Volume 5, Issue 8, August 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Secure Data Aggregation Protocols in Wireless Sensor Networks: A Review

Vasudha Khillan
M.Tech Student of CSE Department
Haryana College of Technology & Management
Kaithal, Haryana, India

Anish Soni
Assistant Professor in CSE Department
Haryana College of Technology & Management
Kaithal, Haryana, India

Abstract-Wireless sensor network is a collection of large number of low cost resource constraint sensor nodes that communicates using wireless medium. Sensor nodes are resource constrained in memory, sensing, communication capability, and battery power. Data communication between nodes consumes a large portion of the total energy consumption of the WSNs. One of the solutions to reduce number of bits transmitted during communication is data aggregation. As wireless sensor networks are usually deployed in remote and hostile environments to transmit sensitive information, sensor nodes are prone to attacks. Thus, security is an important criterion to be considered in WSNs. Many secure data aggregation protocols have been proposed in wireless sensor networks. In this paper, many existing secure data aggregation protocols have been reviewed with their corresponding brief explanation.

Keywords- Wireless Sensor Network, Data Aggregation, Secure Data Aggregation, WSN, Security.

I. INTRODUCTION

Wireless Sensor Network [1] is a network of small sized, low cost sensor nodes, which can sense the environment and communicate the information gathered from the monitored field through wireless links. The data is forwarded via multiple hops to a sinkthat can use it locally or is connected to other networks through a gateway. WSNs are data-centric, application specific, dynamic in nature, and scalable. A sensor node in WSN is made up of four basic components: a sensing unit, a processing unit, a transceiver unit, and a power unit. Sensor nodes may also have application dependent additional components. WSNs have many applications such as in military field surveillance, environment monitoring, health care, accident report, law enforcement, and also in home applications.

WSNs have many issues that affect their design and performance such as deployment, localization, synchronization, energy consumption, calibration, security etc. out of which energy consumption of sensors is one of the major issue as sensors are resource constrained in battery power. Due to dense deployment of sensor nodes in WSN, neighboring nodes can have overlapping range which results in the sensing and transmission of similar data. This increases the amount of energy and bandwidth usage. One of the solutions to this problem is data aggregation. Data aggregation [2] is a process of aggregating data sensed by the sensor nodes by using aggregation functions such as min, max, sum, average etc. and transfer the aggregated data to the higher level aggregation node. Thus, data aggregation technique greatly helps to decrease the number of transmissions in the network, eventually reducing the bandwidth usage, eliminating unnecessary energy consumption and hence to increase the overall network lifetime.

In a hostile environment the aggregated results should be protected from various types of attacks such as sybil attack, selective forwarding attack, sinkhole attack, wormhole attack, node subversion etc. A compromised sensor node generates false reading and false aggregation result. But base station is not capable of detecting the presence of compromised node because attacker behaves in such a way that its incorrect result is easily acceptable by base station. Hence security is necessary to be employed with data aggregation techniques so as to achieve data confidentiality, data integrity, data freshness, data availability and source authentication. Various protocols have been proposed for secure data aggregation in wireless sensor networks.

There are basically two types of secure data aggregation protocols [3] based on the topology used for aggregation. First is tree-based data aggregation protocols in which the intermediate parent nodes in the path from leaf to base station perform data aggregation. The main issue in this type of protocols is to construct an energy-efficient data aggregation tree. Second is cluster-based data aggregation protocols in which sensor nodes are subdivided into clusters. In each cluster, a cluster head is elected to aggregate the data locally and transmit the aggregation result to the base station.

II. SECURE DATA AGGREGATION PROTOCOLS

A. SDA: Secure Data Aggregation [2003]:

The first secure data aggregation (SDA) [4] was proposed by Hu & Evans (2003) who studied the problem of data aggregation once one node is compromised. This protocol aims at providing lightweight security mechanism to effectively detect node misbehavior (dropping, modifying or forging messages, transmitting false aggregate value). It

exploits two main ideas: delayed aggregation and delayed authentication. Therefore, the sensor readings are forwarded unchanged over the first hop and then aggregated at the second hop instead of aggregating at the immediate next hop. This increases the confidence in the sensor readings integrity but data can be altered once a parent and child in the hierarchy are compromised. Even if a compromised node is detected, no practical action can reduce the damaged caused by it. This protocol saves resources by authenticating messages after a time delay instead of authenticating them right away, which enables authentication keys to be symmetric keys. It uses a μTESLA protocol for authentication of messages transmitted by base station and achieves asymmetry from clock synchronization and delayed key disclosure. Before deployment, each sensor node is initialized with a symmetric secret key, K_{AS}, which is shared with base station. Thus, this proposed scheme offers data integrity, freshness and authentication.

