

©2015 IEEE. Reprinted, with permission, from M. Ben Haj Frej, and K.M. Elleithy, " Secure Data Aggregation Model (SDAM) in Wireless Sensor Networks." In Proceedings of 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Miami, FL, 2015. DOI: 10.1109/ICMLA.2015.116.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Bridgeport's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Secure Data Aggregation Model (SDAM) in Wireless Sensor Networks

Mohamed Ben Haj Frej
Computer Science and Engineering
University of Bridgeport
Bridgeport, CT 06604
mbenhaj@bridgeport.edu

Khaled M. Elleithy
Computer Science and Engineering
University of Bridgeport
Bridgeport, CT 06604
elleithy@bridgeport.edu

Abstract— Nowadays, Wireless Sensor Networks (WSN's) are becoming more and more promising and applicable to a variety of fields: military, environmental, medical, wild life habitat, and transportation as well wearable devices, target-tracking. WSN are expected to be a main player in the Internet of Things technology. Power management is very important factor in considering WSN's and it has been demonstrated that communication cost is higher than the computational as nodes consume most of the energy in communication. Added to the fact that sensors could be closely deployed and report the same reading, the data aggregation concept was introduced to resolve those issues and for the sake of better performance at a reduced cost. Nonetheless, sensing devices are prone to failure due to several aspects such as node failure or low batteries as well as being compromised. In this paper, we are introducing a novel method Secure Data Aggregation Model (SDAM) aiming at assuring a secure aggregate communication at a low cost (in terms of resources). Our simulation results showed that implementing SDAM resulted into an increase in the energy efficiency as well as a considerable reduction in cross-layering overhead.

Keywords— availability, base station (BS), cluster head (CH), secure Data Aggregation model (SDAM), sensor node (SN), wireless sensor network (WSN).

I. INTRODUCTION

Following to the digital modernization, there has been a tremendous growth in Wireless Sensor Networks (WSN) field; with many applications such as in the military, environmental monitoring, agriculture production and medical field. WSN consist on a considerable number of motes (sensor nodes) placed in a known topology or on ad hoc basis in the desired location in order to collect information [1]. Each mote is a low end device assuring processing, sensing and communication abilities. Below are the motes' main characteristics:

- Limited battery life with no ability to recharge it in certain situations or because it is not worth it as they are meant to be disposable
- Low processing capabilities, as they should be designed based on required resources only, in order to keep them at a low price.
- Limited data storage capacity

Since the sensors are generally deployed close to each other, they could be reporting the same values, there have been a need to summarize the data then forward it to upper nodes for the sake of energy consumption optimization.

Wireless sensor networks' data aggregation consists on reducing the flow of communication and optimizing the energy consumption.

In data aggregation, designated motes will collect and process the data packets from their surrounding nodes then forward them to the base station making a huge impact on the traffic as well as on the communication time. Thus it does its job in increasing the efficiency. However it comes with its inconvenience as it affects other constraints such as security and delay. The aggregated sensors are vulnerable to intruders which can inject false data in WSN.

Secure communication requires that the encryption will take place before transmitting the data. While for aggregation protocols, it would be preferable to process the data and forward it to the next hop prior to encryption. So it is a challenging task to combine both taking in consideration that we could not apply the computing encryption protocols to wireless sensor networks due to their limitation in resources. Although there could be some compromise while adopting aggregation and encryption, both are deemed essential for a stable and accurately performing network [2].

Amongst every operation that might be performed by a sensor node, gathering and transmitting data are the ones that expend considerably higher vitality when contrasted with other operations [3]. Data accumulation serves to lessen correspondence by joining Data parcels. WSNs are as a rule conveyed in threatening and unattended situations, so the need for secure Data conglomeration is further emphasized.

II. RELATED WORK

There are numerous Secure Data Aggregation plans proposed in writing. The base station (BS) takes wise choices in view of the Data that is assembled amid Data collection stage.

