International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 3 Issue: 12
ISSN: 2454-4248
92 – 96

Data Aggregation and Security Issues in Wireless Sensor Networks:A Survey

Nirbhay Kumar Chaubey
Associate Professor of Computer Science,
S.S.Agrawal Institute of Computer Science, Affliated to Gujarat Technological University,
Navsari, Gujarat, India-396445
e-mail: nirbhay@ieee.org

Abstract-Wireless sensor networks (WSNs) are normally composed of low powered, inexpensive device that is constrained in terms of memory, computation and communication. To reduce communication overhead and resource consumption in WSN, data aggregation is used to increase network lifetime. Hence, the design of an efficient data aggregation protocol is an important issues and inherently challenging task of robust WSN. In this paper, various types of Data Aggregation in WSN, security issues are studied and also author presented possible key research issues of WSN data aggregation.

Keywords- *Survey, Security, Data Aggregation, Secure data aggregation, wireless sensor networks*

*****

I. INTRODUCTION

The wireless sensor network (WSN) is an ad-hoc network. It consists of small light weighted, low powered wireless nodes called sensor nodes with limited memory, computational, and communication resources [1] and it measures physical parameters such as sound, pressure, temperature, and humidity. Wireless sensor network(WSN) led to a variety of applications such as habitat monitoring and target tracking. However, data communication between nodes consumes a large portion of the entire energy consumption of the WSNs. Consequently, data aggregation techniques can significantly help to reduce the energy consumption by eliminating redundant data travelling back to the base station.

Recently, the focus is more on Wireless Sensor Networks. Wireless sensor networks are consisting of numerous light weight and tiny sensor nodes with limited power, storage, communication and computation capabilities. Architecture of Wireless sensor network is shown in the figure 1.

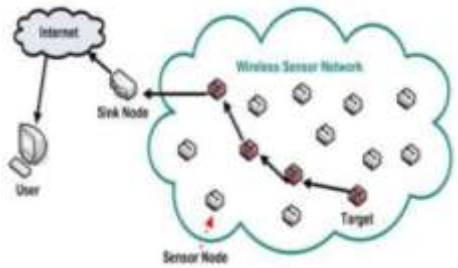

Fig. 1 Wireless sensor networks Architecture [1]

A sensor node has the components include a sensing unit, a processing unit, a transceiver unit and a power unit, additionally, they may also have application dependent components like location finding system, a power generator and a mobilize Users.

The rest of the paper is organized as follows: In section 2, introductory information about data aggregation is given and various existing data aggregation protocol are studied. In section 3, issues of Data Aggregation in WSN are discussed. Section 4 discusses different types of attacks on WSN aggregation and finally, section 5 concludes the paper.

II. DATA AGGREGATION IN WSN

In Wireless sensor network, application specific information is being collected from the environment by a large number of sensor nodes and further transferred to a central base station where it is processed, analyzed, and used by the application.

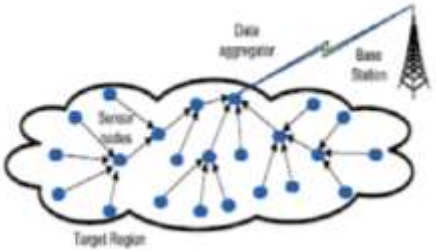

Fig. 2 Data Aggregation in Wireless sensor networks [2]

The general approach in these resource constrained networks is to jointly process the data which is generated by different sensor nodes while being forwarded toward the base station [2]**.** Such kind of distributed in-network processing of data is generally referred as data aggregation and involves combining of data that belongs to the same phenomenon. 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 [1]-[4]. These protocols can be categorized into four parts (i) Centralized Approach (ii) In-Network Aggregation (iii) Tree-Based Approach and and (iv) Cluster-Based Approach.

 ISSN: 2454-4248
92 – 96

Centralized Approach: It is an address centric approach in which data is sent by each node to a central node via the shortest possible route using a multihop wireless protocol.

In-Network Aggregation: The process of gathering and routing information through a multi-hop network and further processing the same data at intermediate nodes to reduce the resource consumption and thereby increasing network lifetime. There are two approaches for in-network aggregation: with size reduction and without size reduction. In-network aggregation with size reduction refers to the process of combining & compressing the data packets received by a node from its neighbors in order to reduce the packet length to be transmitted or forwarded towards sink. In-network aggregation without size reduction refers to the process merging data packets received from different neighbors in to a single data packet but without processing the value of data.

