ISSN 2047-3338

Analyzing the Impact of Standard Encryption Approaches for Data Aggregation in a Wireless Sensor Network

Shivani Desai, Sejal Butani and Sharada Valiveti

***Abstract*—Wireless Sensor Network (WSN) is a variant of Mobile Adhoc Network (MANET) that exhibits strict energy constraints. Recent advances in WSN have lead to promising applications viz. habitat monitoring and target tracking. These applications are based on monitoring and analysis of data received by various sensor nodes. However, data communication between nodes consumes a major portion of the total energy consumption of the WSN. Consequently, data aggregation techniques can efficiently reduce the energy consumption by eliminating redundant data traveling back to the base station. This paper aims to identify an efficient data aggregation protocol and after survey of different data aggregation approaches, LEACH has been selected as a data aggregation protocol which is a cluster based approach. The security issues such as data integrity, confidentiality, and freshness in this data aggregation become crucial when the WSN is deployed in a remote or hostile environment where sensors are prone to node failures and compromises. After comparisons of three methods of encryption like AES, DES and RSA, AES algorithm has been selected which follows the symmetric approach. Here, comparisons are based on the parameters like Execution time, Throughput and Energy consumed.**

***Index Terms*— Wireless Sensor Network, Data Aggregation, Security, Energy Efficiency, Routing and Cluster**

I. INTRODUCTION

WIRELESS Sensor Network consists of small and light weighted large number of wireless nodes called sensor nodes with the ability to communicate among themselves and also to an external sink or a base-station to form a communication network such as a single multi-hop network or a hierarchical organization with several clusters and cluster heads. These sensors are deployed in physical or environmental condition to measure physical parameters such as sound, pressure, temperature, humidity, etc.

Wireless sensor networks (WSNs) have been used for numerous applications including habitat monitoring, building monitoring, health monitoring, military surveillance, target tracking, etc. Generally, WSN have broad applications in either controlled environments (such as home, office, warehouse, etc) or uncontrolled environments (such as hostile or disaster areas, toxic regions, etc). Some important applications are: Area monitoring i.e., gathering information from a region where it is located. Environmental monitoring i.e., measurement of temperature, rainfall etc.

A. Motivation

However, WSNs are resource constrained with limited energy lifetime, slow computation, small memory, and limited communication capabilities. However, data communication between nodes consumes a large portion of the total energy consumption of the WSNs. A sensor node generates data based on its sensing mechanisms and then transmits that sensed data packet to the base station (sink). This process is a direct transmission and it may be possible that the base station may be located very far away from sensor nodes and more energy is required to transmit data over long distances. Hence a better technique is to have fewer nodes send the data to the base station. These nodes are called the aggregator nodes and this process is called data aggregation in wireless sensor network. In this way, data aggregation techniques can greatly help to reduce the energy consumption by eliminating redundant data traveling back to the base station.

Not only the resources limitations affect the WSN performance but the deployment nature does also. Most of the WSNs are deployed in remote or hostile environments and then nodes cannot be protected from physical attacks since anyone can access the deployment area. However, the aggregators are vulnerable to attack especially they are not equipped with tamper-resistant hardware. When an aggregator node is compromised, it is easy for the adversary to change the aggregation result and inject false data into the WSNs. Due to these limitations, security issues such as data integrity, confidentiality, and freshness in data aggregation become crucial in WSN.

The rest of the paper is organized as follows: In section II, data aggregation is explained. Section III, includes Literature survey of various hierarchical routing protocols. Section IV shows the simulation results of LEACH routing protocol. In section V, Security issues and comparisons of different security algorithms are classified. Finally, section VI concludes the paper.

II. DATA AGGREGATION IN WSN

Typically, there are three types of nodes in WSN: normal sensor nodes, aggregators, and a querier. The aggregators collect data from a subset of the network, aggregate the data using a suitable aggregation function like avg, min, max, sum, etc and then transmit the aggregated result to an upper aggregator or to the querier who generates the query. The querier can be the base station or sometimes an external user

who has permission to interact with the network depending of the network architecture.

