An Improved Approach for Secure Data Aggregation in Wireless Sensor Networks

Vivaksha J. Jariwala
Assistant Professor,
CKPCET, Surat

Sankita J. Patel
Assistant Professor,
SVNIT, Surat

ABSTRACT
The Wireless Sensor Networks (WSNs) are composed of sensor nodes that are deployed in remote and hostile environments to sense, process and communicate vital information to the base station. Due to the stringent constraints on the resources in the sensor nodes, it is essential to optimally devise the WSN operational paradigms, to minimize the resource overhead. Since, communications costs always are significantly higher than that for processing the sensed data the WSNs typically, employ in-network processing, so as to minimize effectively, the total number of packets eventually transmitted to the base station. Such in-network processing is largely based on data aggregation operations that aggregate the data into a compact representation viz. a data aggregate for further transmission. However, due to the ubiquitous and pervasive deployment of the sensor nodes, the security concerns in WSNs are anyway critical. Therefore, it is necessary to ensure the security of the data aggregator nodes that depend on various other nodes for the eventual output using carefully designed approaches. In this paper, we investigate various approaches for data aggregation with a view to critically analyze the same and propose a new approach for secure data aggregation.

General Terms
Security, Wireless Sensor Networks.

Keywords
Confidentiality, Integrity, Privacy Homomorphism, Data Aggregation, Secure Data Aggregation.

1. INTRODUCTION
Wireless Sensor Networks (WSNs) consist of hundreds or thousands of tiny sensing devices with restricted memory, computational and communication resources [1]. These devices are severely resource constrained with a typical sensor mote consisting of only 8-bit, 4 MHz processor, 128 kb program flash memory, 512 kb of EEPROM and two AA batteries [2].

The potential applications of the WSNs typically range from those in *defense, military, environmental monitoring, health monitoring, home appliances, civilian societal surveillance* applications etc. [1]. Most of these applications being ubiquitous in nature necessitate that the appropriate mechanisms to ensure the data security and privacy are in place in the protocols in WSNs. In addition, the exigencies on the available resources make it difficult to devise the security protocols for the WSNs that operate with optimum overhead. Hence, novel approaches are used in operational paradigms followed in WSNs – one of them is using data aggregation of the incoming packets to a sensor node, that aims at reducing the communications cost in the WSNs [3]. However, data aggregation in turn necessitates the aggregator node to ensure proper security checks on the data aggregated so that the overall sanctity of the data and the associated operations is maintained.

There are several attempts proposed in the literature to devise Secure Data Aggregation protocols. We aim in this paper to propose a newer approach with the focus on obtaining maximum advantage in terms of the security properties derived, from the underlying algorithm.

The rest of the paper is organized thus: in section 2, we have described data aggregation with its related issues. Various approaches for secure data aggregation in wireless sensor networks are explained in section 3, our proposed approach for secure data aggregation is described in section 4.

1.1 Related Work
Data aggregation algorithms are used to gather and aggregate data in an energy efficient manner to increase the life time of data. A detailed survey of data aggregation algorithm is presented in [3]. However, an important associated concern is the security of the aggregated data as well as that of the data aggregation operation, too. Obviously, the data aggregation techniques *alter* the data during the aggregation operation itself. Hence, any compromise on the security of the input data to the operation or that of the aggregated result or even that of the operation itself, can jeopardize the semantics and utility of the operation itself. Hence, it is necessary to investigate and devise suitable approaches for *secure* data aggregation.

Secure data aggregation is classified generally in two ways viz. a) aggregation on plain sensor data and b) aggregation on encrypted sensor data. In [4], a comprehensive review of the approaches for aggregation on plain sensor data i.e. aggregation employing *in-network* processing is provided; distinguishing approaches based on different layers of the protocol stack as well as that based on data representations used for secure data aggregation.

On the other hand, the approaches ensuring secure data aggregation using encrypted sensor data are based on the property of privacy homomorphism or homomorphic encryption algorithms. The algorithms that use homomorphic encryption can be as usually classified in to those based on either symmetric key based algorithms (as in [1], [5]) or asymmetric key based algorithms (as in [6]). In [7], the authors present a survey of all the technique of secure data aggregation of encrypted sensor data without focusing on those for plain sensor data. A survey on privacy-preserving data aggregation without secure channel is presented in [8].

From our survey of existing Secure Data Aggregation Techniques we recognize that the existing approaches proposed for secure data aggregation do not offer the required security features viz. confidentiality, integrity and

authentication together. Therefore, with the motivation to improve upon the same, we propose a new approach that offers *ALL* of these attributes. Our approach is based on public key cryptography, using homomorphic encryption and additive digital signatures to achieve *confidentiality, message integrity* and *authentication* for data aggregation in wireless sensor networks.

