Secured Data Aggregation in Wireless Sensor Network

BY

Shivani Desai

09MCES03

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AHMEDABAD-382481

May-2012

Secured Data Aggregation in Wireless Sensor Network

Major Project

Submitted in partial fulfillment of the requirements

For the degree of

Master of Technology in Computer Science and Engineering

By

Shivani Desai

09MCES03

Guide

Sharada Valiveti

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AHMEDABAD-382481

May-2012

UNDERTAKING

I, Shivani Desai , Roll No. 09MCES03, give undertaking that the Major Project entitled "Secured Data Aggregation for Wireless Sensor Network" submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in Computer Science and Technology of Nirma University, Ahmedabad, is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Signature

Date:

Place:

Certificate

This is to certify that the Major Project entitled "Secured Data Aggregation in Wireless Sensor Network" submitted by Shivani Desai (09MCES03), towards the partial fulfillment of the requirements for the degree of Master of Technology in Computer Science and Engineering of Nirma University , Ahmedabad is the record of work carried out by her under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination.The results embodied in this major project, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

Prof. Sharada Valiveti
Guide and Asso.Professor,
Department of Computer Engineering,
Institute of Technology,
Nirma University, Ahmedabad

Dr. S. N. Pradhan
Professor and PG-Coordinator(M.Tech(CSE)),
Department of Computer Engineering,
Institute of Technology,
Nirma University, Ahmedabad

Prof. D. J. Patel
Professor and Head(CSE),
Institute of Technology,
Nirma University, Ahmedabad,

Dr K Kotecha
Director,
Institute of Technology,
Nirma University, Ahmedabad

Acknowledgements

With immense pleasure, I would like to present this report on the dissertation work related to "Secured Data Aggregation in Wireless Sensor Network". It gives me great pleasure in expressing thanks and profound gratitude to Prof. Sharada Valiveti, Associate Professor, Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad for her valuable guidance and constant support and interest throughout the project work. I am heartily thankful to her for her time to time suggestions and the clarity of the concepts of the topic that helped me a lot during this study.

I would like to extend my gratitude to Dr.S.N.Pradhan, Programme Co-ordinator M.Tech. CS&E, Institute of Technology, Nirma University, Ahmedabad whose keen interest and excellent knowledge base helped me to finalize the topic of the dissertation work.

My sincere thanks and gratitude to Prof. D.J. Patel, Professor and Head, Computer Science & Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his continual kind words of encouragement and motivation throughout the Dissertation work.

I would like to thank Dr Ketan Kotecha, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for providing basic infrastructure and healthy research environment.

I am thankful to Nirma University for providing all kind of required resources. I would like to thank The Almighty, my family, for supporting and encouraging me in all possible ways. I would also like to thank all my friends who have directly or indirectly helped in making this dissertation work successful.

- Shivani Desai
09MCES03

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/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. The project aims to identify an energy efficient data aggregation protocol. Among various data aggregation techniques LEACH routing protocol has been selected 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. Implementation of security on NS-2 is necessary in network simulation. However, currently, NS-2 does not support these features. Our project aims to solve this issue. The purpose of the project is to find a way to add encryption/decryption features into network simulation program NS-2. For the purposes of this analysis, the assumption is that the key is pre-shared and the encryption/decryption algorithm is illustrative. Based on the experimental results, AES algorithm is selected to encrypt and decrypt algorithms. Finally, integration of LEACH protocol with AES encryption/decryption algorithm has been done to provide secured data aggregation in WSN which facilitates more energy efficiency.

Abbreviation Notation and Nomenclature

WSN ..... Wireless Sensor Network
MANET ..... Monile Adhoc NETwork
LEACH ..... Low Energy Adaptive Clustering Hierarchy
HEED ..... Hybrid Energy Efficient Distributed Clustering Approach
EECS ..... Energy Efficient Clustering Scheme
TDMA ..... Time Division Multiple Access
CDMA ..... Collision Detection Multiple Access
MAC ..... Media Access Control
CH ..... Cluster Head
AMRP ..... Average Minimum Reachability Power
BS ..... Base Station
PEGASIS ..... Power Efficient data GAthering protocol for Sensor Information Systems
EADAT ..... Energy Aware Distributed heuristic
PEDAP ..... Power Efficient Data gathering and Aggregation Protocol
NS ..... Network Simulator
DSDV ..... Destination-Sequenced Distance Vector
DSR ..... Dynamic Source Routing
AODV ..... Ad hoc On-demand Distance Vector routing
TCP ..... Transmission Control Protocol
IP ..... Internet Protocol
UDP ..... User Datagram Protocol
OTcl ..... Object oriented Tool Command Language
AMPS ..... Adaptive Multidomain Power-aware Sensors
SHA-1 ..... Secure Hash Algorithm
MD-5 ..... Message Digest Algorithm
DES ..... Data Encryption Standard
AES ..... Advance Encryption Standard

Contents

List of Tables

List of Figures

Chapter 1

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. The sensors could be scattered randomly in harsh environments such as a battlefield or deterministically be placed at specified locations. The sensors periodically sense the data, process it and transmit it to the base station. The frequency of data reporting and the number of sensors which report data usually depends on the specific application.

1.1 Applications

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: gathering information from a region where it is located. Generally data like heat, pressure, sound, light, vibration, electromagnetic field etc. Environmental monitoring: measurement of tem-

perature, rainfall etc.

1.2 Objective of the Work

The main objective of the work is to increase energy efficiency of WSN.

Specific Objectives

- Increase Network lifetime
- Increase Throughput
- Reduce Latency
- Reduce Communication overhead

1.3 Scope of the work

- In data aggregation process , the cluster head are the data aggregator nodes that send the aggregated data to the base station may be attacked by malicious attacker. If a cluster head is compromised, then the base station (sink) can not ensure the correctness of the aggregated data that has been send to it.
- LEACH assumes that all nodes can communicate with each other and are able to reach the sink. Therefore, it is only suitable for small sized networks.

1.4 Motivation

Sensor nodes 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 transmit that sensed data packet to the base station (sink). This process is a direct transmission and it may also 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. So a better technique is to have

fewer nodes send the data to the base station. These nodes called the aggregator nodes and processes 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 also does. 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. Another factor that affects the WSN performance is communication instability. For example, if two sensors that have same aggregator node start sending packets at the same time, conflicts will occur near the aggregator node and the transfer process will fail. In addition, packets might get dropped at highly congested nodes, since the packet based routing of the WSN is connectionless, which is inherently unreliable. Due to these limitations, security issues such as data integrity, confidentiality, and freshness in data aggregation become crucial when the WSN is deployed in a remote or hostile environment where sensors are prone to node failures and compromises. However, the aggregators are vulnerable to attack. 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. Secure data aggregation scheme is required which is classified in to: hop-by-hop and end-to-end encrypted data aggregation.

1.5 Thesis Organization

The rest of the thesis is organized as follows.
Chapter 2, Literature Survey for Data Aggregation, describes the need for data aggregation in WSN and Survey of different hierarchical data aggregation approaches in WSN in the process of achieving energy efficiency.
Chapter 3, Study of Network Simulator: NS-2, gives the brief introduction of NS-2.
Chapter 4, Simulation of LEACH routing protocol using NS-2, describes installation of NS-2.27 and LEACH extension on Fedora. It also describes performance of WSN in presence of LEACH protocol which is based on simulation results of LEACH protocol.

Chapter 5, Literature survey for adding security, describes why there is a need of security issues like Data Confidentiality, Data Integrity, Data Availability, Authentication, etc in WSN.

Chapter 6, Implementation of security in NS-2, describes how to add encryption/decryption features into network simulation program NS-2 which is currently not available in NS-2. In this section, we presented the implementation of AES, DES and RSA algorithms for different number of nodes and compared the parameter like execution time, throughput and consumed energy to find the suitable method for WSN are presented.

Chapter 7, Integration of LEACH with AES and RSA, shows simulation results of LEACH with security algorithm like AES and RSA by varying the cluster numbers and evaluated the performance of the network in terms of execution time of the sensor network and total energy consumption by the network.

Chapter 8, finally conclude the project work and also shows the future work of this project.

