Privacy and Integrity Preserving Secure Data Aggregation in Wireless Sensor Networks

Vivaksha J. Jariwala[1] and Devesh C. Jinwala[2]

[1] C. K. Pithawalla College of Engineering and Technology, Surat, India
[2] S. V. National Institute of Technology, Surat, India

Abstract. The Wireless Sensor Networks (WSNs) protocols commonly use in-network processing to optimize the communication costs. In-network processing involves processing of the sensed data on-the-fly during the course of the communication to the base station. However, due to the fusion of the data items sourced at different nodes into a single one, the security of the aggregated data as well as that of the aggregating node, demands critical investigation.One of the approaches to ensure secure data aggregation is to use encrypted sensor data for processing, using homomorphic encryption. Our research here is aimed to propose an approach that uses homomorphic encryption and appropriate data integrity mechanisms to offer confidentiality, privacy and data integrityfor secure data aggregation in wireless sensor networks.

1 Introduction

Wireless sensor network is used for numerous applications including military surveillance, facility monitoring and environmental monitoring [1]. Wireless sensor networks (WSNs) often consist of a large number of inexpensive, low-powered sensing devices with limited memory, computational and communication capabilities [1]. Wireless sensor networks have number of sensor nodes that are able to communicate with external node or the base station of network. The sensors communicate between themselves to form a communication network like single-hop network, multi-hop network or a hierarchical network with several clusters and cluster heads. Sensor nodes frequently sense the data, process it and then transmit to the base station. It is inefficient for all the sensor nodes to transmit the data directly to the base station as these sensor nodes are energy constrained. If all the sensor nodes transmit their data to the base station, precious energy of the wireless sensor networks is wasted [2].

Generally, data sensed by neighbouring sensor nodes are always highly correlated and redundant. In addition to that, the amount of data generated in large sensor networks is usually enormous for the base station to process. Hence, we need methods for combining data into high quality information at the sensors or intermediate nodes that can reduce the number of packets transmitted to the base station resulting in conservation of energy and bandwidth. However, primarily due to the severe constraints and secondarily due to the inherent increased costs in the communications relative to that in processing, the WSNs follow in-network processing, wherein the emphasis is on on-the-fly processing of the data packets before being communicated eventually to

Jariwala V. and Jinwala D..
Privacy and Integrity Preserving Secure Data Aggregation in Wireless Sensor Networks.
DOI: 10.5220/0004980100490060
In *Proceedings of the 11th International Workshop on Security in Information Systems* (WOSIS-2014), pages 49-60
ISBN: 978-989-758-031-4

the base station [3]. In in-network processing of the data, each sensor node senses the required measurements and sends the data value to another node up in the hierarchy called the aggregator node [1]. The aggregator node collects measurements from different sensor nodes using that it generates a single representational aggregated value by applying an aggregate function. Subsequently, instead of sending all the messages towards the base station, the aggregator transmits only one aggregated result towards the base station [4].

However, while having advantage in a manner described above, the data aggregation operation certainly gives rise to other consequences. Data aggregator nodes usually collect data from the sensor nodes, apply aggregation operations on it and eventually communicate the processed data to the base station. Obviously, in the presence of the malicious nodes, neither the aggregated data nor the aggregation operation remains trustworthy. Thus, it is essential to ensure that the data aggregation operational paradigm is inherently secure.Thus, protocols for WSNs should be designed to prevent malicious inside nodes from damaging the whole network's functionality or at least constrain their impacts to a reasonable level.

Amongst various security attributes in WSNs, we focus in this discussion on the most important attributes privacy, confidentiality and data integrity. For privacy and confidentiality we investigated various algorithms proposed in [5] for WSNs. We select algorithm EC-OU based on the analysis given in [5]. The approaches to provide data integrity can be either cryptography-based or non-cryptography-based. Our focus here is only on cryptographic approaches. As per our literature survey, we categorize the techniques for supporting data integrity in Secure Data Aggregation into three classes viz. Signature based, Hash function based and Message Authentication Code (MAC) based [6]. The digital signature based approach yields non-repudiation property, however, entails higher overhead as compared to other approaches. To counter the overhead due to the digital signature, our focus here is on the message authentication code based integrity support for secure data aggregation in WSNs.

