A FUZZY LOGIC CLUSTER FORMATION PROTOCOL FOR WIRELESS SENSOR NETWORKS

by

Rogaia M. Mhemed

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
December 2011

DALHOUSIE UNIVERSITY

DEPARTMENT OF ENGINEERING MATHEMATICS

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "A FUZZY LOGIC CLUSTER FORMATION PROTOCOL FOR WIRELESS SENSOR NETWORKS" by Rogaia M. Mhemed in partial fulfilment of the requirements for the degree of Master of Science.

Dated: December 6th, 2011

Co-Supervisor: _____

Co-Supervisor: _____

Reader: _____

Reader: _____

DALHOUSIE UNIVERSITY

DATE: December 6th, 2011

AUTHOR: Rogaia M. Mhemed

TITLE: A FUZZY LOGIC CLUSTER FORMATION PROTOCOL FOR WIRELESS SENSOR NETWORKS

DEPARTMENT OR SCHOOL: Department of Engineering Mathematics

DEGREE: MSc CONVOCATION: May YEAR: 2012

_____

Signature of Author

DEDICATION PAGE

This thesis is dedicated to my dear husband Abdulhakim, and my wonderful children Ahmed and Nour. I dedicate this thesis to my darling parents and all my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The recent advancements in the wireless sensor networks (WSNs) are predominantly motivated by developments in the micro electromechanical systems (MEMS) technology. Typically, a WSN is a collection of a large number of low cost wireless nodes that contain one or more MEMS-based sensors. Integration of sensors to wireless nodes in this manner allows them to interact with the physical world and collect the readings using on-board sensors. Examples of sensors include, light, humidity, motion and GPS. As a result, WSNs could be used for a large number of applications that require data gathering from physical world in an unattended manner. Examples of such applications include environmental monitoring, structural health monitoring, military, and commercial, agriculture, surveillance and security.

Wireless sensor nodes are resource constrained and have limited amount of energy. Therefore, designing protocols that conserve energy is an important area of research. Researchers have investigated architectures and topologies that allow energy efficient operation of WSNs. One of the popular techniques in this regard is clustering. A typical clustering protocol contains two main steps: cluster head election and cluster formation. This thesis is aimed at investigation of the cluster formation process. We propose a Fuzzy Logic based approach that uses three descriptors namely: energy level, distance between cluster-head and base station, and distance between the cluster-head and the sensor's node. We compare our proposed model, FLCFP (Fuzzy Logic Clustering Formation Protocol), with the most popular model, LEACH (Low Energy Adaptive Clustering Hierarchy), which was proposed previously to prolong network lifetime. The FLCFP approach is shown to prolong network lifetime. In addition, it is shown that sensor node energy is consumed in a more uniform fashion.

LIST OF ABBREVIATIONS AND SYMBOLS USED

LIST OF ABBREVIATIONS

ADCs	Analog to Digital Conversion
BS	Base Station
CHEF	Cluster Head Election mechanism using Fuzzy Logic in Wireless Sensor Networks
CHs	Cluster-Heads
COA	Center of Area
CSMA	Carrier Sense Multiple Access
DM	Decision Matrix
FIS	Fuzzy Inference System
FL	Fuzzy Logic
FLCFP	Fuzzy Logic Cluster Formation Protocol
GPS	Global Positioning System
FND	First Node Death
ID	Identification
LEACH	Low Energy Adaptive Clustering Hierarchy
LND	Last Node Death
MEMS	Micro Electro-Mechanical Systems
MFs	Membership Functions
MOECS	Multi-Objective Energy-efficient Clustering Scheme
OM	Options Matrix
PV	Preference Vector
RAM	Random Access Memory
ROM	Read-Only Memory
TDMA	Time Division Multiple Access
WSNs	Wireless Sensor Networks

LIST OF SYMBOLS

$\mu_{\check{A}}(x)$	The degree of membership function of x
p	Probability to become CH
r	The number of the current round
$E_{TX}(l,d)$	Energy consumed in transmitting an *l*-bit packet d meters
$E_{RX}(l,d)$	Energy consumed in receiving an *l*-bit packet d meters
E_{elec}	Electronics energy consumed in transmitting an l-bit packet
ε_{amp}	Energy constant for propagation
λ	The path loss exponent

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor, Dr. Nauman Aslam, and my supervisor Dr. William Phillips. Their wide knowledge has been of great value for me. Their understanding, encouraging and personal guidance have provided a good basis for the present thesis. I am grateful to them for endless hours they are spending with me.

Special thanks to Dr. Frank Comeau for his feedback and the wonderful advice, which helped in the completion of this thesis.

Many thanks to Dr. Mohammed El-Hawary for accept to be in the examiner committee and read my thesis. I would like to thank him for the time that he gave me whenever I need him to help.

I also would like to thank Mrs. Karen Conrod for her patience and all non-thesis help and assistance.

Thanks for all my friends, all Libyan and Muslim community here in Halifax, for their supporting and helping me and my family in the difficult situations. I'm very lucky to meet them and to be one of them.

Finally, I would like to thank my husband for all his patience, pray and endless support. Thanks to my son and daughter they keep my life filled with laughter and love. Last but not the least; I am heartily thankful to all my family in Libya (my parents, my mother-in-law, my brothers and sisters) for their love, prayers and support.

CHAPTER 1 Introduction

1.1 Wireless Sensor Networks Overview

WSNs are usually composed of hundreds or maybe thousands of tiny, inexpensive, low-power sensor nodes with limited memory and processing capability, and one or more controlling Base Station (Sink) [1] [2]. Figure 1.1 illustrates typical cluster-based WSN architecture. While the sensor networks may vary slightly from each other for different applications, they are similar in their general structure and share common technical issues. A number of topologies have been proposed to transmit data from sensor node to the BS, such as tree-based topology and cluster-based topology [3]. In our thesis we will use the cluster-based topology, and the process through the WSNs using cluster-based topology can be explained in an as follows:

The Base Station (BS) selects a number of sensor nodes to act as Cluster Heads (CHs). Each non-cluster head node is associated to a CH thereby forming the clusters. After that, each node senses the environment and measures physical phenomenon of interest (e.g., temperature, pressure, smoke, humidity). Each node then aggregates and transmits its measurements and information to its associated CH. The CH compresses this data in a single signal and transmits it to the BS (sink). The BS serves as a gateway to send the data to another network (Internet or WAN), then to the users.

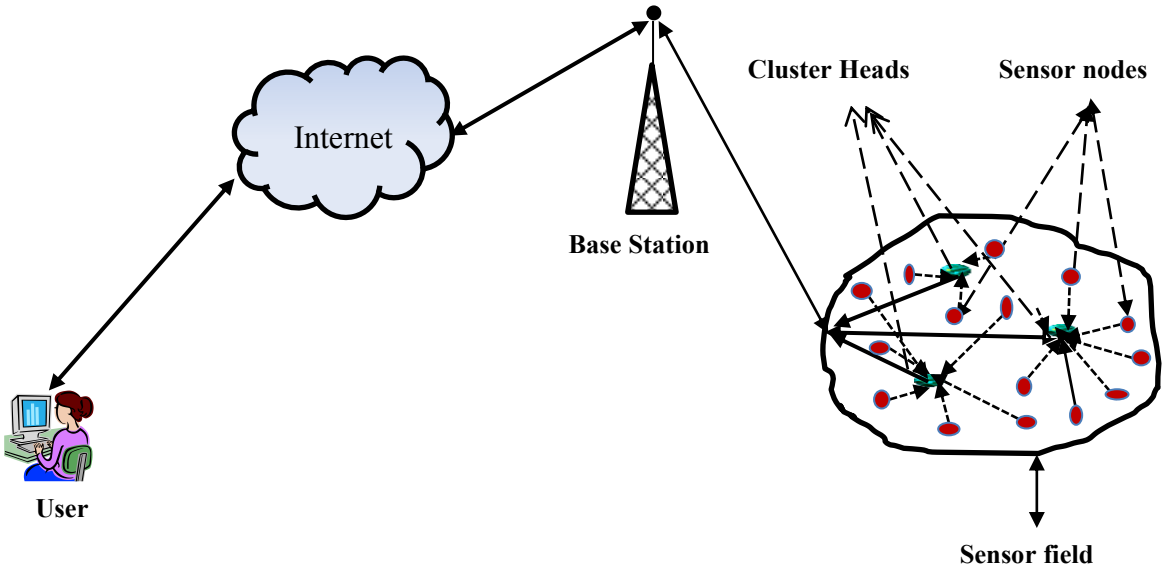

Figure 1.1: Typical cluster-based WSN Architecture

1.2 Sensor Node Hardware Platform

In WSNs, the sensor node is a device that has the ability to transform the parameters, events or phenomena in the physical world, to the electrical signals that can be passed into the computing and control system to be analyzed [4]. Therefore, we can see from Figure 1.2 that the fundamental sensor node hardware is composed of:

- Memory/Storage: is made up of random access memory (RAM) and read-only memory (ROM). With the emergence of flash memory, on-board storage capacity improved dramatically with the passage of time [5] .
- Controller/ Processor: is generally associated with a small storage unit. It manages the procedures that make the sensor node collaborate with the other nodes to carry out the assigned sensing tasks [2].

- Transceiver: achieves the functions of both the transmitter and receiver with a limited transmission range. It connects the node and the network [2].
- Sensing Unit: consist of sensor devices that detect the surrounding environment and perform analog to digital conversion (ADC) [2]. It is possible to have several sensors such as: temperature sensors, light sensors, humidity sensors, pressure sensors etc.
- Power Unit: is used for flexible and random deployment. Nodes are battery powered (e.g. using LiMH AA batteries) [5].

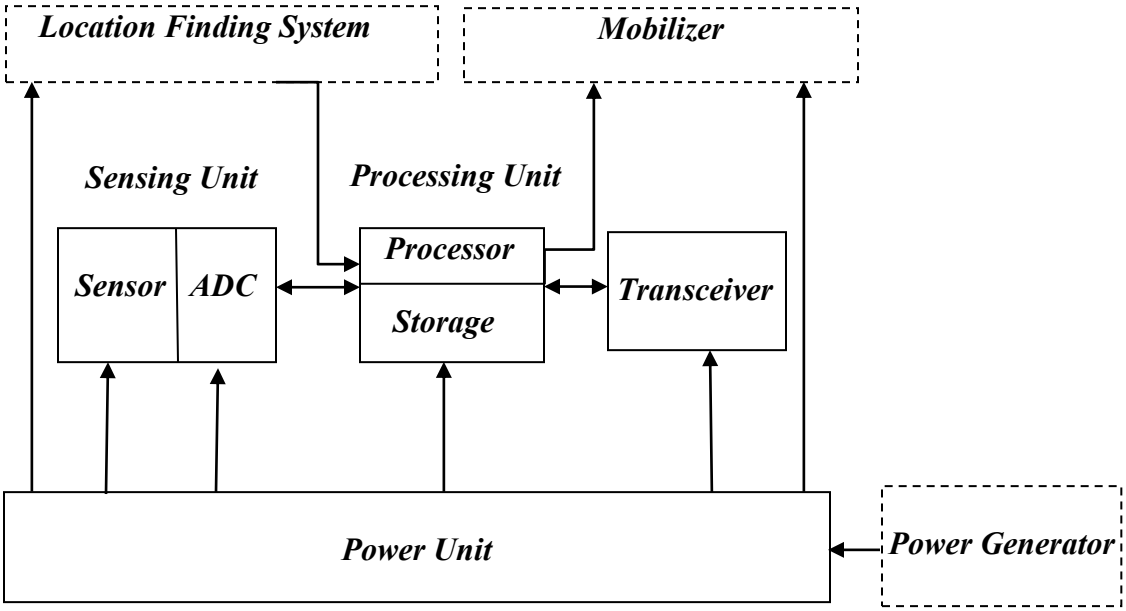

Figure 1.2: Sensor Node Hardware Platform [2]

Furthermore, we can also see that, the sensor nodes could also have additional components for applications such as a location finding system, a mobilizer and a power generator [6].

1.3 Applications

As mentioned in the previous section, the sensor node has a sensor unit that gives it the ability to interrelate with the surround environment. As a result, many WSN applications that appeared and rapidly improved. These applications include: security, surveillance, monitoring, and detection. Some of these applications as reviewed in [2] [6] [7] [8] are listed as follows:

1.3.1 Military Applications:

WSNs are an integral part of military command, control, communications, computing, intelligence, surveillance, and reconnaissance [7]. A number of WSN applications have been found in the military, for instance: battlefield surveillance, nuclear, biological and chemical attack detection and reconnaissance, battle damage assessment [6] .

1.3.2 Environmental Applications:

WSNs are used in many applications for monitoring the environment. For example, there are applications for monitoring the movements of animals, detection of forest fires, floods detection, and surveillance of the environmental factors that affect agricultural crops and livestock [6] [8] [9].

1.3.3 Health Care Applications:

WSNs have been used in a number of health care applications including integrated patient monitoring, diagnostics, drug administration in hospitals, and tracking and monitoring doctors and patients inside a hospital [10].

1.3.4 Home Applications:

As we mentioned, the technology is improving rapidly, and smart sensor nodes and actuators can be found in home appliances, such as vacuum cleaners, micro-wave ovens, and refrigerators. These sensor nodes found inside home devices can interact with each other and with an external network permitting end users to manage home devices locally and remotely [11].

1.3.5 Other Commercial Applications:

Commercial applications of WSNs include monitoring; product quality; constructing smart office spaces; environmental control in office buildings; robot control and guidance in automatic manufacturing environments; interactive toys; detecting and monitoring car thefts; and vehicle tracking and detection [2] [4].

1.4 Problem Statement

WSNs rely on resource-constrained sensor nodes to collect data from the physical environment. Energy efficiency is one of the most important concerns. For an energy efficient operation, optimal cluster formation is necessary to ensure that energy is consumed at a balanced rate. The operation of cluster based WSNs is broken into rounds. Each round is made up of cluster head selection, cluster formation and data transmission. The network lifetime is the number of rounds in which all nodes have non-zero energy. While some studies have focused on the CH selection [12] [13] [14], we have focused our research on the cluster formation process. In the LEACH (Low-Energy Adaptive Clustering Hierarchy) protocol, [15] each node decides which cluster it belongs to by picking the CH that requires the smallest transmission energy. In fact, considering only

the distance between the CH and the node ignores many other factors that affect the energy consumption and the network lifetime. In this thesis we present a Fuzzy Logic approach for the cluster formation process. In this approach we consider three descriptors: the energy level of CH, distance between the BS and the CH and distance between the CH and the node. For cluster formation, each non-CH node applies the three descriptors for each CH in the Mamdani Fuzzy Inference System [16], and joins the CH that has the maximum chance value to form the cluster. We call our approach the Fuzzy Logic Cluster Formation Protocol (FLCFP) for WSN.

To show the features of our FLCFP we compare it with LEACH. Simulation results show that our approach extends the network lifetime significantly as compared to the LEACH protocol. In addition, our simulations show that the nodes consume energy in a more uniform fashion.

1.5 Thesis Outline

Previous Sections provided a broad overview of WSN technology, applications and discussed components of a wireless sensor node.

The rest of the thesis is divided into other four chapters as following:

Chapter 2: introduces the clustering techniques, an overview of Fuzzy Logic system, and discusses some details of the constituent parts of the Fuzzy Logic. We also provide a literature review of related clustering protocols including LEACH [15], and some Fuzzy protocols [12] [13] [14].

