

# Energy Efficient and Fault Tolerant Data Aggregation in Wireless Sensor Networks

**Renu Bala**

*M. Tech. Scholar*

*Department of Computer Science and Engineering  
Sobhasaria Engineering College, Sikar (Rajasthan)*

**Piyush Mathur**

*Assistant Professor*

*Department of Computer Science and Engineering  
Sobhasaria Engineering College, Sikar (Rajasthan)*

## Abstract

Wireless Sensor Networks (WSN) consist of numerous autonomous sensor nodes devices are spatially distributed to sense and monitor various changes of the environment surrounding us. Such devices are also capable to communicate in wireless sensor networks and that can also sense, monitor, transmit, receive or process numerous data like pressure, temperature, sound, motion, humidity etc. Sensor networks are collection of sensor nodes which co-operatively send sensed data to base station. As sensor nodes are battery driven, an efficient utilization of power is essential in order to use networks for long duration hence it is needed to reduce data traffic inside sensor networks, reduce amount of data that need to send to base station. The main goal of data aggregation algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. Power consumption in case of processing of data is less as compared to the transmission of data. Therefore, it is preferable to do in-network-processing of data and reduce packet size. One approach of data aggregation after data gathering is to use distributed system architecture. Wireless sensor networks have limited computational power and limited memory and battery power which leads to increased complexity of applications developed for them. Moreover, this often results in applications that are closely coupled with network protocols. Clustering is a key technique used to extend the lifetime of a sensor network by reducing energy consumption. In this paper, a fault tolerant and energy efficient data aggregation technique on wireless sensor networks is proposed. This work augments the data aggregation scheduling using Connected Dominating Set (CDS) by inculcating energy efficient techniques to enhance the network lifetime through preventing uneven consumption of battery power in the wireless nodes. The simulation is carried out using MATLAB and results produced are in agreement with those of the analytical model.

**Keywords:** Wireless Sensor Network, Data Aggregation Scheduling, Connected Dominating Set

## I. INTRODUCTION

### A. Introduction

With the popularity of laptops, cell phones, GPS devices, PDAs, RFID, and intelligent electronics in the post-PC era, computing devices have become more mobile, cheaper, more distributed, and more pervasive in everyday life. It is now likely to construct, from commercial off-the-shelf (COTS) components, a small size embedded system with the comparable capability of a 90's PC [1]. These embedded systems can be supported with scaled down Windows or Linux operating systems. From this viewpoint, the emergence of wireless sensor networks (WSNs) is fundamentally the latest trend of Moore's Law [1] toward the miniaturization and ubiquity of computing devices. Naturally, a wireless sensor node (or simply sensor node) consists of computing, sensing, actuation, communication, and power components. These components are incorporated on a single or multiple boards, and packaged in a few cubic inches. With state-of-the-art, low-power circuit and networking technologies, a sensor node powered by 2 AA batteries can retain for up to three years with a 1% low duty cycle working mode [2]. A WSN generally consists of tens to thousands of such nodes that communicate through wireless channels for cooperative processing and information sharing. WSNs can be deployed on a universal scale for habitat study and environmental monitoring, over a battle field for military observation and reconnaissance, in developing environments for rescue and search, in factories for condition based maintenance, in buildings for infrastructure health monitoring, in homes to comprehend smart homes, or even in bodies for patient monitoring [3]. After the original deployment (typically ad hoc), sensor nodes are accountable for self-organizing a suitable network infrastructure, often with multi-hop connections between sensor nodes [4]. The onboard sensors then start collecting seismic, acoustic, infrared or magnetic information concerning the environment, by means of either continuous or event driven working modes. Positioning and location information can also be acquired through the global positioning system (GPS) or local positioning algorithms [5]. This information can be collected from across the network and suitably processed to build a global view of the monitoring phenomena or objects. The basic viewpoint behind WSNs is that, while the capacity of each individual sensor node is limited, the aggregate power of the whole network is enough for the required mission.

### B. Applications of WSN

WSNs can be employed in a wide range of applications, ranging from military purposes to habitat monitoring, to fault detection and toxic sensing [3]. Various applications have been proposed and executed using WSN systems in different research for

instance smart dust, industrial applications in manufacturing and supporting tools, and electronic shepherd that supervise flocks' behaviors. Since WSN systems are unattended by human, and they work without maintenance even in most severe conditions, they are used in heavy risk applications such as biomedical applications, radiation and nuclear related systems, habitat monitoring, structural health monitoring and seismic monitoring. Applications such as ecological, habitat, environmental, and soil monitoring; disaster relief management; forecast systems, building, bridges and structural monitoring; smart energy and home application, health care monitoring; and military applications are some examples of applications that can use WSN. WSN systems are also employed in monitoring a broad range of ambient conditions such as: humidity, temperature, lightning condition, movement, pressure, soil makeup, noise levels, the presence and absence of objects, mechanical stress, and current characteristics consisting of direction, speed, and size. Applications such as tracking small animals and birds, chemical and biological detection, monitoring environmental conditions, soil and air pollution detection, and study of environmental phenomena are other examples of environmental application. These types of systems also have been projected for many environmental, inhabitation, and related area systems such as monitoring city water resources, soil monitoring, traffic monitoring and controlling, wild life surveillance and many other systems. One of the most significant applications of WSNs is military application. The fault tolerance, rapid deployment, self-organizing, and maintenance free characteristics of sensor nodes make WSN a superior nomination for unkind and unsupervised conditions like military cases and battle fields. There are many WSN applications in different aspects of human life. The range of the applications is very broad, and as the technology progress and become cheaper, more functional, and better sensors are developed, WSNs are used more in different fields. In addition, Many practical applications and implemented on sensors such as Zigbee [6], WirelessHART [7], and WPAN [8], from different vendors such as Crossbow [9] and Dust Network [10] have been implemented for commercial purposes as well.

### C. Need for Clustering and Data Aggregation in Large WSN

Two approaches for managing power can be considered in WSNs while dealing with data transmission from a Sensor Node to the base station:

- 1) An end-to-end routing protocol which is used in MANETs and clustering using data aggregation. The end-to-end, which is also known as address centric paradigm, is the simplest approach that involves long range data transmission, which is not feasible in vast monitoring area, or it requires a considerable amount of energy consumption.
- 2) Data aggregation, in which data from different sources are combined and then routed, can be considered as a good scheme since it minimizes the number of transmissions and eliminate the common redundancy.