2 Secure Data Aggregation

Data aggregation functions are improving the bandwidth and energy utilization, it also affect negatively to other various performance metrics like delay, accuracy, fault-tolerance and security. As almost all the applications of WSNs demand certain level of security, it is not possible to sacrifice security of WSNs for data aggregation [7]. There is a strong conflict between security protocols and data aggregation protocols. Security protocols require encryption and authentication of data by each sensor nodes to provide security while data aggregation protocols prefer plain data to perform aggregation functions on it, to save energy and bandwidth of WSNs. Moreover, data aggregation also modifies data by applying various aggregation functions i.e. sum, average, min, max on it. Due to these, it is a challenging task to provide source and data authenticity for data aggregation. Therefore, data aggregation and security protocols must be designed together so that data aggregation can be performed without sacrificing security.

2.1 Privacy Homomorphism

The fundamental basis for data aggregations are cryptographic methods that provide privacy homomorphism property. A privacy homomorphism (PH) is an encryption transformation that allows direct computation on encrypted data [8]. With the help of privacy homormophism, data do not need to be decrypted at the intermediate nodes. Hence, they are transmitted to the base station in such a way that information contained in the nodes is not visible or accessible to the intermediate nodes. Hence, privacy of the data is preserved while being transmitted towards the base station. A cryptosystem that supports single operation addition or multiplication is known as partial homomorphic encryption (PHE).

Let Q and R denote two rings, and + and $\oplus$ denote addition operations on the rings. Let k denotes the key space. We denote an encryption transformation $E: K\ X\ Q \rightarrow R$ and the corresponding decryption transformation $D : K\ \ X\ \ R \rightarrow Q$. Given $a, b \in Q$ and $k, k_1, k_2 \in K$,

$a + b = Dk\ (Ek(a) \oplus Ek(b))$

additively homomorphic with a single secret key and

$a + b = Dk\ (k_1,k_2)\ (Ek_1(a)\ \oplus Ek_2(b))$

additively homomorphic with multiple secret keys. We denote an asymmetric additively homomorphic encryption transformation as

$a + b = Dp\ (Ep(a) \oplus Eq(b))$

with (p,q) being a private, public key pair.

With the help of privacy homomorphism secure data aggregation can be achieved by either using symmetric key cryptography or public key cryptography. Symmetric homomorphic encryption requires use of identical key by encryption and decryption while asymmetric homomorphic encryption requires use of public key-private key for encryption and decryption.

2.2 Related Work

In [9], the authors propose secure and efficient scheme for data aggregation in WSNs. The proposed approach uses Castelluccia-Mykletun-Tsudik [10] scheme for encryption. Castelluccia-Mykletun-Tsudik is a symmetric key cryptography based scheme and suffers from the key management issues. The attempt in [11] also follows similar approach of symmetric key cryptography. In [12], the authors propose a scheme for SDA using elliptic curve cryptography that achieves only confidentiality without integrity. None of the approaches discussed so far, support confidentiality as well as message integrity that the authors in [13], attempt to do. In [14], authors propose ECC based SDA scheme that provides confidentiality and integrity, but use EC-EG for encryption. However, as per empirical evaluation in [5], energy consumption as well as decryption time is more for this algorithm. In [15], authors propose scheme for secure data aggregation that focuses only on confidentiality and privacy not on integrity. In [16], authors propose scheme of secure data aggregation but that do not preserve privacy of the sensed data. In [17], authors propose secure data aggregation based on data slicing technique but without integrity support.

As mentioned before, we do not observe any research attempt in using such a benchmarked privacy homormophic encryption algorithm[5] EC-OU in a tree topology, to showcase *truly securedata aggregation*. Thus, we believe ours is the first and unique attempt in integrating EC-OU in the tree topology for providing privacy and confidentiality with MAC based integrity support for secure data aggregation in WSNs.