Chapter 3: presents our novel approach for the cluster formation. The FIS design for the proposed protocol (FLCFP), and the inputs and output parameters are explained. Then we offer an example to show how the Fuzzy Logic calculation steps are done.

Chapter 4: presents the detailed simulation of FLCFP. We compare FLCFP with the LEACH protocol using FNDs metric and analyses the difference among them using MINITAB.

Chapter 5: includes the conclusion of the main idea for our thesis. Some topics for the future work are listed at the end of this chapter.

Appendix A: provides the formulas for Triangle and Trapezoid membership functions used in determine the membership functions degrees.

Appendix B: presents the pseudo code for Set-up stage of our model FLCFP.

CHAPTER 2 Background and Related work

2.1 Introduction

In this chapter we present a detailed background review of related work. We start with a brief overview of clustering techniques used in WSNs. A detailed discussion on fuzzy systems is provided in section 2. Section 3 provides a thorough review of clustering protocols related to our work.

2.2 Clustering Techniques

Clustering is defined as a grouping together of similar data items [17]. More specifically, clustering is a grouping or organizing of objects that share one or more properties. Several clustering strategies have been proposed for WSNs in recent years by many investigators [18] [19]. These clustering strategies add flexibility in achieving many goals such as: energy efficient operation, prolonging the WSN lifetime and decreasing the number of nodes that communicate with the BS. Clustering algorithms in WSN are categorized based on their techniques, motivations and applications [20]. Although there is no agreed upon classification of clustering algorithms, the authors of [18] classify clustering into four categories: Heuristic Schemes, Weighted Schemes, Hierarchical Schemes, and Grid Schemes, as illustrated in Figure 2.1.

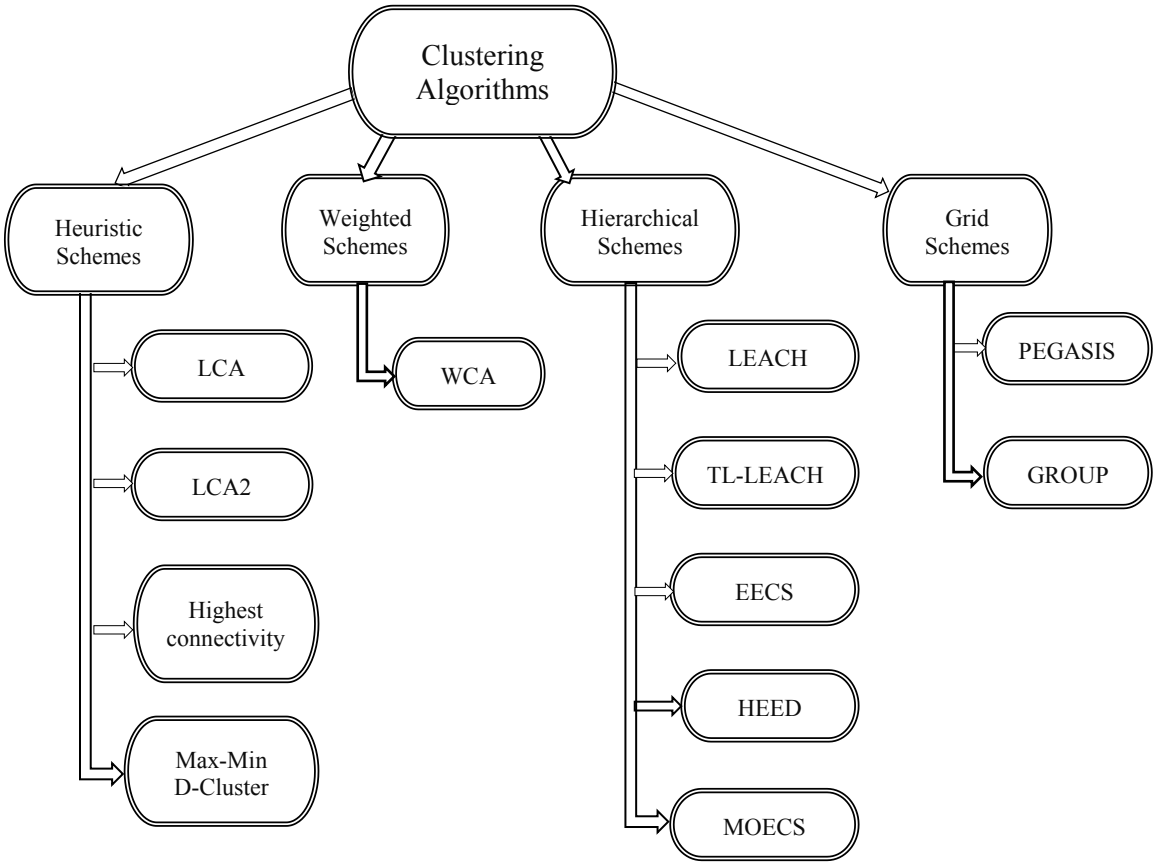

Figure 2.1: Classification of Proposed Clustering Schemes (modified from [18])

Our proposed protocol, the FLCFP, is an extension of the LEACH protocol and therefore fits into the classifications of Figure 2.1 as a Hierarchical Scheme.

2.3 Overview of Fuzzy System

The Fuzzy Logic (FL) concept was introduced by Professor Lotfi Zadeh in the mid-1960's. [21] [22] [23]. Initially, this concept offered an approach for handling data through permitting partial set membership instead of crisp set membership. Zadeh noted: "The closer one looks at a real-world problem, the fuzzier becomes its solution" [21] [24]. We can define Fuzzy Logic as a superset of traditional logic (Boolean) that has

been extended to address the concept of partial truth, meaning the truth values will be between "completely true" and "completely false" [25] [23]. In the other words, classical Boolean logic has just two values, true (always presented numerically as 1) or YES, and false (always presented numerically as 0), or NO. Instead, Fuzzy Logic extends these two values to obtain multi-values between 0 and 1 using the concept of degrees of membership.

In the next subsections, we describe fuzzy sets, membership functions, linguistic variables, fuzzy operations, fuzzy IF-THEN rules, and Fuzzy Inference Systems. These concepts will be used in Chapter 3.

2.3.1 Fuzzy Sets

As we mentioned in Section 2.3, Fuzzy Logic is an extension of classical (Boolean) logic. We readily observe that, the definition of a fuzzy set is simply an extension of the definition of classical sets that allow you to define a distinct function having values between 0 and 1. We can present the classical set as follow:

$$A = \{(x, \mu_A(x)) | x \in U\} \qquad (2.1)$$

Hence, the distinct function of a classical set A, which is $\subseteq U$, $\mu_A(x)$, is the characteristic function defined on the closed interval [0, 1] for set A.

We can then present the fuzzy set definition by extending the above formula as:

$$\check{A} = \{(x, \mu_{\check{A}}(x)) | x \in U\} \qquad (2.2)$$

Where, $\check{A}$ is a fuzzy set in the universe U, and $\mu_{\check{A}}(x)$ is the degree of membership function of x in $\check{A}$.

The $\mu_{\check{A}}(x)$ in the above formula includes all possible real numbers (grades) in the closed interval [0, 1], not just 0 and 1.

2.3.2 Membership Function

The traditional set of a given universe consists of distinct elements that have one feature or more in common, These elements are called characteristic functions, and can be determined by two points: 0 to form non membership for elements that do not belong to the particular set, and 1 to form full membership for elements that do belong to the set. Therefore, we can determine the characteristic function as shown below in equation 2.3

$$\mu_A(x) = \begin{cases} 1 & if\ and\ only\ if\ \ x\ \in A \\ 0 & if\ and\ only\ if\ \ x\ \notin A \end{cases} \qquad (2.3)$$

Hence, $\mu_A(x)$ is a characteristic function for set A, and A is a classical set of the universe.

The fuzzy set is an extension of the classical set; the elements in a fuzzy set extend the notion of a binary characteristic function in a classical set to multiple values on the continuous interval [0, 1].

$$\check{A} = \{(x, \mu_{\check{A}}(x)) | x \in U\} \qquad (2.4)$$

Where, $\check{A}$ is a fuzzy set in the universe U, and $\mu_{\check{A}}(x)$ is a membership function of x in $\check{A}$.

2.3.3 Linguistic Variables

In each model, parameters or variables are used to control the system behavior, and to show their effect on the system performance. These variables called linguistic variables. These linguistic variables may be divided into levels, each one called a Linguistic values or terms. The linguistic values are used to give flexibility to the linguistic variables, to examine or control the system. For instance, in the next chapter we will use three linguistic variables (Energy level of CH, distance To the BS, distance between the CH

and the node) as inputs, each divided into three linguistic values, and linguistic variable (Chance value) as output, divided into nine linguistic values.

2.3.4 Fuzzy Operations

In this subsection we introduce the main operations in both classical and fuzzy sets. Fuzzy sets have the same operations as classical sets: containment, union, intersection, and complement. We therefore assume that $\check{A}$,$\check{B}$ and $\check{C}$ are fuzzy sets in the universe of discourse.

Intersection: contains the shared elements of two sets. In fuzzy sets, the membership $\mu_{\check{A}}(x)$ varies from 0 to 1. For the intersection of sets, we choose the minimum membership value for both sets to find the fuzzy intersection.

So, for fuzzy sets $\check{A}$ and $\check{B}$ the intersection are fuzzy set $\check{C}$ and their membership functions are presented mathematically as [26]:

$$\check{C} = \check{A} \cap \check{B} \tag{2.5}$$

$$\mu_{\check{C}}(x) = min(\mu_{\check{A}}(x), \mu_{\check{B}}(x)) = \mu_{\check{A}}(x) \wedge \mu_{\check{B}}(x) \tag{2.6}$$

Intersection is the AND operator.

Union: The process of collecting the elements of two or more sets in one set is called union. Consequently, the union operator is the opposite of the intersection operator. An OR operator is used for union in Fuzzy Logic Therefore, we choose the maximum membership value for both sets to find the fuzzy union. For fuzzy sets $\check{A}$ and $\check{B}$, the union is fuzzy set $\check{C}$, and their membership function is presented mathematically in equations 2.7 and 2.8 below [26]:

$$\check{C} = \check{A} \cup \check{B} \tag{2.7}$$

$$\mu_{\check{C}}(x) = max(\mu_{\check{A}}(x), \mu_{\check{B}}(x)) = \mu_{\check{A}}(x) \vee \mu_{\check{B}}(x) \tag{2.8}$$

Complement: complement, or negation, is the opposite of the given set, meaning that all elements in universe U and do not belong to that set. The complement operation in Fuzzy sets is the same as the NOT operation in classical sets.

Therefore, in fuzzy set $\check{A}$, the complement is fuzzy set $\overline{\check{A}}$ and its membership function is presented mathematically as [26]:

$$\mu_{\overline{\check{A}}}(x) = 1 - \mu_{\check{A}}(x) \qquad (2.9)$$

In the next chapter we design the FIS using the Mamdani method, which is based on intersection and union. We implement intersection as an AND operator, and union as an OR operator. These operators are used to calculate the chance value in Chapter 3. The next Figure 2.3 illustrates the three previous operators.

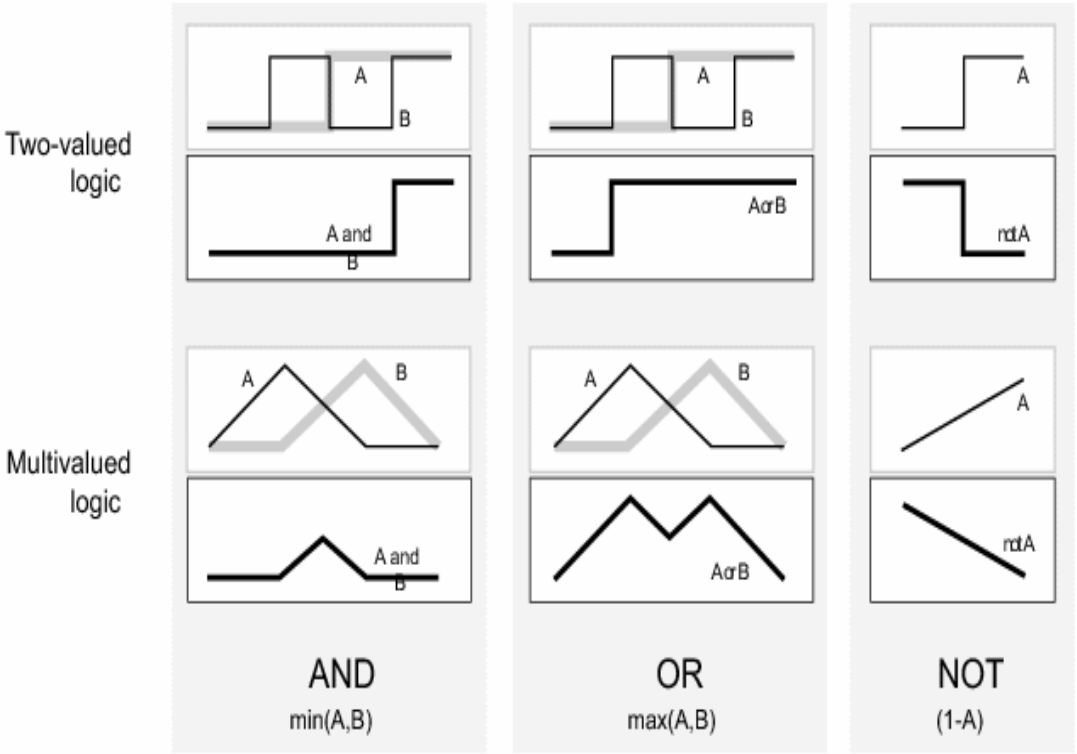

Figure 2.2: Fuzzy Operators (AND, OR, NOT) [30]

Two additional operators are of interest:

Equality: equality of two sets should be equal in all elements, so we can say $\check{A}$ equal $\check{B}$ if and only if all membership functions are equal for both sets. This operation can present mathematically as:

$$\check{A} = \check{B} \Leftrightarrow \mu_{\check{A}}(x) = \mu_{\check{B}}(x) \qquad (2.10)$$

***Containment*:** containment is also called subset; we can say $\check{A}$ is a subset of $\check{B}$ if and only if $\mu_{\check{A}}(x) \leq \mu_{\check{B}}(x)$ and vice versa. Mathematically, we can write:

$$\check{A} \subseteq \check{B} \Leftrightarrow \mu_{\check{A}}(x) \leq \mu_{\check{B}}(x) \qquad (2.11)$$

2.3.5 Fuzzy IF-THEN Rules

Fuzzy rules, or If-Then rules, are statement(s) that consist of three parts: antecedent, fuzzy proposition and consequence(s). More than one antecedent may contain the (AND) or (OR) operator. We can express the fuzzy IF-THEN rule in the following statement:

If *x_1 is A and/or x_2 is B* ***then*** *y is C*

Where, A, B, and C are linguistic values, while x_1, x_2 and y are the linguistic variables.

2.3.6 Fuzzy Inference System

A Fuzzy Inference System (FIS) is the collection of an assortment of fuzzy IF-THEN rules, a database comprising of membership functions of linguistic variables, and a fuzzy reasoning. [27]. In addition, the FIS is a very influential methodology designed for constructing complex and nonlinear relationships among input(s) and output(s). The Mamdani, Sugeno, and Tsukamoto FIS [28] [29] [30] have been used in many different applications. The difference between these three types of FIS lies in the aggregation and defuzzification processes. We chose the Mamdani FIS method because we found that it

has widespread acceptance and is the most commonly used of the three FIS. The Mamdani FIS is popular because it is intuitive [30]. The procedure for implementing the FIS is divided into four steps: fuzzification, rule evaluation, aggregate output(s), and finally defuzzification. These steps will be explained and clarified by a simple example in the next chapter.