A. Advantages and Disadvantages of Data Aggregation in WSN

Advantages:

- It conserves energy of the sensors by eliminating redundant data.
- It reduces the traffic load.
- With the help of data aggregation process we can enhance the robustness and accuracy of information which is obtained by entire network.

Disadvantages:

- The data aggregator nodes send aggregated data to the base station. These aggregator nodes may be attacked by malicious attacker. If any aggregator node is compromised, then the base station (sink) cannot ensure the correctness of the aggregate data that has been sent to it.

B. Performance measure of Data Aggregation

There are very important performance measures of data fusion algorithm. These performances are highly dependent on the desired application [2].

Energy Efficiency: Energy efficiency refers to a reduction in the energy used for a given service (heating, lighting, etc.) or level of activity. By the data-aggregation scheme, we can increase the functionality of the wireless sensor network. In which every sensor nodes should have spent the same amount of energy in every data gathering round. A data aggregation scheme is energy efficient if it maximizes the functionality of the network.

Network lifetime: The network lifetime is defining the number of data fusion rounds till the specified percentage of the total nodes dies and the percentage depend on the application. For example, lifetime is defined as the number of rounds until the first sensor is drained of its energy.

Latency: It can be measured as the time delay between the data packets received at the sink and the data generated at the source nodes.

Communication overhead: It evaluates the communication complexity of the network fusion algorithm.

Throughput: It defines the average data rate of successful message delivery over a communication channel. The throughput is usually measured in bits per second, and sometimes in data packets per second or data packets per time slot.

III. LETERATURE SURVEY OF HIERARCHICAL ROUTING PROTOCOLS

To do data aggregation there is a need of some data aggregation protocol. Following are the various hierarchical data aggregation approaches:

A. LEACH

LEACH [3] stands for Low-Energy Adaptive Clustering Hierarchy and LEACH were one of the first hierarchical protocols. The need of network protocol such as LEACH is due to the fact that a node in the network is no longer useful when its battery dies. LEACH arranges the nodes in the network into small clusters and chooses one of them as the cluster-head. Node first senses its target and then sends the relevant information to its cluster-head. Then the cluster head aggregates and compresses the information received from all the nodes and sends it to the base station. The nodes chosen as the cluster head drain out more energy as compared to the other nodes as it is required to send data to the base station which may be far located. Hence LEACH uses random rotation of the nodes required to be the cluster-heads to evenly distribute energy consumption in the network. It was found that only 5% of the total number of nodes needs to act as the cluster-heads. TDMA/CDMA MAC is used to reduce inter-cluster and intra-cluster collisions. LEACH is suited for applications which involve constant monitoring and periodic data reporting.

LEACH operations can be divided into two phases: 1). Setup phase; 2). Steady phase

In the Setup phase, the clusters are formed and a cluster-head (CH) is chosen for each cluster. While in the Steady phase, data is sensed and sent to the central base station. The steady phase is longer than the setup phase. This is done in order to minimize the overhead cost.

1). Setup phase: During the setup phase, a predetermined fraction of nodes, p, choose themselves as cluster-heads. This is done according to a threshold value, T(n). The threshold value depends upon the desired percentage to become a cluster-head- p, the current round r, and the set of nodes that have not become the cluster-head in the last 1/p rounds, which is denoted by G. The formula is as follows:

$$T(n) = \frac{p}{1-p*\left(r*mod\left(\frac{1}{p}\right)\right)} \quad \forall n \in G \qquad [3]$$

Every node wanting to be the cluster-head and chooses a value, between 0 and 1. If this random number is less than the threshold value, T(n), then the node becomes the cluster- head for the current round. Then each elected CH broadcasts an advertisement message to the rest of the nodes in the network to invite them to join their clusters. Based upon the strength of the advertisement signal, the non-cluster head nodes decide to join the clusters. The non-cluster head nodes then informs their respective cluster-heads that they will be under their cluster by sending an acknowledgement message. After receiving the acknowledgement message, depending upon the number of nodes under their cluster and the type of information required by the system (in which the WSN is setup), the cluster-heads creates a TDMA schedule and assigns each node a time slot in which it can transmit the sensed data. The TDMA schedule is broadcasted to all the cluster-members. If the size of any cluster becomes too large, the cluster-head may choose another cluster- head for its cluster. The cluster-head chosen for the current round cannot

again become the cluster-head until all the other nodes in the network haven't become the cluster-head.