Tree-Based Approach: Earlier research work on data aggregation mainly aimed on improving the existing routing algorithms to make data aggregation possible. So many data aggregation protocols based on the shortest path tree structure are proposed [3], [4] and [5]. In this all nodes are organized in form of tree means hierarchical, with the help of intermediate node we can perform data aggregation process and data transmit leaf node root node. Tree based data aggregation is suitable for applications which involve in network data aggregation. One of the main aspects of tree-based networks is the construction of an energy efficient data-aggregation tree [6–13].

Madden et al. proposed a data-centric data aggregation framework called Tiny Aggregation Service (TAG), which is based on shortest path tree routing [10]. The basic objective of designing TAG was to monitor applications and to permit an adjustable sleep schedule for sensor nodes. Children node must be aware about the waiting time for transmission from the parent node. Also, parent nodes cache their children's data to prevent from data loss. TAG performs data aggregation in two phases (i) distribution phase wherein base station queries are disseminated to the sensor nodes and (ii) collection phase, the aggregated sensor readings are routed up the aggregation tree.

Kiran Maraiya, Kamal Kant, et. al. proposed Directed diffusion (DD) an information aggregation paradigm for wireless device networks [14]. This is a data-centric and application aware paradigm, within the sense that all information generated by the sensor nodes is called by attribute-value pairs. Such a scheme combines the information coming back from totally different sources en-route to the sink by eliminating redundancy and minimizing the amount of transmissions.

A modified version of directed diffusion, called Enhanced Directed Diffusion (EDD), is proposed in [11] which integrates directed diffusion with a cluster-based architecture so that the efficiency of the local interactions during gradient set up phase increases. Another similar protocol is proposed in [12].

In [4], athoor proposed a Power-Efficient GAthering in Sensor Information Systems (PEGASIS) that organizes sensor nodes in a chain for the purpose of data aggregation. In PEGASIS, every data aggregation chain has a leader that is responsible to transmit aggregated data to the base station. In order to evenly distribute the energy expenditure in the network, sensor nodes take turns acting as the chain leader. The chain building process starts from the sensor node furthest from the base station and continues towards the base station. When a node dies, the chain is reconstructed to bypass the dead node. In a sensor node chain, each sensor node receives data from a neighbor and aggregates it with its own reading by generating a single packet that has the same length with the received data. This process is repeated along the chain and the leader adds its own data into the packet and sends it to the base station directly.

In 2003, M. Ding, X. Cheng et. al. proposed Energy-Aware Distributed Aggregation Tree (EADAT) based on an energy-aware distributed heuristic[8]. The base station is the root of the aggregation tree hence it initiates the tree forming by broadcasting a control message has the fields: ID, parent, power, status, and hopcount. EADAT algorithm makes no assumption on local network topology, and is based on residual power. It makes use of neighboring broadcast scheduling and distributed competition among neighbors. Author confirmed through the simulation results that EADAT is a very efficient and effective.

Cluster-Based Approach: In cluster-based data aggregation approach, whole network is divided in to various clusters. In each cluster, a cluster head is elected in order to aggregate data locally and transmit the aggregation result to the base station. Cluster heads can communicate with the sink directly via long range radio transmission. However, this is quite inefficient for energy constrained sensor nodes. Thus, cluster heads usually form a tree structure to transmit aggregated data by multihopping through other cluster heads which results in significant energy savings. In last few year, several cluster based data aggregation protocols have been proposed [15–18].

W.B. Heinzelman, A.P. Chandrakasan proposed a self-organizing and adaptive clustering protocol, called Low-Energy Adaptive Clustering Hierarchy (LEACH)[15]. The advantage of ransomization is taken by the protocol LEACH to evenly distribute the energy expenditure among the sensor nodes. This is a clustered approach where data aggregation points set as cluster heads. This protocol consists of two phases. Cluster structures are formed in the first phase. Then, in the second phase, cluster heads aggregate and transmit the data to the base station.

O. Younis, S. Fahmy proposed HEED: a hybrid, energy-efficient distributed clustering approach for ad hoc sensor networks [16]. For the selection of cluster head, HEED takes

 ISSN: 2454-4248
92 – 96

the advantage of the availability of multiple power levels at sensor nodes. In fact, a combined metric that is composed of the node's residual energy and the node's proximity to its neighbors. The average of the minimum power level required by all sensor nodes within the cluster to reach the cluster head is defined by the HEED. This is called Average Minimum Reachability Power (AMRP) which is used to estimate the communication cost in each cluster.

