International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

An Approach for Improving Security and Efficient Data Aggregation in Wireless Sensor Network

Anil A Kumbar[1], Sagarika[2]

[1]PG[M.tech]Scholar, Department of Computer Science & Engineering, CMRIT, Bangalore-37, India

[2]Assistant Professor, Dept of Computer Science & Engineering, CMRIT, Bangalore-37, India

Abstract: ***Sensor networks are collection of sensor nodes which co-operatively send sensed data to the sink(receiver). As sensor nodes battery driven, an efficient utilization of power is essential in order to use networks for long duration hence it is needed to reduce data traffic inside sensor networks, so amount of data should be reduced and send it to the base station. The aim of data aggregation algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. Wireless sensor networks (WSN) offer an increasingly Sensor nodes need less power for processing as compared to transmitting data. It is mainly deployed in hostile environments, so security is needed. The sensor networks are used in civilian areas, including environment and habitat monitoring, traffic control, home and industrial automation, healthcare application, accident reporting and in security applications such as in military applications. Our main aim 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.***

Keywords: Wireless Sensor Networks; Data Privacy; Data confidentiality; Message authentication; End to end encrypted data aggregation; Privacy homomorphism; Accuracy; Data aggregation; Hop by hop encrypted data aggregation; Data freshness.

1. Introduction

Sensor networks composed of small and cost effective sensing devices equipped with wireless radio transceiver for environment monitoring have become feasible. The key advantage of using these small devices to monitor the environment is that it does not require infrastructure such as electric mains for power supply and wired lines for Internet connections to collect data, nor need human interaction while deploying.

The sensors in the network act as "sources" which detect environmental events and push relevant data to the appropriate subscriber sinks. For example, there may be a sink that is interested in a particular spatio-temporal phenomenon ("does the temperature ever exceed 70 degrees in area A between 10am and 11am ?"). During the given time interval all sensors in the corresponding spatial portion of the network act as event based publishers. They publish information toward the subscriber sink.

However since various sensor nodes often detect common phenomena, there is likely to be some redundancy in the data the various sources communicate to a particular sink. In-network filtering and processing techniques can help conserve the scarce energy resources. So to reduce the number and amount of data transmission, use the data aggregation [2] techniques. It is the process of gathering data from the sensor nodes and aggregate this data using aggregation functions such as MAX, MIN, SUM, AVERAGE, HISTOGRAM, etc , then send the partially aggregated result to the sink node. It improves the energy efficiency, thereby prolong the network lifetime. The extension of this approach is in-network data aggregation [3], which aggregates the data as it passed through the network. The proposed system uses the SUM aggregation function.

Data privacy can be defined as the process in which the adversaries or trusted participating nodes can overhear and decrypt the private data held by each sensor node. But it can still provide a mechanism to prevent them from recovering the sensitive information. To achieve privacy, it is required to protect the transmission trend of a node's private data from its neighbours, because the neighbours know the aggregated sum and the encryption key. This scheme achieves privacy preserved data aggregation through slicing and assembling operation at leaf nodes.

Data confidentiality is achieved through the end to end encrypted data aggregation, based on the homomorphic encryption algorithm. It allows arithmetic operation on encrypted data. Thereby reduces the computational power required to perform the encryption, decryption operations at the aggregator node and also reduces the communication overhead by reducing the delay during data aggregation. Another issue is message authentication. It gives assurance to the receiver that the data is coming from the trusted participating node.

