CHAPTER 4

SELF-CONFIGURATION IN
WIRELESS SENSOR
NETWORKS
CHAPTER 4

SELF-CONFIGURATION IN WIRELESS SENSOR NETWORKS

4.1 Overview of Self-Configuration

A Wireless Sensor Networks once and for all loses possibility of maintenance after its deployment such as node’s battery recharge. Nodes in sensor network which are equipped with battery bounds a life span which lasts from the point of deployment till its battery survives. Thus, from the point of its deployment, battery reserve defines the lifetime of nodes and network and the battery resource becomes a valuable resource because the battery cannot be replenished.

Network lifetime can be defined as the time interval, which the network is capable of performing its intended tasks. In other words, the network lifetime often indicates the time elapsed until the first node drains its battery which is responsible to die down the network. Improving the network lifetime is a challenging issue for system design in sensor networks that can conserve energy resource. Network lifetime is perhaps the significant parameter for evaluation of sensor networks. In a resource-constrained environment, it is necessary to monitor the consumption of energy. However, network lifetime of a Wireless Sensor Networks is the measurement of energy consumption that occupies the exceptional position for formation of upper bound in utility of sensor network. The network easily fulfill the purpose of operation till it is alive and not after that. Hence an indicator for the maximum utility can provide. If the metric is used in an analysis preceding a real-life deployment, the estimated network lifetime can also contribute to justifying the cost of the deployment.

Lifetime is also considered a fundamental parameter in the context of availability and security in networks [Khanand Misic 2008].

Network lifetime strongly depends on the lifetimes of the single nodes that constitute the network. This fact does not depend on how the network lifetime is defined. Each definition can finally be reduced to the question when the individual nodes fail. Thus, if the lifetimes of single nodes are not predicted accurately, it is possible that the derived network lifetime
metric deviates in an uncontrollable manner. It should therefore be clear that an accurate and consistent modeling of the single nodes is very important.

Hierarchical topology control in sensor network using clustering has been accepted for energy conservation in several applications such as data gathering. The clustering maps to the dominating set problem. Finding a small cardinality dominating set maps, that minimizes the dominating set problem. The nodes in a dominating set deals with higher computational and communication loads than other nodes in network resulting to faster energy depletion of dominating set. Often this poses a problem of maximizing the lifetime of dominating set to improve the network lifetime in sensor network.

Quality of Service (QoS) support in WSN is still remained as an open field of research from various perspectives. QoS is interpreted by different technical communities by different ways [105]. In general, QoS refers to quality as perceived by the user or application. In networking community, QoS is interpreted as a measure of service quality that the network offers to the end user or application. Figure 4.1 shows a general QoS model for network which is redrawn from [105]. In RFC 2386 [106], QoS has been defined as a set of service requirements to be fulfilled when transmitting a stream of packets from source to destination.

![Figure 4.1: A simplified QoS Model](image)
In traditional data network, QoS defines certain parameters such as packet loss, delay, jitter, bandwidth etc. However, the QoS requirements in WSN such as data accuracy, aggregation delay, coverage, fault tolerance and network lifetime etc. are application specific and they are different from the traditional end-to-end QoS requirements due to the difference in application domains and network properties. Although, some QoS solutions (like IntServ, DiffServ etc) are developed for traditional networks, these cannot be easily ported in WSN due to Related Work and Motivation 1) severe resource constraints in sensors nodes, 2) large scale and random deployment of sensors nodes and 3) application specific and data-centric communication protocols in WSN.

4.1.1 Challenges in Quality of System

The characteristics of WSN are different from other networks. Such a network requires to sense data from the surrounding environment and finally forwards the sensed data towards a remote and resourceful node called sink or base station. Therefore, QoS provisioning in WSN has some significant challenges. Some of such challenges are as follows.

- **Extreme Resource Constraint:** Some of the very significant resource constraints in WSN are energy, bandwidth, and buffer size and transmission capacity of the sensor nodes. Among these, efficient energy utilization of sensor nodes is a crucial issue as in most of the cases the batteries of the sensor nodes are not rechargeable or replaceable. Efficient bandwidth utilization is also a significant challenge in WSN. The traffics in WSN can be mixture of real time and non real time. So there should be balanced allocation of bandwidth between real time and non real time traffic.

- **Redundant Data:** Since the sensor nodes are densely deployed in a terrain of interest, therefore most of the data generated by sensor nodes are redundant. While this redundancy helps in reliability and fault tolerance of the WSN, it also causes a significant amount of energy wastage. Data aggregation or data fusion is a solution to remove this redundancy. For example image data generated by sensors pointing to the
same direction can be aggregated as those data are less variant. However, data aggregation or data fusion techniques complicate QoS design in WSN.

- Heterogeneity of the Sensor Nodes: Handling heterogeneous data generated by different types of sensor nodes is another challenge in WSN. For instance, there are some applications which require different types of sensors to monitor temperature, pressure and humidity of the surrounding environment, capturing image or video of moving objects. Data generated from these sensors at different rates based on different QoS constraint and delivery models. Therefore, these types of diversified sensor network may impose significant challenges to provide QoS.