Y. Yao, J. Gehrke proposed the Cougar approach to in-network query processing in sensor network that performs periodic per hop data aggregation and it is suitable for applications where sensor nodes continuously generate correlated data[17]. Once the cluster data is aggregated by the cluster heads, they send the local aggregated data to a gateway node.

This approach has a unique cluster head election procedure wherein it selects the cluster heads based on more than one metric and allows sensor nodes to be more than one hop away from their cluster heads. In this technique, synchronization is correctly used to aggregate the data.

S. Chatterjea, P. Havinga, proposed a hybrid approach Clustered Diffusion with Dynamic Data Aggregation (CLUDDA) [18] which combines clustering with diffusion mechanisms. In this approach, base station initiates query definitions inside interest messages . In order to generate a proper response each interest message contains the definition of the query that describes the operations required to be performed on the data components. Interest transformation reduces the processing overhead by utilizing the existing knowledge of queries. In CLUDDA, any cluster head that has the knowledge of query definition can perform data aggregation, and hence the aggregation points are dynamic. Cluster heads also keep a list of the addresses of neighboring nodes from which the data messages originated. These addresses are used to propagate interest messages directly to specific nodes instead of broadcasting.

III. ISSUES IN DATA AGGREGATION

Data Aggregation in wireless sensor network is an important technique as well as security to aggregated data is an important issue. Data aggregation in Wireless sensor Network refers to exploit the sensed data from the sensors to the gateway node.

Some of the important application such as military surveillance and various life critical application data transmission, data aggregation, and data reception must be in a secured and efficient [1], [19-23], [24 –25]. This section discusses the important security issues while doing in network data aggregation in WSN are as follows:

Data Confidentiality: Data confidentiality ensure that information content is never disclosed to unauthorized parties and is most important issue for mission critical application. So, information should be sent in an encrypted form to provide secrecy. And this encryption should be done by the secret key such that intended party that has key can only open and read data. Hop-by-hop encryption and end-to-end encryption are the two methods used for data confidentiality. In the hop-by-hop basis, any aggregator point needs to decrypt the received encrypted data, apply some sort of aggregation function, encrypt the aggregated data, and send it to the upper aggregator point. However, this kind of confidentiality implementation is not practical and very tedious for the WSN since it requires extra computation. On the other basis, the aggregator does not need to decrypt and encrypt data and instead of this, it needs to apply the aggregation functions directly on the encrypted data by using homomorphic encryption.

Data integrity: Data integrity ensures that the message being transmitted has not been modified either maliciously or accidentally. The malicious node can alter the sensed information to affect the overall aggregation results. Moreover, even without the existence of a malicious node, data might be damaged or lost due to the nature of the wireless environment.

Data Accuracy: Any aggregation scheme is to provide an aggregated data as accurately as possible since it is worth nothing to reduce the number of bits in the aggregated data but with very low data accuracy. A trade-off between data accuracy and aggregated data size should be considered at the design stage because higher accuracy requires sending more bits and thus needs more power.

Data freshness: This ensures that the data are recent and no old messages have been replayed, thereby protecting data aggregation protocols against replay attacks. In this kind of attack, the adversary can replay the distributed shared key and mislead the sensor about the current key used to secure sensing information and aggregated results.

Data availability: This gives guarantees that the network is alive and a node has the ability to use the resources. Further, in the presence of malicious nodes, it is highly recommended that the network react to these compromised nodes and eliminate them. Once an attacker gets into the WSN by compromising a node, the attack can affect the network services and data availability, especially in those parts of the network where the attack has been launched. Moreover, the data aggregation security requirements should be carefully implemented to avoid extra energy consumption. If no more energy is left, the data will no longer be available.

Source authentication: It allows a receiver to confirm that whether the received data sent by the actual sender or not. The authentication mechanism is needed to detect maliciously injected and spoofed packet. Without source authentication an adversary could masquerade a node and hence gaining unauthorized access to the resources and sensitive information and it can perform operations to other nodes.

 ISSN: 2454-4248
92 – 96

Non-repudiation: It ensures that a transferred packet has been sent and received by the person claiming to have sent and received the data packet. In secure aggregation schemes, once the aggregator sends the aggregation results, it should not be able to deny sending them. This gives the base station the opportunity to determine what causes the changes in the aggregation results.