2. Steady phase: During the steady phase, the sensor nodes i.e. the non-cluster head nodes starts sensing data and sends it to their cluster-head according to the TDMA schedule. The cluster-head node, after receiving data from all the member nodes, aggregates it and then sends it to the base-station.

After a certain time, which is determined a priori, the network again goes back into the setup phase and new cluster-heads are chosen. Each cluster communicates using different CDMA codes in order to reduce interference from nodes belonging to other clusters.

B. HEED

HEED [3], [4] whose main goal is to form efficient clusters for maximizing network lifetime. The main assumption in HEED is the availability of multiple power levels at sensor nodes. Cluster head selection is based on a combination of node residual energy

of each node and a secondary parameter which depends on the node proximity to its neighbors or node degree. The cost of a cluster head is defined as its average minimum reach ability power (AMRP). AMRP is the average of the minimum power levels required by all nodes within the cluster range to reach the cluster head. AMRP provides an estimate of the communication cost.

C. EECS

An Energy Efficient Clustering Scheme (EECS) [3], [4] is a clustering algorithm in which cluster head candidates compete for the ability to elevate to cluster head for a given round. This competition involves candidates broadcasting their residual energy to neighboring candidates. If a given node does not find a node with more residual energy, it becomes a cluster head. Cluster formation is different than that of LEACH. LEACH forms clusters based on the minimum distance of nodes to their corresponding cluster head. EECS extends this algorithm by dynamic sizing of clusters based on cluster distance from the base station. The result is an algorithm that addresses the problem that clusters at a greater range from the base station requires more energy for transmission than those that are closer. Ultimately, this improves the distribution of energy throughout the network, resulting in better resource usage and extended network life time.

D. PEGASIS

Power Efficient data Gathering protocol for Sensor Information Systems (PEGASIS) [3], [4] is a chain based data aggregation protocol. In PEGASIS, nodes are organized into a linear chain for data aggregation. The nodes can form a chain by employing a greedy algorithm. Greedy chain formation assumes that all nodes have global knowledge of the network. The farthest node from the sink initiates chain formation and at each step, the closest neighbor of a node is selected as its successor in the chain. In each data gathering round, a node receives data from one of its neighbors, fuses the data with its own and transmits the fused data to its other neighbor along the chain. Eventually the leader node which is similar to cluster head transmits the aggregated data to the sink.

The distances that most of the nodes transmit are much less compared to LEACH in which each node transmits to its cluster head. But the main disadvantage of PEGASIS is the necessity of global knowledge of all node positions to pick suitable neighbors and minimize the maximum neighbor distance. Another disadvantage is if the leader node is crash due to some reason then entire aggregated data of the network will be lost.

After study of different hierarchical data aggregation approaches like tree based approach, chain based approach, cluster based approach and grid based approach [1], [3], [4] and [5], LEACH protocol has been selected as a data aggregation protocol which is more energy efficient and cluster based approach.

IV. SIMULATION OF LEACH PROTOCOL USING NS2

Fig. 1 shows the network performance graphs in terms of throughput of the network. Experiment was performed by varying the cluster-heads numbers.

From the results it is clear that performance of the sensor network is better when 5 percent of the nodes are cluster-head in the LEACH protocol. As LEACH assumes that all nodes can communicate with each other and are able to reach the sink. Therefore, it is only suitable for small size networks. If we are going to increase number of cluster heads in a small size of network then collision will occur in the network as data communications to the base station is more. In wireless sensor network data communication consume more energy than computations.