2.4 LEACH Clustering Protocol

In this section, we describe the Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol [15]. LEACH is the original clustering protocol for WSN. It created a foundation for many other approaches such as those proposed by the authors of [12] [31]. LEACH was the first significant protocol that proposed to extend the overall lifetime of the network and to decrease the overall energy consumed by the network [15].

In [32] it is shown that using the LEACH protocol the communication energy decreases as much as eight times as compared to direct transmission.

The LEACH is divided into rounds. Each round consists of a set-up stage and a steady-state stage. The set-up stage consists of CH election and cluster formation. The steady-state stage consists of sensing, and transmission of the sensed data to the CH and then to the BS.

2.4.1 Set-Up Stage

At the start of the first set-up stage and in every subsequent round, every sensor node picks a random number between 0 and 1 to determine it will become a cluster-head or not, then compares this number with threshold value-$T(n)$. If the number chosen by a

particular node is less than the threshold value $T(n)$, the node becomes a CH for that current round. We can compute $T(n)$ as shown below in equation (2.12) [12] [15]:

$$T(n) = \begin{cases} \frac{p}{1-p\times\left(r \bmod \frac{1}{p}\right)} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \qquad (2.12)$$

Where p = k/N; k is the expected number of CHs in the round and N is the number of nodes in the network. The value r is the round number. G is the set of nodes that have not been a CH in the last (r mod (N/k)) rounds.

Hence, from the above formula we can say that, if a particular node does not belong to set G, it cannot become a CH for that round; the threshold *T(n)* for this node will be set to 0. Also, for the first round (r = 0), the probability that a node will become a CH is p for all nodes. For rounds r = 1 to N/k – 1, the nodes, n, that have been a CH have a value of *T(n)* = 0. They cannot become a CH. For the remaining nodes, *T(n)* increases as r increases. Since fewer and fewer nodes have yet to be CHs, the probability that any one will become a CH in a particular round increases. At round N/k – 1, only k nodes on average have not yet been CHs. Their chance of becoming a CH is 1. At round r = N/k, again all nodes have a probability of p of becoming a CH; the process then repeats.

After the CHs have been chosen based on the above technique, each node sends an advertisement message via CSMA-MAC over the network to inform other nodes that it is a CH. All other nodes listen to hear the advertisement messages, and then nodes that receive the advertising message decide which CH to join based on the strongest advertisement message signal. After that, the CHs listen for the join-request message from non-CHs. Each non-CH node sends a join-request message to the CH that will result in the lowest amount of transmission energy. The clusters having been organized, each

CH allocates its TDMA schedule to convey to every member node when it will transmit its data; this approach will decrease the chance of the collisions.

2.4.2 Steady-State Stage

As soon as all the CHs are selected and the clusters are formed, the set-up stage is complete and the steady-state stage starts. In this stage, each CH waits to receive data from all nodes in its cluster depending on the time specified in TDMA schedule. Afterwards, the CHs will process these data, compress them in a single signal, and transmit the result to the BS.

The two following equations (2.13), (2.14), compute the energy that will be consumed during transmission and reception between transmitter and receiver:

$$E_{TX}(l,d) = E_{elec} * l + \varepsilon_{amp} * l * d^{\lambda} \qquad (2.13)$$

$$E_{RX}(l) = E_{elec} * l \qquad (2.14)$$

Where, λ is the path loss exponent, l is a messages size in bits, d is distance between transmitter and receiver, $\varepsilon_{amp}$is energy constant for propagation, and E_{elec} is the electronics energy. For transmissions to the CH λ = 2 and ε_{amp} =10 pJ/bit/m^2. For transmissions to the BS λ = 4 and ε_{amp} =0.0013 pJ/bit/m^4.

2.5 Multi-Objective Cluster Formation Scheme

The work presented in this thesis is inspired by the multi-objective optimization based cluster formation scheme (MOECS) [33]. The MOECS scheme employs multiple metrics for cluster formation that are critical for balanced energy dissipation of the system. In their proposed scheme an arbitrary number of CH properties (remaining energy, distance

from node and distance from base station for example) are taken into consideration, with relative weightings to each other (the preference vector, PV). The cluster formation process starts at each node by constructing an options matrix (OM) as shown in equation (2.15). In the OM, each node records the values of metrics used in the cluster formation process. The rows of the OM represent CHs which are in sensor node's radio transmission range. Each element in any row represents an individual parameter e.g. the element $x_{i,j}$ in the OM represents the j^{th} parameter for the i^{th} CH.

$$OM = \begin{bmatrix} x_{1,1} & x_{1,2} \,..\, x_{1,n} \\ x_{2,1} & x_{2,2} \,..\, x_{2,n} \\ . & . \;..\; . \\ x_{k,1}, & x_{k,2} \,..\, x_{k,n} \end{bmatrix} \qquad (2.15)$$

The OM is converted into a decision matrix (DM) using equation (2.16)

$$DM = \begin{bmatrix} s_{1,1} & s_{1,2} \,..\, s_{1,n} \\ s_{2,1} & s_{2,2} \,..\, s_{2,n} \\ . & . \;..\; . \\ s_{k,1}, & s_{k,2} \,..\, s_{k,n} \end{bmatrix} \qquad (2.16)$$

$$s_{i,j} \leftarrow 2\frac{(x_{i,j}-b_j)}{(a_j-b_j)} - 1\,,\ a_j = max_i(OM_{i,j}),\ b_j = min_i(OM_{i,j}) \qquad (2.17)$$

Finally, obtain the weight vector W by multiplying the decision matrix DM with the preference vector. The C with the highest weight value is chosen.

$$\begin{bmatrix} w_1 \\ w_2 \\ . \\ w_k \end{bmatrix} = \begin{bmatrix} s_{1,1} & s_{1,2} \,..\, s_{1,n} \\ s_{2,1} & s_{2,2} \,..\, s_{2,n} \\ . & . \;..\; . \\ s_{k,1}, & s_{k,2} \,..\, s_{k,n} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ . \\ p_n \end{bmatrix} \tag{2.18}$$

The main difference between work in this thesis and the MOECS is that we used a fuzzy logic inference to derive the chance value for CH selection by a sensor node in the cluster formation process. The other notable difference lies in the CH election process. We follow the same procedure adopted by the LEACH protocol, whereas the MOECS uses a CH candidate based competition. Therefore, performance comparisons are not made.

2.6 Gupta Fuzzy Protocol

A number of researchers have used Fuzzy Logic to extend network lifetime and minimize the energy consumption of the network. The Gupta protocol [12] uses a Fuzzy Logic approach to select CHs. The FIS designer considered three descriptors: energy level, concentration, and centrality, each divided into three levels, and one output which is chance, divided into seven levels. The system also uses 27 IF-THEN rules. In this protocol there are two stages (set-up and steady-state) as in LEACH. The difference between the two protocols lies in the set-up stage where the BS needs to collect energy level and location information for each node, and evaluate them in the designed FIS to calculate the chance for each node to become a CH. The BS then chooses the node that has the maximum chance of becoming a CH.

After the CH selection, everything (advertising message, join CH message, and the steady-state stage) will be the same as in LEACH.

2.7 CHEF Fuzzy Protocol

The CHEF protocol (Cluster Head Election mechanism using Fuzzy Logic in Wireless Sensor Networks) [13], uses a Fuzzy Logic approach to maximize the lifetime of WSNs. It is similar to the Gupta protocol but it does not need the BS to collect information from all nodes. Instead the CHEF protocol uses a localized CH selection mechanism using Fuzzy Logic. Each node chooses a random number between 0 and 1. If this random number is smaller than P_{opt} , the node calculates the chance using an FIS and advertises a candidate message with the chance. The message indicates that the node is a candidate for CH with the value of chance. P_{opt} is calculated as:

$$P_{opt} = p \times \alpha \qquad (2.19)$$

Where p is as in Equation (2.12) and α is a constant value that defines the ratio of the candidate for cluster head. The node then listens for candidate messages from nodes within radius r in equation (2.20).

$$r = \sqrt{\frac{area}{\pi.N.p}} \qquad (2.20)$$

The node with the largest chance is selected as CH.

After the CH selection, everything (advertising message, join CH message, and the steady-state stage) will be the same as in LEACH.

The FIS uses two variables: energy residual and distance between nodes, one output and 9 IF-THEN rules.

2.8 LEACH-FL Protocol

In [14] the LEACH-FL (Improving on LEACH Protocol of Wireless Sensor Networks Using Fuzzy Logic) protocol is proposed. This protocol uses Fuzzy Logic to improve the LEACH protocol by considering three different parameters: energy level, node density, and distance between the CH and the BS. This model is the same as the Gupta protocol with a set-up stage and a steady-state stage, except in the set-up stage it chooses different parameters to apply in the designed FIS to obtain their output called ‘ probability’ for each node. The following formula *G(i)* is to get the value of ‘probability’:

$$G(i) = \frac{\sum_{j=1}^{n} x_j . u(x_j)}{\sum_{j=1}^{n} u(x_j)} \qquad (2.21)$$

Where, $u(x_j)$ is a membership function degree of set j, and x_j is the output ‘probability’ value on x-axis that intersection with $u(x_j)$.

Each node of the WSN computes a *G(i)* value through the FIS that proposed from the researchers. The computed results using the above equation show that the *G(i)* value is between 0.665 and 12.2335.

In the LEACH protocol, every node generate a random number between 0 and 1, and then the nodes that have a number less than the threshold value will be chosen to be the CHs. The researchers converted their *G(i)* to *F(i)* using a linear technique, which is shown below in equation (2.22):

$$F(i) = 1 - \frac{G(i) - 0.665}{12.2335 - 0.665} \qquad (2.22)$$

In every round, the comparison will be between the *F(i)* value instead of *G(i)* and threshold value, and if the *F(i)* of a node is smaller, then the node will be chosen to be the CH. Everything else in both stages will be as in LEACH and Gupta.

2.8 Conclusion

In this chapter we offered a brief presentation about clustering techniques, an overview of Fuzzy Logic and some of the constituent parts of the Fuzzy Logic. This was follow by explanation of the LEACH Protocol because it is the basis of our work. A number of variations on the LEACH protocol using Fuzzy Logic were examined, namely Gupta [12], CHEF [13] and LEACH-FL [14]. For each protocol we described the parameters used in its FIS.

CHAPTER 3 Improving LEACH Protocol Using Fuzzy Logic Approach

3.1 Introduction

In this chapter we present design details for our cluster formation protocol. The FLCFP (Fuzzy Logic Cluster Formation Protocol) is based on a Fuzzy Logic approach to help non-cluster nodes select a cluster head (CH) using three parameters to calculate the CH chance value. The sections of this chapter are organized as follows: in section 3.2 we describe the FLCFP protocol and follow it with an explanation of the FIS design in section 3.3, all inputs and output parameters are discussed in section 3.4, a step by step method to determine the chance value using the centroid method with a simple example to clarify our design system work is discussed in section 3.5. Finally, we summarize our chapter in section 3.6.

3.2 FLCFP Protocol

This section provides details for the network operational model. Our goal is to prolong the lifetime of WSNs by improving the LEACH protocol using a Fuzzy Inference System (FIS), which provides the process of formulating the mapping from a given input(s) to output(s) using Fuzzy Logic [30]. Our fuzzy clustering formation structure is divided into rounds similar to LEACH, each clustering round being composed of a set-up stage and a steady-state stage.

The main difference between FLCFP and LEACH lies in the set-up stage precisely at the period of cluster formation.

As we explained in chapter 2, in LEACH, each non-CH node receives a "join cluster" message from all the CHs and replies to the message that has the strongest signal strength.

In FLCFP the cluster formation phase is different from LEACH. The non-CH nodes compute a chance value for each CH by applying the FIS. In our FIS design three descriptors are considered namely energy level of the CH, the distance between the CH and the BS, and the distance between the CH and the node. Following this step the node joins the CH that has the largest chance value.

The complete implementation details including CH selection and cluster formation in the FLCFP protocol are shown as pseudo code in Appendix B.

3.3 FIS Design

As we mentioned previously in the chapter 2 section 2.3.6, the concept of Fuzzy Logic foundations on four steps: fuzzification, rule evaluation, aggregation, and defuzzification. In our proposed model we used the most frequencies method used in [12] [14] [13], and [31], which is called the Mamdani method in Fuzzy Logic toolbox. This method allows us to describe how our system is working in an easy and simplified mathematical way, as inference techniques and our Fuzzy Logic system design illustrate in Figure 3.1.

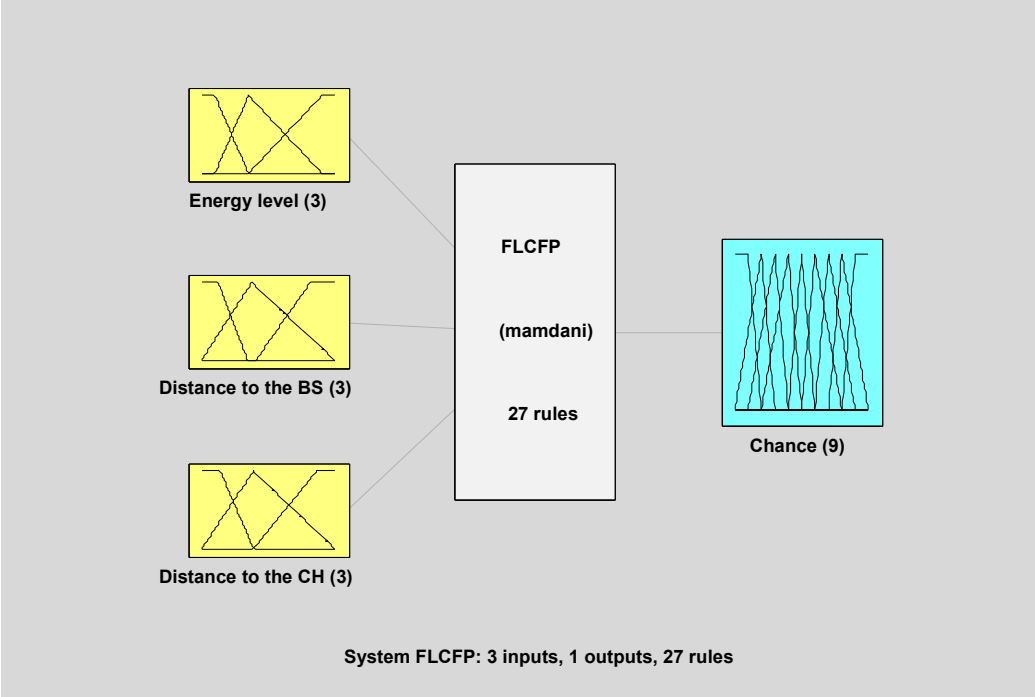

Figure 3.1: Fuzzy Inference System for FLCFP

3.4 Our FIS Parameters and Rules

In our proposed model, FLCFP, we use three parameters: energy level of the CH, distance between the CH and the BS and the distance from non-CHs to the CH. We choose these parameters because of their importance for extending the network lifetime as shown in the previous chapter.

