GAME THEORETIC APPROACH ON ANALYSIS OF WIRELESS SENSOR NETWORKS

Thesis submitted towards the partial fulfilment of requirement for the award of degree of

MASTER OF ENGINEERING

IN

WIRELESS COMMUNICATION

Submitted by:

Divleen Kaur

Roll No: 801363010

Under the guidance of:

Dr. Ravi Kumar

Assistant Professor, ECED

Thapar University, Patiala

ELECTRONICS AND COMMUNICATION ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

(Established under the section 3 of UGC Act, 1956)

PATIALA – 147004, PUNJAB, INDIA

JUNE 2015

DECLARATION

I hereby declare that the work which is being presented in the thesis entitled, **"GAME THEORETICAL APPROACH ON ANALYSIS OF WIRELESS SESNOR NETWORKS"** partial fulfilment of the requirement for the award of degree of Master of Engineering (Wireless Communication) at the Electronics and Communication Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of Dr. Ravi Kumar, Assistant Professor, ECED. The matter presented in this thesis has not been submitted in any other University/Institute for the award of any other degree.

Date: 8/7/2015

Kaur

Divleen Kaur
Roll. No: 801363010

It is certified that the above statement made by the student is correct to the best of my knowledge and belief.

Ravi Kumar
8/7/15

Dr. Ravi Kumar
Assistant Professor
ECED,
Thapar University

Counter signed by:

Dr. Sanjay Sharma
Professor & Head
Thapar University
Patiala-147001

Dr. S. S. Bhatia
Dean of Academic Affairs
Thapar University
Patiala-147001

ACKNOWLEDGEMENT

I take this opportunity to express my profound sense of gratitude and respect to all the people who helped me throughout my thesis work. I am eternally grateful to my research supervisor **Dr. Ravi Kumar, Assistant Professor,** Electronics and Communication Engineering Department, Thapar University, Patiala, for his guidance, valuable suggestions and invigorating support. I wish to express my deep gratitude towards him for providing individual guidance and patiently correcting me with my work.

I would also like to convey my sincere thanks to **Head of the Department, Dr. Sanjay Sharma,** as well as **PG Coordinator, Dr. Amit Kumar Kohli, Associate Professor, ECED**,and the entire faculty and staff of Electronics and Communication Engineering Department for their encouragement and cooperation.

Finally, I am thankful to the almighty God bestowed upon me the ability, strength and confidence to complete my research analysis successfully. I also extend my heartfelt gratitude towards my family and friends who were a constant source of encouragement throughout the time.

Kaur

Divleen Kaur

ABSTRACT

Wireless sensor networks have become increasingly popular due to their wide range of applications. K-means is a typical clustering algorithm for clustering of these wireless sensor networks, and it is widely used for grouping of large sets of data, owing to its ease of computation and implementation. However, due to the limitations and inaccuracy of the algorithm, two novel approaches for clustering of wireless sensor networks have been comprehensively analyzed in this thesis.

In this report, improved methods for clustering of wireless sensor networks for a data set representing the information gathered by the sensor nodes have been provided. Also, the performance and comparison of these methods has been done, based on the experimental results. PCA based K-Means Algorithm proved to be a more efficient approach for clustering, as compared to regular K-Means Clustering. Furthermore, a comprehensive analysis of the regular K-Means Algorithm was carried out, as compared to the Game Theoretic Weighted K-Means Algorithm. Shapley Values were used to perform relative weighting for each element in the original data set, in order to form more optimal cluster memberships based on their weights.

The comparative results demonstrate the robustness and effectiveness of the proposed method, which are further substantiated by values of Davies-Bouldin index values and the energy dissipated after clustering.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1

INTRODUCTION

1.1 Clustering in Wireless Sensor Networks

1.1.1 Introduction to Wireless Sensor Networks

With the recent advancements in wireless communications, highly integrated digital electronics, processor memory, radio, low power and Min-Electro-Mechanical-Systems (MEMS) technology, Wireless Sensor Networks (WSNs) have become one of the most interesting and promising fields of research in the past few years. It has become possible to significantly develop tiny in size, less power and low cost multi-functional wireless sensor nodes. A wireless sensor network is envisioned as an autonomous and a self organising system, composed of a large number of sensor nodes, deployed densely either inside the phenomenon they are observing, or very close to it. Wireless Sensor Networks are generally used for systematic gathering of useful data related to the surrounding environment, like temperature, humidity, seismic data etc, and for transmission of this gathered data to a base station, i.e. sink, for further processing. The numerous real-world applications of WSNs include environmental monitoring, health monitoring, disaster management, precision agriculture, military monitoring, tracking applications etc. These applications are made possible due to the fact that WSN has a short system setup time and sensors can be disposed with acceptable operation cost. Hence, it's expected that in the future, wireless sensor networks will form an essential and integral part of our lives, even more than the present day personal computers. [1]

A sensor node is generally comprised of four basic elements viz. a power unit, a processing unit, a transceiver unit, and a sensing unit. Most of the time the nodes are randomly deployed in inaccessible terrains or in disaster relief applications. This means that sensor network protocols and algorithms must possess self organizing capabilities. Sensor networks contain a unique feature, which is, the cooperative effort of sensor nodes that are fitted with an on-board processor. Instead of sending the raw, unprocessed data to the nodes responsible for the fusion, the processing abilities of sensor nodes enable to carry out simple computations and transmit only the required and partially processed data. [2]

The key challenge in the design and operation of wireless sensor networks (WSNs) is the maximization of system lifetime. Clustering of wireless sensor nodes is commonly considered as one of the most promising techniques for dealing with the given challenge. Wireless sensor nodes are capable of wireless communications, sensing and computation. Hence, it is clear that a wireless sensor network is the result of the combination of distributed information processing, embedded techniques, sensor techniques and communication mechanisms.

1.1.2 The Hardware Structure of a Wireless Sensor Node

A wireless sensor node is an electronic device which is used as an interface between the physical parameters that can be sensed in the sensor field and a data wireless network. The hardware structure of a WSN and the layout of wireless sensor nodes scattered in a field is shown in Figure 1. A sensor node is made up of four main units [3]:

(1) A power unit: It consists of a battery and a number of DC/DC converters. DC/DC converters are also called voltage regulators and provide appropriate voltages to different circuits in the wireless sensor node. Linear regulators have much larger energy losses than switched regulators. Hence, the design of the voltage regulator has a significant impact on power consumption of the node.

(2) A processing unit: It usually consists of a small processor and memory. This unit basically reads the data from the physical sensors, extracts suitable information from the digitized data and implements the network protocols. The processing unit of a sensor node is responsible for determining the energy and computing capabilities of a wireless sensor node.

(3) Physical sensors: Sensors transform a physical or a chemical parameter or variable into a useful electrical signal. Some examples of these variables are distance, acceleration, temperature, pressure, displacement, inclination, pH and

humidity. The electrical signals can be derived from a change like a change in the resistance or capacity, in the sensor.

(4) The transceiver circuit: It is a radio system that is formed by the combination of a transmitter and a receiver.

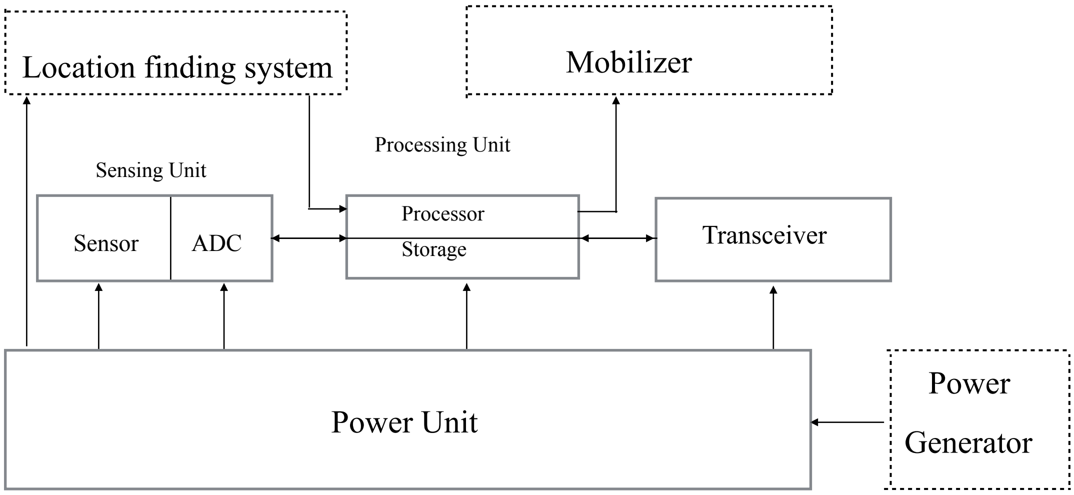

Figure 1.1 Hardware Structure of a Sensor Node (source [1])

1.1.3 The Communication Architecture of Wireless Sensor Networks

As shown above in Fig. 1, a wireless sensor network (WSN) is a network made of a numerous number of sensor nodes with wireless communications, sensing and computation capability. These sensor nodes are placed in a suitable environment, unattended (i.e. sensor field) and situated far away from the user as shown in Fig.2. The main components that build up the architecture of WSNs are [3]:

• The Sensor nodes that constitute the sensor network: These nodes have the main objective of making discrete and local measurements about the phenomenon surrounding these sensors, then to form a wireless network by communicating through a wireless medium, and finally collect the data and then route it back to the user via the sink (Base Station).

• The sink (Base Station): It communicates with the user via internet or satellite communication. It’s usually located near the sensor field of the sensor network. The data collected from the sensor field is routed back to the sink by a multi-hop infrastructure-less architecture through the sink.

• Phenomenon: It is an entity of interest to the user to collect its measurements. The phenomenon is properly sensed and analyzed by the sensor nodes.

- User: The user requires the information about a specific phenomenon to measure/ monitor its behaviour.

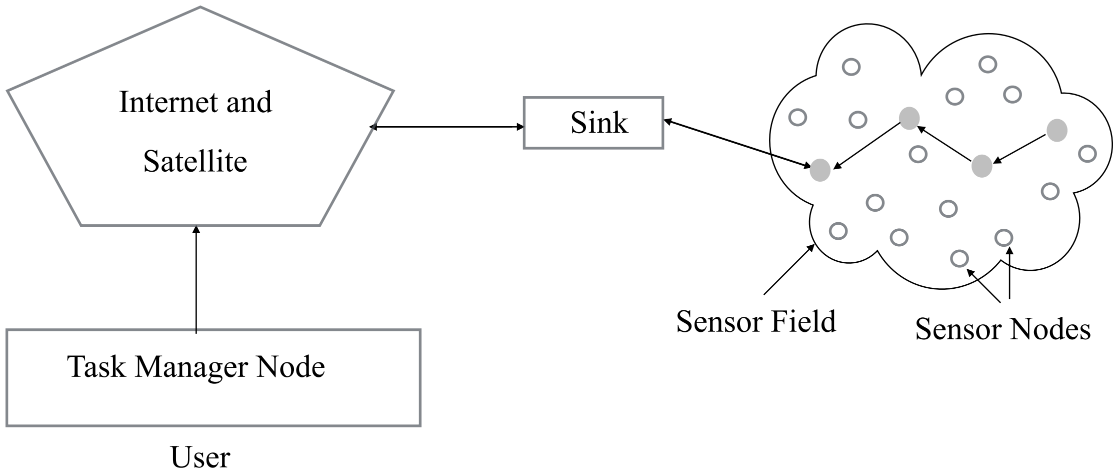

Figure 1.2 Layout of Sensor nodes scattered in a field. (source [1])

1.1.4 Design Parameters and Requirements

The design factors of overall wireless sensor networks communications architecture as well as the design factors of protocols and algorithms for wireless sensor networks (WSNs) are enlisted below. Many researchers in this field have addressed several design factors. These factors serve as hints or guidelines to design a protocol or algorithm for WSNs. [2-3]

- Reliability: Reliability or fault tolerance of a sensor node is the ability to maintain the sensor network functionalities without any interruption due to sensor node failure.
- Density and Network Size/Scalability: Hundreds, thousands or millions of sensor nodes may be deployed to study a phenomenon of interest to users. The density of

these nodes affects the degree of coverage and area of interest. The network size affects reliability, accuracy, and data processing algorithms.