- Dynamic Network Topology and Size: Due to mobility of sensor nodes, link failure and node failure, the topology of the network may get changed. Self reorganizing and making this network adaptable to such changes is a challenging issue in WSN. A typical WSN may consist of hundreds to thousands of densely deployed nodes in a terrain of interest. The number of such sensor nodes may increase even after the initial deployment of the network due to the newly added nodes. Though these nodes are subjected to failure, the QoS should not be affected drastically due to increase or decrease of sensor nodes.

- Less Reliable Medium: The communication medium in WSN is radio. This wireless medium is inherently less reliable. The wireless links are also very much affected by different environmental factors such as noise and cross signal interference.

- Mixed Data Arrival Pattern: In a typical WSN application some sensory data may be created a periodically and these are mainly due to the detection of some critical events at unpredictable times. Again there can be some sensory data which are created at a regular interval of time e.g., continuous real time monitoring of some environmental parameters. Moreover the period of periodic data may or may not be known a priori and this may depend on the kind of application. Therefore data to be handled in a typical WSN may be a mixture of periodic and aperiodic type. This mix nature of data poses significant challenges in designing QoS based schemes (i.e., for guaranteeing timely and reliable delivery) for WSN.
• Multiple Sinks or Base Stations: Even though most of the sensor networks have only single sink or base station, there can be multiple sink nodes depending on the application's requirements. Wireless Sensor Networks should be able to maintain diversified level of QoS support associated with multiple of sinks or base stations.

4.1.2 QoS Requirements in WSN

The requirement of QoS in WSN can be specified from two perspectives [105]. These are application specific QoS and Network QoS.

4.1.2.1 Application Specific QoS

QoS parameters in WSN may vary depending on the application domain. Some of the application specific QoS parameters are data accuracy, aggregation delay, fault tolerance, coverage [185], optimum number of active sensors [186] etc. The application demands certain requirements from the deployment of sensors which are directly related to the quality of application.

4.1.2.2 Network QoS

From the network perspective, it has been considered a show to provide QoS constrained sensor data while optimally utilizing sensor resources. Every class of application has some common requirements in network. The network is concerned with how to transmit the sensed data from the sensor field to the sink node fulfilling the required QoS. There are three data delivery models in sensor network [187]. These are event driven, query driven and continuous. The event driven application in WSN is mostly delay tolerant, interactive and non end-to-end. The sensors detect the occurrence of certain event and to take action accordingly. In one side of the application there is a sink node and the other side a group of sensor nodes which are affected by certain events. The data sent by sensor nodes are highly redundant and has to be sent quickly and reliably to the sink node. The query driven application WSN are
interactive, query based, delay tolerant, mission critical and non-end to end. The queries are generated by the sink node on demand and sent to sensor nodes enquiring occurrence of certain event. In the continuous model sensor nodes send data to the sink node at a pre specified rate. The data can be real time audio, video, image or non real time data as well.

Data fusion is the process of one or several sensors then collect the detection result from other sensor. The collected data must be processed by sensor to reduce transmission burden before they are transmitted to the base station or sink. This technique mostly installed in the dense networks to increase sensor’s reading dependability and achieve the accurate estimation of environment that result in longer sensor life. The simplest data fusion function is duplicate suppression if two sources both send same data, Data fusion node will send only of these forward. Data fusion is very necessary in Wireless Sensor Networks because sensor node have a capability of sense data after sensing it. they are transmit sense data to a base station or sink. it is basically direct transmission and it is expensive since base station may be located very far away and sensor nodes in a network needs more energy power to transmit data over long distance. so that better schemes is that fewer nodes transmit data to this far distance. these fewer nodes called cluster head of individual cluster in Wireless Sensor Networks.

4.2 Motivation

There is evidence of exponential interest in research work for QoS of WSN that provision in the literature survey. However the primary focus of the researchers is bounded to development of cellular industry, wired Internet and MANETs. Some authors proposed the definition of QoS in varying networks in recent years. Crawley et al. [186] stated that, “QoS is the set of service requirements that meet the network while transporting a flow”. The Nikaein et al. [188] in his works proposed that, “QoS (in a MANET) is the provision of a set of parameters in order to adapt the applications to the quality of network while routing them through the network”. Despite of slight variation in definitions proposed for different types of networks, the QoS in any network is generally considered as the ability of system to deliver a guaranteed level of service to its users and/or applications. The service requirements are defined in a set of performance matrix typically computed in three styles: (i) additive (e.g. total delay along a
Self-Configuration in Wireless Sensor Networks

path); (ii) concave (e.g. minimum bandwidth along each link); and (iii) multiplicative (e.g. packet delivery ratio along entire route). Though some performance metrics such as delay, throughput, bandwidth, jitter (delay variance), bandwidth, reliability, packet delivery ratio (PDR) etc are used more significantly than other metrics, each application posses its own set of unique service parameters that are satisfied, while possibly compromising on other sets of metrics. Applications that are loss-tolerant such as multimedia applications does not get adversely affected by occasional data loss and are highly sensitive towards the bandwidth and delay. In contrast, other applications that involve sensitive data integrity like electronic mail and banking transactions require fully reliable transfer of data but may not require any stringent delay constraints to work with elastic bandwidth.