The process of the proposed scheme starts with the leaf nodes which send their sensed data, node id, and message authentication code (MAC) to their parent node. The parent nodes store the message and MAC for a specified time so as to receive messages from all the child nodes and then retransmit the messages and MACs to its parent node. This parent node will aggregate the data received from its grandchildren (via its children) and transmit it to its parent along with the MAC of the aggregation value. The process goes on till messages arrive at the base station. Base station then reveals the temporary node keys along with MAC generated using μTESLA key. After this, nodes and base station advance to their next key in the chain. Thus, the SDA protocol provides integrity, authentication, and data freshness.

B. SIA: Secure Information Aggregation [2003]:

BartoszPrzydatek, Dawn Song, and AndrianPerrig proposed a secure information aggregation (SIA) [5] framework for WSNs called aggregate-commit-prove. This proposed framework provides resistance against a special type of attack called stealthy attack where the attacker's goal is to make the user accept false aggregation results, which are significantly different from true results determined by the measured values, while not being detected by the user. So the security goal is to prevent the user from accepting incorrect results i.e. to prevent the stealthy attacks. This scheme consists of three types of nodes: a home server, a base station, and sensor nodes. SIA assumes that each sensor node has a unique identifier and shares a separate secret cryptographic key with the home server and the aggregator. If data confidentiality is requires, then these keys enable message authentication and encryption. Furthermore, it assumes that a set of uncorrupted sensor nodes in the network can reach each other via paths composed of only uncorrupted sensor nodes. Moreover, it assumes that the home server and base station can have a mechanism, such as TESLA broadcast authentication protocol, to broadcast authentic messages.

The general proposed method for assurance of authenticity and validity of messages sent by aggregator to base station is aggregate-commit-prove. So it consists of three phases: first is aggregation i.e. collecting the data from sensor nodes and locally computing the aggregation result; second is commitment i.e. committing the collected data using Merkle hash-tree construction; third is reporting & proving i.e. reporting the aggregation result while proving the correctness of the result. SIA offers data integrity, authentication, data freshness, and confidentiality (if required).

C. ESPDA: Energy-Efficient and Secure Pattern-based Data Aggregation Protocol [2003]:

H. Cam, S. Ozdemi, P. Nair, and D. Muthuavinashiappan proposed a protocol ESPDA [6] to provide energy-efficient data aggregation together with secure data communication in wireless sensor networks. It is a cluster-based data aggregation protocol. In ESPDA, cluster-head first broadcasts the pattern seed to the sensor nodes and requests them to send the corresponding pattern code for the sensed data. These pattern codes are generated using the secret pattern seed sent by cluster-head which prevents the retrieval of real data from pattern codes and the pattern generation algorithm (PG). These patterns are analyzed by the pattern comparison algorithm at the cluster-head. If multiple sensor nodes send the same pattern code to the cluster-head because of sensing the common data, then only one of them is permitted to send the data to the cluster-head. Thus, data aggregation is performed even before the actual data is transmitted from the sensor nodes.

ESPDA also provides security because it aggregates data by pattern codes, so cluster-heads need not to know the contents of the transmitted data. Thus, the sensor data is transmitted to base station in encrypted form without decrypted anywhere in the transmission path. ESPDA employs a Non-blocking Orthogonal Variable Spreading Factor (NOVSF) code hopping technique. Sensor nodes compute a node-specific-secret-key (NSSK) using their unique secret built-in key and a session key broadcasted by the base station. This NSSK is used to encrypt and decrypt all the data transmissions during a session. Thus, ESPDA is an energy-efficient, bandwidth efficient, and secure protocol which provides data confidentiality, authentication, and data freshness.

D. SecureDAV: A Secure Data Aggregation and Verification Protocol [2004]:

Ajay Mahimkar and Theodore S. Rappaport proposed a protocol [7] that improved the data integrity vulnerability in SDA by signing the aggregated data. SecureDAV is a cluster-based data aggregation protocol. An elliptic curve cryptosystems (ECC) is used for establishing cluster keys in sensor networks using verifiable secret sharing. ECC is used for key management because of its smaller key size, faster computations and reduction in processing power, storage space, and bandwidth. Each sensor within a cluster has a share of the secret cluster key. For each cluster, once the cluster-head receives the sensor readings, it aggregates them and computes its average. It then broadcasts the computed average to all the sensor nodes within its cluster. Then each sensor in the cluster compares its reading with the average value received from the cluster-head. Then, each sensor node partially signs the average value only if the difference between

the received average value and its reading is less than a threshold and then send signed average to the cluster-head. The cluster-head combines the partial signatures to form a full signature of the aggregated result and then send this full signature along with the average reading to the base station. The validity of this signature is verified at the base station who has the corresponding public key. Threshold signature scheme ensures the authenticity of message. While the integrity of the readings is ensured using Merkle Hash Tree avoiding over-reliance on cluster-heads. Thus this protocol ensures that the base station never accepts faulty readings. So we can conclude that SecureDAV provides data confidentiality, data integrity and authentication. There are some drawbacks of this protocol such as it requires high communication costs on data validation, and it supports only the AVG aggregation function.SecureDAV provides data confidentiality, data integrity, and authentication.

