CLUSTER BASED DATA AGGREGATION AND AUTHENTICATION PROTOCOL FOR WSN

Ms.D.Vijayakumari*

Ms.Dhivya Kesavan*

***Abstract*—**

In wireless sensor network the information collected by sensor nodes should be aggregated and communicated without any false data. However Compromised sensor nodes are capable of injecting false data during both data aggregation and data forwarding. An Enhanced cluster based data aggregation and authentication protocol called DAA is proposed, to integrate false detection with data aggregation and confidentiality. In this protocol to perform data aggregation along with false data detection, the monitoring nodes of every data aggregator also conduct data aggregation and compute the corresponding small size message authentication codes for data verification at their pair mates. Cluster based approach is employed to select data aggregator. The data integrity on encrypted data rather than the plain data is verified to support confidential data transmission. Performance analysis shows that DAA detects any false data injected by up to T compromised nodes, and that the detected false data are not forwarded beyond the next data aggregator on the path. It is also considered that enabling every sensor node to be capable of both aggregating and forwarding data in order to improve network security and efficiency.

***Keywords*— Data aggregation, data integrity, network-level security, sensor networks, false detection.**

* Assistant Professor, RMK College of Engineering and Technology, RSM Nagar,Puduvoya-601 206.

I. INTRODUCTION

Wireless sensor networks (WSNs) open up new application areas such as tactical surveillance, intelligent environmental and structural monitoring and target tracking. In a WSN, large numbers of tiny nodes may be deployed in an ad hoc manner. These nodes automatically configure a topology by communicating and coordinating with each other. Nodes assume the roles of both sensing device and router. Messages are relayed to other nodes or to a hub in a multi hop fashion. Multi-hop routing in an energy-constrained WSN has been shown to give rise to significant gains in network performance. With more nodes, the area being monitored can be increased or with the same area, the increase in node density gives more precise and timely data and also provides a degree of operational reliability. In wireless sensor networks it is important to save energy so that the batteries of the sensor nodes last for a long time. This means that computations and communications should be kept at a minimum so that the nodes can sleep as often as possible. On the other hand there is a demand for security which increases the number of clock cycles used for computations and the number of bits sent over communication channels.

II. RELATED WORK

False data injections, data authentication schemes that employ multiple MACs are proposed based on the observations of data integrity. The statistical en-route detection scheme, called SEF, enables relaying nodes and base station to detect false data with a certain probability. In 10 hops, SEF is able to drop 80%–90% of the injected false reports. In the interleaved hop-by-hop authentication scheme, any packet containing false data injected by compromised sensor nodes is detected by those sensor nodes that collaborate to verify data integrity. In the interleaved hop-by-hop authentication scheme, sensor nodes are not allowed to perform data aggregation during data forwarding. The Commutative Cipher based En-route Filtering scheme (CCEF) drops false data en-route without symmetric key sharing[5][6]. Secure data aggregation problem is studied extensively. In [1] the security mechanism detects node misbehaviors such as dropping or forging messages and transmitting false data. Random sampling mechanisms and interactive proofs are used to check the correctness of the aggregated data at base station[2][3]. Several key establishment protocols are developed for sensor networks, which

offer "direct key establishment" for neighboring nodes and "path key establishment"[4] for sensor nodes that are multiple hops away from each other. In path key establishment method, to establish a pairwise key with node, a sensor node needs to find a path between itself and node such that any two adjacent nodes in that path can establish a pair wise key directly.

III. Problem Definition

Power-controlled networks have nodes with variable RF power transceivers that provide greater routing performance at the expense of higher power consumption and costs Fixed-power networks have cheaper motes with fixed-power RF transceivers but may be more prone to communication disruptions. Several routing protocols in fixed-power, multi hop WSNs use shortest-path routing. Since operation is often over long unattended periods, the protocol must be energy efficient. As such, routing protocols must ensure that the WSN can reconfigure, be energy efficient and resilient to failures. These non-trivial requirements pose conflicting demands on protocol design.