Identifying anomaly values in the data frames is a vital errand in choice making for different applications like observing, issue determination and interruption location in WSNs. Exceptions are characterized as "estimations that altogether go a miss from the ordinary example of detected data". The plausible wellsprings of anomalies in WSN incorporate clamor and blunders, occasions, and vindictive assaults [3]. Off base data will antagonistically influence the estimation of the collected data which is utilized by the BS for deciding. For instance, if the sensor nodes are conveyed in backwoods to sense temperature esteem, then an exception in temperature quality can show a flame and the BS will need to call the flame station [4]. Henceforth, anomaly identification is essential for distinguishing crisis in the system. Along these lines, the exception recognition instrument ought to be consolidated with data conglomeration plans [5]. In this paper, they have proposed a secure data conglomeration conspire that recognizes exception values. Their plan distinguishes exception values and in the meantime gives data realness and data uprightness. They have utilized multivariate information investigation method to handle anomaly in associated variables. It permits the BS to identify the false data infusion amid the data conglomeration and to recognize the compromised node.

A broad study of anomaly identification plans has been conducted. The authors have considered a bunch based topology with hubs sorted as: base station (BS), cluster head (CH) and basic sensor nodes (SN). They have utilized multivariate information investigation method to handle anomaly in associated variables. They have utilized the PCA model for flaw discovery [3]. The CH recognizes the arriving information as ordinary or unusual. They have utilized factual methods to discover sensor information to enhance the exception discovery precision. They have utilized a conglomeration tree as a system model. They have planned anomaly recognition plan taking into account commit disseminate system. The authors have presented a model that executes inquiries, distinguishes and sends clients an arrangement of readings that are expected to be exceptions [4]. They have utilized to distinguish a hub as anomaly hub. The authors have proposed a SDA plan which utilizes a bunch based system model. The group head runs the peculiarity recognition calculation and sends the IDs and check of the anomaly values to the guardian group head. Their methodology is taking into account hearty vital segment investigation [4]. To uncover the abnormalities the authors have utilized the relationship among the detected information.

III. CLUSTERING – DATA AGGREGATION

A. *Wireless Sensor networks Limitations*

Since their elaboration, there have been continuous efforts to design efficient and simple WSN's. Hence, there are some constraints to be considered while designing a wireless sensor network with a simple topology [6].

- Energy considerations: nodes can do limited computations as they have less memory and are limited in power supply. Nodes completely depend on their batteries for data transmissions as well as all the other tasks.
- Reliability and Fault Tolerance: the ability to keep transmitting the information even when some of the nodes become unavailable by taking another path.
- Scalability: when the number of nodes increases, it becomes complex to manage and organize the nodes which could affect the overall communication.
- Data Delivery Models: There are different models such as event driven, query driven, and hybrid. These modes control the flow of data from sensors to sink nodes. The type of model used is based upon the applications.
- Quality of Service: there is a set of protocols to measure the energy utilized, time taken for routing the information and to prioritize the data transfer from sensor to sink node.
- Network Dynamics: it is not always the case that all the components in the WSN such as nodes, sink nodes, sensors are static. There is a need to also support mobility.

B. *Clustering process:*

Clustering process steps consists on the following:

- Get ready information: institutionalizing property and characterizing measurement.
- Choosing property: picking viable property from essential properties and putting them away into a vector.
- Collecting property: changing the picked properties to new properties.
- Clustering: picking some sort of separation capacity that has the fitting properties as the estimation of closeness and after that grouping.
- Evaluating on the after effect of grouping: performing assessment of three sorts: outside legitimacy assessment, inside legitimacy assessment and relativity test assessment [7].