Chapter 2

Literature Survey for Data Aggregation

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 on the network architecture. Data communication between sensors, aggregators and the querier consumes a large portion of the total energy consumption of the WSN. The WSN in Figure 1 contains 16 sensor nodes and uses SUM function to minimize the energy consumption by reducing the number of bits reported to the base station. Node 7, 10-16 are normal sensor nodes that are generating the data by its sensing mechanism whereas nodes 1-6, 8, 9 are aggregators that collect data from a subset of the network, aggregate the data using a suitable aggregation function like sum and then transmit the aggregated result to an upper aggregator or to the base station. In this example 16 packets traveled within the network and only one packet is transmitted to the base station.

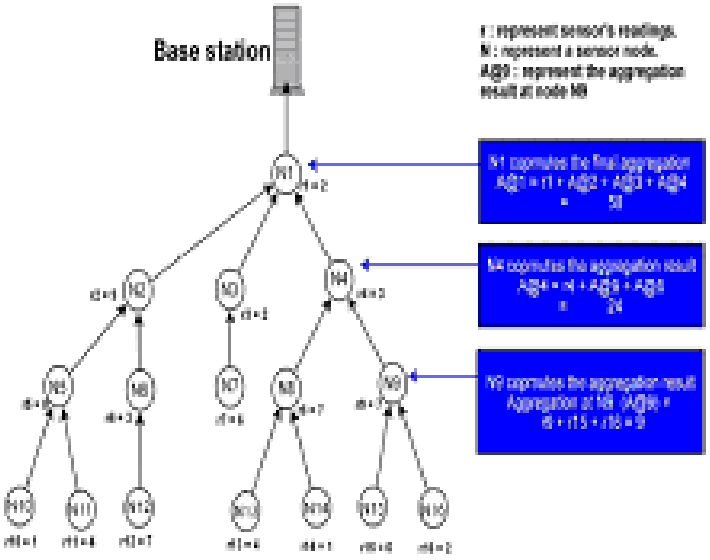


Figure 2.1: An aggregation scenario using sum function
[1]

2.1 Advantage and Disadvantage of Data aggregation in wireless sensor network

Advantage: With the help of data aggregation process we can enhance the robustness and accuracy of information which is obtained by entire network. Certain redundancy exists in the data collected from sensor nodes thus data fusion (data aggregation) processing is needed to reduce the redundant information. Another advantage is that it reduces the traffic load and conserve energy of the sensors.

Disadvantage: The cluster head means data aggregator nodes send these aggregated data to the base station. This cluster head or aggregator node may be attacked by malicious attacker. If a cluster head is compromised, then the base station (sink) cannot ensure the correctness of the aggregated data that has been sent to it.

2.2 Performance measure of data aggregation

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

- Energy Efficiency:

  By the data-aggregation scheme, we can increase the functionality of the wireless sensor network. 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, in applications where the time that all nodes operate together is very important, lifetime is defined as the number of rounds until the first sensor is drained of its energy.

- Latency:

  Latency is defined as the delay involved in data transmission, routing and data aggregation. 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.

- Data accuracy:

  The definition of data accuracy depends on the specific application for which the sensor network is designed. For instance, in a target localization problem, the estimate of target location at the sink determines the data accuracy.

- Throughput:

  It defines the average data rate of successful message delivery over a communication channel. This data may be delivered over a physical or logical link, or pass through a certain network node.The throughput is usually measured in bits per second, and sometimes in data packets per second or data packets per time slot.

$$Throughput = \frac{Receivedpacket'ssize}{stopTime - startTime} \tag{2.1}$$

Table 2.1: Data aggregation in hierarchical networks Vs. flat networks

Hierarchical networks	**Flat networks**
Data aggregation performed by cluster heads or a leader node.	Data aggregation performed by different nodes along the multi-hop path.
Overhead involved in cluster or chain formation throughout the network.	Data aggregation routes are formed only in regions that have data for transmission.
Even if one cluster head fails,the network may still be operational.	The failure of sink node may result in the break down of entire network.
Lower latency is involved since sensor nodes perform short range transmissions to the cluster head.	Higher latency is involved in data transmission to the sink via a multi-hop path.
Routing structure is simple but not necessarily optimal.	Optimal routing can be guaranteed with additional overhead.

2.3 Comparison of architecture of the sensor networks

The architecture of the sensor network plays a vital role in the performance of different data aggregation protocols. This section includes the comparison of two architecture of wireless sensor network which are Flat networks and Hierarchical networks.

- Flat networks:

  In flat networks, each sensor node plays the same role and is equipped with approximately the same battery power. In such networks, data aggregation is accomplished by data centric routing where the sink usually transmits a query message to the sensors, e.g, via flooding. Sensors which have data matching the query would then send response messages back to the sink.

- Hierarchical networks:

  A flat network can result in excessive communication and computation burden at the sink node resulting in a faster depletion of its battery power. The death of the sink node breaks down the functionality of the network.

From the comparison, we conclude that hierarchical networks are good for data aggregation in WSN. Hence, in view of scalability and energy efficiency, several hierarchical data aggregation approaches (protocols) have been proposed. Hierarchical data aggregation involves data fusion at special nodes, which reduces the number of messages transmitted to the sink. This improves the energy efficiency of the network.

2.4 Hierarchical data aggregation approaches

2.4.1 Data aggregation in cluster based approach

In energy-constrained sensor networks of large size, it is inefficient for sensors to transmit the data directly to the sink. 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. In such approach, sensors can transmit data to a local aggregator or cluster head which aggregates data from all the sensors in its cluster and transmits the concise digest to the sink. This results in significant energy savings for the energy constrained sensors. Figure shows a cluster based sensor network organization. As shown in the figure cluster heads can communicate with the sink directly via long range transmissions or multi hopping through other cluster heads.

Disadvantage of this approach, This approach of clustering may introduce overhead due to the cluster configuration and maintenance, but it has been demonstrated that this approach exhibit better energy consumption and it may also possible that cluster head act like a malicious node. So there is a need of security in this kind of approach.

Challenges of clustering: Wireless Sensor Networks present vast challenges in terms of implementation. There are several key attributes that designers must carefully consider, which are of particular importance in wireless sensor networks.

- Cost of Clustering:

  Although clustering plays a vital role in organizing sensor network topology, there are often many resources such as communication and processing tasks needed in the creation and maintenance of the clustering topology. Such costs as the required resources are not being used for data transmission or sensing tasks.

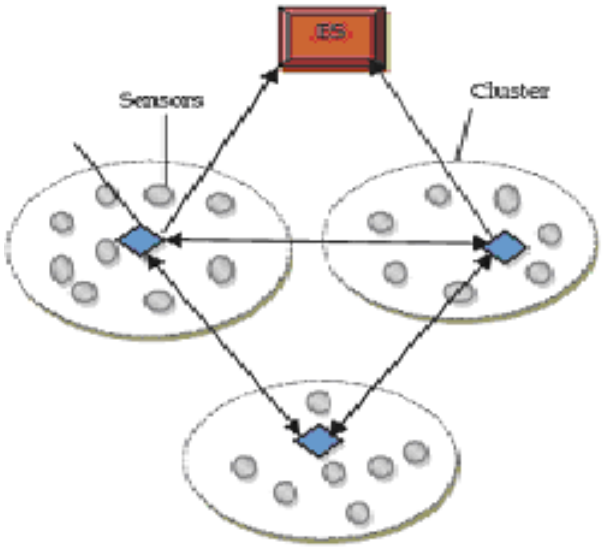


Figure 2.2: cluster based sensor network
[4]

- Selection of Cluster heads and Clusters:

  The clustering concept offers tremendous benefits for wireless sensor networks. However when designing for a particular application, designers must carefully examine the formation of clusters in the network. Depending on the application, certain requirements for the number of nodes in a cluster or its physical size may play an important role in its operation. This prerequisite may have an impact on how cluster heads are selected in this application.

- Real-Time Operation:

  Useful lifetime of data is also a fundamental criterion in designing Wireless Sensor Networks. In applications such as habitat monitoring , simply receiving data is sufficient for analysis, meaning delay is not an important issue. When we look at a military tracking, the issue of real-time data acquisition becomes much more vital. When looking at clustering algorithms, important attention must be paid to the delay created by the clustering scheme itself. In addition, the time required for cluster recovery mechanisms must also be taken into account.