A. False Detection and Authentication

Data confidentiality prefers data to be encrypted at the source node and decrypted at the destination. However, data aggregation techniques usually require any encrypted sensor data to be decrypted at data aggregators for aggregation. The existing false data detection algorithms address neither data aggregation nor confidentiality. Although they could be modified easily to support data confidentiality, it is a challenge for them to support the data aggregation that alters data. For instance, the basic idea behind the false data detection algorithm in is to form pairs of sensor nodes such that one pairmate computes a message authentication code (MAC) of forwarded data and the other pairmate later verifies the data using the MAC, as illustrated in Fig. 1.

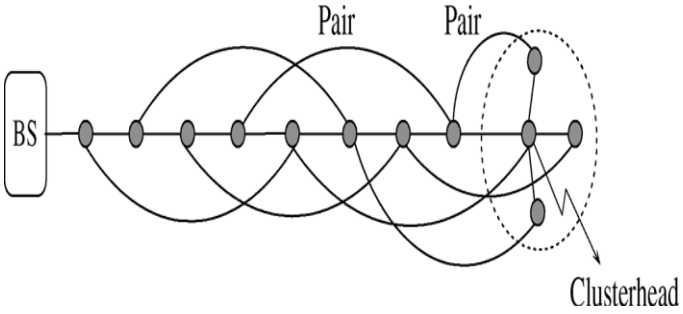


Fig. 1. An example of forming sensor pairs to authenticate data for the false data detection scheme in [3], where data aggregation is not allowed if it requires any change in the data. In this scheme, any data change between two pairmates is considered as false data injection, and therefore, data aggregation is not allowed if it requires alterations in the data. Hence, the false data detection algorithm cannot be implemented when a data aggregator between two pair mates changes the data.

IV. Data Aggregation and Authentication Protocol

This section presents the protocol DAA and its algorithms, namely MNS and SDFC, provides secure data aggregation, data confidentiality, and false data detection by performing data aggregation at data aggregators and their neighbouring nodes and verifying the aggregated data during data forwarding between two consecutive data aggregators. DAA has three steps that are explained in the following subsections.

Input: A Wireless sensor network with densely deployed sensor nodes, some of which are designated as data aggregators. For given valve of ***T,*** data aggregators are already selected in such a way that there exists at least T nodes between any two data aggregators.

Output: Even though the network can have up to T compromised nodes, data are aggregated in data aggregators.

Step-1: T neighbouring nodes of each data aggregator are randomly selected as monitoring nodes to perform the additional data aggregation and to compute sub MAC s of the aggregated.

Step-2: The following 2T+1 pair of nodes are formed by enabling the nodes of every pair to share a distinct symmetric key: (1) one pair is formed by the current and forward data aggregators, (2) T pairs are formed by the monitoring and forwarding nodes of the current data aggregator.

Srep-3: Each data aggregator and its selected T monitoring nodes aggregate data and then compute sub MACs. The aggregated data are encrypted by the current data aggregator. The data aggregator and its monitoring nodes compute two sub MACs: one sub MAC for the encrypted aggregated data and another sub MAC for the plain aggregated data. The current data aggregator constructs two FMACs to forwarding nodes. The integrity of the encrypted data is verified by forwarding pair mates of the selected monitoring nodes of the current data aggregator. The integrity of the plain data is verified by some neighbouring nodes of the forward data aggregator. If the integrity verification of the encrypted or plain data fails at any sensor node, the data are dropped immediately.

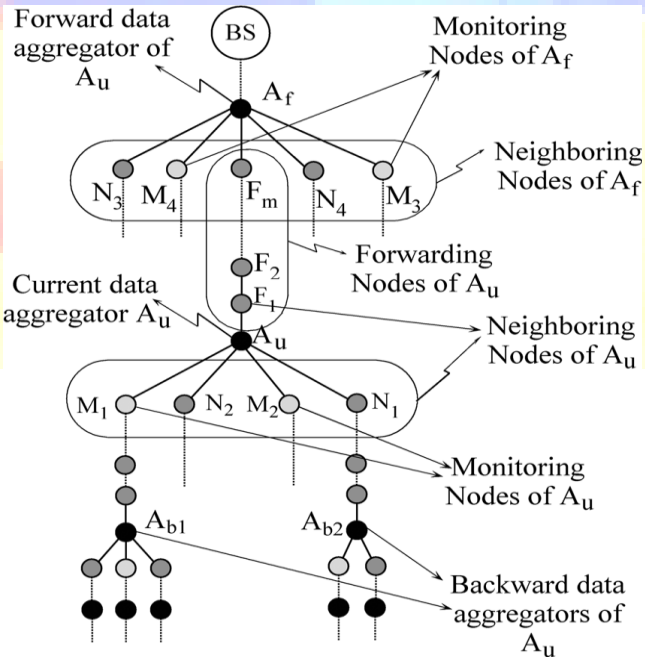


Fig. 2. The system architecture of sensor nodes used by DAA.