To study how much they are effecting the lifetime of the network, and to make these parameters more flexible, we divided each linguistic variable that we used to represent these parameters into three levels: *low*, *medium*, and *high* for energy level of the CH; and *Close, medium*, and *far* for the distance to the BS and the distance between the CHs and the node. Moreover, many types of membership functions are available in the MATLAB Fuzzy Logic toolbox including Triangle, Trapezoidal, sigmoidal, Gaussian, S-shape, and Z-shape. However, the Triangle and Trapezoidal membership functions are more useful

than the other types because their degree is more easily determined. Formulas for Triangle and Trapezoidal membership functions are shown in Appendix A. Therefore, we chose to use them to present our parameters as illustrated in Figures 3.2, 3.3, and 3.4. In these figures we can see that, we present the linguistic variable for middle level (*medium*) by a triangle membership function, while we present both sides levels (*low, high, close, and far*) by a trapezoidal membership function. To give our incidence feature of flexibility we divided the linguistic variable for chance value into 9 levels as follow: *very weak, weak, and little weak, little medium, medium, high medium, little strong, strong, and very strong.* And once again the trapezoidal membership function represents both sides, and triangle membership function represents other chance levels as shown in Figure 3.5.

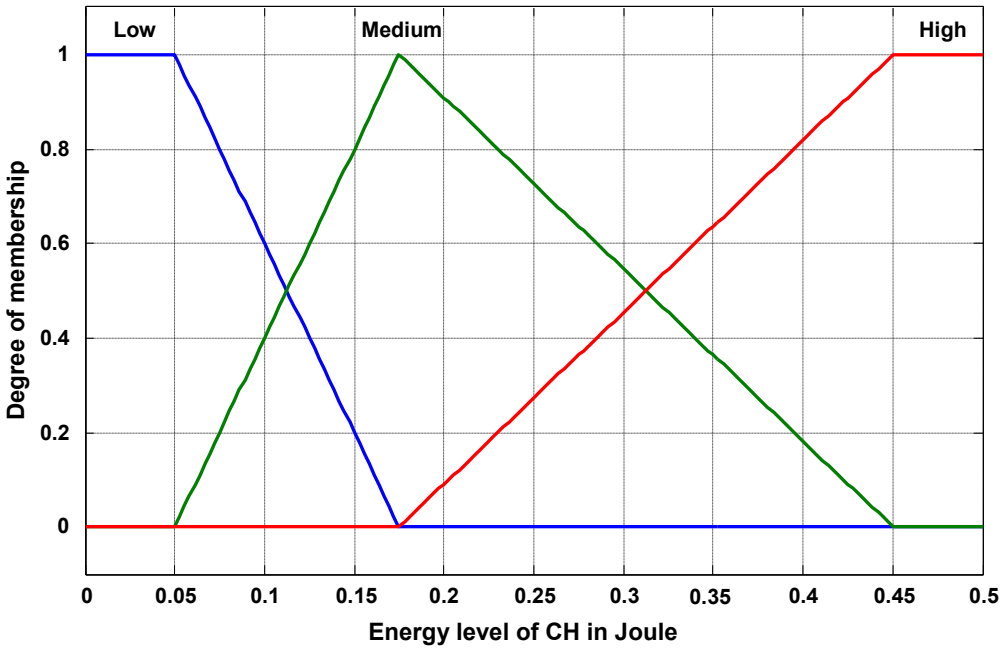

Figure 3.2: Fuzzy set of energy level of the CH

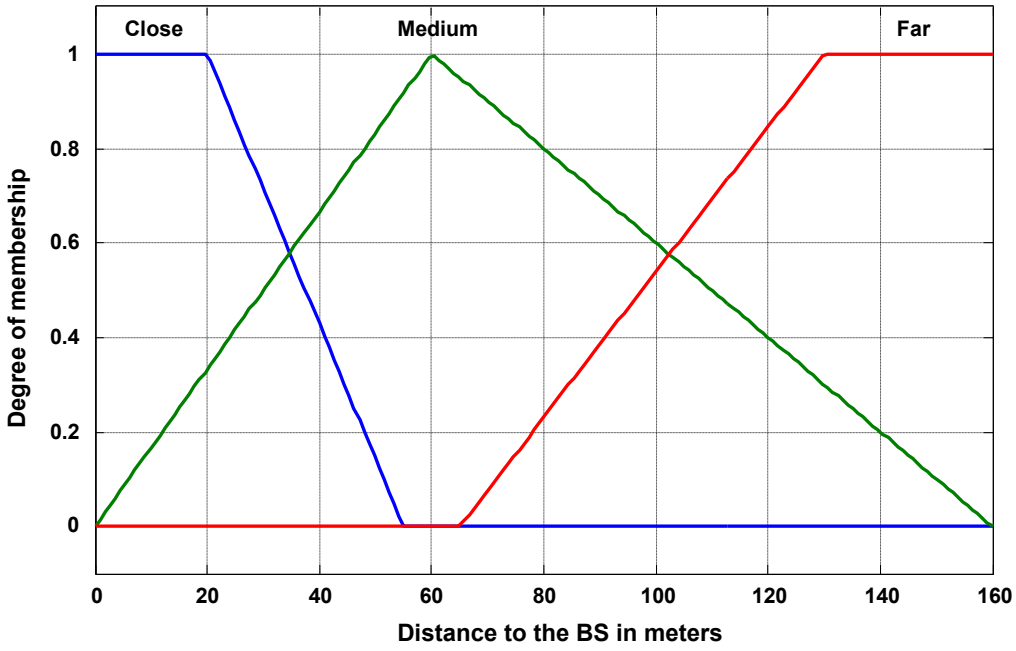

Figure 3.3: Fuzzy set of distance to the BS

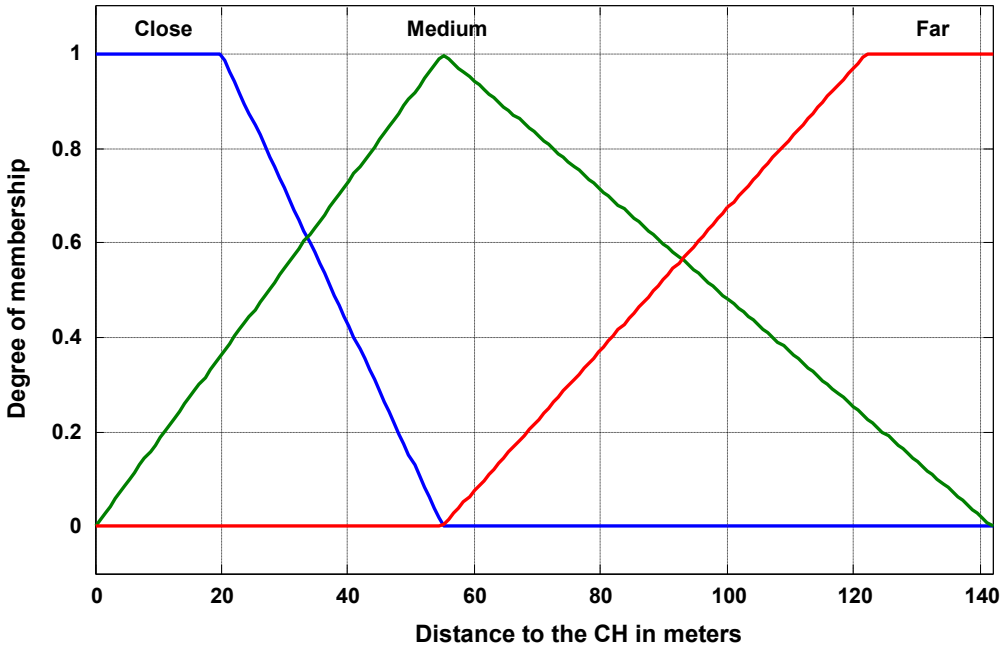

Figure 3.4: Fuzzy set of distance to the CH

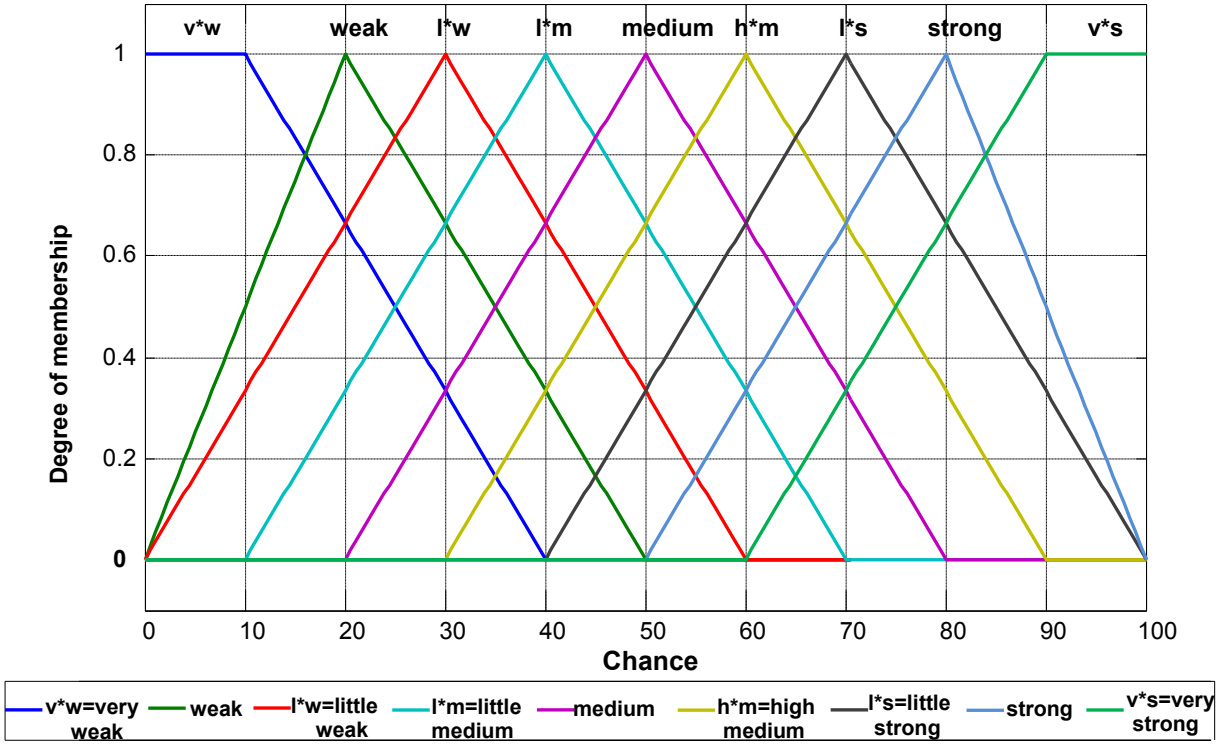

Figure 3.5: Fuzzy set of chance value

To determine the maximum values for our parameters in our FIS model, we have used equations (3.4), (3.5) and (3.6):

$$Max\ energy = initial\ energy \tag{3.4}$$

$$Max\ distance\ to\ the\ BS = \sqrt{(BS_x)^2 + \left(BS_y\right)^2} \tag{3.5}$$

$$Max\ distance\ to\ the\ CHs = \sqrt{xm^2 + ym^2} \tag{3.6}$$

Where, (BS_x, BS_y) is the position of the BS on x and y axis respectively, and (*xm, ym*) is size of the network.

Since we have three parameters, each divided into three levels, we have 3^3=27 possible chance value shown in Table 3.1 below that represents our fuzzy IF-THEN rule. These rules fall between two extremely cases as shown next:

Case (1): *If (energy is low) and (distance to the BS is far) and (distance between the CH and the nod is far) then (chance is very weak)*

Case (2): *If (energy is high) and (distance to the BS is close) and (distance between the CH and the nod is close) then (chance is very strong)*

Table 3.1: Fuzzy Inference System IF-THEN rules

Energy level	**Distance To the BS**	**Distance To the CH**	**Chance**
Low	Far	Far	Very weak
Low	Far	Medium	Weak
Low	Far	Close	Little weak
Low	Medium	Far	Weak
Low	Medium	Medium	Little Weak
Low	Medium	Close	Little medium
Low	Close	Far	Little Weak
Low	Close	Medium	Little medium
Low	Close	Close	Medium
Medium	Far	Far	Little weak
Medium	Far	Medium	Little medium
Medium	Far	Close	Medium
Medium	Medium	Far	Little medium
Medium	Medium	Medium	Medium
Medium	Medium	Close	High medium
Medium	Close	Far	Medium
Medium	Close	Medium	High medium
Medium	Close	Close	Little strong
High	Far	Far	Medium
High	Far	Medium	High medium
High	Far	Close	Little strong
High	Medium	Far	High medium
High	Medium	Medium	Little strong
High	Medium	Close	Strong
High	Close	Far	Little strong
High	Close	Medium	Strong
High	Close	Close	Very strong

Each one of these rules has a number between 0 and 1 called a weight. Generally, this weight is 1 and thus has no effect at all on the implication process. From time to time may weight one rule relative to the others by changing its weight value to something other than 1 [30].

3.5 Determination of Cluster-Head Chance Value:

In this section, we describe how to use a popular fuzzy inference technique called the Mamdani technique that we mentioned and described in chapter 2. To show and clarify how we use FIS to determinate CH chance value by the node, we will consider the simple example for that in the following:

Assume that, we have a CH with energy level (= 0.1 J) and it is located at a distance of (140 m) from the BS. Also the distance between the CH and the node that is used to calculate the chance value for every CH (= 142m). The following four steps provide details for the calculation of chance value in the FIS.

3.5.1 Step 1: Input of Crisp Value and Fuzzification

First we will forward our inputs which are crisp values as we assumed, energy level of the CH (=0.1), the distance to the BS (=140) and the distance to the CH (=142) to our FIS.

Depending on these three crisp numbers, we will determine the value of membership function, which is the intersection point of the value of our parameters (energy level of the CH, the distance to the BS and the distance to the CH) with the degree of the membership function, which we will use in the next step. These membership functions are illustrated in Figures 3.6, 3.7 and 3.8 respectively.