In Address Centric (AC) protocols [11], each node sends its sensed data via the shortest path to the system's base. Although there are many different algorithms for the minimum path in such a graph like system, but there are no advantages for saving a large amount of energy. Clustering is a promising mechanism for improving the power management of WSNs, unless the sensor nodes are deployed sparsely and the distances between nodes are lengthy. Clustering SNs will result both data aggregation and less data transmission with shorter length. Sensed data of SNs within a cluster can be aggregated, and then transmitted to the base station in single-hop or multi-hop manner. Thus, the volume of data to be transmitted is reduced significantly. In addition, SNs transmit their sensed data inside clusters rather than sending it directly to the base station. Clusters are usually formed in a way that SNs access to the cluster head in at most 2-hop manner, which is usually more energy effective than multi-hop data transmission. Data centric protocols, in which data aggregation is a key concept, on the other hand, reduces the quantity of data as well as the number of data transmission. When a node sends data toward the base station, middle nodes (hops) apply an aggregation and consolidation function on data, and send the aggregated with no redundancy to the next node or to the base. This scheme results fewer transmission and thus less load on nodes, which lead to energy saving and increasing the system's lifespan. Different aggregation functions such as max, min, sum, count, mean, or similar functions are used to handle the aggregation of data from different SNs. The aforementioned functions are similar to their corresponding mathematical functions. For instance max aggregation function considers the maximum value of all received data, and sends that as the aggregated value. In addition, there are more complex algorithms which are used for data aggregation. Optimal aggregation which uses minimum Steiner tree [12] and is a NP-hard algorithm and suboptimal including Centre at Nearest Sources (CNS) [13], Shortest Path Tree (SPT), and Greedy Incremental Tree (GIT) heuristics are some example of data aggregation algorithms. Other aggregation algorithms such as Secure Hop-by-Hop Data Aggregation Protocol for Sensor Networks (SDAP) [14] are proposed to take the advantage of secure data aggregation. SDAP is based on a tree topology. Based on the principle of divide-and conquer and commit-and-attest, and use of a probabilistic grouping technique, break the sensor filed into a tree and the similar sub-tree in a dynamical partitioning way. The aim of this approach is to reduce the communication overhead, increase efficiency, and improve the generality by employing a secure data aggregation, eliminating false value, and using various aggregation functions. The strength of this approach is the using of the security issue in data aggregation Also, secure data aggregation approaches have been done on WSNs. Evans and Hu et al. [15] have proposed a secure data aggregation scheme based on hop-by-hop method where just leaf nodes in tree topology sense data while the intermediate nodes won't read the sensed data. Furthermore, Du et al. [16] have done similar studies and proposed a mechanism which let the sink authenticate the aggregated data from several designated aggregators by the endorsement of some witness sensors around the aggregators. Another technique, called Secure Information Aggregation (SIA), is proposed by Przydatek et al. [17] for secure aggregation in WSNs. In SIA, the base station collects the raw data from all sensor nodes, and compute an aggregation based on Merkle hash tree. Other work has been done in the area, which

has proposed aggregation both in secure and unsecure methodology, but we will not look into those deeply since the data security and privacy is out of scope for this work.

#### **D. Motivation**

The wireless sensor network (WSN) has started receiving huge research incentives for its omnipresence in several applications, including environmental monitoring, wildlife exploration, medical supervision and battlefield surveillance. The sensor network is formed with small electronic devices possessing self-configuring capability that are either randomly deployed or manually positioned in huge bulk. It performs activities in several dimensions, for instance identifying the neighborhood, presence of targets or monitoring environmental factors (motion, temperature, humidity, sound and other physical variables). However, owing to limited battery power, the sensor networks demand energy efficient resolutions to enhance the performance of sensor network. Energy consumption problem, being the most visible challenge, is considered central to the sensor research theme. The processing of data, memory accesses and input/output operations, all consume sensor energy. However, the major power drain occurs due to wireless communication. Therefore, attempts require to be carried out to perform as much in-network processing as possible within a sensor or a group of sensors, known as clusters.

This paper is organized as follows:

Section 1 presents an overview of the subject matter and gives the problem statement and the approach for the research. Section 2 provides the detailed overview of various research papers that contributed to the subject. It also makes a background for the concepts and techniques presented in subsequent Sections. Section 3 presents the Dynamic Clustering and Aggregation Policy for Real Time Wireless Sensor Networks. Section 4 gives the simulation results and the plots for various features for similarity measurement. Section 5 concludes the paper.

## **II. PROBLEM STATEMENT AND RESEARCH APPROACH**

In wireless sensor network, data fusion is considered as an essential process for preserving sensor energy. Periodic data sampling leads to enormous collection of raw facts, the transmission of which would rapidly deplete the sensor power. Thus, aggregators are required to aggregate a collection of samples which can be further transmitted to the sink. This aggregator can be implemented in an existing wireless sensor node or can be of a distinct type other than the nodes. The requirement of aggregator is critical for increasing the network lifetime as the battery life of individual nodes can be depleted in sending large amount of raw data without any aggregation.

However, data aggregation comes with the cost of latency in the network and thus, imposes a constraint when deploying the network in a setup with real time processing requirements. Moreover, the data aggregation is a complex procedure as some of the nodes are prone to failures and thus do not contribute to the transmission of data sensed by them. In such a case, efficient techniques are needed for aggregation of data considering the node failure possibility. In this paper, a Connected Dominating Set (CDS) based Data Aggregation Scheduling is proposed so as to ensure an even consumption of battery in almost all the nodes of the network and thereby leading to increased battery life of the nodes and thus increased network lifetime. The analytical model being developed is based on 250kbps, 2.4 GHz networks with AtMega128 microcontrollers with TinyOS. The simulation is carried out in MATLAB and the results are deduced for varying network traffic models and network branching factors.

The network consists of  $n$  sensor nodes and one base station called sink. Each sensor node can send or receive data to or from all directions. The transmission/reception area of each sensor node is roughly a disk centered at the node. It is assumed that all nodes have the same transmission range for simplicity. Let  $G(V,E)$  be the topology graph of a network.  $V$  is the set of all the nodes in the network and  $E$  is the edge set of  $G$ . An edge  $(u, v)$  is in  $E$  if and only if  $u$  is in the transmission area of  $v$ . A node can either send or receive data at a time and it can receive a data packet correctly when it hears only this packet at that moment. Assuming CDS based data aggregation scheduling, the model proposed takes into consideration the residual battery life of the motes and utilize them accordingly so as to maximize the network lifetime.

## **III. PROPOSED WORK**

#### **A. Power Consumption issues in WSN**

The main attractiveness and also the main challenge of IEEE 802.15.4/ZigBee is its potentiality to set up self-organizing networks capable of adapting to diverse topologies, node connectivity and traffic conditions. Typical applications of 802.15.4/ZigBee usually consist of tens or hundreds of simple battery-powered sensor nodes (or motes) which periodically transmit their sensed data to one or several data sinks (acting as the ZigBee Coordinator or intermediate routers in the way to the Coordinator). ZigBee/802.15.4 technology was conceived to minimize the power consumption of these motes. For this purpose, the activity of the motes must be reduced up to a minimum so they can remain most of the time in a low-power (sleep) state. Therefore, a mote just has to wake-up to be active in order to sense and transmit (or receive) the data for a small fraction of time. The general objective is to maximize the lifetime of the battery in the motes and consequently the lifetime of the sensor network itself. In order to predict the battery duration of the devices in a practical implementation of a 802.15.4/ZigBee network, the

activity cycles of the motes must be categorized as well as the current which is drained from the battery during the different operations imposed by the dynamics of 802.15.4 communications, especially those which imply the activation of the radio transceiver (main source of battery consumption in the motes).

