports all types of aggregation function. Each sensor node (all having unique ID) is preloaded with a unique secret number by the sink and it is known only to the sink and the corresponding sensor node. Each node is preloaded with a secure one way hash function, this maps a bit string to a value between to $N-1$.The sink knows the unique number and unique ID of all currently alive sensor nodes.

The sink launches the query message which is a unique noun; this prevents adversaries from launching replay attack. An answer for the particular query is computed based on the histogram. Before transmission, each node first uses an integer range to replace the data. The aggregator plots the histogram for data collected and then estimates the aggregate. Perturbation technique is used to hide the actual individual reading and the actual aggregate results sent by sensor node. If some sensor nodes fail to reply, this will be detected by some upstream nodes on the tree and the IDs of failed nodes will be sent to the sink.

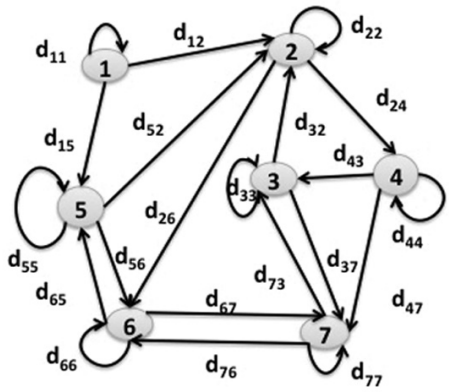
The disadvantages are: first, final aggregate is an approximate value of the sensor data, hence accuracy level is reduced. Second, data are replaced by an integer range, which requires a large size payload. Third, the bandwidth consumption is high as payload increases. Fourth, it consumes large amount of memory because the interval ranges are stored to replace the real data.

3.1.2. Shuffling

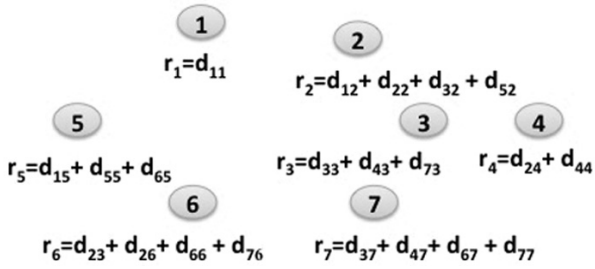
In shuffling, each node slices its data into k numbers. One is kept on the node itself and the remaining $k-1$ slices are encrypted and sent to the $k-1$ neighbors.

1. SMART
The SMART (Slice Mix AggRegate) [26] is an additive slicing based aggregation scheme which guarantees the privacy of individual sensor data in a tree topology. It consists of three steps: Slicing, mixing and aggregation. The aggregation tree is formed using TAG (Tiny Aggregate) protocol. During slicing, every node in the network slices its data into k number of pieces, one piece is kept on the node itself and the remaining $k-1$ slices are encrypted and sent to the $k-1$ neighbor nodes within h hops (for a dense sensor network $h=1$) using the shared key between neighbor nodes. Whenever a node receives an encrypted slice, it

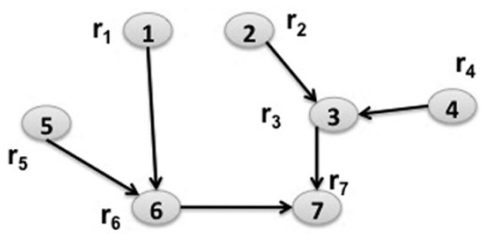
decrypts the data using its shared key with the sender. Upon receiving the first slice, the node waits for a certain time, which guarantees that all slices of this round of aggregation are received, after which each node sums up all the received slices with one of the data slices kept on the node. It encrypts and sends the sum to its parent in the aggregation tree using the shared key. The aggregated data traverses level by level until it reaches BS. The SMART technique reduces the computational overhead, but the communication overhead increases with the increase in the number of transmissions in the network. It can tolerate the collusion up to a threshold. i.e., sum of out degree and in degree minus one.



(a) Slicing ($K = 3, h = 1$) d_{ij} is encrypted and transmitted from node i node j. The d_{ii} is the data piece kept at node i.



(b) Mixing: Each node i decrypts all data pieces received and sums them up including the one kept at itself (d_{ii}) as r_i.



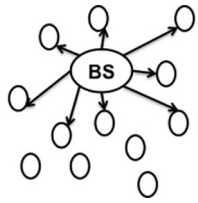
(c) Aggregation (No encryption is needed).

Figure 9. The three steps in the SMART technique.

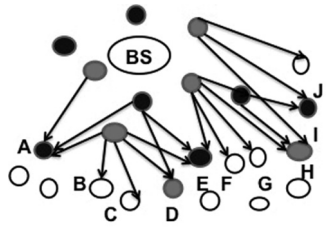
2. IPDA
The iPDA (integrity protecting Private Data Aggregation) [27] addresses the privacy and integrity of the data. In iPDA, data privacy is achieved through slicing and assembling technique and the integrity is achieved through redundancy by constructing two disjoint aggregation trees. Since each node belongs to a single aggregation tree, all malicious nodes can only pollute the aggregation result on aggregation tree it belongs to. Hence by comparing the results from two different aggregation trees, the BS can verify the integrity of the aggregation result. The iPDA consists of three phases: disjoint aggregation tree construction, privacy preserving data report and integrity protecting data aggregation.

