RESEARCH ARTICLE OPEN ACCESS

Secure Data Aggregation in Wireless Sensor Networks

Dr. Debmalya Bhattacharya
Associate Professor, Department of Electronics & Communication Engineering & Assistant Dean of Technology, University of Technology & Management, Shillong, Meghalaya, India

Abstract
The Security in sensor networks has become most important aspect along with low power as the sensors are unattended so there is more possibility of attack in WSN than usual networks, data aggregation security is an important task as if some false node injects a highly odd value it will affect the whole aggregation process, The paper reviews the need of security for data aggregation and propose an architecture which can eliminate the false values injection as well as provides end to end reliability and data freshness, the architecture is also energy optimized.

Keywords- data aggregation, sensor networks, energy efficiency, Hash

I. INTRODUCTION

Wireless sensor networks are normally unattended and self-configurable networks, which are composed of a few to thousands of lightweight and portable tiny sensing nodes. These networks are deployed in remote and hostile environments where these nodes can sense temperature, pressure, vibration, motion, sound and even the pollutant levels in targeted regions [2]-[9]. When the sensor node senses some behavior or phenomenon from the environment, transducer generates the signal for the sense data which is further processed and store by the installed microprocessor and after processing the signal, transceiver transmits this data to base station or some upper level aggregator node, through which the sensor node is wirelessly connected. Performing all this complex functionality, the size of the sensor node just vary from a shoebox to a dust of grain [3]. As the sensor nodes are normally targeted in an uncontrolled environments, physical approaching to deployed sensors is not possible, so the sensor nodes have the built in self-configuration functionality. Sensor nodes make use of multi hop routing algorithms, through which multiple sensing nodes simultaneously send the data to some other nodes or base station. Wireless sensor networks can use infrared or laser light, but these technologies require clear line of sight. So wireless sensor networks use radio communication technology for communication purpose, which does not require any clear line of sight [3]. The sense data from the

leaf nodes is routed through multiple routing paths and intermediate nodes and at the end it reaches to a base station. The base station may be a simply a computer or some specialized hardware which can perform some computation on it, store and forward data and can also respond to the data sending nodes.

The emergence of senor node architecture with its advance capabilities to control the different hardware units and its enhancements in low power and affordable computational devices has made sensor networks from dream to reality. Sensor networks are deployed in widespread targeted areas where they can work for many years and can sense the environment behavior and itsdifferent variables without any need to externally charge their installed batteries. Early sensor networks were used only for military purposes such as in battlefield surveillance and keeping track the motion and other different activities of the enemy in remote and unreachable areas. Due to the recent developments and enhancements in sensor networks technology, wireless sensor networks are not now merely limited to military applications.

Its typical applications are habitat monitoring, object tracking, remote controlling, traffic monitoring [4] and its most complex and mission critical application is to monitor and control the nuclear reactors and making preemptive actions in case of any problem occurs during its normal working functionality [5].

Sensor nodes are deployed normally in remote and unattended environments where their maintenance and battery changing or charging is not possible, so saving the battery power of the sensor node increases its lifetime. From energy perspective, communication is the most costly and expensive process in wireless sensor networks. Because of limitation of its design structure, sometimes it may possible that more than one sensor nodes have overlapping targeted regions which causes generation of redundant data packets. Data is correlated in terms of time and apace whenever it is sensed from the sensor nodes, so sending duplicate data values again and again may lead to expire sensor nodes early [8].

In wireless sensor networks, wireless medium is used for data transfer from source to destination, so the security is the prime factor that should be considered during the transmission of data, because anybody tuned to that particular frequency gain access to the sensed data [6]. Wireless sensor networks are normally deployed in unattended, untrusted and hostile environments where the sensor nodes and wireless communication links can be eavesdropped easily. Adversary by compromising only one sensor node or a wireless link can easily forge or alter the data [7]. Sensor nodes have limited computational power, small memory size and low storage capacity as compared to other networks so during the design of any security and routing protocol for wireless sensor networks, the resource limitation factor is always considered. Specially the energy is the scarcest resource of the wireless sensor networks which is directly concerned with the life time of the senor node, so normal public/private key algorithms, which require complex mathematical computation are not feasible for these networks [9]. The security of wireless sensor networks is broadly categorized into two main types, internal security and external security. Internal security is called data privacy as well, and it is maintaining the privacy of sensed data from the trusted sensor nodes inside the network. The second type of security is called external security or data security, in which sensed data is protected from outsiders such as adversaries or eavesdroppers [7].

