IMPLEMENTATION OF FUZZY C-MEANS CLUSTERING ALGORITHM FOR FINGERPRINT-BASED INDOOR LOCALIZATION TECHNIQUE IN WIRELESS SENSOR NETWORKS

DWI JOKO SUROSO

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF ENGINEERING IN COMPUTER ENGINEERING
INTERNATIONAL COLLEGE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2012

KMITL - 2012 - IC - M – 006 – 003

Thesis Title	Implementation of Fuzzy C-Means Clustering Algorithm for Fingerprint-based Indoor Localization Technique in Wireless Sensor Networks
Student	Mr. Dwi Joko Suroso
Student ID.	53601152
Degree	Master of Engineering
Program	Computer Engineering
Year	2012
Thesis Advisor	Asst. Prof. Dr. Panarat Cherntanomwong
Co-advisor	Prof. Dr. Jun-ichi Takada

ABSTRACT

Fuzzy C-Means (FCM) is known as one of the clustering algorithm that can handle the uncertainty data. In the wireless node localization, the data can be obtained as the power in the receiver through the received signal strength indicator (RSSI) in the device. One of the most common techniques used in the localization technique is Fingerprint-based technique. This technique allows us to have the database of information stored inside the system. One of the benefits of this technique is its accuracy to estimate the target position. The density of sensor nodes and the quality of the database are the main important points in this technique. The implementation of the FCM as the clustering algorithm in the database of fingerprint helps the faster calculation in the nearest neighbor algorithm as the pattern matching algorithm. Technically speaking, the nodes that are suspected as the position of the target node are clustered in a cluster. It allows the pattern matching algorithm to calculate within members in that cluster.

The validation of the proposed method is verified by conducting the indoor experiment. The result of clusters and the accuracy of FCM to cluster the target node and sensor nodes are the main concern. The accuracy of the system will be shown as the estimated target position error as the result of pattern matching algorithm calculation. From the indoor experiment results, the FCM can cluster the target node in almost of target position together with fingerprint nodes in 2 Dimensional (2D) and 3 Dimensional (3D), respectively. From the results, we can conclude that the FCM can be implemented as supporting algorithm in the fingerprint-based indoor localization system. The power and issues of time consumption in the proposed technique are composed as the future works for our research.

ACKNOWLEDGEMENTS

I am sincerely and heartily grateful to my advisor, Asst. Prof. Dr. Panarat Chentanomwong for her support, invaluable advice, and continuous guidance throughout my study in the past two years. Without her patients, valuable suggestions, and comments, I would not be able to work in this challenging and interesting topic. For conducting this research successfully, I would also like to show my thoughtful gratitude to Prof. Dr. Jun-ichi Takada (Tokyo Institute of Technology, Japan) who served as my co-advisor, and for his constructive discussion, comments, and his point of views as well as provision in conducting experiments. This thesis would not have been possible without his insight and expertise.

I gratefully acknowledge ASEAN University Network/Southeast Asia Engineering Education Development Network (AUN/SEED-Net) project and the Japan International Cooperation Agency (JICA) for financial support for Master Degree within 2 years. Furthermore, I also extend my sincere appreciation to KMITL for giving me the great opportunity to do research in a warmly and friendly environment.

I would like to thank to Assoc. Prof. Pitikhate Sooraksa and his lab members for the grateful ideas and for kind friendship during my study. I would like to thank my laboratory mates, especially Mr. Nuttapon Nakarach and his team for conducting experiment and fruitful discussion, Mr. Sisongkham Phimmasean, Mr. Thanapong Chuenurajit for their help during experiment. Moreover, my sincere gratitude also goes to Mrs. Arunee Suvarnatarn, who helps and gives me a guideline during the whole period of my study. To all my friends and colleagues, I wish to express many thanks for the enjoyable and stimulating atmosphere that they provide with their companion and friendship. I am also grateful to friends in the Indonesia Student Association in Thailand for making my study life so enjoyable and colorful.

Finally, I would like to dedicate this thesis to my beloved family for their unflagging love, encouragement, inspiration, dedication and persistent supports. It is their tolerance and love that carry me through the peaks and troughs of my entire life. I would also like to thank my beloved lover, Indrawati Mei Purwandari, who always supports me every time.

Bangkok, Thailand　　　　Dwi Joko Suroso

26 September 2012

Table of Contents

List of Tables

List of Figures

Figure **Page**

Chapter 1

Introduction

1.1 Background

The advancement of wireless sensor node technology leads the rapid research in the wireless sensor network (WSN) application. Node localization on WSN-based technique is a particular area in the wireless communication. The wireless node localization can be defined as verifying the location of a sensor node within the network with respect to a reference. Localization usually refers to the process of dynamically determining the location(s) of one or more target node(s). The various techniques are implemented during the development of localization system. Localization techniques can be classified into several categories, i.e., received signal strength (RSS), time of arrival (ToA) or time difference of arrival (TDoA), angle of arrival (AoA), location fingerprint based techniques, and their combinations [1]-[4]. However, the RSSI and time based techniques require at least three base stations (BSs) for the system. For AoA, the accuracy of the localization depends only from the accuracy of angle estimation. By considering the equipment used and the accuracy, the fingerprint technique is the localization technique that can make the localization system simple and accurate since only a single BS and comparatively simple equipment can be employed.

Location fingerprint technique as pattern matching-based localization is commonly used as non-real time localization technique. This technique allows us to have the signal information recorded as database. In order to estimate the target location, its current signal information is compared with those in the database. In our work, fingerprint-based technique is utilized. Received signal strength indicator (RSSI) from ZigBee hardware is used as the fingerprint information. As the IEEE 802.15.4 standard for wireless communication, ZigBee has various advantages, i.e. cost-effective, low-power consumption, security, robustness, reliability and it can support low data rates. ZigBee has parameters which can be deployed as estimation parameters to find the location of the unknown sensor nodes. The useful parameters of this wireless standard for localization are RSSI and link quality indicator (LQI) [5]. The disadvantages of the fingerprint technique are time consuming and requiring a lot of effort during the process of collecting data [6].

This thesis proposes a solution to the problems using clustering data technique for fingerprint-based technique. The Fuzzy C-means (FCM) algorithm is applied. FCM algorithm is first proposed by Bezdek [7] to be used in cluster analysis, pattern recognition, and image processing. By applying the

FCM method, a partition of the feature vectors into different regions can be performed. In the case of our work, FCM algorithm is applied to cluster the fingerprint database. Each fingerprint node as transmitter which provides the RSSI information from reference nodes (receiver) is assigned a degree of belonging to cluster head rather than completely being a member of just one cluster. Therefore, the nodes close to the boundary of a cluster may become members of the cluster with a degree approximating the degree of belonging to the neighbor clusters.

The improvement of accuracy can be achieved while the RSSI values as database are clustered into vectors that have similarity among others. In this thesis, we focus on an indoor localization system using ZigBee standard as the sensor (reference) nodes and the target node, respectively. As we know, the accuracy of the indoor localization is affected by many factors i.e., propagation in the media, wall, etc. By using the location fingerprint technique, the accuracy to estimate the target node location is expected to be improved. Nearest neighbor algorithm is used as pattern matching algorithm. We present the approach in the clustering technique using FCM as the new method for measuring the target location in the fingerprint technique. The clustering technique is also well-known as the technique which has the spatial advantages related to scalability and efficient communication [8]

This chapter begins with some background issues in indoor localization system and the motivations of the work are presented in Section 1.2. The objectives of the thesis are well described in Section 1.3 and finally the outline of this thesis will be given in Section 1.4 as the last section in this chapter.

1.2 Motivations

Figure 1.1 depicts the four major problems in accurate and robust localization system. One of them is methodology/positioning algorithm which becomes our motivation to do this thesis work. As we know, positioning algorithm especially in indoor environment has many factors affecting the accuracy of the system. In this work, the radio localization using smart sensor nodes in WSN is proposed. This method is supported by the advancement of micro-mechanical systems and development of digital electronics technology. Radio localization is realized to be an important application for daily life and WSN is popularly selected to be used for a localization system. The localization system which provides the acceptable accuracy of position estimation, relatively cheap in cost, low power consumption, small in size, efficient, and simple in application is required. Regarding the accuracy of the system, many issues such as diffraction at edges, refraction by media with different propagation velocity and reflection in metallic objects is affected the localization system performance.

The smart, small, relatively cheap, and low power consumption-sensor node is deployed in our localization system. The accuracy of the system is achieved by the improving the quality of database in the fingerprint technique. The clustering technique is applied to obtain the efficient and effective communication system. The FCM is related with the grouping of fingerprint nodes which have the similarity with the target node in the area of interest. The estimated position accuracy is validated using the nearest neighbor algorithm. The position accuracy, simple and effective system are become our motivation conducting this research.

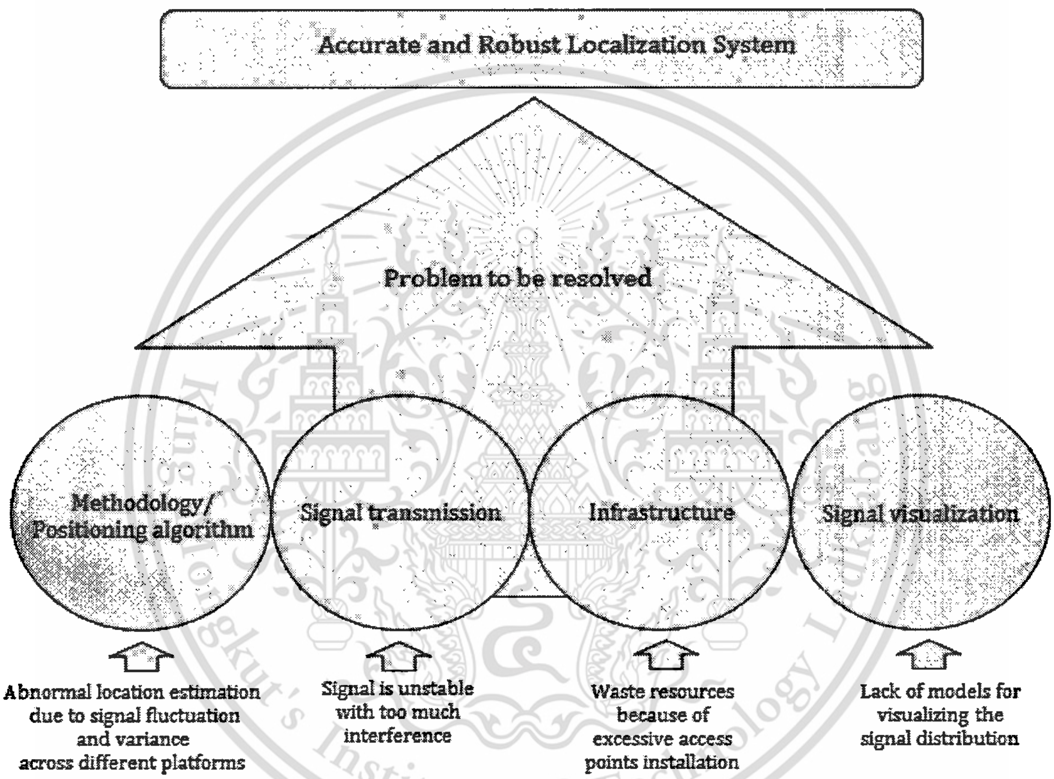


Fig. 1.1 Four major problems of the localization system

1.3 Objectives

In this research, the Fuzzy C-Means clustering algorithm is proposed to be implemented as the pre-processing method for the fingerprint-based technique in indoor localization scenario to investigate the possibility of reducing the power and achieving effective communication in wireless sensor network-based localization system. Implementation of FCM to cluster the fingerprint database and the target node in the area of interest is emphasized. The Nearest neighbor algorithm as pattern matching algorithm in fingerprint based technique is employed. So, the investigation of the FCM accuracy to cluster the fingerprint database is achieved.

1.4 Thesis Outline

The thesis is organized in five chapters and is briefly summarized as follows.

Chapter 1 introduces some background information and previous works about the research in the localization technique and the application of wireless sensor networks. The main idea in this chapter is to show the prospective benefit of implementing the clustering technique in the fingerprint-based localization technique.

Chapter 2 describes the literature reviews of WSN and indoor localization system. The basic concept of WSN related with the communication, platform and parameters are discussed. Several localization techniques are described in this chapter such as the range-based and fingerprint-based localization techniques. Fuzzy C-Means as one of the fuzzy clustering method and its implementation in our work are presented.

Chapter 3 presents methodology of the research including how to obtain database and to validate the performance of the fingerprint database. The simulation system parameter and hardware components for the research are also given.

Chapter 4 gives the simulation results and analysis of fingerprint based technique using FCM clustering algorithm for indoor localization system.

Chapter 5 offers the conclusion of the thesis by reviewing and highlighting its contributions which have been presented in this thesis. The recommendations for future works related to this research are given as well.

Chapter 2

Literature Reviews

The key interest in this thesis is the implementation and analysis of the Fuzzy C-Means clustering algorithm in the fingerprint-based localization technique. Some essential preliminaries are presented in this chapter, which will be described in the following chapter. Section 2.1 describes the overviews of a wireless sensor network (WSN) and its applications. Section 2.2 introduces the overview of radio localization system, while Section 2.3 reviews the algorithms of indoor localization techniques, Section 2.4 explains data clustering. Section 2.5 shows the ZigBee and Section 2.6 briefly explains Xbee as the hardware used in this research. And finally Section 2.7 presents the one of the classifications of localization technique, received signal strength (RSS) using fingerprint-based technique based on ZigBee IEEE 802.15.4 wireless communication standard.

2.1 Wireless Sensor Networks

2.1.1 The Definition of Wireless Sensor Networks

Wireless Sensor Networks (WSNs) is a small number of sensor nodes used to measure the properties of the environment and processing that information to generate new knowledge about the environment around it. WSNs become a reality because of the integration of three technologies:

(1) Micro electro-mechanical system (MEMS), which could make the sensor properties (mechanical parts), fit into very small chip.

(2) Digital electronics, which make the small chip (integrated with microcontroller) powerful enough to handle the incoming sensor data i.e., data fusion, data compression, and networking operations.

(3) Wireless communications, which relay the sensor data among many sensors.

WSN technology is a combination of embedded and wireless communications. The data transmission between the sensor nodes from a network of ad-hoc features of the sensor network based protocol ad-hoc is not required to install equipment for the network as a network in WLAN or GSM. In addition, sensor nodes in WSN are small in size and low power consumption. It can be installed in various environments. Sensor network technology is expected to be a propulsion technology in the era of computers everywhere (ubiquitous computing, pervasive computing) by creating artificial environments around us.