E. SRDA: Secure Reference-Based Data Aggregation Protocol [2004]:

H. OzgurSanli, SuatOzdemir, and Hasan Cam developed a new data aggregation technique called SRDA [8] that sends the differential data i.e. difference between the sensed data and the reference value instead of the raw sensed data. Reference value is taken as the average value of previous sensor readings. Each sensor node first sense the data from environment, then computes the differential data, encrypts it, and send it to the cluster-head. SRDA provides a key distribution scheme with low memory overhead to establish secure communication links in the network and to save energy it implements variable strength security at different levels of the clustering hierarchy i.e. the security level of the network is gradually increased as the data is traveled to higher level cluster-head. Increasing security levels are implemented by using a cryptographic algorithm RC6 with adjustable parameters such as the number of rounds to achieve different level of security in the WSN. Increasing or decreasing the number of rounds changes the security strength of the RC6 that can be measured by a parameter called Security Margin. The security margin is the deviation of the actual number of rounds from the minimum number of rounds for which the algorithm is considered to be secured. The SRDA uses a higher security margin at higher level cluster-heads compared to low level cluster-heads. Thus, SRDA incorporates both data aggregation and security concepts together in cluster-based wireless sensor network. SRDA provides data confidentiality, data freshness, and authentication.

F. CDA: Concealed Data Aggregation [2005]:

Joao Girao, Markus Schneider, and Dirk Westhoff addressed the problem of aggregating encrypted data in WSN. They proposed a protocol called CDA [9] which uses an additive and multiplicative homomorphic encryption scheme that allows the aggregator to aggregate encrypted data. In this approach, every sensor node shares a same key with the base station. So it does not provide guarantee for the privacy of individually sensed data because once a sensor node is compromised, it leads to the decryption of other sensors data. In this protocol, each sensor node splits its data into 'd' parts (where d<=2) and then encrypt them using the common key shared with the base station and send it to the aggregator. Aggregator aggregates the encrypted sensor data with other sensors encrypted data because of privacy homomorphism property and finally sends this aggregated result to the sink. This aggregated data is decrypted at the sink using the same key used for the encryption.

The developers of this protocol have applied the privacy homomorphism (PH), proposed by Domingo Ferrer, which is suitable to aggregation function average and movement detection. There are some disadvantages of this protocol such as its vulnerability to replay attack and malicious aggregation, expensive encryption, additional communication overhead, and also this protocol does not address the problem of non-response ID. The authors of this protocol argued that the security level of this protocol is reasonable, but Wagner proved that PH is unsecure against chosen plain text attacks. Thus, CDA ensures only data confidentiality.

G. SDAP: Secure Hop-by-Hop Data Aggregation Protocol [2006]:

Yi Yang, Xinran Wang, Sencum Zhu, and Guohong Cao proposed SDAP [10] protocol which can tolerate more than one compromised node. The design of SDAP is based on two principles: divide-and-conquer and commit-and-attest. This general purpose data aggregation protocol has three steps. First step is tree construction and query dissemination, in which an aggregation tree is constructed and thereby all nodes identify their parents, after which the base station disseminates the aggregation query message through the tree. Second step is probabilistic grouping and data aggregation, in which SDAP uses the divide-and-conquer principle to divide the network tree into multiple logical subtrees based on a probabilistic grouping technique which depends on group leader selection. This increases the number of aggregators and reduces the number of nodes in each subtree and hence, the damage caused by a compromised aggregator of a subtree is reduced. Then it generates one group aggregate from each group by hop-by-hop aggregation. The other principle of SDAP i.e. commit-and-attest, enhances an ordinary hop-by-hop aggregation protocol with commitment capability, which ensures that once a group commits its aggregate then this group cannot deny it later. The third step is verification and attestation, in which the content of data packet and the authenticity of leader are verified first. Then the base station identifies the suspicious groups based on group aggregates by using a bivariate multiple-outlier detection algorithm. These suspected groups then participate in an attestation process to prove the correctness of their group aggregate. Finally the aggregate is calculated over all the group aggregates that have passed the attestation procedure.

This protocol has advantages such as it is applicable on multiple aggregation functions, it has adjustable detection rate, and also this protocol provides data integrity, data confidentiality, and source authentication. But energy utilization and transmission overhead of this protocol is high.