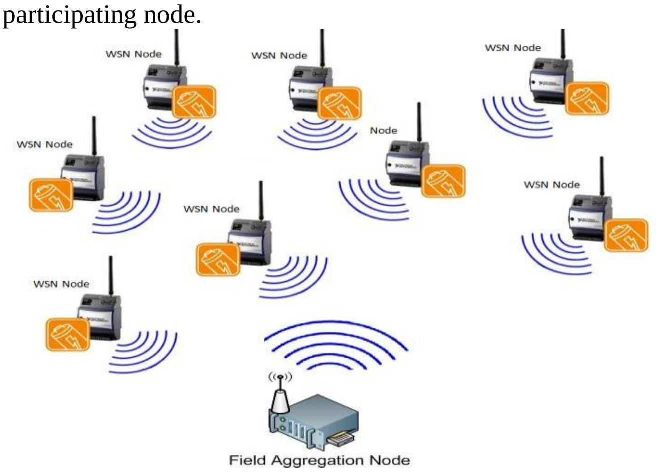

Figure 1: Wireless Sensor Network

In this scheme, message authentication is achieved using the secret key and ID pair of each node. The intruder can still attack the network by replaying old data to the network. This scheme achieves data freshness using the varying encryption key for each session.

2. Related Works

Based on the type of nodes in the sensor networks, privacy preserving protocols 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 performs sensing, aggregation and forwarding of the aggregated result to the sink. All sensor nodes can play the role of aggregator. In heterogeneous protocols, more than one type of sensor node exists and aggregator is considered as a special node. i.e, aggregator play the role for aggregation and forwarding the aggregated value to the sink, but not used for sensing.

The secure aggregation protocols are divided into two types; end to end encrypted data aggregation and hop by hop encrypted data aggregation protocols. Most of the end to end encrypted aggregation protocols achieve end to end confidentiality by allowing aggregation to be carried out on encrypted data rather than plain text in hop by hop encrypted data aggregation protocols. So the end to end privacy can achieve in the end to end protocols but, compromising of aggregator leads to the loss of data privacy at aggregator in hop by hop encrypted protocols.

The homogeneous and heterogeneous protocols are of different types: perturbation, privacy homomorphism, hybrid, shuffling. In perturbation, each sensed data is customized using the encryption key and public or private seed generated by randomization technique [5] to hide the data. In privacy homomorphism [6], the arithmetic operations are done on the encrypted data without decryption, so it reduces the energy consumption at the aggregator node. The hybrid approach uses more than one privacy preserving technique to achieve privacy preserving data aggregation. In shuffling each node slices its data into k number. One piece is kept on the node itself, and the remaining k-1slices are encrypted and send to the k-1 neighbors.

The ESPDA (Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks) [7] technique improves the energy efficiency by sending pattern code instead of actual data. The privacy is achieved using the end to end encryption key of each node. It also provides confidentiality and message authentication for the data.

The CDA (Concealed Data Aggregation) [8] uses the end to end encrypted aggregation using DF approach [9] to reduce high computational overhead of a hop by hop aggregation. All sensor nodes share a common encryption key with the BS. So the compromise of one sensor node leads to loss of privacy between the sensor nodes. But in EAED (Efficient Aggregation of Encrypted Data in wireless sensor network) [10], each node's share a unique key with BS. So it achieves data privacy among sensor nodes, but it is not scalable in the large network because BS wants to know the keys of all aggregated packets. So it causes the transfer of nodes ID.

The RCDA (Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks) [11] uses elliptic curve based additive privacy homomorphism technique to achieve end to end privacy and confidentiality. The recoverability of individual sensing data at the BS helps to overcome the limitation of BS on aggregation function and to verify integrity, authenticity of sensing data using the aggregated signature scheme.

The PIA (Privacy Preserving Integrity Assured data aggregation) [12] addresses the integrity assured data aggregation with efficiency and privacy as joint objective. The PIA proposed four symmetric key solutions for the single aggregator model for the integrity and privacy protection.

In the first solution, it combines the homomorphism and MAC to construct an authenticated encryption scheme for the aggregator node. It only supports aggregation functions such as average and standard deviation. The second solution uses the Order Preserving Encryption Scheme (OPES) [13] to preserve the privacy of distribution of data. OPES only verifies the integrity of comparison based aggregation. The third solution uses a Secure Hierarchical In networking Aggregation (SHIA) [14] scheme for adapting distributed. This scheme supports any aggregation function. The fourth solution used to improve the privacy and integrity of the third solution by using a logical aggregation tree within the aggregator node. It only supports decomposable functions such as mean, standard deviation, count, MIN/MAX.