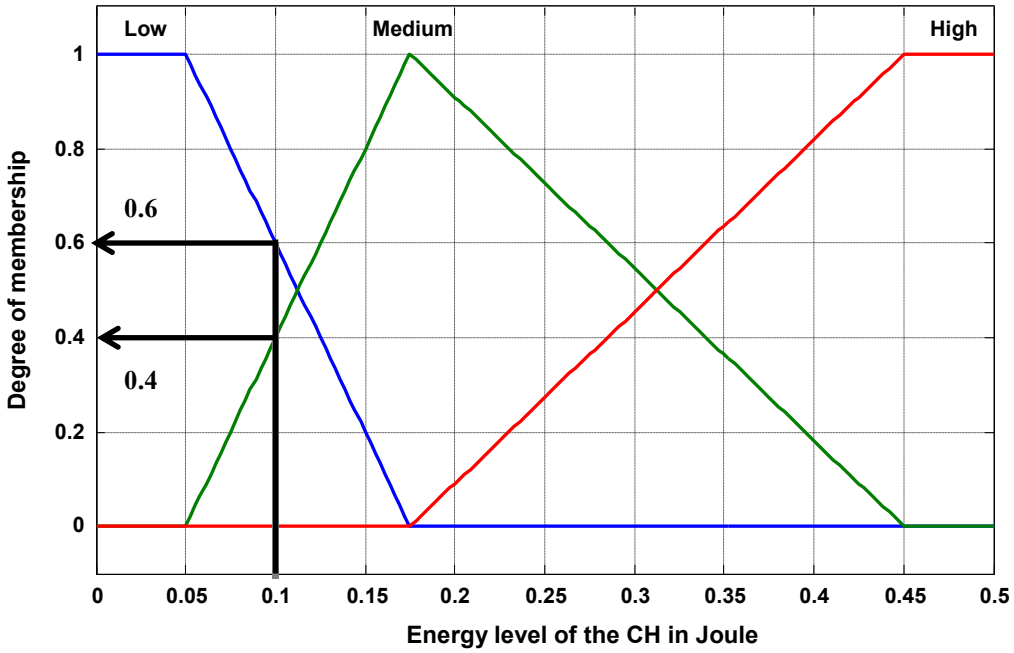

Figure 3.6: Fuzzification of crisp Energy level (=0.1J)

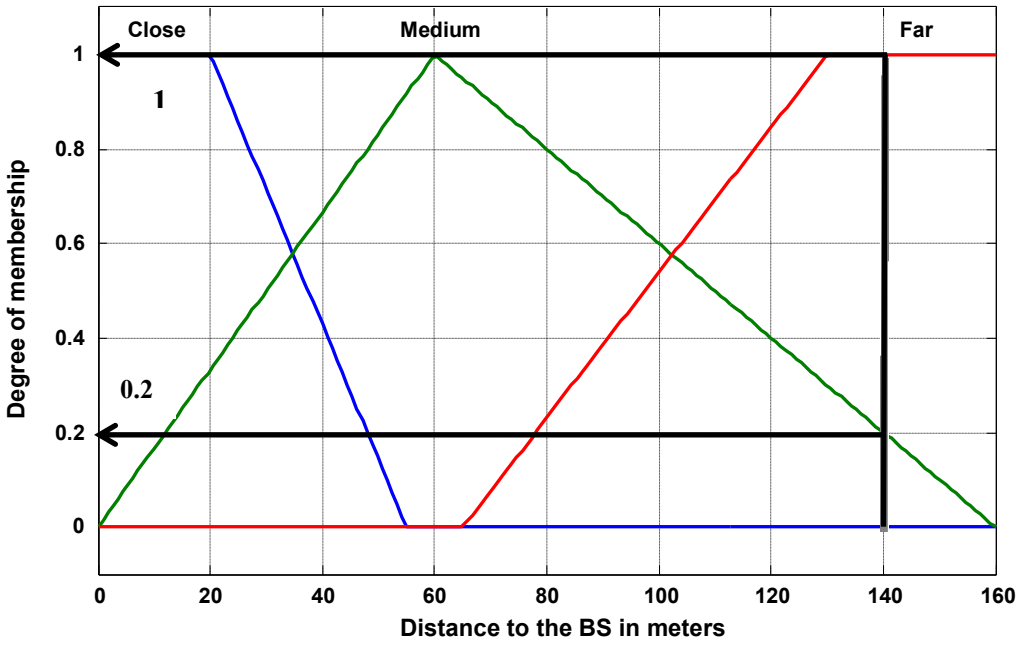

Figure 3.7: Fuzzification of crisp distance to the BS (=140m)

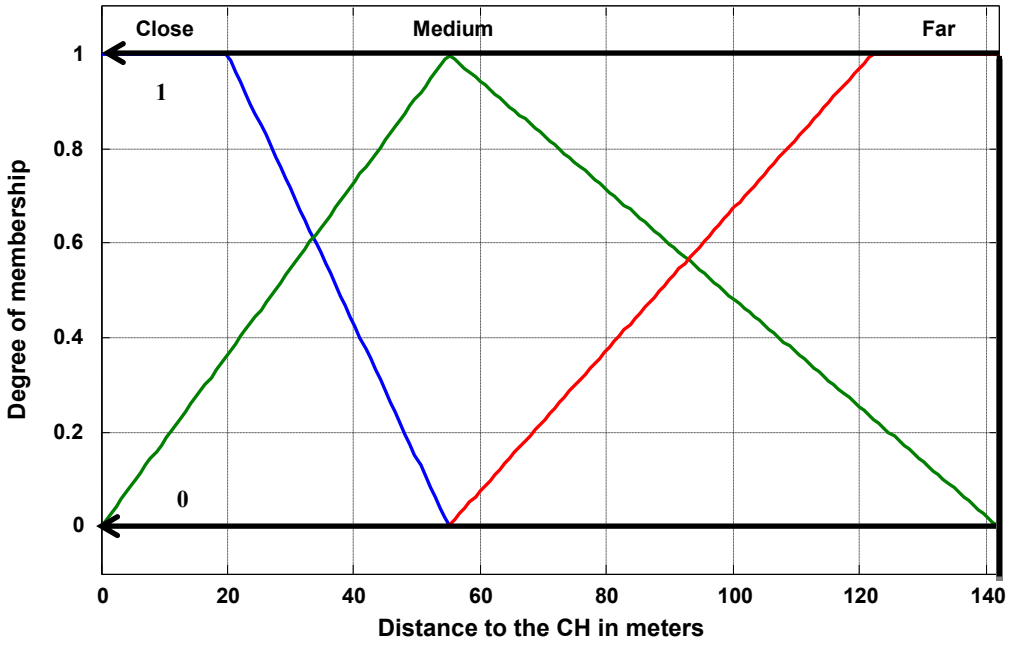

Figure 3.8: Fuzzification of crisp distance to the CH (=142m)

3.5.2 Step 2: Rule Evaluation

After the fuzzification step has been completed and the membership values have obtained, we supply/feed these values to our IF-THEN rules to determine our new fuzzy output set. Where, our fuzzy IF-THEN rules have multiple entrances, which are the three variables we have identified previously, and the fuzzy operator (AND), which simply selects minimum of our three membership values is used to get a single number as we show in the Table 3.2.

Table 3.2: Evaluation of fuzzy IF-THEN rules

Rule No.	Energy level	Distance to the BS	Distance to the CH	Chance
1	Low (=0.6)	Far (=1)	Far (=1)	Very weak (min=0.6)
2	Low (=0.6)	Far (=1)	Medium (=0)	Weak (min=0)
3	Low (=0.6)	Far (=1)	Close (=0)	Little weak (min=0)
4	Low (=0.6)	Medium (=0.2)	Far (=1)	Weak (min=0.2)
5	Low (=0.6)	Medium (=0.2)	Medium (=0)	Little Weak (min=0)
6	Low (=0.6)	Medium (=0.2)	Close (=0)	Little medium (min=0)
7	Low (=0.6)	Close (=0)	Far (=1)	Little Weak (min=0)
8	Low (=0.6)	Close (=0)	Medium (=0)	Little medium (min=0)
9	Low (=0.6)	Close (=0)	Close (=0)	Medium (min=0)
10	Medium (=0.4)	Far (=1)	Far (=1)	Little weak (min=0.4)
11	Medium (=0.4)	Far (=1)	Medium (=0)	Little medium (min=0)
12	Medium (=0.4)	Far (=1)	Close (=0)	Medium (min=0)
13	Medium (=0.4)	Medium (=0.2)	Far (=1)	Little medium (min=0.2)
14	Medium (=0.4)	Medium (=0.2)	Medium (=0)	Medium (min=0)
15	Medium (=0.4)	Medium (=0.2)	Close (=0)	High medium (min=0)
16	Medium (=0.4)	Close (=0)	Far (=1)	Medium (min=0)
17	Medium (=0.4)	Close (=0)	Medium (=0)	High medium (min=0)
18	Medium (=0.4)	Close (=0)	Close (=0)	Little strong (min=0)
19	High (=0)	Far (=1)	Far (=1)	Medium (min=0)
20	High (=0)	Far (=1)	Medium (=0)	High medium (min=0)
21	High (=0)	Far (=1)	Close (=0)	Little strong (min=0)
22	High (=0)	Medium (=0.2)	Far (=1)	High medium (min=0)
23	High (=0)	Medium (=0.2)	Medium (=0)	Little strong (min=0)
24	High (=0)	Medium (=0.2)	Close (=0)	Strong (min=0)
25	High (=0)	Close (=0)	Far (=1)	Little strong (min=0)
26	High (=0)	Close (=0)	Medium (=0)	Strong (min=0)
27	High (=0)	Close (=0)	Close (=0)	Very strong (min=0)

3.5.3 Step 3: Aggregation of the Rule Outputs

After the fuzzification and rule evaluation have been done, the aggregation step will start. The aggregation is a process of the union of all the outputs obtained from applying all rules (27 rules in our FIS model). Since we are looking at aggregating all our rules we have used an (OR) Fuzzy Logic operator. The OR operator simply selects the maximum of our rule evaluation values, to generate the new aggregate fuzzy set that we will use in next step. The Figure 3.9 below illustrates the aggregation output of the rules.

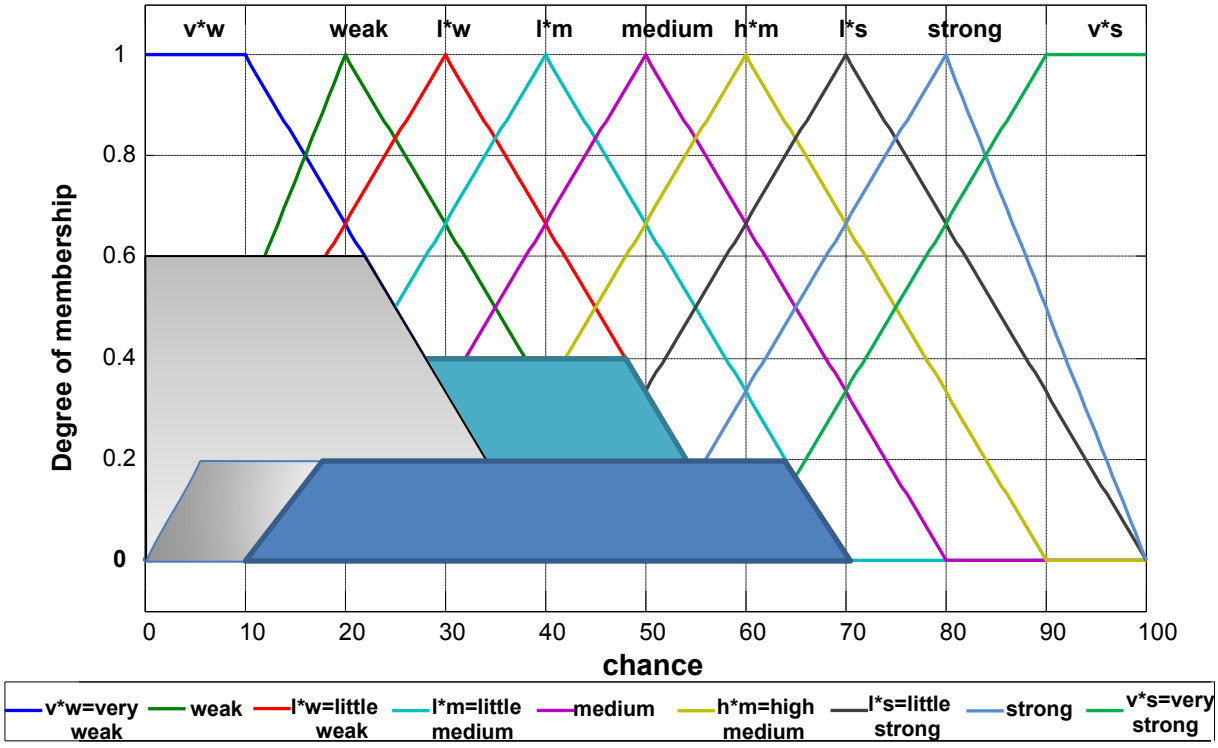

Figure 3.9: Output of evaluation of fuzzy IF-THEN rules

3.5.4 Step 4: Defuzzification

The last step is defuzzification, where we will obtain our chance value. As mentioned in section 3.3, we have used the Mamdani technique to calculate the implication value, and

the Centroid defuzzification method to find the final crisp number, which represents the CH election chance value to form a cluster formation. Therefore, the Center Of Area (COA) will be used in the centroid defuzzification, which we can compute by the following equation (3.7).

$$COA = \int \mu_A(x) * xdx / \int \mu_A(x)dx \qquad (3.7)$$

Where, ∫ denotes an algebraic integration, and $\mu_A(x)$ is degree of membership function of set A.

By applying the values we got from step 3 previously in equation (3.7) and calculating the algebraic integration, we determine the chance value for electing a CH to form a cluster formation approximately equal to (= 26.7). This amount is equal to the amount we got when applying our parameters, that we assumed above in our example in our designed FIS for FLCFP and the Figure 3.10 illustrate the centroid point.

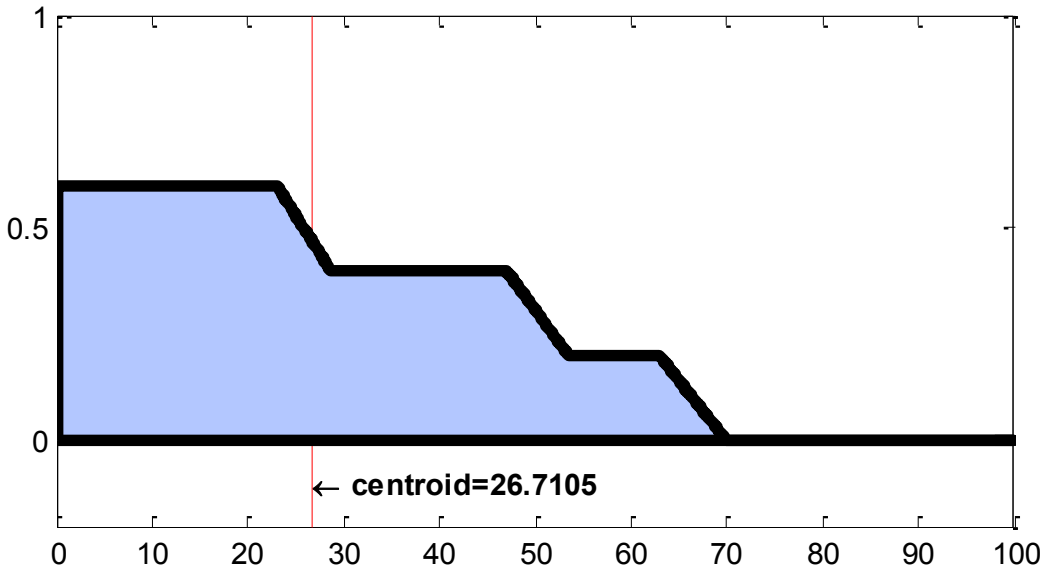

Figure 3.10: The Centroid point

In order, perhaps the most popular defuzzification method is Centroid defuzzification, which is frequently used in [34] [17]. So we chose it to calculate our chance value for clustering formation. Therefore, Centroid defuzzification returns the Center Of Area under the curve. The centroid is the point along the x-axis about which this shape would balance [30].

However, we can use methods other than the centroid method. But, because it is useful and easier than other methods, we have been used it as we mentioned previously.

For completeness, we list the five defuzzification methods that are provided in the BATLAB Fuzzy Logic toolbox:

- **Centroid:** centroid of area.
- **Bisector**: bisector of area.
- **mom:** mean value of maximum.
- **som:** smallest (absolute) value of maximum.
- **lom:** largest (absolute) value of maximum.

Additionally, we provide some examples for chance value calculated by the same way above in the Table 3.3 below:

Table 3.3: Examples for Chance value

No.	Energy level	Distance to the BS	Distance to the CH	Chance (%)
1	0.25	80	71	50.7%
2	0.05	30	20	42.2%
3	0.45	30	20	77%
4	0.5	10	50	75.1%
5	0.5	100	130	56.9%
6	0.15	100	130	36.1%
7	0.2	17	15	64.5%
8	0.25	25	15	64.5%

If a particular node calculates the same chance value, as in case no. (7) and (8) in Table 3.3 , that node will chose to join the CH which has the higher level of energy, since the main purpose of our research is energy conservation and prolonging the lifetime of network. If the two CHs have the same energy, to break the tie between the CHs chance we choose the closer one to the BS. Then we use the distance to the CH.