- Sensor Network Topology: The topology of a network affects many of its characteristics such as latency, capacity, and robustness. Also, the complexity of data routing and processing depends on the network topology.
- Energy Consumption: One of the components of sensor nodes is the power source which is limited enough. A sensor node is battery-operated. Hence, the life time of a sensor node depends strongly on the battery life time, especially where no power source replenishment is possible in certain application scenarios.
- Data Aggregation/Data Fusion: It is the task of reducing data size by summarizing the data into a set of meaningful information via computation while data are propagating through the wireless sensor network. As sensor networks are made of large number of sensor nodes, this can easily congest the network and flood it with information. Hence, a solution to data congestion in sensor networks is to use computation to aggregate or fuse data within WSN, then transmit only the aggregated data to the controller.
- Security: Security aspects in WSNs have been focused on the centralized communications approaches. Some of the threats to a WSN are as follows: Node Outage, Passive Information Gathering, False Node, Node Malfunction, Message Corruption, Denial of Service, Supervision of a Node and Traffic Analysis. There is a need to develop distributed security approaches for wireless sensor network.
- Quality of Service: For some applications, data delivery within a bounded latency (i.e., time constrained applications) is of great importance; otherwise, the sensed data that delivered after certain latency will be useless. In other applications like not time-constrained applications, power conservation is more important than the quality of the sent data. Hence, there is a tradeoff between the quality of service/the quality of data sent and the energy conservations or consumption depending on the applications.
- Node Deployment: The node Implementation in Wireless Sensor Networks depends on the kind of application being used and it directly affects the performance of routing protocol. On one hand, it could be a deterministic distribution, where the

sensors are placed manually and data is routed using default routes. On the other hand, it could be a random distribution, where the output distribution of the nodes is not uniform. It always aims to find the optimal clustering which allows the utmost best connectivity. But sometimes, it works on the assumption that the network has an energy-efficient behaviour. Because communication between the sensor nodes is mostly limited in bandwidth and in the packet's delivery time, the most probable routes can easily be formed by multi-hop wireless paths.

- Fault Tolerance: Some wireless sensor nodes might fail and stop the data transmission because of physical damage, power shortage or environmental interference. Such node failures should not interfere with purpose of the network. Thus, MAC layer protocols and routing protocols should be able to adapt to the formation of new routes and links The network should always remain functional and should be able to continue data transmission. In case, if there are many node failures, implementation of redundancy techniques at various levels might be necessary to get a suitably good level of fault tolerance.
- Coverage: The sensor node's view of the environment that it is situated in is limited both in range and in accuracy. This means the ability of sensor nodes to cover physical area of the environment is limited.
- Connectivity: A permanent connection between any two individual sensor nodes that are densely deployed in a sensor network defines the network connectivity. This connectivity is extremely important, since it influences communications protocols' design and data dissemination techniques.

1.1.5 Clustering in Wireless Sensor Networks

Clustering is a kind of unsupervised learning, i.e. no assumption on the information on the class is made, and the goal is to find all the "natural" groups occurring in the data. Wireless sensor networks clustering involves grouping of sensor nodes into a set of clusters on the basis of similarities and dissimilarities among the data elements, in such a manner so that the similarity measure between the sensors within a cluster is higher than the similarity measure between sensors of the other clusters. The emergence of cluster structure depends on various choices, such as the choice of a similarity measure and clustering algorithm, data representation and normalisation.

The widely used clustering approach could be either (i) hierarchy based or (ii)partition based. Both of these approaches have their own advantages and shortcomings in terms of cluster size, number of clusters, shape of clusters, separation between clusters, etc. The hierarchy based algorithms work in a bottom up manner, with each pattern belonging to a separate cluster. The clusters are then iteratively merged, according to a suitable criterion. The divisive algorithms start from the entire data set in a single cluster and work in a top-down manner by iteratively dividing every cluster into two components till all clusters are singletons in the end. The partition based algorithms provide a partition of the dataset into a number of clusters. Partitional algorithms usually have input parameters which control the number of clusters that are produced.A few other approaches are based on hybridisation of the various clustering techniques. Many clustering algorithms use the center based cluster criterion. The center of a cluster is called a centroid, which is the average of all the data points in a cluster. [4]

The main challenge in the operation of wireless sensor nodes is the maximization of system lifetime. Clustering of wireless sensor nodes has been considered as one of the most promising solutions for dealing with the said challenge. The sensor nodes are grouped into a set of disjoint clusters, each having a designated node as a cluster head. This cluster head forms the network backbone and is responsible for all the inter-cluster communication for that particular cluster. Nodes in one cluster don't transmit their gathered data directly to the data sink but to their respective cluster heads, which transmits it to the data sink, which reduces the amount of energy spent. If a cluster is too large the intra-cluster point-to-point communication overhead is very high and the nodes may drop out of the network. Alternatively, if the cluster is too small, the nodes will get to get the cluster head frequently, causing the power to dissipate rapidly due to long distance communication overhead, hence it will drop out of the network. However, if the clusters are approximately equipartitioned, the performance of the network improves significantly. Using the Game theoretic K-Means Algorithm, the ad hoc network can be partitioned into load balanced as well as compact partitions [5].

Due to the power in batteries having limited availability in sensors and high communication they must be divided into teams to optimise the two most important objectives:

1) Compaction: To minimise power dissipation in intra-cluster communication

2) Equipartitioning: To form clusters with uniform power distribution, so that the clusters do not drop out of the system due to rapid power exhaustion.

One solution to achieve these objectives is by using K-Means Clustering Algorithm.

1.2 K-Means Clustering

1.2.1 Introduction to K-Means Clustering

The term "k-means" was first used by James MacQueen in 1967 , although the idea goes back to 1957. [6] The standard K-means algorithm was first proposed by Stuart Lloyd in 1957 as a technique for carrying out pulse-code modulation, but it was not published till 1982. K-means is a popularly used partitional clustering method in industries.

K-Means Clustering algorithm is a partitioning clustering algorithm widely used because of its ease and simplicity of implementation, which uses an iterative scheme to find a local minima. Also it is extremely efficient in terms of the computation and running time of the algorithm. The K-means algorithm is an evolutionary algorithm that gains its name from its method of operation. Its goal is to partition samples into clusters, where samples belonging to the same cluster are similar and those belonging to the other clusters are distinct. It minimises the Euclidean distance between the sensor nodes and the cluster heads or centroids of a cluster. Each centroid of a cluster is represented by the node closest to mean value of the nodes in the relevant cluster [7].

The K-means algorithm works in a way so as to create “k” partitions or disjoint sets of the data, where each set represents a cluster and $k \leq n$, where “n” represents the number of data objects. It usually classifies the entire given data into k groups, which collectively satisfy the following criteria:

(i) Each group/cluster should contain at least one object each

(ii) Each object should only belong to exactly one cluster.

1.2.2 K-Means Algorithm

The algorithm consists of two individual phases. In the first phase, K centers are randomly selected, where the value of 'k' is fixed in advance. In the next phase, each data object is assigned to the nearest center, based upon the object's proximity to the centroid of the cluster. Euclidean distance is generally considered as a parameter to determine distance between each data object and the cluster centers. All the data objects when included in different clusters, the initial step of the algorithm gets completed and an early grouping of the clusters is performed. Recalculating the average of these early-formed clusters, the centroids of the clusters are recomputed and updated. This iterative algorithm continues repeatedly until the criterion function minimises. [8-9]

The process of K-Means Algorithm is described as follows [10-11]:

1. Randomly select 'c' centroids as initial cluster heads for 'k' clusters, from the input dataset D={d_1, d_2, …d_n}.
2. Calculate euclidean distance between each data point, from the input dataset, and all 'c' centroids.
3. Assign each data object to its nearest centroid based on the minimum euclidean distance. The euclidean distance (x_i,y_i) can be calculated as follow:

$$d(x_i,y_i)=\left[\sum_{i=1}^{n} (x_i - y_i)^2\right] \quad (1.1)$$

4. For each cluster, recalculate the centroid by calculating the mean of each cluster.
5. Repeat steps from 2 to 4 until no further change in cluster centroids.

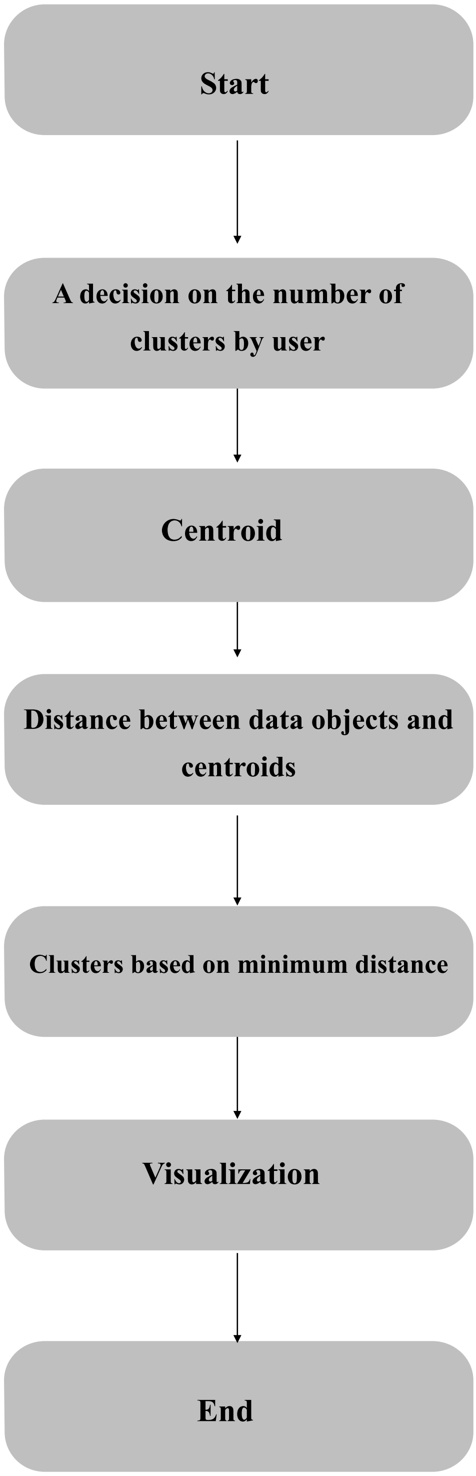

Figure 1.3 Flowchart of K-Means Algorithm Process (Source [10])

1.3 Shortcomings of Conventional K-Means Algorithm

Despite the ease of implementation, the K-Means Algorithm presents some limitations which are listed below [11]:

- K-Means Algorithm has to calculate the distance from each data object to every cluster centre in each iteration. However, by experiments we find that it is not necessary for us to calculate the distance each time. Thus it takes up a long execution time thus affecting the efficiency of clustering.
- Secondly, the clustering results are related with the input sequence of objects and the initial cluster enters which are chosen at random.
- A major problem of this algorithm is that it is sensitive to the appointing of the initial partition and hence may converge to the local optima.
- Another difficulty of clustering with K-Means is that it fails to identify clusters with large variation in sizes since original large clusters tend to split. If the clusters are of different size, densities and non globular shapes, then the resulting outputs are not optimal.
- Also, It does not work well if the clusters which are in the original data, are of different size and density.

1.4 Organization of the Thesis

This remainder of this report consists of 6 chapters, which are organized as below:

- Chapter 2: Literature review: In this chapter, study of the work that has been done regarding wireless networks, game theory and K-Means has been mentioned. Also gaps in the study have been discussed.
- Chapter 3: Dimensionality reduction in Wireless Sensor Networks: In this chapter, the concept of Principal Component Analysis and its implementation with K-Means Algorithm has been explained.
- Chapter 4: Game Theoretic K-Means: In this chapter, basics of Game Theory have been explained. Along with it, Weighted K-Means Algorithm and its implementation has been mentioned.

- Chapter 5: Cluster Validity Measures: In this chapter, the cluster validity measures like Dunn's Index and Davies-Bouldin Index have been discussed.
- Chapter 6: Results and Discussion: In this chapter, the results of the thesis have been discussed as performed in MATLAB.
- Chapter 7: Conclusion: In this chapter, the entire work has been concluded on the basis of results and observations.

CHAPTER 2

LITERATURE SURVEY

Sandra Sendra *et al.* [3] propose power saving and energy optimization techniques for wireless sensor networks. A survey of energy optimization and power saving techniques for WSNs are presented in this paper. This survey improves the techniques in use and introduces one to some of the well known methods available to save energy. These are analysed from various perspectives : transmission, routing protocols, device hardware and MAC. The main causes of energy loss in wireless sensor nodes are also discussed.

