International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 5 (2018) pp. 3027-3033

OSDAP- Optimized and Secure Data Aggregation Protocol for Wireless Sensor Networks

Anish Soni[1], Dr. Rajneesh Randhawa[2]

Punjabi University, Patiala (Punjab), India.

Abstract

Maximum utilization of limited resources (energy, bandwidth and memory) in wireless sensor networks is a challenge for the researcher community. Addition to this, security in data transmission is also a big constraint which make WSN research more and more attractive. Data aggregation is one of the technique for energy preservation which minimize the computing overhead by eliminating the redundant data. However, some security attacks makes the data aggregation readings false, thus not providing accurate results. Also, there is a tradeoff between security and energy consumption of WSN. If we go for higher security, it also increases the energy consumption (more encryption and decryption causes more energy consumption) and if we try to preserve the energy, we have to compromise security somewhere. In this paper, an optimized and secure protocol of data aggregation is proposed which preserve energy while not compromising the security. Encrypting the data only at leaf nodes, use of privacy homomorphism technique and slicing the data ensures secure and accurate data aggregation. Our theoretical analysis and simulation shows that our protocol is more energy efficient and secure than existing protocol EESSDA.

Keywords:

Data Aggregation, Security, Secure Data Aggregation Protocol, WSN,OSDAP, EESSDA

INTRODUCTION

Energy consumption is one of the main issues in wireless sensor network because once node is dead, other issues have no meaning. Extensive research has been done by different researchers to minimize the energy consumption of nodes after they have been deployed. Sensor nodes in WSN are very close to each other and therefore nodes residing in the nearby area sense and transmit similar data thereby wasting bandwidth and energy both. In [1], experiments shows that energy consumption in transmission of single bit is same as in computing 800 instructions. Although there are many solutions proposed by different authors but the most trusted one is data aggregation. Data aggregation [2-6] is a process of eliminating the redundant data in network by aggregating the data before transmission, thus saving energy and increasing the overall network lifetime.

Whereas aggregation reduces consumption of energy , different attacks[7] on this aggregated data put question mark on the accuracy of data . There must be some provisions for protection of the nodes from different attacks. A compromised sensor node generates false reading and aggregation result. Base station(Sink) is not capable of detecting the presence of compromised node because attacker presents themselves in a manner that base station easily accept their incorrect results also. Therefore it becomes necessary to employ security with data aggregation so as to achieve data accuracy, integrity, confidentiality and authentication.

RELATED WORK

Different protocols have been proposed by different authors to achieve security in data aggregation. Hu & Evans in *Secure Aggregation for Wireless Network[8],* used a lightweight security mechanism for the detection of misbehavior in nodes effectively. In *SIA[9]* authors propose a mechanism to resist the stealthy attack in which attacker forces the user to accept the false readings resulting wrong aggregation values. In *ESPDA[10]*, author proposed a protocol to provide secure communication between nodes while consuming very less energy. This protocol is based on clustering technique and it uses the concept of pattern matching. A pattern seed is broadcasted by cluster-head to the sensor nodes. In return, these nodes sent back the corresponding pattern code to the cluster head. S. Rappaport in *SecureDAV[11]*, improve the data integrity by signing the data after aggregation. Clustering mechanism is used for data aggregation and for establishing the cluster keys, an elliptic curve cryptosystems(ECC) was applied. H. Ozgur Sanli et. al. in *SRDA[12]*, use the concept of differential data. Average of some previous readings by sensor nodes was taken as a reference and they calculate the difference between this value and the actual data currently sensed by the nodes. This value is sent to the cluster head rather than the raw data. SRDA provides data authentication, confidentiality and freshness. In *CDA [13]*, first time aggregation was applied on data which was already encrypted. Authors used a homomorphic encryption scheme which is both additive and multiplicative in nature. Yi Yang et. al. in *SDAP[14]* proposed a protocol that aggregates the data hop-by-hop and follows the principles of commit/attest and divide/conquer. Suat Ozdemir in *Secure and Reliable Data Aggregation for Wireless Sensor Networks [15]*, ensures the transmission of aggregated data which is secure even if some node are compromised by the attackers. Miloud Bagaa et. al. propose *SEDAN[16]* in which each node in the network is having the capability of immediately verifying the aggregation of the next neighbor and the integrity of its two hops neighbor's data. Thus the