- Synchronization:

One of the primary limitations in Wireless Sensor Networks is the limited energy capacity of nodes. Slotted transmission schemes (such as TDMA), allow nodes to regularly schedule sleep intervals to minimize energy used. Such schemes require synchronization mechanisms to setup and maintain the transmission schedule. When considering a clustering scheme, synchronization and scheduling will have a considerable effect on network lifetime and the overall network performance.

- Data Aggregation:

  One major advantage of wireless sensor networks is the ability for data aggregation to occur in the network. In a densely populated network there are often multiple nodes sensing similar information. Data aggregation allows the differentiation between sensed data and useful data. In-network processing makes this process possible and now it is fundamental in many sensor network schemes , as the power required for processing tasks is substantially less than communication tasks. As such, the amount of data transferred in-network should be minimized. Many clustering schemes provide data aggregation capabilities , and as such, the requirement for data aggregation should be carefully considered when selecting a clustering approach.

- Repair Mechanisms:

  Due to the nature of Wireless Sensor Networks, they are often prone to node mobility, node death and interference. All of these situations can result in link failure. When looking at clustering schemes, it is important to look at the mechanisms in place for link recovery and reliable data communication.

- Quality of Service (QoS):

  From an overall network standpoint, we can look at QoS requirements in Wireless Sensor Networks. Many of these requirements are application dependant (such as acceptable delay and packet loss tolerance), and as such, it is important to look at these metrics when choosing a clustering scheme. Implementations can vary widely in terms of these metrics, and as a result, the design process should consider these aspects.

Recently, several cluster based network organization and data aggregation protocols have been proposed. In this section we discuss three such protocols which are Low Energy Adaptive Clustering Hierarchy (LEACH), Hybrid Energy Efficient Distributed Clustering Approach (HEED) and Energy Efficient Clustering Scheme (EECS).

LEACH

LEACH stands for Low-Energy Adaptive Clustering Hierarchy and LEACH was 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.

Operation

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 formulae is as follows:

$$T(n) = \frac{p}{1 - p \times (r \times mod\frac{1}{p})} \forall n \in G \tag{2.2}$$

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.

HEED

HEED 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. Clus-

ter 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 reachability 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.

The simulation results show that HEED improves the network lifetime over gen-LEACH. In gen-LEACH the selection of cluster heads is random which may result in rapid death of certain nodes. However, in HEED the cluster heads are selected such that they are well distributed with minimum communication cost. In addition, the energy dissipated in clustering is less in HEED compared to gen-LEACH. This is due to the fact that gen-LEACH propagates residual energy. To conclude, HEED prolongs network lifetime and achieves a geographically well-distributed set of cluster heads.

EECS

An Energy Efficient Clustering Scheme (EECS) 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. EECS is a LEACH-like clustering scheme, where the network is partitioned into a set of clusters with one cluster head in each cluster. Communication between cluster head and BS is direct (single-hop).In the network deployment phase, the BS broadcasts a "hello" message to all the nodes at a certain power level. By this way each node can compute the approximate distance to the BS based on the received signal strength. It helps nodes to select the proper power level to communicate with the

BS. Also this distance is used to balance the load among cluster heads. In cluster head election phase, well distributed cluster heads are elected with a little control overhead. And In cluster formation phase, a novel weighted function is introduced to form load balanced clusters.

2.4.2 Chain based data aggregation

In cluster-based sensor networks, sensors transmit data to the cluster head where data aggregation is performed. However, if the cluster head is far away from the sensors, they might expend excessive energy in communication. Further improvements in energy efficiency can be obtained if sensors transmit only to close neighbors. The key idea behind chain based data aggregation is that each sensor transmits only to its closest neighbor. Power Efficient data Gathering protocol for Sensor Information Systems (PEGASIS) 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. Figure shows the chain based data aggregation procedure in PEGASIS. Nodes take turns in transmitting to the sink. The PEGASIS

protocol has considerable energy savings compared to LEACH. The distances that most of the nodes transmit are much less compared to LEACH in which each node transmits to its cluster head. The leader node receives at most two data packets from its two neighbors. In contrast, a cluster head in LEACH has to perform data fusion of several data packets received from its cluster members. 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.

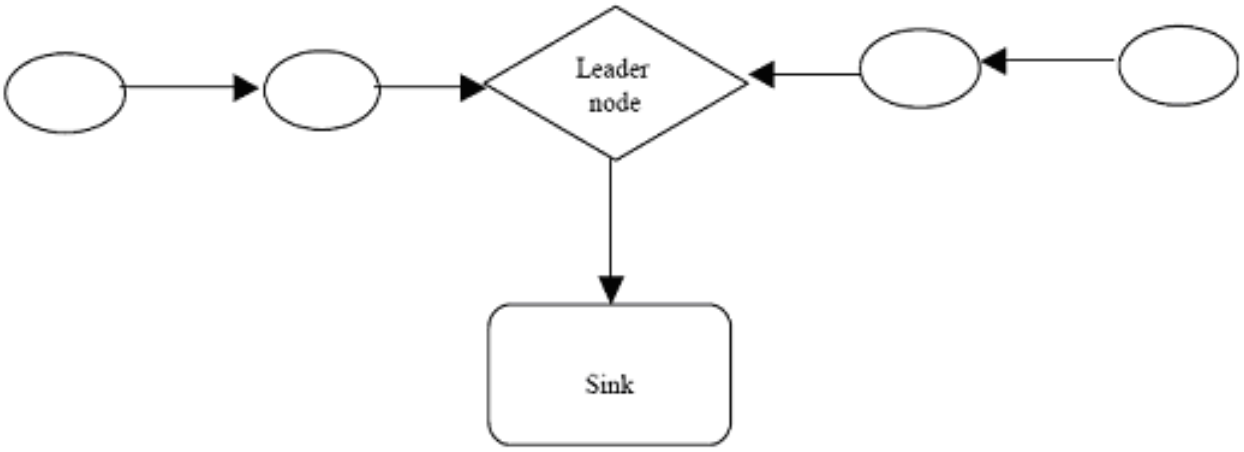


Figure 2.3: Chain based organization in sensor network

2.4.3 Tree based data aggregation

The tree based approach is defining aggregation from constructing an aggregation tree. The form of tree is minimum spanning tree, sink node consider as a root and source node consider as a leaves. Information flowing of data start from leaves node up to root means sink(base station). Tree based data aggregation is suitable for applications which involve in-network data aggregation. An example application is radiation level monitoring in a nuclear plant where the maximum value provides the most useful information for the safety of the plant. Tree based protocol called an energy aware distributed heuristic (EADAT) is

used to construct and maintain a data aggregation tree in sensor networks. In applications where each sensor node has data to send to the sink in every round of communication, it is essential to maximize the network lifetime. In such scenario, a power efficient data gathering and aggregation protocol (PEDAP) is used. The goal of PEDAP is to maximize the lifetime of the network in terms of number of rounds, where each round corresponds to aggregation of data transmitted from different sensor nodes to the sink. PEDAP is a minimum spanning tree based protocol which improves the lifetime of the network even when the sink is inside the field. In contrast, LEACH and PEGASIS perform poorly when the sink is inside the sensor field.

Disadvantage of this approach: as we know like wireless sensor network are not free from failure .So in case of data packet loss at any level of tree, the data will be lost not only

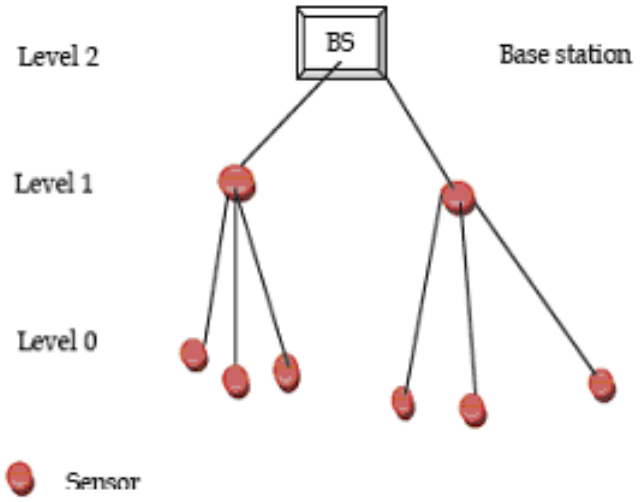


Figure 2.4: Tree based data aggregation
[4]

for single level but for whole related sub tree as well.