H. SELDA: Secure and Reliable Data Aggregation [2007]:

SuatOzdemir proposed a Secure and rELiable Data Aggregation protocol (SELDA) [11] which is based on trustworthiness of sensor nodes and data aggregators. This protocol ensures security and reliability of aggregated data in the presence of compromised sensor nodes. The first step in the protocol is generating the web of trust. Sensor nodes observe actions (sensing, routing, and availability) of their neighboring nodes to compute the reputation values of sensor nodes using Beta Distribution Function. Compromised sensor nodes of the sensor network have lower reputation values than honest sensor nodes. These reputation values are exchanged among sensor nodes to form a web of trust that allows them to determine reliable data aggregators and secure paths to those data aggregators. The second step is secure and reliable data aggregation, in which the data of each sensor node is weighted based on its reputation value with respect to the data aggregator by using Reliable Data Aggregation (RDA) algorithm, thereby reducing the effect of compromised sensor nodes on the aggregated data. The final step is the multi path data transmission which resists the forgery and selective forwarding attacks by compromised nodes. A secure multi path data transmission algorithm is used to select some paths based on their reliability and keeps the quantity and identity of the selected paths secret. The sensor node transmits its data to data aggregator over those selected secure paths which ensure the secure data delivery to data aggregators. Thus, SELDA increases the reliability of the aggregated data at the expense of a tolerable communication overhead. SELDA provides data integrity, authentication, and freshness.

I. SEDAN: Secure and Efficient protocol for Data Aggregation [2007]:

MiloudBagaa, NoureddineLasla, AbdelraoufOuadjaout, and YacineChallal proposed a protocol [12] that provides secure data aggregation for wireless sensor networks. This protocol is based on a two hops verification mechanism of data integrity in which each node can verify immediately the integrity of its two hops neighbors' data, and the aggregation of the immediate neighbors. Thus this protocol avoids the useless transmission of bogus data and hence reduces the energy consumption. A new type of key called "two hops pair-wise key" is used in it which allows sharing a secret between any two hops neighbors, unknown by the intermediate node. When a node transmits its data, it calculates a MAC using two hops pair-wise key shared with the grandparent. Since this key is unknown by the next hop, the integrity of data is preserved and any modification in value can be detected by the grandparent.

SEDAN assumes a tree communication topology and its process consists of five steps. The first step is "key establishment", in which all the needed pair-wise keys are established. Second step is "data authentication", in which when a node wants to send its data, it sends its ID, sequence number of data, One hop MAC (computed using the pair-wise key shared with the parent), and Two hops MAC (computed using the pair-wise key shared with the grandparent) along with the data to its parent. The third step is "one-hop data integrity verification", in which when a node receives a data packet from its child, it verifies its one hop MAC to validate the origin of the packet and stores the rest of the information. The fourth step is "authentication of aggregated data", in which the node aggregates all the data received from its child nodes and then calculates the one hop MAC and two hops MAC of this aggregated result and send this aggregated data along with these two MACs and its own ID and sequence number of data to its parent. The final step is "two-hops data integrity verification", in which the grandparent node verifies the two hop MACs of each of its grandchildren and also computes the correct aggregation value of its child node using the grandchildren's data. Then it compares its calculated value with the value generated by its child node. Thus, SEDAN can detect faulty aggregation immediately without any delay. SEDAN provides data integrity, freshness, and authentication.

J. RSDA: Reputation-based Secure Data Aggregation [2008]:

Hani Alzaid, Ernest Foo, and Juan Gonzalez Nieto proposed a new protocol RSDA [13] that integrates the aggregation functionalities with the advantages that are provided by a reputation system in order to enhance the network lifetime and the accuracy of aggregated data. RSDA is composed of two types of identities: a base station and normal sensor nodes. The target terrain, where RSDA is implemented, is divided into smaller non-overlapping cells of equal areas. During the bootstrap period when the network is not vulnerable to any type of attacks, each sensor node discovers its neighboring nodes and computes its cell key and shared keys with neighboring cells. Each sensor node monitors the behavior of other nodes within the same cell and then calculates the reputation value for them. Based on the calculated reputation values, one of the sensor node is selected to be the Cell Representative. The aggregation process begins when the base station initiates a query message to all the cells. The cell representative is responsible for confirming its cell reading, aggregating it with other readings (if the cell is an intermediate cell), and forwarding it to the upper stream cell. The base station then processes the received answer for its query and derives meaningful information that reflects the events in the target field.

RSDA is capable of detecting compromised nodes and then black list them which helps to reach its two main goals: extend the network lifetime and protect the accuracy of the aggregated data. Besides data accuracy and availability, it also provides other security services such as data integrity, freshness and authentication. It provides resistance against selective forwarding attack, replay attack, and stealthy attack. But it suffers from the node compromise attack.

K. SEEDA: Secure End-to-End Data Aggregation [2010]:

A.S.Poornima and B.B.Amberker proposed a secure data aggregation scheme [14] which provides end-to-end data privacy. There are two confidentiality requirements considered in this protocol. First is generic confidentiality i.e. sensors not participating in aggregation mechanism have no access to contents of the data. Second is end-to-end confidentiality i.e. sensors actively participating in aggregation mechanism have no access to the data that is already aggregated. There

are three types of nodes considered in this protocol: sink node, sensor nodes, and aggregator nodes. These nodes are organized as a m-ary tree where sink node is at the root of tree.