Secure Node Localization: Node localization is very important issue in WSN so it should be kept secure and should not be accessed by malicious node. If location of sensor node is revealed to malicious node then all routing information also revealed. Further, wireless sensor network always needs location information accurately and automatically. However, there is a chance that attacker can manipulate no secured location information by reporting false signal strengths and replaying signals, etc. easily.

Time Synchronization: The majority of wireless sensor network applications rely on some form of time synchronization. In order to conserve power, an individual sensor‘s radio may be turned off for periods of time. Furthermore, sensors may wish to compute the end-to-end delay of a packet as it travels between two pair-wise sensors. A more collaborative sensor network may require group synchronization for tracking applications, etc.

Self-Organization: A wireless sensor network requires that every sensor node be independent and flexible enough to be self-organizing and self-healing according to different situations. In WSN, no fixed infrastructure available for the purpose of network management and this inherent feature brings a great challenge to wireless sensor network security as well.

IV. ATTACKS ON WSN AGGREGATION

WSNs are vulnerable to various types of attacks because of the character of transmission medium, remote and hostile deployment location, and moreover lack of physical security in each node[26] – [29]. Attacks that might affect the aggregation in the WSN are discussed in this section.

Denial of Service Attack: This attack on the WSN by transmitting radio signals that interfere with the radio frequencies used by the WSN. When the attacker capability increases, then it can affect larger portions of the network. Denial of Service Attack (DoS) can be an aggregator that refuses to aggregate and prevents data from traveling into the higher levels.

Node Compromise: In this type of attack, the attacker gain control over the deployed sensor node and extract the information stored on it which is sometimes called supervision attack. Considering the data aggregation scenario, once a node has been taken over, all the secret information stored on it can be extracted.

Node Subversion: Capturing of one node cause the reveal of the secret information and it may cause the compromise of the whole sensor network.

Sybil Attack: In this attack, attacker can make multiple identity within the network. It affects aggregation technique in many ways. An adversary may create multiple identities to generate additional votes in the aggregator election phase and select a malicious node to be the aggregator. Consequently, the aggregated result may be affected if the adversary is ble to generate multiple entries with different readings.

Selective Forwarding Attack: In this attack, a compromised sensor node may refuse to forward received messages. It is up to the attackers that control the compromised node to either forward the received messages or not.

In the aggregation context, any compromised intermediate nodes have the ability to launch the selective forwarding attack and this subsequently affects the aggregation results.

Replay Attack: In replay attack, an attacker records some traffic from the network without even understanding its content and replays them later on to mislead the aggregator and consequently the aggregation results will be affected.

Stealthy Attack: In this attack, the attacker node inject false data into the network without revealing its existence. In a data aggregation scenario, the injected false data value leads to a false aggregation result. A compromised node can report biased values, and perform a Sybil attack to affect the aggregation result.

Injection Attack: In this attack, adversary injects the wrong data into the network, consequently, in the process of aggregation this wrong data will result in false aggregated data.

Sinkhole Attack: In this attack, Sink is a high capability resource node. So attacker places himself in a network with high capability resources in order to confuse other nodes. As a result all data passed to attackers.

Wormhole Attack : A wormhole is low latency link between two portions of a network over which attacker replays network message.

Hello flood Attack : In this attack, attacker broadcasts HELLO packets with high transmission power to sender or receiver. The node receiving the message assumes that the sender node is nearest to them and send packet by this node. By this attack congestion occur in network.

Passive Information Gathering : An attacker with powerful resources (such as powerful receiver well designed antenna) can pick off the data stream strong encryption is the one of the solution to prevent from this attack.

V. CONCLUSION

This paper present detailed reviewed of wireless sensor network, architecture concept of data aggregation and extensive research survey of various data aggregation

 ISSN: 2454-4248
92 – 96

protocols of wireless sensor network. Security issues and attacks in data aggregation of WSN are summarized in the paper.

REFERENCES

[1] Suat Ozdemir and Yang Xiao, Secure data aggregation in wireless sensor networks: A comprehensive overview, Elsevier, Computer Networks 53 (2009) pp. 2022–2037.

[2] R. Rajagopalan, P.K. Varshney, Data aggregation techniques in sensor networks: a survey, IEEE Commun. Surveys Tutorials 8 (4) (2006).

[3] 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.

[4] S. Lindsey, C. Raghavendra, K.M. Sivalingam, Data gathering algorithms in sensor networks using energy metrics, IEEE Trans.Parallel Distrib. Sys. 13 (9) (2002) 924–935.