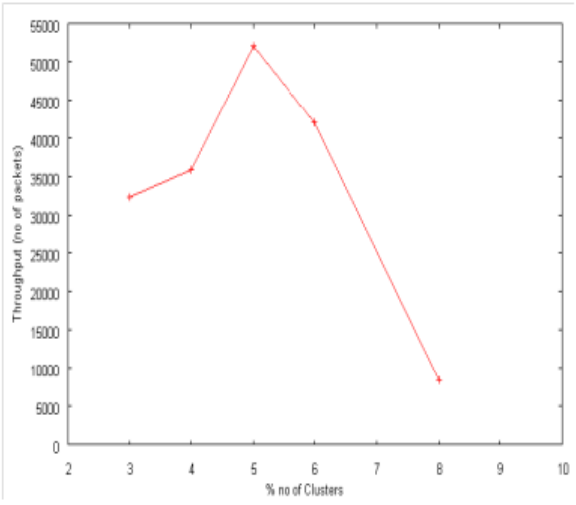

Fig. 1: Number of clusters Vs Throughput of the Network

V. SECURITY ISSUES IN WSN

Most existing proposals for data aggregation are subject to attack. Once a single node is compromised, it is easy for an adversary to inject false data into the network and mislead the aggregator to accept false readings. Because of this, the need for secure data aggregation is raised.

The required security properties to strengthen the security in aggregation schemes are defined as follows:

i). Data Confidentiality ensures that information content is never exposed to anyone who is not authorized to receive it.

ii). Authentication allows the receiver to verify if the message is sent by the claimed sender or not.

iii). Data Integrity ensures that the content of a message has not been altered, either maliciously or by accident, during transmission process.

iv). Data Freshness ensures that the data is recent and that no old messages have been replayed to protect data aggregation schemes against replay attacks.

Data confidentiality can be achieved by using cryptography algorithm like AES, DES, RSA, etc. By using any hashing function like MD5, SHA-1, HMAC, etc we can achieve data integrity.

A. Comparison of AES, DES and RSA

After study of some algorithms [6], [7] here is comparison of two symmetric key algorithms which are AES, DES and one asymmetric key algorithm which is RSA algorithms for different number of nodes.

Method	Approach	Execution. Time(s)	Throughput(bits/s)	Energy Consumed(J)	Security
AES	Symmetic	Faster	Highest	Less	Moderate
DES	Symmetric	Faster	High	Less	Moderate
RSA	Asymmetric	Slow	Moderate	High	Highest

Table I: Performance analysis and comparison of AES,DES and RSA algorithms

Here, compared parameter are Execution time, Throughput and Energy consumed to find the suitable method for WSN.