This protocol is a hybrid scheme which adopts the best features of end-to-end aggregation scheme and hop-by-hop aggregation scheme. Thus it ensures end-to-end data privacy and the number of bits transmitted is almost same as that of hop-by-hop aggregation scheme. The process of this protocol starts with the deployment of nodes in tree structure and numbering the levels of the tree from 0,1,…,h where h is the height of tree. Sink node is at level 0. Node at level h senses the data and encrypts it using secret key. This encrypted data (cipher text) is forwarded to next higher level where the aggregator adds the cipher text of responding nodes whereas for non-responding nodes message value 0 is added to the aggregated cipher text. The count of non-responding nodes is appended to this aggregated message and then the message is forwarded to next higher level. This process is repeated for all the higher levels till level 1. Finally the sink node decrypts the message by subtracting all the keys from the cipher text and computes the average based on counter value, which gives information regarding the non-responding nodes. Thus the average number of bits transmitted per node is reduced in this protocol. SEEDA provides data confidentiality, authentication, and data freshness.

L. EEHA: Energy-Efficient and High-Accuracy Secure Data Aggregation [2011]:

Hongjuan Li, Kai Lin, and Keqiu Li proposed a protocol EEHA [15], in which accurate data aggregation is achieved without releasing private sensor readings and without introducing significant overhead on the battery-limited sensors. The main focus of this protocol is on the defense of eavesdropping attack in which an attacker tries to overhear the transmission over wireless links to obtain private information. The design objective of this scheme is to achieve accurate data aggregation with moderate extra communication overhead to preserve data privacy. This protocol consists of three types of nodes: base station, intermediate nodes, and leaf nodes. The process starts with the construction of aggregation tree, which is a directed tree formed by the union of all paths from the sensor nodes to the base station. Then the leaf nodes adopt "slicing and mixing" technique in which they slice their private data into pieces, and send these pieces to different neighbors while one piece is kept by itself. All the leaf nodes wait for a certain time and then mix (or sum) all the received slices and the slice left by itself to get a new result. This result is encrypted and sent to the intermediate node. The intermediate node aggregates the received data and its own sensor reading, & then forwards it to its parent. The result is propagated level by level up the tree and reaches the root, at which the final data is the summation of all the sensors data. In this protocol, the "slicing and assembling (or mixing)" technique is implemented only to leaf nodes of the aggregation tree, hence the communication overhead is greatly reduced because of less number of nodes that slice their data. Less communication overhead leads to less message collision, which results in high level of aggregation accuracy. EEHA is very energy-efficient because of the low communication overhead. Data privacy is achieved by using different schemes for leaf nodes and intermediate nodes. Leaf nodes use slicing and assembling technique for data privacy while intermediate nodes use aggregation functions. EEHA provides data confidentiality, authentication, and data freshness.

M. IPHCDA: Integrity Protecting Hierarchical Concealed Data Aggregation [2011]:

SuatOzdemir, Yang Xiao proposed this protocol which allows hierarchical aggregation of data encrypted with different keys while providing data integrity and confidentiality. This protocol [16] employs a privacy homomorphic encryption scheme (based on elliptic curve cryptography) and message authentication codes (MAC) to achieve hierarchical data aggregation. IPHCDA assumes a group based network deployment in which the network area is divided into different regions and a public/private key pair (public key to sensor nodes and private key to base station) is assigned to each region. Each sensor node of a region also shares a unique MAC key with the base station.

In this protocol, each sensor node encrypts its sensed data using the public key of its corresponding region and sends it to the data aggregator of its region. The data aggregator aggregates the received encrypted data from the sensor nodes in its region and then computes the MAC of the aggregated data using a unique symmetric key that it shares with the base station. The encrypted data of several regions is hierarchically aggregated into a single piece of data without violating data confidentiality and the MAC of each region is combined using the XOR function and then sent to the base station. During the decryption, the base station can classify the aggregated data based on the encryption keys and verify the MAC of the aggregated data, thereby achieving data integrity. IPHCDA provides resistance to various attacks such as ciphertext analysis, known plaintext attack, replay attack, malleability, unauthorized authentication, forge packets, and physical attacks. IPHCDA provides data confidentiality, data integrity, data freshness, and authentication.

N. RCDA: Recoverable Concealed Data Aggregation for Data Integrity [2012]:

Chien-Ming Chen, Yue-Hsun Lin, Ya-Ching Lin, and Hung-Min Sun proposed this protocol [17] in which the base station can recover all sensing data generated by sensors even if these data has been aggregated by cluster-head. This property is called "recoverable". With these individual data, the base station can verify the integrity and authenticity of all sensing data and also the base station can perform any aggregation function on them. RCDA is a cluster-based aggregation protocol, in which all the sensor nodes in the network are divided into several clusters. Each cluster has a cluster-head responsible for collecting and aggregating sensing data. In this protocol two RCDA schemes have been proposed: RCDA-HOMO and RCDA-HETE for homogeneous and heterogeneous WSN respectively.