2.4.4 Grid based data aggregation

In grid-based data aggregation, the data aggregator is fixed in each grid and it aggregates the data from all the sensors within the grid. This is similar to cluster based data aggregation in which the cluster heads are fixed. Grid-based data aggregation is suitable for mobile environments such as military surveillance and weather forecasting and adapts to dynamic changes in the network and event mobility.

2.5 Summary

After study of different hierarchical data aggregation approaches LEACH protocol has been selected which is cluster based approach as it has been prove that it is more energy efficient and easy to implement.

Protocol	Organization type	Objectives	Characteristics
LEACH	cluster	Network lifetime: number of nodes that are alive, latency	Randomized cluster head rotation, non-uniform energy drainage across different sensors.
HEED	cluster	Lifetime: number of rounds until the first node death	Assumption: Multiple power levels in sensors. Cluster heads are well distributed. Achieves better performance than LEACH
PEGASIS	chain	Lifetime: average energy expended by a node	Global knowledge of the network is required. Considerable energy savings compared to LEACH.
Hierarchical chain based protocols	chain	Energy× delay	Binary chain based scheme is eight times better than LEACH and the three level scheme is 5 times better than PEGASIS.
EADAT	tree	Lifetime: number of alive sensors at the end of simulation time	Sink initiated broadcasting approach. It is not clear how to choose the threshold power (P_{th}) for broadcasting help messages. No comparisons made with other existing aggregation algorithms.
PEDAP-PA	tree	Lifetime: time until the death of last node	Minimum spanning tree based approach. Achieves two times performance improvement compared to LEACH, PEGASIS.

Figure 2.5: Summary of hierarchical data aggregation protocols

Chapter 3

Study of Network Simulator: NS-2

NS-2 is an open source system that is developed using C++ and Tool Control Language TCL. Researchers can freely add new components to the system to serve their own purposes. It provides support for IP protocols suite and many standard routing protocols for wire and wireless networks. There are unicast and multicast routing protocols for wire network and DSR, DSDV, AODV for wireless networks.

The network simulator (NS), which is a discrete event simulator for networks, is a simulated program developed by VINT (Virtual InterNetwork Testbed) project group (A Collaboration among USC/ISI, Xerox PARC, LBNL, and UCB). It supports simulations of TCP and UDP, some of MAC layer protocols, various routing and multicast protocols over both wired and wireless network etc. The basic structure of NS-2 is as shown in figure.

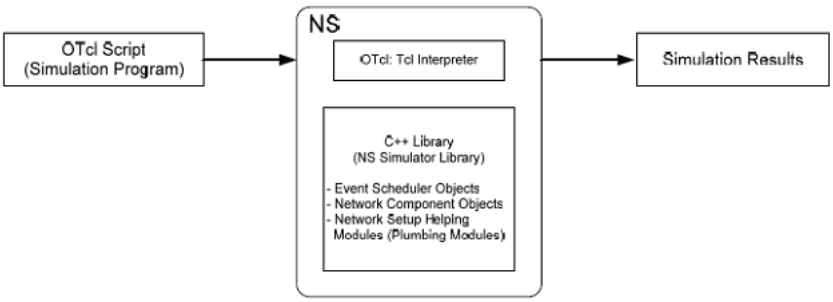


Figure 3.1: Basic Structure of NS

To setup and run a simulation, a user writes an OTcl script, which is a simulation program to initiate an event scheduler, set up the network topology using the network

objects and plumbing functions in the library, and to tell traffic sources when to start and stop transmitting packets through the event scheduler. When NS-2 which works as OTcl interpreter receives the OTcl script, it will set environment parameters following the received script. If a user wants to make a new network object, it will be easy to make a compound object from the object library, and plumb the data path through the object rather than write a new one. When the simulation is finished, the simulation results are produced in one or more text-based output files that contain detailed simulation data, which can be used to analyze directly.

3.1 MIT's extension to NS-2

The MIT uAMPS (Adaptive Multidomain Power-aware sensors) extensions to ns add support for large scale wireless sensor networks. These extensions include models for node energy dissipation and node state, as well as several routing protocols. It describes the uAMPS additions to ns and gives details on how to use the code and run the simulator for wireless sensor networks. It is an extension for the implementation of LEACH routing protocol.

3.2 Summary

Here,NS-2 is selected as a network simulator tool,it is an open source system that is developed using C++ and Tool Control Language.So it is easy for user to add specific components to the system to serve their own purposes.

Chapter 4

Simulation of LEACH routing protocol using NS-2

4.1 Installing NS-2.27 and LEACH extension on Fedora 10

Step 0: Prepare neccesary files for installation

- NS-2.27 package: ns-allinone-2.27.tar.gz
- Patch for compiling NS-2.27 with GCC 4.1.x: ns-2.27-gcc410.patch
- MIT's LEACH extension: mit.tar.gz
- LEACH's Makefile patch: leach_makefile-2.27.patch

Step 1: Download NS-2.27, apply ns-2.27-gcc410.patch, and install it under your home directory (/root):

- wget http://www.isi.edu/nsnam/dist/ns-allinone-2.27.tar.gz
  OR wget http://www.internetworkflow.com/downloads/ns2leach/ns-allinone-2.27.tar.gz
- tar zxvf ns-allinone-2.27.tar.gz
- wget http://www.tekno.chalmers.se/ yusheng/reports/ns-2.27-gcc410.patch
- LEACH's Makefile patch: leach_makefile-2.27.patch

- patch -p0 < ns-2.27-gcc410.patch
- cd ns-allinone-2.27/
- ./install

Step 2: Set the environment variables to make NS-2.27 works

- cd /root
- gedit .bashrc

  # LD_LIBRARY_PATH

  OTCL_LIB=/root/ns-allinone-2.27/otcl-1.8

  NS2_LIB=/root/ns-allinone-2.27/lib

  X11_LIB=/usr/X11R6/lib

  USR_LOCAL_LIB=/usr/local/lib

  export

  LD_LIBRARY_PATH=$LD\_LIBRARY\_PATH$:OTCL_LIB:$NS2\_LIB$:X11_LIB:USR_LOCA

  $\#TCL\_LIBRARY$

  $TCL\_LIB = /root/ns - allinone - 2.27/tcl8.4.5/library$

  $USR\_LIB = /usr/lib$

  $exportTCL\_LIBRARY$ =TCL_LIB:USR_LIB

  $\#PATH$

  $PATH$ =PATH:/root/ns-allinone-2.27/bin:/root/ns-allinone-2.27/tcl8.4.5/unix:/root/ns-allinone-2.27/tk8.4.5/unix

- source .bashrc

Step 3: Download, copy, and extract MIT's LEACH extension

- wget http://www.internetworkflow.com/downloads/ns2leach/mit.tar.gz
- cp mit.tar.gz /root/ns-allinone-2.27/ns-2.27/
- cd /root/ns-allinone-2.27/ns-2.27/
- tar xzvf mit.tar.gz

- rm mit.tar.gz

Step 4: Modify NS-2 source code

- gedit /root/ns-allinone-2.27/ns-2.27/mac/wireless-phy.cc
  Goto line 59, that is after line 58:
  #define max(a,b) (((a)<(b))?(b):(a))
  Insert:
  #define min(a,b) (((a)>(b))?(b):(a))

- cp mit.tar.gz /root/ns-allinone-2.27/ns-2.27/
- cd /root/ns-allinone-2.27/ns-2.27/
- tar xzvf mit.tar.gz
- rm mit.tar.gz

Step 5: Add enviroment variables for LEACH extension

- gedit /root/.bashrc
  Goto line 59, that is after line 58:
  export RCA_LIBRARY=$NS/mit/rca$
  $exportuAMPS\_LIBRARY$ =NS/mit/uAMPS
- source /root/.bashrc

Step 6: Download and apply patch for Makefile.vc, edit Makefile and re-compile NS-2.27 with LEACH extension

- wget http://voyager.ce.fit.ac.jp/wiki/tool/leach_makefile-2.27.patch
- patch -p0 < leach_makefile-2.27.patch
- gedit Makefile Add -DMIT_uAMPS to the DEFINE list
  Add -I./mit/rca -I./mit/uAMPS to the INCLUDE list
  Add the following just prior to the line gaf/gaf.o

newline mit/rca/energy.o mit/rca/rcagent.o
newline mit/rca/rca-ll.o mit/rca/resource.o
newline mac/mac-sensor-timers.o mac/mac-sensor.o mit/uAMPS/bsagent.o
newline

- make clean
- make

Step 7: Test and debug

- ./test

4.2 Performance comparison based on simulation result of LEACH Protocol

This chapter will includes the various kind of analysis on the results obtained after the simulation of LEACH protocol. To simulate the LEACH protocol, MITs NS2 extension for LEACH simulator [11] is used.
After installation of LEACH protocol successfully,a random test network was used having 100 nodes with base station located at (50,175) (not shown) as shown in figure.For the experiments, each nodes was initial given 2J of energy.