2.1.2 Wireless Sensor Network Architecture

The sensor nodes are usually scattered in a sensor field as shown in Figure 2.1. Sensor network architecture consists of three parts: a wireless gateway and the sensor unit including a base station.

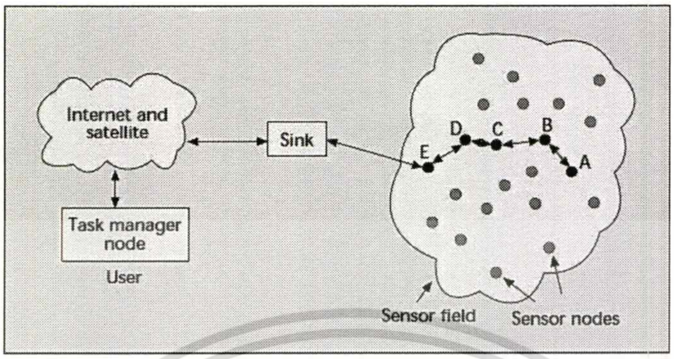


Fig. 2.1 Sensor nodes scattered in a sensor field [9]

The sink node is used as a bridge between sensor nodes and the internet and satellite. Each of these scattered sensor nodes has the abilities to collect data and route it back to the sink. Data are routed back to the sink by a multi-hop infrastructure-less architecture through the sink as shown in Figure 2.1. Wireless sensor network architecture consists of three parts: the sensor unit includes a gateway and base station (base station) as shown in Figure 2.1 with the large number of sensors embedded to store the data. Each unit includes a sensor which communicates wirelessly with this unit. Each unit includes a sensor control. It is able to control and manage its own work (self-organize). Each unit includes a sensor interface can work together (collaboration) and can transmit data to each other. Although the sensor unit cannot be contacted directly with the source, the sensor unit is able to send information to the other sensor units. The base station is responsible for storing the measured data from a sensor network [9].

2.1.3 Properties of a Sensor Unit in a Wireless Sensor Network

In order to build a practical WSN application, a sensor unit for WSN should have these following criteria: tiny (small) size, low cost, and low power/energy consumption [10].

(1) Tiny size: A WSN sensor should be handy to reach large-scale, convenient deployment. For example, if we want to deploy a large-scale sensor nodes in the mountainous area. It will be easier to use the small sensor rather than the big one. Not only how we carry such a huge sensor, but also it will be more convenient to have the small sensor that "hiding" in the environment. Making it is safe and clean in the environment.

(2) Low cost: A WSN should appropriate the budget needed in the deployment of hundreds or thousands sensor node. So, the relatively cheap sensor nodes are required.

(3) Low power consumption: in a large-scale WSN, it will give more benefit if the sensor unit deployed there has low power consumption. Because we need not replace batteries one by one in such of WSN.

A sensor unit is designed by four basic components: a sensing unit, a processing unit, a transceiver unit, and a power unit as shown in Figure 2.2.

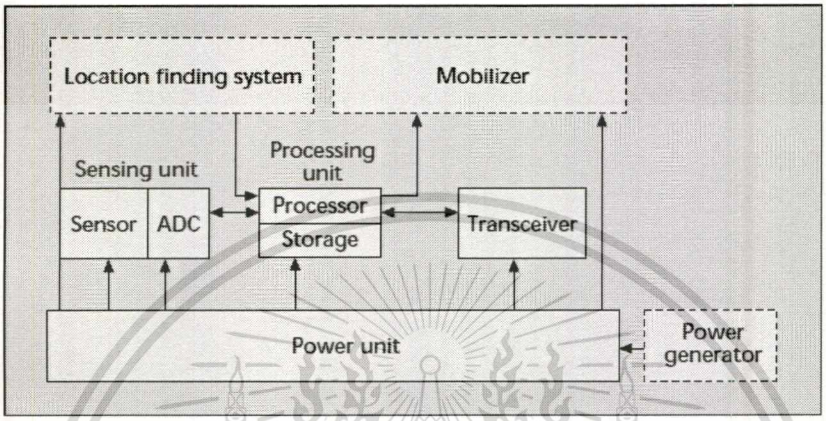


Fig. 2.2 The components of a sensor node [9]

2.1.3.1 The basic components

(1) Sensing unit, a sensor which measures the value of the type of environment such as temperature, humidity, light intensity, smoke, vibration, acceleration, the depth of the acid or alkaline, etc.

(2) Transceiver units, transferring wireless data (to get - to send) information between the sensor nodes.

(3) Processing unit, to process the information and stored it into memory. Requesting for information or send information through the transceiver unit. Processor may receive data from the location to assist in processing of various control functions. In addition, the processor can do processing network and route the transmission of the sensor unit.

(4) Power unit, energy storage and energy provided to all parts of the wireless sensor networks components.

2.1.3.2 The additional components

a. Location finding system (positioning unit), as the location of the sensor unit includes a GPS position data to be used as a way to process the data. Location for the motion of the sensor unit.

b. Mobilizer (mobilizing unit), available for various purposes such as tracking the network topology for the communication signal.

c. Power generator unit, as power generation from wind energy, heat of environment, etc. for storing or for further use.

2.1.4 Application of Wireless Sensor Networks

The network of wireless sensor is divided into two types. There are the type questions (querying) and the type of division of labor (tasking). A query (querying type) is when the sensor unit is asked to find out any information in to the wireless sensor network. For instance, in the agriculture application when the sensor unit will be asked to use a humidity and temperature. The sensor unit will respond with data stored in each sensor unit.

The division of labor (tasking) is used for monitoring events. The sensor unit includes a wireless sensor network have been assigned to monitor the heat and smoke. The sensor unit will be notified immediately when the heat and smoke levels increased or decreased based on the setting [11].

2.2 Overview of Radio Localization

Radio localization can be defined as a process to measure the radio signal parameters or information which is consequently used to estimate the location of mobile terminal. In this subsection, the radio localization techniques and its classification will be concisely given. Moreover, the examples of the existing algorithm are explained.

2.2.1 Radio Localization Technique

The two main categories of radio localization technique are mobile-based and network-based localization technology [6].

1) Mobile based Localization Technique

For example, in the global positioning system (GPS), the system has base station (BS) and mobile station (MS) or mobile terminal. The MS estimates its own position based on signals received from BSs. This form is usually defines as a self-location system. For this localization technique, the MS needs to be modified, so that it is able to obtain highly accurate location estimation. This matter causes to increase cost, size and battery (power) consumption. Furthermore, network synchronization is required.

2) Network based Localization Technique

In the network based localization technique, the BSs measure the MS signal information and relay them to a central site for further estimation of MS position. Unlikely the previous category, this form of localization technique is defined as a form of remote location system. One of the advantages of this system is the MS is not involved in the location-estimation process, so this technology does not need the modifications to the existing mobile handsets. The location estimation using the network-based localization technology requires two operations as the BSs: (1) the BSs measure signal parameters such as RSS, AOA, TOA, TDOA, etc. from MS, and (2) the measured signal parameters are combined in a data fusion stage to give the final location estimation.

2.2.2 Taxonomy of indoor positioning technique

Indoor positioning technique can be classified into three general types: (1) scene analysis, (2) triangulation, and (3) proximity as shown in Figure 2.3 [10].

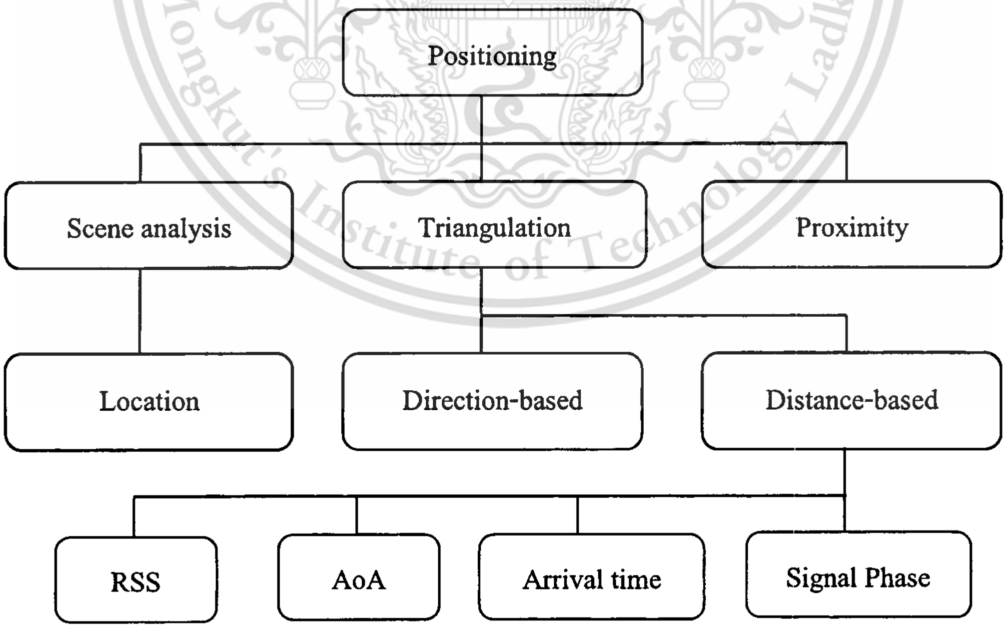


Fig. 2.3 Three types of indoor positioning techniques

1) Scene analysis

Scene analysis positioning techniques collect and extract features from observed scene. The scene could be the radio frequency waves, acoustic sound, visual images or measurable physical phenomena which usually exist near the object. The typical technique of radio frequency based scene analysis is location fingerprinting (LF). The observed features (fingerprints) are usually specific and unique. The fingerprints are used to estimate the location of observer or of observed objects in the scene. The distance can be estimated by matching the similarity of features.

2) Triangulation

Triangulation uses the geometric properties of triangles to estimate the target location. It can be divided into subcategories of lateration, using distance measurement, and angulations, using primarily angle or bearing measurement. Lateration calculates the target position by measuring its distance from multiple reference points. Since it uses the distance for estimating position, this technique is called range measurement technique. The distance can be measured by many parameters such as received signal strength (RSS), time of arrival (TOA), or time difference of arrival (TDOA), and the distance is derived by computing the attenuation of the emitted signal strength or by multiplying the radio signal velocity and the travel time. Roundtrip time of flight (RTOF) is also called received signal phase applied for range estimation in some system.

Angulation locates an object by calculates the angle relative to multiple reference points such as in angle of arrival (AOA).

3) Proximity

The proximity technique usually provides symbolic relative location information. This technique identifies an object with a tag and a dense grid antenna in a known location detects the tag. If the tag is detected by a single antenna, the location of an object refers to 'near' with that antenna. For more than one antenna detect the mobile object, it is considered to be estimated as those the one receives the strongest signal. The example of using proximity technique such as radio frequency identification (RFID), infrared radiation (IR), pressure and touch sensors, and capacitive field detectors.

2.3 Algorithms for Indoor Localization

Wireless sensor networks technology has been used to monitor and measure the environment change. WSNs technology is also appropriate for application of location finding in indoor and outdoor, respectively. For locating and tracking the object or target, there are system and algorithms that have been extensively researched to achieve the reliable and effective system.

2.3.1 Received Signal Strength

The most commonly used parameter to find the location of the target is received signal strength (RSS). The RSS based localization method uses a mathematical model which describes the path loss attenuation with distance. To find the target, the intensity of received signal from the reference points is used to measure the distance between the mobile station (MS) and those base stations (BSs). This is also usually called as beacon node localization technique [6]. Since the distance between the MS and each BS is known using the signal strength measurement, the location of MS can be estimated on a circle which the BS is at the center. In this method, at least three BSs have to be installed. Figure 2.4 shows the geometry of RSS based localization technique.

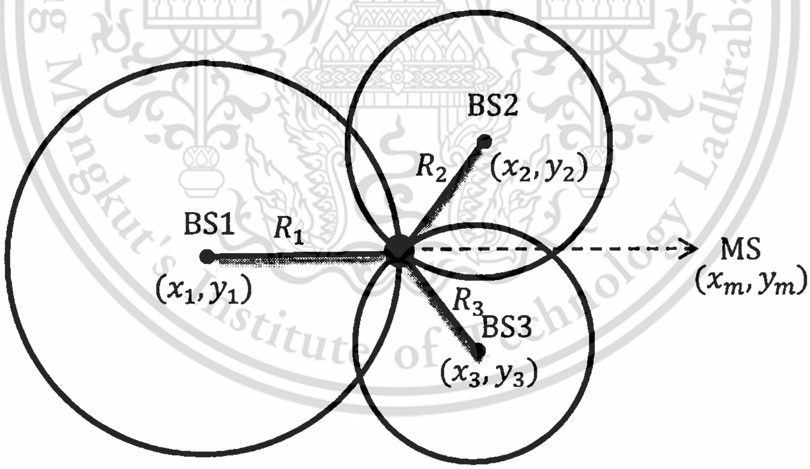


Fig. 2.4 Location estimation using the RSS measurement

To measure the distance, the signal strength is converted to the voltage. Due to the variance of the signal environments, this algorithm is not used for the application which needs high accuracy in the result. This is because RSS often gives the error its estimation. An error that occurred may be caused by following factors:

1. Non Line of Sight (NLOS)

   The communication is not in a straight line of the measured angle. So, the distance will increase.

2. Multiple access interference is a problem that occurs in any of the Code Division Multiple Access (CDMA) system.

3. Fluctuation in signal propagation

   Speed of sound wave propagation is disturbed by external factors such as wind or temperature, and the moisture in the air. These factors can cause high error rates at the greater distances.

To measure the received signal strength indicator (RSSI) in the hardware, especially for ZigBee as IEEE wireless communication standard 802.15.4, the equation of RSSI linearly with the distance is shown in (2.1). The radio signal strength is measured in decibel mill watts (dBm).

$$RSSI = -10n\,log_{10}(d\,) + A, \tag{2.1}$$

where, n is the path loss exponent or the signal propagation constant, d is the distance from transmitter in meter, and A is the received signal strength at 1 meter distance, in dBm. From (2.1), if the distance (d) increases, it will result in a decrease in the signal strength. This result can be used to determine the distance and position of the object [11].

2.3.2 Angle of Arrival (AOA)

AOA is a technique to find the location using the angle of the signal sent from the object (MS) to BSs. It can be defined as AOA based localization technique determines the location of the MS by first calculating the AOAs of a signal received from the MS at several BSs using the array antenna. By considering a two dimensional geometry, the AOAs at only two BSs is enough to estimate the location as shown in Figure 2.5.