During slicing, each node randomly selects l neighbor nodes including itself from the two aggregation trees, a node from an aggregation tree select $l-1$ other neighbor nodes from the same aggregation tree and l other neighbor nodes from the other aggregation tree. Each node slices its data into l pieces, one is kept on the node itself and the remaining $l-1$ pieces are encrypted and sent to $l-1$ neighbor nodes in the same aggregation tree using the shared key with neighbors. Independently, the node slices its data into l pieces and sends it to l neighbor nodes in the different aggregation trees, each node takes $2l-1$ transmission during slicing. After receiving all the encrypted slices corresponding to it, each node decrypts it using the key shared with the neighbors. Then the aggregator sums up the data slices received from the neighbors, the aggregated data received from the child nodes in the aggregation tree it belongs to and one of the data slices is kept in the node, then encrypted and forwarded to its parent within the same aggregation tree. After receiving the aggregation result from two disjoint aggregation trees, compare the values with each other. If too much difference is obtained due to the pollution attack,

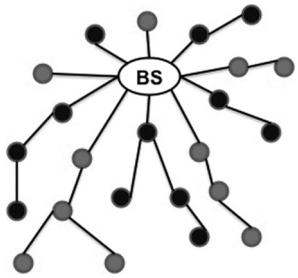
reject it else accept even if small difference obtained because of data loss. Figure 11 shows the steps in privacy preserving data report.



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

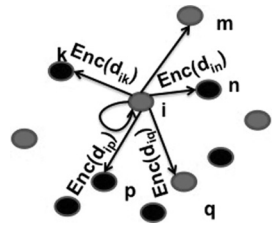


(b) Nodes A, D, E, H, I receive HELLO messages from both black and red aggregators, then they randomly select their roles. ie, black aggregator and red aggregator. Nodes B, C, F, G, J only receive HELLO from red aggregators, so they should wait until they receive HELLO message from both black and red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

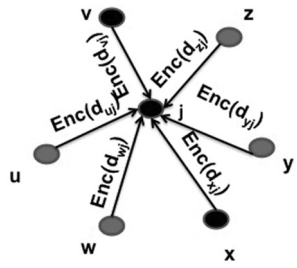


(c) As the disjoint aggregation tree construction procedure continues, two disjoint aggregation trees rooted at the BS are formed. Blue aggregator and red aggregator interleave.

Figure 10. Procedure for constructing disjoint aggregation trees.



(a) Slicing step by node *i*.



(b) Assembling at node *j*.

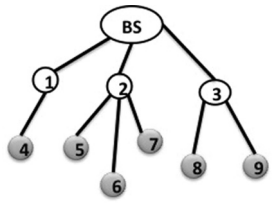
Figure 11. Privacy preserving data report.

The advantages of iPDA are the following; first it can build on any key management scheme and can detect the pollution attack by comparing the aggregation result along the two disjoint aggregation trees. It can tolerate the collusion up to a certain threshold. Communication overhead is increased by increasing the number of slices to achieve privacy preservation. When the transmission increases, collision and packet loss occur, which decreases accuracy.

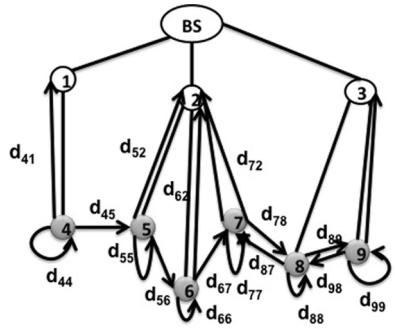
3. EEHA
The EEHA (Energy Efficient and High Accuracy secure Data Aggregation) [28] scheme provides an energy efficient and high accuracy secure additive data aggregation without releasing individual sensor reading and introducing extra overhead on the battery limited sensors. The EEHA consists of four steps: aggregation tree construction, slicing, mixing and aggregation.

The aggregation tree is constructed using the TAG protocol, represents the path from sensor node to BS. It reduces communication overhead of SMART technique by slicing only in the leaf node, which reduces message transmission and degree of collision, without compromising accuracy and energy efficiency. Each intermediate node decrypts

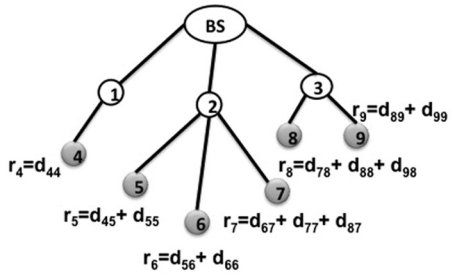
the aggregated data received from the child nodes and sums up the aggregate, forwards it to its parent in the aggregation tree. After receiving all the partially aggregated data from the aggregator, the BS sums up and generates the final aggregation result.



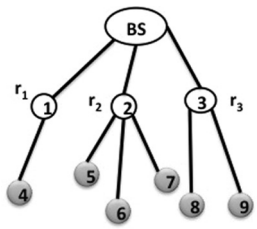
(a) Aggregation tree construction.



(b) Slicing.