P. Kumarawadu *et al.* [4] propose algorithms for node clustering in wireless sensor networks. A complete study of clustering algorithms for WSNs are provided in this paper. These are then classified on cluster formation parameters and election criteria for cluster heads. Key design challenges and performance related issues of clustering algorithms are furthered discussed as well.

Upavan Gupta [5] proposes a Game Theoretic framework for spatial clustering on the basis of multiple conflicting criteria. A Nash-equilibrium-based methodology is used to derive solutions that are socially fair for every player. However, The clusters are updated, after each step using the K-Means algorithm, and the processe is repeated until the stopping criteria are satisfied.

Shi Na *et al.* [9] propose an improved k-means Clustering Algorithm. In this paper the standard K-means clustering algorithm is discussed and the shortcomings are analysed. An improved K-Means algorithm is proposed thus. In the improved method the repeated computation of the distance of each data object to the cluster centre is avoided, thus saving running time. The improved method in experimental results has shown that it can improve the accuracy and speed of clustering.

Kitti Koonsanit *et al.* [10] present a simple, parameter-free K-means technique for performing K-means in satellite imagery clustering applications, in order to determine the initial number of clusters with image processing algorithms, based on co-occurrence matrix method. The authors proposed a maxima technique to automate counting number of peaks in occurrence matrix as number of clusters. The parameter-free technique was tested with multispectral imagery and hyperspectral imagery. The results confirmed the effectiveness of the proposed technique.

Ahamed Shafeeq B M *et al.* [11] present a modified K-Means algorithm with the aim of improving cluster quality and to fix the optimal number of clusters. In practical scenario, it's very difficult to fix the number of clusters in advance in K-Means Algorithm. The proposed method works for both the cases. The user has the choice to either fix the number of clusters or input the minimum number of clusters that are required. In the latter case, the algorithm computes new cluster centers by incrementing the cluster counter by one in each iteration until it satisfies the validity of cluster quality. The authors show how the modified k-mean algorithm will increase the quality of clusters compared to the K-means algorithm. The modified approach assigns the data points to their appropriate class or cluster more effectively

Dr. M.P.S Bhatia *et al.* [12] propose experimental study of data clustering using k-means and modified algorithms. This paper contains the results acquired from the experimental study of the various different approaches to K-means clustering. The results are compared on different datasets using original k-Means along with some other modified algorithms that have been implemented. *S*ome performance measures are used to calculate the results, such as no. of iterations, Silhouette validity index , execution time, no. of points misclassified and accuracy.

Katsuhiro Honda *et al.* [14] extended the PCA-guided k-Means procedure to a situation in which a few observations are missing. Principal component values that can be identified using a rotated solution of cluster indicators of k-Means clustering algorithm, are estimated in an iterative process. While solving eigenvalue problem of covariance matrices, k-Means like partitions are also derived through lower rank approximation of data matrix, ignoring the missing elements. Experimental results demonstrated that the PCA-guided process is much more robust to initialization problems, despite being on iterative optimization, just as the standard k-Means procedure is.

Tomohiro Matsui *et al* [16] propose cluster validation in K-Means clustering based on PCA-guided K-Means and procrustean transformation of PC scores. In this paper a new technique is introduced by using Procrustean transformation to estimate a rotation matrix of principal component scores to be able to select from among multiple solutions, the optimal solution derived by K-Means,and calculates the deviation between re-constructed membership indicator matrix and K-Means solutions by proposing a cluster validation measure.

Shokri Z. Selim *et al.* [18] propose a generalized convergence theorem and characterization of local optimality. The paper addresses various questions related to K-means algorithm. It presents that the algorithm may fail to converge to a local minimum under certain conditions and it converges to a Kuhn-Tucker point under differentiability conditions. Eventually a means to obtain a local-minimum solution is given.

Shigen Shen *et al.* [19] carried out a survey of game theory in wireless sensor networks security. This paper mainly focuses on WSNs security by game theory. The various existing game theoretic approaches for the security of WSNs are overviewed and the pros and cons of each approach is discussed. Future research areas in WSNs security, based on game theory have also been pointed out.

Fatemeh Kazemeyni *et al.* [21] propose group selection by nodes in wireless sensor networks using coalitional game theory. An algorithm is proposed in this paper for nodes to choose the best group that are in its' signal range, by implementing coalitional game theory and determining what is beneficial to power consumption. By implementing the Maude tool, the protocol is formalized in rewriting logic. Using the Maude tool, the proposed protocol is proved correct, by trying to detect all possible scenarios of failure of the protocol. The results proved that grouping nodes is done correctly in all reachable states from a set of initial states of the model. The results showed the power efficiency to be improved significantly.

Afrand Agah *et al.* [24] propose a game theory based approach for security in wireless sensor networks. In this paper a game is defined between the sensor nodes and concentrate on three fundamental parameters: quality of security, reputation, and cooperation. More stronger cooperation between the nodes implies there is more reliable data communication between them, and as a result, better is its reputation. By instilling these factors the nodes are clustered in such a way that the payoff function of all sensor nodes are close to each other where payoff is the largest possible individual gain for each sensor according to a defined utility metric.

Shamik Sengupta *et al.* [25] apply game theory in order to solve the power control problem in a CDMA-based sensor network. They formulate a noncooperative game and analyse the existence of Nash equilibrium. With the help of the equilibrium, they devise a distributed algorithm, which helps in optimal power control and prove that the system is power stable only if the nodes have certain transmit power thresholds. The power level at which a node must transmit, so as to maximize its utility, is evaluated. They also compare the utilities when nodes are allowed to transmit with continuous and discrete power levels. They define a distortion metric, which gives a quantitative measure of having finite power levels and then find those levels that minimize the distortion. Results demonstrate that the proposed algorithm achieves the best possible payoff/utility for the sensor nodes, while consuming less power.

Ravi Kumar [30] presents a novel approach for identification of odors/gases, using game theory. To classify the sampled data into four classes, a coalitional game is modelled, in which each sensor of the array acted like a player and formed coalitions with other players. With each coalition of players a pay-off function is associated with higher pay-offs being assigned to coalitions, which maximize class separability of data. Shapley value is utilised to quantify contribution of each player. A weighting criterion for relative scaling of the samples is also proposed. The results proved that 89% of the samples were correctly identified using the technique, thus proving its efficiency.

J.Z. Huang et al. [31] propose a k-means type clustering algorithm, which can automatically calculate variable weights. The authors introduce a new step to the k-means clustering process, in order to iteratively update the relative weights, based on current partition of the data and a formula for weight calculation. The variable weights produced by the algorithm measure the priority of variables in clustering and can be used for variable selection in data mining applications, where complex and large real data are usually involved. Experimental results showed that the new algorithm outperformed the standard k-means algorithm in recovering clusters in data.

Sanghamitra Bandyopadhyay *et al.* [32] discuss the variable string length genetic algorithm (GA), which is used for developing a nonparametric clustering technique, when number of clusters is not fixed *a priori*. Chromosomes of the same population might then have different lengths as they encode different number of clusters. The authors redefine the crossover operator in order to tackle the concept of variable string length. Cluster validity index has been used as a measure of fitness of a chromosome. Performance of several cluster validity indices, such as Davies–Bouldin (DB) index, Dunn's index, and two of its generalized version, are also compared.

Dan Tudose *et al.* [34] discuss an energy consumption model for radio transceivers in wireless sensor networks, since they are subjected to severe energy consumption constraints and thus extending the sensor node battery life is of paramount importance for network autonomy. They proposed a model for estimating the energy consumption of a sensor node's radio transceiver and then evaluate its parameters for both single-hop and multi-hop wireless sensor network architectures.

Research Gaps

From the literature reported above following gaps have been identified:

- There are various shortcomings and limitations associated with the conventional K-Means Algorithm as have been mentioned in chapter 1. These limitations thus render the use of standard k-means not very beneficial for implementing an effective clustering scheme.
- There has not been any previous attempt has to combine Principal Component Analysis with a clustering paradigm.
- Cluster validity indices have not been used very often to measure the effectiveness and the measure of compactness empirically in clustering algorithms .
- No attempt has been made to quantify the contribution of sensors in the evolution of clusters during the execution of the K-Means Algorithm.

CHAPTER 3

DIMENSIONALITY REDUCTION IN WIRELESS SENSOR NETWORKS

3.1 Principal Component Analysis

Principal Component Analysis (PCA) is a popular technique for dimensionality reduction and feature extraction. It is a widely used statistical technique which uses orthogonal transformation for reduction of dimensionality of a data set consisting of a large number of correlated variables while retaining maximum of the variation present in the data set. The necessary operations are then carried out in the new feature space, where the data samples are more separable. This reduction is obtained by converting the set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables, which are known as principal components. The principal components have a number of less than or equal to the number of original variables. Cluster formations are then carried out by using the K-Means Algorithm, in which a continuous solution of membership indicators in K-Means are principal components in PCA. [13]

PCA is designed to reduce the dimensionality of a large data matrix while retaining most of the original variability in the data. This is achieved by converting a set of observations, which consists of possibly correlated variables into a set of linearly uncorrelated variables, which are called principal components. The first principal component represents the maximum variation in the original data. Then, every succeeding component accounts for as much of the remaining variation, subject to being uncorrelated with the previous component [14].

The covariance or correlation matrix that is derived from the data matrix plays an important role in PCA in order to calculate its eigenvalues and eigenvectors to obtain the associated components, which account for most of the variations in the data.

Feature extraction or feature selection is a necessity requirement in statistical pattern classification. Feature selection often requires that the data space is transformed into a feature space in such a way that the transformed data has the same dimension as the input data but the data space can be represented by a fewer number of selected features instead of the complete feature space, yet containing the most important information from the input data. [15]

Say, we have a problem where the input data that is nonlinearly separable is of a very high dimensionality. According to Cover theorem for separability of patterns [15], the mapping of data nonlinearly to a higher dimension classifies a nonlinear data linearly. However, if we map a data that already is of such a large dimension, we will require a massive neural network with a large number of hidden layers and neurons. Therefore it is required to decrease the dimensionality of the input data. We look for the redundancy in the 'm' dimensional data, so that we can select 'm_o' dimensions such that (m_o<m), taking into consideration that the dimensions that we are discarding must contain unimportant data. Truncating the data causes error. The mean square error hence obtained will be equal to the sum of the variance of the elements that are eliminated. We require the mean square error (MSE) to be as small as possible. Hence the data with the maximum information is restored.

In the pattern space of raw data, we cannot decide what data points hold the required information hence it is to be decided that which data points should be truncated. Therefore, we map the vector into another space, which we obtain by transforming it using a transformation matrix **T**. The transformation matrix **T** is an m by m dimensional matrix and we pre multiply it with the input vector, to form a new vector such that it will be easier for us to decide what data points should we discard so as to maximise the rate of decrease of variance.

$$\dot{x} = (x_1, x_2, x_3 \ldots x_{m_0}\, x_{m_{0+1}}) \tag{3.1}$$

$$\dot{x}_{new} = T\,\dot{x} \tag{3.2}$$

It is assumed that the vector $\dot{x}$ is a zero mean random vector and if it is not a zero mean vector, the mean can be subtracted from all the elements of the vector in order to get a zero mean vector. Consider an m-dimensional unit vector q, and it is required to project the vector $\dot{x}$ onto that vector in order to have a scalar value of projection as A.

$$E(\dot{x}) = 0 \tag{3.3}$$

$$\|\dot{q}\| = 1 \tag{3.4}$$

$$A = \dot{x}^T \dot{q} \tag{3.5}$$

Consider the expectation and variance of the vector in the projected space. Now since the unit vector is not a random vector, it comes out of the expectation operator making the expectation of the projection zero.

$$E(A) = E(\dot{q}^T \dot{x}) = \dot{q}\, E(\dot{x}) = 0 \tag{3.6}$$

Thereafter we need to find the variance of the projection *A*, which is denoted as $\sigma^{2.}$

$$\sigma^2 = E(A^2) = E[(\dot{q}^T \dot{x})(\dot{x}^T \dot{q})] = \dot{q}^T E(\dot{x}^T \dot{x})\dot{q} = \dot{q}^T \dot{R}\dot{q} \tag{3.7}$$