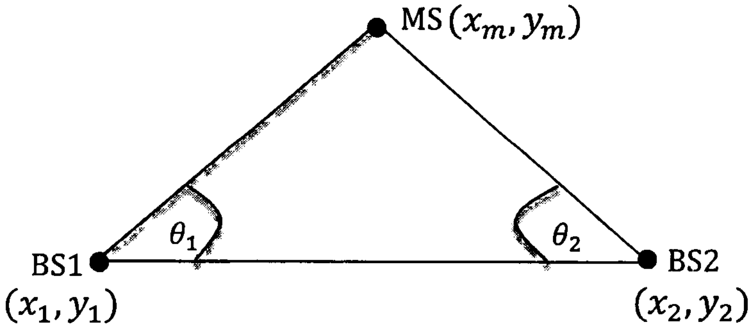


Fig. 2.5 Location estimation using the AOA estimation method

Given that, θ_1 and θ_2 are AOAs of the signal from mobile terminal arriving at BS1 and BS2, respectively, the position of MS can be calculated by following equation:

$$y_m - y_1 = \tan(\theta_1)\,(x_m - x_1) \tag{2.2}$$

$$y_m - y_2 = \tan(\theta_2)\,(x_m - x_2) \tag{2.3}$$

From (2.2) and (2.3), we can obtain:

$$\theta_1 = tan^{-1}\left(\frac{|y_m - y_1|}{|x_m - x_1|}\right) \tag{2.4}$$

$$\theta_2 = tan^{-1}\left(\frac{|y_m - y_2|}{|x_m - x_2|}\right) \tag{2.5}$$

By substituting (2.4) and (2.5), the position of MS (x_m, y_m) will be determined. In this technique, the similar issue in RSS based localization technique is also happened. It is the absence of a LOS signal component between the MS and BSs. It results the estimated angles do not correspond to the actual direction from MS to BSs. Moreover, the multipath still affects the AOA estimation eve the system is in LOS between MS and BSs. The error in AOA estimation results in the error in the location estimation of the MS. Because of it, the high accuracy of the AOA estimation is needed to achieve a high accuracy of location estimation [11], [12].

2.3.3 Time of Arrival (TOA)

The TOA based localization method is the method that position of MS can be estimated based on TOAs of the signal of MS arriving at least in three different BSs [12]. If t_i is the TOA of the signal of MS at BS_i , the distance between the MS and BS_i can be calculated by this following equation

$$R_i = (t_i - t_0)c, \tag{2.6}$$

where t_0 is the time instant at which the mobile terminal begins transmitting the signal, and c is the signal propagation speed (speed of light: $c = 3 \times 10^8\ m/s$).

object and BS position i cross to speed of light and R_j is TDOA between the object and BS position j cross to speed of light. The equation (2.7) will become

$$R_{i,j} = (t_i - t_0)c - (t_j - t_0)c, \\ = (t_i - t_j)c, \quad (2.8)$$

where $t_i - t_j$ is the TDOA between BS_i and BS_j. The hyperboloids that explain the range difference $R_{i,j}$ between BSs are given by

$$R_{i,j} = \sqrt{(x_i - x_m)^2 - (y_i - y_m)^2 - (z_i - z_m)^2} - \sqrt{(x_j - x_m)^2 - (y_j - y_m)^2 - (z_{ij} - z_m)^2}, \quad (2.9)$$

where (x_m, y_m, z_m) is the unknown coordinates of the object (mobile terminal), (x_i, y_i, z_i) and (x_j, y_j, z_j) are the locations of BS_i and BS_j, respectively. Similar case in the TOA based localization method, if the number of unknowns or coordinates of the object is equal to the number of equations, or measurements of range differences, the system is consistent and gives a unique solution. On the contrary, the system may be inconsistent and a unique solution may or may not exist. The advantage of TDOA based localization method is that TDOA does not need to know the instant time beginning to transmit the source signal as TOA method does [6], [10], [11], [12].

2.3.5 Location Fingerprinting

The location fingerprinting is one of the localization methods that based on the RSS measurement or other signal information to form a database of known signal-location information. In this fingerprint method, a two-phase process is involved. First, during the off-line calibration phase the received signals at selected locations are recorded in a database. This database is also referred to as the radio map. Then, the second phase, called the on-line, pattern matching algorithms are used to infer a target's location by comparing the current observed signal features to the pre-recorded values in the database. The key points of this technique are the selection of the spatial signature and the method in constructing database. The location fingerprint is also referred in the literature as pattern matching, pattern recognition, and database correlation.

As mentioned before, to estimate the location of MS, firstly, the system needs to measure the signal information at particular locations and uses some localization algorithms to determine the MS location. The location of MS can be estimated by comparing the current received signal information with the previously recorded as database through pattern matching algorithms. The location in the database that gives the best match or the minimum error between the fingerprint and the current observed signal information is represented as the estimated location of MS. Figure 2.7 shows the block diagram of the fingerprint based localization system.

Fingerprint Database

Known Location	**Fingerprint Information**
Location 1	Information 1
Location 2	Information 2
....	....
Location L-1	Information L-1
Location L	Information L

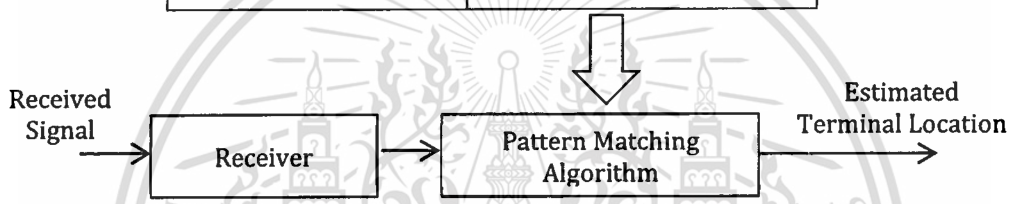


Fig. 2.7 Location estimation using the location of fingerprint technique [12]

The advantage of the fingerprint based localization technique is that it requires only 1 BS and does not need the additional or modified hardware or software within MS. Furthermore, it is more appropriate to estimate the location of MS in the multipath environment where the geological method such as AOA, TOA, TDOA, and RSS often fail for locating the location estimation. However, the disadvantages of this technique are time consuming and requiring a lot of effort during the process of constructing database [12].

2.4 Data Clustering

Clustering refers to an unsupervised learning task that purposes at decomposing a given set of objects into subgroups or clusters based on similarity. The aim is to divide the data-set in such way that objects (or example cases) belonging to the same cluster are as similar as possible, whereas objects belonging to different clusters are as dissimilar as possible. In other words, clustering is the process of grouping a data set in a method that the similarity between data inside a cluster is maximized, on the other hand, the similarity between data of different clusters is minimized.

Clustering method also can be used as data reducing method. It means that the similarity among data can be grouped and the others data that have dissimilarity can be ignored. So, the single object data that wanted to be analysis only focus on the certain cluster. Therefore, clustering method is endowed by measuring the distance to show the similarity and dissimilarity in the data [13]. In this subsection, the fuzzy-based clustering technique is presented.

2.4.1 Introduction of Fuzzy Set Theory

In the beginning of 1970s, Lotfi Zadeh, a professor at the University of California at Barkeley introduced the concept of fuzzy sets theory. Fuzzy sets theory was presented as a way of processing data by allowing partial set membership rather than crisp set membership or non-membership. Fuzzy sets provide a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information [14].

Fuzzy set theory is an extension of classical theory where elements have varying degree of membership. A logic based on the truth values, *True* and *False*, is sometimes not enough to describe human reasoning. Fuzzy logic applies the overall interval between 0 (false) and 1 (true) for describing human justification. A fuzzy set is any set which allows its members to have different degree of membership called membership function, in the interval [0, 1]. The degree of membership or truth is not same as probability:

(a) Fuzzy truth is not likelihood of some events or conditions,

(b) Fuzzy truth represents membership in vaguely defined sets.

Fuzzy logic is derived from fuzzy set theory dealing with justification that is approximate rather than precisely deduced from classical predicate logic. Fuzzy logic is capable of handling inherently imprecise concepts. Fuzzy logic allows in linguistic form the set membership values to imprecise concepts like “slightly”, “quite”, and “very”. Fuzzy set theory defines Fuzzy operators on Fuzzy sets.

Fuzzy logic requires some numerical parameters in order to operate such as what is considered significant error and significant rate-of-change-of-error, but exact values of these numbers are usually not critical unless very responsive performance is required in which case empirical tuning would determine them. Some of the advantages of fuzzy logic such as fuzzy logic is inherently robust since it does not require precise, noise-free inputs and can be programmed to fail safely i.e., in control system if a feedback sensor quits or is destroyed. Fuzzy logic can control nonlinear system that would be difficult or impossible to model mathematically.

2.4.2 Fuzzy C-Means (FCM)

Fuzzy C-Means (FCM) is one of clustering technique using fuzzy sets theory. In the fuzzy clustering, fuzzy cluster analysis allows gradual membership of data points to cluster measured as degree in [0, 1]. This gives the flexibility to express that data points can belong to more than one cluster. Moreover, these membership degrees offer a much finer degree of detail of the data model. Apart from assigning a data point to clusters in shares, membership degree can also express how ambiguously or definitely a data point should belong to a cluster [15].

Fuzzy clustering allows fine grained solution spaces in the form of fuzzy partitions of the set of given examples $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. Whereas the clusters Γ_i of data partitions have been classical subsets so far, they are represented by the fuzzy sets $\mu\Gamma_i$ of the data-set X in the following. Complying with fuzzy set theory, the cluster assignment, u_{ij}, is membership degree of datum $\mathbf{x}_j$ to cluster Γ_i, such that: $u_{ij} = \mu\Gamma_i(\mathbf{x}_j) \in [0, 1]$. Since memberships to clusters are fuzzy, there is not a single label that is indicating which cluster a data points belongs. Instead, fuzzy clustering methods associate a fuzzy label vector to each data point $\mathbf{x}_j$that states its membership to c clusters:

$$\mathbf{x}_j = (u_{ij}, \dots, u_{cj})^T \tag{2.9}$$

The $c \times n$ matrix $\mathbf{U} = (u_{ij}) = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ is then called a fuzzy partition matrix. Based on the fuzzy set notion, it is now better suited to handle ambiguity of cluster assignments when clusters are badly delineated or overlapping. So far, the general definition of fuzzy partition matrices leaves open how assignments of data to more than one cluster should be expressed in form of membership value. In FCM, the probabilistic partition is essential. Some of literatures only call fuzzy partition and ignore the 'probabilistic'. In this explanation, we use the subscript f for the probabilistic approaches and p for probabilistic model.

Let $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be the set of given examples and let c be the number of clusters $(1 < c < n)$ represents the fuzzy sets $\mu\Gamma_i$, $i = (1, \dots, c)$. Then we call $U_f = (u_{ij}) = (\mu\Gamma_i(\mathbf{x}_j))$ a probabilistic cluster partition of X if

$$\sum_{j=1}^{n} u_{ij} > 0, \quad \forall_i \in \{1, \dots, c\}, \tag{2.10}$$

$$\sum_{j=1}^{n} u_{ij} = 1, \quad \forall_i \in \{1, \dots, n\}, \tag{2.11}$$

The $u_{ij} \in [0, 1]$ is interpreted as the membership degree of datum $\mathbf{x}_j$ to cluster Γ_i relative to all other clusters. Equation (2.10) guarantees that no cluster is empty. This corresponds to the requirement in classical cluster analysis that no cluster, represented as (classical) subset of X is empty. Equation (2.11) makes sure that the sum of the membership degrees for each datum equals to 1. It means that each datum receives the same weight in comparison to all other data and, therefore, that al data are (equally) included into the cluster partition. Moreover, Equation (2.11) corresponds to a normalization of the memberships per datum. Thus the membership degrees for a given datum *formally resemble* the probabilistic of its being a member of the corresponding cluster.

Hence the problem to divide a given data-set into c clusters can be stated as minimizing the squared distance of the data point to their cluster centers, since we want to maximize the degree of membership. The probabilistic fuzzy objective function J_f is thus based on the least sum of squared distance. Formally, a fuzzy cluster model of a given data-set X into c cluster is defined to be optimal when it minimizes the objective function:

$$J_f(\mathrm{X}, \mathrm{U}_f, \mathrm{C}) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d_{ij}^{2}, \tag{2.12}$$

under the (2.10) and (2.11) that must be satisfied for probabilistic membership degrees in U_f. The condition (2.10) avoids the trivial solution of minimization problem i.e., $u_{ij} = 0,\ \forall i, j$. The normalization constraint (2.11) leads to a 'distribution' of the weight of each data point over the difference clusters. Since all data points have the same fixed amount of membership to share between clusters, the normalization condition implements the known partitioning property of any probabilistic fuzzy clustering algorithm. The parameter $m, m > 1$ is called *fuzzifier* or *weighting exponent*.

The exponentiation of the memberships with m in J_f can be seen as function g of membership degrees, $g(u_{ij}) = u_{ij}^{m}$, that leads to generalization of the well-known least squared error functional. The actual value of m determines the 'fuzziness' of the classification. It has been show in the case $m = 1$ (when J_h and J_f become identical), that clusters assignments remain hard when minimizing the target function, even though they are allowed to be fuzzy, i.e., they are not constraint in $\{0, 1\}$. For achieving the desired fuzzification of the resulting probabilistic exponents $m > 1$ that leads to fuzzy memberships has been proposed in [15]. With the higher values for m the boundaries between clusters become softer, with lower values they get harder. Usually $m = 2$ is chosen. Aside from the standard weighting of the membership with u_{ij}^{m}, other functions g that can serve as fuzzifiers have been explored.

The objective function J_f is alternately optimized, i.e., first the membership degrees are optimized for fixed cluster parameters, then the cluster prototypes are optimized for fixed membership degrees:

$$\mathrm{U}_\tau = j_{\mathrm{U}}(\mathrm{C}_{\tau-1}),\ \tau > 0 \tag{2.13}$$

and

$$\mathrm{C}_\tau = j_{\mathrm{C}}(\mathrm{U}_\tau). \tag{2.14}$$

In each of two steps the optimum can be computed directly using the parameter update equations j_{U} and j_{C} for the membership degrees and the cluster centers, respectively. The update formula is derived by simply setting the derivative of the objective function J_f. The resulting equations for the two iterative steps form the fuzzy C-means algorithm. The membership degrees have to be chosen according to the following update formula is independent of the chosen distance measure [15]:

$$u_{ij} = \frac{1}{\sum_{l=1}^{c}\left(\frac{d_{ij}^2}{d_{lj}^2}\right)^{\frac{1}{m-1}}} = \frac{d_{ij}^{\frac{-2}{m-1}}}{\sum_{l=1}^{c} d_{lj}^{\frac{-2}{m-1}}} \tag{2.15}$$

In this case, there exists a cluster i with zero distance to datum $\mathbf{x}_j$, $u_{ij} = 1$ and $u_{lj} = 0$ for all other clusters $l \neq i$. The (2.14) shows the relative character of the probabilistic membership degree. It depends not only on the distance of the datum $\mathbf{x}_j$ to cluster *i*, but also on the distance between this data point to other clusters. The update formula, j_{C} for the cluster parameters depend on the parameter used to describe a cluster (location, shape, and size) and on the chosen distance measure. Therefore the update formula cannot be given. In this case of the basic FCM model the cluster center vectors serve as prototypes, while an inner product norm induced metric is applied as distance measure. Consequently the derivation of J_f:

$$\mathbf{c}_i = \frac{\sum_{j=1}^{n} u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{2.16}$$

The choice of the optimal cluster center points for fixed memberships of the data to the clusters has the form of a generalized mean value computation for which fuzzy C-means algorithm has its name.