For many applications in wireless ad hoc and sensor networks, bare dominating sets are not a well suited organization. Often dominating set needs to fulfill the additional criteria is to be connected. Several applications in sensor networks such as routing and aggregation often requires a backbone based on connected dominating sets. This poses a maximum lifetime connected dominating set problem for improving the lifetime of sensor networks.

The reliability of WSN Jian Wan et al. [107] is affected by faults that may occur due to various reasons. In ZigBee sensor networks, ZigBee specification presents a self-configuration process which relies on its tree topology. When a ZigBee node failed, all of the descendants of the fault node should rejoin the network. Instead, authors present an efficient scheme for repair itself after node failure that just one or some of the descendants should rejoin the network. It allows a subnet of fault node rejoin the network with minimum message exchanges and, therefore, saves energy. Simulation results show that most of subnet rejoin the network in the first level of sub-tree, and confirm the efficiency of the proposed scheme.

DARPA recently announced about the self-configuration as the ability of the network that defines the combat coverage in efficient manner along with network disconnection caused by routing holes as the most important and desirable properties of WSN. Though the concept of self-configuration in various literature for the mobile nodes is relatively overviewed (Vlajic Moniz et al. [108]) in the area of research. This research focus mainly on two segments of holes i.e. the routing holes in which the energy aspect of combating holes via deployment of single node (also called as 'super') is discussed. The specific contribution of the paper is the self configuration of the nodes. The nodes with the high amount of initial energy are route to
Self-Configuration in Wireless Sensor Networks

the data for the long paths and the low energy nodes dissipate the low amount of energy (shorter path or dense node density). The mobile application in some manner is the high in justification of energy in the entire square like shaped and circle shaped holes, regardless to the actual size and number of end nodes that actively coordinate in the process of path routing. The need of consideration for the other parameters like for the example static-node failure or overall transmission delay during the decision of deployment of mobile is demonstrated. OPlaMoN is the simple distribution algorithm proposed for Optimal Placement of Mobile Node determination within the routing hole of any random topology. The algorithm is a complex optimization problem which is solved by division and solving of simpler components by individual nodes. The terminating solution is reached via cooperative communication of decision – making algorithm assuming that a minimum information exchange among effected nodes will be established. The algorithm provides excellent energy conservation and is highly configurable in WSN environments.

Faults in WSN are not a new problem and appears in almost every domain of the system. The pervasive applications for the end-users need the autonomic self-configuration. Wireless Sensor Networks are considered for the critical applications that are often unattended largely and requires reliability in results over the years of deployment. The real world applications like communication, clock drift, sensing and failure realities, system performance etc. degrade significantly over time. It is necessary that the natural deteriorations are continuously monitored and corrected with self-configuration when necessary. In their paper Jingyuan Li et al. [110], introduce a dependency constraint directed self-configuration scheme for WSN, including sensing, communication and tracking. The research indicates that, following the dependency constraints in self-configuration design is not only a must for the correctness of self-configuration services, but is also a key to energy efficient self-configuration.

In the last decade, there was a great technological advancement in the areas of sensors, integrated circuits and wireless communication, which led to creation of WSN. This type of network is being used in areas such as tourism, education, medical, military and others. The applications, in that type of network, impose specific requisites related to the energy consumption and reliable delivery. The routing protocols, for the WSN, must have self-configuration characteristics aimed at discovering the best way for transferring the information, with delivery assurance and with minimum energy consumption, among the
Self-Configuration in Wireless Sensor Networks

nodes that compose the network. Araujo H. et al. [111] proposed a modification of the Directed Diffusion Routing Protocol so as to reduce the energy consumption in the network when there is an occurrence of failures. The proposal utilizes a Geocast approach to repair broken paths by constructing a new routing tree. The performed simulations demonstrated that our proposal is more efficient from the energy viewpoint.

In publish/subscribe (pub/sub) communication scheme under s WSN, there exist inherent tradeoffs among conflicting objectives in event publication. To address this issue, Boonma P. [112] in their paper investigates pub-/sub middleware for WSN, called Tiny DDS. With its self-configuring event routing protocol, Tiny DDS adaptively performs event publication according to dynamic network conditions and autonomously balances its performance among conflicting objectives. Tiny DDS leverages an evolutionary multi-objective optimization mechanism to seek the optimal tradeoffs among objectives and adjust parameters in its event routing protocol. Simulation results validate the ability of tiny DDS to tune its event publication against dynamic network conditions. Tiny DDS is implemented lightweight and efficient enough to run on resource-limited nodes.

A sensor network application for the dense deployment selects specific nodes and the configuration topologies that results in the construction of a long-lived network. The active nodes collaborate for participation in message transmission via virtual and random active route called as routing path; nodes at rest are turned off i.e. breathe at low power state. The researchers approach to design network topologies that can minimize number of nodes in active route and thus save critical and constrained energy resources of nodes. These schemes preclude the existence of multiple active nodes in a specific area. The crowded active nodes are depreciating parameters of the energy efficient network cause of frequent message collisions. Hoseung et al. [113] in their work designed a self-configuration network topology for the system to achieve efficiency, robustness and scalability based on adjacent distance between two consecutive nodes. The scheme uniformly distributes the active nodes and gets considerable energy utilization in sensor field.

