Volume 3, Issue 7, July 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Secure Reference Based Data Aggregation Protocol for Wireless Sensor Networks

Miriyala Markandeyulu
M. Tech in CSE
Vignan's Lara Institute of Technology & Science
Vadlamudi, Guntur, India

Guttikonda Prashanti
Asst. Professor in CSE
Vignan's Lara Institute of Technology & Science
Vadlamudi, Guntur, India

Abstract. In many sensor applications, the data has been collected from the individual nodes and it is aggregated at a base station or host computer. To reduce the energy consumption, many systems also can perform in-network aggregation of sensor data at the intermediate nodes enroute to the base station. The most existing aggregation algorithms and systems do not include any provisions for providing security, and consequently these systems are vulnerable to a large variety of attacks. In particular, the compromised nodes can be used to inject the false data that leads to incorrect the aggregates being computed at the base station. We are discussing the security vulnerabilities of data aggregation for systems, and present a survey of robust and secure aggregation protocols that are resilient to false data injection attacks.

Keywords. Sensor networks, aggregation, security.

1. Introduction

The Wireless sensor networks are usually composed for hundreds or thousands of less expensive, low-powered sensing devices with limited memory, communication resources and computational. These networks offer potentially low-cost solutions for an array of problems in both civilian applications and military including battlefield target tracking, environmental and health care monitoring, wildfire detection, and traffic regulation. Due to the low deployment cost requirement of wireless sensor networks, sensor nodes have simple hardware and severe resource constraints. Hence, it is a challenging task to provide efficient solutions to data gathering problem. Among these constraints,in designing wireless sensor network protocols ‘‘battery power” is the most limiting factor To reduce the power consumption of wireless sensor networks, several mechanisms are proposed like control packet elimination, radio scheduling, topology control,and most important one is data aggregation. The aim of Data aggregation protocols is to combine and summarize the data packets of several sensor nodes so that the amount of data transmission has reduced. An example for the data aggregation scheme is presented in Fig. 1 where a group of sensor nodes have collect information from the target region.

When a base station queries a network, instead of sending each sensor node's data to base station, one of its sensor nodes, called data aggregator, collects the information from its neighboring nodes, and aggregates them,and sends the aggregated data to the base particular station.over a multihop path. As illustrated by the example, data aggregation reduces the number of data transmissions thereby improving the bandwidth and energy utilization in the network.

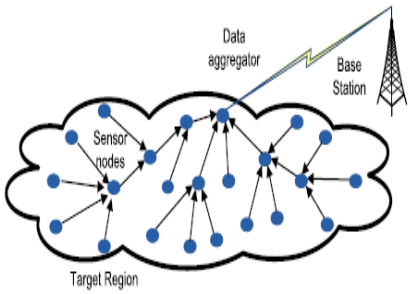

Fig. 1. Data aggregation in a wireless sensor network.

In wireless sensor networks, the benefit of data aggregation increases if the intermediate sensor nodes perform data aggregation incrementally when data are being forwarded to the base station.This continuous data aggregation operation improves the bandwidth and also energy utilization, it may affect other performance metrics negatively such as accuracy,delay, fault-tolerance, and also security. As the majority of wireless sensor network applications require a certain level of security, it is not possible to sacrifice the security for data aggregation. In addition, there is a strong conflict between security and data aggregation protocols. Security protocols require the sensor nodes to encrypt and

authenticate any sensed data prior to its transmission and prefer data to be decrypted by the base station. On the other hand, data aggregation protocols prefer plain data to implement data aggregation at every intermediate node so that energy efficiency is maximized. Moreover, data aggregation results in alterations in sensor data and it is a challenging task to provide source and data authentication along with the data aggregation.Hence due to this conflicting goals,the data aggregation and security protocols must be designed together, so that the data aggregation can be performed without sacrificing security.The necessity of implementing data aggregation and security together have led many researchers to work on secure data aggregation problem. In this paper, we aim to provide an extensive overview of secure data aggregation concept in wireless sensor networks by defining the main issues and covering the most important work in the area.Compared to general data aggregation problem which is a well researched topic in wireless sensor networks, secure data aggregation problem still has the potential to provide many interesting research opportunities. Hence, we also aim to give a starting point for researchers who are interested in secure data aggregation problem by presenting the open research areas and future research directions in the field.