3 Proposed Approach of Privacy Preserving Secure Data Aggregation

3.1 Algorithms

In our scheme, we propose two algorithms, first one for the sensor node and the other one for the base station. We discuss the algorithms with the help of their pseudo code further. Each sensor node will execute SensorNodeAlgo( ) that takes plain text message and private key as input to the algorithm. Next, each sensor node will compute its public key by multiplying its private key to the base point of elliptic curve E. Subsequently, each sensor node comprise two large prime number p & q and calculate n as public key and calculate H=nG with point multiplication. Next, sensor node will compute the cipher text and transmit it to its parent node.
ciphertext C = mG + rH

The parent node receives cipher text from the child nodes, performs the summation of all the received cipher texts and finally transmits the aggregated cipher text to the base station.The pseudo code for the proposed algorithm is shown below:

```
Algorithm 1:SensorNodeAlgo()
// Maps its reading $m_i$ on the elliptic curve E
// Elliptic Curve Parameters E = (q, a, b, G, p, h)
// Each sensor node will computes following
1. Public Key:  $n = p^2q$, G, H, Q=pG
2. Encryption:  plaintext $m < 2^{k-1}$, $r \bullet_R 2^{2k}$,
Ciphertext C = mG + rH
3. if sensor is a parent (Aggregator Node)
c = $\bullet\, c_i$ // combines all cipher texts into one cipher text
end if
```

The base station algorithm takes cipher text as input to the algorithm. Base station then decrypts the incoming packets and calculates plain text that is the summation of all the messages from all the sensor nodes.

```
Algorithm 2: BaseStationAlgo( )
// Maps its reading $c_i$ from the elliptic curve E
// Elliptic Curve Parameters E = (q, a, b, G, p, h)
// Base station will computes following
1. Public Key:$n = p^2q$, G, H, Q=pG
2. Private Key:    p
3. Compute$m = \frac{\bullet p}{\bullet}$ (mod p)where, $\bullet p(x, y) = -\frac{x}{y}$(mod $p^2$)
4. Base station can get $m = \bullet m_i$
```

3.2 Example

Example of secure data aggregation of our propose approach is there in figure 1. In this, nodes 3 and 4, first encrypt messages 3 and 5, respectively using EC-OU privacy homormophic encryption function and send resultant cipher texts to node 1. Node 1 receives cipher text from nodes 3 and 4, aggregates them using the addition operation (SUM), and forwards it to node 0 that acts as a base station.

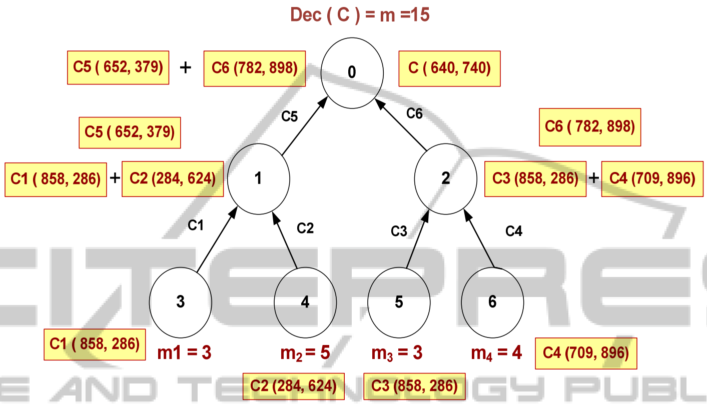

Fig. 1. Example of Privacy Preserving Secure Data Aggregation

In the same manner, nodes 5 and 6 encrypt messages 3 and 4, respectively using EC-OU privacy homomrophic encryption function and forward them to node 2. Node 2 receives both the encrypted data, perform aggregation function, and send the resultant aggregated result to node 0. After receiving both the aggregated result from node 1 and 2, node 0 performs the aggregation operation. After aggregation, node 0 applies EC-OU decryption function to decrypt the aggregated result. As shown in fig. 1, decrypted output is message 15 that is the summation of messages 3,5,3,4 of nodes 3, 4, 5, and 6 respectively. Hence, data are transmitted to the base station in such a way that their privacy is preserved.

4 Proposed Approach of Integrity Preserving Secure Data Aggregation

In this section, we discuss MAC based hop-by-hop integrity support for secure data aggregation in WSNs. In hop-by-hop integrity, if malicious adversary insert any false data in the networks that will be detected immediately at next hop. Therefore, that malicious data do not travel towards the base station. In contrast to that, in end-to-end integrity, false data are detected only at the end i.e. at the base station, wasting energy and lifetime of sensor networks. Thus, our proposed approach of hop-by-hop integrity in secure data aggregation save bandwidth of the sensor networks and increase the lifetime as well as security of the sensor nodes.