Authors in [114] examines one of the simplest such heterogeneous scenarios - in which sensors are equipped with different battery power in a clock-driven sensor network. In this paper the lifetime of a sensor network is defined as the expected lifetime of any given sensor in the network. In a densely deployed sensor network this definition can be easily extended to
be the time until a certain percentage of the sensors died. Thus in prolonging the lifetime of a sensor network, it is important to balance the power depletion from one sensor to another. The authors focus on the clustering approach and examine the use of a heterogeneous structure where some sensors carry more power than others, and thus naturally become cluster heads. For network model the authors considered a square sensing field with each side measuring meters. The coordinates of the field are as shown in Figure 4.2, where crosses represent overlay sensors, and circles represent normal sensors. All data collected by the sensors is to be sent to a receiver/collector located outside the sensing field. The collector is located at (0, -D), and is thus D meters away from the sensing field. This location is assumed to be fixed. Authors assume all sensors are aware of the location of the receiver via some type of pre-configuration or self-configuration. There are a total of n normal sensors in the field.

Figure 4.2: Heterogeneous network

The results show how the analysis can be used to determine the appropriate number of heads in the network. In summary, regardless of whether the network is homogeneous or heterogeneous, under the considered parameters, the optimal number of clusters is between 4 and 10. The exact number varies slightly depending on the size of the field, the location of the receiver, etc. (due to space limit authors are not able to show more results). However, when the distance to the receiver is very small comparing to the size of the field, it seems the best choice is to use as many clusters as the receiver can handle.
In [115], Qunfeng Dong et al. try to present formal analysis of a variety of network lifetime maximization problems in different energy consumption models. An analysis of energy consumption in WSN leads to two energy consumption models for formal analysis, i.e., the time based model and the intensively researched packet based model. Various network lifetime maximization problems are identified in individual models. The complexities of these problems are formally analyzed. In particular, authors identify different energy consumption models, define a variety of fundamental network lifetime maximization problems in individual energy consumption models, and formally analyze their complexity. Polynomial time algorithms are presented for tractable problems, and NP-hardness proofs are presented for intractable problems. Stationary sensor nodes are assumed to be equipped with an omnidirectional antenna. A WSN is denoted by a weighted directed graph $G = (V, A)$, where $V$ is the set of sensor nodes and $A$ is the set of directed links. Each node is labelled with a unique ID $i \in [1..|V|]$ and has a maximum transmission power of $P_{\text{max}}^i$. Let $P_{ij}$ denote the minimum transmission power required to maintain a reasonably good quality link from node $i$ to node $j$. $G$ contains link $(i, j)$ (i.e., the link from node $i$ to node $j$) if and only if $P_{ij} \leq P_{\text{max}}^i$. Initially, each sensor node $i \in V$ has an energy of $P_i$. Time is divided into discrete time slots, denoted by $t \geq 1$, $t \in \mathbb{Z}^+$. In the time based model, authors study the problem of maximizing network lifetime while preserving connectivity and prove that it is NP-hard. In the packet based model, authors formally define the following problems: broadcast lifetime, multicast lifetime, many-to-many unicast lifetime, many-to-one unicast lifetime, one-to-many unicast lifetime and one-to-one unicast lifetime. Broadcast lifetime and multicast lifetime are NP-hard, even if each node has a fixed transmission power. Authors show that the unicast lifetime problems are NP-hard in both the multiple commodity models and the single commodity model. However, authors show that in cases where each node has a fixed transmission power, many-to-one unicast lifetime, one-to-many unicast lifetime, and one-to-one unicast lifetime are polynomially solvable. Many-to-many unicast lifetime is also polynomially solvable in the single commodity model, but remains NP-hard in the multiple commodity models.

Yunxia Chen et al. [116] proposed a medium access control protocol that exploits both the channel state information and the residual energy information of individual sensors. Referred to as the max-min approach, this protocol maximizes the minimum residual energy across the
network in each data collection. They derived a general expression for the lifetime of WSN which holds regardless of the underlying network model. This formula provides insights on lifetime-maximizing protocol design. It reveals that a lifetime-maximizing protocol should exploit both CSI and REI of individual sensors. Based on this formula, we propose a greedy approach to lifetime maximization which achieves considerable improvement in lifetime performance.

Kewei Sha et al. [117] proposed a novel model to formally define the lifetime of a WSN based on energy by considering the relationship between individual sensors and the whole sensor network, the importance of different sensors based on their positions, the link quality, and the connectivity and coverage of the sensor network. Using the proposed model, they have compared two types of query protocols, the direct query protocol and the indirect query protocol, in terms of both mathematical analysis and comprehensive simulation. The lifetime model was formalized from the view of energy, the most important consideration in the design of the sensor network. In the lifetime model, authors combined several significant factors such as the relationship between the individual sensor and the whole sensor network, the importance of the sensor based on their location, the link quality of the wireless communication, the connectivity and coverage of the sensor network. Considering the real usage of the proposed lifetime model, they can have either centralized or localized algorithm to calculate the value of LSN. In the centralized approach, authors assumed that there is a program running at the sink, which collects the corresponding information from the sensor network periodically, so the powerful sink can easily detect the coverage of the sensor network and calculate the RLSN based on the formula to make a decision on the time when the lifetime of the sensor network is over. In the real implementation, authors do not want this operation consumes extra energy. Thus, this piece of information can be piggybacked by regular data messages. On the other hand, it is also easy to develop a decentralized algorithm to calculate the lifetime because the RLIS and the link quality are properties of the individual sensors, which is always available to each sensor; the importance of a sensor is decided by the location of the sensor, which is easy to get by proposed localization approaches. After calculating the RLIS of each individual sensor, it will send it back (using aggregation) to the sink so that the LSN of the sensor network can be calculated. Based on these models authors
compared two query protocols; both theoretical and simulation results show that IQ balances the load so that extends the lifetime of the sensor network.