Following command is used to get the simulation result. In following table, Simulation results shows how the performance of the sensor network using LEACH protocol varies as the percent of the nodes that are cluster-heads is changed.By these experiments, we concluded the optimal percent of nodes that should be cluster-heads.

Following figure shows the network performance graphs in terms of throughput of the network.

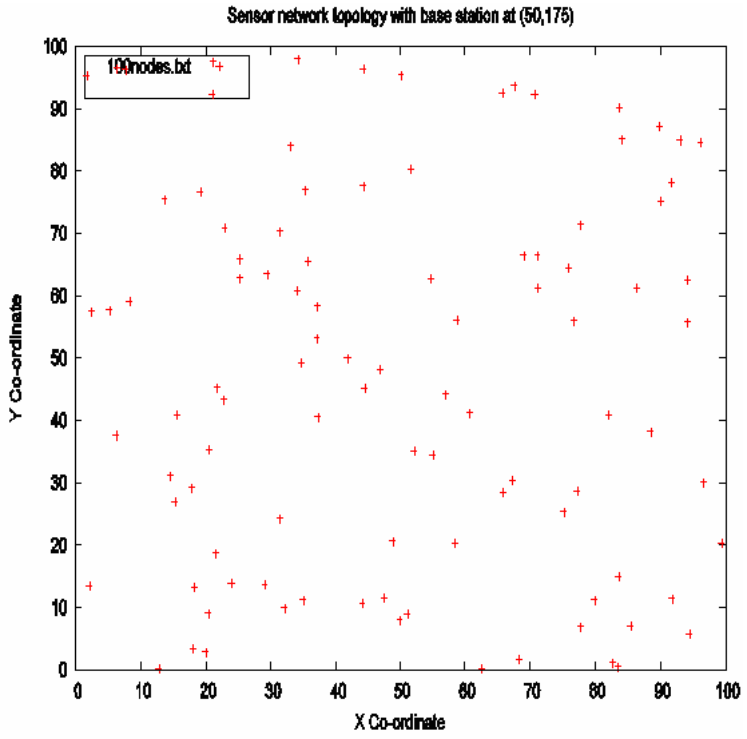


Figure 4.1: Sensor network topology with base station at(50,175)

4.3 Summary

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. Despite the significant overall energy savings, however, the various assumptions made by LEACH are as follows:

- LEACH assumes that all nodes begin with the same amount of energy and that the amount of energy a CH consumes is more than that of a non-cluster node.

Table 4.1: Simulation Results of LEACH Protocol

% Number of clusters	Throughput
3	32279
4	35897
5	52127
6	42041
8	8301

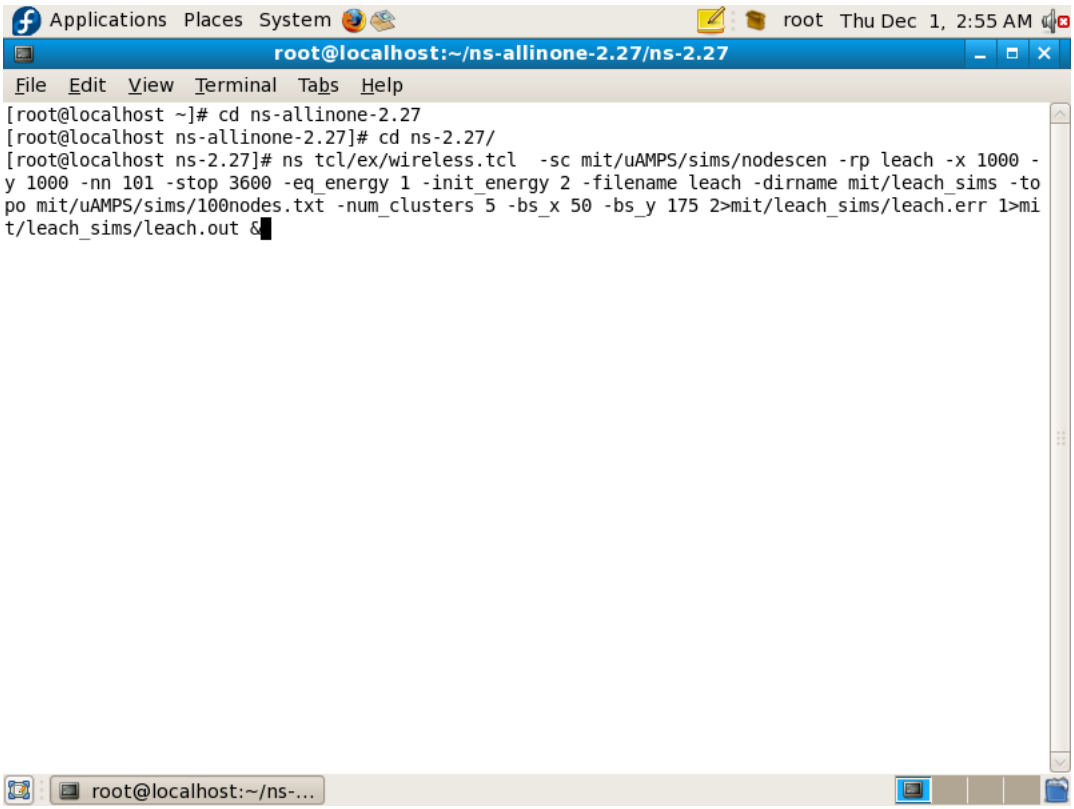


Figure 4.2: Command to get simulation results

- 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.

- LEACH requires that all nodes are continuously listening. This is not realistic in a random distribution of the sensor nodes, for example, where cluster-heads would be located at the edge of the network.

- Finally, there is no mechanism to ensure that the elected cluster-heads will be uniformly distributed over the network. Hence, there is the possibility that most of the cluster-heads are concentrated in one part of the network, which is comparable with the problem like local minima.