2.4.3 Implementation of Fuzzy C-Means in Fingerprint Technique

In this subsection, the parameter for FCM in this research will be described. The implementation of clustering technique will speed the processing of fingerprint-based localization technique. As pre-processing method before pattern matching between the current signal information (from object) and those in fingerprint database, FCM clustering algorithm will be useful to limit the scope of data in matching process. FCM has many benefits to be applied in the system which has multi channels data, in our case is RSSI data. As explained before, the number of fingerprint locations as M and $\mathbf{D}_m$ as fingerprint information (RSSI values for Nnodes) at position $\mathbf{F}_m$ (m $= 1, 2, \ldots, M$) and r represents RSSI values from each node. With $M = 12$ and number of reference node is $N = 6$ nodes for 2D and 3D scenario, respectively. The FCM will be deployed to minimize the partition of RSSI data set.

FCM is the method for data clustering which allows one piece of data belongs to two or more clusters (groups) [22]. The FCM algorithm minimizes the objective function for the partition of the data set, $\mathbf{P} = [\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3, \ldots, \mathbf{D}_M]^T$ and center of clusters, $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_C]^T$ given by:

$$J_w(\mathbf{U}, \mathbf{V}) = \sum_{c=1}^{C} \sum_{m=1}^{M} u_{cm}^{w} \|\mathbf{D}_m - \mathbf{v}_c\|^2 \tag{2.17}$$

where, J_m is the objective function, M is the rows in the matrix in the vector $\mathbf{P}$, C is the number of clusters $(1 \leq C \leq M)$, u_{cm} is the element of partition matrix $\mathbf{U}$ of size $(C \times M)$ containing the membership function, $\mathbf{v}_c$ is the center of the C cluster, and w is a weighting factor that controls fuzziness of the membership function. The matrix $\mathbf{U}$ is constrained to contain elements in the range of $[0,1]$ such that $\sum_{m=1}^{M} u_{cm} = 1$ for each $u_{cm}(1 \leq C \leq M)$. The norm $\|\mathbf{D}_m - \mathbf{v}_c\|$ is the Euclidean distance between the sample $\mathbf{D}_m$ and the clusters center $\mathbf{v}_c$ [9].

There are two steps for obtaining the fuzzy cluster, first step is called initialization step. In this step, the fuzzy system will randomly pick the cluster centers based on number of fuzziness (w), basically use $w = 2$ and number of cluster, C. Second step is called fuzzy computation step. Here, the process to get the fuzzy cluster or group is utilizing.

FCM algorithm is used to cluster the fingerprint database and the target in 2D and 3D scenario, respectively. In this work, the data set, $\mathbf{P}$ is formed as $A \times B$ matrix. A represents the number of fingerprint database and B represents the number of reference nodes. For instance, in 2D scenario the matrix size of fingerprint database is 36×4 for 4 reference nodes and 36×6 for 6 reference nodes. In 3D scenario, the 12 fingerprint database is applied with 6 reference nodes, 12×6 matrix. In order to find the location of the target, the clustering technique has to have the target node's information clustered among the fingerprint database. In this manner, the size of matrix will be added by 1 row of target node. As known, the target signal information has the 1×6 matrix. So, in the data set, **P** the size of matrix will become $(A+1) \times B$ [7], [20], [21].

The membership function is the most essential part in any fuzzy clustering methods. Figure 2.8 shows the typical graph of membership function in FCM [23].

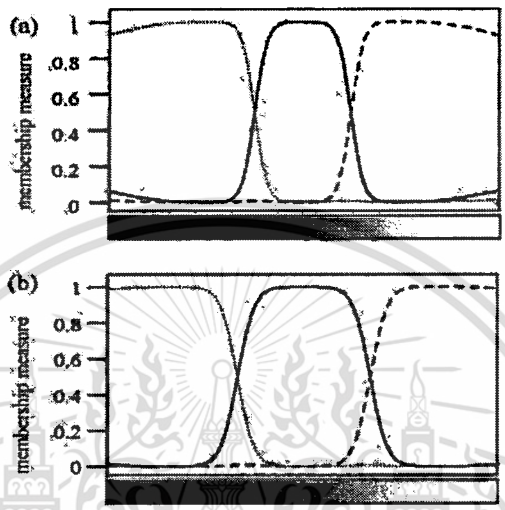


Fig. 2.8 (a) Fuzzy C-Means membership function (b) Fuzzy Generalized membership function

From Figure 2.8 dot line represents hard truth and dashed line represents partial truth-the truth values between completely true and completely false. Given an example, When **P** is a fuzzy set and **u** is a relevant object, the proposition "**u** is a member of **P**" is not necessarily either true or false, as required by the two-valued logic, but it may be true only to some degree, the degree to which **u** is actually a member of **P**, is a real number in the interval [0, 1]. Theoretically, if **U** is a collection of objects denoted generically by **u**, then a fuzzy set **P** in **U** is a set of ordered pairs, $\mathbf{P} = \{(\boldsymbol{u}, \mu\mathbf{P}(\boldsymbol{u})) \mid \boldsymbol{u}\ \varepsilon\ \mathbf{U}\}$, $\mu\mathbf{P}(\boldsymbol{u})$is called the membership function (or grade of membership) of **u** in **P** that maps **U** to the membership space **M**. The range of the membership function is a subset of the nonnegative real numbers whose supremum is finite [24].

Figure 2.9 shows the flowchart of FCM deployed in this work. This flowchart describes of two steps in FCM clustering algorithm, initialization and verification step.

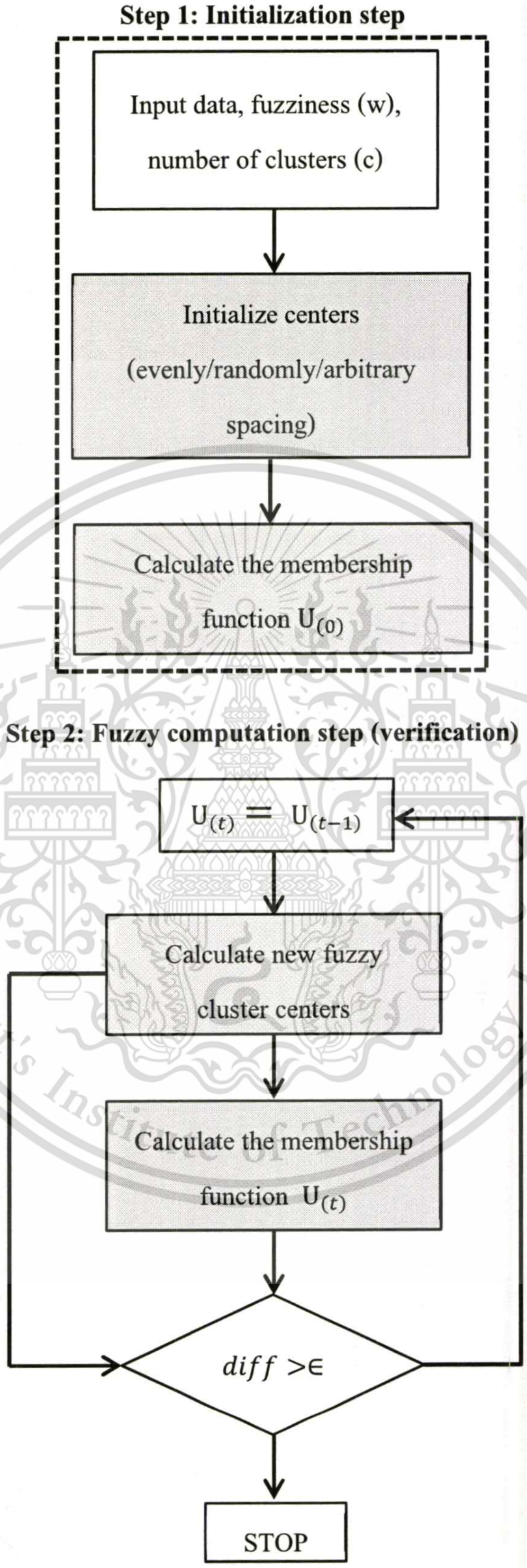


Fig. 2.9 Flowchart of FCM clustering algorithm

As illustrated in Figure 2.9, FCM clustering algorithm can be summarized by the following steps [5]:

Step 1: Initialization (Iteration 0)

Scan the RSSIs data row by row to construct the vector **P** containing all the certain levels in the data set. Randomly initialize the centers of the classes vector $\mathbf{V}^{(0)}$. From the iteration $t = 1$ to the end of algorithm:

Step 2: Calculate the membership matrix $\mathbf{U}^{(\mathrm{t})}$of element u_{ik} using:

$$u_{cm} = \left(\sum_{c=1}^{C} \left(\frac{\|D_m - v_c\|}{\|D_m - v_C\|} \right)^{\frac{2}{w-1}} \right)^{-1} \tag{2.18}$$

$\mathbf{v}_c$ $(c = 1,2, \ldots ., C)$ is the cluster center in the iteration, $t = 0$ (the superscript, t is the iteration index). Then, the $\mathbf{v}_1$, $\mathbf{v}_2$ until $\mathbf{v}_C$ can be obtained for the next step.

Step 3: Calculate the vector $\mathbf{V}^{(\mathrm{t})} = [\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_C]$ using:

$$\boldsymbol{v}_c = \frac{\sum_{m=1}^{M} u_{cm}^{w} D_m}{\sum_{m=1}^{M} u_{cm}^{w}} \tag{2.19}$$

Step 4: Convergence test: if $\left\|\mathbf{V}^{(\mathrm{t})} - \mathbf{V}^{(\mathrm{t}-1)}\right\| > \epsilon$, then increment the iteration t, and return to the **Step 2**, otherwise, stop the algorithm. Note: ϵ is a chosen positive threshold.

Although the FCM can find a partition of data for a fixed number of clusters (objects), one objective of cluster validity procedure is to determine automatically the optimal number of clusters. This is desired, for example, when the FCM is used for image segmentation purposes and the number of objects in the image is unknown. Validation of a generated fuzzy partition by the FCM can be achieved by a cluster validity index. If c_{min} is defined as the minimum number and c_{max} as the maximum number of clusters, respectively, then for each $c \in [c_{min}, c_{max}]$ a partition can be generated by the FCM. The value of a cluster validity index can then be calculated for the partitions of data for each c. By comparing all values of an index for all possible number of clusters, one can determine the optimal number of clusters. This can be achieved, for example, by selecting the number of clusters for which an index is minimized [25].

2.5 ZigBee

Zigbee is a wireless networking standard developed by ZigBee Alliance. It is aimed to be applied in remote control and sensor applications. It is suitable for operation in harsh radio environments and in isolated locations. Zigbee as IEEE standard 802.15.4, it defines the application and security layer specifications enabling interoperability between products from different manufacturers. In this way, ZigBee is a superset of the 802.15.4 specification. Low power consumption is the essential feature of ZigBee that enables the power supply or battery can be operated for a long period of time. The bandwidth of ZigBee is relatively low compared to wireless LAN. In some types of ZigBee, the distance between 2 nodes can be up to 50 meters [16].

2.5.1 Hardware Physical and MAC Layer

One of the most benefit aspects of ZigBee is its frequency band. It is a license free band at 2.4 GHz. It allows a ZigBee product to be used all over the world. Table 2.1 shows the few differences between the bands in ZigBee [11], [17].

Table 2.1 The few differences between the bands in ZigBee

Frequency	868 MHz	915 MHz	2.4 GHz
Bandwidth	20 kbps	40 kbps	250 kbps
No. of channels	1	10	16

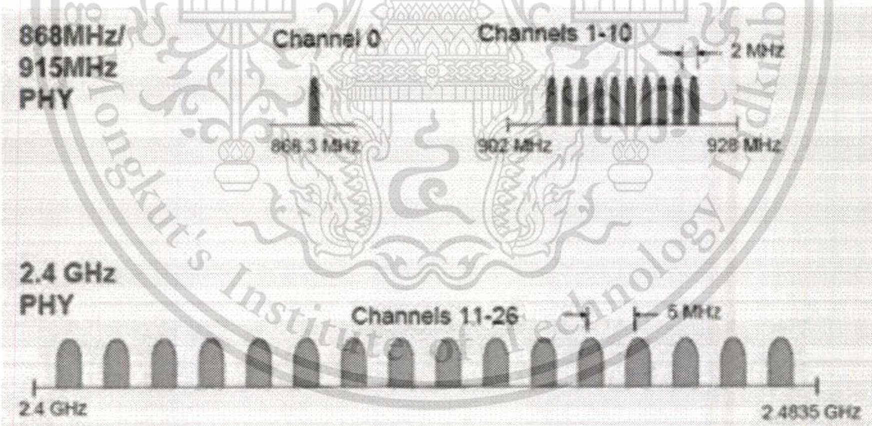


Fig. 2.10 Channel division [17]

In all bands, Direct Sequence Spread Spectrum (DSSS) is used. 868 and 915 MHz are using Binary Phase Shift Keying (BPSK) and 2.4 GHz uses Offset Quadrature Phase Shift Keying (O-QPSK). In total, there is room for a maximum of 104 bytes. This is obtained by used of a guaranteed time slot mechanism so that the high priority messages can be sent as fast as possible. ZigBee applies 2 kinds of addressing. There is a 64-bit IEEE address that can be compared to the IP address on the internet. There is also a 16-bit short address. The short addresses are used once a network in setup period, so this makes a total of 2^{16} = ~64000 nodes within one network possible.

2.5.2 ZigBee Architecture

ZigBee is a wireless multi-hop routing, which can send data to each other sensors. It is one of the essential features that Bluetooth does not have. ZigBee communication helps extend the transmission range of information is out indefinitely. ZigBee is known because of its mesh network architecture but it does also support a star topology or cluster tree or hybrid architecture. Physical Layer and MAC Layer of ZigBee and IEEE 802.15.4 standard defines a wireless communication WPAN (Wireless Personal Area Network) to work in such a layer at a lower level of the signal, link quality, access control, and security. Many aspect of the network are specified in this layer such as application profiles, security settings and the messaging. ZigBee Application Layers of the architecture shown in Figure 2.11[17].