bogus data is controlled and comparatively less amount of energy is used in data transmission. Hani Alzaid et. al. in *RSDA [17]* proposed another protocol that integrates reputation system in data aggregation functionalities so as to upgrade the system lifetime and the exactness of aggregated information. A. S. Poornima, B. B. Amberker in *SEEDA[18]* provides end to end privacy for the data. There are two confidentiality requirements considered in this protocol. First is generic confidentiality i.e. those sensor nodes which are not participating actively in data aggregation process, do not have access to the data. Second is end-to-end confidentiality i.e. sensors nodes which are actively participating in aggregation mechanism are having no access to the data which is already aggregated. Hongjuan Li, Kai Lin, and Keqiu Li in *EEHA[19]*, design a protocol that prevents the private data readings sensed by nodes from being disclosed and thus achieve accurate data aggregation results. The main focus of this protocol was to protect the data from eavesdropping attack. Chien-Ming Chen et.al. in *RCDA[20]* propose a protocol in which the sink node is having the capability of recovering sensed data even after the process of data aggregation. Joyce Jose et. al. in *PEPPDA[21]* provide an energy efficient and secure scheme for data aggregation that guarantees the data freshness, its authenticity and privacy. Also the aggregated data is accurate and maintains integrity, confidentiality. Taochun Wang et. al. presented *EESSDA[22]* in which secure channel is established between the sensor nodes before data transmission. After that data is divided into pieces using the slicing technique and transmitted to the neighbors, thus making the transmission more secure. This protocol adopts a random key distribution mechanism.

SYSTEM MODEL

Network Model

WSN is a network of small nodes in large numbers having very less number of resources and targeted for a specific application. Considering a tree structure for the network, three types of nodes are there ie nodes at leaf level, nodes at intermediate levels and sink node. Data sensed by leaf nodes is partitioned into pieces (Slicing) for privacy preservation, encrypted using homomorphic encryption and then transferred to parent nodes. The intermediate nodes get the encrypted data from their corresponding child nodes and without decrypting this data, aggregates it with its own sensed data. This is now the sink node's responsibility to process the received aggregated data, generate the required result for the targeted application and verify the integrity of data.

Attack Model

It is dangerous to transmit the data from one node to other in sensor network without implementing any security features because it can lead to the leakage of important data to the attackers. Also intruders can deploy their own sensor nodes which becomes the part of sensor network and leads to the incorrect results of aggregation. One method to avoid this is to provide a unique identifier to each node and a secret key before deployment so that to preserve authentication property of security.

Design Objectives

We are assuming that base station is having enough resources in terms of hardware and is trustworthy (cannot be attacked). The sensor nodes on the other hand are very poor in resources and are not trustworthy(can be attacked). Data aggregation technique used must be energy saving and achieve the following security goals.

Data Privacy:

Privacy of data is must in secure applications because compromised data leads to the incorrect aggregation results. For privacy of data in our protocol, slicing operation is performed on the data sensed by the nodes at leaf level.

Data Confidentiality:

Data is assumed to be confidential only if an authorized node can view this. We authorize only base station to be such node to which the aggregated data(partially or fully) is visible. To achieve this, we use end to end encryption. In the intermediate levels, nowhere data is being decrypted.

Data Authentication:

Data is assumed to be authentic only if the claimed sender has sent the data otherwise it is assumed as unauthenticated and ignored. We achieve data authentication using a unique identifier (UID) and a separate secret key for each sensor node (NSK).

Energy Efficiency:

In WSNs, nodes are having very limited battery life and therefore it becomes necessary for the data aggregation process to be energy efficient without compromising the security goals. More the number of encryption and decryption processes, larger the consumption of energy. Therefore, We use end to end encryption for less energy consumption and no decryption is being done at intermediate nodes.

Aggregation Accuracy:

All critical decisions made by the base station depends upon the accuracy of aggregation result. It is worth nothing if we save energy but inaccurate data is received at sink node . Avoiding the decryption process at each intermediate level improves the accuracy of data as there remains less chances for an intruder to falsify the data.