(c) Mixing.



(d) Aggregation.

Figure 12. Four operations in EEHA.

The EEHA can only be used for additive aggregation. It uses the hop by hop aggregation, which increases computational overhead, thereby energy consumption is increased.

4. ICPDA
The ICPDA (Integrity-protecting CPDA) [29] protocol is the extension of CPDA to detect both collaborative and individual attackers. Here all nodes can participate in the calculation of intermediate aggregation result and can monitor the behavior of the CH. If a cluster node detects that the CH pollutes the result, it reports the malicious behavior to the BS. It has four phases: cluster formation, calculation of the aggregate result within clusters, cluster data aggregation and misbehavior reporting. The nodes in the network are randomly grouped to form clusters using the procedure in CPDA.

The calculation of aggregate result can be done in two ways: algebraic properties of polynomials or Secure Multi party Calculation (SMC). Algebraic properties of polynomials are same as CPDA. SMC uses the slicing and assembling techniques, both ways it broadcasts the assembled data of each node in a cluster. Here every node can calculate the intermediate aggregation result.

When the CH receives aggregation results from its downstream nodes, the CH should make these results accessible to all its cluster members by broadcasting the message within the cluster along with its ID. The aggregated result is sent to the BS with signature. If a node reports the malicious behavior to BS, it contains reporter ID, attacker ID and value. The reporters should select unicast routes to send the report to avoid the malicious aggregator in the route. After BS receives the report, it will decide whether to accept the aggregation result or not.

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

3.1.3. Privacy Homomorphism

In privacy homomorphism, the arithmetic operations are done on the encrypted data without decryption which reduces the energy consumption.

1. CDA
   The CDA (Concealed Data Aggregation) [30] scheme uses the homomorphic encryption scheme that allows to perform efficient aggregation on encrypted data without decryption. The CDA uses the Domingo Ferrer's approach [31] to conceal the privacy of data. Every leaf node in a network shares a common secret key with the BS. In DF approach, each leaf node splits its data into d parts ($d \geq 2$), encrypts it using the symmetric key shared with the BS and sends it to the aggregator. The aggregator performs aggregation on encrypted data and sends it to BS, it decrypts the aggregated data by using the symmetric key shared with the source node. The disadvantages of the CDA are the following: if an adversary can get the encryption key and access to the encrypted data by compromising any one sensor node, it will reveal the private data of every node. The CDA has the following advantages: first it reduces the computational overhead of hop by hop encryption, thereby improving the energy efficiency. Secondly, it prevents the passive adversaries that eavesdrop the communication link by encrypting the data.

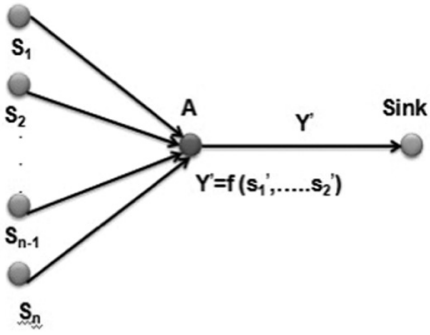


Figure 13. Data aggregation in CDA.

2. EAED
   The EAED (Efficient Aggregation of Encrypted Data in wireless sensor network) [32] additive homomorphic scheme provides a secure communication channel between the source node to the BS. Each sensor node in a network shares a separate key with the BS, the leaf node encrypts its sensed data and forwards the aggregated data to upper aggregator or BS, where it will decrypt the aggregated data by using the sum of shared keys of every participant node. The encryption, decryption and aggregation operations are given below.

$$\begin{aligned} \textit{Encryption } c &= (m+k) \mod M \\ \textit{Decryption } m &= (c-k) \mod M \\ \textit{Aggregation } C_{12} &= (c_1+c_2) \mod M \end{aligned}$$

   Where M is a large integer, i.e., $M = n \times t$, n represents the total number of nodes in the network and the t represents the maximum value which appears in the sensed data. If M is smaller than the sum of all sample values and encryption keys, the sink fails to reproduce the real sum, instead it produces a smaller number than M.

   It minimizes the bit transmission between sensor nodes and the sink by preserving end to end privacy. It is not scalable since BS has to store keys of every node in the network, to decrypt the aggregated value received at the sink. It transmits the ID and corresponding secret key of node increasing communication cost and vulnerable to malicious modification of data by adding the natural numbers to cipher the text which does not provide integrity checking mechanism.

3. RCDA
   The RCDA (Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks) [33] uses elliptic curve based additive privacy homomorphism algorithm for achieving end to end confidentiality along with in-network data aggregation. The problem in other privacy homomorphism based techniques is that BS can only retrieve aggregated result rather than individual data. Thus the BS is limited in using aggregation function (such as MAX, MIN) and it cannot verify integrity and authenticity of all sensing data. The RCDA provides a mechanism to recover all sensing data from aggregated result reached at the BS, which overcomes the limitations of the BS in other privacy homomorphism

techniques. The RCDA provides two solutions for homogeneous wireless sensor network (RCDA-HOMO) and other for heterogeneous wireless sensor network (RCDA-HETE)

- RCDA-HOMO

This technique works through four phases: setup, encrypt-sign, aggregate and verify. In the setup phase, BS prepares the entire system by generating keys (public key, private key) for itself and sensor nodes. Further, BS loads the private key of sensor node, hash function along with its public key to all sensor nodes. In the encrypt-sign phase, every sensor node generates cipher text and signature for its message, sends it to CH as a pair. The cipher text is generated using the mapped value corresponding to the message and BS public key. The signature is generated using original message and its private key. In the aggregate phase, CH aggregates the incoming cipher text, signature from the sensor node is separately passed to BS. In the final verification phase, the BS confirms the integrity of all sensed data using the recovered data and public key of sensors.