Where $\dot{R}$ is a correlation matrix of dimension m by m. Since it is a matrix of two same vectors multiplied after a transpose, they will be symmetric, which implies

$$\dot{R}T = \dot{R} \tag{3.8}$$

And it can also be shown that if a and b are any m by 1 vectors, thus

$$\dot{a}^T \dot{R}\dot{b} = \dot{b}^T \dot{R}\dot{a} \tag{3.9}$$

The variance σ2 has to be minimised, but the value of $\dot{R}$ cannot be controlled, as it depends upon the input vector $\dot{x}$, which is a random variable. Thus, a probe is used in order to calculate the minimum variance. Variance of the projection A is a function of the vector *q* and hence may be called as a variance probe.

$$\varphi(\dot{q}) = \sigma^2 = \dot{q}^T \dot{R}\dot{q} \tag{3.10}$$

Next, those unit vectors have to be found such that the extreme values can be calculated or the local minima can be found, having the constraint on the vectors q of having a Euclidean norm constraint. This problem can be solved by considering the eigen values of the correlation matrix Ṙ. If we have a unit vector q such that there is minima at the variance probe, then a small change or perturbation in the vector will cause no change in the value of the variance probe, such as

$$\varphi(\dot{q} + \delta\dot{q}) = \varphi(\dot{q}) \tag{3.11}$$

$$\varphi(\dot{q} + \delta\dot{q}) = (\dot{q} + \delta\dot{q})^{T}\dot{R}(\dot{q} + \delta\dot{q}) = \dot{q}^{T}\dot{R}\dot{q} + 2\delta\dot{q}^{T}\dot{R}\dot{q} + \delta\dot{q}^{T}\dot{R}\delta\dot{q} \tag{3.12}$$

Using the property in equation 3.9 in equation 3.11 we get the equation 3.12. It can be observed that the last term is negligible and hence can be neglected. Using equation 3.10 in equation 3.12 we get the following results:

$$\varphi(\dot{q} + \delta\dot{q}) = \varphi(\dot{q}) + 2\delta\dot{q}^{T}\dot{R}\dot{q} \tag{3.13}$$

But $\varphi(\dot{q} + \delta\dot{q}) = \varphi(\dot{q})$ to the first order of approximation, hence we can say

$$\delta\dot{q}^{T}\dot{R}\dot{q} = 0 \tag{3.14}$$

We have restricted ourselves to only those perturbations in which the euclidean norm remains to unity such that,

$$\|\dot{q} + \delta\dot{q}\| = 1 \tag{3.15}$$

$$(\dot{q} + \delta\dot{q})^{T}(\dot{q} + \delta\dot{q}) = 1$$

$$\dot{q}^{T}\dot{q} + 2\delta\dot{q}^{T} + \delta\dot{q}^{T}\delta q = 1$$

Taking the last term s negligible and using the result of equation 3.4, we can say that

$$\delta\dot{q}^{T}\dot{q} = 0 \tag{3.16}$$

The elements of the vector q are dimension less, hence if we have to combine equation 3.14 and equation 3.15 we require a scalar of size m by m such that,

$$\delta\dot{q}^{T}\dot{R}\dot{q} - \lambda\delta\dot{q}^{T} = 0 \tag{3.17}$$

Or equivalently,

$$\delta\dot{q}^{T}(\dot{R}\dot{q} - \lambda\dot{q}) = 0$$

$$\dot{R}\dot{q} - \lambda\dot{q} = 0 \tag{3.18}$$

The equation 3.17 can be seen as an eigen value formulation. The problem has non-trivial solution for some particular values of λ and those values are called the eigen values of the correlation matrix $\dot{R}$. Particular to those values of λ we have particular values of vector q and these values are called the eigen vectors. As the correlation matrix is symmetric the eigen values are non negative and real, and assuming the eigen values are unique the eigen vectors are also different. If the eigen values and the corresponding eigen vectors are arranged in matrix form it can be expressed in following form

$$\dot{R}\dot{q}_j = \lambda_j \dot{q}_j \qquad for\ j = 1,2,...m \tag{3.19}$$

The corresponding eigen values are arranged in the decreasing order as

$$\lambda_1 > \lambda_2 > \ ...\ > \lambda_j > \ ...\ > \lambda_m \tag{3.20}$$

Such that λ1 is λmax. Let the resulting eigen vectors be constructed by an m by m matrix.

$$Q = [q_1,\ q_2, ...q_j...q_m] \tag{3.21}$$

Hence we can combine the equation 3.18 with 3.19 and 3.20 as

$$\dot{R}\dot{Q} = \dot{Q}\Lambda \tag{3.22}$$

Where $\wedge$ is a diagonal matrix with diagonal elements the eigen values of the correlation matrix R.

$$\wedge = diag\ [\ \lambda_1 > \lambda_2 > \ ...\ > \lambda_j > \ ...\ > \lambda_m\ ] \tag{3.23}$$

The matrix Q is a unitary matrix, such that the column vectors of Q satisfy the orthonormality condition.

$$q_i^T\ q_j = \begin{cases} 1, & j=1 \\ 0, & j\neq 1 \end{cases} \tag{3.24}$$

The above equation gives us different eigen values. Hence, we may write

$$\dot{Q}^T\ \dot{Q} = 1 \tag{3.25}$$

From which we can deduce that the inverse of the Q matrix is the transpose of the matrix itself.

$$\dot{Q}^T = \dot{Q}^{-1} \tag{3.26}$$

It can also be written in the form,

$$q_j^T \dot{R} q_k = \begin{cases} \lambda_k, & k=j \\ 0, & k \neq j \end{cases} \tag{3.27}$$

PCA and the decomposition of matrix R is one and the same thing for solving the probe function. That shows the variance probe values and the eigen values are one and the same thing.

$$\varphi(\dot{q}) = \lambda_j \qquad j = 1, 2, ..., m \tag{3.28}$$

3.2 Clustering using PCA

Cluster analysis is used in several disciplines like biology, marketing and hydrology to partition observations of similar patterns to the same cluster and dissimilar patterns to different cluster. Principal component analysis is a reduction dimension technique that is used often as a pre-processing method in order to guide the process of grouping data in so as to improve the accuracy and efficiency of cluster solutions. The basic idea of Principal Component Analysis is to reduce the dimensionality of a data set that consists of a large number of interrelated variables, while retaining as much variation present in the data set, as possible. Generally, Pearson correlation is used in Principal Component Analysis in order to provide an eigen analysis to obtain the associated components which account for most of the variations in the data. PCA basically constructs a set of uncorrelated directions, which are ordered by their variance. In a lot of cases, directions with the maximum variance are the ones that are most relevant to the clustering. Also, PCA is the basis for various variable selection techniques, such as variables that have a large component in the low variance directions are discarded. PCA is also very useful as a visualisation tool as it can provide us with a low dimensional summary of the entire data, perform quality control and help detect outliers. A large number of attributes or dimensions of a particular data set can create problems in data clustering. The first problem is that the efficiency of some distance measures like Euclidean Distance become meaningless. The second problem is regarding data visualisation. Also, linear dimensionality reduction techniques have been proven to be ineffective in improving data clustering results when number of dimensions of data set are more than 30. Thus, applying PCA on the data set initially leads to better clustering of the data than by performing clustering on the data set without PCA [16].

3.3 K-Means with PCA

Recently, Ding and He [17] had proposed a PCA-guided K-Means process, in which a continuous solution of the cluster membership indicators in k-Means can be derived in a Principal Component Analysis-guided manner. In the procedure, the within-cluster-errors criterion used in K-Means is redefined by a centroid-less formulation and the relaxed indicator vectors, which represent cluster memberships, are calculated as principal components in PCA. The indicator vectors are then identified with the eigenvectors of a within-cluster or inner-product similarity matrix. However, the cluster indicator that is derived is a rotated solution, and the rotation matrix cannot be estimated explicitly. To visually access the cluster structures, an approach such as visualisation by the ordering of samples in connectivity matrices is applied.

As mentioned earlier, the conventional K-Means algorithm typically converges to a local minimum of mean squared error or MSE [18]. Even though the algorithm is quite simple and useful in a lot of cases, it's possible to obtain multiple solutions with different initialisations. The K-Means algorithm is usually initialised with a randomly chosen partition. However, in this sense, there is no guarantee of convergence to the global optima. PCA is applied in the estimation of the suboptimal initial partition. The nonlinear principal components are constructed by performing PCA in the higher dimensional feature space. Application of dynamic programming in each nonlinear principal direction leads to an optimal partition of data samples in the projection subspace. This strategy leads to **a** smaller distortion between the suboptimal initial partition and the globally optimal solution.

In the PCA-guided K-Means, the objective function of K-Means clustering is redefined by a centroid-less formulation and the relaxed indicator vector representing cluster memberships are calculated by utilising PCA, in which the indicator vectors may be identified with the eigenvectors of a within-cluster or inner-product similarity matrix, i.e., a continuous solution of the cluster membership indicators in K-Means is principal components in PCA. Multiple clusters are extracted in a batch process by calculating multiple eigenvectors in a single step while cluster structure can be captured only after an orthogonal transformation.

CHAPTER 4

GAME THEORETIC K-MEANS

4.1 Introduction to Game Theory

4.1.1 Basics

Game theory is increasingly attracting more attention as a mechanism to solve various problems in WSNs. Game theory can be defined as a theory of decision making, under certain conditions of interdependence and uncertainty. It is a mathematical model, which describes the phenomenon of cooperation and conflict between intelligent and rational decision makers. Game theory offers models for distributed allocation of resources and thus provides a suitable way for exploring several characteristics of wireless sensor networks. Game Theory is mostly used extensively in the context of political science, economics, and psychology. Initially, it just addressed to zero-sum games, such that the gains of one player would equal exactly to the net losses of the other player(s). However, today, game theory applies to a variety of behavioural relation applications, for both non-humans (e.g. computers) and humans. It is a powerful mathematical tool to predict and analyse the behaviour of selfish and rational entities.

Due to the interesting and sometimes unexpected results that game theory produces, its popularity spread to the filed of networking and communications technology. In recent years, the theory has been proven to be highly useful in the field of wireless sensor networks. It can be used mathematically to analyse system operations in self-organizing and decentralized networks. Game theory basically describes the behaviour of players, equivalently sensors in WSNs, in a game. These players (sensors) could be either cooperative or non-cooperative, while striving for maximisation of the outcomes from their game. In this respect, wireless sensors manage their operations in regard with power resources devoted to communicating and sensing amongst themselves, and also with a global controller, so that the assigned task can be completes effectively as desired [19].

Generally, a game consists of [20]:

(1) a set of players,

(2) a set of strategies for every player, and

(3) a set of corresponding utility functions.

The general form game of a WSN of n sensor nodes is represented by a three-tuple

$$G = \langle N, S, U \rangle \tag{4.1}$$

Here, G is a particular game, where $N = \{n_1, n_2, \ldots, n_n\}$ is a finite set of the sensor nodes or players. $S = \{S_1, S_2, \ldots, S_n\}$ is the strategy space of the sensor node i, which is represented by S_i ($i = 1, 2, \ldots, n$). $U = \{u_1, u_2, \ldots, u_n\}$ is the corresponding payoff or utility function of node i represented by u_i ($i = 1, 2, \ldots, n$). u_i is the utility value of each node received at the end of an action.

A strategy for a player is a complete plan of actions in all possible situations in the game. To maximise their consequences the players try to act selfishly according to their preferences. The payoff functions have to be formulated in a way that will help node i to select a strategy S_i that represents the best response to the strategies selected by the other n-1 nodes.

A utility function describing player preferences for a given player assigns a number for every possible outcome of the game with the property that a higher number implies that the outcome is more preferred.

4.1.2 Types of Game Theory

1. Cooperative Game Theory [21]

To reduce the whole WSN's energy consumption and prolong its lifetime, some of these nodes will form a coalition by cooperation. Cooperative game theory is also known as coalitional game theory. For WSNs to comply with cooperative game theory, cooperating groups are formed and strategies are formulated to maximize their own groups' utility by the players. A reduction of power consumption in WSN by forming coalitions is possible due to Coalitional game theory.

We can group the nodes in two ways for different applications: (1) All the sensor nodes could be placed in the same group having similar sensed data, for example, sensing application. (2) The sensor nodes with the shorter distances between them could compose a same group, such as, sending data to the sink from a source node.