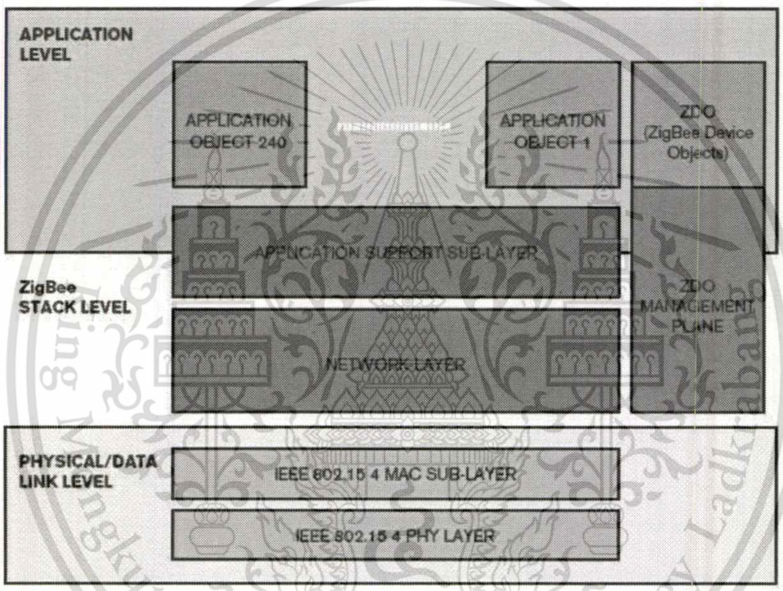


Fig. 2.11 ZigBee and IEEE 802.15.4 protocol stack [17]

The characteristic of ZigBee is the entrance channel using the Carrier Sense Multiple Access with Collision Avoidance (CSMA - CA) or the entrance channel. A way to avoid the collision-type devices, ZigBee has two types of model devices; Physical Device and Logical Device. Physical Device can be classified into two types.

a. Full Function Device (FFD)

FFDs are equipped with a full set of MAC layer functions, which enables them to act as a network coordinator or a network end device. FFDs acting as network coordinator will have the ability to send data to beacon, offer synchronization, communication and network join services.

b. Reduced Function Device (RFD)

RFD is the best way to connect to a network of battery power to transmit data from other devices. RFDs can only act as end devices and are equipped with sensors/actuators like transducers light switches, lamps, etc. It may only interact with a single FFD.

ZigBee is based on the behavior of the three types:

1. Coordinator is responsible for the communication. It is an FFD managing the whole network.
2. End Device is equipped with the destination. It corresponds to an IEEE RFD or FFD acting as a simple device.
3. Router is responsible for routing data in a variety of networks, which is comparable to FFD. It is an FFD with routing capabilities.

2.5.3 Network Topologies

Network topology is defined as a connection of a network in WSN. In the ZigBee standard, it supports three main topologies i.e., star topology, peer-to-peer topology, and tree topology. A star topology is simple. All nodes directly communicate with one central node (like a star). The mesh topology is more complicated; each node may communicate with any other node within range. It gives many possible routes through the network. The cluster tree topology is basically a combination of star and mesh topologies [17].

2.5.3.1 Star Topology

In the star topology, the communication is established between devices and a single central controller, called the PAN coordinator. Applications that benefit from this topology include home automation, PC peripherals, toys and games. After an FFD is activated for the first time, it may establish its own network and become the PAN coordinator. Each start network chooses a PAN identifier, which is not currently used by any other network within the radio sphere of influence. This allows each star network to operate independently [11], [16], [17].

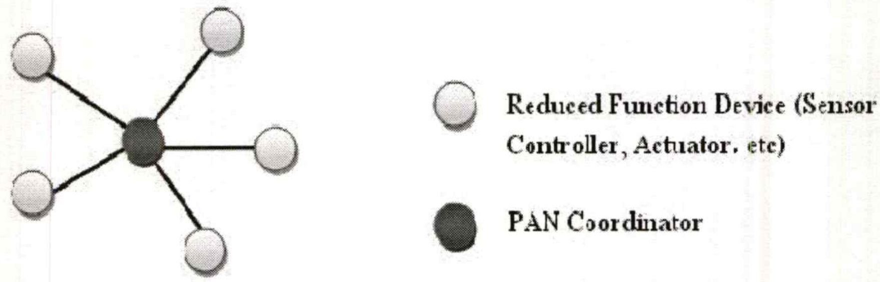


Fig. 2.12 Star topology [26]

2.5.3.2 Peer-to-peer topology

Figure 2.13 shows the illustration of peer-to-peer topology in ZigBee standard. In peer-to-peer topology, there is also one PAN coordinator. In contrast to star topology, any device can communicate with any other device as long as they are in range of one another. A peer-to-peer network can be ad hoc, self-organizing and self-healing. Applications using this topology are mostly in monitoring such as industrial control and monitoring, wireless sensor networks, asset and inventory tracking. It also allows multiple hops to route messages from any device to any other device in the network. It can provide reliability by multipath routing [16].

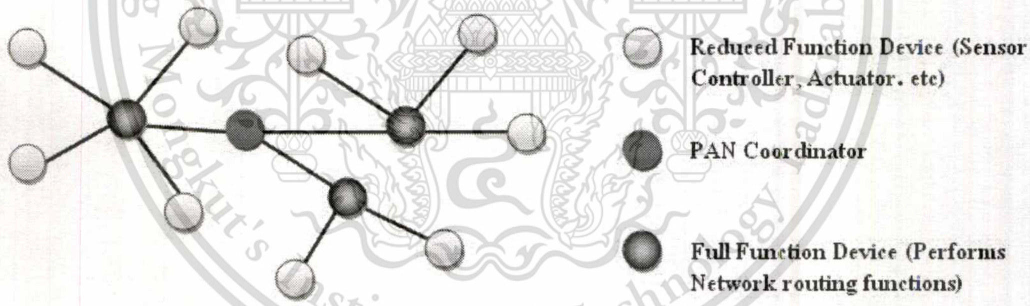


Fig. 2.13 Peer-to-peer topology [26]

2.5.3.3 Cluster tree topology

The special case of peer-to-peer network is cluster-tree network. In cluster-tree topology, most devices are FFDs and an RFD may connect to a cluster-tree network as a leave node at the end of a branch. Any of the FFD can act as a coordinator and provide synchronization services to other devices and coordinators. Only one of these coordinators becomes the PAN coordinator.

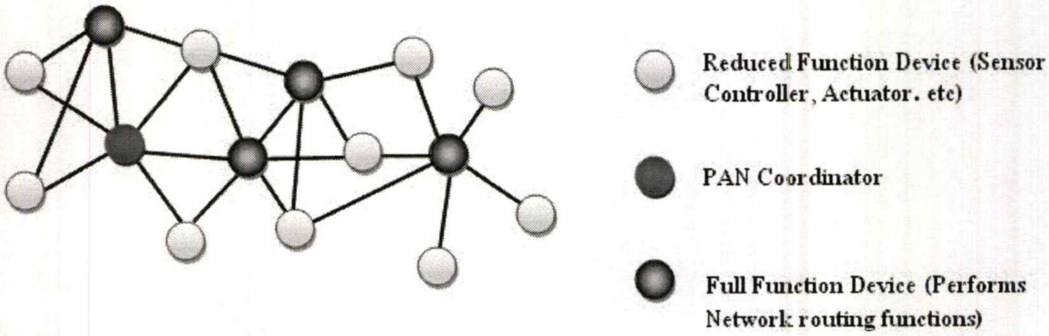


Fig. 2.14 Cluster tree topology [26]

The PAN coordinator forms the first cluster by establishing itself as the cluster head (CLH) with a cluster identifier (CID) of zero, choosing an unused PAN identifier, and broadcasting beacon frames to neighboring devices. A candidate device receiving a beacon frame may request to link the network at the CLH. If the PAN coordinator permits the device to link, it will add this new device as a child device in its neighbor list. The newly joined device will add the CLH as its parent in its neighbor list and begin transmitting periodic beacons such that other candidate devices may then join the network at that device. Once application or network requirements are met, the PAN coordinator may command a device to become the CLH of a new cluster adjacent to the first one. The advantage of this clustered structure is the increased coverage area at the cost of increased message latency [16] [17].

2.6 Xbee

Xbee is used wireless data transmission. The network was created by a standard based on IEEE 802.15. Xbee is considered as one of the ZigBee families [17], XBee has a feature that can be easily connected to the RS232 which allows any microcontroller or microprocessor to immediately use the services of the ZigBee protocol. All a ZigBee hardware designers have to be concern of the serial port logic levels that are compatible with the Xbee standard, 2.8- to 3.4-V logic levels. Xbee can be divided as functionality of the various forms of communication such as the Coordinator, Router or End Device. Xbee is equipped with microcontroller and RF IC inside to act as a transceiver and a Half Duplex Frequency 2.4 GHz is managed by utilizing a simple interface to send and receive data. It worthy to mention some advantages of the Xbee chip that are motivated us to select working on this chip.

Here are some of the main networking features of the device:

- 128-bit Encryption
- Point-to-point, point-to-multipoint and peer-to-peer topologies supported
- 13 direct sequence channels, each with over 65,000 unique network addresses available

- Self-routing/Self-healing mesh networking
- 2.4 GHz frequency band
- Manufactured under ISO 9001:2000 registered standards
- Supply voltage: 3.4-2.8 V
- Power down current is less than 10 μA

2.6.1 Simplified Block Diagram

Data is presented to the Xbee module through its DIN pin, and it must be in the asynchronous serial format, which consists of a start bit, 8 data bits, and a stop bit. Because the input data goes directly into the input of a UART (RS232) within the Xbee module, no bit inversions are necessary within the asynchronous serial data stream. All of the required timing and parity checking is automatically taken care of by the Xbee. Xbee module produces a received data asynchronous serial data stream for the host on its DOUT pin for more details, please see [17].

Fig. 2.15 Xbee device [27]

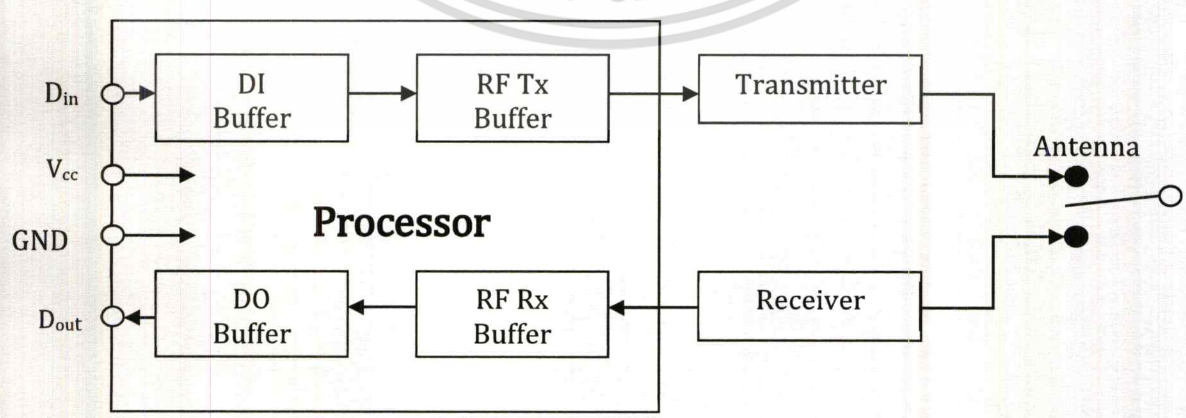


Fig. 2.16 Simplified block diagram of the Xbee modules

The processor used in Xbee module is the well-known MC9S08GT60 which contains main blocks shown in Figure 2.16. When serial data enters the RF module through the DI pin (pin 3), the data is stored in the DI Buffer until it can be processed. If the module is receiving a continuous stream of RF data, any serial data that arrives on the DI pin is placed in the DI Buffer. The data in the DI buffer will be transmitted over-the-air when the module is no longer receiving RF data in the network.

On the other hand, when RF data is received, the data enters the DO buffer and is sent out the serial port to a host device. Once the DO Buffer reaches capacity, any additional incoming RF data is lost. Two cases wherein the DO Buffer may become full and possibly overflow, the two cases are:

1. If the RF data rate is set higher than the interface data rate of the module, the module will receive data from the transmitting module faster than it can send the data to the host.
2. If the host does not allow the module to transmit data out from the DO buffer because of being held off by hardware or software flow control.

2.6.2 Xbee Topology

In a wireless network, ZigBee must contain at least two of the coordinator nodes and one client (Router / End device) to communicate and work in the form of PAN (Personal area network). The three types of Xbee networks are star, cluster tree, and mesh network [11], [17].

2.6.2.1 Star (Broadcast) Network

In the star topology, the communication is established between devices and a single central controller, called the PAN coordinator. It allows the user to get information any time.

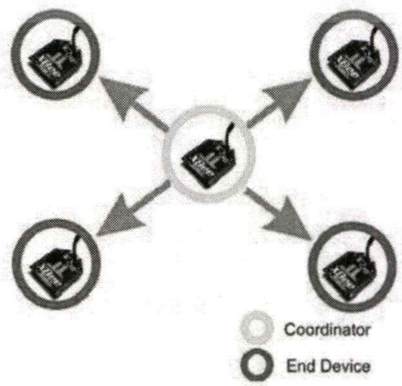


Fig. 2.17 Star (broadcast) network [11]

2.6.2.2 Cluster Tree (Tree) Network

In the cluster tree (tree) network, aside using the coordinator as the broadcaster, in this network we can use the other node as buffer or media to contact each other. As shown in Figure 2.18 of cluster tree (tree) network.

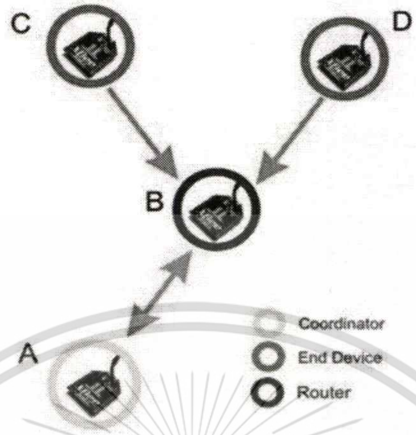


Fig. 2.18 Cluster tree (tree) network [11]

In Figure 2.18, the transfer of data transmission, such as A to communicate with C, but C is far from A and A cannot communicate with C, but it is not with B (between A and C). So, in this network, B will be used as media to contact (repeater) between A and C.

2.6.2.3 Mesh Network

The mesh network has been popular and applied in many applications because of its advantages. In this network, data can be sent to multiple targets. It can transmit data to many destinations that allows us to apply it to many design application.