- RCDA-HETE

This technique helps the BS to perform any aggregation function on sensed data and allows the BS to confirm data integrity in heterogeneous wireless sensor networks. It consists of two types of sensor nodes, H and L sensors. The protocol contains five procedures: setup, intra cluster encrypt, inter cluster encrypt, aggregate and verify. In the first phase the BS generates public key, private key for every H sensor and for itself. In addition, the BS loads its public key, private key of H sensors, necessary aggregation to all H sensors and public key of BS to all L sensors. This phase prepares the entire system for other phases. In the second phase, each sensor node encrypts its data using the key shared with the CH and passes it along with MAC of the data to the CH, to decrypt and check the integrity using MAC. In the third phase, the CH performs aggregation, encrypts the aggregated data using BS public key, after which it encodes and maps the aggregated result to elliptic curve point. The CH also generates digital signature for individual aggregated data using its private key and passes it along with cipher text as a pair to the upper H sensors. In the fourth phase, the cipher text, signature pairs coming from all low level H sensors are aggregated and sent to upper H sensors. This procedure is continued until aggregate result reaches the BS. In the last phase, the BS checks the integrity of all individual aggregated results, using the public key of H sensors.

3.2. Heterogeneous Protocols

The heterogeneous sensor nodes in the WSN are different in terms of resources.

3.2.1. Perturbation

This technique is same as perturbation in homogeneous networks. The perturbation techniques in heterogeneous WSNs are explained below.

1. MPDA
   The MPDA (Multidimensional Privacy-preserving Data Aggregation scheme for Wireless Sensor Networks) [34] can combine more than one data into a single result which improves the energy efficiency and privacy preservation. Figure 14 shows the multidimensional aggregation in WSN's. The MPDA scheme uses a super increasing sequence and a perturbation technique for multidimensional additive aggregation. Every sensor node shares a symmetric key with its neighbor aggregator. The pairing operation is more time consuming than key distribution scheme [35], however the neighbor key establishment saves the storage space. Whenever a sensor node receives a query from the sink, the sensor node customizes its data using the private key of the node and a sequence received from sink and generates a hash, sends to aggregator along with timestamp, finally aggregator checks the timestamp for transmission delay. If it is within the period, aggregator node verifies the message by calculating the hash using the neighbor key, if it is correct, accept the encrypted message, or else reject it. After receiving k valid encrypted data from the sensor nodes, the aggregator node decrypts to calculate the aggregated result, encrypts it along with public key of BS and sends to the sink; private

key is known only to BS to decrypt the data and compares the timestamp with no delay. If accept, then compute hash, compare the computed hash with the hash received from the aggregator if it is the same accept it, else reject.

The MPDA scheme preserves privacy against passive attack and node compromise attack. A sensor node compromise does not disclose the data sensed by other nodes or aggregator nodes.

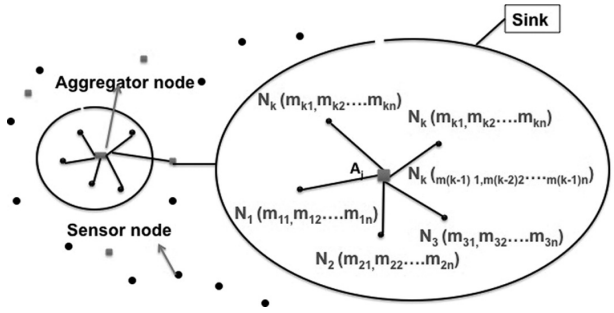


Figure 14. Multidimensional Privacy-preserving data aggregation.

2. VPPRQ
   In VPPRQ (Verifiable Privacy Preserving Range Query in two tiered sensor networks) [37] provide a privacy preserving storage scheme based on the bucketing scheme [38, 39]. Figure 15 shows the two tiered sensor networks, consist of three types of sensor nodes, they are powerful sink, regular sensor nodes and storage node with large storage capacity. 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 (interval between two submissions). Every sensor node has a unique ID and a secret key shared with the sink, compromising of one sensor node does not affect the security of other sensor data. Before sending data to a storage node, encrypt the data using shared key between the sensor node and the sink, attach a tag to the encrypted data based on which bucket the data falls into. Upon receiving the encrypted data from the source node, the storage node divides the data into multiple buckets and each bucket is assigned a tag. The sensors and the sink have agreed on the same bucket partition. Whenever 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 returns 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, and obtains the real sensor data. The major predicament will be the transmission delay. i.e., the time required to send the data from the source to the sink, however this cannot be implemented for applications like event detection and tracking objects which require immediate response. This scheme focuses on the problem when storage node and sensor node are compromised. To prevent the storage node from dropping the data, an encoding number is generated on each sensor if no data in a query range are collected on the sensor. The VPPRQ technique prevents the false reply.