In several previous work, the lifetime of the sensor network is defined as the time for the first node to run out of power such as in Refs.[119-122] or a certain percentage of network nodes to run out of power as in Refs.[123, 124]. We think that these definitions of the lifetime of the sensor network are not satisfactory. The former is too pessimistic since when only one node fails the rest of nodes can still provide the whole sensor network appropriate functionality while the latter does not consider the different importance of the sensors in a sensor network.

In the work of Refs [125, 128], the lifetime of the sensor network is defined as the time when the sensor network first loose connectivity or coverage. The rationale of their definition is based on the functionality of the sensor network, which is similar to our definition. However the way to detect the termination of the sensor network is different. Blough and Santi34 define it by checking the connectivity of a graph; Mhatre et al. use a connectivity and coverage model to describe it; while we define it as the time when the remaining lifetime of the whole sensor network starts to keep constant as losing connectivity or the sensor network loses coverage.

Xue and Ganz study the lifetime of a large scale sensor network in Ref.[129]. They explore the relationship between the lifetime of a sensor network with the network density, transmission schemes and maximum transmission range. Their work is based on a general cluster-based model, and does not consider the importance of different sensors. They also aim to explore the fundamental limits of network lifetime. Compared with their work, our model is more general which can be used not only for cluster based model. Furthermore, because we take more factors into consideration in our model, our model is more useful and flexible, in which the lifetime is calculated according to the really energy consumption.

Bhardwaj et al. define upper bounds on the lifetime of the sensor network in Refs.[125, 126]. They explore the fundamental limits of data gathering lifetime that previous strategies strive to increase. One of their motivations is to calibrate the performance of collaborative strategies and protocols, but they just give out an upper bound of the lifetime rather than the actually lifetime model for different strategies. Besides, our model can also guide the design of the low-level protocols.
In [130], the authors perform a comprehensive study of the problem of scheduling the communication between the central controller and other wireless nodes, with focus on energy conservation. The paper contributes three directory protocols that may be used by the central node to coordinate data transmissions considering multiple factors such as traffic-type (e.g. downlink, uplink, peer-to-peer) and the effects of packets errors.

In [131], the authors propose an energy conservation technique for Wireless Sensor Networks that works by selecting and successively activating mutually exclusive sets of sensor nodes, where every set completely covers the entire monitored area. Their method achieves energy savings by increasing the number of disjoint covers. The authors propose a heuristic solution to this problem.

In [132], the authors proposed a new multi-access protocol, PAMAS, based on MACA [8], with the addition of a separate signalling channel. PAMAS achieves energy savings by powering off the nodes which are not actively transmitting or receiving packets.

Mihaela Cardei et al. [118] proposed an efficient method to extend the sensor network operational time by organizing the sensors into a maximal number of disjoint set covers that are activated successively. Only the sensors from the current active set are responsible for monitoring all targets and for transmitting the collected data, while nodes from all other sets are in a low-energy sleep mode. In this paper we address the maximum disjoint set covers problem and we design a heuristic that computes the sets. Authors considered that a large number of sensors are dispersed randomly in close proximity to a set of objectives and send the monitored information to a central processing node. Every target must be monitored at all times by at least one sensor and every sensor is able to monitor all targets within its operational range. The disjoint sets in the approach are modeled as disjoint set covers, where every cover completely monitors all the target points. Authors assume that the targets have fixed locations, so the algorithm for computing the covers is executed only once by a central node after the location for all sensors has been determined. After the wireless sensors are deployed, they activate their positioning service and send their location information to the central node. Based on this information, the central node computes the disjoint set covers and sends membership information back to every sensor. Knowing the set it belongs to and the number of covers, every sensor is then able to identify the time periods when it has to be active or in the sleep state. Authors assume that a time synchronization service is available to
sensors, most likely facilitated by periodic beacon messages from the central node or on-board GPS receivers. Authors proposed an efficient heuristic \textit{MC-MIP} with a mixed integer programming formulation.

4.3 Data Fusion Technique

Data-aggregation focuses on removing duplicate data which can result while aggregating information that is very much essential for energy-constrained WSN. Aggregators are vulnerable to attack when it is comprised by injecting false data in the sensor network. Another possible attack is to compromise a sensor node and inject forged data through a sensor node. Without authentication, the attackers can fool the aggregators into reporting false data to the base station. Secure data aggregation requires authentication, confidentiality, and integrity. Moreover, secure data aggregation also requires the cooperation of sensor nodes to identify the compromised sensors.