2. Non-Cooperative Game Theory [22]

Non-cooperative game theory studies strategies between interactions among competing players. In a non-cooperative game Each player is selfish but rational. Non-cooperative game theory uses a utility function to find the Nash Equilibrium. The main application of Non-cooperative game theory is in spectrum sharing in cognitive radio, power control, congestion control, distributed resource allocation and many others.

3. Repeated Game Theory [23]

In this, a player has to take into account the impact of its current action on the future actions of others. A penalizing mechanism based on repeated game theory to prevent the non-cooperative selfish behaviour of decreasing the contention window without permission was proposed. A repeated game theoretic model is based on cooperative packet forwarding under the conditions of selfish and rational nodes for improving energy efficiency and sensor networks payoff. By implementing the punishment mechanism, the NE can be propelled by this repeated game model and decrease the defection possibility of selfish nodes.

4.1.3 Game Theory in Applications of WSNs

1. Clustering of the sensor nodes [5]

A unique property of game theory is social equity or social fairness , which ensures that each player in the game (i.e. every sensor node in the network) is satisfied and the overall goals are reached. In a multi- objective clustering problem with compaction and equipartitioning as the objectives, other optimization methods intend to identify solutions targeting the overall system optimization, rather than the optimization of individual objectives. Instead, a game theoretic modeling ensures that each metric is optimized with respect to the other metrics.

2. Game Theory for WSN Security [24]

Due to the limited capabilities of sensor nodes in terms of energy, communication and computation, providing security to WSN has increasingly become one of the most interesting areas of research in recent years. WSN security is a primarily important and critical issue before WSN can be widely used. There generally exist two mechanisms of intrusion, prevention and detection in WSN security. GT provides a mathematical method of analyzing and for modeling WSN security problems, for it considers scenarios where multiple players with contradictory objectives compete with each other.

3. Game Theory for Wireless Network Management in WSN [19]

The design of a wireless network and optimization of its performance is a non-trivial and complicated process. The WSN has to fulfill straight requirements imposed from a set of operation goals. Game theory thus plays a supportive and critical role in designing and operating a WSN. When the WSN is put into operation, it is always susceptible that the network may be attacked by hackers where data could be intercepted and retrieved illegally. To this end, the WSN has to be designed for deny of service (DoS) prevention and the incorporation of intrusion detection capability.

4. Power Management—Energy Saving and Power Control [25]

In WSN, energy is a limited resource and must be used judiciously. Currently, the energy problem remains one of the major obstacles somehow preventing the complete exploitation of WSN technology. Energy saving and power control strategies should be devised at sensor nodes as well as in the network to prolong the network lifetime. In practice, the Game-theoretical Total Link (GTL) algorithm which sets each node's energy range is usually better than Critical Transmitting Range (CTR) in energy saving according to the topological changing.

5. Routing Protocol Designs [26]

Routing protocol design is concerned with the efficiency that data, acquired by a sensor node, can be channeled through other nodes and either directly or indirectly reach the base station for analyze. Protocols used in WSN, on the other hand, have an integrated effect on the power consumed in radio transmission as well as the economical use of radio spectrum. Game theory can be satisfactorily employed in the design of routing protocols that it is able to account for difficulties in node behaviour, energy balance, dynamic route allocation and others.

6. Target Tracking [27]

A method for target tracking based on multi-agent and game theory was proposed. When a target appears in the sensing field of WSN, sensor nodes begin to form coalition dynamically and then they start to negotiate with game theory. Coalition is formed to track it with the target moving. Utilising multi-agent method and game theory in WSN enables nodes to perform tasks coordinately to achieve some desired objectives.

7. Packet Forwarding [28]

In WSNs, selfish nodes refuse to forward packets for other nodes in order to save energy which causes the network to enter into a faulty state. At the same time, some nodes may be malicious, whose aims are to damage the network. The fault tolerance and security problem can be modelled as a non-cooperative game in which each player maximises its own utility function.

4.2 Role of Coalition Game Theory in Clustering

A coalition or cooperative game theoretical approach is applied for the optimisation of clusters. Non-cooperative games assume competition between individual players for maximisation of gains whereas cooperative game theory is based upon competition between different "coalitions" of players rather than between individual players. In cooperative games, groups of players are formed, called coalitions. Players try to find a coalition to strengthen their position in the game and make an agreement to act as a unified simple entity. Coalitional games have proved beneficial to design robust, fair, and efficient cooperation strategies in communication networks. In a general coalitional game (N, v) with N players, the coalition value or utility function of a coalition is determined by a characteristic function v : $2^N \rightarrow R$ which applies to coalitions of players. The core of the coalitional game (N,v) guarantees that none of the players has any incentive to leave N to form another coalition on their own. [29]

Cooperative game theoretic approach has been chosen over non cooperative approach because the former reduces and balances the energy consumption as the neighbouring nodes can cooperate and choose a more suitable and shorter path to reach the cluster head or destination, hence increasing the energy efficiency of the system. In wireless sensor networks, sensor nodes are selfish and self organised and aim at maximising their own interest and having local information. As an important generalisation of Nash Equilibrium, Coalition Game Theory can directly consider the ability of players to coordinate actions. Thus, there is a higher degree of cooperation and better solution as compared to non-cooperative Nash Equilibrium.

The initial clusters identified using K-Means are models as players and resources, different combinations of data objects requested by each player from different resources are modelled as strategies, and a function of competing objectives: compaction and equipartition, as the payoff.

4.3 Introduction to Shapley Value

Shapley value is derived from the solution concept in cooperative game theory which was defined by Shapley. It is one of the most important solution concepts in cooperative game theory and a representative single-valued solution concept in the theory of cooperative games . An agent's Shapley value gives an indication of its prospects of playing the game in cooperative game theory. It is useful when there is a need to allocate the worth that a set of players can achieve if they agree to cooperate. The Shapley value was defined for TU games and NTU games in regard to conflicts among players.

A coalition game is a type of cooperative game defined by a pair (S, V) where $S = \{1,2,\ldots,n\}$ is the set of all players, and V (ξ) is the objective function, quantifying the contribution of coalition ξ towards the total benefits achieved for every $\xi \subseteq S$. Thus, the objective function $V(\xi)$ is also knows as "worth of coalition ξ ". In the next step, the contribution of each player to the game is represented by constructing a value function of "payoff" which assigns a real value to every individual player, by taking into account the "marginal contributions" of that player in all the coalitions of which he was a part. Shapley analysis is a popular and effective method of assigning a real value to a player in a n player cooperative game. The Shapley value is defined as follows. [30]

Let the "marginal contribution" of player i to a coalition, with $i \notin \xi$, be:

$$\Delta_i(\xi) = V(\xi \cup \{i\}) - V(\xi) \tag{4.2}$$

Then, the Shapley value is defined by the payoff

$$\phi_i\,(V) = \frac{1}{n!}\,\Sigma_{\mu\in\pi}\;\Delta_i\,(\xi_i\,(\mu)) \tag{4.3}$$

where, ∏ is the set of permutations over S, and $\xi_i\ (\mu)$ is the set of players appearing before the i^{th} player in permutation μ. The Shapley value assigned to a player is nothing but a weighted mean of its marginal contributions over all possible subsets of players.

4.4 Weighted K-Means Algorithm

It is necessary to introduce weights for K-Means Algorithm to decrease the effects of irrelevant attributes while clustering and ensure maximization of connectedness. Thus. to generate clusters of appropriate population, a weighting system is added to the general K-Means Algorithm. Every element is associated with a real valued weight, representing its mass or importance. Every element of the data set is associated with a real valued weight, representing its mass or importance. Adding relative weights associated with the data elements generalizes the notion of element duplication. The weighted setting enables a convenient method for prioritizing certain data elements. In sum, it is essential to assign larger weights to the sensor nodes with higher frequency of usage and smaller weights to less frequently used nodes to differentiate different clusters, or partition more frequently used nodes into the same cluster [31].

Steps:

Input: a set of n data points, and the number of clusters K

Output: centroid of the K clusters

1. Using the standard K-means algorithm, assign points to clusters and centers to appropriate positions.
2. Repeat

Assign each data point to its nearest cluster centre according to its relative weight

3. Recompute the cluster centres using the current cluster memberships and weights
4. Until there is no further change in the assignment of data points to new cluster centers.

By adjusting the weights, we are able to control the growth or decay of the clusters. If the weight of a cluster increases, data points are more likely to be grouped in other clusters. Similarly, decreasing the weight helps to increase the population of a cluster. Thus the weight function is crucial in the performance of the algorithm.

4.5 Relatively Weighting the sensor nodes

In accordance with the Weighted K-Means Algorithm, it is required that each data sample from the original data set be weighted by the appropriate coefficient, for achievement of optimal clustering. Let a_1, a_2 …a_n represent the relative weights corresponding to n-dimensions of Class 1 which is formed as a result after standard K-Means Clustering . Then, these coefficients are given by [30]:

$$a_n = \left| \frac{\phi_{s_n}(V)}{\sum_{n=1}^{3} \phi_{s_n}(V)} \right| \tag{4.4}$$

Similarly, the coefficients b_1, b_2 …b_n represent the relative weights corresponding to n-dimensional data belonging to Class 2, and are given by:

$$b_n = \left| \frac{\phi_{sn}(V)}{\sum_{n=1}^{3} \phi_{sn}(V)} \right| \tag{4.5}$$

And similarly, the coefficients c_1, c_2 …c_n represent the relative weights corresponding to n-dimensional data belonging to Class 3, and are given by:

$$c_n = \left| \frac{\phi_{sn}(V)}{\sum_{n=1}^{3} \phi_{sn}(V)} \right| \tag{4.6}$$

CHAPTER 5

CLUSTER VALIDITY MEASURES

Clustering is more of an unsupervised technique hence the evaluation of the clustering algorithms is of great importance. In the clustering process there are no predefined classes therefore it is difficult to find an appropriate metric for measuring if the evolving cluster configuration is acceptable or not. The result of a clustering algorithm can be very different from each other on the same data set as the other input parameters of an algorithm can extremely modify the behavior and execution of the algorithm. The aim of the cluster validity is to find the partitioning that best fits the underlying data. Usually 2D data sets are used for evaluating clustering algorithms as the reader easily can verify the result. But in case of high dimensional data the visualization and visual validation is not a trivial tasks therefore some formal methods are needed. The process of evaluating the results of a clustering algorithm is called cluster validity assessment.

Distance Parameters: In clustering algorithms distance measure refers to the measure of similarity between different data patterns. Distance parameters can be classified as Inter cluster distance and Intra cluster distance.

- *Intra Cluster distance:* The intra cluster distance can be defined as the distance between the members of a cluster and it should be minimum possible.
- *Inter Cluster distance:* The clusters themselves should be widely separated. Inter cluster distance can be defines as the distance between the centroids of all possible pairs of clusters and it should be maximum possible.

5.1 Dunn's Index

The Dunn's Index (DI) is a metric, which is used to evaluate the effectiveness of a clustering algorithm. The numeric result is based on the clustered nodes or data itself. It helps identifying the clusters that are compact with little variance between the cluster members, and well separated so that the centroids of the clusters are sufficiently far apart. For a given set of clusters, the higher the Dunn's index better is the clustering [32].

Let S and T be two non empty sets of R^N. Then the diameter Δ of S and set distance ∂ between S and T are

$\Delta(S) = max\ \{d(x,y)\}$, where $x,y \in S$

$\partial\ (S,T) = min\ \{d(x,y)\}$, where $x \in S,\ y \in T$

where $d\{x,y\}$ is the distance between points x and y. For any partition, Dunn defined the following index:

$$v_D = min_{1 \leq i \leq K} \left\{ min_{1 \leq i \leq K,\ j \neq i} \left\{ \frac{\delta\left(C_i, C_j\right)}{max_{1 \leq k \leq K}\left\{\Delta\left(C_k\right)\right\}} \right\} \right\} \quad (5.1)$$

Larger values of the Dunn's Index correspond to the good clusters, and number of clusters that maximizes the index is taken as the optimal number of clusters.