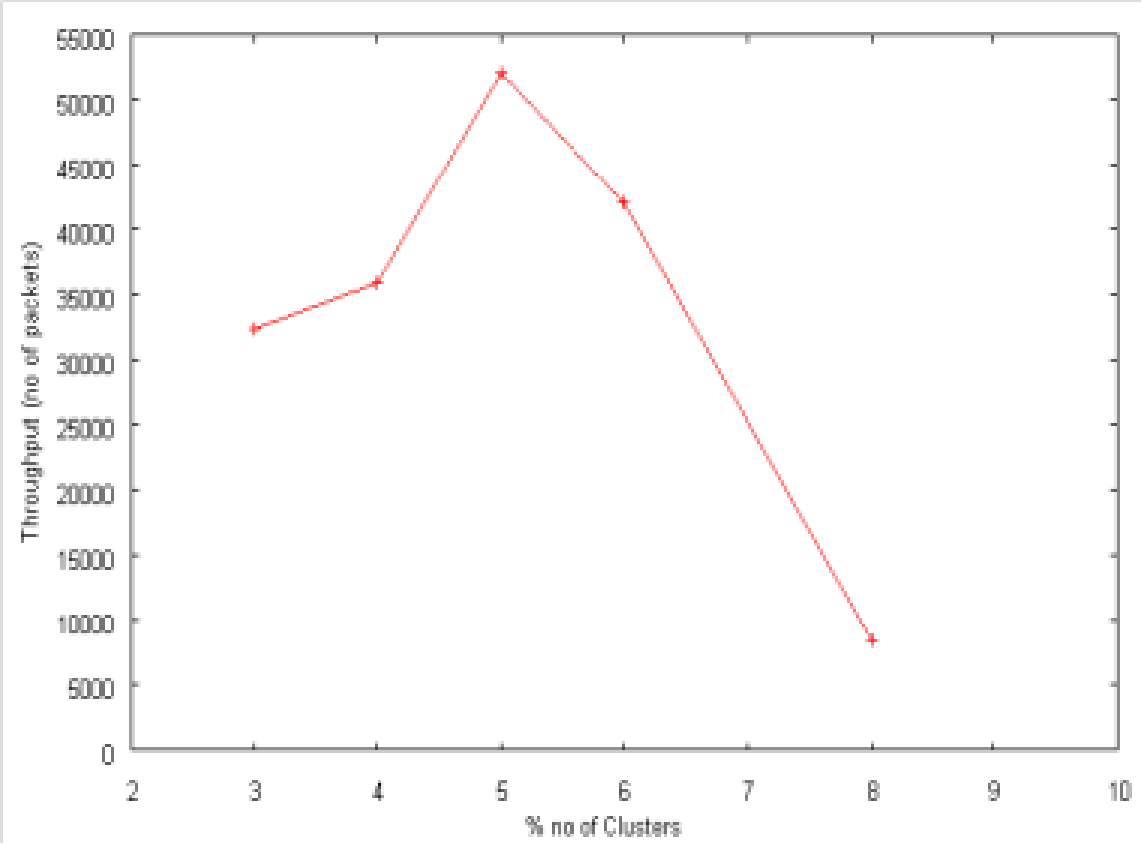


Figure 4.3: Number of clusters Vs Throughput of the Network

Literature survey for adding security

Chapter 5

Literature survey for adding security

5.1 Security goals 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. A general definition for secure data aggregation is the efficient delivery of the summary of sensor readings that are reported to an off-site user in such a way that ensures these reported readings have not been altered. Moreover, a detailed definition of secure data aggregation is proposed as the process of obtaining a relative estimate of the sensor readings with the ability to detect and reject reported data that is significantly distorted by corrupted nodes or injected by malicious nodes . However, rejecting reported data that is injected by malicious nodes consumes the network resources, specifically the nodes' batteries, since each time the suspicious packet will be processed at the aggregator point. The damage caused by malicious nodes or compromised nodes should be reduced by adding a **self-healing property** to the network. This property helps the network in learning how to handle new threats through extensive monitoring of network events, machine learning and network behavior modeling. Consequently, it is believed that a secure data aggregation scheme for the WSN should have the following properties:

- Ability to reduce the size of the data transmitted through the network.
- Data freshness and integrity are important and should be included in the scheme. However, the application type of the WSN affects the scheme designer's decision regarding whether to add the data confidentiality and availability or not.
- Dynamic response to attack activities by executing of a self-healing mechanism.
- Dynamic aggregator election/rotation mechanism to balance the workload at aggregators.

These properties should work together to provide accurate aggregation results securely without exhausting the network.

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

- Data Confidentiality: ensures that information content is never exposed to anyone who is not authorized to receive it. It can be divided (in secure data aggregation schemes) into a hop-by- hop basis and an end-to-end basis. 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. This kind of confidentiality implementation is not practical for the WSN since it requires extra computation. On the other hand in end-to-end 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: ensures that the content of a message has not been altered, either maliciously or by accident, during transmission process. Confidentiality itself is not enough since an adversary is still able to change the data although it knows nothing about it. Suppose a secure data aggregation scheme focuses only on data confidentiality. An adversary near the aggregator point will be able to change the aggregated result sent to the base station by adding some fragments or manipulating the packet's

content without detection. Moreover, even without the existence of an adversary, data might be damaged or lost due to the wireless environment.

- Data Freshness : ensures that the data are recent and that no old messages have been replayed to protect data aggregation schemes against replay attacks. In this kind of attack, it is not enough that these schemes only focus on data confidentiality and integrity because a passive adversary is able to listen to even encrypted messages transmitted between sensor nodes can replay them later on and disrupt the data aggregation results. More importantly when the adversary can replay the distributed shared key and mislead the sensor about the current key.

- Data Availability: ensures that the network is alive and that data are accessible. It is highly recommended in the presence of compromised nodes to achieve network degradation by eliminating these bad nodes. Once an attacker gets into the WSN by compromising a node, the attack will 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. When the adversary is getting stronger, it is necessary that a secure data aggregation scheme contains some of the following mechanisms to ensure reasonable level of data availability in the network:

  -Self-healing that can diagnose, and react to the attacker's activities especially when he gets into the network and then start corrective actions based on defined policies to recover the network or a node.

  - Aggregator rotation that rotates the aggregation duties between honest nodes to balance the energy consumption in WSN.

- Authentication: There are two types of authentication; entity authentication, and data authentication. Entity authentication allows the receiver to verify if the message is sent by the claimed sender or not. Therefore, by applying authentication in

the WSNs, an adversary will not be able to participate and inject data into the network unless it has valid authentication keys. On the other hand, data authentication guarantees that the reported data is the same as the original one. In a secure data aggregation, both entity and data authentication are important since entity authentication ensures that some exchanged data between sensors. For example, electing an aggregator point or reporting invalid aggregated results are authenticated using their identity while data authentication ensures that raw data are received at the aggregators at the same time as they are being sensed.

- Non-repudiation: ensures that a transferred packet has been sent and received by the person claiming to have sent and received the 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.

- Data Accuracy: One major outcome of 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.

5.2 Attacks On Routing Protocol

WSNs are vulnerable to different types of attacks due to the nature of the transmission medium (broadcast), remote and hostile deployment location, and the lack of physical security in each node. Following are the different kinds of attacks that might affect the aggregation in the WSN.

- Denial of Service Attack(DoS): is a standard attack on the WSN by transmitting radio signals that interfere with the radio frequencies used by the WSN and is sometimes called jamming. As the adversary capability increases, it can affect larger portions of the network. In the aggregation context, an example of the DoS can be an aggregator that refuses to aggregate and prevents data from traveling into the higher levels.

- Node Compromise: is where the adversary is able to reach any deployed sensor and extract the information stored on it which is some times 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.

- Sybil Attack: is where the attacker is able to resent more than one identity within the network. It affects aggregation schemes in different ways. Firstly, an adversary may create multiple identities to generate additional votes in the aggregator election phase and select a malicious node to be the aggregator. Secondly, the aggregated result may be affected if the adversary is able to generate multiple entries with different readings. Thirdly, some schemes use witnesses to validate the aggregated data and the data is only valid if n out of m witnesses agreed on the aggregation results. However, an adversary can launch a Sybil attack and generate n or more witness identities to make the base station accept the aggregation results.

- Selective Forwarding Attack: With no consideration about security, it is assumed in the WSN that each node will accurately forward received messages. However, a compromised node may refuse to do so. It is up to the adversary that is controlling 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 this case 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: The adversary aims to 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 significantly biased or fictitious values, and perform a Sybil attack to affect the aggregation result.

5.3 Summary

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.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. Due to these limitations, security issues such as data integrity, confidentiality, and freshness in data aggregation become crucial when the WSN is deployed in a remote or hostile environment.

Chapter 6

Implementation of security in NS-2

Implementation of security on NS-2 is necessary in network simulation. However, currently, NS-2 does not support these features. Our project will aim to solve this issue.The purpose of the project is to find a way to add encryption/decryption features into network simulation program NS-2. For the purposes of this analysis, the assumption is that the key is pre-shared before deployment of the node and the encryption/decryption algorithm is illustrative.

Our approach is to build a new protocol at network layer - IP layer. We also define new packet format to represent new protocols. The new protocol is represented by a class derived from built-in class in NS-2. Within new derived class we will add encryption and decryption for data field in the data packet. We will also implement message digest generation function to ensure the integrity of data packet during transmission. We consider our data as plain text. The cryptography algorithm are AES,DES and RSA.The programming language is C++ and NS-2 version is 2.27.

6.1 Implementation process

Marc Greis's tutorial shows how to add a new packet protocol to NS-2 (As shown in Figure 7.1).

The new packet class is created in the folder ../apps. After that, the new packet name has

Figure 6.1: Flow chart of adding new protocol to NS-2

to register to the packet.h. Of course, the makefile has to be modified so that the new class is complied. At the TCL layer, the new packet must be declared by adding the name and default packet size value to the ns-default.tcl file. Finally, we have to make an entry for the new packet in the ns-packet.tcl file. After recompile the ns-2, we can use the new packet for our simulation.