Data Integrity:

Data integrity is the property which ensures that data will not be altered while transmission from the sensor nodes to the sink node. MAC (message authentication codes) are used for this purpose. At the base station, MAC generated by sensor nodes is compared with the MAC generated from the aggregation result at sink node. If both MACs are same, integrity is achieved otherwise compromised.

Data Freshness:

It guarantees that final data that reaches the base station is fresh and there is no effect of the replay attack. We change the encryption keys for each session for achieving data freshness.

PROPOSED SCHEME: OSDAP

Our proposed protocol OSDAP consists of seven steps:

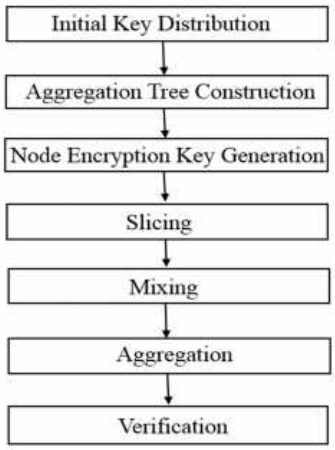

Figure 1. OSDAP Steps

Initial Key Distribution

The key distribution used is similar to the ESPDA [10]. Before deployment , we assign a secret key which is common (CSK), a secret key which is node specific (NSK), a unique id (UID) to each node.

CSK	NSK	UID

Figure 2. Keys Stored in a Node

Similarly CSK, a session key (SK) and combination of UID and node specific secret key (NSK+UID) of all nodes is kept with base station before deployment.

CSK	SK	NSK+UID

Figure 3. Secret Information Stored in Sink

Aggregation Tree Construction

Constructing an aggregation tree is one most common method of data aggregation. TAG[23] protocol for tree construction is used here. According to this, base station is the root of the tree. All the sensor nodes are connected to each other is such a way that each node has the shortest path to the root. All the leaf nodes will sense the data and forward it to the aggregator node which aggregates the received data with its own data and further transmit it to the upper level nodes upto the sink.

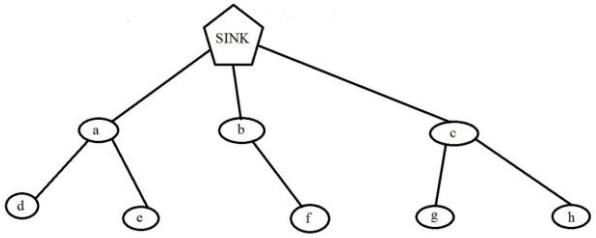

Figure 4. Aggregation Tree Construction

Node Encryption Key Generation

When a node senses some data from the environment , it request for the same to the sink node. A session key(SK) is then generated by sink node for that particular node and encrypted with the CSK before broadcasting. Each node will receive the encrypted session key and then decrypt it with same CSK. Now an encryption key is generated by each node by applying XOR operation on the session key obtained from base station and node specific key of the node itself. ie EKi=SK+NSKi

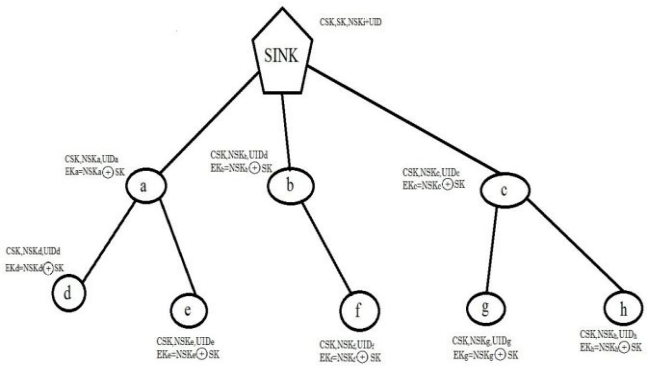

Figure 5. Encryption Key at each node

Slicing

For privacy preservation, slicing is applied to the data sensed by leaf nodes. Before slicing, MAC of this data is calculated for integrity check at base station. Slicing technique similar to EPSDA [24] is used in which data sensed by each leaf node is divided into m different slices. After that, using the encryption key generated in step 3 above, encryption is performed on all slices of data. After keeping one slice with itself, remaining m-1slices are appended with UID of corresponding node and sent to all neighbors. This one slice

which is kept with the node itself is only transmitted to the parent node(aggregator)with other slices.