5.2 Davies-Bouldin Index

Clustering of wireless sensor networks into different clusters or classes with the objective of maximum connectedness ensures a more efficient network. This is done by minimising the inter-cluster and intra-cluster distances. Davies-Bouldin (DB) Index is a clustering validity measure which evaluates the goodness of clustering results [33]. This performance index is a function of the ratio of the sum of the within cluster scatter corresponding to intra cluster distance to the inter cluster distance. The scatter within the ith cluster is formulated as [32]

$$S_{i,q} = \left(\frac{1}{|c|_i} \sum_{x \in c_i} \left\{ ||x - z_i||^q \right\} \right)^{\frac{1}{q}} \quad (5.2)$$

And the inter cluster distance between and is formulated as

$$d_{ij,t} = ||z_i - z_j||_t \quad (5.3)$$

$S_{i,q}$ is the q[th] root of the q[th] moment of the $C_{i,j}$ points in cluster C_i with respect to their mean z_i, and is a measure of the dispersion of the points in the cluster. Specifically $S_{i,1}$ used, is the average Euclidean distance of the vectors in class i to centroid of the class i. $d_{ij,t}$ represents the Minkowski distance of order t between the centroids z_i and z_j that characterize clusters C_i and C_j []. Subsequently

$$R_{i,qt} = \max_{j,j \neq i} \left\{ \frac{S_{i,q} + S_{j,q}}{d_{ij,t}} \right\} \tag{5.4}$$

The Davies–Bouldin index is then formulated as

$$DB = \frac{1}{K} \sum_{i=1}^{K} R_{i,qt} \tag{5.5}$$

The objective is to minimize the DB index for achieving proper clustering. Therefore, lesser the value of DB index, more optimal is the clustering.

5.3 Custer Validity using First Order Radio Communication Model

There are several methods which can be used to judge the effectiveness of a clustering algorithm. One of these methods is calculation of the average energy dissipated to send the gathered information by the sensor nodes in the WSN. Lower the mean dissipated energy, better is the effectiveness of the clustering algorithm. The first order radio model is used to calculate this energy.

A lot of research is ongoing in the area of low-energy radio for integrated circuits and is motivated mainly by the mobile and embedded market applications. Different types of assumptions about the radio characteristics, which usually include energy dissipation in the transmitting and receiving modes, could change the advantages of different algorithms. In [34], the authors proposed an energy consumption model for radio transceivers, designed specifically for WSNs. The main aim is to estimate the energy needed to send a package of n bits of data from the transmitter to the receiver, as in Figure 5.1.

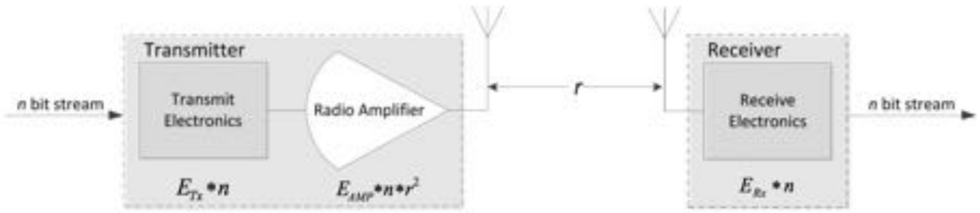

Figure 5.1 Radio model for the transmission of n bits of information (Source [34])

In order to transmit a package of n bits at a distance r, radio transmitter will consume following amount of energy:

$$E_{Tx}(n,r) = E_{tc}(n) + E_{amp}(n,r) \qquad (5.6)$$

where $E_{tc}(n)$ is the energy that a radio circuitry needs to expend in order to process n bits, and $E_{amp}(n,r)$ is the energy needed by the radio amplifier circuit to send n bits at r meters.

We can further refine (1) by elaborating on the formula for $E_{amp}(n,r)$:

$$E_{Tx}(n,r) = E_{tc}(n)+E_{amp}(n,r) = n\cdot E_{trans} + n\cdot\varepsilon_{amp}\cdot r^{\gamma} \qquad (5.7)$$

where E_{trans} is the amount of energy needed to process a single bit using the radio transmission circuits, ε_{amp} is the transceiver's energy dissipation and γ represents the path loss exponent.

Path loss exponents are linked to a medium of propagation [35] and usually range from 2 to 4, where 2 is the path loss of free space propagation and 4 is the path loss exponent for lossy environments such as buildings or stadiums.

CHAPTER 6

RESULTS AND DISCUSSION

The entire procedure has been implemented using MATLAB R2013a software and the dataset used for the same is three-dimensional and is randomly generated. The data represents empirical values of the information collected by the sensor nodes in a wireless sensor network at a given location. The data set consists of 100 three-dimensional data points. The results of the two methods have been discussed as follows:

6.1 K-Means Algorithm

As stated in the K-Means Algorithm, centroids are initially randomly selected and then updated during the algorithm.Therefore, ’k’ centers are initialised and chosen randomly as shown in Figure 6.1. In accordance to the Game Theoretic approach, three players have been assumed for the clustering problem and thus, the value of ‘k’ has been fixed at 3 and thus the data set will be grouped in three clusters or classes as a result.

K-Means Algorithm is then applied to the given data. Since the centroids are updated after every iteration of the algorithm, thus 16 iterations of the algorithm have been fixed, so as to stabilise the centroid updation by the 16^{th} iteration. The output after the complete K-Means Algorithm is shown in Figure 6.2. Also, after each iteration, the DB Index of the cluster is calculated as stated in Chapter 5. These DB index values are shown graphically in Figure 6.6. As we can see, since we chose the value of k=3, thus three distinct classes are formed in the output, each class having its respective cluster head.

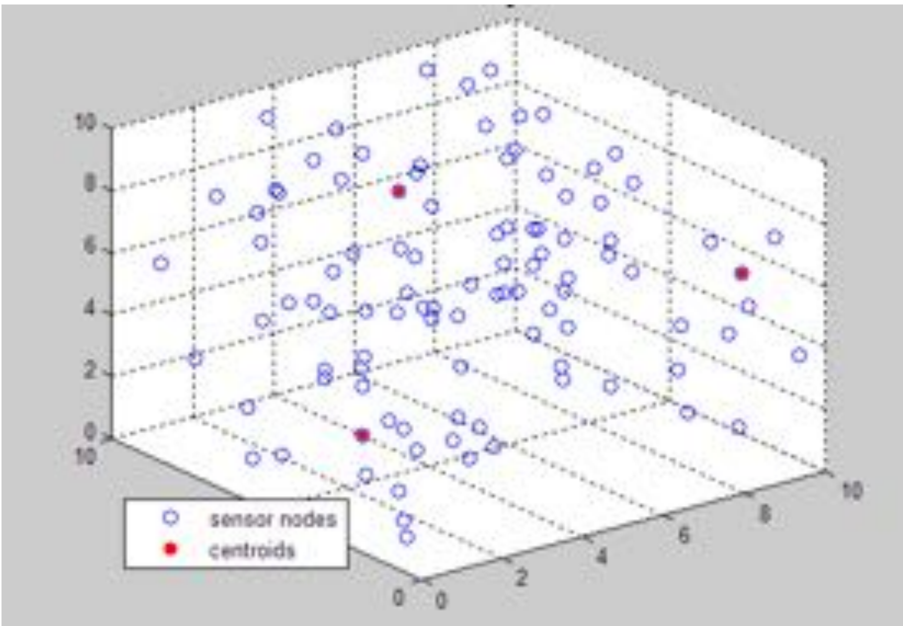

Figure 6.1 3-D Scatter Plot of Original Data

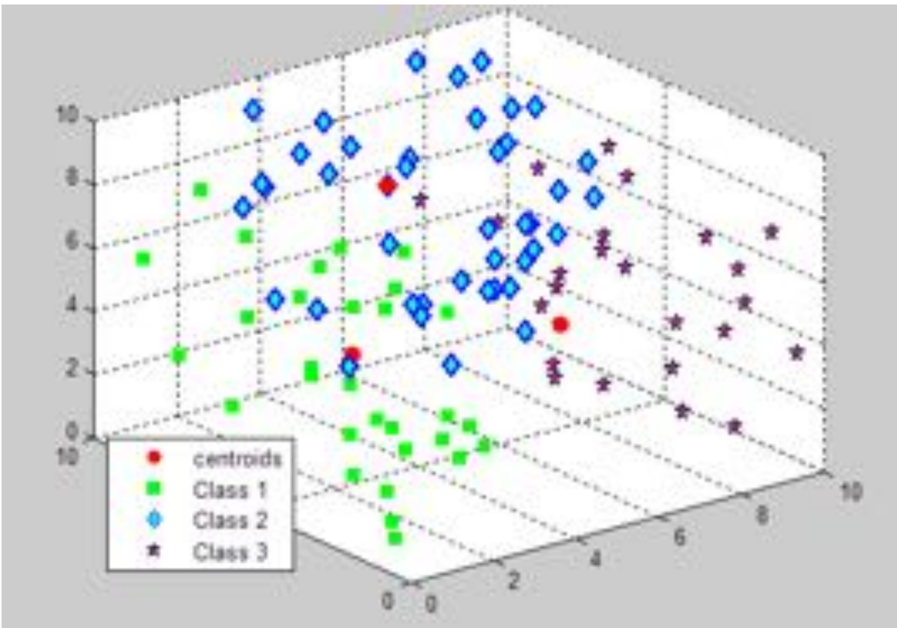

Figure 6.2 3-D Scatter plot of data after regular K-Means Clustering

6.2 PCA based K-Means Algorithm

Furthermore, we have proved that PCA based K-Means Algorithm leads to better and more optimum data clustering than regular K-Means Clustering. For this, Principal Component Analysis is performed on the same data set as shown in Figure 6.1. The output is shown in Figure 6.3.

Thereafter, K-Means Algorithm is applied to the data altered by applying PCA. Three centroids are randomly initiated in order to form three clusters. 16 iterations of the K-Means Algorithm are applied in both the cases: to the original data, and the data altered by PCA. The output shows the three clusters that are formed, as shown in Figure 6.4 and Figure 6.4.

To evaluate the validity of clusters, DB Index values for both the cases have been calculated after each iteration of each algorithm. The lower the value of the Davies-Bouldin Index, more optimal and better is the clustering. The comparison between values of DB Index for each iteration are shown in Figure 6.5. Another data set is taken and the entire procedure is repeated in order to see the trend in the comparison between the DB Index values for both the cases. This plot of comparison, for another randomly generated data set is shown in Figure 6.6. It is seen that the final stabilised value of DB Index is lesser in the case where PCA based K-Means Algorithm is used, as compared to the regular K-Means Algorithm. The empirical values of the DB Index values, calculated after each iteration of regular K-Means Clustering and PCA based K-Means Clustering are shown in Table 6.1. It can be observed that after an average of 5 iterations, the DB Index value for each case reaches a stable value.