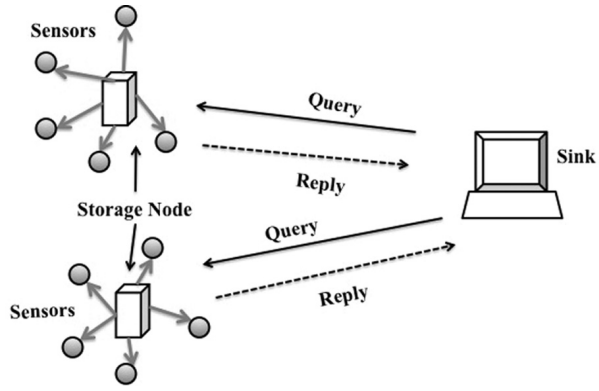


Figure 15. Two tier model.

3.2.2. Hybrid

It uses more than one privacy preserving technique to do the privacy preservation.

1. PIA
   The PIA (Privacy Preserving Integrity Assured data aggregation) [40] addresses the integrity assured data aggregation with efficiency and privacy as joint objective. Based on who will perform the data aggregation, there are two types of Integrity Verification (IV) in WSNs, centralized and distributed. In both types, the integrity of aggregate is verified by recomputing the aggregation

function on the raw data. In centralized IV, after receiving data from sensor node, the server verifies the aggregate. In distributed IV, the sensor node recomputes the aggregation function using the data of other sensors. The results obtained from sensor nodes are the same, aggregation is considered secure. It distributes the communication throughput in the network. The PIA proposed four symmetric key solutions for a single aggregator model.

In the first solution, homomorphic encryption is used to hide the data and MAC [41] to construct an authenticated encryption scheme for preserving the privacy of individual nodes' data. It only supports aggregation functions such as average and standard deviation. A trusted third party generates and distributes an encryption key (K_i^e) and MAC key (K_i^m) to every sensor node. Every leaf sensor node sends the data and MAC to the aggregator. The aggregator calculates the data aggregate (y) and MAC aggregate (σ) using the XOR function and forwards the data aggregate, MAC aggregate and all the cipher text to a user. The user gets the data aggregate by decrypting the sum of cipher text using master key which is the sum of all encryption keys. Compare the data aggregate and MAC to ensure the aggregator has not changed the data. The first solution uses the centralized IV.

The second solution uses the Order Preserving Encryption Scheme (OPES) [42] to preserve the privacy of data distribution. It is used to verify the integrity of comparison based aggregation function. Sensors encrypt the data by using OPES with master secret key shared by all sensors. The aggregator calculates the aggregate of the encrypted data set and sends it to the user. The user decrypts the data set and calculates the plain text of the aggregate using centralized IV.

The third solution uses a Secure Hierarchical In-networking Aggregation (SHIA) [43] scheme for adapting distributed IV. This scheme supports any aggregation function because the sensor nodes have access to all raw data. The communication overhead of each node is $O(N)$, where N is the total number of sensor nodes in a network. This scheme preserves privacy without using any additional privacy preservation mechanism such as encryption. A commitment is constructed during the aggregation. After calculating the aggregate, each sensor node independently reconstructs the commit tree and ensures that the data is not modified or discards the contribution of the node by any adversary.

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

4. Comparative Study of Different PPDA Protocol

1. *Network Topology (NT):* It determines which topology is used in the data aggregation. Either tree or cluster.
2. *Communication Cost (CMC):* 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.
3. *Computational Cost (CPC):* This is the processing overhead 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, encryption or decryption.
4. *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.
5. *Accuracy (AUC):* The aggregated data received at the sink divided by the sum of actual data gives the accuracy level. The Values are Low, Medium and High.

Topology	Protocols	CMC	CPC	EC	AUC	DLR	DLY	PLS	MC	Aggregation functions
Cluster	CPDA	M	H	M	H	N	M	M	M	Additive aggregation functions
	REBIVE	M	H	M	H	Y	M	M	M	Additive aggregation functions
	PPDA	H	M	M	L	N	L	S	H	SUM
	ESPDA	M	M	M	L	N	H	G	H	Additive aggregation functions
	SRDA	M	M	M	L	N	M	S	L	SUM
	DADPP	H	H	H	H	Y	M	M	M	SUM
	PRDA	H	H	H	H	Y	H	M	M	SUM
	ICPDA	M	L	M	H	N	M	M	M	Additive aggregation functions
	CDA	L	L	L	H	Y	L	S	L	Additive aggregation functions
	RCDA	M	M	M	H	N	L	S	L	All
Tree	IPSDA	L	L	H	H	N	M	G	H	SUM
	KIPDA	L	L	H	H	N	L	G	H	MIN/MAX
	PHA	M	M	M	L	Y	L	G	H	All
	SMART	H	L	H	H	N	M	S	M	Additive aggregation functions
	IPDA	H	M	H	H	N	H	S	M	Additive aggregation functions
	EEHA	M	L	L	H	N	L	S	L	SUM
	EAED	L	L	L	H	N	L	M	L	SUM
	MDPA	L	H	M	H	N	H	G	H	All
	VPPRQ	L	M	M	H	Y	M	S	M	All
	PIA	L	M	L	M	Y	M	S	M	Solution 1- AVG, Standard deviation Solution 2- Comparison based aggregation function Solution 3- All Solution 4- Decomposable functions

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

Table 1. Comparison result of different PPDA protocols for WSNs.