Mixing

Once the slicing operation is done, every node waits before mixing for a pre defined time period. This time period is given to make it sure that all the transmissions of slices between different nodes are completed. After that mixing operation is performed by every leaf node.

Aggregation

In aggregation process, initially each leaf node sums up all slices received from different neighbors, encrypts it and transfer it to its parent node along with its own encrypted slice , MAC calculated before slicing and UID. Every node at the intermediate level then calculates the aggregate of encrypted data received from child nodes, encrypted slice received from the leaf nodes and its own sensed data. This aggregated result is further encrypted and transmitted to the upper layers till it reaches the sink node. Since each intermediate node has to receive data from all child nodes as well as possibly from leaf node, it requires more time to wait than its child nodes. This time interval difference dt can be set at the time of tree construction and using this time out t_o of each node can be computed. t_o is the maximum time upto which a node in the network waits for receiving all the required data for aggregation. When t_o finishes, this partially aggregated data(along with MAC) is transferred to the upper aggregator node until it reaches to the base station. Base station now, after seeing the UID, identifies the node specific secret key and decrypts the data using the decryption key which can be generated by XORing the NSK with the SK.

Verification

After generating all decryption keys, base station sums up them and calculates the MAC of this aggregated result. Now a comparison is made between this calculated MAC and network's MAC. If there is a minor difference between these two, data is verified and accepted otherwise rejected.

OSDAP ALGORITHM

1. Assign CSK, NSK and UID to every sensor node except sink node.
2. Assign CSK, SK and UID+CSK pairs of all nodes to the sink node.
3. Using TAG, construct an aggregation tree.
4. The sink node broadcasts session key (SK) to the network after encrypting it using CSK.
5. Each node in the network receives this encrypted SK from the base station.
6. Each node now decrypts SK with the help of CSK and generates new encryption key by XORing SK with NSK

$$EKi = SKi \oplus NSKi$$

7. A time interval difference dt is set between child nodes and parent nodes to calculate the timeout t_o for each node.
8. if LeafNode then
   i. Calculate MAC of sensed data and make m slices of data.
   ii. using EKi , each node encrypts the slices.
   iii. m-1 slices and UID are appended together and transmitted to neighboring nodes excluding parent node
   iv. remaining one slice is held with the node itself.
   v. if t_o
      a. Mixing is performed by every node. In this operation each node sums up all received slices and slice of its own sensed data which was kept by the node in step iv.
      b. Transmit this data along with MAC to the parent nodes.

   end if
9. if IntermediateNode then

   Using privacy homomorphism technique, calculate aggregated MAC for the sensed data and child node's data.

   ii. Apply encryption to the data sensed by the node using EKi

   iii. if t_o
      a. Data encrypted in step ii is summed up with the encrypted aggregated data received from the child nodes.
      b. This data with the node's UID is transmitted to the upper aggregator until it reaches to the sink.

   end if
10. if SinkNode then
    i. seeing the UID identifies the node specific secret key and decrypts the data using the decryption key which can be generated by XORing the NSK with the SK.
    ii. After generating all decryption keys, sums up them and calculate the MAC of this aggregated result.

iii. Compare this calculated MAC and network's MAC. If there is a minor difference between these two, data is verified and accepted otherwise rejected.

end if

SIMULATION SCENARIO

OSDAP protocol is implemented using MATLAB. A Window 7 PC with i3 Processor and 2GB RAM is used for simulation. Network of 100 sensor nodes is deployed randomly over an area of 100 x 100 m^2 . The maximum range over which a node can transmit the data is 50 m.. The energy consumed in transmitting 1 bit of data is assumed to be 0.72µ Joules and for receiving same amount of data, this consumption is 0.81µ Joules. For encrypting/decrypting 1 bit of data, energy consumption is assumed to be 0.8892µ Joules.