3.6 Conclusion

In this chapter we discussed the design details of our proposed FLCFP. We also provided a thorough review of Fuzzy Logic approach with all particulars including FIS design, three usage parameters, and the procedure to determine the chance value for CH.

CHAPTER 4 Simulation and Results

In this chapter, we analyze the performance of our algorithm via simulation conducted in MATLAB. We also compare our clustering algorithm with the LEACH algorithm. Since energy conservation is the primary objective of our work, performance metrics such as network lifetime, energy consumed per round, and the residual energy level of sensor nodes are of particular interest.

We use a network operation model as defined in [15]. In this model, the network operation progresses in rounds. Each round in turn consists of a cluster set-up and data transmission phase. In the cluster set-up phase, a set of new CHs is elected from the active nodes and the remaining nodes become cluster members. In the data transmissions phase each sensor node sends a fixed amount of data to its CH, which is later forwarded to the BS after aggregation. Using this network operation model allows the network lifetime metric to be measured in data collection rounds till the very first node runs out of its energy. The event of first node death (FND) is has been used extensively in literature [33] [36] [37]. Other metrics for measuring network lifetime such as a percentage of nodes deaths and last node death (LND) are also cited in literature. For the work presented in this thesis, we will focus on FND metric to test and analyze our algorithm.

4.1 Assumptions

The proposed system model uses the assumptions listed below:

1- All the nodes in WSNs are having the same hardware, communication, and computation capabilities.

2- The nodes are deployed randomly in a 2-D plane using uniform distribution.

3- All the nodes have the equal initial energy.

4- The base station position is located outside of the WSNs.

5- Nodes consume energy according to the model described in subsection 2.4.2.

6- Nodes are location unaware, i.e. they are not equipped with any global positioning system (GPS) device.

4.2 Scenarios

Three different scenarios are considered for the performance analysis of our proposed protocol. Most of the experiments in all scenarios used general parameters described as follow: a network of *N* sensor nodes each with a 0.5J initial energy deployed in an area 100×100 meters^2. We simulate the algorithm for ten runs each with a number of rounds. So in every round in our simulation each node in the clusters collects the 4000 bits data packets and transmits them to its CH until it runs out of energy. The CH carries the aggregated message to the BS, located outside of the WSNs at $50, 150$ meters. Unless otherwise stated, all results are represented using an average taken over ten independent runs. General configuration parameters are shown in Table 4.1.

Table 4.1: Configuration parameters

Parameters	Value
Network size	(100 x 100)m^2
Base station location	(50 , 150) m
Number of nodes	100
Initial Energy	0.5 J
Data packet size	4000 bits
Probability to become CH	0.1

4.2.1 Scenario1

The configuration parameters for this scenario are shown in Table 4.1. Based on the above network features and parameters, our algorithm was implemented and examined using MATLAB.

Figure 4.1 shows the number of alive nodes for both algorithms and it is clear that our FLCFP outperforms the LEACH algorithm.

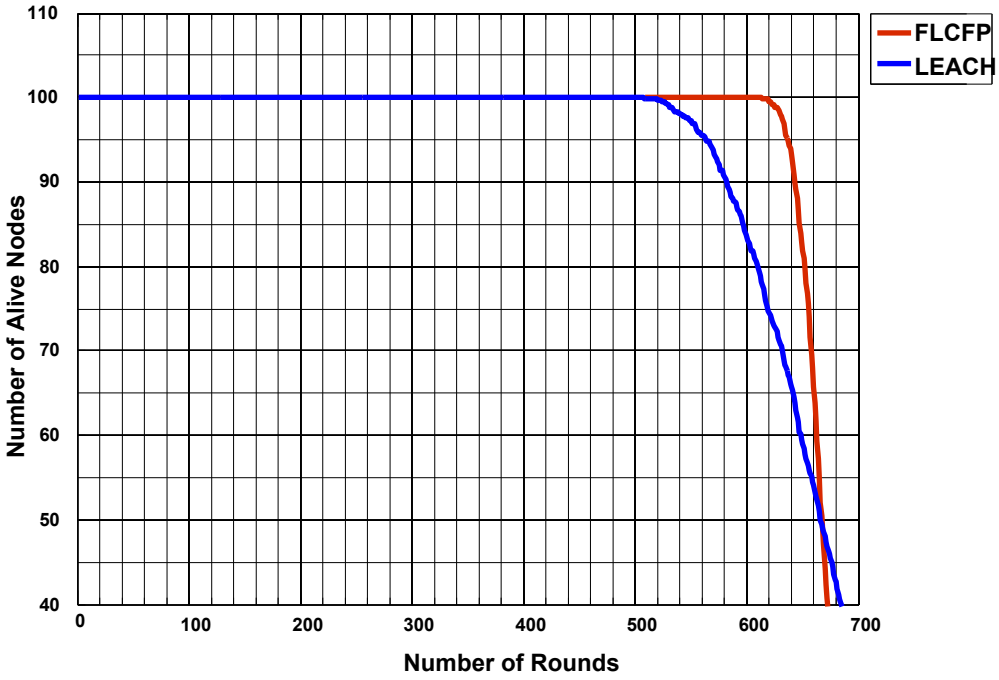

Figure 4.1: Scenario 1: Number of alive nodes according to Number of Rounds

To clarify the significance of our results and to show the improvement over LEACH, the FND metric is plotted for both protocols. Table 4.2 shows the FND result for both protocols for the ten runs. We observe that the average value for FND for LEACH is 541, whereas, in the FLCFP, the average is 634. The same data is plotted in Figure 4.2 for clarity.

Table 4.2: scenario1: FND metrics for both protocols

Run	FLCFP	LEACH Protocol	Difference
	FND	FND	
1	622	535	87
2	634	520	114
3	638	556	82
4	637	552	85
5	630	508	122
6	623	527	96
7	638	548	90
8	645	545	100
9	632	555	77
10	642	572	70
Average	634.1	541.8	92.3

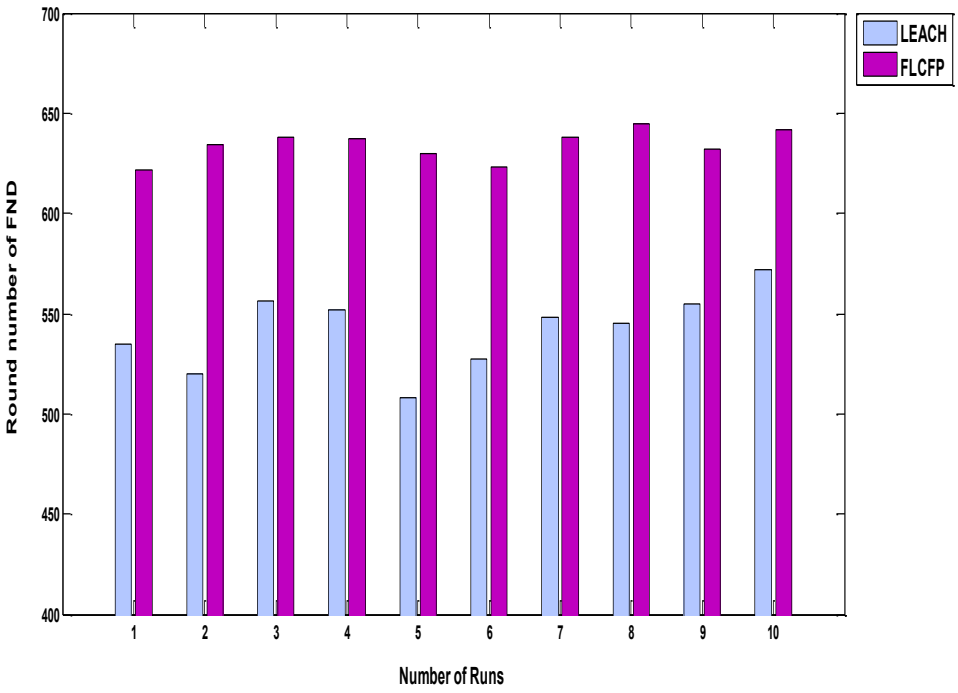

Figure 4.2: FND values for both protocols

We confirmed the results for the FND metric using statistical analysis that was based on testing a hypothesis that there is difference between both algorithms. The MINITAB software was used to conduct a Paired T-test. Our testing hypothesis will be:

H_0: *Difference = 0*

H_1: *Difference ≠ 0*

Paired T-Test and CI: FLCFP, LEACH

```
Paired T for FLCFP - LEACH

             N    Mean  StDev  SE Mean
FLCFP       10  634.10   7.53     2.38
LEACH       10  541.80  19.22     6.08
Difference  10   92.30  16.16     5.11


95% CI for mean difference: (80.74, 103.86)

T-Test of mean difference = 0 (vs not = 0): T-Value = 18.06  P-Value = 0.000
```

The MINITAB results show a p-value of zero indicating that the FLCFP network lifetime is significantly longer as compared to the LEACH protocol.

From the above results we can see that our proposed algorithm delays FND approximately 17% in average as compared with LEACH, a significant improvement in network lifetime. The improvement is attributed to that fact that LEACH uses only one parameter (the sensor node's local distance to the CH) in the cluster formation process. If a node receives multiple advertisements from neighboring CHs, it will choose the one located at the smallest distance. Conversely, a sensor node in our FLCFP considers three parameters (energy level of the CH, distance of the CH to the BS and the distance between the CH and the node) to calculate the chance value for each CH. Then it will choose to join the CH with the greatest chance value.

To understand the energy consumption behavior of the sensor nodes, we monitor the residual energy level of the nodes just before the FND in FLCFP (round 620).

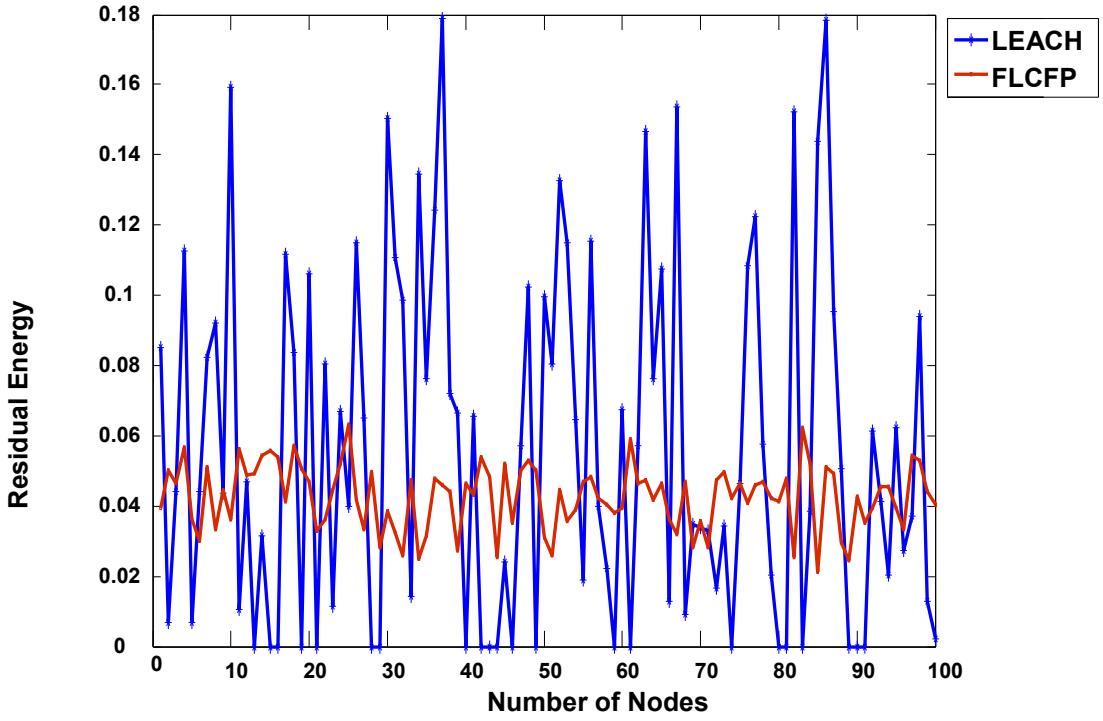

Figure 4.3: Residual energy of the nodes

From Figure 4.3, we can see that our FLCFP helps the nodes to consume energy in more uniform way. In the other words, the energy of the nodes in LEACH is quite variable with some nodes with high energy and some dead while in the FLCFP the node energies do not have these extremes.

4.2.2 Scenario 2

This scenario will be the same as the first one. All settings are exactly the same as in previous scenario except for the probability to become CH, which is '*p*' value. We selected a value of '0.05' for this scenario to examine the effect of the number of CHs on the FLCFP.

In Figure 4.4 we can see the number of alive nodes for both algorithms; and it is evident that our FLCFP performance is much better than the LEACH performance.

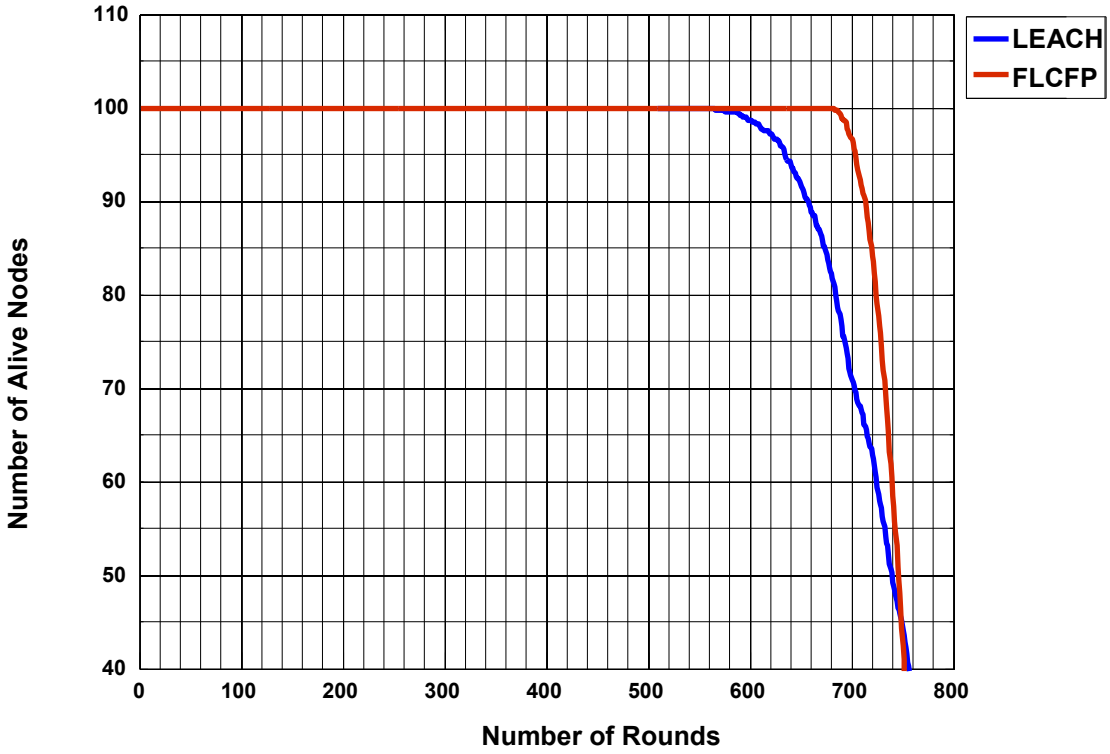

Figure 4.4: Scenario 2: Number of alive nodes according to the Number of Rounds

Furthermore, to display the differences and to clarify our results we have Table 4.4 and Figure 4.5 to present the FND metric and plot respectively for both protocols. Table 4.4 provides the FND values for ten independent runs for both protocols. We observe that the average FND value for LEACH which is 601, whereas, the average is 688 in the FLCFP.