To support false data detection, secure data aggregation, and confidentiality against up to T compromised sensor nodes, DAA forms 2T+1 pairs of sensor nodes by the neighbouring and forwarding nodes of Au and Af.

A. Data aggregation

Data aggregation results in better bandwidth and battery utilization, which enhances the network lifetime because communication constitutes 70% of the total energy consumption of the network. This paper introduces a data aggregation and authentication protocol (DAA) to provide false data detection and secure data aggregation against up to **T** compromised sensor nodes, for **T>1**. The value of **T** depends on security requirements, node density, packet size, and the amount of tolerable overhead. We assume that some sensor nodes are selected dynamically as data aggregators, and the nodes between two consecutive data aggregators are called forwarding nodes simply because they forward data. To detect false data injected by a data aggregator while performing data aggregation, some neighbouring nodes of the data aggregator also perform data aggregation and compute MACs for the aggregated data to enable their pair mates to verify the data later. DAA also provides data confidentiality as data are forwarded between data aggregators. To provide data confidentiality during data forwarding between every two consecutive data aggregators, the aggregated data are encrypted at data aggregators, and false data detection is performed over the encrypted data rather than the plain data. Whenever the verification of encrypted data fails at a forwarding node, the data are dropped immediately to minimize the waste of resources such as bandwidth and battery power due to false data injection.

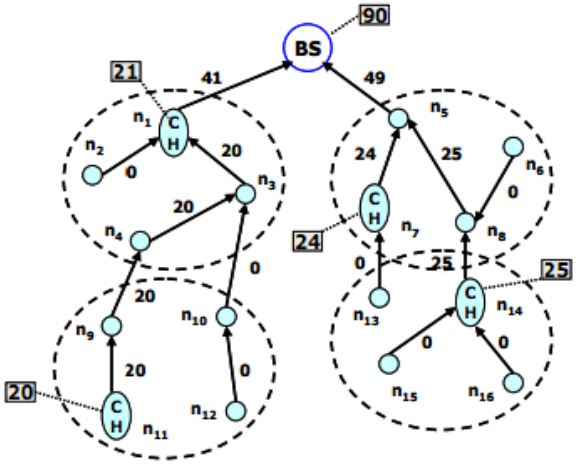


Fig.3 Aggregation of four cluster aggregates

B. Cluster based Approach for Data Aggregators

In cluster-based approach, whole network is divided in to several clusters. Each cluster has a cluster-head which is selected among cluster members. Cluster-heads do the role of aggregator which aggregate data received from cluster members locally and then transmit the result to sink. The advantages and disadvantages of the cluster-based approaches are very much similar to tree-based approaches form a direct link with cluster-head. Phase I of this scheme is similar to various scheme used for clustering but differ in one way that the cluster-head rotation is localized and is done based on the remaining energy level of the sensor nodes which minimize time variance of sensors and this lead to energy saving from unnecessary cluster-head rotation. In phase II, each node within the cluster searches for a neighbour closer than cluster-head which is called data relay point and setup up a data relay link. Now the sensor nodes within a cluster either use direct link or data relay link to send their data to cluster head which is an energy efficient scheme. The data relay point aggregates data at forwarding time to another data relay point or cluster-head.