Our contribution in this paper is twofold. First, we look at the data aggregation problem from the security perspective by giving a comprehensive literature survey. Second,based on the observations from the state-of-the-art secure data aggregation protocols, we discuss the open research areas and future research directions. Although there are couple of existing survey papers on data aggregation in wireless sensor networks [7,8], to the best of our knowledge, this is the first survey paper that focuses on solely secure data aggregation concept. In this paper, we cover many secure data aggregation protocols that are not covered by the previous survey papers. In addition, the open research areas and future research directions presented in this paper do not appear in the existing survey papers either. Nonetheless,we believe our paper will serve as a useful guide and starting point for the researchers who are interested in conducting research in the secure data aggregation area. The organization of the paper as follows: Section 2 starts with a brief summary of security requirements of wireless sensor networks and show how they relate with data aggregation process. Section 3 gives introductory information about data aggregation and summarize the most important work in the area. Section 4 presents ‘‘state-of-the-art’’ secure data aggregation protocols in wireless sensor networks.

In this section, a broad overview of secure data aggregation is given by evaluating each protocol based on the security requirements of wireless sensor networks. Section 5 defines open research areas and future research directions in secure data aggregation. Section 6 concludes the paper by emphasizing our contributions in this paper.

2. Security requirements of wireless sensor networks

Due to hostile environments and unique properties of wireless sensor networks, it was a challenging task to protect sensitive information transmitted by wireless sensor networks[1]. In addition to that, wireless sensor networks have security problems that traditional networks do not face these types of problems.Therefore,the security is an important issue for wireless sensor networks and there are many security considerations that should be investigated. In this section, we present the essential security requirements that are raised in a wireless sensor network environment and explain how these requirements relate with data aggregation process. Fig. 2 illustrates the interaction between wireless sensor network security requirements and data aggregation process.

2.1. Data confidentiality

In wireless sensor networks, data confidentiality ensures that secrecy of sensed data is never disclosed to unauthorized parties and it is the most important issue in mission critical applications. Authors of [4] state that a sensor node should not leak its readings to neighboring nodes. Moreover, in many applications, sensor nodes transmit highly sensitive data, e.g., secret keys; and therefore it is extremely important to build secure channels among sensor nodes.

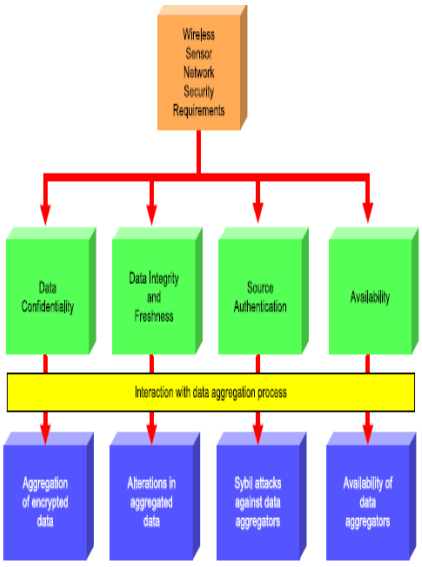

Fig. 2. Interaction between wireless sensor network security and data aggregation process

Public sensor information, such as sensor identities and public keys, should also be encrypted to some extent to protect against traffic analysis attacks. Furthermore, routing information must also remain confidential in certain cases as malicious nodes can use this information to degrade the network's performance. The standard approach for keeping sensitive data secret is to encrypt the data with a secret key that only intended receivers possess, hence achieving confidentiality. However, data aggregation protocols usually cannot aggregate encrypted data.