In wireless sensor networks data from source to destination is transferred through some intermediate aggregator nodes and these nodes act like a dominating nodes. From security point of view these intermediate aggregation nodes hold the most critical position in wireless sensor networks. For example wireless sensor network is deployed in some temperature critical environment, which is divided into different zones and in each zone some sensor nodes are deployed and one of them acts like a aggregate cluster node.

In each zone of deployed environment, it is required to maintain the average temperature value to 25^{o}C, and if the average temperature value in any zone is greater or lesser than 25^{o}C, a temperature control system which is associated with base station begins to start working for maintaining the average temperature value. Leaf sensing nodes which are well equipped with some powerful security mechanism, after sensing the temperature value, perform some cryptographic function on it and send this cipher text to aggregate cluster node. Aggregate cluster node decrypts this cipher text into plain text, performs average aggregation function on the decrypted data and sends this aggregation value to the base station. If this trusted aggregate cluster node is compromised, the adversary can easily forge or alter the temperature values and disturb the whole temperature of the environment by sending fake and bogus aggregation results to temperature control system.

Providing a secure data aggregation, while preserving the security and privacy of the sense data is still challenging task because traditional security algorithms based on cryptography and public/private keys are very expensive and are not feasible in wireless sensor networks [9].

Depending on the encryption schemes, secure data aggregation protocols in wireless sensor networks are grouped into two main categorizes, Hop-by-Hop and End-to-End encryption protocols [7]. In hop-by-hop secure data aggregation, sensor nodes sense the data, encrypt it and then send this cipher text to the aggregator node, which then decrypts it, perform some data aggregation function on the received data from its child nodes and then again encrypts this aggregated data and sends it to the upper level node or base station. But in End-to-End secure data aggregation, the aggregation nodes do not have the decryption keys so data aggregation function at intermediate nodes is performed on the encrypted data [12]-[13].

II. LITERATURE REVIEW

Wireless sensor networks are emerging technologies currently being deployed in seismic monitoring, wild life studies, manufacturing and performance monitoring. These networks are densely deployed in some predetermined geographical area to self-organize into ad-hoc wireless network together and aggregate data [16]. These networks are composed of tiny sensor nodes which are powered by a small built in batteries and are deployed in hostile and uncontrolled environments where these networks has to work for many years[1]. Because of their limitations in energy, storage capacity, computation and communication, these networks pose unique security challenges [2]. [1] Proposed a strong security architecture using network coding and the use of hash functions, the architecture is energy optimized, but the architecture can work only for single hop network, our proposed architecture is an extension of [1] for multi hop networks.

Data aggregation can also reduce data packet size, number of data transmissions and the number of nodes involved to gather data [7]. So to save energy in sensor networks, data aggregation is put forward as an in-network processing, the data from source to destination is transferred through intermediate aggregator nodes, which perform aggregate function on the data and then sends aggregation results to a higher level aggregators [17]. As sensor networks use radio signals for communication purposes, anybody tuned to that particular frequency can forge or alter the data, so the

confidentiality of the transmitted data can be considered as the most critical [6]-[7]. Confidentiality in wireless sensor networks is provided through Homomarphic encryption and allows network data aggregation. Homomarphic encryption provides the way to calculate the aggregate over the encrypted values. But Homomarphic encryption does not provide data integrity, so by using public key elliptic curve cryptography, digital signatures can be used to provide data integrity as well [6]. The sensed data is not only to be protected externally but from internal nodes as well. So privacy of sensor data is another security issue in sensor networks. For preserving the privacy of sensed data and hiding its details from the trusted nodes, the algorithm uses the additive property of complex numbers. Sensor node is pre-deployed with the two numbers, one is real private seed and other is imaginary private seed. When the sensor node sense the data, it first adds it with the real private seed and gets 'a' then again it adds the obtained 'a' with the imaginary private seed 'bi' and data is transformed into complex number form, encrypts complex number form by using symmetric key and then transmits this cipher text to its parent [7]. Data involved in aggregation purposes should not only belong to trusted nodes but also follow reliable links. Reputation values are used to ensure that the data sending nodes are trusted and are not compromised. The nodes and the aggregators can sense the behavior of the neighboring sensor nodes in the environment. Depending on the sense behavior of sensor nodes a reputation value is calculated, which depends on the sensor ability of sensing, routing and availability. This web of trust relationship provides secure and reliable path for data transfer from sensor to aggregate node [8]. Algorithms proposed for security purposes, focus on secure data transfer from source to destination, but at the same time it is also needed that proposed algorithm sure to decrease the communication overhead by transferring less number of bits through network. Secure End-to-End data aggregation protocol uses the additive homomarphic encryption and performs its functionality by calculating the aggregation value from cipher text received from the leaf sensing nodes. This protocol is basically proposed to decrease the transmission of extra bits for the nodes that are not participating in the aggregation. The nodes that are not responding or have no data values for sending, use "0" as their sense value and sends this cipher text to the aggregate node. One variable is used for keeping track the quantity of the non-responding nodes [10].