V. MONITORING NODE SELECTION

Aggregator Au requests it's every neighbouring node to send two random numbers along with its node ID number. Each neighbouring node of Au generates two random numbers(Ra and Rb) using its key that it shares with Au. Ra, Rb and MAC(Ra|Rb) are sent to Au. When Au finishes receiving random numbers and node IDs from nits neighbouring nodes, it labels them Ni in the receiving order of their random numbers. Then Au sorts all random numbers in ascending order and computes MAC using Kgroup. All such information are broadcasted by Au. Each Ni verifies the broadcast numbers by checking whether two random numbers that it sent earlier to Au match two of the random numbers that Au has broadcasted. If the verification is successful Ni encrypts the MAC Kgroup using the key it shares with Au and sends it to Au. Else Ni informs its neighbouring nodes and Au about it, along with a request of re-starting the monitoring node selection. To determine the indices of the T monitoring nodes, each Ni runs the following modulus function Ik. Any Ni whose index I happens to be equal to an Ik is selected as a monitoring node. If there is a duplicate Ik value, modulus function is run by increasing the K value by 1.

$$I_k = \left[\left(\sum_{j=k}^{n-1+k} R_j + K^n_{group} \right) mod(n) \right] + 1$$

A. SECURE DATA AGGREGATION AND FALSE DETECTION-INTEGRATION.

To provide data confidentiality, transmitted data are always encrypted and forwarding nodes perform the data verification over the encrypted data. Prior to this step of DAA, monitoring nodes of every data aggregator are selected, and **2T+1** pairs are formed. To verify data integrity and detect false data injections, one pair mate computes a sub MAC, and the other pair mate verifies the sub MAC. Sub Macs are computed for both plain and encrypted data. Sub Macs of plain data are used to detect false data injections during data aggregation, whereas subMACs of encrypted data are used to detect false data injections during data forwarding. To detect any false data that the current data aggregator can inject during data aggregation, the monitoring nodes of also aggregate the incoming data of and compute subMACs for the plain aggregated data, so that the forward data aggregator and its neighboring nodes verify the subMACs. Similarly, to detect those false data that can be injected during data forwarding, the monitoring nodes of compute subMACs for the encrypted aggregated data and then their pairmates of forwarding nodes verify

the subMACs. Main steps of SDFC are: 1) whenever some data are received by a data aggregator, the authenticity of data is verified by the data aggregator and its neighboring nodes; 2) the data aggregator and its monitoring nodes aggregate the data independently of each other; 3) each monitoring node computes one subMAC for the encrypted data and the other subMAC for the plain data; 4) the data aggregator collects these subMACs from its monitoring nodes to form the FMACs of the encrypted and plain data, appends the FMACs to the encrypted data, and transmits them; 5) the forwarding nodes verify the data integrity of the encrypted data; and finally 6) the neighboring nodes of the next aggregator verify the integrity of the plain data. In Algorithm SDFC, each data aggregator forms two FMACs: one FMAC for the encrypted data, and the other FMAC for the plain data. Each FMAC consists of **T+1** subMACs computed by the data aggregator and its **T** monitoring nodes. The FMACs of encrypted and plain data are forwarded along with the encrypted data. In the formation of FMACs, data aggregator determines the order of subMACs in anyway and informs each forwarding node about its subMAC location individually.

VI. Performance Evaluation

Integration of False detection and data aggregation providing confidentiality results in improvising network security. The quality of the data obtained is better vivid through the performance analysis iv various parameters such as packet delivery ratio, End-End latency, Energy Consumption, Throughput.

Packet delivery ratio (PDR): It measures the percentage of data packets generated by nodes that are successfully delivered, expressed as

$$\frac{\text{TOTAL NUMBER OF PACKETS SUCCESFULLY DELIVERED}}{\text{TOTAL NUMBER OF PACKETS SENT}} X100\%$$

End-End latency: It measures the average time it takes to route a data packet from the source node to the hub it is expressed as

$$\frac{\sum \text{INDIVIDUAL DATA PACKET LATENCY}}{\text{TOTAL NUMBER OF PACKETS DELIVERED}}$$

Energy consumption: It measures the energy expended per delivered data packet. It is expressed as

$$\frac{\sum ENERGY\ EXPENDED\ BY\ EACH\ NODE}{TOTAL\ NUMBER\ OF\ PACKETS\ DELIVERED}$$

Throughput: It is defined as the number of packed at destination side at a particular time.