4.1 Algorithms

In our proposed approach, algorithm 1 is to be implemented on the leaf node. Each sensor node computes M_i on the outgoing message m_i.

M_i= MAC(m_i)

Parent of sensor node receives M_i and m_i and computes M_i on received m_i. If received M_i is the same as computed M_i, parents accept the message and apply aggregation on the message. After aggregation, parents compute MAC on aggregated message and send it further.

```
Algorithm 1:Leaf Node()
// Each leaf node will computes following
MAC Generation: Each sensor computes
$M_i$= MAC($m_i$)
Append this $M_i$to message and
send it to parent node
```

Algorithm 2 is to be implemented on the aggregator node and the base station. Aggregator node or base station will receive aggregated message and MAC of sensor nodes. Aggregator node or base station again compute MAC on the received message and verify that weather received MAC is same as computed MAC or not. If it is not same, aggregator node simply discard the message. Thus, our approach ensures hop-by-hop integrity through MAC.

```
Algorithm 2: Aggregator Node and Base Station()
// Aggregator Node and Base station will computes following
Verification of MAC:
Aggregator or Base station receive $M_i$ and $m_i$
Compute $M_i$= MAC($m_i$)
Verify Computed $M_i$ = received $M_i$
Base station can get m = • $m_i$
```

4.2 An Illustration

In our approach of hop-by-hop MAC based integrity, each leaf node shares secret key with parent and each leaf node generates MAC using SHA-1on the outgoing message with the help of the shared key. Similarly, after receiving the message and computed MAC from child, aggregator node or Base station will again compute MAC with the key shared with child and verify the integrity of the message. For example, in our approach, node 3 generates MAC on data of node 3 so M3 (MAC 3) is generated that is received by node 1. Similarly node 4 generate MAC on data of node 4 so M4 (MAC 4) is generated that is also received by node 1. Now node 1 have M3 (MAC 3) and M4 (MAC 4), so node 1 verify M3 and M4. If it is verified, then only node 1 will accept message from node 3 and 4 and apply aggregation function on it and generate m1. Otherwise, node 1 will not accept the messages from node 3 and node 4 and simply discard the messages. If node 1 has accepted messages from node 3 and node 4 and generated aggregated message, node 1 apply MAC on aggregated message and generate M1 (MAC 1). Same way node 0 (Base station) accept messages from node 1 and node 2 only if M1 (MAC 1) and M2 (MAC 2) is verified. In fig. 2, this scenario is shown.

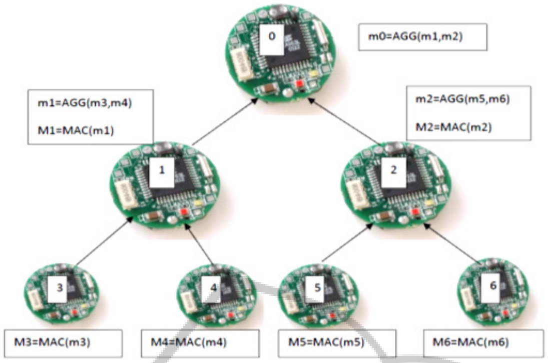

Fig. 2. Example of Integrity Preserving Secure Data Aggregation.

5 Combined Approach of Privacy and Integrity Preserving Secure Data Aggregation

In this section, we discuss integration of MAC based Hop-by-Hop integrity with privacy and confidentiality.

5.1 Pseudo Code

In our proposed approach of secure data aggregation algorithm 1 is to be implemented on the sensor node. Each sensor node generates cipher text c_i of the message.

C_i=Enc(m_i)

In this, each sensor node computes M_i on the outgoing message C_i.