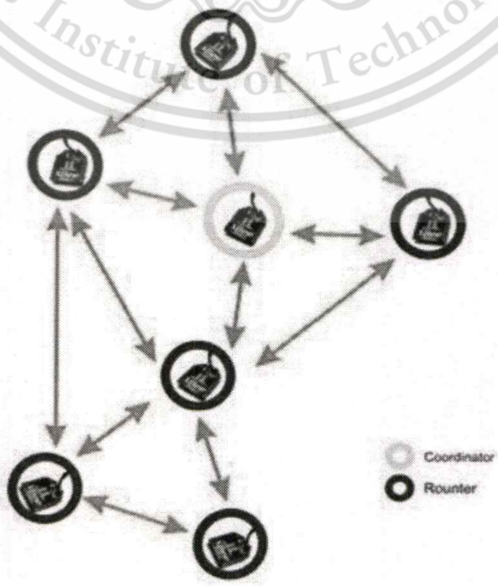


Fig. 2.19 Mesh network [11]

2.7 Indoor Localization using ZigBee based on RSSI [18]

Basically, a wireless localization system comprises at least two separate hardware components; a measurable unit that usually carries the major part of the system intelligence and a signal transmitter/receiver. The transmitter in the simplest case is just a node. Figure 2.20 represents basic hardware model of localization system. Figure 2.20 shows the location sensing devices detect the signals transmitted by or received at known reference pints using radio sensing technologies. The sensing technique is based on the signal strength level that converts the sensed signal to RSSI. Given a set of reference points, the relative position of the MS can be derived from the characteristic signal of RSSI. The signal characteristics of RSSI are unique for a location. Then, the microcontroller processes the position algorithm and estimates the target position and displays it in a PC.

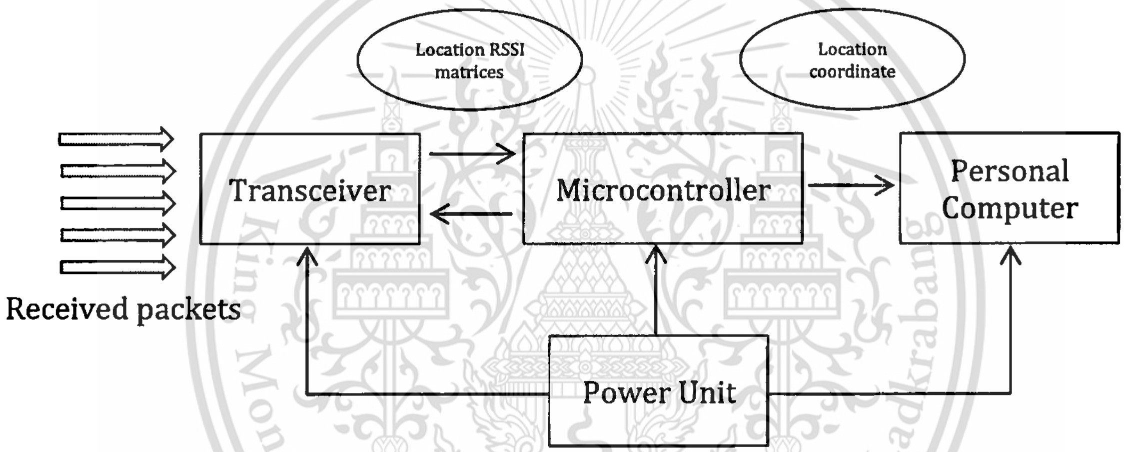


Fig. 2.20 Block diagram representing hardware module

In this subsection, relevant RF-based indoor positioning system which can be used to estimate MS location is reviewed. The impressive of development of IEEE 802.15.4 in recent years leads an interesting future for the location fingerprinting technique. The following discussion is divided according to the localization system component. First, the common form of location fingerprinting used RSSI and its relationship with the physical locations are explained. Finally, the location estimation algorithm: the Euclidean distance method is briefly explained for the RSSI based localization method.

2.7.1 Location Fingerprint

A location fingerprint based on RF characteristics such as RSSI is the basic for representing a unique location. It is designed as each location represents the position in a building (indoor) that has unique RF signature. Generally, a fingerprint F is labeled with location information L. The location fingerprint and its information are maintained as the fingerprint database and used during the on-line phase to estimate location as shown previously in Figure 2.7.

The location and fingerprint are usually denoted as a tuple of (L, F). In reference [18], the authors point out that the location information L for indoor location can be recorded in two forms as either a tuple of coordinates or an indicator variable. The tuple of real coordinates can be varied from one dimension to five dimensions which include the three dimension space and two orientation variable expressed in the spherical coordinates. For example, the location information of a 2D system with an orientation, can be expressed as triplet, $L = \{(x, y, d) \mid x, y \in R^2, d \in \{North, East, South, West\}\}$.

The RSSI is found to be more location-dependent than the signal to noise ratio (SNR) since the noise component is rather random in nature. However, the RSSI fluctuates over time for each location and access point. Each RSSI element can be considered as a random variable; therefore it can be stored by recording its descriptive statistics parameters, maintaining the whole measurement dataset or approximating each distribution. The location fingerprint is usually referred as an array or vector of signal strength received at any position in the location-based service area. The size of vector is based on the number of access point or reference points.

To create fingerprint-based received signal strength such as in a number of samples of vectors of signal strength are collected over a window of time for each position. This is called a prototype. After that, the averages of RSSI of each access or reference point is captured and recorded as an element in the location fingerprint. For an area that can receive signal from N reference points, the location fingerprint can be expressed as a vector of average RSSI elements, β_i

$$F = (\beta_1, \beta_2, \ldots., \beta_N)^T \tag{2.19}$$

Besides the basis location fingerprint, the samples of location fingerprint measured during the on-line phase are also essential for the mobile target system. The number of samples should be selected appropriately to represent the location fingerprint for mobile application. A step in statistical analysis method called pre-processing refers to a step that cleans the raw data (training set/database) before any further operations or analysis. The cleaning may comprise of encoding, dimensionality reduction, feature extraction/selection, clustering, and outlier elimination. In reference [18] stated that the pre-processing enables faster location estimation and reduce the noise from the training data. In this work, the clustering technique is applied as pre-processing step in the fingerprint based indoor localization technique.

2.7.2 Location Estimation Algorithm: Euclidean distance [18]

Localization algorithms are procedures which exploit dependency between location information and location fingerprint basis in order to determine a location from samples RSSI signals. The Euclidean distance method as the positioning algorithm used in this thesis is briefly discussed as follow; **Euclidean distance method** is deterministic approach of location estimation since they require only a set of constant location fingerprints which includes mean vectors. In order to determine the location, a form of discriminate function is used to classify a sample of RSSI fingerprint into a position. The average RSSI vector is a center of a mass that represents each class of location fingerprint. Suppose a set of l location fingerprints is denoted by $\{F_2, F_2, ..., F_l\}$ exist and each fingerprint has a one-to-one mapping to a set of locations $\{L_2, L_2, ..., L_l\}$. A sample of an RSSI fingerprint measure during on-line phase is denoted as S which can be another average RSSI vector for small window of RSSI samples. Assuming that an indoor localization system only considers the average RSSI from N reference points as a location fingerprint, the sample of RSSI vector is $S = (s_1, s_2, ..., s_N)^2$ and each location fingerprint i in the database can be expressed as $F_i = (\beta_1, \beta_2, ..., \beta_N)^T$. The simplest closeness metric is a distance measurement in signal space denoted as $Dist(.)$ function.

Chapter 3

Research Methodology

In this chapter, the research methodology is presented. Fuzzy C-Means (FCM) clustering algorithm will be implemented as pre-processing process in fingerprint based localization system. Pre-processing process even known as process which minimizing the data object in the system. The previous researches of fingerprint based localization technique in indoor localization have been presented in [19], [20]. The implementation of FCM in the fingerprint technique can be referred as simplifying method in pattern matching algorithm. In this research, the performance of localization technique in indoor environment is analyzed by measuring the accuracy of FCM to cluster the object/target and the accuracy to estimate the target position.

3.1 Indoor Localization Using Fingerprint Based Localization Technique

In this thesis, the indoor localization using fingerprint technique to find the location of an object (wireless device) is utilized. The implementation of clustering technique using FCM as the pre-processing before pattern matching algorithm of fingerprint technique is proposed. In this subsection, the explanation of the experiment setup and fingerprint database will be described.

For indoor localization system, both 2D and 3D indoor localization system, the Xbee devices are scattered in area of interest as illustrated in Figure 3.1.

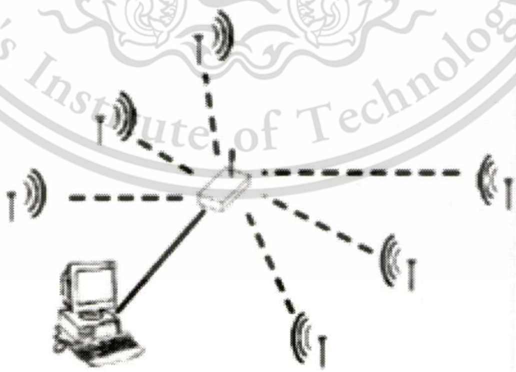


Fig. 3.1 Illustration of Xbee system

From Figure 3.1, the Xbee system is illustrated by the antenna symbols, a computer and a transmitter device. The antenna symbols represents the six location of Xbee, these locations refer to the reference nodes. The transmitter is directly connected to a PC to observe and record the Xbee parameter, in this research is RSSI.

3.1.1 Fingerprint Database

In order to record database, the RSSI parameter from Xbee has to be acquired. To get the RSSI values of the Xbee device is arranged as illustrated in Figure 3.1. In this scenario, the target node sends the command to the reference nodes to resend the intensity values that will be converted to RSSI in dBm. In this manner, the reference node will know the amount of transmitted power from the target node. In this process, the reference node will send back the information to the target node. As illustrated in Figure 3.1, the target node is directly connected to PC through serial communication RS-232 to record the RSSI values from the given RSSI packets.

3.1.1.1 ZigBee System

The Transceiver radio signal is a module Xbee-24ZB of the ZigBee alliance. It is featured on the DIGI and registered in the ZigBee standard to create routing. It has networks topologies and easy to use and to configure. The microcontroller is connected via USART module that can be configured by the manufacturer of the X-CTU software. Overview of the Xbee system is shown in Figure 3.2 [11], [17]

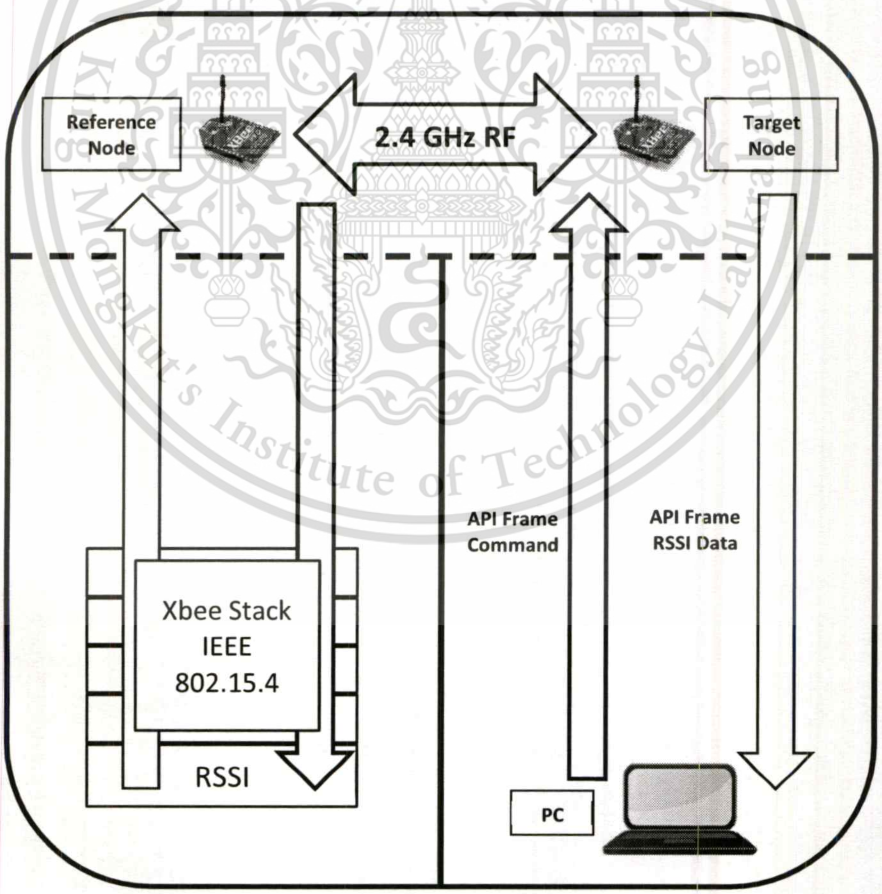


Fig. 3.2 Overview of the system

The database will be stored as fingerprint information using RSSI package from reference nodes received by transceiver to PC. To send the packet and get the RSSI value from Xbee, the API frame is applied. The detail of received package will be explained in the next subsection.

3.1.1.2 Frame Format Used to Communicate with The Network Layer

Frame format used to communicate with the network layer below:

- Header size is 1 Byte for Byte Stuffing.
- Frame length scale for the size of the frame is 2 Byte.
- Frame type 1 Byte size is used for the type of frame.
- Frame ID 1 Byte size of the frame number.
- DATA, depending on the type of the frame.
- Checksum size is 1 Byte Checksum of the Byte to 4.

Header 1 Bytes	Frame Length 2 Bytes	Frame ID 1 Bytes	Data Depend on Frame type	Checksum 1 Bytes

Fig. 3.3 Detail of the package size of 8 Byte

3.1.1.3 Types of Frame

Type of frame that used in this work is comprised of:

1. Remote AT Command Request (0x17).

   By using API mode, it is possible to send AT commands from a transmitting radio to a receiving radio. This allows module parameter registered on remote devices to be queried or set. One benefit application of this feature is to toggle an IO on a remote radio from a high to a low state. In this manner the radios can be used as a wireless relay to control a wide variety of remote devices. Remote AT command request table of features is shown in Figure 3.4.
2. ZigBee Transmit Request (0x10)

   ZigBee Transmit Request (0x10) table of features is shown in Figure 3.5.
3. Explicit Addressing ZigBee Command Frame (0x11)

   Sending the API frame 0x11 (Explicit Addressing Command Frame) allows the device to override these defaults. Explicit Addressing ZigBee Command Frame (0x11) table of features is shown in Figure 3.6.