**B. First Order Radio Energy Dissipation Model**

Currently, there is a great deal of research in the area of low-energy radios. Different assumptions about the radio characteristics, including energy dissipation in the transmit and receive modes, will change the advantages of different protocols. In this work, a simple model is assumed where the radio dissipates  $E_{elec} = 50nJ/bit$  to run the transmitter or receiver circuitry and  $\epsilon_{amp} = \frac{100pJ}{bit} / m^2$  for the transmit amplifier to achieve an acceptable  $\frac{E_b}{N_o}$ . These parameters are slightly better than the current state-of-the-art in radio design. It is also assumed that an  $r^2$  energy loss due to channel transmission. Thus, to transmit a  $k$ -bit message a distance  $d$  using this radio model, the radio expends:

$$E_{Tx}(k, d) = E_{Tx-elec}(k) + E_{Tx-amp}(k, d)$$

$$E_{Tx}(k, d) = E_{elec} * k + \epsilon_{amp} * k * d^2 \quad \text{equation 3.1}$$

and to receive this message, the radio expends:

$$E_{Rx}(k) = E_{Rx-elec}(k)$$

$$E_{Rx}(k) = E_{elec} * k \quad \text{equation 3.2}$$

For these parameter values, receiving a message is not a low cost operation; the protocols should thus try to minimize not only the transmit distances but also the number of transmit and receive operations for each message. An assumption is also made that the radio channel is symmetric such that the energy required to transmit a message from node A to node B is the same as the energy required to transmit a message from node B to node A for a given SNR. For the analytical model, it is also assumed that all sensors are sensing the environment at a fixed rate and thus always have data to send to the end-user.

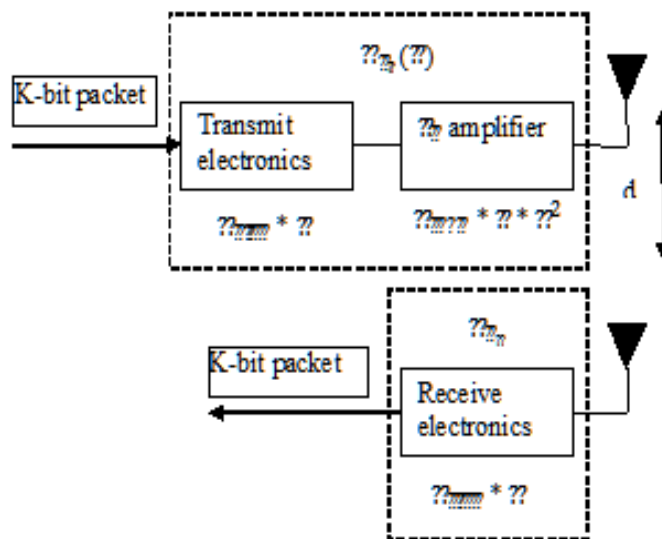


Fig. 1: Transmitter and Receiver System

**C. Data Aggregation Scheduling in WSN**

Data Aggregation is one of the most critical aspect of wireless sensor networks as the in-network data processing reduces the bulk data transmission which results in increase of network lifetime. Also, as the hardware nodes are prone to failure or depletion of battery, a fault tolerant data aggregation is essential for efficient functioning of WSN.

A fault tolerant data aggregation scheduling consists of two parts:

- 1) Basic Data Aggregation Scheduling
- 2) Amendment Strategy after a Node Fails

The amendment strategy further consists of two parts:

- 1) Aggregation Tree Repairing Strategy
- 2) Data Aggregation Rescheduling

The first step is performed so as to compute the optimal path from the sensor node to the cluster head. The latter part computes the collision free schedule so that the data is received by cluster head without packet collision occurrence as a result of multiple senders competing at the same time. The simplest and widely used clustering strategy is LEACH. However, the most efficient data aggregation scheduling strategy is Connected Dominating Set (CDS). The amendment strategy must be carefully designed with the following features:

- 1) The amendment strategy should repair the aggregation tree quickly using only the local information.

- 2) It should repair the tree under low computational cost.
- 3) It should not corrupt the properties of original aggregation tree.
- 4) It reschedule the nodes to achieve the interference free aggregation scheduling after amendment.

In this paper, a fault tolerant data aggregation scheme for wireless sensor networks is proposed. It needs local information to repair the aggregation tree and automatically reschedules nodes to achieve interference free aggregation scheduling after the amendment.

Compared to rebuilding the aggregation tree and rescheduling the entire network from scratch, the proposed strategy greatly reduces both message and time costs, which is also verified by extensive simulation results.

#### D. Fault Tolerant Data Aggregation Scheduling in WSN

Let  $G(V,E)$  denote the sensor network that consists of  $V = N+1$  sensor nodes including 1 base station, and  $E$  is the set of wireless links between the sensor nodes. For any node of the network, the set of edges connected though it connects it to all the other sensor nodes lying within its transmission range. Thus, the edges connects a node to all the nodes lying within the circle spanned by the transmission range, with centre on the node under consideration.

Suppose there are two sets of sensor nodes  $A$  and  $B$  ( $A, B \in V$ ) and  $A \cap B = \phi$ . Data are aggregated from  $A$  to  $B$  in one time-slot if all nodes in  $A$  can transmit data simultaneously in this time slot and all data are successfully received by nodes in  $B$  without interference.

Thus, the data aggregation schedule can be defined as a sequence of sender sets  $S_1, S_2, \dots, S_l$ , satisfying the following conditions, where  $l$  is referred to as the aggregation latency:

- 1)  $S_i \cap S_j = \phi$  for all  $i \neq j$ .
- 2)  $\sum S_i = V$
- 3) At time slot  $k$ , data is aggregated from  $S_k$  to  $S_l \in V$ .

Suppose that there are two nodes, called  $a$  ( $a \in A$ ) and  $b$  ( $b \in B$ ). If data is aggregated from  $a$  to  $b$  during the data aggregation process, then  $a$  is called the child of  $b$ , and  $b$  is the parent of  $a$ . A node can only have one parent at most, nevertheless, it may has multiple children. Then it can be assumed that each node maintains a list to record its children.

When a node is broken, every one of the children nodes should find a new parent. Nevertheless, the parent node of the corrupted node just needs to remove it from the children list. Thus, the children or parent nodes of some nodes may be changed after the amendment process.

For ease of description, one can classify nodes into several types according to their locations and functions during the amendment process.

- 1) Directly Affected Node (DAN) A node is called a DAN node if the corrupted node is either its parent or its child node.
- 2) Indirectly Affected Node (IAN) A node is called an IAN node if it is not a DAN, and either its parent or its children list is changed after the amendment process.
- 3) Non-Affected Node (NAN) A node is an NAN if neither its parent nor its children list is affected by the corruption or amendment process.
- 4) Milestone A milestone node is the IAN node on the farthest end of the new aggregation path of a child of the corrupted node.