4.4 Methodology

The solution adopted by Yu and Prasanna [17], which relied on minimizing the energy dissipation in overall sensors along a path spending same energy in all the nodes, is not optimal due to, as we described in our scenario that different nodes can have more traffic load than others. Therefore the node’s consumption of energy inside a path can be different to the consumption of the other nodes.

Let communication path involves \( n \) nodes, \( S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow \ldots \rightarrow S_n \) and consider a packet size \( s \) is needed to transmit from \( S_1 \) to \( S_n \) in a given time interval \([0, T]\).

By fixing the symbol rate \( R_i \), for each hop, the time duration for transmission over can be modeled as: \( S_1 \rightarrow S_{i+1}, \tau_i \) can be modeled as:

\[
\tau_i = \frac{s}{b_i \cdot R_i}
\]  

(4.1)

Further, the corresponding energy dissipation for sending the packet \( E_i \) can be modeled as:

\[
E_i = (P_i + D_i \cdot R_i) \cdot \tau_i = \left[C_i \cdot (2^{b_i} - 1) + D_i\right] \cdot \frac{s}{b_i}
\]

(4.2)
Where $P_i$ is the transmission power; $C_i$ is determined by the quality of transmission, in terms of Bit Error Rate, and the noise power; $D_i$ is a device-dependent parameter that determines the power consumption of electronic circuitry; $b_i$ is the constellation size in number of bits per symbol.

By substituting $b_i = \frac{s}{\tau_i R_i}$, in above equation it can be shown that $E_i$ is non-negative, monotonically decreasing function of $\tau_i$.

Let $\omega_i(\tau_i)$ denotes energy function. $\bar{\tau} = \{\tau_i, i = 1, 2, 3, ..., n-1\}$, is feasible if the transmission latency, $\sum_{i=1}^{n-1} \tau_i$, is within $T$. Consider that the energy functions of all the hops are known a priori. Then the algorithm can be defined as:

Given a series of consecutive single-hop communication links $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow \ldots \rightarrow S_n$ and packet size $s$, find a feasible schedule of packet transmission, so as to minimize the maximal energy dissipation over all nodes:

$$OBF_T = \max_{i=1}^{n-1} \omega_i(\tau_i) = \max_{i=1}^{n-1} \left[ C_i \left( \frac{s}{2\tau_i R_i} - 1 \right) + D_i \right] \cdot \tau_i R_i$$

(4.3)

The node $S_i$ that satisfies $\omega_i(\tau_i) = OBF_T$ is called the critical node for $\tau_i$ or simply critical node. In general, critical node may not be unique.

The necessary and sufficient conditions for optimality are:

1. $\sum_{i} \tau_i = T$; and
2. $\omega_1(\tau_1) = \omega_2(\tau_2) = \ldots = \omega_{n-1}(\tau_{n-1})$

The value of $b_i$ can be derived from equation 1 based on the final value of $\tau_i$. Hence by variations in latency modulation parameter can be changed.

### 4.4.1 Tree-Based Topology

Many researchers have worked in WSN and presented a variety of considerations in traffic pattern, many-to-one or in other words the “tree topology”. The topology here performs the test of our algorithm because of their non-uniform distribution of energy characteristics along a path. Nodes located near to the sink or the main node consumes a high amount of energy in
comparison with the other nodes that are closer to the lowest levels of the tree. This section discuss about the spanning tree with data fusion.

### 4.4.2 Data Aggregation Tree

31 nodes with 5 levels of a spanning-tree is depicted to test the protocol in figure 4.3.

![Figure 4.3 Spanning Tree](image)

Usually this spanning tree is based on the data fusion and data aggregation. The nodes of the tree gather the information by aggregating themselves from its children nodes and route the packet to sink. As for the reason that the correlation level between the information a packet of reduced size is generated. The correlation level \( k \) between information of the range \([0,1]\); is the reasoning factors in the generation protocols along with the number of source nodes \( d \) in the sub tree routed. The relationship among the previous packets \( s \) and the new packets \( s' \) are abstracted in following equation [287]:

\[
s' = \frac{ds}{d^k - k + 1}
\]

(4.4)

For our analysis it is also assumed the following abstraction:

- First, all sensor nodes (16, 17, 18, 19 ... 31) broadcast packets at equivalent times with a latency limitation.
- Secondly, if \( k = 0 \) each node will broadcast two times higher the sum of packets than its children nodes do. This means that double amount of energy will be depleted.
Therefore, if \( k \neq 0 \) a factor \( \left( \frac{t}{k} \right) \) will determine the sum of energy exhausted for propagation of each packet.

The motive is try-to-emulate the consumption of real node in this abstraction for whole network. The nodes gains packets from multiple packets occurs due to reason that previous simulations state about the protocol of one packet transmission via each node from a single source i.e. not from the multiple sources.

The simulation of one path for the spanning tree is performed by choosing the path composed by nodes 1, 2, 4, 8 and 16. Equation 4.4 stats about the energy consumption factors \( Z_j \) for \( k = 0.5 \).