	Frame Fields		Offset	Example	Description
API Packet	Start Delimiter		0	0x7E	
	Length		MSB 1	0x00	Number of bytes between the length and the checksum
			LSB 2	0x10	
	Frame-specific Data	Frame Type	3	0x17	
		Frame ID	4	0x01	Identifies the UART data frame for the host to correlate with a subsequent ACK (acknowledgement). If set to 0, no response is sent.
		64-bit Destination Address	MSB 5	0x00	Set to the 64-bit address of the destination device. The following addresses are also supported: 0x0000000000000000 - Reserved 64-bit address for the coordinator 0x000000000000FFFF - Broadcast address
			6	0x13	
			7	0xA2	
			8	0x00	
			9	0x40	
			10	0x40	
			11	0x11	
			LSB 12	0x22	
		16-bit Destination Network Address	MSB 13	0xFF	Set to the 16-bit address of the destination device, if known. Set to 0xFFFE if the address is unknown, or if sending a broadcast.
			LSB 14	0xFE	
		Remote Command Options	15	0x02 (apply changes)	Bitfield to enable various remote command options. Supported values include: 0x01 - Disable ACK 0x02 - Apply changes on remote. (If not set, AC command must be sent before changes will take effect.) 0x40 - Use the extended transmission timeout for this destination. Setting the extended timeout bit causes the stack to set the extended transmission timeout for the destination address (see chapter 4). All unused and unsupported bits must be set to 0.
		AT Command	16	0x42 (B)	Name of the command
			17	0x48 (H)	
		Command Parameter	18	0x01	If present, indicates the requested parameter value to set the given register. If no characters present, the register is queried.
	Checksum		19	0xF5	0xFF - the 8 bit sum of bytes from offset 3 to this byte.

Fig. 3.4 Package Remote AT Command Request (0x17)[17]

	Frame Fields		Offset	Example	Description
API Packet	Start Delimiter		0	0x7E	
	Length		MSB 1	0x00	Number of bytes between the length and the checksum
			LSB 2	0x16	
	Frame-specific Data	Frame Type	3	0x10	
		Frame ID	4	0x01	Identifies the UART data frame for the host to correlate with a subsequent ACK (acknowledgement). If set to 0, no response is sent.
		64-bit Destination Address	MSB 5	0x00	Set to the 64-bit address of the destination device. The following addresses are also supported: 0x0000000000000000 - Reserved 64-bit address for the coordinator 0x000000000000FFFF - Broadcast address
			6	0x13	
			7	0xA2	
			8	0x00	
			9	0x40	
			10	0x0A	
			11	0x01	
			LSB 12	0x27	
		16-bit Destination Network Address	MSB 13	0xFF	Set to the 16-bit address of the destination device, if known. Set to 0xFFFE if the address is unknown, or if sending a broadcast.
			LSB 14	0xFE	
		Broadcast Radius	15	0x00	Sets maximum number of hops a broadcast transmission can occur. If set to 0, the broadcast radius will be set to the maximum hops value.
		Options	16	0x00	Bitfield of supported transmission options. Supported values include the following: 0x01 - Disable ACK 0x20 - Enable APS encryption (if EE=1) 0x40 - Use the extended transmission timeout for this destination Enabling APS encryption decreases the maximum number of RF payload bytes by 4 (below the value reported by NP). Setting the extended timeout bit causes the stack to set the extended transmission timeout for the destination address. (See chapter 4.) All unused and unsupported bits must be set to 0.
		RF Data	17	0x54	Data that is sent to the destination device
			18	0x78	
			19	0x44	
			20	0x61	
			21	0x74	
			22	0x61	
			23	0x30	
			24	0x41	
	Checksum		25	0x13	0xFF - the 8 bit sum of bytes from offset 3 to this byte

Fig. 3.5 Features of ZigBee Transmit Request (0x10)[17]

	Frame Fields		Offset	Example	Description
API Packet	Start Delimiter		0	0x7E	
	Length		MSB 1	0x00	Number of bytes between the length and the checksum
			LSB 2	0x1A	
	Frame-specific Data	Frame Type	3	0x11	
		Frame ID	4	0x01	Identifies the UART data frame for the host to correlate with a subsequent ACK (acknowledgement). If set to 0, no response is sent.
		64-bit Destination Address	MSB 5	0x00	Set to the 64-bit address of the destination device. The following addresses are also supported: 0x0000000000000000 - Reserved 64-bit address for the coordinator 0x000000000000FFFF - Broadcast address
			6	0x00	
			7	0x00	
			8	0x00	
			9	0x00	
			10	0x00	
			11	0x00	
			12	0x00	
		16-bit Destination Network Address	MSB 13	0xFF	Set to the 16-bit address of the destination device, if known. Set to 0xFFFE if the address is unknown, or if sending a broadcast.
			LSB 14	0xFE	
		Source Endpoint	15	0xA0	Source endpoint for the transmission.
		Destination Endpoint	16	0xA1	Destination endpoint for the transmission.
		Cluster ID	17	0x15	Cluster ID used in the transmission
			18	0x54	
		Profile ID	19	0xC1	Profile ID used in the transmission
			20	0x05	
		Broadcast Radius	21	0x00	Sets the maximum number of hops a broadcast transmission can traverse. If set to 0, the transmission radius will be set to the network maximum hops value.
		Transmit Options	22	0x00	Bitfield of supported transmission options. Supported values include the following: 0x01 - Disable ACK 0x20 - Enable APS encryption (if EE=1) 0x40 - Use the extended transmission timeout for this destination. Enabling APS encryption decreases the maximum number of RF payload bytes by 4 (below the value reported by NP). Setting the extended timeout bit causes the stack to set the extended transmission timeout for the destination address. (See chapter 4.) All unused and unsupported bits must be set to 0.
		Data Payload	23	0x54	Data that is sent to the destination device
			24	0x78	
			25	0x44	
			26	0x61	
			27	0x74	
			28	0x61	
	Checksum		29	0x3A	0xFF - the 8 bit sum of bytes from offset 3 to this byte.

Fig. 3.6 Package Explicit Addressing ZigBee Command Frame (0x11)[17]

3.2 2D and 3D Environment for Indoor Localization

Fingerprint technique will apply two phases process. First phase is called the offline phase. In this phase, fingerprint information will be recorded as database. Second phase is called the online phase. In this phase, the current signal information will be compared with those in the database (RSSI). For example, let the number of reference nodes as N, number of fingerprint locations as M and $\mathbf{D}_m$ as fingerprint information (RSSI values for N nodes) at position $\mathbf{F}_m(m = 1,2, \ldots ., M)$ and r represents RSSI values from each node. Hence, the fingerprint technique can be expressed as

- Fingerprint location:

$$\mathbf{F}_m = [x_m \ , y_m] \tag{3.1}$$

- Fingerprint information:

$$\mathbf{D}_m = \left[r_{\{1,m\}}, r_{\{2,m\}} \ldots \ldots r_{\{N,m\}}\right]. \tag{3.2}$$

From here, the database consists of fingerprint location as $\mathbf{F}_m$ and fingerprint information as $\mathbf{D}_m$ is acquired. In the 2D scenario, 36 fingerprint locations are used as the fingerprint database, and 12 fingerprint locations for 3D scenario [21]. For the full explanation of FCM is described in Chapter 2.

3.2.1 2D Scenario

In 2D scenario, we simply record database by placing the transceiver in area of interest 5m × 5m. The layout of experiment setup using sets of 4 and 6 reference nodes is shown in Figure 3.7.

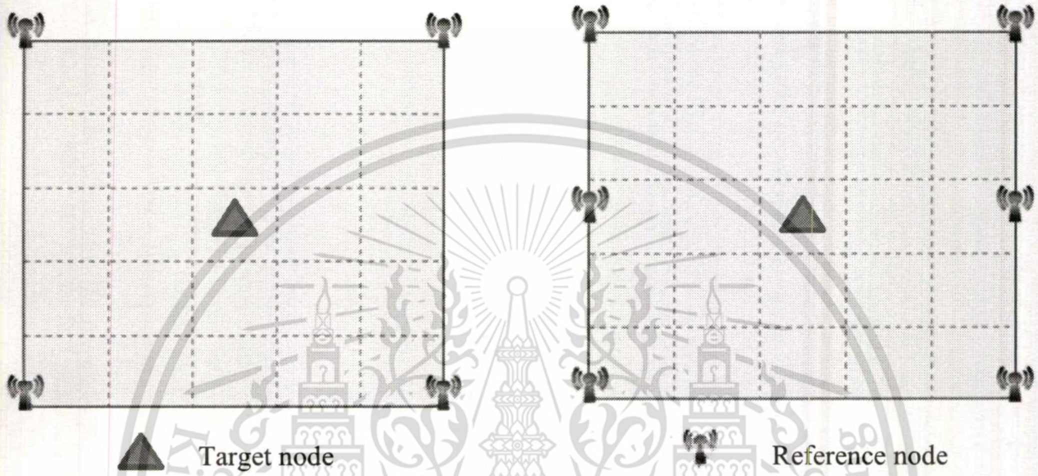


Fig. 3.7 Experiment setup for 2D scenario

In this work, the XBee-24ZB is used as reference nodes and target node, respectively. Xbee-24ZB is ZigBee module with the operating frequency of 2.4 GHz. The RSSI provided by ZigBee device is employed as parameter to estimate the location of the target node. The basic idea behind RSSI is the configured transmission power at the transmission device (TX) directly affects the receiving power at the receiving device (RX). In our scheme, the reference nodes receive the request message, and then will send intensity signal values to the target node. The target node will collect the data (RSSI) from reference nodes and forward it to the PC for estimating the target location.

3.2.1.1 2D Target Position Testing

In 2D scenario, the diagonal, horizontal and vertical testing positions are applied in order to validate the performance of fingerprint database. The illustration will show the system using 4 set and 6 set of reference nodes, respectively.

3.2.1.1.1 Diagonal Position Testing

The first experiment to obtain the validation of our proposed method, the diagonal position testing is applied. Figure 3.8 shows the fingerprint nodes and the target node at diagonal positions on the observed area for 4 set and 6 set of reference nodes. The target node is stationary placed for each experiment. For the left diagonal testing, the target node is first placed at (0 m, 0 m) and increase continuously by 0.33 m. For instance, the target node location for the diagonal left is moved from (0 m, 0 m) and continued until reaching the target position at (5 m, 5 m). For the right diagonal position, the target node location is moved from coordinates (5 m, 0 m) to (0 m, 5 m). The brown box shows in Figure 3.8 is the expected cluster in nearest location expectation of the target and yellow box is cluster in exact location of fingerprint location.

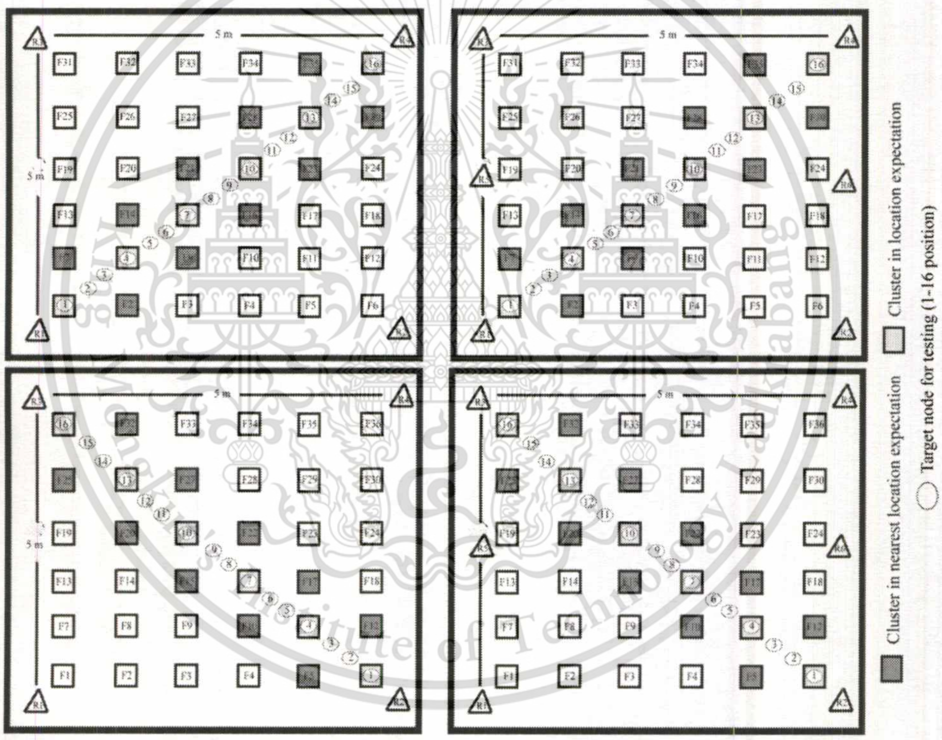


Fig. 3.8 Left and right diagonal testing

3.2.1.1.2 Horizontal and Vertical Testing

The other types of testing position are horizontal and vertical position. Figure 3.9 shows the measurement setup for horizontal and vertical of target node testing position. For the horizontal and vertical position testing, the target node is started to be placed from (0 m, 2.5 m) for the horizontal position; with the increasing distance are 0.25 m and 0.5 m. For example, position number 1 is (0 m, 2.5 m), number 2 (0.25 m, 2.5 m), number 3 (0.75 m, 2.5 m), and number 4 (1 m, 2.5 m). The position is continued until (5 m, 2.5 m). For the vertical position, it is begun at (2.5 m, 0 m) and ended at (2.5 m, 5 m).

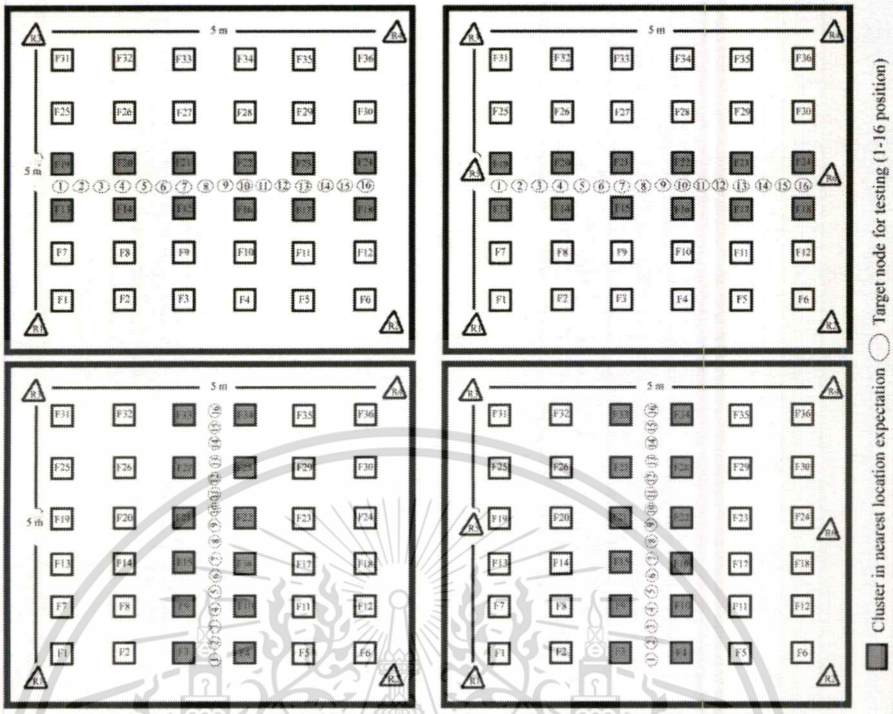


Fig. 3.9 Horizontal and vertical diagonal testing

3.2.2 3D Scenario

In 3D scenario, the real system consists of a bookshelf as the area of interest is applied. In this system, the realistic environment such as human body effect and propagation effect of media (furniture, books, etc.) are analyzed. As known in 3D scenario, the height or latitude difference is attracted many researchers to develop the indoor localization that relatively simple and effective. Based on previous explanation, in this work, the implementation of the FCM as clustering technique in the fingerprint based 3D indoor localization system is proposed.