The security issues such as authentication, integrity and data freshness become very crucial when sensors are deployed in hostile environments where they are prone to node failures and compromised [9]. Secure data aggregation (SDA) scheme is resilient to intruder devices and single node compromise. It is based on delayed aggregation and delayed authentication. Senor nodes need to buffer the data to authenticate it once the shared key is revealed by the base station. It also provides integrity and data freshness but data can be altered once the parent and child hierarchy is compromised [18]. Authentication can be provided both for end to end and hop by hop. For this purpose a pair of aggregated values is used, one is for authentication of hop-by-hop and the other one is for authentication of end-to-end. Besides such a pair, the ID list representing sensor nodes involves in a pair of aggregated values is produced and used to regenerate MAC to check integrity along the way at the base station. However, the concatenation of IDs will occur the overhead of data aggregation, particularly in scalability [19]. Node compromise in data aggregation is threatening as compromised sensing nodes will inject false data in aggregation. Injected false data and adversary node is identified by specified external nodes in parent child network hierarchy. The role of the external node is just to monitor the parent and children nodes and then report to the base station. Keys are pre-deployed and shared among nodes for authentication purpose and the judgment of false data nodes is done through decision of majority [20].

III. PROPOSED ARCHITECTURE

Here we will propose a security mechanism in which first odd data values are identified and then the fault or compromised nodes are permanently removed from the network through hash base authentication at aggregate node.

The proposed architecture is shown in Figure 1. Initially the value P_0 is predeployed in all the sensor nodes and this value is shared with the aggregator node as well. When the leaf sensing node A senses some data A_1 from the environment, it concatenates A_1 with predeployed P_0 and calculates hash on it, and this calculated hash and the sense data A_1 is sent to next sensing node B.

The format of the data from the node A to B is like

$$\mathbf{A_1 // H (A_1 // P_0)}$$

When the sensing node B receives the data from the node A, it first calculates the difference of B_1 and A_1 like $(B_1\text{-}A_1)$, and then concatenates it with the predeployed P_0, calculates its hash like H $(B_1\text{-}A_1$ // $P_0)$, and appends the difference of B_1 and A_1 and calculated hash with the old data packet received from the sensor node A, and sends this data to the next sensing node C.

The format of the data packet from node B to C will be like

$$(A_1 // H (A_1 // P_0)) // (B_1\text{-}A_1 // H (B_1\text{-}A_1 // P_0))$$

When the data is reached at the node C, it then calculates the difference of C1 and A1 (C1-A1) by using network coding and also concatenates C1-A1 with P0, calculates it hash and appends it with the packet received from the node B. The format of the data packet from node C to D will be like this.

$$(A_1 // H (A_1 // P_0)) // (B_1\text{-}A_1 // H (B_1\text{-}A_1 // P_0)) // (C_1\text{-}A_1 // H (C_1\text{-}A_1 // P_0))$$

The proposed architecture is shown in figure 1.

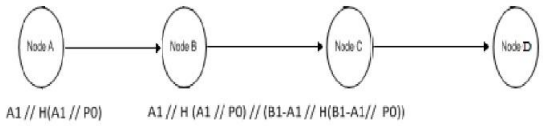

Figure 1: Proposed Architecture

At the end the data will reached to the aggregate node. Aggregate node will first check the received readings of sensor nodes, if the values lie in some acceptable range, then it will calculate hash on it and compare it with the received hash. If the result is same the received values will be included in the final aggregation.

As the aggregate node will again calculate the hash function on the received data values so it can easily detect odd / false data values and will not include it in the aggregation.