Therefore, such data aggregation protocols must decrypt the sensor data to perform data aggregation and encrypt the aggregated data before transmitting it.This decryption/encryption of sensor data at data aggregators not only results in delay and energy consumption but also prevents end-to-end data confidentiality.

2.2. Data integrity and freshness

Although data confidentiality guarantees that only intended parties obtain the un-encrypted plain data, it does not protect data from being altered. Data integrity guarantees that the message being transferred is never corrupted. A malicious node may just corrupt messages to prevent network from functioning properly. In fact, due to unreliable communication channels, data may be altered without the presence of an intruder. Thus, message authentication codes or cyclic codes are used to prevent data integrity.

Data aggregation results in alterations of data; therefore,it is not possible to have end-to-end integrity check when data aggregation is employed. Moreover, if a data aggregator is compromised, then it may corrupt sensor data during data aggregation and the base station has no way of checking the integrity of this aggregated sensor data. Providing data integrity is not enough for wireless communication because compromised sensor nodes are able to listen to transmitted messages and replay them later on to disrupt the data aggregation results. Data freshness protects data aggregation schemes against replay attacks by ensuring
that the transmitted data is recent.

2.3. Source authentication

Since wireless sensor networks use a shared wireless medium, sensor nodes need authentication mechanisms to detect maliciously injected or spoofed packets. Source authentication enables a sensor node to ensure the identity of the peer node it is communicating with. Without source authentication, an adversary can masquerade a node,thus gaining the unauthorized access to resource and sensitive information and interfering with the operation of other nodes. Moreover, a compromised node may send data toits data aggregator under several fake identities so that the integrity of the aggregated data is corrupted. Faking multiple sensor node identities is called Sybil attack and it poses significant threat to data aggregation protocols [5]. If only two nodes are communicating, authenticationcan be provided by symmetric key cryptography. The receiver and the sender share a secret key to compute the message authentication code (MAC) for all transmitted data. However, data aggregators may need broadcast authentication which requires more complex techniques,such as lTESLA [5].

2.4. Availability

Availability guarantees the survivability of network services against Denial-of-Service (DoS) attacks. A DoS attack can be launched at any layer of a wireless sensor network and may disable the victim node(s) permanently. In addition to DoS attacks, excessive communication or computation may exhaust battery charge of a sensor node.Consequences of availability loss may be catastrophic. For example, in a surveillance, application of battle field if the availability of some sensor nodes cannot be provided, this may lead to an enemy invasion. Wireless sensor networks are deployed with high node redundancy to tolerate such availability losses. Since data aggregators has collect the data of a number of sensor nodes and sends the aggregated data to the base station, availability of data aggregators is most important than regular sensor nodes. Thus, in wireless sensor networks, intruders launch DoS attacks with the aim of preventing data aggregators from performing their task so that some part of a network losses its availability.

3. Data aggregation

In the typical wireless sensor network, a large number of sensor nodes are collecting application specific information from the environment and this information is transferred to a central base station where it is processed, analyzed, and used by the application. In these resource constrained networks, the general approach is to jointly process the data generated by different sensor nodes while being forwarded toward the base station [8]. Such distributed in-network processing of data is generally referred as data aggregation and involves combining the data that belong the same phenomenon. The main objective of data aggregation is to increase the network lifetime by reducing the resource consumption of sensor nodes (such as battery energy and bandwidth). While increasing network lifetime, data aggregation protocols may degrade important quality of service metrics in wireless sensor networks, such as data accuracy, fault-tolerance, latency and security. Therefore, the design of an efficient data aggregation protocol is an inherently challenging task because the protocol designer must trade off between energy efficiency,latency, data accuracy , security and fault-tolerance . In order to achieve this trade off, data aggregation techniques are very tightly coupled with how packets are routed through the network. Hence, the architecture of the sensor network plays a vital role in the performance of different data aggregation protocols. There are several protocols that allow routing and aggregation of data packets simultaneously.