M_i= MAC(C_i)

Parent of sensor node receives M_i and C_i and parent computes M_i on received C_i. If received C_i is same as computed C_i, parent accepts the message and applies aggregation on message. After aggregation, parent compute MAC on aggregated message and send it further.

```
Algorithm 1: SensorNodeAlgo()
// Maps its reading $m_i$ on the elliptic curve E
// Elliptic Curve Parameters E = (q, a, b, G, p, h)
// Each sensor node will computes following
Public Key:  $n = p^2q$, G, H, Q=pG
Private Key:  p
Encryption: plaintext $m_i < 2^{k-1}$, $r \bullet_R 2^{2k}$,
            Cipher text $C_i = m_iG + rH$
MAC Generation: Each sensor computes $M_i$= MAC($C_i$)
if sensor is a parent (Aggregator Node)
Compute $M_i$= MAC($C_i$)
Verify Computed $M_i$ = received $M_i$
if verified go to next step
C = $\bullet$ $c_i$ //combines all cipher texts into one cipher text and
send it to parent node
End if
```

Algorithm 2 is to be implemented on the base station. Base station will receive aggregated cipher text and MACof the sensor nodes. Base station again compute MAC on the received message and verify that weather received MAC is same as computed MAC or not. If it is same then apply decryption function on aggregated cipher text and get the original message that is summation of all the plain text messages. Thus, our approach ensures confidentiality and privacy through privacy homomorphic encryption and in addition to that, it ensures hop-by-hop integrity through MAC.

```
Algorithm 2: BaseStationAlgo()
// Maps its reading $c_i$ from the elliptic curve E
// Elliptic Curve Parameters E = (q, a, b, G, p, h)
// Base station will computes following
1. Public Key:n = $p^2$q, G, H, Q=pG
2. Private Key:p
3. Verification of MAC: Compute $M_i$= MAC($C_i$)
   Verify Computed $M_i$ = received $M_i$
   if verified go to next step
4. Compute m = $\frac{\Psi p((p+2)C)}{\Psi p((p+2)G)}$ (mod p)
•p(x, y) = - $\frac{x}{y}$(mod $p^2$)
5. Base station can get m = • $m_i$
```

5.2 An Illustration

Example of our framework of secure data aggregation is there in figure 3. The notations used are as follows:

Msg = Message of sensor node
MAC_i = MAC computed using SHA-1by i^{th} node
C_i = Cipher text generated by i^{th} node

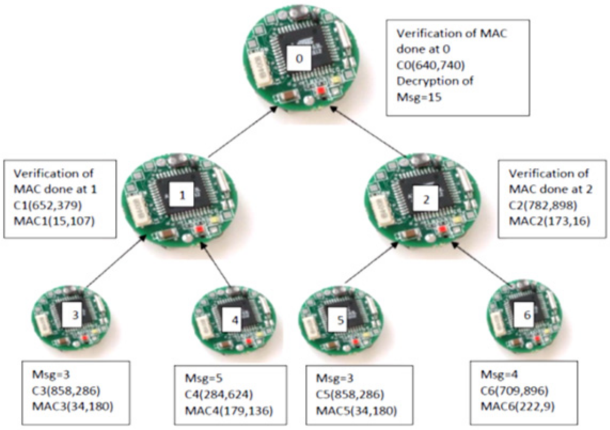

Fig. 3. Example of Privacy and Integrity Preserving Secure Data Aggregation.

6 Implementation Methodology and Results

6.1 Platform and Tools Used

We devise an application testECOUM for ECC based privacy homomorphic encryption algorithm in the TinyOS 1.x operating environment [18] using TinyECC [19] library, with nesC [20] as the language for implementation. Our evaluation is based on a two-step approach: a) we use TOSSIM [21] as the WSN simulator; using that we also obtain the estimates of the storage requirements of the respective implementation, b) energy consumption in Joules using the Avrora emulator [22].

6.2 Results

Table 1. % Increase Compared to without SDA.

Framework	ROM in bytes	RAM in bytes	Energy Consumption in μJoules
Confidentiality and privacy preserving SDA	4.45%	7.59%	0.002%
Confidentiality, Privacy and Integrity preserving SDA	20.06%	36.45%	0.008%

Table 1 shows the summary of our results. It shows percentage increase in results compared to that of without secure data aggregation. Confidentiality and privacy preserving approach of SDA requires 4.45 % more ROM and 7.59 % more RAM compared to that of without SDA. Our MAC based hop-by-hop approach requires 20.06 % more ROM and 36.45 % more RAM.

Figure 4 shows ROM requirements of data aggregation with various options. (A) without SDA (only data aggregation) (B) data aggregation that supports privacy and confidentiality (C) data aggregation that supports MAC based hop-by-hop integrity in addition of privacy and confidentiality.