#### E. Connected Dominating Set

A connected dominating set (CDS) is a subset of the nodes such that it forms a Dominating Set (DS) and all the nodes in the DS are connected. An example of CDS is shown in figure 2

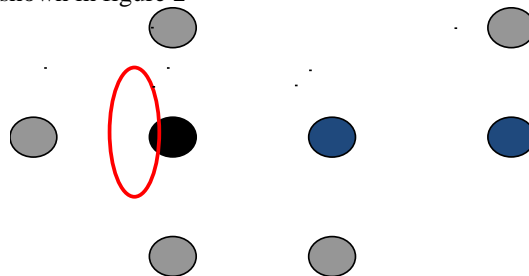


Fig. 2(a): CDS in a Graph

CDS is used as a virtual backbone in Wireless Sensor Networks.

##### 1) CDS Construction Algorithm

Given a graph  $G$  and a DS  $C$ , all nodes in  $G$  can be divided into three classes.

- Black nodes: Nodes belong to  $C$ .
- Grey nodes: Nodes are not in  $C$  but adjacent to  $C$ .
- White nodes: Nodes are neither in  $C$  nor adjacent to  $C$ .

One example of this illustration is as shown in the figure 3.2 below:

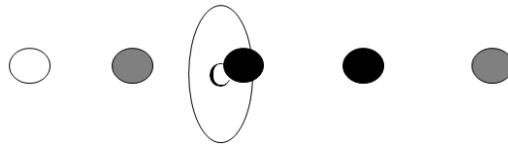


Fig. 2(b): Illustration of CDS

The problem of finding the connected dominating set is NP hard. However, approximation algorithms are there which could find the sufficiently Minimal Connected Dominating Set. One simple greedy approach to create a CDS is the Greedy Algorithm, This strategy can be explained as follows:

- Iteratively select the node with the max number of white neighbors as a dominating node. Label all such nodes as grey nodes. Maintain a list of all the white nodes connected through grey nodes. The first phase terminates when there are no white nodes.
- Color some grey nodes black to connect all the black nodes.

The step by step operation of the data aggregation algorithm can be depicted as shown in the following figures:

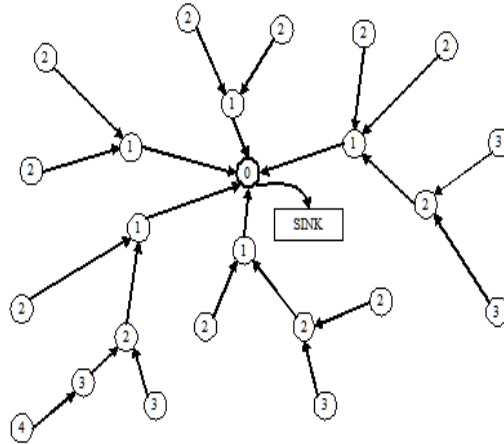


Fig. 3: Initial Spanning Tree Topology with node levels

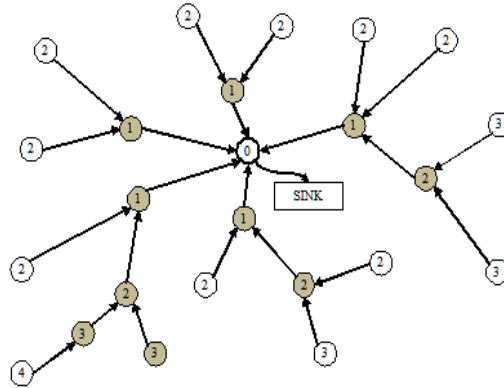


Fig. 4: Marking Grey Levels as per Greedy Algorithm

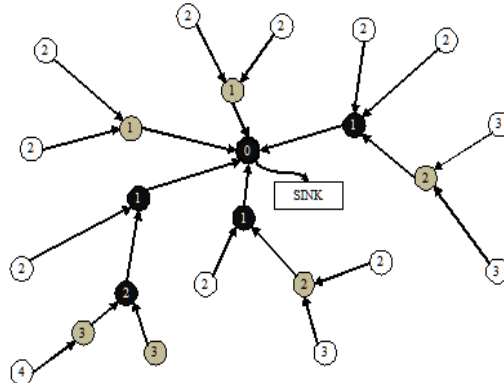


Fig. 5: CDS Construction

**F. Proposed Fault Tolerant and Energy Efficient Data Aggregation Algorithm**

As sensor network nodes are prone to failure, a fault tolerant data aggregation strategy is critical for efficient working of WSN. Moreover, a fault can occur in any of the nodes irrespective of whether it belongs to white, grey or CDS class. The proposed scheme works as follows:

The sink maintains a list of the CDS of the entire network and periodically verifies the working of the nodes belonging to the CDS through beacon signals. Each CDS node maintains the list of grey nodes transferring data to it. Each Grey node maintains a list of white nodes aggregating and transferring data to it.

If a fault occurs in a white node, then no CDS reconstruction is needed. One can simply ignore the failure and the corresponding grey node can generate alarm to notify the black node which sends the message to sink regarding the failure of the node.

If a fault occurs in a Grey node, then the entire list of white nodes attached to it needs to be transferred to some other grey or black node. This fault is alarmed by the corresponding CDS node to the sink. The CDS node then schedules the amendment strategy as follows:

Let B refers to the set of nodes of CDS and G refers to the set of Grey nodes which are in the transmission range of a white node. The selection of the node (Black or Grey) by any of the white node is made on the basis of the residual energy and other parameters of the node so as to maximize the network lifetime. Let  $E_n$  refers to the residual energy in node labeled n. The residual lifetime of any node can be computed using the first order radio electronics.

Let N be the average number of data packets send by the white node per unit time and  $n_b$  be the number of bits in the data packet. The energy consumption of a parent node in processing this data is

$$\text{Per unit time energy consumption due to one white node} = E_{Tx-elec}(N * n_b)$$

It is assumed that the packet size of the grey node is same as that of the white node. This assumption is critical for many WSN as white nodes can be associates with grey nodes as well as nodes of the CDS (black nodes) and thus, the packet formats needs to be identical for both white and grey nodes. Also, the packet size of aggregated data is same, thus the transmission energy required to be independent of the number of child nodes.

Let  $N_w$  and  $N_G$  be the number of white and grey nodes associated with a particular node. If this node is itself a grey node, then  $N_G = 0$ . Then, one can compute residual life of a node using the following formula:

$$\text{Residual life of node (RL)} = \frac{E_n}{(N_w + N_G) * E_{Tx-elec}(N * n_b)}$$

The amendment strategy proceeds in such a way so as to ensure that there residual life of all the remaining nodes can be maximized.

After the death of a grey node, each white node associate with it has been assigned a new parent. This process is done sequentially. The first white node in the list is assigned to the grey or black node in the transmission range for which the RL is maximum. This parameter is then reset and fresh computed for the next node. This process is done iteratively until all the white nodes have been assigned new parents.

If the fault occurs in a black node, then the all the Grey nodes associated to it needs to be assigned some new parent. In this case, the sinks finds out a new CDS as per the Greedy Approach and prepares a fresh aggregation schedule for efficient data aggregation.