3. System Model

3.1. Network Model

In this aggregation is performed on the aggregation tree routed at the BS. The BS is a powerful node with enough resources. Three types of nodes are present in a WSNs; these are Base Station (BS or sink or Query Server), the intermediate node (aggregator), and leaf node (normal sensor node). The BS is a node where the aggregation results are destined, and it is responsible for processing the received data from the sensor network and derives the meaningful information reflecting the events in the target field. The intermediate node performs sensing, aggregation and forwarding of data from the leaf node to upper aggregator or to sink depending on the type of sensor networks. The leaf node performs sensing, aggregation and forwarding of data. In addition it performs the slicing and assembling operations to achieve privacy preserved aggregation. This scheme focuses on the SUM data aggregation function.

International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

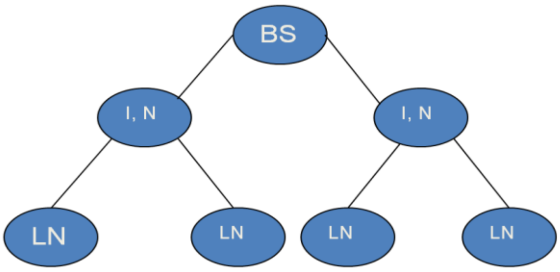

Figure 2: Aggregation tree

3.2. Attack Model

Security is becoming a more and more important concern with the extensive application of sensor networks. A malicious attacker can launch a variety of attacks to break the data security. And privacy concern is one of the major obstacles to apply the wireless sensor network to civilian applications. In this paper, we mainly focus on the defense of eavesdropping to protect data privacy in wireless sensor networks. One approach to protect the sensor network against an intruder sensor is to initialize each node with a Unique identifier and a secret key before it is deployed.

In an eavesdropping attack, the attacker tries to overhear the transmission channel to get the sensitive information. The eavesdroppers are two types: inside eavesdroppers and outside eavesdroppers. The inside eavesdroppers are intruder or compromising node. They can get the private data destined for others. But, by using privacy preservation technique, prevent them from getting the private data of individual sensors. The outside eavesdroppers can prevent by using encryption.

4. Security and Efficiency for Aggregation of Data

It guarantees the energy efficiency, accuracy, privacy preservation, end to end confidentiality, data freshness and message authentication during data aggregation. There are four steps in our scheme, i.e, aggregation tree construction, slicing, mixing and aggregation.

4.1. Aggregation Tree Construction

A common technique for data aggregation is to build an aggregation tree which is the directed tree formed by the union of all the paths from the sensor nodes to the base station. These paths may be arbitrarily chosen and are not necessarily shortest paths. The optimization of the aggregation tree structure is out of the scope of this paper. There are various methods for constructing the aggregation tree according to different application requirements. One method for constructing an aggregation tree is described in TAG [10].

4.2. Slicing

Here the slicing operation is done on the leaf node only. First, each leaf node slices its sensed data into m number of pieces. Then encrypt each slice using the encryption key generated by the node after it receives the session key from the BS. One of the m encrypted slices is kept on the node itself and the remaining m-1 encrypted slices are appended with the node ID and transmitted to m-1 neighbour nodes within the h hop (for a dense network h=1) except the encrypted slices to its parent. The encrypted slice with its ID to the parent is appended with the encrypted slice kept on the node, and it is transmitted to its parent with the aggregation result from leaf node.

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. So, in order to avoid this problem take M as large M=n*t, where n is the number of nodes and t=max (di), i.e, maximal value, which may appear in the measurement. Figure3 shows the slicing operation. The leaf node sliced its data into m pieces (m=3) and encrypted all the m pieces using the encryption key of leaf nodes. One of the encrypted slices is kept in the node itself and the remaining m-1 pieces are appended with node ID and sent to m-1 neighbour nodes except the piece to its parent. Once slicing of data is done each sliced data is encrypted using public key encryption algorithm(RSA). And each encrypted slice is sent to its parent node and its siblings.