C. *Data aggregation process:*

As previously mentioned, wireless sensor networks consist on tiny sensors deployed in large number in the field. Each node is a low end device that integrates data processing, wireless communication and sensing abilities. Setting a base node and data aggregation has to be done towards other nodes in the network hierarchy and leads into saving energy. In case of Multi hops network the data aggregation will consist on getting the min, max and sum and sending it to the upper node [8]. The data aggregation process totally depends on mathematical models. It will decide the following:

- How to locate a bunching inclination in the information.

- How to locate a superior approach to seek the bunch and in addition the gathering hubs.
- How to discover a test system in order to demonstrate that the segments are right and secure [9].

IV. SECURE WIRELESS SENSOR NETWORKS

A wireless sensor network is considered to be secured if it offers all the following services:

- **Authentication and Data Confidentiality:** nodes can carry confidential information of an organization. In order to prevent any hacking attacks, nodes should communicate within secured channels after the authentication. Before giving access, WSN should confirm the identity of all its nodes. The best way to protect the data is to send it in an encrypted form by using encoding and decoding techniques.
- **Data Integrity:** providing exact data sent by the sensors without any alterations.
- **Data Freshness:** Data transferred should be fresh and new. Nodes get busy unnecessarily if same message is sent repeatedly, this could reduce the battery life of the node and reduce the efficiency of the system.
- **Availability:** A node is completely dependent on its battery life. A node should be able to send a message if its battery life is going down and let the other nodes take another path to communicate.
- **Self-Organization:** all the dynamic nodes should organize themselves to form a multi path communication system making the fault tolerance of the wireless network more efficient.
- **Secure Localization:** Location of all the sensors in a wireless sensor network should be traced accurately.

All these features make an ideal wireless sensor network. But in reality, it is a bit difficult to design a WSN which could strictly follow all the protocols. Formally a WSN should authenticate all the nodes, detect and prevent any malicious attacks, and have a data recovery solution in case one or more nodes have been compromised

V. PROPOSED SOLUTION – SECURE DATA AGGREGATION MODEL

In this paper we are proposing a new secure data aggregation consisting of a sub layer between the MAC and the network layers. SDAM's primary objective is to ensure secure communication by eliminating the dilemma of having the data encrypted before aggregation while the aggregation protocols perform better on unencrypted data.

A. SDAM Components:

The SDAM consists of three modules: Operations sub-layer, Data aggregation module and Communications sub-layer (please refer to figure 1).

- **Operations Sub-layer** responsible on checking based on the MAC address of the device to determine if it is coming from a legitimate or illegitimate node. If the node is deemed illegitimate, the packets will be dropped.
- **Data Aggregation Module:** responsible for removing the repetition by applying aggregation algorithms as well as deciding on the size of the packets.
- **The Communications sub-layer** is responsible to decide on the number of packets that are needed to aggregate and then forward them to the MAC layer. Both incoming and outgoing traffic are sent by the MAC layer and then forwarded to the SDAM module.

Below is the diagram of the proposed solution:

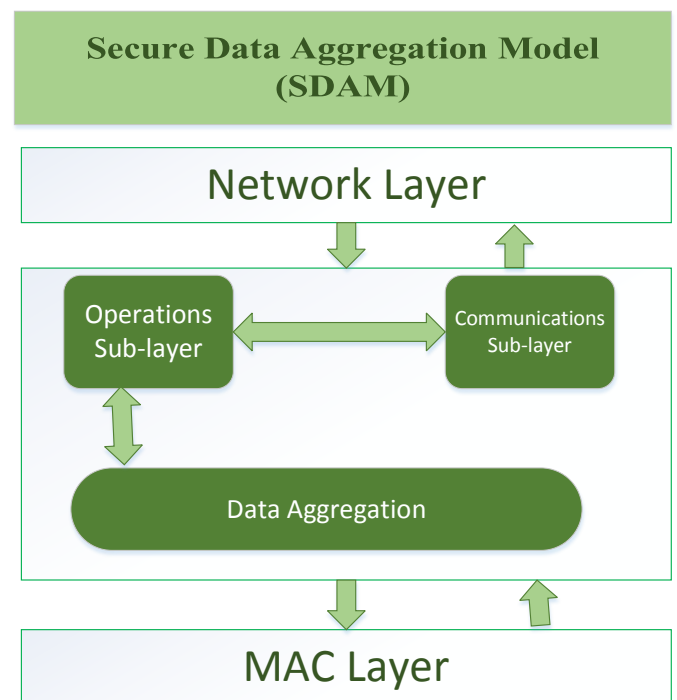


Figure 1: Secure Data Aggregation Model

VI. SIMULATION RESULTS

This section shows the experimental performance of our proposed SDAM. The performance of the proposed model was measured by using Network Simulator-2 (NS2) on Ubuntu 14.2 operating system. The measurements are done on a network size of 700 m x 700 m. We have assumed that there are two types of nodes in the network:

- Homogenous

- Heterogeneous.

The homogenous nodes are distributed in the network that sense and relay the data. The heterogeneous nodes are head nodes in each cluster with more power and bandwidth. Each homogeneous node initially uses 4 joules energy, whereas each heterogeneous node uses 16 joules energy.

We have implemented our proposed model on each cluster head node and assumed that the topology is static. Our goal in this simulation is to apply the secure aggregation. The base station is located at point (0, 850). The size of the packets is 128 bytes. The remaining parameters are given in Table 1.

PARAMETERS	VALUES
Size of network	1600 × 1600 square meters
Number of nodes	180
Number of aggregation generated	1050
Energy of homogenous node	4 joules
Energy of heterogeneous node	16 Joules
Packet size	128 bytes
Data Rate	280 kilobytes/second
Sensing Range of node	35 meters
Simulation time	450 Seconds
Average Simulation Run	05
Base station location	(0,850)
Transmitter Power	13.2 mW
Receiver Power	11.4 mW

Table 1: Simulation Parameters

Based on the initial simulation results, we will be focusing on evaluating the following metrics:

- *Energy Efficiency*
- *Cross-Layering Overhead*

A. Energy Consumption

Our generated scenario consists of 180 nodes that are randomly distributed. 10% of the nodes are malicious constitute a hurdle on the way of aggregation. Our proposed SDAM is implemented on each cluster head node. Once, the nodes collect the data from the sensing event that are forwarded to the cluster head node, which is responsible to apply aggregation and balance the communication between MAC sub-layer and network layer.

Based on the results on Figure 2. Based on the results, we observed that 1.34 joules energy is consumed when using proposed model whereas, during the simulation period, while 1.53 joules energy is consumed without using SDAM.

Applying SDAM made the network more efficient in terms of energy consumption and more secure as it detects the malicious nodes and avoids data loss. Without SDAM, the packets are captured by the malicious nodes because the sensor nodes do not have the capability to identify the possible malicious activities so that the captured packets are lost and retransmitted.

As a result, the additional energy is consumed for retransmission the packets. We have proved that our proposed model saved 0.19 joules energy that is approximately 9.8% saving in the entire network and that our proposed model could be used to extend the network lifetime.