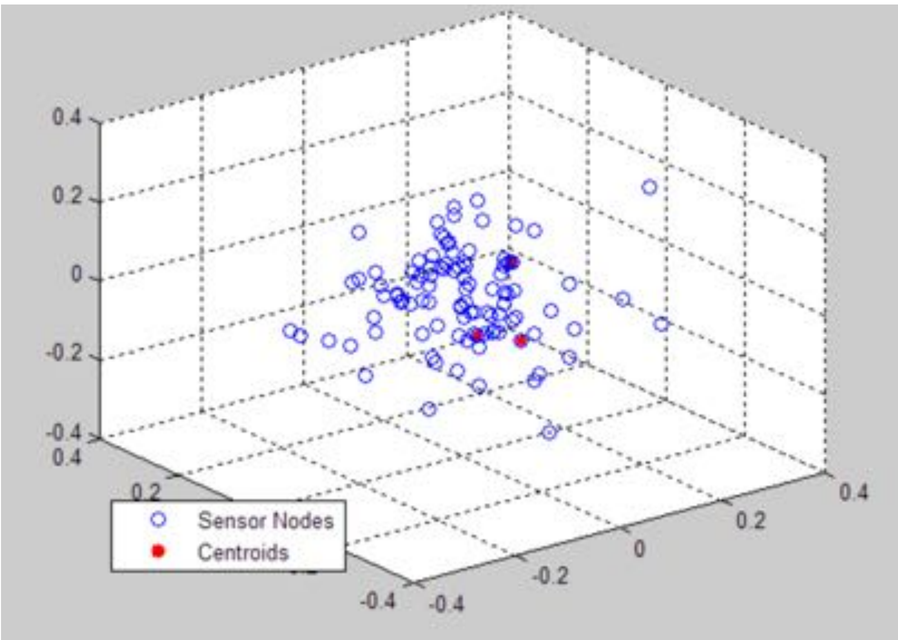

Figure 6.3 3-D Scatter plot of data after applying PCA

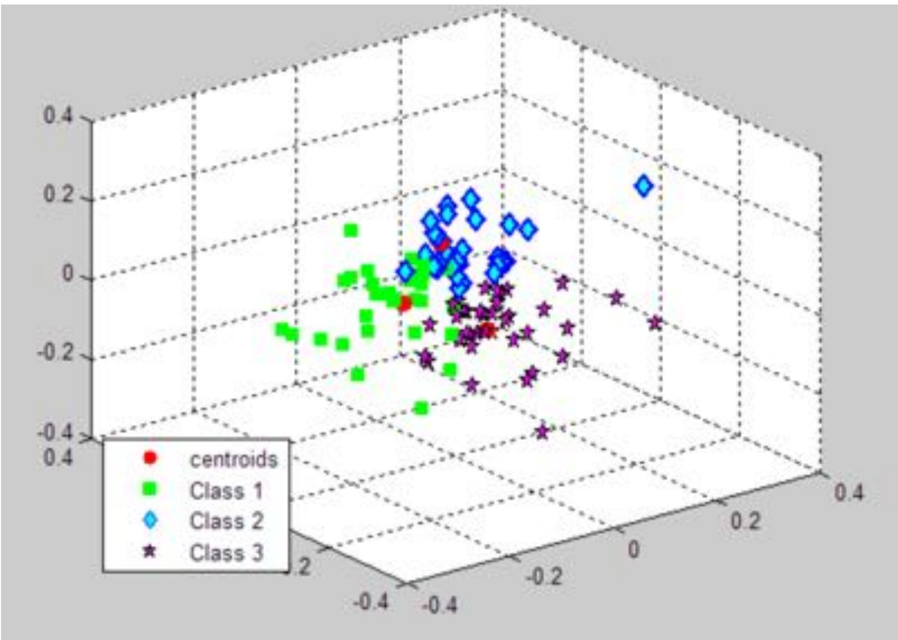

Figure 6.4 3-D Scatter plot of data after PCA based K-Means Clustering

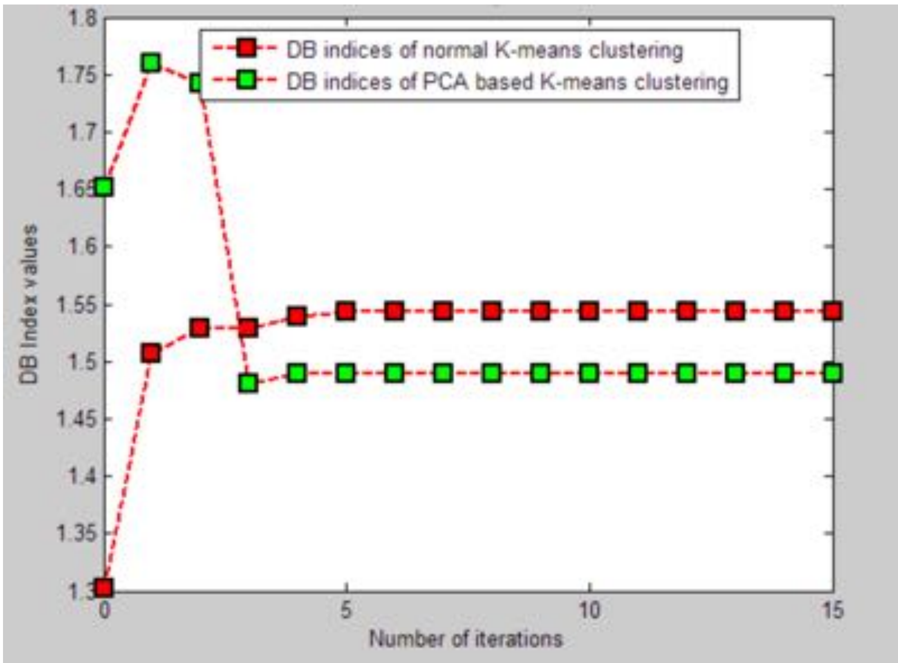

Figure 6.5 Plot of comparison of DB Index values for the two cases

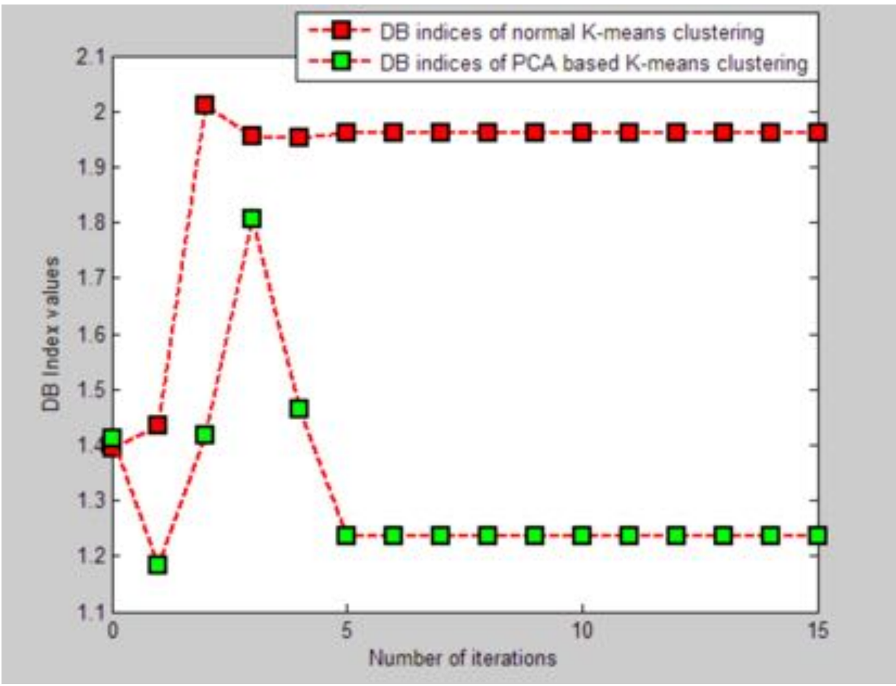

Figure 6.6 Plot of comparison of DB Index values for the two cases for a different data set

Number of iterations	DB Index Values for K-Means Clustering	DB Index Values for PCA based K-Means Clustering
1	1.30283407419202	1.65196402316232
2	1.50693504018069	1.75906671919414
3	1.52869729481111	1.74200473834610
4	1.52921648282933	1.48059992326934
5	1.53891858845080	1.48952561999990
6	1.54334760281319	1.48952561999990
7	1.54334760281319	1.48952561999990
8	1.54334760281319	1.48952561999990
9	1.54334760281319	1.48952561999990
10	1.54334760281319	1.48952561999990
11	1.54334760281319	1.48952561999990
12	1.54334760281319	1.48952561999990
13	1.54334760281319	1.48952561999990
14	1.54334760281319	1.48952561999990
15	1.54334760281319	1.48952561999990
16	1.54334760281319	1.48952561999990

Table 6.1 Comparison of Numerical Values of DB Indices for K-Means and PCA based K-Means Clustering

Now, the mean energy dissipated by the sensor nodes after the entire clustering procedure is calculated, as stated in Chapter 5. We have assumed a simple model where the radio dissipates E_{trans} = 50 nJ/bit to run transmitter or receiver circuitry. ε_{amp} = 100 pJ/bit for transmit amplifier to achieve an acceptable bit error rate. A model that is nearer to reality can be obtained if we modify γ=3, which is typical for environments or stores. The calculated results are as follows:

1. Mean Transmitted Energy after regular K-Means Algorithm = 5.7293e-07 J
2. Mean Transmitted Energy after PCA based K-Means Algorithm = 4.0001e-07 J

As it can be seen, the results demonstrate that the average energy dissipated in PCA based K-Means Algorithm is lesser than that dissipated in regular K-Means clustering.

6.3 Game Theoretic Weighted K-Means Algorithm

We have also shown that Weighted K-Means Algorithm using Shapley values leads to better clustering than regular K-Means Algorithm. Weighted K-Means Algorithm is to be applied on the data set. For that, calculation of weights has to be carried out, for which Shapley Values are required. Thus, Shapley Value of the data is calculated for all the three players, as mentioned in Chapter 5, taking all the three centroids, respectively. The calculated values are shown in Table 6.2.

Shapley Values	**Player 1**	**Player 2**	**Player 3**
φ	103.1627	57.4934	109.0555
φ	73.5918	85.6455	88.3008
φ	80.944	57.4609	48.537

Table 6.2 Resultant Shapley Values

Relative weights are then calculated using these Shapley Values, as stated in Chapter 5, in order to perform weighted K-Means Algorithm. The resultant numerical values of these relative weights are shown in Table 6.3.

Relative Weights		
a	a_2	$a_3 = 0.3141$
b_1	b_2	b_3
c_1	c_2	$c_3 = 0.1974$

Table 6.3 Value of Relative Weights for Game Theoretic Weighted K-Means Algorithm

Since, the coefficients a_1, a_2, a_3 represent the weighted means of the marginal contributions of the three-dimensional data points belonging to Class 1, which is formed as a result of regular K-Means Clustering, hence these relative weights are multiplied to each dimension of the data objects belonging to Class 1, respectively. On similar lines, b_1, b_2, b_3 are multiplied with each of dimension of Class 2, respectively. Similarly, c_1, c_2, c_3 are multiplied to each dimension of Class 3, respectively. Thus, a new data set is formed, on which the Weighted K-Means Algorithm is applied. 16 iterations of the algorithm are applied to get stabilised classes without any further updation of the centroids. The resultant clustering is shown in Figure 6.7.

To check the effectiveness of clustering empirically, Davies-Bouldin Index was calculated after every iteration of K-Means Algorithm and Weighted K-Means Algorithm. The lower the value of the DB Index, lesser are the inter-cluster and intra-cluster distances, better is the compactness and separation of the clusters and thus more efficient is the clustering. It is observed that the final stabilised value of DB Index is lesser in the case where Weighted K-Means Algorithm is used, as compared to the regular K-Means Algorithm. The comparison between the values for K-Means Clustering and Weighted K-Means Clustering are shown graphically and numerically in Figure 6.8 and Table 6.4.

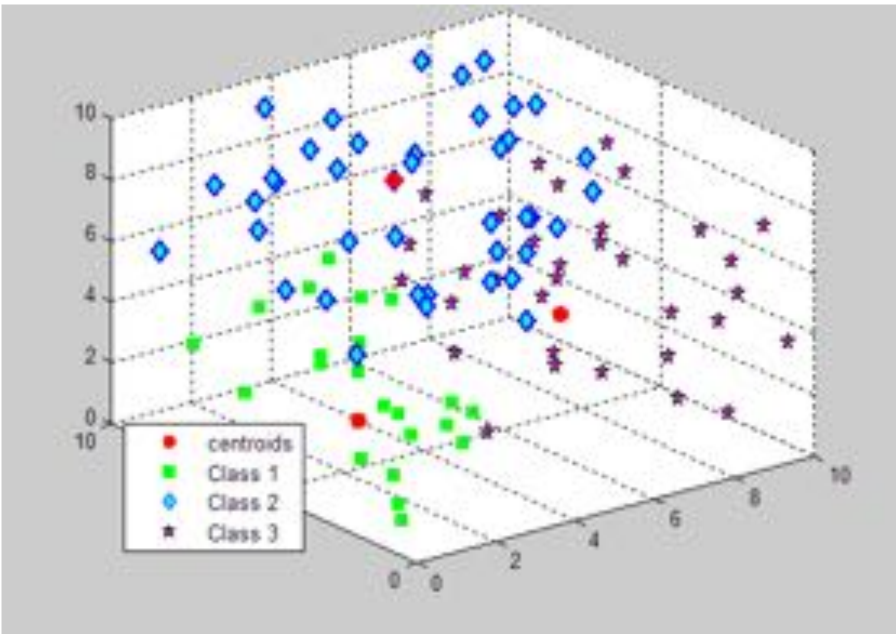

Figure 6.7 3-D Scatter plot of the data after Game Theoretic Weighted K-Means Clustering