6. *Data Loss Resiliency (DLR):* Protocol can compute the correct aggregate, even in some data loss or some nodes fail to participate.
7. *Delay (DLY):* The time taken to get the sensed data from the source to the sink.
8. *Aggregation Function (AF):* To determine how many protocols support Sum, Average, Count, Standard deviation, Max, Min, Variance, Histogram and Median. The two classes: Numerous and Few.
9. *Pay Load Size (PLS):* It determines the real size of the message after applying privacy preserving operation.
10. *Memory Consumption (MC):* This is the amount of memory needed to store the keys, integer ranges and other variables.

4.1. Security Features of PPDA Protocols

1. *Type of encrypted data aggregation:* It determines the type of encrypted data aggregation, either end to end or hop by hop encrypted data aggregation.
2. *Data Integrity:* It ensures that the data has not been modified or altered during data aggregation and also determines whether the protocol achieves data integrity during the data aggregation.
3. *Authenticity:* It ensures that the data is received from an authenticated node to protect from the intruder and determines whether the protocol achieves authenticity during data aggregation.
4. *End to end confidentiality:* It ensures that the partially or fully aggregated data is only known to the BS, no matter how many nodes have been compromised. It checks whether the protocol guarantees end to end confidentiality or not.
5. *Data Freshness:* It protects from replay attack and checks whether the protocol achieves data freshness during data aggregation.

Techniques to achieve privacy	Protocols	Type of encrypted data aggregation	Data Integrity	Authenticity	End to end confidentiality	Data Freshness	Privacy vs Outsiders	Privacy vs Insiders
Perturbation	CPDA	End to end	No	No	No	No	Yes	Yes
	REBIVE	Hop by Hop	No	No	No	Yes	Yes	Yes
	PPDA	Hop by Hop	No	No	No	Yes	Yes	Yes
	ESPDA	End to End	Yes	Yes	Yes	Yes	Yes	Yes
	SRDA	Hop by hop	No	No	No	No	Yes	Yes
	DADPP	Hop by hop	No	No	No	No	Yes	Yes
	PRDA	Hop by hop	No	No	No	No	Yes	Yes
	PHA	End to end	No	Yes	No	Yes	Yes	Yes
	IPSDA	Hop by hop	Yes	No	No	No	Yes	Yes
	KIPDA	Non cryptographic	No	Yes	No	No	Yes	Yes
	VPPRQ	End to end	Yes	Yes	Yes	Yes	Yes	Yes
	MPDA	Hop by hop	Yes	Yes	No	Yes	Yes	Yes
Shuffling	SMART	Hop by hop	No	No	No	No	Yes	Yes
	IPDA	Hop by hop	Yes	No	No	No	Yes	Yes
	EEHA	Hop by hop	No	No	No	No	Yes	Yes
	ICPDA	Hop by hop	Yes	Yes	No	Yes	Yes	Yes
Privacy homomorphism	EAED	End to end	No	Yes	Yes	No	Yes	Yes
	CDA	End to end	No	No	Yes	No	Yes	No
	RCDA	End to end	Yes	No	Yes	No	Yes	Yes
Hybrid	PIA	Hop by hop	Yes	Yes	No	No	Yes	Yes

Table 2. Security features of different PPDA protocols for WSNs.

6. *Data Privacy:* It ensures that the data is only known to the owner, even if adversaries get the secret information. It checks whether the protocol achieves data privacy from adversaries and trusted participating nodes.

5. Conclusion

Privacy preserving data aggregation is the main issue in WSNs. It guarantees that the sensed data is known only to the owner. This paper presents twenty different techniques for the privacy preserving data aggregation. The PPDA protocols are classified based on the type of aggregator, technique used to achieve data privacy and topology used for data aggregation. This paper also discusses some of the application areas of PPDA protocols and, finally, we compare the existing PPDA 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 that this paper will light in the direction of future research.

References

[1] V. PANDEY, AMARJEET, N. CHAND, A review on data aggregation techniques in wireless sensor network. *Journal of Electronics & Electrical Engineering*, **1**(2) (2010), 1–8.

[2] N. S. PATIL, P. R. PATIL, Data Aggregation in Wireless Sensor Network. *IEEE International Conference on Computational Intelligence and Computing Research*, (2010).

[3] K. AKKAYA, I. ARI, In-network Data Aggregation in Wireless Sensor Networks. *Handbook of Computer Networks* (H. BIDGOLI, Ed.), **2** (2008) pp. 1131–1146. John Wiley & Sons.

[4] N. LI, N. ZHANG, S. K. DAS, B. THURAISINGHAM, Privacy preservation in wireless sensor networks: A state of the art survey. *Ad Hoc Networks*, **7** (2009), 1501–1514.