$$\frac{NUMBER\ OF\ PACKET\ RECEIVED}{TIME\ (Sec)}$$

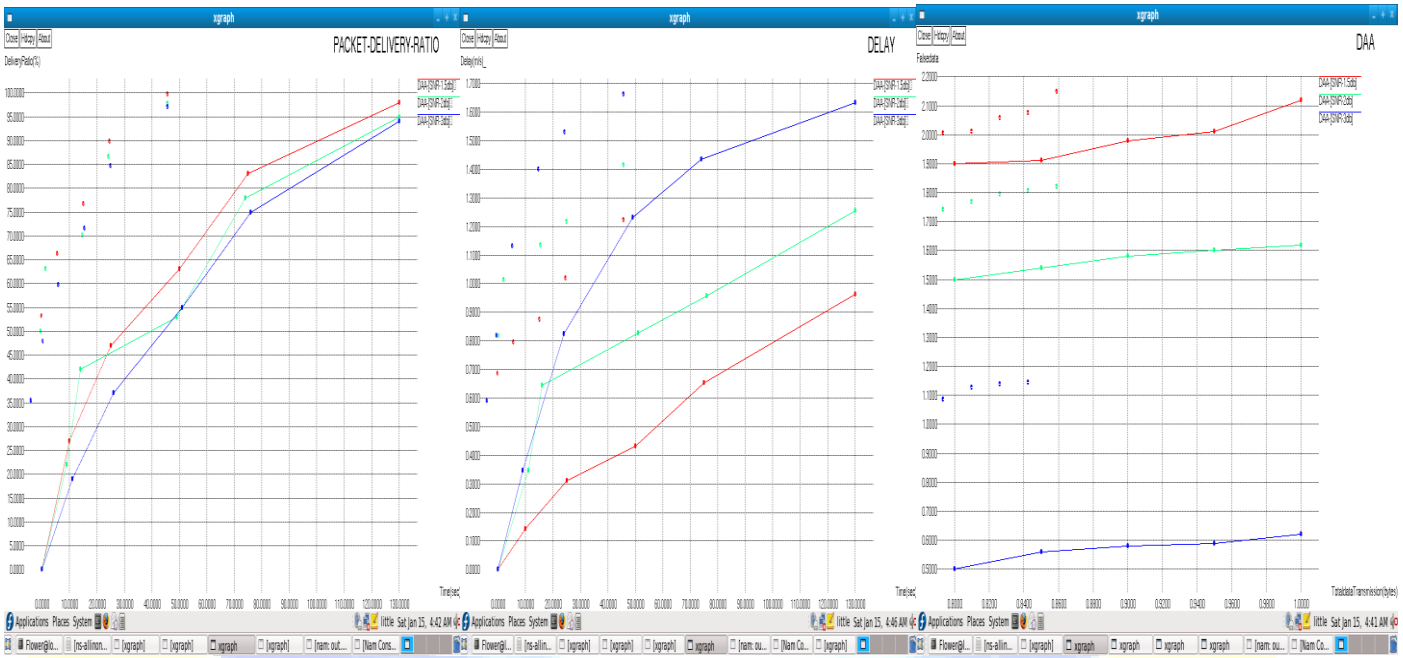
VII. CONCLUSION

Data integrity distorted by compromised sensor nodes is prevented through the security protocol which detects false data during both data forwarding and in data aggregation, improving the network security and quality of received data. Significant Bandwidth utilization, Energy consumption, and improved data accuracy is achieved. Cluster based approach employed in selection of data aggregators adds with reducing end to end delivery. As for the future research, we consider of enabling every sensor node to be capable of both aggregating and forwarding data simultaneously in order to improve network security and efficiency.