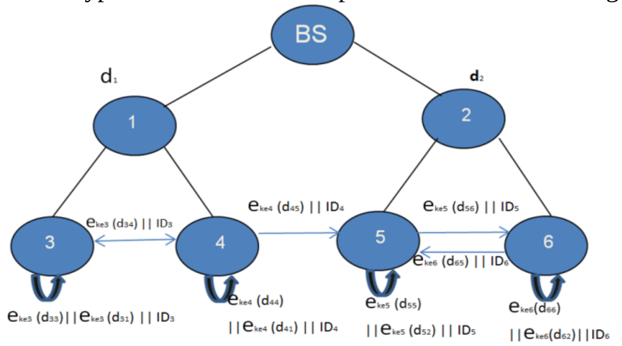

Figure 3: Slicing

4.3. Mixing

First, all leaves of the aggregation tree wait for certain time, which guarantees that all slices are received.

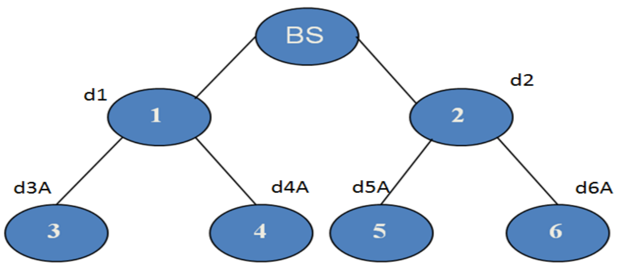

Figure 4: Mixing

Then, each leaf decrypts the data using its shared key with the sender, sums up all the received slices and the slice left by itself to get a new result ri . Figure. 4 shows the mixing step on leaf nodes.

Here,

$d_{3A}=EK_{e3}(d_{33}) \| EK_{e3}(d_{31}) \| ID_3\|ID_3$

$d_{4A}=EK_{e4}(d_{44}) + EK_{e3}(d_{34}) \| ID_3\| EK_{e4}(d_{41}) \| ID_4\|ID_4$

$d_{5A}=EK_{e5}(d_{55}) +EK_{e4}(d_{45}) + EK_{e6}(d_{65}) \| ID_4\| ID_6\|EK_{e5}(d_{52}) \| ID_5\|ID_5$

$d_{6A}=EK_{e6}(d_{66}) +EK_{e5}(d_{56}) \| ID_5\| EK_{e6}(d_{62}) \| ID_6\|ID_6$

4.4. Aggregation

During data aggregation, each leaf node sends the aggregated result and the encrypted slice appended with the encrypted slice kept in the node if any, to its parent node, after appending its ID. After receiving the aggregated result from all its child nodes, the intermediate node encrypts its data using its own encryption key and sum up it with the aggregated result received from all its child nodes using privacy homomorphism. It then appends with the intermediate node ID and send to the upper aggregator or sink. Each intermediate node takes longer time to aggregate than its child nodes. So the difference between the times is measured as Δt. Then each node can find out its timeout ti. The timeout ti is elapsed, the partially aggregated result is sent to upper aggregator. The aggregation result goes level by level and finally reaches the BS. The final aggregation result fA is the encrypted sum of all sensed data. After receiving the aggregated encrypted result, the BS decrypts it by using the decryption key and generates the aggregated result fR.

$$d1A=d3A+d4A+EKe1\ (d1) + EKe3\ (d31) + EKe4\ (d41)\ ||ID1$$
$$d2A=d5A+d6A+EKe2\ (d2) + EKe5\ (d52) + EKe6\ (d62)\ ||ID2$$
$$fR = d1A+d2A$$

5. Simulation Results

We consider a wireless sensor network with 15 sensor nodes. These are randomly deployed over an area of 1500 m × 500 m. One of the node is taken as BS and the remaining nodes form a tree routed at the BS. Each sensor node is assigned with 100J of energy. The other parameters are transmission power =0.660W, receiving power =0.395W, idle power =0.035W, simulation time =20ms.