These protocols can be categorized into two parts: tree-based data aggregation protocols and cluster-based data aggregation protocols. Earlier work on data aggregation focused on improving the existing routing algorithms so as to make data aggregation possible. As a result, many data aggregation protocols based on shortest path tree structure have been proposed [10,17,46]. To reduce the latency due to tree-based data aggregation, recent work on the data aggregation tends to group sensor nodes into clusters so that data are aggregated in each group for improved efficiency.

3.1. Tree-based data aggregation protocols

The simplest way to achieve distributed data aggregation is to determine some data aggregator nodes in the network and ensure that the data paths of sensor nodes include these data aggregator nodes. Such tree-based data aggregation techniques have been extensively studied in the literature [9–18]. The main issue of tree-based data aggregation protocols is the construction of an energy efficient data aggregation tree. Fig. 3 illustrates an example of tree-based data aggregation. Greedy Incremental Tree (GIT) [9] is a data-centric routing protocol that allows data aggregation based on Directed Diffusion [10]. In [11] GIT is compared with two other data-centric routing schemes, namely Center at Nearest Source (CNS) [3] and Shortest Path Tree (SPT) [10]. The simulation results show that GIT performs the best in terms of average number of transmissions Another SPT based data aggregation protocol that promotes the parent energy-awareness is proposed in [12].

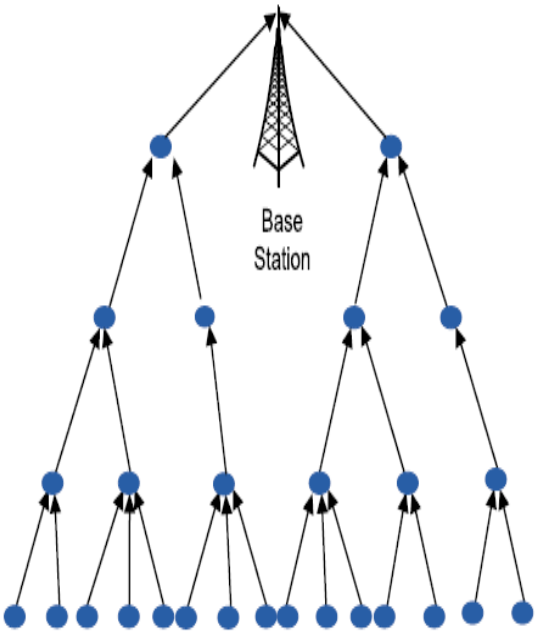

Fig. 3. Tree-based data aggregation.

In this protocol, parent selection is based on sensor nodes' distance to the base station and their residual energy level. There are also data aggregation protocols that consider information theory as routing metric. For example, [13] proposes a centralized approach that routes the packet based on their joint entropies. However, this protocol is not feasible as it depends on the global knowledge of the information entropy of each sensor node as well as the joint entropy of each node pair. In the rest of this subsection,we present some of the important work in tree-based data aggregation in detail.

In [14], Madden et al. proposed a data-centric data aggregation framework called Tiny AGgregation Service (TAG), which is based on shortest path tree routing. TAG is specifically designed for monitoring applications and allows an adjustable sleep schedule for sensor nodes. To achieve this, parent nodes let their children know about the waiting time for transmission. Also, parent nodes cache their children's data to prevent from data loss. TAG performs data aggregation in two phases. In the first phase, called distribution phase, base station queries are disseminated to the sensor nodes, then in the second phase, called collection phase, the aggregated sensor readings have routed up the aggregation tree. During the distribution phase, a message is broadcasted by the base station requiring sensor nodes to organize a routing tree so that the base station can send its queries. Each message has the field that specifies the level or distance from the root of the sending node (the level of the root is equal to zero). When a node that does not belong to any level receives this message, it sets itsown level by incrementing the current level in the message by one and assigns the sender as its parent.

This process continues until all sensor nodes in the network joins the tree and have a parent. This messaging periodically repeated to keep the tree structure updated. Once the tree is formed, then the base station queries the network via the aggregation tree. Sensor nodes use their parents when replying to base station queries. TAG employs an SQL like language to query the network. Each query specifies the quantity that needs to be collected, aggregation function and the sensor nodes that need to perform the data collection.