2. SECURE DATA AGGREGATION

One of the issues that is encountered when devising the security protocols for the resource constrained environments is the associated overhead. In order to perform data aggregation hop-by-hop, intermediate nodes have to decrypt each received message, then aggregate the messages according to the corresponding aggregation function, and finally encrypt the aggregation result before forwarding it. Clearly, this is not an energy efficient way of performing secure data aggregation and it may result in considerable delay. In addition, this process requires neighboring data aggregators to share secret keys for decryption and encryption. So in order to achieve end-to-end data confidentiality and data aggregation together without requiring secret key sharing among data aggregators privacy homomorphism has been used in the literature [9, 10].

Naturally, when using secure data aggregation on the plain sensor data, the number of rounds of encryption/decryption operations (and therefore the associated overhead) drastically increases. In addition, such multiple rounds of decryption-encryption operations at each intermediate node also make the entire protocol more vulnerable to security threats. Hence, it is necessary to explore the probable approaches for devising secure data aggregation based on encrypted sensor data [4][5]. Obviously, the encryption/decryption operations are done at the end-to-end level in this approach and hence reduces the associated overhead.

In this section we discuss the various Secure Data Aggregation schemes in general.

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 [10].

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 $E : K \times Q \rightarrow R$ and the corresponding decryption transformation $D : K \times R \rightarrow Q$. Given $a, b \in Q$ and $k, k1, k2 \in K$, we term $a + b = Dk (Ek(a) \oplus Ek(b))$ additively homomorphic with a single secret key and $a + b = Dk (k1,k2) (Ek1(a) \oplus Ek2(b))$ 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.

2.2 Encryption Transformations

Encryption transformations are based on symmetric homomorphic encryption transformations and asymmetric homomorphic encryption transformations.

Symmetric Homomorphic Encryption Transformations

Symmetric homomorphic encryption requires identical keys to be used for encryption and decryption. All the symmetric homomorphic encryption transformations are summarized in Table 1.

Table 1. Summary of SDA using Encrypted (Symmetric Key) Sensor Data

No.	Approaches	Citation	Remark
1	Domingo-Ferrer	[11]	First approach for symmetric homomorphism of encrypted data
2	Energy efficient and high accuracy approach	[12]	Secure data aggregation without releasing private sensor readings and without imposing overhead on sensor nodes.
3	PH Probabilistic	[13]	Probabilistic means encryption transformation involves some randomness that chooses the ciphertext corresponding to given plaintext.
4	Castelluccia, Mykletum, Tusdik (CaMyTs)	[14]	It uses only modular additions; it is very well suited for CPU constrained devices. Dis : Sensor node identity also sent to base station along with aggregate data. This involves lot of overhead and scalability problem.
5	Pair - Wise key approach	[15]	Sensor sends data to its parent, it encrypts it n time using additively homomorphic schemes as [13]
6	Dynamic Cluster	[16]	Security equal to one time pad is achieved and also combined with dynamic cluster head recycling. Balance energy

			consumption of the whole network and prolong its life.
7	Combined Cryptosystem	[17]	Combination of [10] and [13]. Combine the advantages of both crypto schemes.

Asymmetric Homomorphic Encryption Transformations

All the asymmetric homomorphic encryption transformations are summarized in Table 2.

Table 2. Summary of SDA using Encrypted (Asymmetric Key) Sensor Data

No.	Approaches	Citation	Remark
1	Discrete Logarithms	[18]	OU- As secure as factoring and based on the ability of computing discrete logarithms in particular subgroup.
2	Probabilistic Cryptosystem	[19]	Benaloh - Encryption cost is dependent on the size of the plaintext. Additive homomorphic property is achieved through the multiplication of cipher texts.
3	Elliptic Curve Cryptography	[20]	EC-OU - Discrete logarithms are easy to compute in curves Ep(ap,bp) over Fp. EC-NS - factoring based algorithms are exported to particular families of EC. EC-P - is not as efficient and requires too much computation. EC-EG - additively homomorphic, and ciphertexts are combined through addition.
4	Group Based	[21]	HCDA - sensor nodes in a group use the same public key. It is based on Elliptic curve cryptography so it is not affected by node compromise attacks whereas symmetric key based protocols are significantly affected from these attacks.
5	Hierarchical	[22]	Confidentiality and integrity

3. PROPOSED APPROACH FOR SECURE DATA AGGREGATION

In this research exercise, we attempt to explore further the feasibility of the PKC schemes using privacy homomorphism in WSNs. As we explain further, in this section, we employ digital signatures to achieve secure data aggregation, providing confidentiality and integrity of the data. Our proposed approach combines the idea from the PKC based Discrete Logarithms approach OU proposed in [18] and ECDSA proposed in [23] to achieve confidentiality and integrity of data to be aggregated and passed on to the base station.