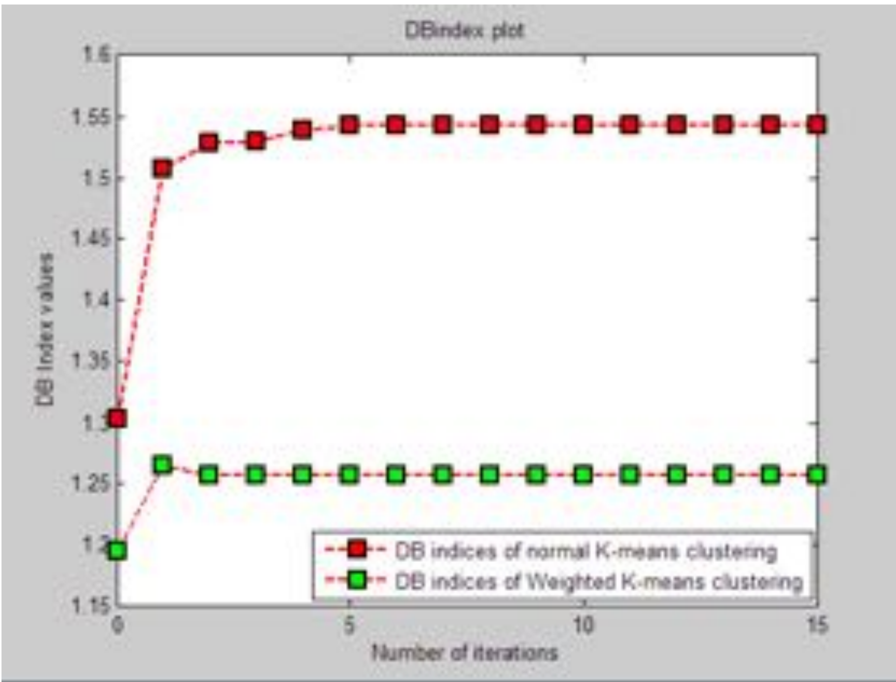

Figure 6.8. Plot of comparison of DB Index values for K-Means and Weighted K-Means Clustering

Number of iterations	DB Index Values for K-Means Clustering	DB Index Values for Weighted K-Means Clustering
1	1.30283407419202	1.19594042856272
2	1.50693504018069	1.26499656064031
3	1.52869729481111	1.25733362412308
4	1.52921648282933	1.25733362412308
5	1.53891858845080	1.25733362412308
6	1.54334760281319	1.25733362412308
7	1.54334760281319	1.25733362412308
8	1.54334760281319	1.25733362412308
9	1.54334760281319	1.25733362412308
10	1.54334760281319	1.25733362412308
11	1.54334760281319	1.25733362412308
12	1.54334760281319	1.25733362412308
13	1.54334760281319	1.25733362412308
14	1.54334760281319	1.25733362412308
15	1.54334760281319	1.25733362412308
16	1.54334760281319	1.25733362412308

Table 6.4 Comparison of Numerical Values of DB Indices for K-Means and Weighted K-Means Clustering

Finally, the mean energy dissipated by the sensor nodes is calculated. The results calculated are as follows:

1. Mean Transmitted Energy after regular K-Means Algorithm = 5.6950e-07
2. Mean Transmitted Energy after Weighted K-Means Algorithm = 5.6668e-07

It was observed in the above results that the average energy dissipated in Weighted K-Means Algorithm is lesser than that dissipated in regular K-Means clustering.

CHAPTER 7

CONCLUSION AND FUTURE SCOPE

7.1 Conclusion

The algorithms are implemented using MATLAB and experiments are conducted and results are evaluated based on performance parameters. The value of the cluster validity measure, Davies-Bouldin Index backs up the claims that both the methods reported in this thesis, viz. PCA based K-Means Algorithm, and Game Theoretic Weighted K-Means Algorithm using Shapley Value, outperform the conventional K-Means Algorithm. This hence proves that performing data reduction using principal component analysis and then applying the standard K-Means Algorithm is an improved approach for clustering of wireless sensor networks, as opposed to simply applying the conventional K-Means Algorithm. Moreover, relatively weighting the data elements using Shapley Values as the weighting criterion and then performing the K-Means clustering, is also a more superior technique for clustering. In addition, the values of average energy dissipated by the wireless sensor networks, using the radio communication model come out to be lesser in the case of both the reported methods, as compared to using standard K-Means algorithm, thus proving the efficiency of the proposed techniques.

7.2 Future Scope

There is a lot of scope of extending the Weighted K-Means approach reported in this work, to application specific and real world scenarios, typically in remote environment monitoring in areas where providing electrical power is difficult, such as flood detection, precision agriculture, biocomplexity mapping of environment, monitoring of inventory location, monitoring friendly military forces, equipment and ammunition, battlefield and terrain surveillance, and also to detect foreign chemical agents in the air and the water, etc. More experiments can be conducted with natural datasets and different features.

LIST OF PUBLICATIONS

- Divleen Kaur and Ravi Kumar, “A PCA based K-Means Clustering Algorithm for Wireless Sensor Nodes”, *Journal of Web Engineering & Technology, STM Journals,* 2015. (**Accepted**)

REFERENCES

[1] I.F. Akyildiz, Weilian Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks", *IEEE Communications Magazine, V*ol. 40, Issue:8, pp. 102-114, 2002.

[2] Savo G. Glisic, "Advanced Wireless Networks", *John Wiley & Sons Ltd,* 2006.

[3] Sandra Sendra, Jaime Lloret, Miguel García and José F. Toledo, "Power saving and energy optimization techniques for Wireless Sensor Networks", *Journal of Communications*, Vol. 6, No. 6, pp. 439-459, 2011.

[4] P. Kumarawadu, D. J. Dechene, M. Luccini, and A. Sauer, "Algorithms for Node Clustering in Wireless Sensor Networks: A Survey" , Information and Automation for Sustainability, 2008. ICIAFS 2008. 4th International Conference , pp.295-300 , 2008.

[5] Upavan Gupta, "A Game Theoretic Approach for Simultaneous Compaction and Equipartitioning of Spatial Data Sets", IEEE Transactions on knowledge and data engineering", Vol. 22, No. 4, pp. 465-478, 2010.

[6] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations", *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability,* University of California Press., pp. 281–297, 1967.

[7] Jiawei Han, Micheline Kamber, "Data Mining: Concepts and Techniques", *Morgan Kaufmann Publishers*, second Edition, 2006.

[8] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman and Angela Y. Wu, "An Efficient k-Means Clustering Algorithm: Analysis and Implementation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, NO. 7, pp. 881 - 892, 2002.

[9] Shi Na, Guan Yong and Liu Xumin, "Research on k-means Clustering Algorithm An Improved k-means Clustering Algorithm", *Intelligent Information Technology and Security Informatics (IITSI)*, Third International Symposium, pp.63-67, 2010.

[10] Kitti Koonsanit, Chuleerat Jaruskulchai and Apisit Eiumnoh, "Parameter-Free K-means Clustering Algorithm for Satellite Imagery Application", *Information Science and Applications (ICISA)*, pp. 1-6, 2012.

[11] Ahamed Shafeeq B M and Hareesha K S , "Dynamic Clustering of Data with Modified K-Means Algorithm", *International Conference on Information and Computer Networks, V*ol. 27 , pp. 221-225, 2012.

[12] Dr. M.P.S Bhatia and Deepika Khurana, "Experimental study of Data clustering using k- Means and modified algorithms", *International Journal of Data Mining & Knowledge Management Process,* Vol.3, No.3, pp. 17-30, 2013.

[13] Katsuhiro Honda, Ryoichi Nonoguchi, and Akira Notsu, Hidetomo Ichihashi, "PCA-guided k-Means clustering with incomplete data", *IEEE International Conference on Fuzzy Systems*, pp. 1710 - 1714, 2011.

[14] Raimir Holanda Filho and José Everardo Bessa Maia, "Network Traffic Prediction using PCA and K-means", *Network Operations and Management Symposium (NOMS)*, IEEE, pp. 938 - 941, 2010.

[15] Simon S. Haykin, “Neural Networks: A Comprehensive Foundation”, *Prentice Hall*, Second edition, 1998.

[16] Tomohiro Matsui, Katsuhiro Honda, Chi-Hyon Oh, Akira Notsu, and Hidetomo Ichihashi, “Cluster Validation in *k*-Means Clustering Based on PCA-guided *k*-Means and Procrustean Transformation of PC Scores”, *IEEE International Conference on Fuzzy Systems*, pp. 1546-1550, 2009.

[17] C. Ding and X. He, “K-means clustering via principal component analysis,” *Proc. of Int’l Conf. Machine Learning*, pp. 225-232, 2004.

[18] Shokri Z. Selim and M. A. Ismail, “K-Means-Type Algorithms: A Generalized Convergence Theorem and Characterization of Local Optimality”, *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Vol. Pami-6, No.1, pp.81-87, 1984.

[19] Shigen Shen, Guangxue Yue and Qiying Cao, “A Survey of Game Theory in Wireless Sensor Networks Security” , *Journal Of Networks*, Vol. 6, No. 3, pp. 521-532 , 2011.

[20] Fotini-Niovi Pavlidou and Georgis Koltsidas, “Game Theory for Routing Modeling in Communication Networks — A Survey”, *Journal of Communications and Networks*, Vol. 10, Issue: 3, pp. 268 - 286, 2008.

[21] Fatemeh Kazemeyni, Einar Broch Johnsen, Olaf Owe and Ilangko Balasingham, “Group Selection by Nodes in Wireless Sensor Networks Using Coalitional Game Theory”, *16th IEEE International Conference on Engineering of Complex Computer Systems*, pp.254-262, 2011.

[22] Li Hui, Liu Chun and Li Hua, “Power Control Based on Non-cooperative Game in Wireless Sensor Networks”, International Conference on Intelligent Networks and Intelligent Systems, pp. 135 - 138, 2013.

[23] Mi Yan, Liang Xiao, Liang Du and Lianfen Huang, “On Selfish Behavior in Wireless Sensor Networks: A Game Theoretic Case Study”, *International Conference on Measuring Technology and Mechatronics Automation*, Vol. 2, pp. 752-756, 2011.

[24] Afrand Agah, Sajal K.Das and Kalyan Basu, “A game theory based approach for security in wireless sensor networks”, *IEEE International Conference on Performance, Computing, and Communications,* pp. 259-263, 2004.

[25] Shamik Sengupta, Mainak Chatterjee, and Kevin A. Kwiat, “A Game Theoretic Framework for Power Control in Wireless Sensor Networks”, *IEEE Transactions on Computers*, Vol. 59, NO. 2, pp. 231-242, 2010.

[26] E. Karapistoli, A.A. Economides, “Routing in Wireless Sensor Networks: An approach using Stackelberg games”, *IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications*, pp. 322 - 327, 2014.

[27] Dongmei Yan, Jinkuan Wang, Li Liu and Jing Gao, “Target tracking based on cluster and game theory in wireless sensor network”, *IET 2nd International Conference on Wireless, Mobile and Multimedia Networks,* pp. 45 - 48, 2008.

[28] Jia Peng, Wang Zhe, Yan Yujie and Wang Shu, “Cooperative Packet Forwarding in Wireless Sensor Networks”, *International Conference on Communication Technology*, pp. 1-5, 2006.

[29] Walid Saad, Zhu Han, Merouane Debbah, Are Hjørungnes, and Tamer Basar, “Coalitional game theory for communication networks: A tutorial”. *IEEE Signal Processing Magazine*, pp. 77–97, 2009.

[30] Ravi Kumar, “Game Theoretic Pattern Analysis for Identification of Odors/ Gases Using Response of a Poorly Selective Sensor Array”, *IEEE Sensors Journal*, Vol. 13, No. 3, pp. 1110-1116, 2013.

[31] J.Z. Huang, M.K Ng, Rong Hongqiang and Li Zichen, “Automated variable weighting in k-means type clustering”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 657 - 668, 2005.
[32] Sanghamitra Bandyopadhyay and Ujjwal Maulik, “Nonparametric Genetic Clustering: Comparison of Validity Indices”, *IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews*, Vol. 31, No. 1, pp. 120-125, 2001.

[33] D.L. Davies, D.W. Bouldin, “A cluster separation measure”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 224-227, 2009.

[34] Dan Tudose, Laura Gheorghe and Nicolae Tapus, “Radio transceiver consumption modeling for multi-hop wireless sensor networks,” *UPB Scientific Bulletin, Series C*, Vol. 75, No. 1, pp. 17–26, 2013.

[35] G. Mao, B. D. O. Anderson, and B. Fidan, “Path loss exponent estimation for wireless sensor network localization,” *The International Journal of Computer and Telecommunications Networking*, Vol. 51, No. 10, pp. 2467-2483, 2007.