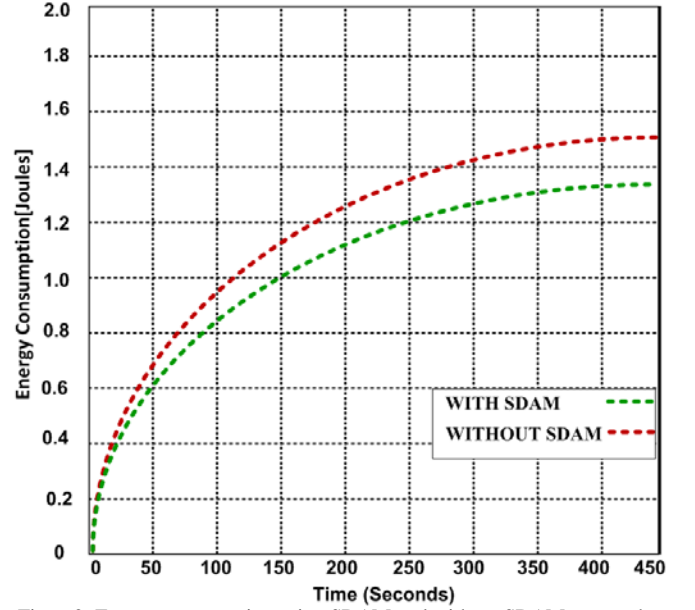


Figure 2: Energy consumption using SDAM and without SDAM approaches during the entire simulation time

B. Cross-Layering Overhead

Using the same simulation setup we have measured the overhead cross-layering for 1800 aggregations (Figure 3-a) and for 3600 aggregations (Figure 3-b).

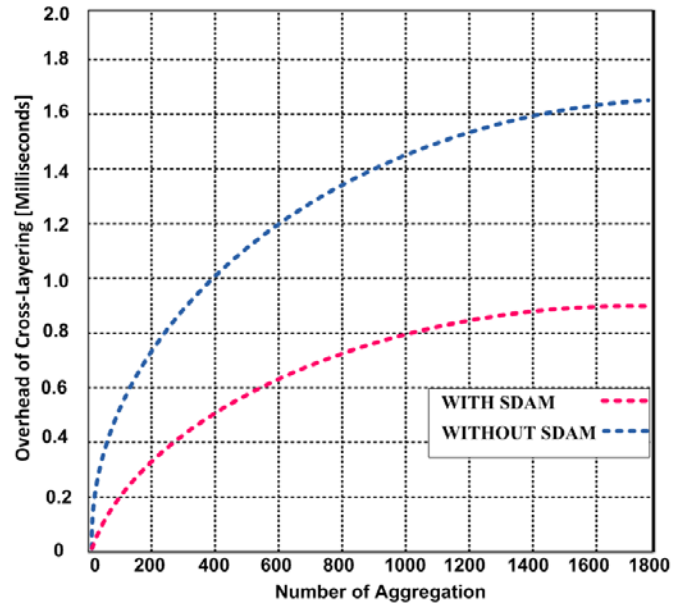


Figure 3a: Cross-layering Overhead With and Without SDAM

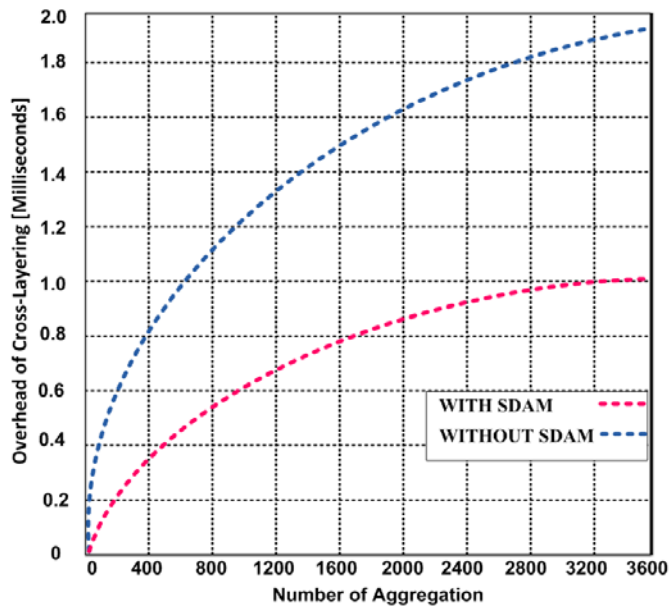


Figure 3b: Cross-layering Overhead With and Without SDAM- Doubling the number of Aggregations

Our simulation results shows the following: the overhead with SDAM is: 0.7 milliseconds for 1800 aggregations and 0.8 for 3600 aggregations Without the SDAM, it is 1.64 milliseconds for 1800 aggregations and 1.92 milliseconds for 3600 aggregations.