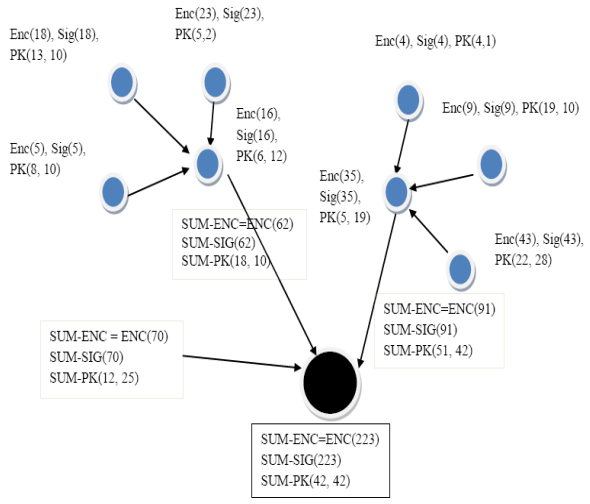

Fig 1: An Example

In our scheme, we have proposed two algorithms, first to be implemented on the sensor node and the other one to be implemented on the base station.

Our proposed algorithm for sensor node generates unique signature for each outgoing message say Sig(x). The data from each sensor node is encrypted using the base station's public key viz. Enc(x). Thus, every node communicates the encrypted message, its signature and its public key to its parent. After receiving the message from every child, the parent combines signature, public key and encrypted message and sums it. The message so received can be decrypted only

by the base station that has the corresponding private key. Thus, the base station can now verify the sum of the signatures given the sum of the public keys.

Our proposed approach is based on Elliptic Curve Digital Signature Algorithm [23] and Elliptic Curve Okamoto Uchiyama algorithm [18]. Both approaches are based on elliptic curve cryptosystem. Elliptic curve cryptosystem is also having its own advantages like (a) their use of small keys which lead to short cipher texts, (b) the smaller real-estate required for hardware implementations (number of gates) and (c) a better security-per-bit ratio. By considering advantage and efficiency of OU and Elliptic curve cryptosystem, we have used EC-OU [18] for asymmetric homomorphic encryption transformation that gives us the confidentiality of data. EC-OU is provably secure and has many properties like: (a) Its trapdoor technique is essentially different from any other asymmetric schemes. (b) It is a probabilistic encryption scheme. (c) It can be proven to be as secure as the intractability of factoring $n = p^2q$ (d) It is semantically secure under the p-subgroup assumption, which is comparable to the quadratic residue and higher degree residue assumptions. (e) It has homomorphic property [18]. We have used ECDSA to provide integrity of data for data aggregation. ECDSA is assumed to be secure under the assumption that the underlying group is generic and that a collision resistant hash function has been used.

4. CONCLUSION AND FUTURE WORK

We propose a novel approach for secure data aggregation in WSNs. The proposed approach uses homomorphic encryption EC-OU algorithm to achieve data confidentiality while allowing in-network aggregation. We have used an additively digital signature algorithm based on Elliptic Curve Digital Signature Algorithm (ECDSA) to achieve integrity of the aggregate. To the best of our knowledge, combining preeminent feature of EC-OU and ECDSA to attain confidentiality and integrity and digging up benefit of both is new and hence it would be worth exploring. Our future work will include analysis of our novel approach and implementing the same in TinyOS/TOSSIM[24] .

5. REFERENCES

[1] F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci. 2002. Wireless sensor networks: a survey. *Computer Networks*, 38(4):393 – 422, 2002. ISSN 1389-1286. doi: DOI:10.1016/S1389-1286(01)00302-4.

[2] Crossbow, mica, wireless measurement system , datasheet. http://www.xbow.com/products/product_pdf_files/wireless_pdf/mica2_datasheet.pdf

[3] Ramesh Rajagopalan and Pramod K. Varshney. 2006. Data aggregation techniques in sensor networks: A survey. *Comm. Surveys Tutorials*, IEEE, 8:48–63.

[4] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi. 2007. In-network aggregation techniques for wireless sensor networks: a survey. *IEEE Wireless Communication*, 14(2):70–87.