<table>
<thead>
<tr>
<th>Node</th>
<th>16</th>
<th>8</th>
<th>4</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z )</td>
<td>1</td>
<td>1.33</td>
<td>1.78</td>
<td>2.37</td>
<td>3.16</td>
</tr>
</tbody>
</table>

Table 4.1: Energy Consumption Distribution along the Path

For instance, the node with factor 2.37 stats the amount of energy it will spend i.e. \( E \times 2.37 \) when all the sensor nodes transmit one packet at same time due to correlation among packets sent from discrete sources.

The parameters for the simulation were the following ones:

- \( \Omega = \) variable and \( \gamma = 3 \) (static constraints)
- 5 nodes
- \( s = 100 \) bits; \( R = 1 \) MHz and \( T = \) variable
- \( C = [0.8 \times 10^{-7}, 1.2 \times 10^{-7}] \) \( (\text{uniform distribution}) \)
- \( D = [0.8 \times 10^{-7}, 2.2 \times 10^{-7}] \) \( (\text{uniform distribution}) \)
- Maximum battery level \( \epsilon_{\text{max}} = 10 \) Joules

At this point it is difficult to determine whether the algorithm should be initiated from the starting of path or in else condition is efficient to employ the optimal values of modulation until the energy level of certain nodes goes down beyond the threshold values. Hence, the power aware protocol and the Balancing algorithm were mixed. When one node acquires a
specific starting level, the power-aware protocols are initiated from the starting point. The energy above the threshold value initiates this condition.

4.4.3 Optimal Selection of Modulation Scheme

\[ [E_{opt}, Tn] = \text{optimal}_\text{energy}(T, n, C, D, R, s, PR) \]

This function is based on the work done by Yu and Prasanna [288]. Given a group of nodes the aim is to find a group of points \((E(i); \tau(i))\) where the following constraints must be satisfied:

\[
E(i) = E(i + 1); \text{for } i = 0, 1, ..., n - 1; \quad (4.5)
\]

\[
\sum(\tau_i) = T; \quad (4.6)
\]

This means that to find a solution where all nodes consume the same amount of energy, the latency constraint must be satisfied.

---

Figure 4.4: Optimal Selection of Modulation Scheme
4.4.4 Balancing Process of Modulation Scheme

Given a group of nodes (hereafter balancing group), the aim of balancing leads in finding the proper modulation parameters for balancing group nodes in order to achieve an energy saving in the node selected as low battery level node (hereafter target node) while satisfying a latency constraint.

The most important inputs are the optimal modulation parameters and omega. As said before, omega is the amount of energy that we want to spend in the target node. It is comprised between 0 and 1. For instance, omega = 0.63 would be a 63% of energy saving (figure 4.5).
4.4.5 General Description of Main Block

Firstly, the optimal modulation node parameters of the whole path are computed together in order to know the theoretical optimal transmission time of each node. Secondly, put in order the nodes from lower values of battery level to higher ones. Thus it is easier to manage groups in order to perform the balance of energy. The strategy is to take one node from the beginning of the list as a target, and then start to take nodes from the end of the list as balancing nodes. Thirdly, go into a loop where a target is selected and several balancing nodes are assigned to it.

The amount of balancing nodes depends on the following aspect. It is needed to achieve the stated saving (defined by omega) in target node. Always a group of nodes is selected before performing the balance of energies. This group is composed of the target node and the balance group of nodes. It searches a combination of modulation parameters such the target node satisfies omega. Furthermore, γ has to be satisfied in all nodes that belong to the balance group. Therefore, if the energy spent ratio of some node of balance group overcomes gamma factor it is needed to add a new node to the group. The higher number of nodes in balance group the lower energy requirements per each node for achieving stated saving in target node.
Figure 4.6: Block diagram of top architecture for balancing modulation order
4.5 Results and Discussions

The total number of transmissions for the Power-aware protocol was:

\[ N_{tx} = 18013 \]

And for the non-aware protocol:

\[ N_{tx} = 3245 \]

The best results were observed with omega equal to 0.4 for static constraints. The highest value for the life-time starts at the level of 70% however, difference among the different starting levels is negligible. The expanse of lifetime is approx. 41% higher regarding the Power-aware protocol.

If \( \text{batlevel} < 10 \)

\[ \omega = 0.4; \, \gamma = 1; \]

else if \( \text{batlevel} > 50 \)

\[ \omega = -2; \, \gamma = 3; \]

else

\[ \omega = 0.15; \, \gamma = 2.7; \]

end
Figure 4.7 Behavior of dynamic constraints

Figure 4.8 Behavior of dynamic constraints

Figure 4.9 Behavior of dynamic constraints
Figure 4.10: Behavior of static constraints ($\omega=0.2$, $\gamma=1$)

Figure 4.11: Behavior for static constraints ($\omega=0.4$, $\gamma=3$)
Figure 4.12: Behavior for static constraints (omega=0.2, gamma=2)

Figure 4.10, figure 4.11 and figure 4.12 elaborate the management of nodes in the proposed algorithm when dynamic parameters are employed. To the surprise, the lifetime of results was similar or even worse. The simulation of other topologies the conclusion is sentenced that the use of static constraints is better than running the algorithm with dynamic constraints. The reason that the dynamic constraints in the network show more conserving nature, than the static ones present. There exist the three levels of modulation parameters to be applied to nodes. The nodes that have lower modulations (around 40 nodes) are target nodes that save around 40% of energy. The nodes with high modulations (around 50) are achieved. They are balancing nodes and thus expend around 200% of energy.