We propose to build a new packet class carrying data. The methods of the class include some algorithms of encryption and decryption as well as generate messages digest functions for integrity. AES,DES and RSA are selected to demo encrypt and decrypt algorithms. The message digest generator is the hash function in C++.

In the next section, we presented the implementation of AES,DES and RSA algorithms for different number of nodes and compared the parameter like execution time,throughput and consumed energy to find the suitable method for WSN.

DES Algorithm DES is a block cipher, with a 64-bit block size and a 56-bit key. DES consists of a16-round series of substitution and permutation. In each round, data and key bits are shifted, permutated, XORed, and sent through,8 s-boxes, a set of lookup tables that are essential to the DES algorithm. Decryption is essentially the same process, performed in reverse [13].

AES Algorithm AES uses 10, 12, or 14 rounds. The key size that can be 128,192 or 256 bits depends on the number of rounds. AES uses several rounds in which each round is made of several stages. To provide security AES uses types of transformation. Substitution permutation, mixing and key adding each round of AES except the last uses the four transformations [14].

RSA Algorithm RSA is a commonly adopted public key cryptography algorithm [15]. The

first, and still most commonly used asymmetric algorithm RSA is named for the three mathematicians who developed it, Rivest, Shamir, and Adleman. RSA today is used in hundreds of software products and can be used for key exchange, digital signatures, or encryption of small blocks of data. RSA uses a variable size encryption block and a variable size key. The keypair is derived from a very large number, n, that is the product of two prime numbers chosen according to special rules. Since it was introduced in 1977, RSA has been widely used for establishing secure communication channels and for authentication the identity of service provider over insecure communication medium. In the authentication scheme, the server implements public key authentication with client by signing a unique message from the client with its private key, thus creating what is called a digital signature. The signature is then returned to the client, which verifies it using the server's known public key [16].

6.2 Experimental Results and Analysis

Experimental result for Encryption algorithm AES, DES and RSA are shown in following figures, which shows the comparison of three algorithm AES, DES and RSA for different number of nodes and compared the following parameter:

- Execution time
- Throughput
- Consumed Energy

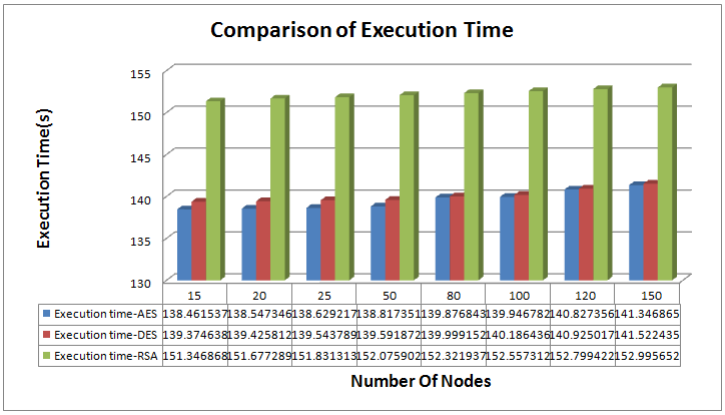
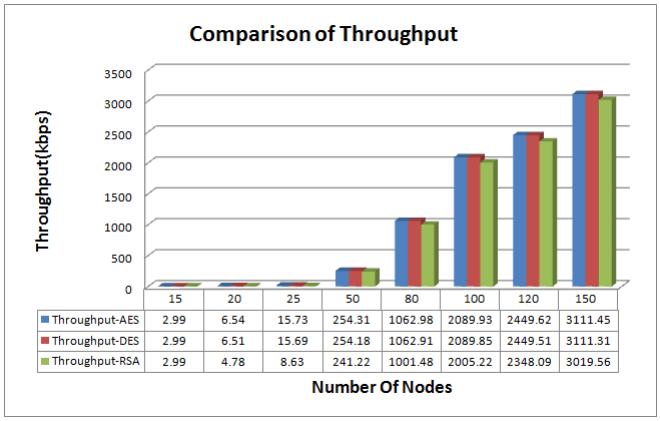
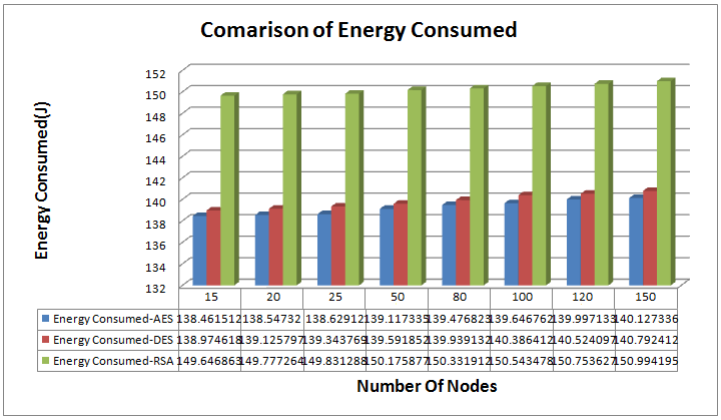


Figure 6.2: Comparison of Execution Time



	15	20	25	50	80	100	120	150
Throughput-AES	2.99	6.54	15.73	254.31	1062.98	2089.93	2449.62	3111.45
Throughput-DES	2.99	6.51	15.69	254.18	1062.91	2089.85	2449.51	3111.31
Throughput-RSA	2.99	4.78	8.63	241.22	1001.48	2005.22	2348.09	3019.56

Figure 6.3: Comparison of Throughput



	15	20	25	50	80	100	120	150
Energy Consumed-AES	138.461512	138.54732	138.62912	139.117335	139.476823	139.646762	139.997133	140.127336
Energy Consumed-DES	138.974618	139.125797	139.343769	139.591852	139.939132	140.386412	140.524097	140.792412
Energy Consumed-RSA	149.646863	149.777264	149.831288	150.175877	150.331912	150.543478	150.753627	150.994195

Figure 6.4: Comparison of Energy Consumed

6.3 Summary

Method	**Approach**	**Execution Time(s)**	**Throughput (bits/s)**	**Consumed energy(J)**	**Security**
AES	Symmetric	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.

Based on Experimental result of Encryption algorithm AES, DES and RSA, AES is selected to encrypt and decrypt algorithms as it consumes less energy and it also takes less time to execute.

Chapter 7

Integration of LEACH with AES and RSA

We simulated LEACH with security algorithms like AES and RSA using MIT's NS2 extension for LEACH by varying the cluster numbers and evaluated the performance of the network in terms execution time of the sensor network and total energy consumption by the network. For the experiment, number of cluster heads were changed as 3,4,5,6 and 8 for different readings. Following figure shows the simulation graphs for various cluster head numbers verses execution time and average energy consumption respectively.

Despite the significant overall energy savings, however, the various assumptions made by LEACH protocol raise a number of issues:

- LEACH assumes that all nodes begin with the same amount of energy and that the amount of energy a CH consumes is more than that of a non cluster node. It also assumes that the amount of energy consumed by cluster heads in every cluster round is constant. This assumption is however not realistic.

- 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.

- LEACH requires that all nodes are continuously listening. This is not realistic in a random distribution of the sensor nodes, for example, where cluster-heads would be

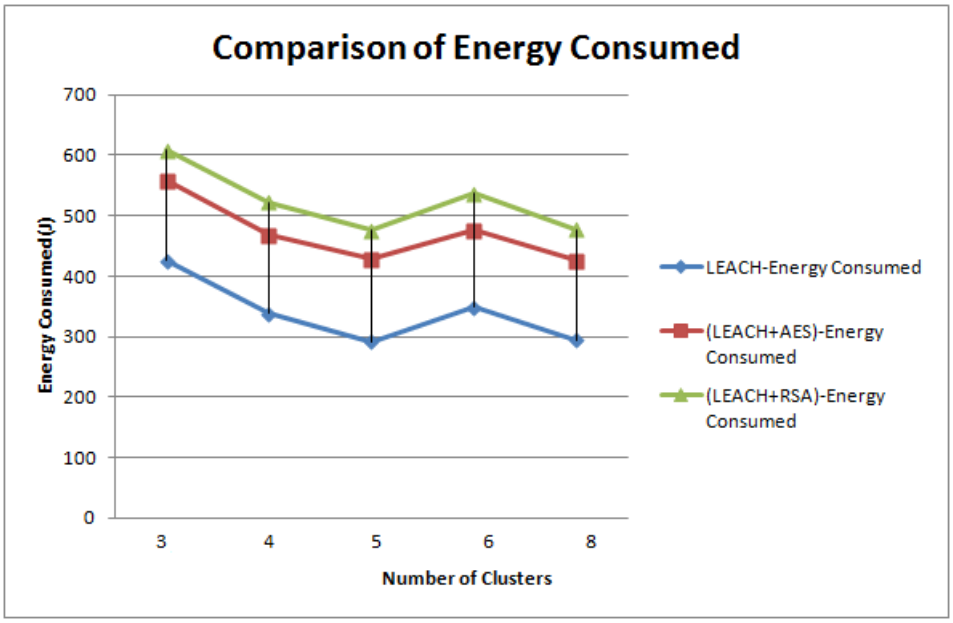


Figure 7.1: Comparison of Energy Consumed

located at the edge of the network.