RCDA-HOMO is composed of four procedures. The first is "Setup" in which all the necessary secrets for the base station and sensors are prepared and installed. The second procedure is "Encrypt-Sign" which is performed by sensors before sending their sensed data to cluster-head. The third procedure is "Aggregate" which is performed by cluster-head

once it receives all results from its members, and then sends the final result to base station. The last procedure is "Verify" in which the base station first extracts individual sensing data by decryption of aggregated data and then verifies the authenticity and integrity of decrypted data.

RCDA-HETE has two types of sensors: L-Sensors (low-end sensors) and H-Sensors (high-end sensors). H-sensors are capable of stronger computation ability and stable power supply and hence act as cluster-heads. RCDA-HETE is composed of five procedures. The first is "Setup" procedure in which necessary secrets are loaded to H-Sensor and L-Sensor. The second is "Intracluster Encrypt" which is employed when L-Sensors desire to send their sensing data to the corresponding H-Sensor. The third is "Intercluster Encrypt" in which each H-Sensor aggregates the received data and then encrypts and signs the aggregated result. The fourth is "Aggregate" procedure which is activated if an H-Sensor receives ciphertexts and signatures from other H-Sensors on its routing path. The last is "Verify" which ensures the authenticity and integrity of each aggregated result. Thus, RCDA provides data confidentiality, data integrity, data freshness, and authentication.

O. CRSR Secure Data Aggregation Algorithm [2013]:

Mrs. R. Lathamanju and Dr. P. Senthilkumar proposed this secure data aggregation algorithm [18] to improve the network lifetime and provide security. This algorithm employs cluster based data aggregation technique by using LEACH-KED algorithm to form clusters. CRSR performs the four stages of operation i.e. C- Challenge the friends, R-Rate the friends, S- Share friends, R- Routing through friends. The first three stages of the algorithm are periodically repeated whereas the last stage is executed whenever needed. This algorithm starts with the challenging process in which trusted nodes are identified by sending an initial challenge. The nodes which complete the challenge find place in the friend list and rest are shifted to the question mark list which contains information about the malicious nodes. Then the second stage comes in which the friends are rated on a scale of zero to ten on the basis of the amount of data they transfer through themselves and according to the rating of other friends which is obtained during the friend list sharing process. Then the stage of sharing the friends comes, in which any node can ask for a friend sharing request, and after friend sharing challenges are initiated for those nodes which were not in the friend list. The last stage is routing through friends in which, when a node wants to transmit data, it initiates a route request message. Each intermediate node forwards the request only if the sending node is not in the question mark list. When the destination node receives the route request, it sends route reply and its public key. On receiving the route reply message, the source node evaluates the best route with the greatest number of trusted friends to transmit the data. An ad-hoc on demand distance vector (AODV) routing protocol is used for transmitting data securely to the destination. The data to be sent is encrypted by using Diffie-Hellman algorithm to enhance secure routing. Thus the chances of man-in-the-middle attack and eavesdropping are greatly reduced. Hence, this algorithm provides an efficient way of data transmission even in the presence of malicious nodes by identifying them and putting into question mark list. CRSR provides data confidentiality, data freshness, and authentication.

P. PEPPDA: Power Efficient Privacy Preserving Data Aggregation [2013]:

Joyce Jose, M. Princy, and Josna Jose proposed this protocol which is best suited for time critical and secure applications such as military applications. The main aim of this protocol [19] is to provide a secure data aggregation scheme which guarantees the privacy, authenticity and freshness of individual sensed data as well as the accuracy and confidentiality of the aggregated data without introducing a significant overhead on the battery limited sensors. The privacy is achieved through using slicing and assembling operation at leaf nodes. Data confidentiality is achieved through the end-to-end encrypted data aggregation. Message authentication is achieved using the secret key and ID pair of each node. Data freshness is achieved by using the varying encryption key for each session. The security protocol used in PEPPDA is similar to the ESPDA key distribution used in cluster topology.

Three types of nodes are considered in this protocol: base station (sink or query station), intermediate nodes (aggregators), and leaf nodes (normal sensor nodes). This protocol consists of four steps. First is aggregation tree construction, in which aggregation tree rooted at base station is constructed using TAG protocol. The second step is slicing, in which each leaf node slices its sensed data into pieces and then encrypt them using encryption key generated by node after it receives the session key from the base station. One of the slice is kept to itself and other are appended with node ID and transmitted to neighbors. The third step is mixing, in which the node sums up all the received encrypted slices using privacy homomorphism technique. The final step is aggregation, in which leaf node sends the aggregated result and the encrypted slice appended with encrypted slice kept by itself to its parent node. The intermediate node encrypts its own data and then sum up it with the aggregated results received from all its child nodes. It also appends its ID with it and then send it to the upper aggregator. The aggregation result goes level by level and finally reaches the base station where it is decrypted by the base station using the decryption key and the aggregated result is generated. PEPPDA provides data confidentiality, data freshness, and authentication.