Performance Analysis

Computational Overhead

In EESSDA, the computational overhead is due to large number of power consuming encryption and decryption operations while getting keys from the key pool, checking shared keys, and finally establishing secure channels between nodes. In OSDAP protocol, encryption is done at only leaf nodes and there is no decryption process in the intermediate levels. Therefore less number of encryption and decryption operations are performed in the proposed protocol. The total number of computations in existing is 579 and in proposed is 202.

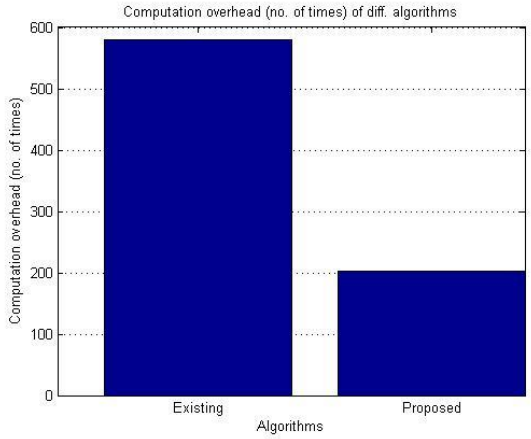

Figure 6. Comparison of Computational Overhead of Existing & Proposed Protocols

Communication Overhead

In wireless sensor networks, energy consumed in data transmission is much more than computation. Communication overhead can be determined by the number of transmissions which takes place in a given time period. Larger the number of transmissions, higher is the communication overhead. In EESSDA, communication occurs in two stages for a single data aggregation process. First for secure channel establishments and then for actual data transmission. While in proposed algorithm, no communication overheads for secure channel establishments, but same amount of communication is required for encryption key generation for the first time. So the data transmission in both the protocols is almost same for single message. For more than one messages, our protocol is having less communication overhead because once encryption key is generated, it can be use for all messages but in case of EESSDA, each time a new secure channel has to be built before transmission.

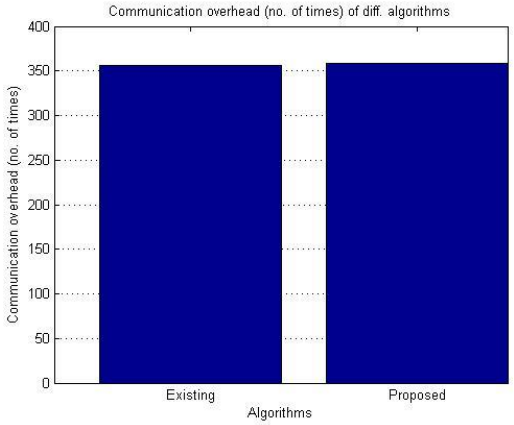

Figure 7. Comparison of Communication Overhead for single data aggregation

Energy Consumption

Wireless sensor network nodes operate on limited battery power , therefore energy consumption of sensor nodes is a very important factor which must be taken into consideration while designing the network. The Figure 8 shows the comparison of energy consumption of existing and proposed protocol with respect to number of encryption and decryption operations performed. The energy consumed in EESSDA is 6900.19µ J and in OSDAP is 711.36µ J.

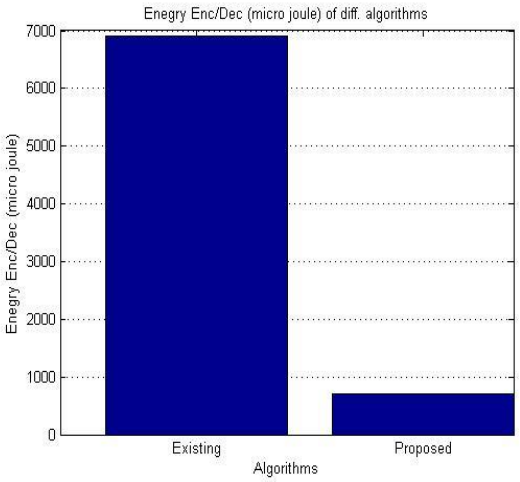

Figure 8. Comparison of Energy Consumption with respect to Number of Encryption/Decryption Operations Used

The Figure 9 below shows the comparison of energy consumption according to the number of transmissions and receptions. The number of data transmissions in both the protocols is almost same. In EESSDA it is 5446.80µ J while in our case it is 5477.40µ J