Next time each node will use the H (A1 // P0) as a value of P1, as this value is changed in each session of communication so data freshness is maintained and there is no need to transmit the value of P again and again to all the sensor nodes which will cause communication overhead, as only difference of the sensed values is sent to the aggregate nodes so less number of bits are used so communication overhead is decreased.

IV. CONCLUSION

Securing the data aggregation in wireless sensor is very important for accuracy sensitive applications like surveillance systems etc. We have proposed a network coding based architecture for providing security for data aggregation, the architecture provides end to end reliability without the use of acknowledgements, and the architecture can work in noisy and malicious environments.

References

[1] Shehzad Ashraf Ch, ZahidMehmood, Rashid Amin, Dr. Mohammad Alghobiri, Tahir Afzal Malik "Ensuring Reliability & Freshness in Wireless Sensor Networks" 2010 International Conference on Intelligent Network and Computing (ICINC 2010)

[2] Tanveer Zia and Albert Zomaya "Security Issues in Wireless Sensor Netwoks"

[3] Kay Romer and FriedemannMattern "The Design Space of Wireless Sensor Networks" NCCR-MICS, grant no. 5005-67322. IEEE Wireless Communicationss, Dec. 2004

[4] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson,"Wireless Sensor Networks for Habitat Monitoring," WSNA'02, Atlanta, Georgia, September 2002.

[5] N. Xu, S. Rangwala, K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A Wireless Sensor Network for Structural Monitoring," in Proceedings of the ACM Conference on Embedded Networked Sensor Systems, Baltimore, MD, and November 2004.

[6] Julia Albath and Sanjay Madria"Secure Hierarchical DataAggregation in Wireless Sensor Networks", IEEE Communications Society.

[7] RabindraBista, Kyoung-Jin Jo and Jae-Woo Chang "A New Approach to Secure Aggregation of Private Data in Wireless Sensor Networks" 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing.

[8] SuatOzdemir "Secure and Reliable Data Aggregation for Wireless Sensor Networks" H. Ichikawa et al. (Eds.): UCS 2007, LNCS 4836, pp. 102–109, 2007.Springer-Verlag Berlin Heidelberg 2007.

[9] Hani Alzaid, Ernest Foo, Juan Gonzalez Nieto "Secure Data Aggregation in Wireless Sensor Networks:a survey" 6th Australasian Information Security Conference (AISC 2008), Wollongong, Australia, 2008.

[10] A.S.Poornima and B.B.Amberker "SEEDA: Secure End-to-End Data Aggregation in Wireless Sensor Networks" 978-1-4244-7202-4/10/©2010 IEEE.

[11] Tanveer Zia and Albert Zomaya "A Security Framework for Wireless Sensor Networks" IEEE Sensors Applications Symposium Houston, Texas USA, 7-9 February 2006.

[12] C. Castelluccia, E. Mykletun, and G. Tsudik, "Efficient aggregation of encrypted data in

wireless sensor networks,"in The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, 2005, pp. 109–117.

[13] J. Girao, D. Westhoff, and M. Schneider, "CDA: Concealed Data Aggregation in Wireless Sensor Networks," 40th IEEE International Conference on Communications, May 2005.

[14] J. M. Kahn, R. H. Katz, and K. S. J. Pister, "Next century challenges: mobile networking for smart dust," MobiCom 99: Proc. 5th ACM/IEEE Intl. Conf. on Mobile computing and networking, pp. 271–2781, 1999.

[15] A. Ephremides and B. Hajek," Information theory and communication networks: an unconsummated union,"IEEE Trans. on Inform. Theory, vol. 44, no. 6, pp. 2416–2434, 1998.

[16] N. Hu, Randy R. K. Smith and P. G. Bradford 'Security for Fixed Sensor Networks' Proceedings of the 42nd annual Southeast regional Information Sciences, Prince Sultan College for Tourism conference, ACM Press, 2004, NY, USA .

[17] Ying Sang, Hong Shen 'Secure Data Aggregation in Wireless Sensor Networks' IEEE Proceedings ofthe Seventh International Conference.

Dr. D. Bhattacharya,finished his Master Degree in 2002 from Calcutta University in Electronics. He obtained his PhD from UK in 2007. Presently, he is working as Associate Professor in the Department of ECE and Assistant Dean of Technology at University of Technology & Management, Shillong, and Meghalaya. He has more than 7 years of experience in the field of Engineering Education and 5 years of experiences in Research. His area of research work in Antenna Design & Application of Digital Signal Processing in Bio-Medical Electronics research.