The simulation results shows 43 % improvement for 1600 aggregations and around the same improvement (42 %) with 3600 aggregations.

These results indicate that our proposed SDAM solution shows considerable improvement in the cross-layering overhead along with its energy efficiency optimization.

VII. CONCLUSION

Considering the limited resources of the wireless sensor networks and the impact of the use of encryption on aggregated data; In this paper we have proposed a simplistic and efficient method. The approach consisted On adding modules able to authenticate non-legitimate nodes in the network as well as facilitate the communications, in an energy efficient way as shown by our simulation results. Our simulation results also showed an around 40% improvement in the cross-layering overhead.

ACKNOWLEDGMENT

Special thanks to Dr. Abdul Razaque, from the Cleveland State University, who has been with a help especially in the simulation section.

REFERENCES

- [1] Razaque, Abdul, and Khaled Elleithy. "Modular Energy-Efficient and Robust Paradigms for a Disaster-Recovery Process over Wireless Sensor Networks." *Sensors* 15, no. 7 (2015): 16162-16195.
- [2] Karthikeyan, B., Velumani, M., Kumar, R., and Inabathini, S.R.: 'Analysis of data aggregation in wireless sensor network', in Editor (Ed.)^(Eds.): 'Book Analysis of data aggregation in wireless sensor network' (IEEE, 2015, edn.), pp. 1435-1439.
- [3] Rezvani, M., Ignjatovic, A., Bertino, E., and Jha, S.: 'Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks', *Dependable and Secure Computing*, IEEE Transactions on, 2015, 12, (1), pp. 98-110.
- [4] Bharuka, K., and Jinwala, D.C.: 'A secure data aggregation protocol for outlier detection in wireless sensor networks using aggregate Message Authentication Code', in Editor (Ed.)^(Eds.): 'Book A secure data aggregation protocol for outlier detection in wireless sensor networks using aggregate Message Authentication Code' (2014, edn.), pp. 1-6.
- [5] Sahana, S., and Amutha, R.: 'Data aggregation in wireless sensor networks', in Editor (Ed.)^(Eds.): 'Book Data aggregation in wireless sensor networks' (2014, edn.), pp. 1-6.
- [6] Jose, J., Manoj Kumar, S., and Jose, J.: 'Energy efficient recoverable concealed data aggregation in wireless sensor networks', in Editor (Ed.)^(Eds.): 'Book Energy efficient recoverable concealed data aggregation in wireless sensor networks' (2013, edn.), pp. 322-329.
- [7] Durrani, N.M., Kafi, N., Shamsi, J., Haider, W., and Abbasi, A.M.: 'Secure multi-hop routing protocols in Wireless Sensor Networks: Requirements, challenges and solutions', in Editor (Ed.)^(Eds.): 'Book Secure multi-hop routing protocols in Wireless Sensor Networks: Requirements, challenges and solutions' (IEEE, 2013, edn.), pp. 41-48.
- [8] Mingxin, Y., Jingsha, H., and Xuguang, S.: 'Research on Secure Data Aggregation in Wireless Sensor Networks Based on Clustering Method', in Editor (Ed.)^(Eds.): 'Book Research on Secure Data Aggregation in Wireless Sensor Networks Based on Clustering Method' (2011, edn.), pp. 1-4.
- [9] Ben Othman, S., Trad, A., Youssef, H., and Alzaid, H.: 'Secure data aggregation in wireless sensor networks', in Editor (Ed.)^(Eds.): 'Book Secure data aggregation in wireless sensor networks' (IEEE, 2013, edn.), pp. 55-58.
- [10] CAO, X.-m., LI, W.-l., and YANG, G.: 'Research on Secure Data Aggregation in Wireless Sensor Network', *Computer Technology and Development*, 2012, 11, pp. 060.