Finally another group of middle nodes (around 10) have more rounded optimal modulation values. From the node point of view that only belong to one path is optimal. However, the nodes that belong to several paths is not optimal.

Therefore, the high remaining energy can be seen when the path dies or is not used for helping target nodes.
Figure 4.13: Energy curves for long-range communications

The approach here is the long-range communications (hops from 10 to 50 meters). The study elaborates the reason by energy curves that indicates the difficulty in achievement in energy saving for short range communication (Figure 4.13)

Our approach is for long-range communications (hops from 10 to 50 meters). The reason comes from a study of the energy curves performed, where as it can be observed in Fig. 4.14 and it is difficult to achieve energy savings in short range communications Fig 4.16.
The savings in the previous graph considers that transmission time of a node can be changed for continuous range of time. However, the fact that transmission time is directly bounded to bounded to modulation parameter is responsible for modulation parameter to adopt an integer value in the discrete range of time periods. This indicates about the limitations of user to use the transmission times that can be computed via different values of modulation parameters (figure 4.14 shows that modulation order increases with consumption in energy).

Figure 4.14 Energy curve with respect to different modulation scheme
The total number of transmissions for the power-aware protocol was:

\[ N_{tx} = 18013 \]

And for the non-aware protocol:

\[ N_{tx} = 3245 \]

The best results were observed with omega equal to 0.4 for static constraints. The highest value for the life-time starts at the level of 70% however; difference among the different starting levels is negligible. The expanse of lifetime is approx. 41% higher regarding the power-aware protocol.
Fig 4.16 No. of transmissions for $T = 15 \times 10^{-5}$ ($\omega = 0.06$)
Fig 4.17: No. of transmissions for $T = 15 \times 10^{-5}$ ($\omega = 0.04$)
Fig 4.18: Analysis of life time with variable latency
The total number of transmissions for the Power-aware protocol was:

\[ N_{tx} = 77043 \]

And for non-aware protocols:

\[ N_{tx} = 3245 \]

The best result for a starting level equal to 100% was achieved for a static parameter equal to 0.2. The increase of the lifetime regarding the Power-aware protocol is approximately 34% higher (1.34 times higher).

The tables 5.2 and 5.3 states about the proper starting level in case when all the nodes are fully charged i.e. at 100% battery level. Also, it is observed that reduction is observed on overall increase in lifetime of nodes as for reason of different latency employed. Hence, the simulations of different nature are required to be performed so as to visualize the dependence of saving along with latency path \( T \). The values responsible to supplement the lifetime regarding Power-aware protocol for several values of latency in the modulation parameter in the range of \( 3 < b < 9 \).
Self-Configuration in Wireless Sensor Networks

The suitable choice for the strong dependence is base for the application. The application that operates in the environment of high modulation parameters (low latency) acquires favorable increase in lifetime. However, the applications that assist low modulations or low energy consumption shows smaller increase in the lifetime of network with the proposed algorithm. Hence to measure the best option, best way is the assumption of latency analysis i.e. it has same probability of usage. Hence the computation of mean value for every lifetime increment provides us the best choice of combinational choices.

Thus,

- $\text{mean (}= 0.2) = 46:9692 \Omega$
- $\text{mean (}= 0.4) = 44:1369 \Omega$
- $\text{mean (dynamic)} = 42:4862 \Omega$

If we use a combination of omegas we obtain the best choice:

- $\text{Omega} = 0.4 \in T = [5.55 \times 10^{-5}, 11.1 \times 10^{-5}]s$
- $\text{Omega} = 0.4 \in T = [11.1 \times 10^{-5}, 16.65 \times 10^{-5}]s$

Then we obtain:

- $\text{mean (Ω}_{\text{MIXED}}) = 51:3646\% (1.51 \text{ times higher})$

4.6 Conclusion

We carried out a performance comparison between our protocol and the Power-aware and Non-power-aware methods. On the one hand, the results regarding the Non-power-aware algorithm were not commented due to the really huge increases of lifetime obtained. Lifetime was improved up to 40 times with our approach compared to Non-power aware algorithm, depending on the latency constraint applied. On the other hand, many comparisons regarding the power-aware protocol were presented. The increase of lifetime ranged from 50% to 120% more transmissions.

On all topologies tested, we observed a tendency of having less increase of lifetime as the latency grew up (modulation parameter decreased). The reason can be explained with the Fig. 6.2. As it can be seen, when we are working with low modulation parameters ($b= 2, 3, 4$) the curves become flatter. To make matters worse, the latency separation between different
modulations is bigger. On the contrary, when we are working with high modulation parameters (7, 8, 9), the latency separations between different modulations are smaller. Also, the curves have a higher inclination. These characteristics make it more difficult to reach a specific saving in the low modulation area than in the high modulation one and therefore, the performance decreases as the latency rises up.

We have to mention that to our surprise, the results for dynamic power constraints were worse than static constraints. The reason is due to the fact that dynamic constraints are more conservative than static ones and therefore more unused energy remains in the nodes after a path is "lost" due to power-outage in one or more nodes.