4. Secure data aggregation

Like any other wireless sensor network protocol, data aggregation protocols must satisfy the security requirements explained in Section 2. However, the resource constrained sensor nodes and necessity of plain data for aggregation process pose great challenges when implementing security and data aggregation together. Security requirements of wireless sensor networks can be satisfied using either symmetric key or asymmetric key cryptography.

Due to resource constraints of sensor nodes, symmetric key cryptography is preferable over asymmetric key cryptography. Hence, the necessity of implementing the data aggregation and security using symmetric key cryptography algorithms have led many researchers to work on secure data aggregation problem [26–34]. In these protocols,security and data aggregation are achieved together in a hop-by-hop fashion. That is, data aggregators must decrypt every message they receive, aggregate the messages according to their corresponding aggregation function, and encrypt the aggregation result before forwarding it. In addition, these schemes require data aggregators to establish secret keys with their neighboring nodes. Therefore, hop-by-hop secure data aggregation protocols cannot provide data confidentiality at data aggregators and result in latency because of the decryption/encryption process. Inorder to mitigate the drawbacks of hop-by-hop secure data aggregation protocols, a set of data aggregation protocols is proposed [36–41]. The proposed protocols perform data aggregation without requiring the decryption of the sensor data at data aggregators. While some of these protocols use symmetric cryptography, others employ asymmetric key cryptography functions, such as [42,43], that are suitable for resource constrained sensor nodes.

As data aggregators do not have to decrypt sensor data to perform aggregation, the protocols proposed in [36–41] provide end-to-end data confidentiality and result in less latency compared to hop by- hop secure data aggregation protocols. On the other hand, the downside of the data aggregation protocols that do not require the decryption of sensor data is that they are applicable to only a set of aggregation functions, such as sum and average. In what follows, we classify and explain the secure data aggregation protocols based on the requirement of decrypting sensor data at data aggregators.

5. Open research issues and future research directions

In this paper, we present a comprehensive overview of secure data aggregation concept in wireless sensor networks. We survey the state-of-the-art data aggregation protocols and categorized them based on network topology and security.The presented research addresses so many problems of data aggregation, there are still many research areas that needs to be associated with the data aggregation process, especially from the security point of view. As for the general data aggregation concept, the relation between routing mechanisms and data aggregation protocols have been well studied as they are highly correlated topics. In addition to diffusion and tree-based data aggregation protocols, many cluster-based, the data aggregation protocols are route aggregated data over cluster heads have been proposed. Although, these protocols shown to be very efficient in static networks in which the cluster structures do not change for a sufficiently long time, in dynamic networks they perform quite poorly. Hence, data aggregation in dynamic environments is a possible future research direction. The impact of the sensor node heterogeneity over the data aggregation protocols is another unexplored research area [40]. The protocols that use powerful sensor nodes as data aggregators presented promising results.However, determining locations of these powerful nodes for the best data aggregation results needs further research.

In order to provide end-to-end security, privacy homomorphism based secure data aggregation protocols have drawn considerable attention recently. However, the design and implementation of resource efficient privacy homomorphic aggregation functions yet to be explored. Many existing public key cryptography based privacy homomorphic functions are not feasible for resource limited sensor nodes. Hence, in some secure data aggregation schemes elliptic curve cryptography is employed [36].However, these elliptic curve cryptography based privacy homomorphic functions can only work for some specific query-based aggregation functions, e.g., sum, average, etc.

Therefore, design of efficient privacy homomorphic functions that are able to work with all types of data aggregation functions needs to be explored. In addition, for certain wireless sensor network settings where real-time data delivery is demanded, symmetric key cryptography based privacy homomorphic encryption schemes are recommended [38,37]. But, there are not many symmetric key based privacy homomorphic schemes. Hence, exploration of symmetric key cryptography based privacy homomorphic functions in the secure data aggregation concept is another promising research area. Using ‘‘digital watermarking” schemes to replace the expensive privacy homomorphic functions is a newly introduced concept in secure data aggregation [41]. However, this method allows only one way authentication of sensor data at the base station.Hence, investigation of two-way authentication by using watermarking techniques that will allow in-network secure data aggregation in the network may be a good research direction.