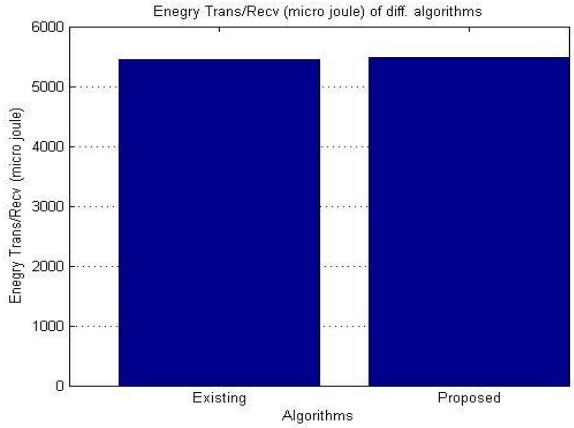

Figure 9. Comparison of Energy Consumption with respect to Number of Transmissions/Receptions Required

In Figure 10 overall energy consumption of the existing and proposed protocols are compared. The overall energy consumed in OSDAP is 6188.76µ J, which is less as compared to EESSDA where it is 12346.99µ J. Thus the overall energy consumption is reduced to 50.12%.

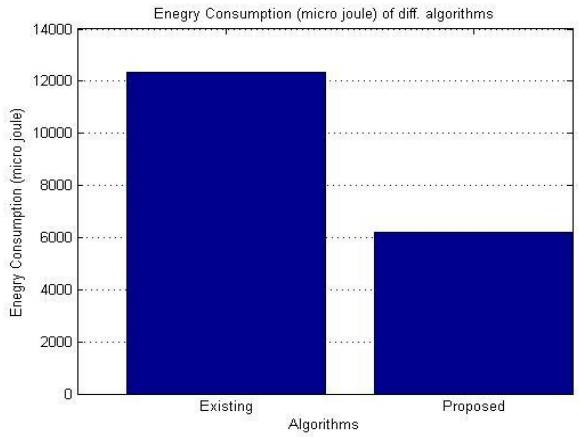

Figure 10. Comparison of Overall Energy Consumption

Table 1 Comparison of Different Parameters of Existing and Proposed Protocols

	Comm. Overhead (Key distribution)	Computational Overhead	Energy Consumption (Encryption/ Decryption)	Energy Consumption (Transmission / Reception)	Overall Energy Consumption
EESSDA	356	579	6900.19	5446.80	12346.99
OSDAP	358	202	711.36	5477.40	6188.76
% Reduction	-	65.11%	89.6%	-	49.87%

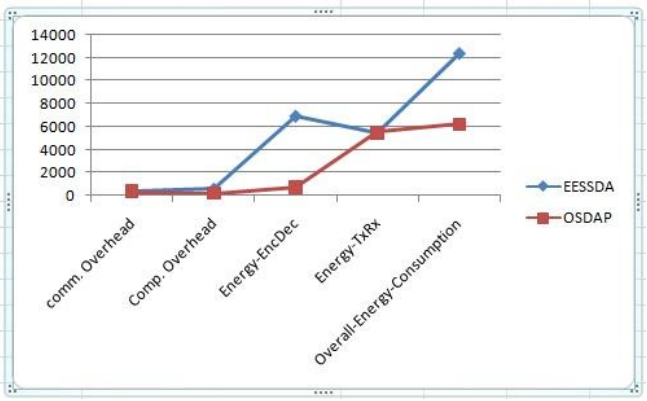

Figure 11. Overall Comparison

CONCLUSION

The proposed protocol OSDAP is a secure and energy efficient protocol for wireless sensor networks. The encryption algorithm used is based on privacy homomorphism in which aggregation can be performed even on encrypted data, thus making the protocol more secure . It provides all the essential security requirements with minimal communication and computational overheads. We compared the performance of existing protocol EESSDA and proposed protocol OSDAP. The simulation results show that our protocol reduces the energy usage by 50.12% .