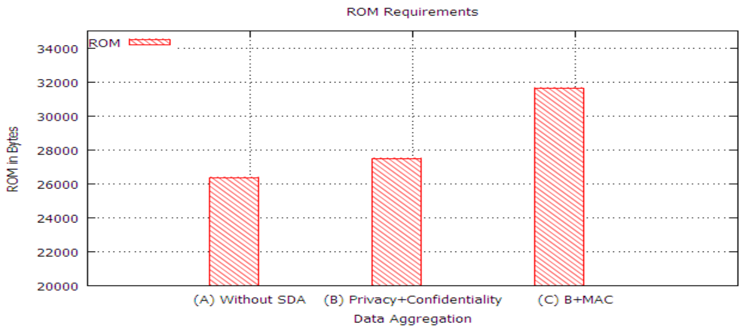

Fig. 4. ROM Requirements.

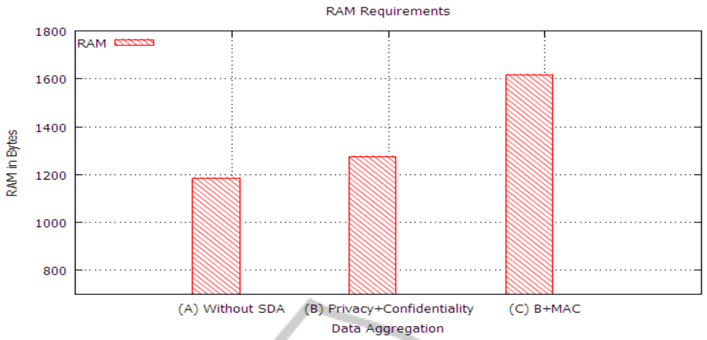

Fig. 5. RAM Requirements.

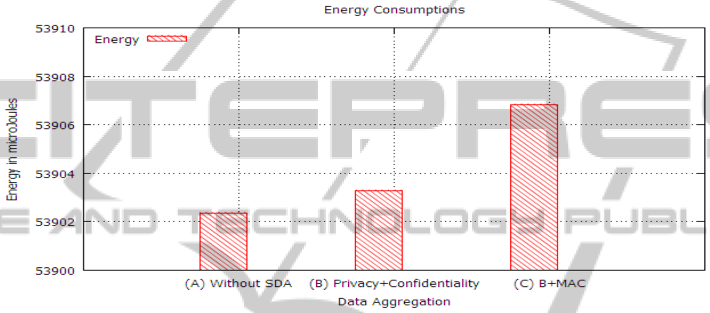

Fig. 6. Energy Consumptions.

Figure 5 shows RAM requirements and figure 6 shows energy consumptions. Figures clearly shows that if we add integrity in addition of privacy and confidentiality, then resource consumption and energy consumption increases. This is because the computation involves in computing MAC on each message but that is at the advantage of providing support for all the required security attributes.

7 Conclusion

In this research exercise we propose privacy, confidentiality and integrity preserving secure data aggregation in Wireless Sensor Networks. From our experimentations, it is observed that all the necessary security attributes for secure data aggregation are preserved within life time of the resource constrained environment of WSNs. Our approach requires less than 21% ROM, less than 37% RAM and less than 1% increase in energy consumption. Our results show that our approach of integrity, privacy and confidentiality preserving secure data aggregation is feasible in WSNs.

References

1. E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi. In-network aggregation techniques for wire

less sensor networks: a survey. Wireless Communications, IEEE, Vol.14, No.2, pp.70,87, April 2007.