The proposed work suggests a Fault Tolerant and Energy Efficient Data Aggregation scheme that is useful in many real life wireless sensor networks. Section 4 gives the simulation results followed by Section 5 which concludes the paper.

**IV. ANALYSIS OF PROPOSED WORK**

**A. Network Parameters**

The network simulation for modeling the scheme proposed in Section 3 is assumed to have following parameter. These parameters correspond to TinyOS on ATmega128L microcontroller:

Table - 1  
Parameter Specification Of Values Of Variables In The Analytical Analysis

Parameter Symbol and Description.	Value
Data Rate, Frequency Band	250kbps, 2.4 GHz
Transmission Range	10m/20m
N; Number of Nodes in the Network, considering all as the active nodes	1000
R; Bits per Packet	114*8 (TinyOS Specification)
$e_{t1}$ ; energy dissipation per bit in transmitter circuitry before transmission	50nJ/bit
$e_{d1}$ ; energy dissipation per bit in sending a bit over d distance.	10pJ/bit/m <sup>2</sup>
$e_{r1}$ ; energy dissipation per bit in receiver circuitry.	50nJ/bit
J; Energy Consumption of ATmega128L per CPU cycle. Being 8 bit microcontroller, it processes 8 bits per unit time.	100nJ/bit

The following results are obtained corresponding to the set of values as specified in table 1

Figure 6 depicts the average life of a wireless sensor node, powered on a battery of average 1.5 volts, as a function of mean number of packets transmitted per unit time. The packet size specification is that of Tiny OS which is one of the most common Operating System for wireless motes. It is evident that more is the number of packets transmitted, the less is the average life of the node.

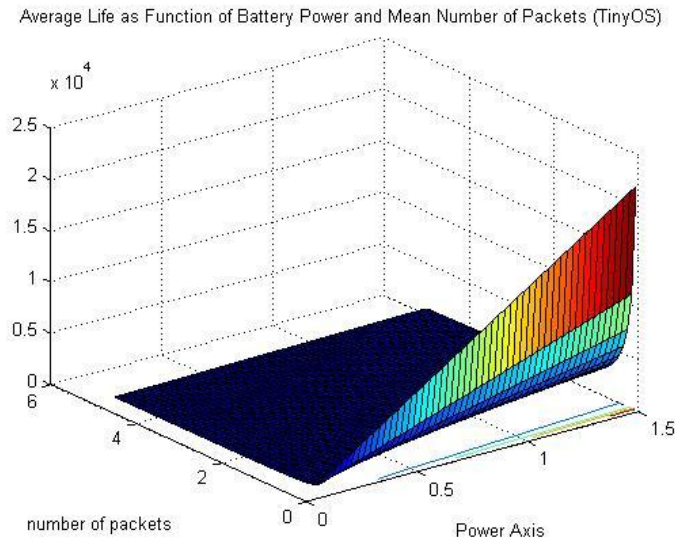


Fig. 6: Average Node Life as a function of Initial Power and Mean Number of Packets Transmitted

Figure 7 depicts the average node life as a function of number of White and Grey nodes attached to it. The white nodes are the nodes which senses the data and the pass this data to grey or black nodes for data aggregation. Grey nodes aggregate the data and sends the data packet to black nodes which forms the Connected Dominating Set (CDS). As evident from the figure 7, as more and more nodes of white and black category gets attached to a particular node, the average life of node decreases as more power is utilized in data aggregation which leads to rapid depletion of battery power, thereby causing the device to deplete it battery life rapidly. This eventually leads to node failure as a result of battery power and thus, resulting in small network lifetime. For a maximized network lifetime, it is required that the energy consumption in all the nodes of WSN occur in evenly manner.

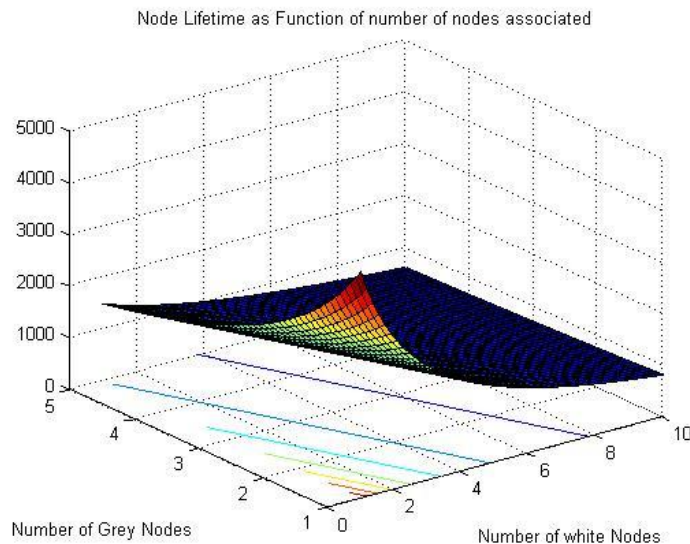


Fig. 7: Average Node Life as a function of Number of Black and White nodes attached

Figure 8 shows the average lifetime of a node as a function of mean number of nodes attached to it and the packet size. It turns out that larger packets results in much more energy depletion as compared to smaller size packets. Moreover, the energy dissipation rate of a node is directly proportional to the number of nodes attached to it. The more is the number of nodes attached to a particular node, the rapid is the energy consumption. However, in case of Tiny OS, the packet size of the node is fixed, which is equal to 114 bytes. However, for analytics purposes, the node lifetime is plotted as function of packet size. It is important to node that if the message size is large, then larger size data packet is suitable as packets of smaller size would induce

more overhead of address and other fields. At the same time, if the network is unreliable, then the use of smaller size data packets is preferred as it induces less retransmission cost if the packet is lost during transmission.

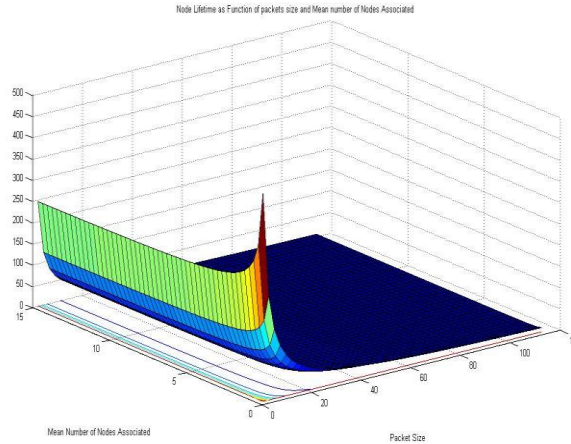


Fig. 8: Node life as a function of mean number of nodes associated and the packet size.

Figure 9 shows the plot of node lifetime as a function of transmission range and the mean number of nodes attached to a particular node. It is evident from the figure that node lifetime decreases exponentially as the transmission range is increased. The result is obvious as per the first order radio electronics model. The energy needed to transmit a bit over some distance is directly proportional to the square of the distance of transmission.