FUTURE SCOPE

In proposed protocol OSDAP, we have used end-to-end data aggregation in which nodes between leaf and sink aggregate the data without decrypting it and ensuring end-to-end data confidentiality, less computation and transmission cost.. But this technique also has a limitation. In this technique we do not decrypt the received messages, rather we concatenate the received encrypted messages. So the length of the packet to be transmitted is increased, which further increases the bandwidth required for transmitting data packets. Usually we have limited bandwidth available for

communication, so in such case the packet delivery ratio will reduce, which results in low throughput. In future, some work can be done on this end-to-end data aggregation technique to remove this limitation.

REFERENCES

[1] R. Szewczyk and A. Ferencz, "Energy Implications of Network Sensor Designs", *Berkeley Wireless Research Center Report*, Berkeley, California, USA, 2000.

[2] D. Estrin et. al., "Next Century Challenges: Scalable Coordination in Sensor Networks", in Proceedings of *ACM Mobicom*, Seattle, Washington, USA, ACM, August 1999, pp. 263-270.

[3] J. Heidemann et. al., "Building Efficient Wireless Sensor Networks with Low-level Naming", in: Proceedings of the 18th *ACM Symposium on Operating Systems Principles*, 2001, pp. 146–159.

[4] B. Krishnamachari, D. Estrin, S. Wicker, "The Impact of Data Aggregation in Wireless Sensor Networks", in Proceedings of the *ICDCS Workshops*, 2002, pp. 575–578.

[5] Y. Yu, B. Krishnamachari, V.K. Prasanna, "Energy-Latency Tradeoffs for Data Gathering in Wireless Sensor Networks", in Proceedings of the 23rd Conference of *IEEE Communication Society (INFOCOM)*, Hong Kong, SAR China, March 2004.

[6] K. Akkaya, M. Demirbas, and R. S. Aygun, "The Impact of Data Aggregation on the Performance of Wireless Sensor Networks", *Wireless Communications and Mobile Computing*, Volume 8, no. 2, pp. 171–193, 2008

[7] C. Karlof and D. Wagner," Secure Routing in Wireless Sensor Networks: Attacks and Countermeasures," *AdHoc Networks Journal*, Volume 1, no. 2–3,pp. 293–315, September 2003.

[8] Hu L., Evans D., "Secure Aggregation for Wireless Networks", in *International Symposium on Applications and the Internet* ,Orlando, Florida, USA, pp. 384-391, 27-31 January 2003.

[9] Przydatek B., Song D., Perrig A., "SIA: Secure Information Aggregation in Sensor Networks", in proceedings of the *1st International Conference on Embedded Networked Sensor Systems*, Los Angeles, CA, USA, pp. 255-265, November 05 – 07, 2003.

[10] Cam H. et al. , "ESPDA: Energy-Efficient and Secure Pattern-Based Data Aggregation for Wireless Sensor Networks", in *Computer Communications, Elsevier*, Volume 29, Issue 4, pp. 446–455, February 2006.

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

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

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

[14] Yang Y. et al., "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,Article No. 18, New York, USA, July 2008.

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

[16] Bagaa M. et al., "SEDAN: Secure and Efficient Protocol for Data Aggregation in Wireless Sensor Networks", in proceedings of *32nd IEEE Conference on Local Computer Networks,* pp. 1053-1060, 15-18 October 2007.

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

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

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

[20] Chen C. M. et al., "RCDA: Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks", in *IEEE Transactions on Parallel and Distributed Systems*, Volume 23, Issue 4, pp. 727-734,August 2011.

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

[22] Wang T., Qin X., Liu L., "An Energy-Efficient and Scalable Secure Data Aggregation for Wireless Sensor Networks" in *International Journal of Distributed Sensor Networks, Hindawi Publications*, Article ID 843485, Volume 2013(2013).

[23] S. Madden, M. J. Franklin, J. M. Hellerstein et al., "TAG: a tiny aggregation service for ad-hoc sensor networks," *ACM SIGOPS Operating Systems Review*, Volume 36, no. I, pp. 131–146, 2002.

[24] [J Jose, M Princy, J Jose ," EPSDA: Energy Efficient Privacy preserving Secure Data Aggregation for Wireless Sensor Networks" in *International Journal of Security and Its Applications* Volume 7, No. 4, July, 2013