2. Chris Karlof, Naveen Sastry, David Wagner. TinySec: A link layer security architecture for wireless sensor networks. InProceedings of the 2nd international conference on Embedded Networked Sensor Systems SenSys '04, ACM, New York, NY, USA, pp. 162-175, 2004.
3. Tieyan Li, Hongjun Wu, Xinkai Wang, Feng Bao. SenSec: Sensor Security Framework for TinyOS. In Proceedings of the Second International Workshop on Networked Sensing Systems, San Diego, California, USA, pp.145-150, 2005.
4. Ramesh Rajagopalan and Pramod K. Varshney. Data aggregation techniques in sensor networks: A survey. Comm. Surveys Tutorials, IEEE, Vol. 8, pp. 48–63, 2006.
5. VivakshaJariwala, Asha Munjpara, DeveshJinwala, Dhiren Patel. Comparative Evaluation of ECC Based Homomorphic Encryption Algorithms in TOSSIM for Wireless Sensor Networks. In Proceedings of the National Workshop on Cryptology, Cryptology Research Society of India and VIT, Vellore, TN, pp. 1-14, 2012.
6. Don Johnson, Alfred Menezes and Scott Vanstone.The Elliptic Curve Digital Signature Algorithm (ECDSA).International Journal of Information Security, Vol. 1, No. 1, Springer-Verlag, pp. 36-63, 2001.
7. Alzaid Hani, Foo Ernest, and Nieto Juan Gonzalez. Secure data aggregation in wireless sensor network: a survey. In Proceedings of the sixth Australasian conference on Information security (AISC '08), Australian Computer Society, pp. 93–105, 2008.
8. Josep Ferrer and Domingo.A new privacy homomorphism and applications.Inf. Process. Lett., Vol. 60, No. 5, pp. 277–282, 1996.
9. Xiaoyan Wang; Jie Li; Xiaoning Peng; BeijiZou.Secure and Efficient Data Aggregation for Wireless Sensor Networks. In Proceeding of the Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd , pp.1-5, 6-9 Sept. 2010.
10.ClaudeCastelluccia, EinarMykletun, and Gene Tsudik. Effcient aggregation of encrypted data in wireless sensor networks.In Proceeding of the MobiQuitous, IEEE Computer Society, pp. 109–117, 2005.
11.Poornima, A.S.; Amberker, B.B. SEEDA: Secure end-to-end data aggregation in Wireless Sensor Networks. In Proceeding of the 7th International Conference of Wireless And Optical Communications Networks (WOCN), pp.1-5, 6-8 Sept. 2010.
12. Jacques M. Bahi, Christophe Guyeux, and AbdallahMakhoul. Efficient and Robust Secure Aggregation of Encrypted Data in Sensor Networks. In Proceedings of the 2010 Fourth International Conference on Sensor Technologies and Applications (SENSORCOMM '10), IEEE Computer Society, Washington, DC, USA, pp. 472-477, 2010.
13.SuatOzdemir and Yang Xiao. Integrity protecting hierarchical concealed data aggregation for wireless sensor networks.Comput.Netw., Vol. 55, No. 8, pp. 1735-1746, 2011.
14.SuatOzdemir and Yang Xiao. Hierarchical Concealed Data Aggregation for Wireless Sensor Networks, 2009.
15.AjayJangra, Priyanks and Richa. Cb-SDA: Cluster-based Secure Data Aggregation for Private Data in WSN. Wireless and Mobile Technologies, Vol. 1, No. 1,Science and Education Publishing pp. 37-41, 2013.
16 .Shaik Mohammad Rafi and K. Subbarao. Secure Data Aggregation In Wireless Networks. International Journal of Research in Computer and Communication Technology, Vol 3, Issue 1, pp. 87-93, January- 2014
17.TaochunWang,Xiaolin Qinand Liang Liu. An Energy-Efficient and Scalable Secure Data Aggregation for Wireless Sensor Networks.International Journal of Distributed Sensor Networks.Volume 2013.
18.J. Hill, et al. System Architecture Directions for Networked Sensors. In Proceedings of 9th Intl. Conf. on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2000), ACM Press, pp. 93-104, 2000.
19.A. Liu, P. Kampanakis, P. Ning. TinyECC: Elliptic curve cryptography for sensor networks. In Proceedings of the International Conference on Information Processing in Sensor Net

works, pp. 245-256, 2008.

20. David Gay, Phil Levis, Rob von Behren, Matt Welsh, Eric Brewer, and David Culler. The nesC language: A holistic approach to network embedded systems. In Programming Language Design and Implementation (PLDI), June 2003.

21. Philip Levis and Nelson Lee. TOSSIM: A Simulator for TinyOS Networks Version 1.0 June 26, 2003.

22. Ben L. Titzer, Daniel Lee, and Jens Palsberg. Avrora: Scalable sensor network simulation with precise timing. In Proceedings of the 4th Intl. Conf. on Information Processing in Sensor Networks (IPSN), Los Angeles, CA, April 2005.