[5] R. BISTA, J.-W. CHANG, Privacy-Preserving Data Aggregation Protocols for Wireless Sensor Networks: A Survey. *Sensors*, **10** (2010), 4577–4601.

[6] R. AGRAWAL, R. SRIKANT, Privacy-preserving data mining. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, (2000) Dallas, TX, USA, pp. 439–450.

[7] H. KARGUPTA, Q. W. S. DATTA, K. SIVAKUMAR, On the privacy preserving properties of random data perturbation techniques. In *Proceedings of the IEEE International Conference on Data Mining*, (2003) Melbourne, FL, USA, pp. 99–106.

[8] W. HE, X. LIU, H. NGUYEN, K. NAHRSTEDT, T. ABDELZAHER, PDA:Privacy-preserving data aggregation in wireless sensor networks. In *Proceedings of 26th IEEE International Conference on Computer Communications (Infocom 2007)*, (2007) Anchorage, Alaska, USA, pp. 2045–2053.

[9] S. MADDEN, M. J. FRANKLIN, J. M. HELLERSTEIN, TAG:A Tiny Aggregation Service for Ad-Hoc Sensor Networks. *OSDI* (2002).

[10] F. RAHMAN, MD. E. HONQUE, S. I. AHAMED, Preserving Privacy in Wireless Sensor Networks using Reliable Data Aggregation. *Applied Computing Review*, 52–60.

[11] SHAMIR, How to share a secret. *Commun. ACM*, **22**(11) (1979), 612–613.

[12] A. UKIL, Priacy Preserving Data Aggregation in Wireless Sensor Networks. In *Proceeding of IEEE ICWCMC*, (2010) Valencia, Spain.

[13] O. GOLDREICH, Secure multi-party computation. Working Draft, Version 1.3, 2011.

[14] H. CAM, S. OZDEMIR, P. NAIR, D. MUTHUAVINASHIAPPAN, H. OZGUR SANLI, Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks.

[15] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST), DES mode of operation. *Federal Information Processing Standards Publication*, 81 (FIPS PUB 81) (1981).

[16] H. ÇAM, Nonblocking OVSF Codes and Enhancing Network Capacity for 3G Wireless and Beyond Systems. Special Issue of Computer Communications on *3G Wireless and Beyond For Computer Communications*, **26**(17) (2003), 1907–1917.

[17] H. ÇAM, K. VADDE, Performance Analysis of Nonblocking OVSF Codes in WCDMA. In *Proc. of the 2002 International Conference on Wireless Networks*, (2002), pp. 50–55.

[18] H. OZGUR SANLI, S. OZDEMIR, H. ÇAM, SRDA: Secure Reference-Based Data Aggregation Protocol for Wireless Sensor Networks.

[19] J. YAO, G. WEN, Protecting classification privacy data aggregation in wireless sensor networks. In *Proceedings of the 4th International Conference on Wireless Communication, Networking and Mobile Computing, WiCOM*, (2008) Dalian, China, pp. 1–5.

[20] M. SHAO, S. ZHU, W. ZHANG, G. CAO, Security and privacy support for data-centric sensor networks. In *Proceeding of 26th IEEE International Conference on Computer Communications, INFOCOM*, (2007) Anchorage, AK, USA, pp. 1298–1306.

[21] M. CONTI, L. ZHANG, S. ROY, R. DI PIETRO, S. JAJODIA, L. V. MANCINI, Privacy-preserving Robust Data Aggregation in Wireless Sensor Networks. *Secure Communication Networks*, **2** (2009), 195–213.

[22] H. CHOI, S. ZHU, P. LA, F. THOMAS, SET: Detecting node clones in sensor networks. In *Proceedings of IEEE 3rd International Conference on Security and Privacy in Communication Networks, SecureComm*, (2007) Nice, France, pp. 341–350.

[23] R. BISTA, H.-K. YOO, J.-W. CHANG, A New Sensitive Aggregation Scheme for Protecting Integrity in Wireless Sensor Networks. In *Proceedings of 10th IEEE International Conference on Computer and Information Technology (CIT 2010)*.

[24] M. M. GROAT, W. HE, S. FORREST, KIPDA: *k*-Indistinguishable Privacy-preserving Data Aggregation in Wireless Senor Networks. In *Proceedings of the 30th International Conference on Computer Communications*, (2011).

[25] W. S. ZHANG, C. WANG, T. M. FENG, GP2S: Generic privacy-preservation solutions for approximate aggregation of sensor data, concise contribution. In *Proceedings of the 6th Annual IEEE International Conference on Pervasive Computing and Communications, PerCom*, (2008) Hong Kong, China, pp. 179–184.

[26] W. HE, X. LIU, H. NGUYEN, K. NAHRSTEDT, T. ABDELZAHER, PDA:Privacy-preserving data aggregation in wireless sensor networks. *Proceedings of 26th IEEE International Conference on Computer Communications (Infocom 2007)*, (2007) Anchorage, Alaska, USA, pp. 2045–2053.

[27] W. HE, X. LIU, H. NGUYEN, K. NAHRSTEDT, T. ABDELZAHER, iPDA: An Integrity – Protecting Private Data Aggregation Scheme for Wireless Sensor Networks. *IEEE MILCOM*, (2008), 1–7.