Table 4.3: scenario2: FND metrics for both protocols

Run	FLCFP	LEACH Protocol	Difference
	FND	FND	
1	683	613	70
2	687	602	85
3	690	564	126
4	692	634	58
5	663	588	75
6	687	591	96
7	688	630	58
8	705	598	107
9	690	588	102
10	695	606	89
Average	688	601.4	86.6

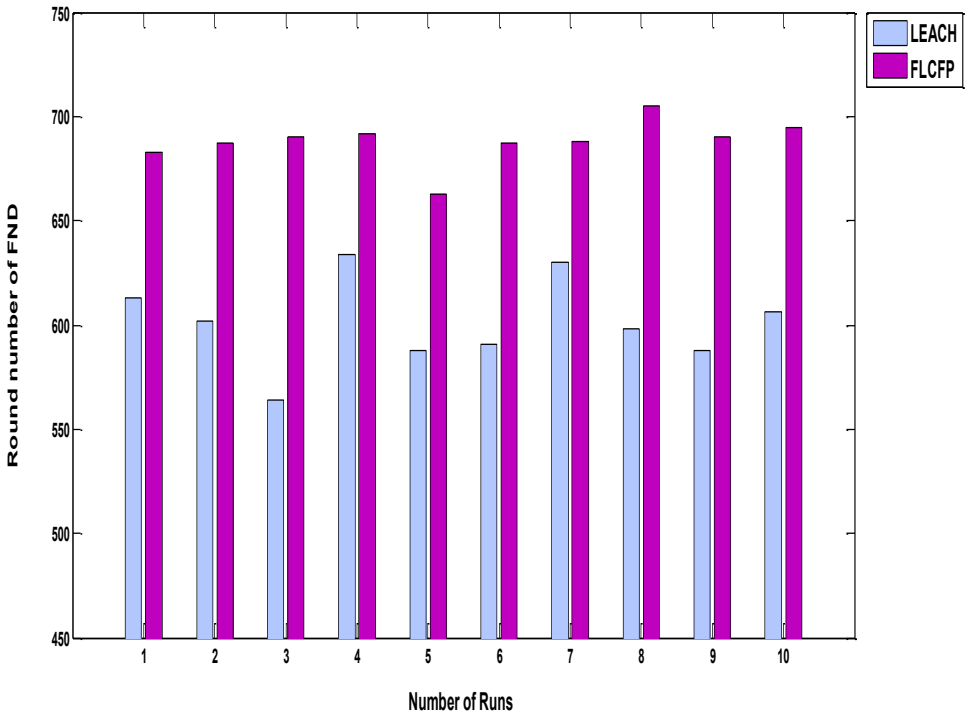

Figure 4.5: FND values for both protocols

After that, we will follow the same steps as the first scenario; to get some statistics we have used MINITAB, for testing hypothesis we assumed to prove that there is a difference between both protocols, and we will use Paired T-Test.

Our testing hypothesis will be the same for all our scenarios which are:

H_0: *Difference = 0*

H_1: *Difference ≠ 0*

Paired T-Test and CI: FLCFP, LEACH

```
Paired T for FLCFP - LEACH

             N    Mean  StDev  SE Mean
FLCFP       10  688.00  10.61     3.36
LEACH       10  601.40  20.86     6.59
Difference  10   86.60  21.98     6.95


95% CI for mean difference: (70.88, 102.32)

T-Test of mean difference = 0 (vs not = 0): T-Value = 12.46  P-Value = 0.000
```

The MINITAB results show a p-value of zero indicating that the FLCFP network lifetime is significantly longer as compared to the LEACH protocol.

From the above results we can see that our proposed algorithm delays FND approximately 14% as compared with LEACH, a significant improvement in network lifetime. As we mentioned in the scenario1 the improvement is attributed to that fact that LEACH uses only one parameter. Whereas, a sensor node in our FLCFP considers three parameters that effect energy consumed.

To understand the energy consumption behavior of the sensor nodes for this scenario and how the probability to become CH affects it, we monitor the residual energy level of the nodes just before the FND in FLCFP (round 685).

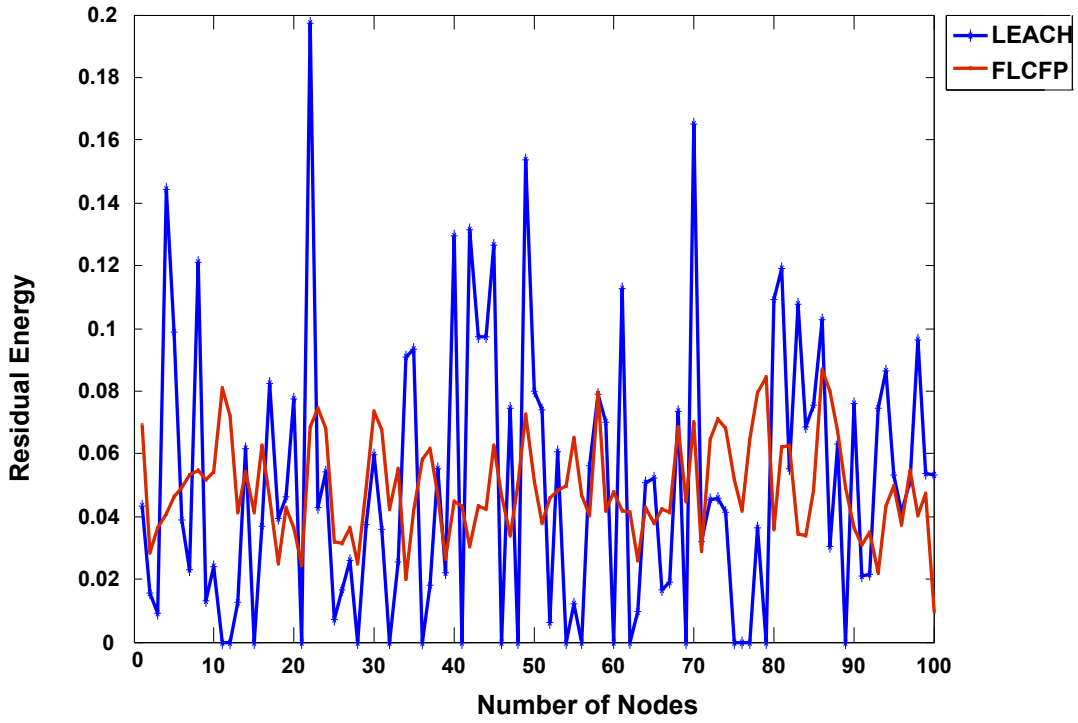

Figure 4.6: Residual energy of the nodes

As a result, the Figure 4.6 shows us that, the energy consumption behaviors are similar to that in scenario1. That means the energy of the nodes in LEACH is quite variable with some nodes with high energy and some dead while in the FLCFP the node energies do not have these extremes.

4.2.3 Scenario 3

The scenario3 is similar to the first scenarios. It is using the configuration parameters that listed in Table 4.1 except the number of nodes N. We selected 200 nodes to be deployed. This change allowed us to study the effect of the nodes density on the FLCFP behavior.

In Figure 4.7 we can see that the number of alive nodes for both algorithms; and it is obvious that our FLCFP performance is much better than the LEACH performance.

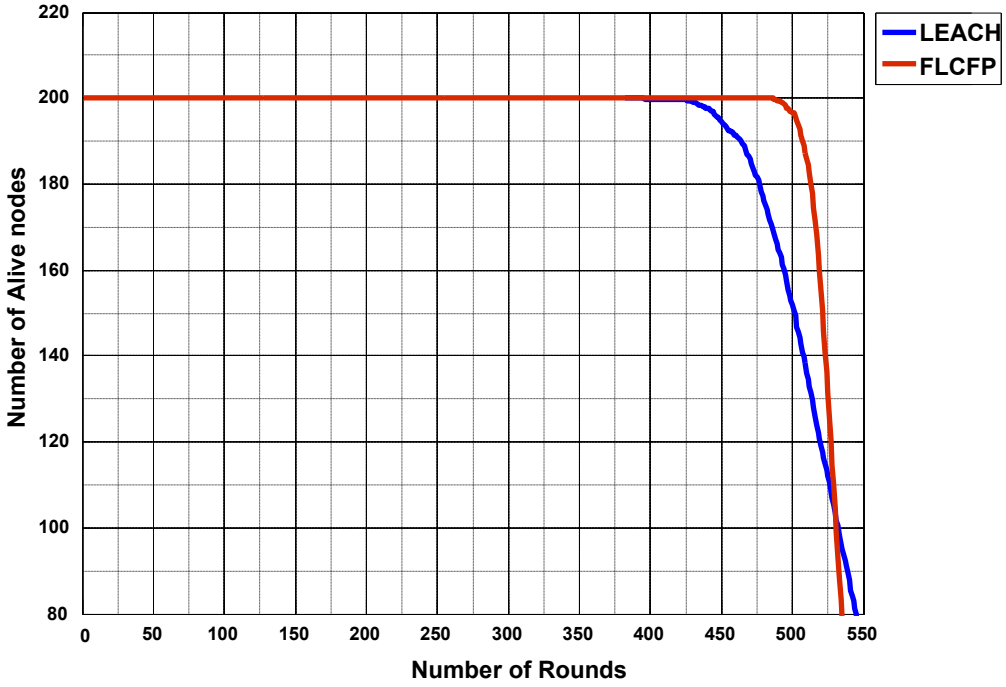

Figure 4.7: Scenario 3: Number of alive nodes according to the Number of Rounds

Moreover, to make the results more clear we present the FND metric and plot it for both protocols to show the differences. From the Table 4.6 below we can determine the average of the FND for ten independent runs for LEACH is 440, whereas, in the FLCFP the average is 496.

Table 4.4: scenario3: FND metrics for both protocols

Run	Proposed Protocol	LEACH Protocol	Difference
	FND	FND	
1	491	429	62
2	500	449	51
3	487	430	57
4	506	442	64
5	487	453	34
6	492	452	40
7	504	459	45
8	506	449	57
9	496	443	53
10	498	396	102
Average	496.7	440.2	56.5

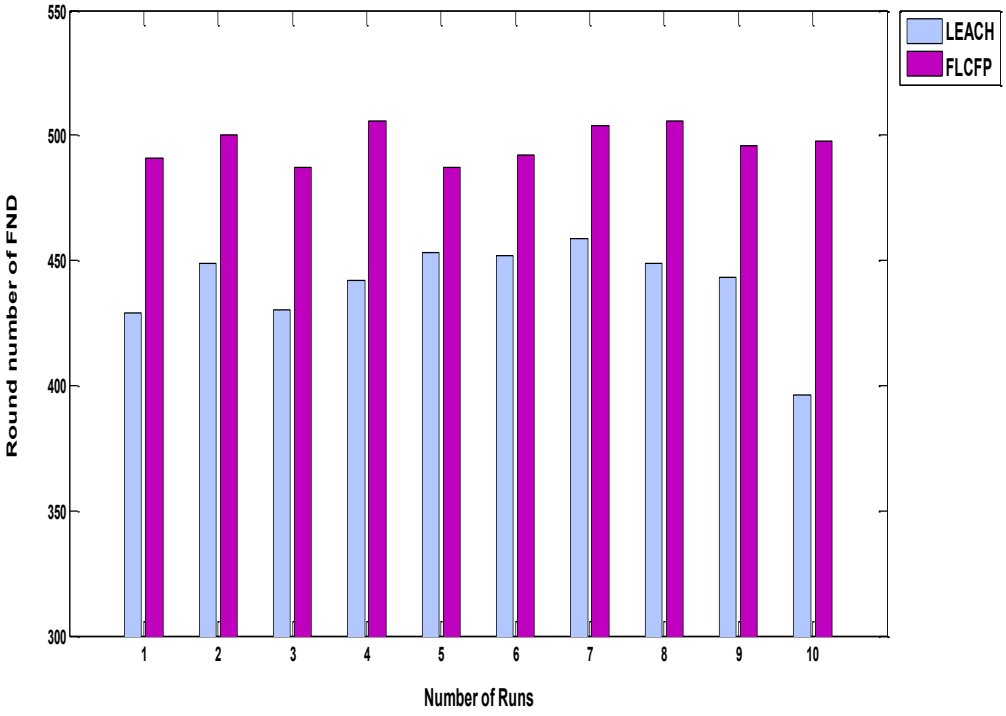

Figure 4.8: FND values for both protocols

Once again we have used MINITAB to get some statistics and for testing hypothesis. We assumed that there is a difference between both protocols, and Paired T-Test will be used in this exam.

Our testing hypothesis will be:

H_0: *Difference = 0*

H_1: *Difference ≠ 0*

Paired T-Test and CI: FLCFP, LEACH

```
Paired T for FLCFP - LEACH

             N   Mean  StDev  SE Mean
FLCFP       10  496.70   7.32     2.31
LEACH       10  440.20  18.27     5.78
Difference  10   56.50  18.59     5.88


95% CI for mean difference: (43.20, 69.80)

T-Test of mean difference = 0 (vs not = 0): T-Value = 9.61  P-Value = 0.000
```

For this scenario also, the MINITAB results show a p-value of zero indicating that the FLCFP network lifetime is significantly longer as compared to the LEACH protocol.

From the above results we can see that our proposed algorithm delays FND approximately 12.84% as compared with LEACH, a significant improvement in network lifetime. For the same reason that we mentioned in the two previous scenarios, the improvement is attributed to that fact that LEACH uses only one parameter in the cluster formation process. However, a sensor node in our FLCFP considers three parameters in the configuration the clusters.

To understand the energy consumption behavior of the sensor nodes, we monitor the residual energy level of the nodes just before the FND in FLCFP (round 495).

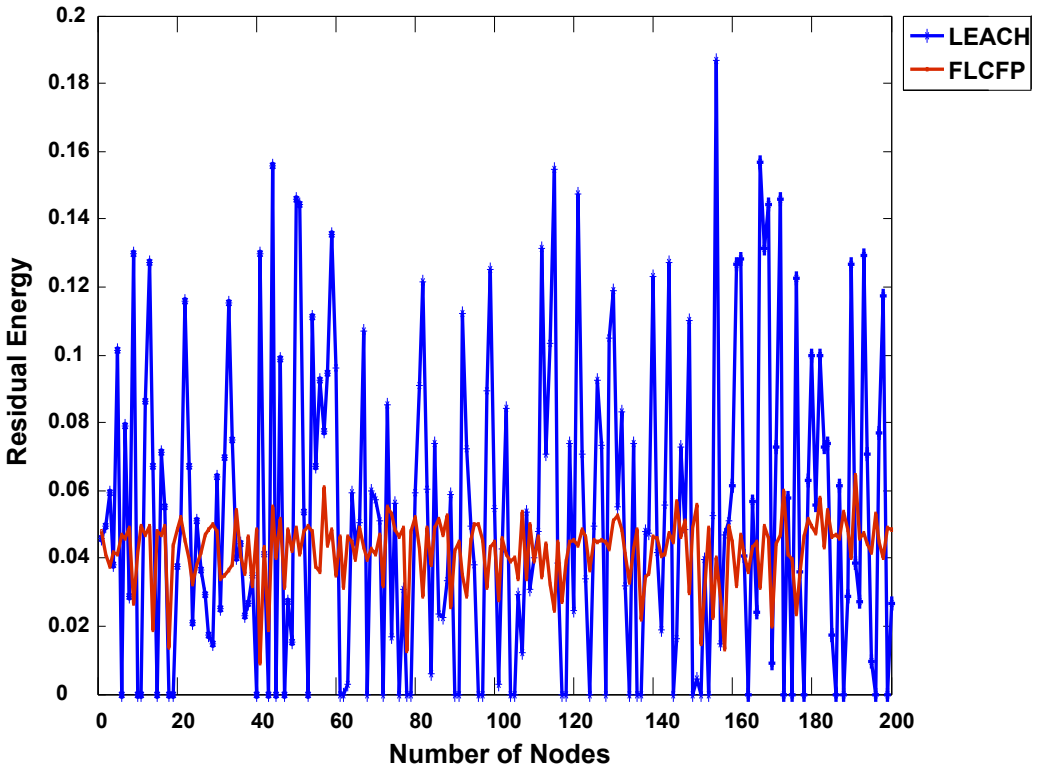

Figure 4.9: Residual energy of the nodes

According to Figure 4.9, we can see that the energy consumption behavior of the sensor nodes in our FLCFP is still more uniform than in LEACH, even with the increased density of nodes.