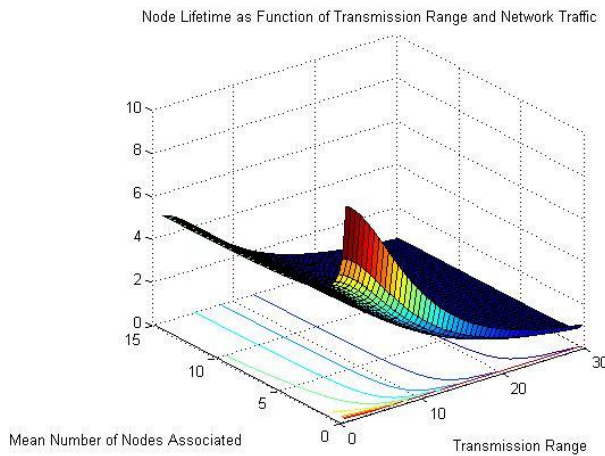


Fig. 9: Node life as a function of Transmission Range and Network Traffic.

**B. Simulation of WSN for CDS Based Energy Efficient Data Aggregation**

The simulation model for WSN is depicted as shown in Fig 10 shown below. White node corresponds to the nodes which are in the outermost circle of the sensor network.

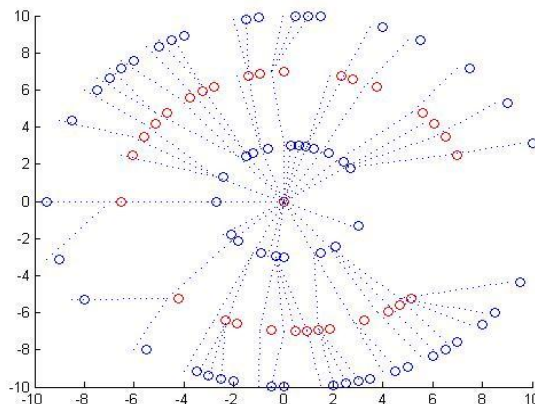


Fig. 10: Simulation of WSN CDS based Data Aggregation (The horizontal and Vertical Axis shows the X and Y coordinates of the WSN)

These nodes are attached to the grey nodes which aggregates the data from the white node and transmits the aggregated data to black nodes which corresponds to the CDS. The probable CDS for the WSN is depicted in figure 9 shown below:

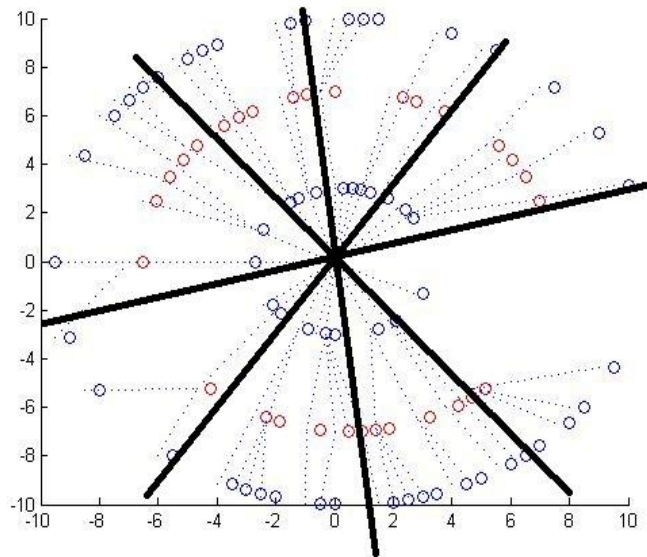


Fig. 11: CDS Construction for the WSN depicted in Figure 10. The Black lines shows the path and the nodes which could form CDS for this WSN. (The horizontal and Vertical Axis shows the X and Y coordinates of the WSN)

Figure 12 shows the energy dissipation stats of the wireless sensor node. As the nodes are prone to failure, the group of white or black nodes needs to be re-associated to some grey or black node. This reassignment is done on the basis of a nearby node in the proximity which could be assigned the nodes whose parent is failed using only local information. However, no botheration is given to the node's residual energy while eventually leads to uneven residual energy of the nodes leading to early failure of some of the nodes and consequently, reduced network lifetime.

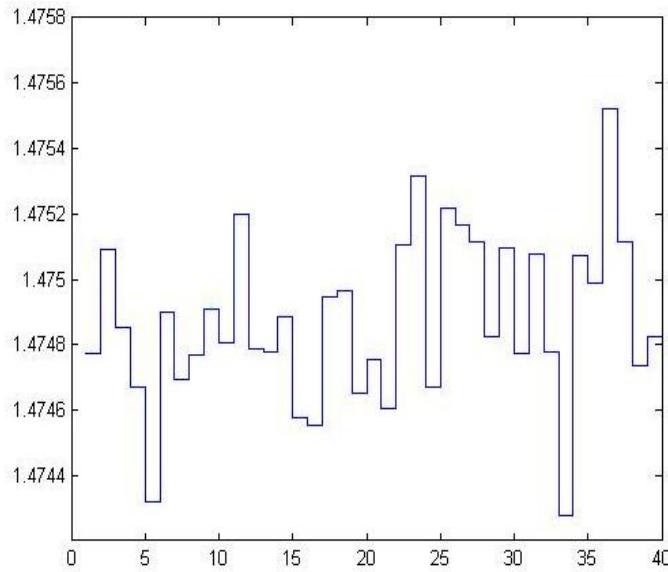


Fig. 12: Plot of Residual Battery Life of White node after 1000 rounds (Medium Traffic). Vertical Axis shows the residual battery and horizontal axis shows the node count.

Figure 13 shows the residual energy of nodes under the conditions of heavy traffic. It is evident from the figure that this leads to even more uneven distribution as compared to the condition of low network traffic. This result is obvious as under the conditions of heavy network traffic and node failure, more rapidly energy dissipation takes place of nodes which have been assigned new nodes as per the amendment strategy.

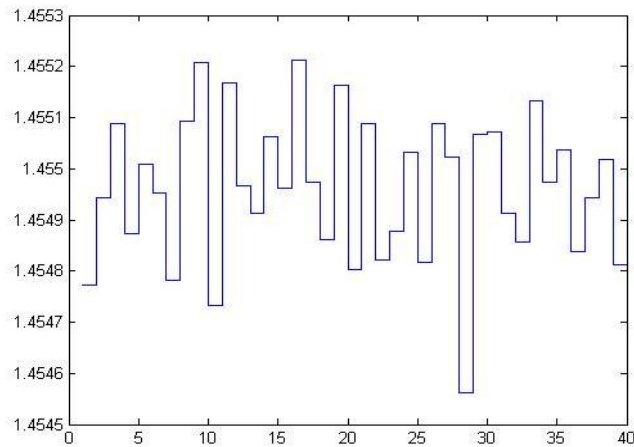


Fig. 13: Plot of Residual Battery Life of White node after 1000 rounds (Heavy Traffic). Vertical Axis shows the residual battery and horizontal axis shows the node count.