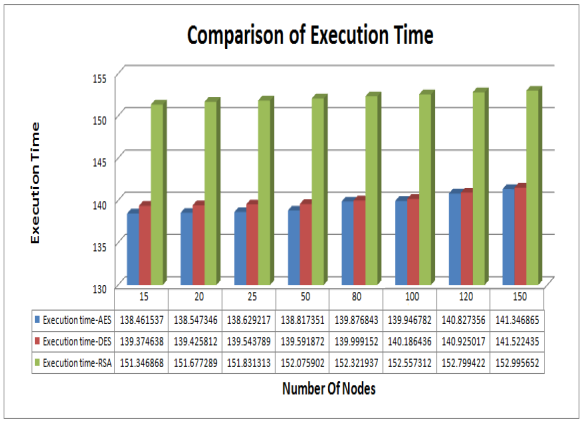

Fig. 2: Comparison of Execution Time

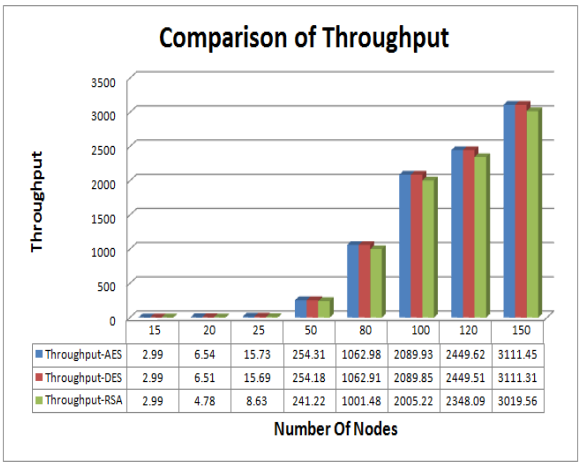

Fig. 3: Comparison of Throughput

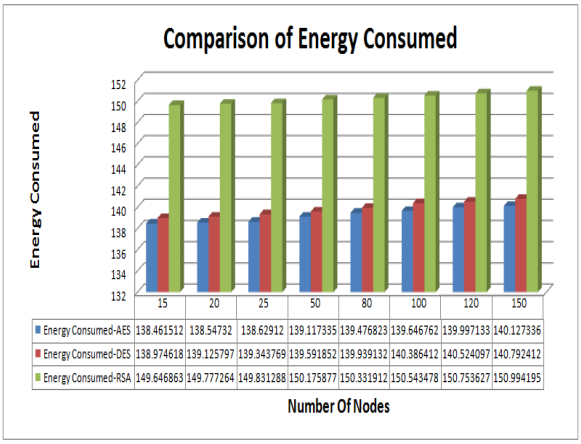

Fig. 4: Comparison of Energy Consumed

VI. CONCUSION AND FUTURE WORK

After survey of different data aggregation approaches in wireless sensor networks which are focusing on optimizing important performance measures such as network lifetime, data latency, throughput and energy consumption, LEACH protocol has been selected. It achieves greater network lifetime, throughput and also consumes less energy when 5 percent of the nodes are cluster-head in the LEACH protocol.

Based on Experimental results of Encryption algorithms AES, DES and RSA, AES are selected to encrypt and decrypt algorithms. RSA is a asymmetric cryptography scheme which involve computationally intensive mathematical functions and WSN are highly resource constraint with having limited energy lifetime, small memory and limited computation and communication capabilities.

In future work addition of Key Distribution Center (KDC) at each data aggregator node in WSN is possible to make AES algorithm more secure.

REFERENCES

[1] Kiran Maraiya, Kamal Kant, Nitin Gupta "Wireless Sensor Network: A Review on Data Aggregation" International Journal of Scientific & Engineering Research Volume 2, Issue 4, April - 2011 ISSN 2229-5518.

[2] Hani Alzaid Ernest Foo Juan Gonzalez Nieto "Secure Data Aggregation in Wire-less Sensor Network: a survey" Information Security Institute Qeensland University of Technology, PO Box 2434, Brisbane, Queensland 4001.

[3] Ankita Joshi and Lakshmi Priya.M "A Survey of Hierarchical Routing Protocols in Wireless Sensor Network" MES Journal of Technology and Management.

[4] D. J. Dechene, A. El Jardali, M. Luccini, and A. Sauer, "A Survey of Clustering Algorithms for Wireless Sensor Networks".

[5] Suraj Sharma and Sanjay Kumar Jena "A Survey on Secure Hierarchical Routing Protocols in Wireless Sensor Networks" ICCCS'11 February 12-14, 2011, Rourkela, Odisha, India.

[6] Shashi Mehrotra Seth and Rajan Mishra "Comparative Analysis Of Encryption Algorithms For Data Communication" IJCST Vol. 2, Issue 2, June 2011.

[7] NeetuSettia. "Cryptanalysis of modern Cryptography Algorithms". International Journal of Computer Science and Technology, December 2010.

Shivani Desai B.E in Computer Engineering from Government. Enginnering. College, Gandhinagar, Gujarat University, Gujarat, India, Year 2009. Working in area of Wireless Sensor Network. Pursuing M.tech from Nirma University, Ahmedabad, Gujarat, India. Field of research is Wireless Sensor Network and Security.

Sejal L. Butani B.Tech. Information and Technology from UVPCE,Ganpat university , Mehsana, Gujarat, India, Year 2008. Working in the area of Wireless Ad hoc Network. Pursuing M.Tech from Nirma University, Ahmedabad, Gujarat, India. Field of research is Wireless Network and Security.

Sharada Valiveti M.E. in Computer Engineering from ISTAR, SP University, Vallabh Vidyanagar, Gujarat, India in the year 2005. Working in the area of Ad Hoc Network Security. Pursuing Ph.D. from Nirma University, Ahmedabad, Gujarat, India. Field of research is Enhancing Security with Effective Routing Mechanisms for Ad Hoc Networks.