[5] Alzaid Hani, Foo Ernest and Nieto Juan Gonzalez. 2008. Secure data aggregation in wireless sensor network: a survey. In: *Proceedings of the sixth Australasian conference on Information security (AISC '08)*, pages 93–105. Australian Computer Society, Inc., ISBN 978-1-920682-62-0.

[6] Einar Mykletun, Joao Girao, and Dirk Westhoff. 2006. Public key based crypto schemes for data concealment in wireless sensor networks. In *IEEE International Conference on Communications.*

[7] S. Peter, D. Westhoff, and C. Castelluccia. 2010. A survey on the encryption of convergecast traffic with in-network processing. *IEEE Transactions on Dependable and Secure Computing*, 7(1).

[8] Kshitija Nandgaonkar, Swarupa Kamble. 2016. A survey on privacy-preserving data aggregation without secure channel. *International Research Journal of Engineering and Technology (IRJET)*, 3(1).

[9] D. Westhoff, J. Girao, and M. Acharya. 2006. Concealed data aggregation for reverse multicast traffic in sensor networks: encryption key distribution and routing adaptation. *IEEE Trans. Mobile Comput.*, 5(10):1417–1431.

[10] Josep Ferrer and Domingo. 1996. A new privacy homomorphism and applications. *Inf. Process. Lett.*, 60(5):277–282, ISSN 0020-0190.

[11] Domingo-Ferrer and Joseph. 2002. A provably secure additive and multiplicative privacy homomorphism. In *ISC '02: Proceedings of the 5th International Conference on Information Security*, pages 471–483, London, UK, Springer-Verlag. ISBN 3-540-44270-7.

[12] Hongjuan Li, Kai Lin, and Keqiu Li. 2010. Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks. *Computer Communications*, In Press, Corrected Proof, 2010. ISSN 0140-3664.

[13] J. Girao, D. Westhoff, and M. Schneider. 2005. Cda: Concealed data aggregation for reverse multicast traffic in wireless sensornetworks. In *IEEE Intl Conf. Comm.*

[14] Claude Castelluccia. 2005. Efficient aggregation of encrypted data in wireless sensor networks. In *In MobiQuitous*, pages 109–117. IEEE Computer Society.

[15] M. Oenen and R. Molva. 2007. Secure data aggregation with multiple encryption. In *Proceedings of Fourth European Conference on Wireless Sensor Networks*.

[16] Xinyang Huang, Ming Yang, and Yong Tong. 2007. An efficient and secure aggregation of encrypted data for wireless sensor network based on dynamic cluster. In *SpringSim '07: Proceedings of the 2007 spring simulaiton multiconference*, pages 51–57, SanDiego, CA, USA, 2007. Society for Computer Simulation International. ISBN 1-56555-312-8.

[17] S. Peter, P. Langendo, and K. Piotrowski. 2007. On concealed data aggregation for wireless sensor networks,. In *Fourth IEEE Consumer Comm. and Networking Conf. (CCNC).*

[18] Tatsuaki Okamoto and Shigenori Uchiyama. 1998. A new public-key cryptosystem as secure as factoring. In *Eurocrypt '98, LNCS 1403*, pages 308–318. Springer-Verlag.

[19] Josh Benaloh Clarkson. 1994. Dense probabilistic encryption. In *Proceedings of the Workshop on Selected Areas of Cryptography*, 120–128.

[20] David Naccache and Jacques Stern. 1998. A new public key cryptosystem based on higher residues. In *CCS '98: Proceedings of the 5th ACM conference on Computer*

and communications security, pages 59–66, New York, NY, USA. ACM. ISBN 1-58113-007-4. doi: http://doi.acm.org/10.1145/288090.288106.

[21] Suat Ozdemir and Yang Xiao. 2009. Hierarchical concealed data aggregation for wireless sensor networks. In *Proc. of Embedded Systems and Communications Security Workshop in conjunction with IEEE SRDS 2009*.

[22] Julia Albath and Sanjay Madria. 2009. Secure hierarchical data aggregation in wireless sensor networks. In *Proceedings of the 2009 IEEE conference on Wireless Communications and Networking*.

[23] Don Johnson, Alfred Menezes and Scott Vanstone. The elliptic curve digital signature algorithm (ecdsa).

[24] P. Levis, N. Lee, M.Welsh, and D. Culler. 2003. Tossim: accurate and scalable simulation of entire tinyos applications. In *SenSys '03: Proceedings of the 1st international conference on Embedded networked sensor systems*. New York, NY, USA: ACM Press, 126–137.