Q. EESSDA: Energy-Efficient and Scalable Secure Data Aggregation [2013]:

Taochun Wang, Xiaolin Qin, and Liang Liu proposed an energy-efficient, secure, highly accurate and scalable protocol for data aggregation (EESSDA) [20]. This protocol does not need encryption and decryption operations during data aggregation and the secure data aggregation is achieved by establishing secure channel and slicing technology. There are three types of nodes considered in this protocol: the Sink, intermediate nodes, and leaf nodes. This protocol adopts a random key distribution mechanism which consists of three phases: key pre-distribution (each node selects k

keys from key-pool to form a key ring), shared-key discovery (a secure link is established between neighbors which share common key), and path-key establishment (secure link is established by two or more multihop if no common key is shared between neighboring nodes).

EESSDA consists of five steps. The first step is "aggregation tree construction", in which the network is organized as a tree rooted at the sink node where each sensor node has a shortest routing path to the sink and also all the parent-child nodes share a common key. The second step is "secure channel establishment", in which each intermediate node establishes a secure channel with its parent or child node by sharing a common secret random number and also each leaf node establishes a secure channel with its parent node and neighbors. The third step is "slicing", in which leaf node slices its data into pieces before sending to parent node so as to ensure the confidentiality of data but one of the slices is kept at the leaf node itself. The fourth step is "assembling and mixing", in which all nodes wait for a certain time to receive all slices and then each leaf node aggregates the received slices and the slice kept by it to get a new result. The final step is "aggregation", in which leaf node sends the new result to its parent (intermediate node) through secure channel. After receiving all data from child and leaf nodes, intermediate node performs an aggregation function to get a new result which is further forwarded to its parent through secure channel. The process goes on and hence the final aggregation result reaches the sink.

EESSDA provides privacy by the use of secure channel and slicing & assembling technology. It is an energy efficient protocol as it does not require the encryption/decryption in the processing of data aggregation. The amount of traffic is reduced in this protocol which results in high accuracy of aggregation because the chances of data packets collision are reduced. The advanced deployment of shared information between nodes is not required in this protocol, which increases the scalability of the network. EESSDA provides data confidentiality.

R. ECIPAP: Efficient Confidentiality and Integrity Preserving Aggregation Protocol [2014]:

Liehuang Zhu, Zhen Yang, JingfengXue, and Cong Guo proposed an efficient confidentiality and integrity preserving aggregation protocol [21] based on homomorphic encryption and result-checking mechanism. In this protocol it is assumes that an aggregation tree is already set up in the deployment phase or if not already set up, then TAG can be used to build such tree-based network. There are three types of nodes in tree: base station, intermediate nodes, and leaf nodes. Before the deployment of sensor nodes in the monitory area, each sensor node shares a private key, a large integer, and a unique ID with base station.

This protocol has three phases: query dissemination, data aggregation, and result-checking. In the query dissemination phase, base station broadcasts a query message along with a random number to the whole network by using an authenticated method μTESLA. On receiving the query message, sensor nodes store it in their RAMs and then start the data aggregation phase. In the data aggregation phase, sensor node collects environment data value such as temperature and set the values of some parameters. Next, it generates temporary keys so as to encrypt these parameters. Then message authentication code is computed by the sensor node and a data tuple is prepared which contains the encrypted parameters and MAC of that node. This data tuple is sent to the parent node. The parent node then aggregates the data tuples received from its child nodes along with the data tuple created by itself. This result is then transmitted to the next higher level of the tree and eventually the final aggregation result is transmitted to the base station. In the result-checking phase, the base station first decrypts the message and then broadcasts the aggregated tuple down to the whole network using authenticated method. Every sensor node can verify if its own data was added to the aggregation data by comparing its own data to the data sent by parent nodes. The intermediate sensor nodes aggregates the received authentication messages and aggregate them using MAC aggregation function. The base station also calculates this authentication message with its own data stored before network deployment. These two messages are then compared to verify if all the sensing data is added to the final aggregation result. The base station accepts this aggregation result only if it passes the verification phase. ECIPAP provides data confidentiality, data integrity, data freshness, and authentication.

III. CONCLUSION

This paper provides brief description of various secure data aggregation protocols in wireless sensor networks. Additional with this, various design issues such as data confidentiality, data integrity, data freshness, and authentication, of these secure data aggregation protocols are also discussed. By using this data, required secure data aggregation protocols for various wireless sensor network applications can be easily chosen.

REFERENCES

[1] J. Jose, J. Jose, and F. Jose, "A Survey on Secure Data Aggregation Protocols in Wireless Sensor Networks", in International Journal of Computer Applications, Volume 55, October 2012.