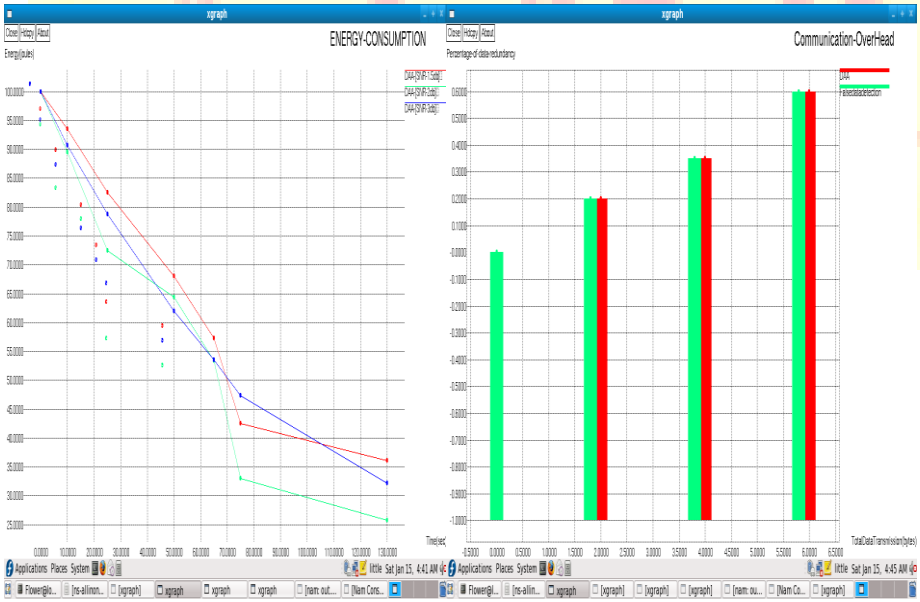
Fig.4 Performance Analysis

(a) Packet delivery ratio, (b) Delay, (c) DAA compared with different S/N ratio, (d) Energy Consumption, (e) Communication overhead-Data aggregation Vs transmitted bytes

(a) (b) (c)



(d) (e)



REFERENCES

[1] K. Wu, D. Dreef, B. Sun, and Y. Xiao, "Secure data aggregation without persistent cryptographic operations in wireless sensor networks," *Ad Hoc Netw.*, vol. 5, 2007, pp. 100–111.

[2] R. Rajagopalan and P. K. Varshney, "Data aggregation techniques in sensor networks: A survey," *IEEE Commun. Surveys Tutorials*, vol. 8, no. 4, 4th Quarter 2006.

[3] H.Çam, S.Ozdemir, P.D.Muthuavinashiappan, and H. O. Sanli, "Energy-efficient and secure pattern based data aggregation for wireless sensor networks," *Comput. Commun.*, vol. 29, no. 4, pp. 446–455, Feb. 2006.

[4] D. Liu, P. Ning, and R. Li, "Establishing pairwise keys in distributed sensor networks," *ACM Trans. Inf. Syst. Security*, vol. 8, no. 1, pp. 41–77, Feb. 2005.

[5] S. Xu, "On the security of group communication schemes based on Symmetric key cryptosystems," in *Proc. ACM Workshop Security Ad hoc Sensor Netw.*, 2005, pp. 22–31.

[6] P. Gauravaram, W. Millan, J. G. Nieto, and E. Dawson, "3C—A provably Secure pseudorandom function and message authentication code: A new mode of operation for cryptographic hash functions," Cryptology ePrint archive, Rep., 2005.

[7] C. Karlof, N. Sastry, and D. Wagner, "TinySec: A link layer security architecture for wireless sensor networks," in *Proc. 2nd ACM Conf. Embedded Netw. Sensor Syst.*, 2004, pp. 162–175,

[8] D. Seetharam and S. Rhee, "An efficient pseudo random number generator for low-power sensor networks," in *Proc. 29th Annu. IEEE Int.Conf. Local Comput. Netw.*, 2004, pp. 560–562.

[9] F. Ye, H. Luo, S. Lu, and L. Zhang, "Statistical en-route detection and filtering of injected false data in sensor networks," in *Proc. IEEE INFOCOM*, 2004, vol. 4, pp. 2446–2457.

[10] H. Yang and S. Lu, "Commutative cipher based en- route filtering In wireless sensor networks," in *Proc. IEEE VTC*, 2004, vol. 2, pp. 1223–1227.

[11] W. Du, J. Deng, Y. S. Han, and P. K.Varshney, "A witness-based approach for data fusion assurance in wireless sensor networks," in *Proc. IEEE GLOBECOM*, 2003, pp. 1435–1439.

[12] W. Du, J. Deng, Y. S. Han, and P. K. Varshney, "A pairwise key predistribution scheme for wireless sensor networks," in *Proc. 10th ACM CCS*, 2003, pp. 42–51.

[13] L. Hu and D.Evans, "Secure aggregation for wireless networks," in *Proc. Workshop Security Assurance Ad hoc Netw.*, Orlando, FL, Jan. 28, 2003, pp. 384–394.