- LEACH assumes that all nodes have data to send and so assign a time slot for a node even though some nodes might not have data to transmit.
- LEACH assumes that all nearby nodes have correlated data which is not always true.
- Finally, there is no mechanism to ensure that the elected cluster-heads will be uniformly distributed over the network. Hence, there is the possibility that all cluster-heads will be concentrated in one part of the network.

LEACH doesn't consider data rate generation from a region while selecting cluster-heads. Hence non CH nodes that belong to the regions that have a higher data rate generation, which are expected to transmit frequently, dissipate more energy in transmitting data to a remote CH located far. This leads to uneven energy dissipation over the network thereby reducing the network lifetime. Secondly, LEACH assumes that every time a node becomes a CH, it dissipates an equal amount of energy. This is incorrect, as cluster-heads located far from the base station spend more energy in transmitting data those located near to the base station. To ensure an even energy load distribution over the whole network, additional parameters including the residual energy level of candidates relative to the network and their

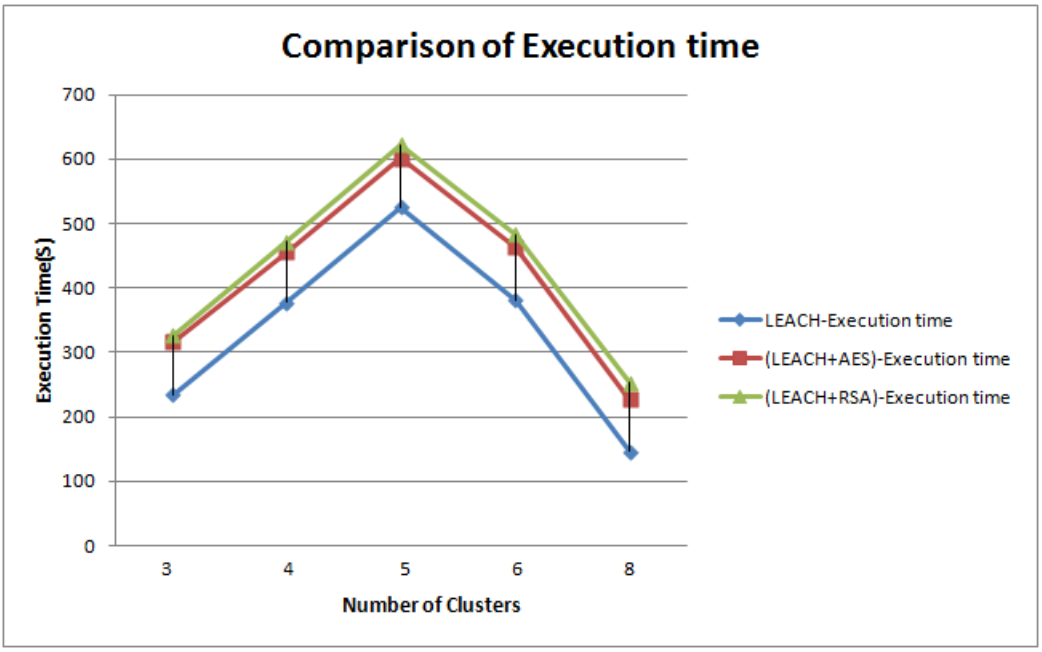


Figure 7.2: Comparison of Execution Time

data rate generation value should be considered to optimize the process of CH selection. Another improvement over LEACH is to prevent energy depletion by selecting one or more nodes are chosen as CH backup node(s) which substitute the CH in some failure/energy depletion of current CH to avoid a complete cluster set-up phase. The "number-of-hops" metric can be introduced, which indicates how far the CH is from the sensing node. This allows nodes to:

- Select the nearest CH node, which saves energy and reduces messaging needed to bridge the distance between the CH and the sensor node.
- Allows a node to learn the shortest path to the selected CH.

Some other proposals are:

- The sensor nodes can send the information regarding their energy level and location to the base station (BS) so that BS can later on decide which nodes will be elected as cluster heads. This is better as only the nodes with the required energy level and location are needed to be considered for cluster head formation.
- The clusters need not be formed every time and the clusters with ample energy be

retained. This eliminates another overhead.

7.1 Summary

Based on simulation results, AES algorithm is selected to provide security with LEACH protocol. This satisfies the title of my project work which is "Secured Data Aggregation in WSN".

Chapter 8

Conclusion and Future Scope

8.1 Conclusion

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 result of Encryption algorithm AES, DES and RSA, AES is selected to encrypt and decrypt algorithms. RSA is an asymmetric cryptography scheme which involves computationally intensive mathematical functions and WSN are highly resource constraint with having limited energy lifetime, small memory and limited computation and communication capabilities. Based on simulation results, AES algorithm is selected to provide security with LEACH protocol. This satisfies the title of my project work which is "Secured Data Aggregation in WSN".

8.2 Future Scope

- There are certain optimizations possible in the LEACH protocol as discussed earlier which may be integrated in the basic LEACH algorithm.

- Security is moderate in AES algorithm , so we can make it more secure by adding Key Distribution Center(KDC) at each data aggregator node.

References

Web References

[13] http://www.isi.edu/nsnam/ns/tutorial/

[14] http://s.pudn.com/

[15] http://www.isi.edu/nsnam/ns/tutorial/nsscript5.html

[16] http://www.evanjones.ca/ns2/

[17] http://mohit.ueuo.com/AWK_Scripts.html

[18] http://en.wikibooks.org/wiki/LaTeX/Mathematics

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] K. Dasgupta, K. Kalpakis, and P. Namjoshi, "An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks", IEEE 2003.

[3] K. Dasgupta et al., "Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks", In Proc. of IEEE Networks'02 Conference, 2002.

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

[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] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, Energy Efficient communication Protocols for Wireless Sensor Networks In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (HICSS), 30053014, Big Island, Hawaii, USA, January 2000.

[7] L. B. Oliveira E. Habib H. C. Wong A. C. Ferreira, M. A. Vilaa and A. A. Loureiro. Security of cluster-based communication protocols for wireless sensor networks. In 4th IEEE International Conference on Networking (ICN05), volume Lecture Notes in Computer Science, pages 449,Washington, DC, USA, 2005.

[8] Jamal N. Al-karaki and Ahmed E. Kamal. Routing techniques in wireless sensor networks: A survey. IEEE Wireless Communications, 11:6-28,2004.

[9] C. Shen, C. Srisathapornphat, and C. Jaikaeo, "Sensor information networking architecture and applications," IEEE Personnel Communications,Aug. 2001, pp.52-59.

[10] Haowen Chan Adrian Perrig Dawn Song , "Random Key Predistribution Schemes for Sensor Networks" Carnegie Mellon University.

[11] Wendi Heinzelman, Anantha Chandrakasan and Hari Balakrishnan,"The MIT uAMPS ns Code Extensions", Massachusetts Institute of Technology Cambridge,MA02139, A technical report, Jane 2000.

[12] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey", IEEE Wireless communication 2007.

[13] Erik Olson, Woojin Yu, "Encryption for Mobile Computing"

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

[15] R.Rivest, A. Shamir, L.Adleman. "A method for obtaining digital signatures and public-key cryptosystems"z. Communications of the ACM, Feb 1978.

[16] Andrea Pellegrini, Valeria Bertacco, Todd Austin on topic Fault-Based attack of RSA Authentication.

Index

Index