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

[3] S. Ozdemir, Y. Xio, "Secure Data Aggregation in Wireless Sensor Networks: A Comprehensive Overview", in Journal of Computer Networks, Elsevier, Volume 53, Issue 12, 13 August 2009, pp. 2022–2037.

[4] L. Hu, D. Evans, "Secure Aggregation for Wireless Networks", in Symposium on Applications and the Internet Workshops, 27-31 January 2003, pp. 384-391.

[5] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure Information Aggregation in Sensor Networks", in proceedings of the 1st International Conference on Embedded Networked Sensor Systems, 2003, pp. 255-265.

[6] H.Cam, S.Ozdemir, P. Nair, and D. Muthuavinashiappan, "ESPDA: Energy-Efficient and Secure Pattern-Based Data Aggregation for Wireless Sensor Networks", in Computer Communications, Elsevier, Volume 29, Issue 4, February 2006, pp. 446–455.

[7] A. Mahimkar, T. S. Rappaport, "SecureDAV: A Secure Data Aggregation and Verification Protocol for Sensor Networks", in IEEE Conference on Global Telecommunications, Volume 4, 29 November – 3 December 2004, pp. 2175-2179.

[8] H. OzgurSanli, S. Ozdemir, and H. Cam, "SRDA: Secure Reference-Based Data Aggregation Protocol for Wireless Sensor Networks", in IEEE 60th Conference on Vehicular Technology, VTC2004-Fall, Volume 7, 26-29 September 2004, pp. 4650–4654.

[9] J. Girao, M. Schneider, and D. Westhoff, "CDA: Concealed Data Aggregation in Wireless Sensor Networks", in IEEE International Conference on Communications, Volume 5, 16-20 May 2005, pp. 3044-3049.

[10] Y. Yang, X. Wang, S. Zhu, and G. Cao, "SDAP: A Secure Hop-by-Hop Data Aggregation Protocol for Sensor Networks", in Journal of ACM Transactions on Information and System Security (TISSEC), Volume 11, Issue 4, July 2008, Article No. 18, New York, USA.

[11] S. Ozdemir, "Secure and Reliable Data Aggregatiob for Wireless Sensor Networks", in proceedings of 4th International Symposium, UCS 2007, Tokya, Japan, 25-28 November 2007, pp. 102-109.

[12] M. Bagaa, N. Lasla, A. Ouadjaout, and Y. Challal, "SEDAN: Secure and Efficient Protocol for Data Aggregation in Wireless Sensor Networks", in 32nd IEEE Conference on Local Computer Networks, 15-18 October 2007, pp. 1053-1060.

[13] H. Alzaid, E. Foo, and J. G. Nieto, "RSDA: Reputation-based Secure Data Aggregation in Wireless Sensor Networks", in 9th IEEE International Conference on Parallel and Distributed Computing, Applications and Technology, 1-4 December 2008, pp. 419-424.

[14] A. S. Poornima, B. B. Amberker, "SEEDA: Secure End-to-End Data Aggregation in Wireless Sensor Networks", in 7th IEEE International Conference on Wireless and Optical Communications Networks (WOCN), 6-8 September 2010, pp. 1-5.

[15] H. Li, K. Lin, K. Li, "Energy-Efficient and High-Accuracy Secure Data Aggregation in Wireless Sensor Networks", in Journal of Computer Communications, Elsevier, Volume 34, Issue 4, 1 April 2011, pp. 591–597.

[16] S.Ozdemir, Y.Xiao, "Integrity Protecting Hierarchical Concealed Data Aggregation for Wireless Sensor Networks", in Journal of Computer Networks, Elsevier, Volume 55, Issue 8, 1 June 2011, pp. 1735–1746.

[17] C. M. Chen, Y. H. Lin, Y. C. Lin, and H. M. Sun, "RCDA: Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks", in IEEE Transactions on Parallel and Distributed Systems, Volume 23, Issue 4, 18 August 2011, pp. 727-734.

[18] R. Lathamanju,P. Senthilkumar, "CRSR Algorithm: A Secure Data Aggregation Algorithm in WSN", in International Journal of Advanced Research in Electronics and Communication Engineering, Volume 2, Issue 9, September 2013.

[19] J. Jose, M. Princy, and J. Jose, "PEPPDA: Power Efficient Privacy Preserving Data Aggregation for Wireless Sensor Networks", in IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology, 25-26 March 2013, pp. 330-336.

[20] T.Wang, X.Qin, and L.Liu, "An Energy-Efficient and Scalable Secure Data Aggregation for Wireless Sensor Networks" in International Journal of Distributed Sensor Networks, Hindawi Publications, 2013.

[21] L. Zhu, Z. Yang, J. Xue, and C. Guo, "An Efficient Confidentiality and Integrity Preserving Aggregation Protocol in Wireless Sensor Networks", in International Journal of Distributed Sensor Networks, Hindawi Publications, 2014.