[28] H. LI, K. LIN, K. LI, Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks. *Computer Communication*, **34** (2011), 591–597.

[29] W. HE, ICPDA: Integrity protecting Private Data Aggregation, 2008.

[30] J. GIRAO, D. WESTHOFF, M. SCHNEIDER, CDA: Concealed Data Aggregation for Reverse Multicast Traffic in Wireless Sensor Networks. In *Proc. 40th International Conference on Communications, IEEE ICC*, (2005).

[31] J. DOMINGO-FERRER, A provably secure additive and multiplicative privacy homomorphism. In *Proceedings of the 5th International Conference on Information Security*, (2002) Sao Paulo, Brazil, pp. 471–483.

[32] C. CASTELLUCCIA, E. MYKLETUN, G. TSUDIK, Efficient aggregation of encrypted data in wireless sensor networks. In *Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, MobiQuitous*, (2005) San Diego, CA, USA, pp. 109–117.

[33] C.-M. CHEN, Y.-H. LIN, Y.-C. LIN, H.-M. SUN, RCDA:Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks. *IEEE Transactions on Parrallel and Distributed Systems*, **23**(4) (2012).

[34] X. LIN, R. LU, X. S. SHEN, MPDA: Multidimensional privacy-preserving aggregation scheme for wireless sensor networks. *Wireless Communications and Mobile Computing*, (2010), pp. 843-856.

[35] D. LIU, P. NING, Establishing pairwise keys in distributed sensor networks. *Proceeding of the 10th ACM Conference on Computer and Communication Security (CCS 2003)*, (2003) Washington, DC, pp. 52–61.

[36] X. LIN, R. LU, X. SHEN, Y. NEMOTO, N. KATO, SAGE: A strong privacy preserving scheme against global eavesdropping for health systems. *IEEE Journal of Selected Areas of Communications* (in press).

[37] B. SHENG, Q. LI, Verifiable privacy-preserving range query in two-tiered sensor networks. In *Proceedings of the 27th IEEE International Conference on Computer Communications, INFOCOM*, (2008) Phoenix, AZ, USA, pp. 457–465.

[38] H. HACIGUMUS, B. R. IYER, C. LI, S. MEHROTRA, Executing SQL over encrypted data in the database service provider model. In *Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data*, (2002) Madison, WI, USA, pp. 216–227.

[39] B. HORE, S. MEHROTRA, G. TSUDIK, A privacy-preserving index for range queries. In *Proceedings of the 28th Very Large Database Conference, VLDB*, (2004) Toronto, Canada, pp. 720–731.

[40] G. TABAN, V. D. GLIGOR, Privacy-preserving integrity-assured data aggregation in sensor networks. In *Proceedings of the International Symposium on Secure Computing, SecureCom*, (2009) Vancouver, Canada, pp. 168–175.

[41] M. BELLARE, R. CANETTI, H. KRAWCZYK, Keying hash functions for message authentication. In *Proceedings of Annual International Cryptology Conference, Crypto*, (1996) Santa Barbara, CA, USA, pp. 1–16.

[42] R. AGRAWAL, J. KIERNAN, R. SRIKANT, Y. XU, Order preserving encryption for numeric data. In *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data*, (2004) Paris, France, pp. 63–574.

[43] H. CHAN, A. PERRIG, D. SONG, Secure hierarchical in-network aggregation in sensor networks. In *Proceedings of the 13th ACM Conference on Computer and Communications Security*, (2006) Alexandria, VA, USA, pp. 278–287.

Received: December, 2013
Revised: March, 2014
Accepted: March, 2014

Contact addresses:

Joyce Jose
Post Graduate Scholar
Dept. Information Technology
Karunya University
Karunya Nagar
Coimbatore-641114
India
e-mail: joycejose1990@gmail.com

Josna Jose
Post Graduate Scholar
Dept. Information Technology
Karunya University
Karunya Nagar
Coimbatore-641114
India
e-mail: josnajose1990@gmail.com

M. Princy
Lecturer
Dept. Information Technology
Karunya University
Karunya Nagar
Coimbatore-641114
India
e-mail: princym@karunya.edu

JOYCE JOSE received the Bachelor of Technology (B Tech) degree in Computer Science & Engineering from Viswajyothi College of Engineering & Technology, Vazhakulam, Kerala, India in the year 2011, and Master of Technology (M Tech) degree in Network and Internet Engineering from Karunya University, Coimbatore, Tamilnadu, India in 2013. She has published 3 International Journals and 2 IEEE Conferences papers. Her research interest is in the area of privacy preserving data aggregation in Wireless Sensor Networks.

JOSNA JOSE received the Bachelor of Technology (B Tech) degree in Computer Science & Engineering from Viswajyothi College of Engineering & Technology, Vazhakulam, Kerala, India in the year 2011, and Master of Technology (M Tech) degree in Network and Internet Engineering from Karunya University, Coimbatore, Tamilnadu, India in 2013. She has published 3 International Journals and 2 IEEE Conferences papers. Her research interest is in the area of secure 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 a project on sleep scheduling algorithm for power efficiency in Sensor Networks and guiding projects on power efficient aggregation and routing in Sensor Networks and she plans to pursue her doctorate in the same area.