4.3 Conclusion

In this chapter, results were presented to demonstrate the performance of the FLCFP and its effects on energy conservation and network lifetime. We used performance metrics including FND, residual energy, and energy consumed per round to evaluate the behavior of FLCFP and LEACH. Simulation experiments were conducted using three different scenarios using two different values of p and the number of nodes. We observed that in all experiments our FLCFP improves network life in the range of 12% to 19% with respect the FND metric. The results were statistically tested using a paired T-test. We also observed that FLCFP optimizes the energy consumption behavior of sensor nodes such that the dissipation rate is balanced among all nodes in the network. In conclusion, we demonstrated that FLCFP achieves significant energy savings and enhances network lifetime as compared to the LEACH protocol.

We would like to emphasize that explicit comparison with the other protocols including Gupta, CHEF and LEACH-FL were not made due to the following two reasons. One, the LEACH protocol is considered the baseline for performance analysis. Two, Gupta, CHEF and LEACH-FL modify the LEACH protocol in terms of CH selection, whereas FLCFP modifies LEACH in terms of cluster formation.

CHAPTER 5 Conclusion and Future work

In this thesis, we presented a novel approach for cluster formation in WSNs using Fuzzy Logic to enhance the network lifetime. We have analyzed the performance of our protocol through simulations, and compared its performance with the LEACH protocol. Our conclusions from the performance analysis are articulated in Section 5.1, followed by the directions of future research in Section 5.2.

5.1 Conclusion

- Our approach improved the network lifetime in the range of 12% to 19% compared to LEACH. This improvement is attributed to the fact that our proposed protocol used three parameters in the cluster formation process compared to LEACH that uses only one. Relying on one parameter (distance from node to the CH) is not suitable to produce optimal clusters because the energy consumption behaviour in WSNs is a very complex phenomenon. Factors such as remaining energy of CHs and distance of CH from the sink also contribute to energy dissipation in WSNs. We have focused on combining diverse parameters using an FIS which produces a chance value for sensor nodes to decide which CH is the most suitable for overall energy efficiency
- We also demonstrated that by using three parameters, that the energy is consumed in a balanced fashion in the network. This behaviour was confirmed by monitoring the residual energy levels of sensor nodes. As a result, our protocol

does not suffer from an early node death problem that could happen in the case where some nodes consume energy at a higher rate than the others.

- Similar to LEACH, our protocol operates in a distributed manner where decisions are made based on the local information only. That is, each node makes its decision in the cluster formation process using the messages received from CHs in its communication range. No global state maintenance is required.
- Our FLCFP shows how fuzzy logic can be used in the cluster formation process to distribute the tasks and energy consumption over all the nodes in a WSN.

5.2 Future work

In future, some extensions of the FLCFP approach can be applied. The next list contains some topics in this area:

- A natural extension of this work will be to combine FLCFP approach with one of the Fuzzy Logic approaches for CH selection, to minimize the overall energy consumption, and extend the network lifetime.
- By adjusting the shape of each fuzzy parameter, maybe we can achieve additional improvement in the network lifetime and energy consumption.
- In future work, we may choose different parameters such as centrality or/and density of nodes instead one or more of the chosen parameters. Choosing various parameters may further improve the performance of our work.
- Experiments with the BS located inside the network as well as mobile sensors, may be included in the future work in this area.

BIBLIOGRAPHY

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks,* vol. 52, no. 12, pp. 2292 - 2330, 2008.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Communications,* vol. 38, no. 4, pp. 393 - 422, 2002.

[3] Holger Karl, Andreas Willig, Protocols and architectures for wireless sensor networks, 2005.

[4] Waltenegus Dargie, Christian Poellabauer, Fundamentals of Wireless Sensor Networks: Theory and Practice, August 2010.

[5] Wu, Ruomin, "Maximum Lifetime Routing and Resource Allocation in Wireless Sensor Networks," 2011.

[6] Rajashree.V.Biradar, V.C .Patil, Dr. S. R. Sawant, Dr. R. R. Mudholkar, "CLASSIFICATION AND COMPARISON OF ROUTING PROTOCOLS IN WIRELESS SENSOR NETWORKS," *UbiCC Journal,* vol. 4, pp. 704-711.

[7] Kazem Sohraby, Daniel Minoli, Taieb F. Znati, Wireless sensor networks: technology, protocols, and applications, 2007.

[8] M. Kuorilehto, "A Survey of Application Distribution in Wireless Sensor Networks," *EURASIP Journal on Wireless Communications and Networking,* pp. 774-788, 2005.

[9] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, 2002.

[10] D. Malan, T. Fulford-Jones, M. Welsh, and S. Moulton, "CodeBlue: An Ad Hoc Sensor Network Infrastructure for Emergency Medical Care," in *Proceedings of the MobiSys Workshop on Applications of Mobile Embedded Systems (WAMES 04)*, 2004.

[11] T. Arampatzis, J. Lygeros, and S. Manesis, "A Survey of Applications of Wireless Sensors and Wireless Sensor Networks," *Intelligent Control, Proceeding of the 2005 IEEE International Symposium on, Mediterrean Conference on Control,* pp. 719 - 724, 2005.

[12] I. Gupta, "” Cluster head election using fuzzy logic for wireless sensor networks"," Halifax, Nova Scotia, March 2005.

[13] J. Kim, S. Park, Y. Han and T. Chung, "“CHEF: Cluster Head Election mechanism using Fuzzy logic in Wireless Sensor Networks”,," in proceeding of the ICACT, 2008..

[14] Ran, H. Z., "Improving on LEACH Protocol of Wireless Sensor Networks Using Fuzzy Logic," *Journal of Information & Computational Science,* 2010.

[15] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan,, "“An application-specific protocol architecture for wireless microsensor networks”," in *IEEE Transactions on Wireless Communica-tions, 1(4):*, 2002.

[16] E. H. Mamdani, "Application of fuzzy logic to approximate reasoning using linguistic synthesis," *IEEE Transactions on Computers,* Vols. C-26, pp. 1182-1191, 1977.

[17] Fung, Glenn, "A Comprehensive Overview of Basic Clustering Algorithms," June 22, 2001.

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

[19] Ameer Ahmed Abbasi, Mohamed F. Younis, "A survey on clustering algorithms for wireless sensor networks.," *Computer Communications,* Vols. 30(14-15):, pp. 2826-2841, 2007.

[20] Kim, Wooyoung, "Survey topic: Parallel Clustering Algorithms Detailed Annotated Bibliography and Classification of the Results," 2009.

[21] S. N. Sivanandam, S. Sumathi, S. N. Deepa, ” Introduction to fuzzy logic using MATLAB”, 2007.

[22] [Online]. Available: http://www.seattlerobotics.org/encoder/mar98/fuz/fl_part1.html.

[23] "The Net's Original Fuzzy Logic Archive - Since 1994," [Online]. Available: www.austinlinks.com/Fuzzy/.

[24] Timothy J. Ross ,, ”Fuzzy Logic with Engineering Applications”, 3rd edition, 2010., 2010..

[25] [Online]. Available: http://www.cs.cmu.edu/Groups/AI/html/faqs/ai/fuzzy/part1/faq-doc-2.html.

[26] Hans-Jürgen Zimmermann, “Fuzzy set theory--and its applications”, 4th edition., 2001.

[27] Stefania Montani, Lakhmi C. Jain, ” Successful Case-based Reasoning Applications-1”, 2010.

[28] Guney, K, "Comparison of Mamdani and Sugeno Fuzzy Inference System," *Progress In Electromagntics Research,* vol. 12, pp. 81-104, 2009.

[29] R. WOODMAN, "Razor Robotics," [Online]. Available: http://www.razorrobotics.com/articles/fuzzy-control-system/.

[30] [Online]. Available: http://www.mathworks.com/ Fuzzy Logic Toolbox user's guide.

[31] Bagci, Hakan, ""Fuzzy unequal clustering in wireless sensor networks",," January 2010.

[32] W. Heinzelman, A. Chandrakasan and H. Balakrishnan,, "“Energy-efficient communication protocol for wireless microsensor networks,”," in *in Proc. of the 33rd Annual Hawaii nternational Conference on System Sciences (HICSS), Maui, HI,*, Jan. 2000.

[33] Nauman Aslam, W. Robertson, W. Phillips and S. Sivakumar, "A Multi-Criterion Optimization Technique for Energy Efficient Cluster Formation in Wireless Sensor Networks," *Information Fusion - Elsevier Publication on Multi-Sensor Multi-Source Information Fusion,* 2011.

[34] Bergbreiter, Sarah Elizabeth, "Autonomous Jumping Microrobots," December 18, 2007.

[35] H. Karl and A. Willing, "“Protocols and Architectures for Wireless Sensor Networks”," John Wiley & Sons, Ltd Paperback edition, 2005.

[36] Fahmy, O. Younis and S., "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks," *Transactions on Mobile Computing,* vol. 3, no. 2004, pp. 660-669.

[37] M. Ye, C. F. Li, G. H. Chen, and J. Wu, "EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks," in *IEEE Int'l Performance Computing and Communications Conference (IPCCC)*, 2005.

[38] A.K. JAIN, M.N. MURTY, and P.J. FLYNN, "Data Clustering: A Review," *ACM Computing Surveys,* vol. Vol. 31, no. No. 3, September 1999.

[39] Estivill-Castro, Vladimir, "Why so many clustering algorithms — A Position Paper," vol. 4, no. 1, p. 65.

[40] G. Ran, H. Zhang, and S. Gong, ""Improving on LEACH Protocol of Wireless Sensor Networks Using Fuzzy Logic”," *Journal of Information & Computational Science,* 2010.

[41] II, Rui Xu and Donald Wunsch, "Survey of clustering algorithms," in *Neural Networks, IEEE Transactions on,*, May 2005.

[42] M. Bani Yassein, A. Al-zou'bi, Y. Khamayseh, W. Mardini, "Improvement on LEACH Protocol of Wireless Sensor Network (VLEACH)," *International Journal of Digital Content Technology and its Applications,* vol. Volume 3, Number 2, June 2009.

[43] O. Younis, M. Krunz, and S. Ramasubramanian, "Node clustering in wireless sensor networks: recent developments and deployment challenges," *IEEE Network,* vol. 20, pp. 20-

25, 2006.

[44] Waltenegus Dargie, Christian Poellabauer, Fundamentals of Wireless Sensor Networks: Theory and Practice, August 2010.

[45] Sitar, D. M. Doolin and N., "Wireless sensors for wildfire monitoring," 2005.

[46] Simoes, Marcelo Godoy, "Introduction to Fuzzy Control," Colorado School of Mines, Engineering Division, Golden, Colorado.

[47] Mehdi Saeidmanesh, Mojtaba Hajimohammadi, and Ali Movaghar, "Energy and Distance Based Clustering: An Energy Efficient Clustering Method for Wireless Sensor Networks," 2009.

[48] M.J. Handy, M. Haase and D. Timmermann,, "“Low energy adaptive clustering hierarchy with deterministic cluster-head selection,”," in *in Proc. 4th International Workshop on Mobile and Wi reless Communications Network,*, Sept. 2002.

[49] 15 APR 93. [Online]. Available: http://www.cs.cmu.edu/Groups/AI/html/faqs/ai/fuzzy/part1/faq-doc-2.html.

APPENDIX A: MEMBERSHIP FUNCTIONS

Triangle membership function

$$\mu_A(x) = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ \frac{c-x}{c-b}, & b \le x \le c \\ 0, & c \le x \end{cases} \qquad \text{(A.1)}$$

Or, more efficiently, by

$$\mu_A(x) = \max\left(min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \qquad \text{(A.2)}$$

Where, the points (*a* and *c)* locate the "feet" of the triangle and the point (*b*) locates the "top" of the triangle.

Trapezoidal membership function

$$\mu_A(x) = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \\ \frac{d-x}{d-c}, & c \le x \le d \\ 0, & d \le x \end{cases} \qquad \text{(A.3)}$$

Or, more efficiently, by

$$\mu_A(x) = \max\left(min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right) \qquad \text{(A.4)}$$

Where, the points (*a* and *d*) locate the "feet" of the trapezoid and the points (*b* and *c*) locate the "shoulders" of the trapezoid.

APPENDIX B: PSEUDO CODE FOR FLCFP

The operation of the set-up stage in the FLCFP is outlined in the Algorithm (1). In each round there is CH selection and cluster formation. The CH selection takes place in lines 1-11. Initially, all nodes set their states to "PlainNode". The threshold, T, is set for the current round using equation 2.12. Each node generates a random number between 0 and 1. If the generated number is less than the threshold value, the node will be a CH, change its state to "CLUSTERHEAD" and advertise a CH message to the network. Lines12-27 describe the cluster formation process. The nodes receive all CHs messages. The non-CHs nodes compute the chance value for each CH, then set their CHID with the CH that has the maximum chance value ID, and their CHJoinStatus = 1. In lines 28-35 the non-CH node sends a join message to the chosen CH and becomes part of that cluster.

```
1: nodeState⟵PlainNode
2: cluster ⟵ empty
3: T ⟵threshold for current round
4: for each node
5:          temp_rand⟵rand (0,1)
6:          if temp_rand < T then
7:                nodeState⟵CLUSTERHEAD
8:                cluster = cluster+1
9:                Advertise CH Message (ID)
10:         end if
11: end for
12: On receiving all CH Messages
13: for each node (i)
14:         if nodeState = PlainNode then
15:               Chancefuzzy (length (C),1) = 0
16:               For each CH (j)
17:                     Get energy, distance to the BS, and
                        distance to the CH
18:                     Compute the chance value using fuzzy IF-
                        THEN mapping rules
19:                     chancefuzzy (j,1) = chance value
20:               end for
```

```
21:              temp = max (chancefuzzy)
22:              id1 = Find (chancefuzzy ==temp)
23:              node(i).CHID=id1;
24:              node(i).CHJoinStatus=1
25:              end for
26:         end if
27:   end for
28: For each node (i)
29:         For each CH (j)
30:              if CH (ID) = node(i).CHID
31:                   Send CH join Message (ID) to this CH
32:                   add node ID to the cluster Members
                      list
33:              end if
34:         end for
35: end for
```

Algorithm (1): FLCFP set-up stage