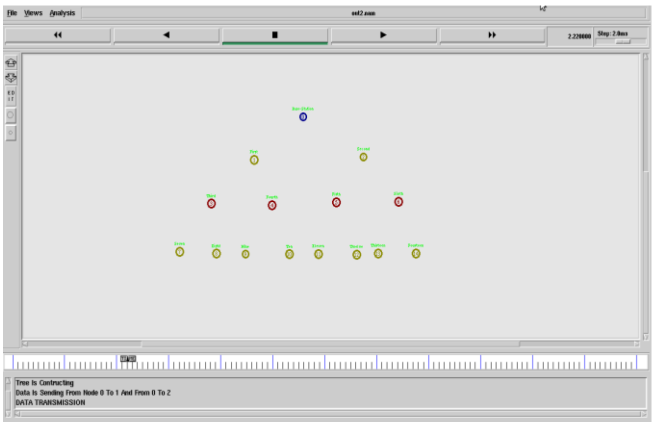

Figure 5: Aggregation tree structure.

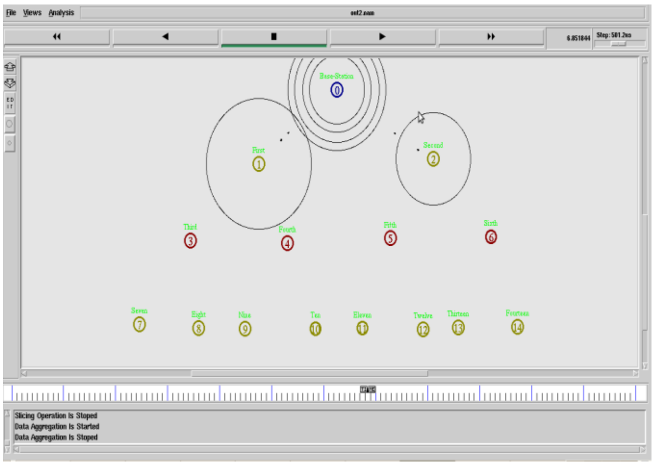

Figure 6: Aggregation at Base station

- **Throughput**

Network throughput is the average rate of successful message delivery over a communication channel. This data may be delivered over a logical or physical link, or may pass through a certain network node. The throughput is generally measured in bits per second (bit/s or bps), and sometimes in data packets per time slot or data packets per second.

Throughput is the amount of data received by the destination. The Average Throughput is the throughput per unit of time

Example: Suppose a TCP receiver receives 60 M Bytes of data in 1 min, then:
1. The throughput of the period is 60 M Bytes
2. The average throughput is 60 M Bytes/min or 1 M Bytes/sec

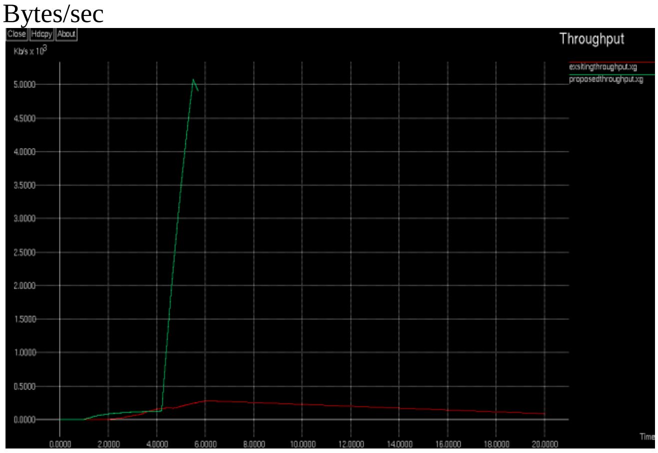

Figure 7: Graph showing comparison of throughput between existing and proposed system.

- **Packet Delivery Ratio**

The ratio of the number of delivered data packet to the destination. This illustrates the level of Packet delivery ratio of the number of delivered data packet t data to the destination.