Figure 14 shows the distribution of residual battery life in conditions of low traffic. It is evident from the figure that under the conditions of low traffic, the distribution of residual battery life is still uneven but the overall curve for battery life remains above the battery life in case of conditions of high traffic. The plot shows the residual battery life after 1000 rounds. In actual practice, the number of rounds is dependent on the duty cycle of the nodes chosen as per the application requirement in such a way so as to maximize the battery life. In most of the applications, the duty cycle (wake up period/ sleep period) is chosen in such a way so at a single AA battery of 1.5 volts can operate upto 6 months.

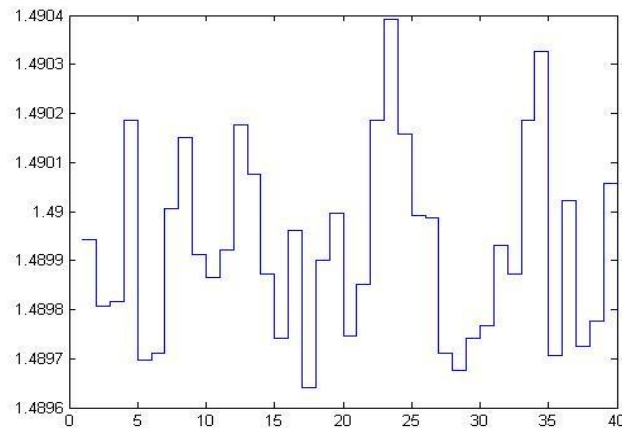


Fig. 14: Plot of Residual Battery Life of White node after 1000 rounds (Light Traffic). Vertical Axis shows the residual battery and horizontal axis shows the node count.

Figure 15 shows the plot of residual energy distribution under energy optimization scheme. The blue curve shows the residual energy distribution in CDS based data aggregation scheduling whereas the red curve shows the residual energy distribution under the energy efficient CDS based data aggregation scheme. It is clearly evident that an energy efficient CDS based data aggregation leads to a much more evenly distributed residual energy and consequently an enhanced network lifetime.

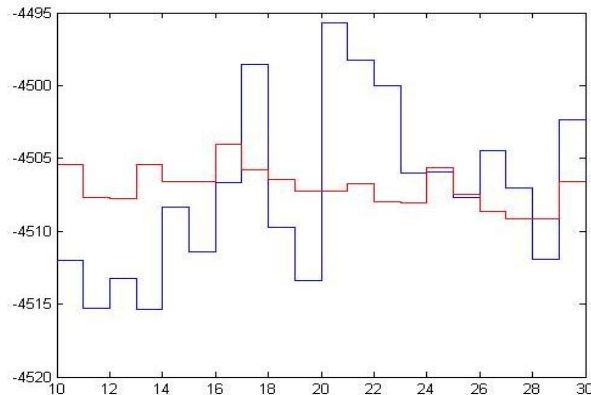


Fig. 15: Plot of Residual Battery Life of Nodes under CDS Scheme (Blue Curve) and Energy Efficient CDS Scheme (Red Curve)

**C. Comparison of CDS based Data Aggregation and CDS based Energy Efficient Data Aggregation**

Network lifetime can be measure subjectively with the distribution of the residual battery life of the nodes under the two schemes as depicted in the figure 15 depicted above.

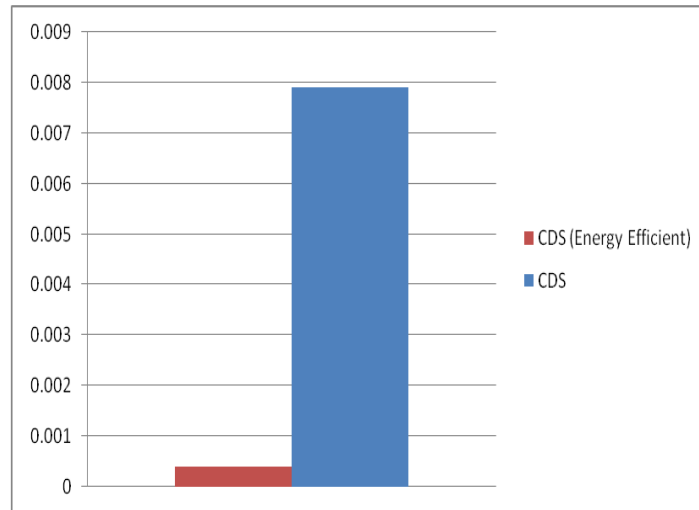


Fig. 16: Energy level difference between the weakest power and strongest power node of WSN at the end of 100 rounds.

The calculation for network lifetime can be done in a simple manner as shown:

TABLE 4.2  
ENERGY DISSIPATION USING CDS AND ENERGY OPTIMIZED CDS

<i>CDS Based Energy Efficient Data Aggregation</i>	
<i>Energy Dissipation in 1000 rounds</i>	<i>0.0004 Volts</i>
<i>Number of rounds accommodated by 1.5 volts</i>	<i>3750000</i>
<i>Assuming 1 round every 5 seconds, number of days of network before the first node fails</i>	<i>= 3750000 / (24 * 60 * 60) = 217.013889 days</i>
<i>CDS Based Energy Efficient Data Aggregation</i>	
<i>Energy Dissipation in 1000 rounds</i>	<i>0.0007 Volts</i>
<i>Number of rounds accommodated by 1.5 volts</i>	<i>2142857.14</i>
<i>Assuming 1 round every 5 seconds, number of days of network before the first node fails</i>	<i>= 2142857.14 / (24 * 60 * 60) = 124.007937 days</i>

The graphical illustration of Energy Efficient CDS based Data Aggregation is depicted in Figure 4.12 shown below:

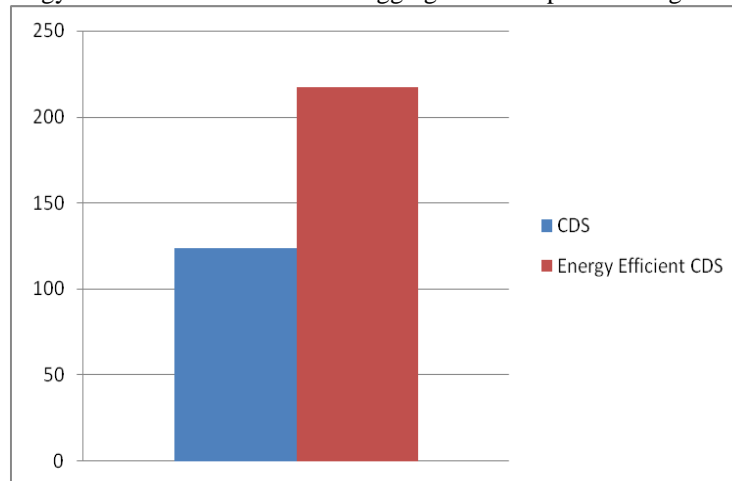


Fig. 17: Comparison of Network Lifetime with CDS based approaches of Energy Efficient Data Scheduling and Random Scheduling.

It is evident that energy efficient CDS based data aggregation scheduling gives a much more increased lifetime those random approaches. Section 5 discusses the results and concludes the paper.