[5] Y. Xu, J. Heidemann, D. Estrin, Geography-informed energy conservation for ad hoc routing, in: Proceedings of the CM/SIGMOBILE MobiCom, 2001, pp.70–84.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] S. Madden et al., TAG: A Tiny AGgregation Service for Ad Hoc Sensor Networks, OSDI, Boston, MA, 2002.

[11] B. Zhou et al., A Hierarchical Scheme for Data Aggregation in Sensor Network, IEEE ICON 04, Singapore, 2004.

[12] 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.

[13] G. Di Bacco, T. Melodia, F. Cuomo, A MAC Protocol for Delay-Bounded Applications in Wireless Sensor Networks, Med-Hoc-Net,Bodrum, Turkey, 2004.

[14] Kiran Maraiya, Kamal Kant, Nitin Gupta "Architectural Based Data Aggregation Techniques in Wireless Sensor Network: A Comparative Study", International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 3 Mar 2011

[15] W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, IEEE Trans. Wireless Commun. 1 (4) (2002) 660–670.

[16] O. Younis, S. Fahmy, HEED: a hybrid, energy-efficient distributed clustering approach for ad hoc sensor networks, IEEE Trans. Mobile Comput. 3 (4) (2004) 366–379.

[17] Y. Yao, J. Gehrke, The Cougar approach to in-network query processing in sensor networks, ACM SIGMOD Rec. 31 (3) (2002) 9–18.

[18] S. Chatterjea, P. Havinga, A dynamic data aggregation scheme for wireless sensor networks, in: Proceedings of the Program for Research on Integrated Systems and Circuits, Veldhoven, The Netherlands, 2003.

[19] 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.

[20] Perrig, R. Szewczyk, D. Tygar, V. Wen, D. Culler, SPINS: security protocols for sensor networks, Wireless Networks J. (WINE) 2 (5) (2002) 521–534.

[21] Xiangqian Chen, Kia Makki, Kang Yen, and Niki Pissinou, " Sensor Network Security: A Survey", IEEE Communications Surveys & Tutorials, Vol. 11, No. 2, Second Quarter, 2009.

[22] Yong Wang, Garhan Attebury, and Byrav Ramamurthy,"A Survey Of Security Issues In Wireless Sensor Networks", IEEE Communication ,2nd quarter, volume 8, NO. 2,2006.

[23] Shen Xueli , Wu Wenjum ,"The Research Of Data Aggregation In Wireless Sensor Networks", International Forum Of Information And Technology, IEEE Computer Society, 2010.

[24] Alzaid, Hani, Ernest Foo, and Juan Gonzalez Nieto. "Secure data aggregation in wireless sensor network: a survey." In Proceedings of the sixth Australasian conference on Information security-Volume 81, pp. 93-105. Australian Computer Society, Inc., 2008.

[25] Jha, Mukesh Kumar, and T. P. Sharma. "Secure data aggregation in wireless sensor network: a survey." International Journal of Engineering Science and Technology 3, no. 3 (2011).

[26] Dr.G.Padmavathi,Mrs.D.Shanmugapriya, 2009, A Survey of Attacks, Security Mechanisms and Challenges in Wireless Sensor Networks, (IJCSIS) International Journal of Computer Science and Information Security.

[27] Y.E.Aslan and E.Kayaaslan, Security in wireless sensor network, JOURNAL OF CS514 CLASS FILES, VOL.1, NO.1, JANUVARY 2008.

[28] A.Pandey and R.C Tripathi, A Survey on Wireless Sensor Networks Security, International Journal of Computer Applicationsc (0975-8887), Volume 3-No.2, June 2010.

[29] Akyildiz, Ian F., Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci. "A survey on sensor networks." Communications magazine, IEEE 40, no. 8 (2002): 102-114.

Author

Nirbhay K. Chaubey, Ph.D. (Senior Member of IEEE, Senior Member of ACM, Life Member of CSI) working as an Associate Professor, S.S. Agrawal Institute of Computer Science, Gujarat Technological University, Gujarat, India and a Ph. D. supervisor (Computer Science and Engineering), Gujarat Technological University. His research interests lie in the areas of Computer Networking, Wireless Networks (Protocol Design, QoS, Routing, Mobility, and Security), Cloud Computing and Sensor Network, etc. He has published several research papers in peer reviewed International Journals, International, and National Conferences.