∑Number of packet receive / ∑ Number of packet send

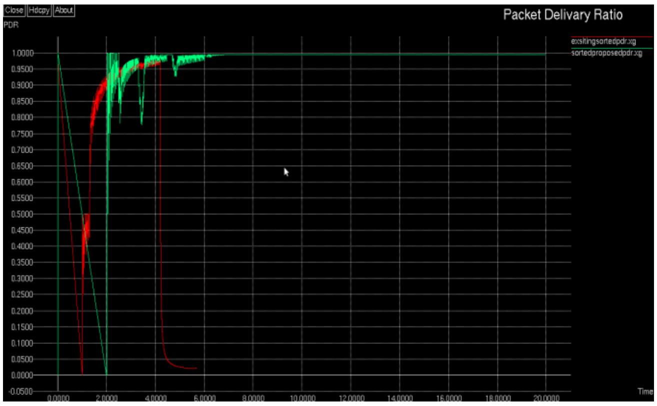

Figure 8: Graph showing comparison of Packet Delivery Ratio between existing and proposed system.

- **Control Overhead**

Sending a payload of data (reliably) over a communications network requires sending more than just the desired payload data, itself. It also involves sending various controls and signaling data (TCP) required achieving the reliable transmission of the desired data in question. The control signaling is overhead.

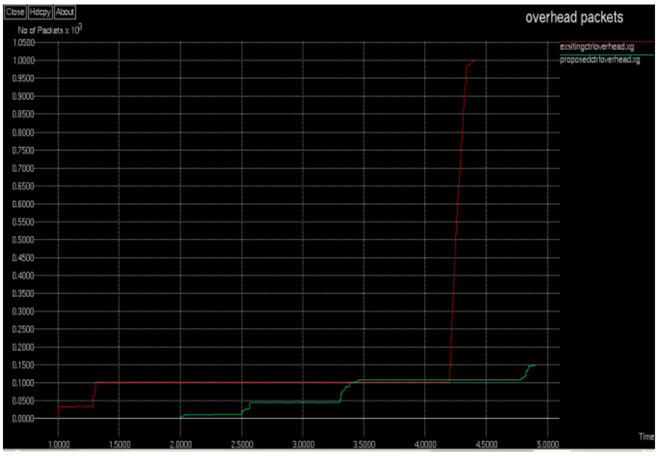

Figure 9: Graph showing comparison of control overhead between existing and proposed system.

6. Conclusion

The data collected from sensor network are correlated, so the direct transmission of data from the sensor node to sink wastes too much energy. So to reduce the number and amount of data transmission, use the data aggregation techniques. Here in This scheme tree based approach is used and also in this scheme slicing and mixing operation is used so security is increased by encrypting the sliced data. And it also provides privacy, confidentiality, authentication, data freshness and accuracy at low communication and computation overhead during data aggregation. Thus it overcomes the energy burden imposed by the hop by hop based privacy preservation protocol on an aggregation node by allowing aggregation on encrypted data.

Besides, future works will also consider the heterogeneous wireless sensor network and also planning to include the integrity checking mechanism into our system. For verifying the correctness of the final aggregated result without introducing a significant overhead.

References

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

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

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

[4] Hongjuan Li, Kai Lin, Kequi Li, "Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks", Computer Communication,34 (2011);591-597.

[5] Kargupta, H.; Datta, Q.W.S.; Sivakumar, K, "On the privacy preserving properties of random data perturbation techniques", In Proceedings of the IEEE International Conference on Data Mining, Melbourne, FL, USA, November 19–22, 2003; pp. 99–106.

[6] S. Peter, D. Westhoff, and C. Castelluccia, "A Survey on the Encryption of Convergecast-Traffic with In-Network Processing", IEEE Transactions on Dependable and Secure Computing, vol. 7, no.1.

[7] Hassan Cam, Suat Ozdemir, Prashant Nair, Devasenapathy Muthuavinashiappan, H.Ozgur Sanli, "Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks".

[8] 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,May 2005.

[9] Domingo-Ferrer J. A provably secure additive and multiplicative privacy homomorphism. In Proceedings of the 5th International Conference on Information Security, Sao Paulo, Brazil, September 30–October 2, 2002;pp. 471–483.

[10] Castelluccia, C.; Mykletun, E.; Tsudik, G. 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, San Diego, CA, USA, July 17–21, 2005; pp. 109–117.

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

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

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

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

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