**V. CONCLUSION AND FUTURE SCOPE**

The problem of fault tolerant data aggregation scheduling for wireless sensor networks is investigated in this paper. A Connected Dominating Set based Energy Efficient Data Aggregation Scheduling scheme is proposed so as to make even consumption of battery power by all the nodes in the WSN. The performance of the proposed protocol is evaluated through both analytical and simulation models. Using Analytical model and first order Radio Electronics model, the distribution of network lifetime is

investigated under various data traffic models. The same is also investigated as a function of data packet size, transmission range and node branching factor.

A model of WSN is implemented using MATLAB simulation implementing the leaf nodes (white nodes), the intermediate nodes (grey nodes) and the nodes belonging to CDS (black nodes). The Network lifetime is computed as a function of number of white and grey nodes attached to a particular node under given traffic conditions. It was proved that energy efficient CDS leads to a wise distribution of battery consumption leading to increased battery life of all the nodes involved in the data communication, thereby providing an increased network lifetime. In conclusion, the proposed strategy improves the network lifetime in the CDS based data aggregation compared to rebuilding the aggregation tree and rescheduling the entire network

#### A. Future Scope

As a future scope of this work, a stochastic model for a WSN application will be considered and packet transmission probability and residual life of nodes will be computed based on Poisson Probability Distribution. The complete model will be evaluated using Discrete Time Markov Chain (DTMC) model. This will more appropriately model the type of wireless sensor network deployment under consideration. Moreover, the computational complexity and thereby, the latency can be computed much more accurately with known probability distribution of the packet transmission and the branching factor. The computation of latency in WSN is a critical aspect for a larger class of WSN called the Real time Wireless Sensor Networks. Thus, stochastic modeling using appropriate traffic modeling can more appropriately model the Quality of Service for the WSN under consideration..

### REFERENCES

- [1] Roza, E., "Systems-on-chip: what are the limits?," *Electronics & Communication Engineering Journal* , vol.13, no.6, pp.249,255, Dec 2001. doi: 10.1049/ecej:20010602.
- [2] Rout, R.R.; Ghosh, S.K., "Enhancement of Lifetime using Duty Cycle and Network Coding in Wireless Sensor Networks," *Wireless Communications, IEEE Transactions on* , vol.12, no.2, pp.656,667, February 2013 doi: 10.1109/TWC.2012.111412.112124
- [3] Murillo, A.F.; Pena, M.; Martinez, D., "Applications of WSN in health and agriculture," *Communications Conference (COLCOM), 2012 IEEE Colombian* , vol., no., pp.1,6, 16-18 May 2012 doi: 10.1109/ColComCon.2012.6233678
- [4] Jianlin Guo; Orlik, P.; Jinyun Zhang; Ishibashi, K., "Reliable routing in large scale wireless sensor networks," *Ubiquitous and Future Networks (ICUFN), 2014 Sixth International Conf on* , vol., no., pp.99,104, 8-11 July 2014 doi: 10.1109/ICUFN.2014.6876758
- [5] Wenming Shi; Chuanhe Huang; Mingkai Shao; Dandan Song; Yong Cheng; Ying Long, "A Novel Positioning Algorithm in Wireless Sensor Networks Based on Indoor Connective Information," *Sensors Applications Symposium, 2007. SAS '07. IEEE* , vol., no., pp.1,5, 6-8 Feb. 2007 doi: 10.1109/SAS.2007.374364
- [6] Vljajic, N.; Stevanovic, D., "Performance Analysis of ZigBee-Based Wireless Sensor Networks with Path-Constrained Mobile Sink(s)," *Sensor Technologies and Applications, 2009. SENSORCOMM '09. Third International Conference on* , vol., no., pp.61,68, 18-23 June 2009 doi: 10.1109/SENSORCOMM.2009.114
- [7] Habib, G.; Haddad, N.; El Khoury, R., "Case study: WIRELESSHART vs ZIGBEE network," *Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2015 Third International Conference on* , vol., no., pp.135,138, April 29 2015-May 1 2015 doi: 10.1109/TAEECE.2015.7113614
- [8] IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirement Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)," *IEEE Std 802.15.4a-2007 (Amendment to IEEE Std 802.15.4-2006)* , vol., no., pp.1,203, 2007 doi: 10.1109/IEEESTD.2007.4299496
- [9] Hegde, N.; Manvi, S.S., "Implementation of security mechanism in Wireless Sensor Network using Crossbow nodes," *Advances in Electronics, Computers and Communications (ICAECC), 2014 International Conference on* , vol., no., pp.1,6, 10-11 Oct. 2014 doi: 10.1109/ICAECC.2014.7002422
- [10] Sharifi, M.; Kashi, S.S.; Ardakani, S.P., "LAP: A Lightweight Authentication Protocol for smart dust wireless sensor networks," *Collaborative Technologies and Systems, 2009. CTS '09. International Symposium on* , vol., no., pp.258,265, 18-22 May 2009 doi: 10.1109/CTS.2009.5067489
- [11] Srinath, R.; Reddy, A.V.; Srinivasan, R., "AC: Cluster Based Secure Routing Protocol for WSN," *Networking and Services, 2007. ICNS. Third International Conference on* , vol., no., pp.45,45, 19-25 June 2007 doi: 10.1109/ICNS.2007.26
- [12] Maharjan, N.; Shrestha, A.; Tamrakar, A.; Panday, S.P., "Solving Minimum Steiner Tree based on behavior of ants," *Internet (AH-ICI), 2012 Third Asian Himalayas International Conference on* , vol., no., pp.1,4, 23-25 Nov. 2012 doi: 10.1109/AHICI.2012.6408443
- [13] Tamilselvan, G.M.; Kiruthika, S., "An energy efficient data aggregation based medium access control protocol using centre at nearest source approach for sensor networks in a lattice topology," *Computing, Communication and Applications (ICCCA), 2012 International Conference on* , vol., no., pp.1,6, 22-24 Feb. 2012 doi: 10.1109/ICCCA.2012.6179146.
- [14] Mlaih, E.; Aly, S.A., "Secure hop-by-hop aggregation of end-to-end concealed data in wireless sensor networks," *INFOCOM Workshops 2008, IEEE* , vol., no., pp.1,6, 13-18 April 2008 doi: 10.1109/INFOCOM.2008.4544601
- [15] L. Hu and D. Evans. Secure aggregation for wireless networks. In *Workshop on Security and Assurance in Ad hoc Networks*, 2003
- [16] Du, W., J. Deng, Y. S. Han, S.Chen and P.K. Varshney, 2004. A Key management scheme for wireless sensor networks using deployment knowledge. *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies*, Mar. 7-11, IEEE Xplor Press.
- [17] DOI: 10.1109/INFCOM.2004.1354530
- [18] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure information aggregation in sensor networks," *ACM SenSys 2003*, Nov. 2003.
- [19] Martirosyan, A., Boukerche, A., and Pazzi, R. W. N. "Energy Aware and Cluster-Based Routing Protocols for Large-Scale Ambient Sensor Networks," In *Proceedings of the First International Conference on Ambient Media and Systems*, Quebec, Canada, 2008