6. Conclusion

This paper provides a detailed review of secure data aggregation concept in the wireless sensor networks. To give the motivation behind the secure data aggregation, first, the security requirements of wireless sensor networks are presented and the relationships between data aggregation concept and these security requirements are explained.Second, an extensive literature survey is presented by summarizing the state-of-the-art data aggregation protocols. Based on this extensive literature survey, open research areas and future research directions are given.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, IEEE Commun. Mag. 40 (8) (2002) 102–114.

[2] J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey,Comput. Networks 52 (12) (2008) 2292–2330.

[3] K. Akkaya, M. Demirbas, R.S. Aygun, The Impact of Data Aggregation on the Performance of Wireless Sensor Networks, Wiley Wireless Commun. Mobile Comput. (WCMC) J. 8 (2008) 171–193.

[4] J. Newsome, E. Shi, D. Song, A. Perrig, The Sybil attack in sensor networks: analysis and defenses, in: Proceedings of the Third IEEE/ACM Information Processing in Sensor Networks (IPSN'04), 2004, pp.259–268.
[5] A. Perrig, R. Szewczyk, D. Tygar, V. Wen, D. Culler, SPINS: security protocols for sensor networks, Wireless Networks J. (WINE) 2 (5) (2002) 521–534.
[6] Crossbow Technologies Inc. <http://www.xbow.com>.
[7] E. Fasolo, M. Rossi, J. Widmer, M. Zorzi, In-network aggregation techniques for wireless sensor networks: a survey, IEEE Wireless Commun. 14 (2) (2007) 70–87.
[8] R. Rajagopalan, P.K. Varshney, Data aggregation techniques in sensor networks: a survey, IEEE Commun. Surveys Tutorials 8 (4) (2006).
[9] C. Intanagonwiwat, D. Estrin, R. Govindan, J. Heidemann, Impact of network density on data aggregation in wireless sensor networks, in: Proceedings of the 22nd International Conference on Distributed Computing Systems, 2002, pp. 457–458.
[10] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, F. Silva, Directed diffusion for wireless sensor networking, in: IEEE/ACM Transactions on Networking, vol. 11, 2003, pp. 2–16.
[11] B. Krishnamachari, D. Estrin, S. Wicker, The impact of data aggregation in wireless sensor networks, in: Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops, 2002, pp. 575–578.
[12] M. Ding, X. Cheng, G. Xue, Aggregation tree construction in sensor networks, in: Proceedings of the 58th IEEE Vehicular Technology Conference, vol. 4, 2003, pp. 2168–2172.
[13] R. Cristescu, B. Beferull-Lozano, M. Vetterli, On network correlated data gathering, in: Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 4, 2004, pp. 2571–2582.
[14] S. Madden et al., TAG: A Tiny AGgregation Service for Ad Hoc Sensor Networks, OSDI, Boston, MA, 2002.
[15] B. Zhou et al., A Hierarchical Scheme for Data Aggregation in Sensor Network, IEEE ICON 04, Singapore, 2004.
[16] M. Lee, V.W.S. Wong, An Energy-Aware Spanning Tree Algorithm for Data AggrAegation in Wireless Sensor Networks, IEEE PacRrim, Victoria, BC, Canada, 2005.

Authors Profile

Miriyala Markandeyulu Persuing M.Tech CSE in Vignan's Lara Institute of Technology & Science, Vadlamudi, Guntur. Email-markyellow9@gmail.com

G.Prashanthi Working As Asst. Professor in Vignan's Lara Institute of Technology & Science, Vadlamudi,Guntur. Email-prashantiguttikonda77@gmail.com