The area of interest in this research is a 92 cm × 25 cm × 152 cm bookshelf in the clean environment. The 4 levels of this bookshelf are assumed as the location of target in 3D scheme. Each level has 35 cm height as shown in Figure 3.10. The room is empty room in corridor of a lobby and near the elevator in our institute. Target node is a wireless device which receives a packet from 6 reference nodes, which measure the received power. For validating the proposed method performance, the 16 target positions are deployed. The target is stationary placed between the fingerprint nodes as depicted in Figure 3.11.

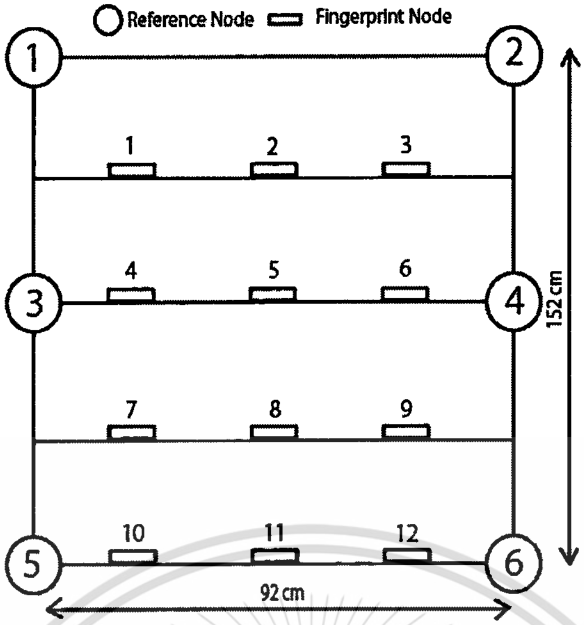


Fig. 3.10 Illustration of the experiment setup

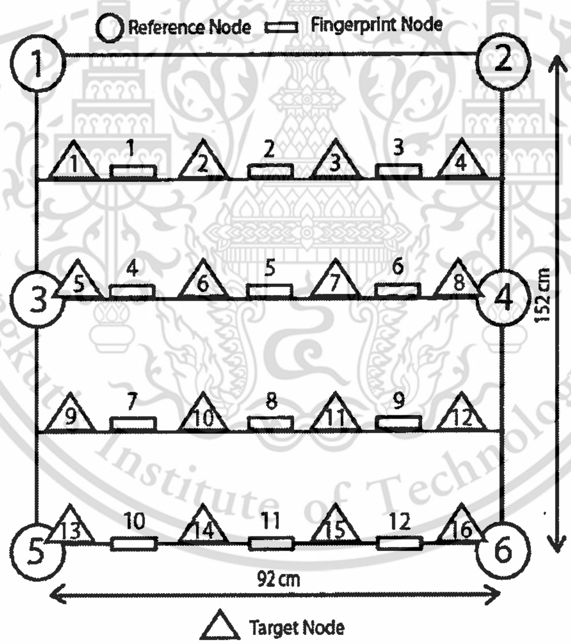


Fig. 3.11 Target node position

3.3 Pattern Matching Algorithm: Nearest Neighbor

In this work, the implementation of FCM is assigned as the pre–processing method before location estimation process in the pattern matching algorithm. In this work, Nearest Neighbor algorithm is deployed as pattern matching algorithm. The basic idea of this algorithm is to find the maximum similarity or the minimum error between target node and the fingerprint node (inside FCM cluster or group). Since the RSSI parameter is used in this work, the minimum error distance of RSSI value between target node and fingerprint node inside of FCM member is assigned as the location of the target.

When the target node in the area of interest collects the RSSI values from all reference nodes as

$$\mathbf{D}_t = \left[r_{\{1,t\}}, r_{\{2,t\}} \dots \dots r_{\{N,t\}}\right] \tag{3.3}$$

In the simulation, FCM cluster information as G is

$$\mathrm{G} = \left\{\mathbf{D}_{g_1}, \mathbf{D}_{g_2}, \mathbf{D}_{g_3}, \dots\dots, \mathbf{D}_{g_K}\right\} \tag{3.4}$$

and

$$\mathbf{D}_{g_k} = \left[\mathrm{r}_{\{1,g_k\}}, \mathrm{r}_{\{2,g_k\}}, \mathrm{r}_{\{3,g_k\}}, \dots\dots, \mathrm{r}_{\{N,g_k\}}\right] \tag{3.5}$$

FCM clusters the fingerprint information and the target signal information assigned as $\mathbf{D}_t$. RSSI values are represented by r from each reference node. The similarity between the target information with one of members in the cluster is verified by using Norm equation. For example, let the group of fingerprint locations as FCM clustering result as G, $\mathbf{D}_{g_k} (k = 1,2,3, \dots\dots, K)$ is the one member of G consists of RSSI values according to (3.5), and K is the number of members of G. The member of cluster is varying. From here, $\mathbf{D}_t$ and $\mathbf{D}_{g_k}$ are two points in Euclidean N-Space. In this manner, error distance, e_k is obtained by

$$e_k = \left\|\mathbf{D}_t - \mathbf{D}_{g_k}\right\|. \tag{3.6}$$

The estimated location will be obtained by choosing the minimum error from norm equation between target signal information and one of member in FCM cluster information which represents as the estimated location of the target.

3.4 Error Distance

The error distance is analyzed to validate the accuracy of our proposed method. This distance is between the true location (exact location) of the target node and the estimated location of the target node from our proposed localization system. In our work, the norm equation is used to find the error distance of the target location estimation. Like previously described in the Chapter 2, the error distance is acquired by the Euclidean distance estimation. These equations for 2D and 3D environment are

$$error = \sqrt{(x - x')^2 + (y - y')^2}, \tag{3.7}$$

and

$$error = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}, \tag{3.8}$$

where (x, y) and (x, y, z) is the true location of the target node (true testing position) and (x', y') and (x', y', z') are the estimated location of the target node (the actual position). From these equations, the accuracy of the system will be obtained by observing the error distance between the true location and the estimated location of the target.

Chapter 4

Results and Analysis

This section describes the methodology used in this research. The Fingerprint technique using Fuzzy C-Means (FCM) clustering algorithm for 2D and 3D scenes is evaluated. The performance evaluation is the clustered fingerprint nodes in the desired cluster with the target node. In addition, the error distance from true location of the target with the estimated location from localization system will be shown as the accuracy of the system. The data for evaluation is obtained by conducting indoor experiment and simulating it to get the validation of the system performance.

4.1 Simulation Systems and Parameters

In this research, simulation has been conducted to observe the performance of our proposed method. Table 4.1 shows the experiment system for 2D indoor localization system including the number of fingerprint, number of desired cluster in FCM and how we get the error distance from location estimation. For the 3D localization system, the description of experiment component is given in Table 4.2.

Table 4.1 Experiment systems for 2D indoor localization

Name	Number and Explanation
Fingerprint Node	36 fingerprint nodes, 6 nodes Each row.
Reference Node	4 set and 6 set, to know the scalability differences.
Area of Interest	5m × 5m
Grid Size (Fingerprint's Grid Size)	1m × 1m
Hardware	ZigBee (Xbee 24-ZB series), USB cable and connector, PC, power supply (battery).
Frequency	2.4 GHz
FCM Cluster	*C* or number of cluster equals to 10

Table 4.2 Experiment component for 3D indoor localization

Components	Explanation
6 Xbee 24-ZB	As reference nodes including its power source
1 Xbee 24-ZB	As target node
USB Cable	Connection between target node and PC
Bookshelf	92 cm × 25 cm × 152 cm
PC	Laptop Toshiba Satellite L745

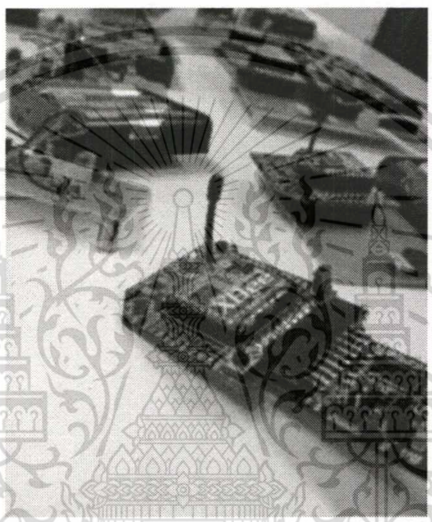


Fig. 4.1 Xbee 24-ZB as the target and reference node

In this simulation, RSSI information of the target node becomes the indoor localization parameter. The RSSI is stored in the same directory with the fingerprint database. The simulation is done based on the target location and the scenarios that have been explained in the previous chapter.

4.2 System Performances Results

4.2.1 2D Indoor Localization

In this subsection, the result of simulation system for 2D indoor localization will be given.

4.2.1.1 Fingerprint Database

In the fingerprint technique, the offline phase is conducted to acquire the fingerprint database. In this phase, signal information i.e., RSSI in this work is stored. Table 4.3 shows you the deployed component in this offline phase.

Table 4.3 Experiment component for 2D indoor localization

Component	Explanation
Xbee 24-ZB	4 and 6 set of reference nodes
Xbee 24-ZB	1 target node
USB Cable	Connection between target node and PC
PC	Laptop Toshiba Satellite L745
Fabric (5m × 5m)	Grid for fingerprint database in 1m × 1m

Fig. 4.2 Indoor experiment for 2D scenario

Fingerprint database for 2D indoor localization using 4 reference nodes is shown in Table 4.4.

Table 4.4 Fingerprint database for 4 reference nodes

Fingerprint Location	Reference 1 (-dBm)	Reference 2 (-dBm)	Reference 3 (-dBm)	Reference 4 (-dBm)
(0m,0m)	26.0	44.9	47.0	52.0
(1m,0m)	32.5	42.8	47.7	49.9
(2m,0m)	38.3	41.0	48.5	49.7
(3m,0m)	41.3	36.9	48.3	46.8
(4m,0m)	44.3	31.4	50.9	48.9
(5m,0m)	44.2	26.0	26.0	53.2
(0m,1m)	31.2	46.6	45.5	53.2
(1m,1m)	32.3	45.8	45.4	49.3
(2m,1m)	37.0	39.0	47.9	47.6

Table 4.4 Fingerprint database for 4 reference nodes (cont.)

Fingerprint Location	Reference 1 (-dBm)	Reference 2 (-dBm)	Reference 3 (-dBm)	Reference 4 (-dBm)
(3m,1m)	40.4	36.1	49.4	46.8
(4m,1m)	44.9	33.6	48.5	45.3
(5m,1m)	48.9	32.1	51.4	44.8
(0m,2m)	35.1	46.4	42.9	48.8
(1m,2m)	34.4	44.3	42.7	48.1
(2m,2m)	39.6	41.1	45.2	45.9
(3m,2m)	40.8	39.5	46.8	43.0
(4m,2m)	44.8	35.2	44.8	42.0
(5m,2m)	49.2	36.3	51.3	39.2
(0m,3m)	39.3	48.2	37.5	47.4
(1m,3m)	38.6	45.3	38.4	45.2
(2m,3m)	40.2	44.3	41.2	44.1
(3m,3m)	41.5	41.2	43.2	41.1
(4m,3m)	46.4	40.3	45.6	37.5
(5m,3m)	50.3	40.3	48.8	36.5
(0m,4m)	41.2	50.4	32.6	48.8
(1m,4m)	42.0	49.3	35.8	45.1
(2m,4m)	42.0	45.0	37.4	42.7
(3m,4m)	44.8	44.5	40.1	38.3
(4m,4m)	47.4	44.5	44.9	35.2
(5m,4m)	48.9	42.7	47.7	31.7
(0m,5m)	47.4	48.7	26.0	48.3
(1m,5m)	46.0	49.2	31.1	44.2
(2m,5m)	46.7	48.6	36.5	41.7
(3m,5m)	48.5	46.9	41.3	37.5
(4m,5m)	48.6	47.6	41.7	33.5
(5m,5m)	49.4	49.0	45.3	26.7

Table 4.5 Fingerprint database for 6 reference nodes

Fingerprint Location	Ref. 1 (-dBm)	Ref. 2 (-dBm)	Ref. 3 (-dBm)	Ref. 4 (-dBm)	Ref. 5 (-dBm)	Ref. 6 (-dBm)
(0m,0m)	26.0	46.1	46.2	55.8	37.3	53.9
(1m,0m)	31.4	41.8	47.6	51.6	40.1	46.9
(2m,0m)	36.4	39.1	48.3	50.1	41.2	43.2
(3m,0m)	40.8	36.3	48.0	47.2	46.0	40.2
(4m,0m)	43.0	31.1	53.5	51.1	49.1	38.6
(5m,0m)	42.7	26.0	53.7	43.9	54.8	34.7
(0m,1m)	31.8	45.7	45.7	53.7	36.8	49.3
(1m,1m)	33.4	45.4	45.0	51	36.9	44.7
(2m,1m)	36.2	39.1	47.7	48.6	40.4	41.1
(3m,1m)	39.9	36.4	48.9	46.2	45.3	37.4
(4m,1m)	45.0	33.6	49.7	45.6	46.9	33.9
(5m,1m)	47.4	31.5	50.6	44.2	50.7	32.2
(0m,2m)	35.1	46.7	41.6	49.6	30.2	46.2
(1m,2m)	34.2	44.8	42.3	48.4	34.9	43.8
(2m,2m)	41.0	42.1	44.2	44.9	38.4	40.4
(3m,2m)	41.2	39.7	46.3	43.0	41.7	35.7
(4m,2m)	44.9	35.1	47.6	41.9	44.8	31.8
(5m,2m)	48.4	34.6	50.6	40.9	47.9	28.1
(0m,3m)	39.6	47.8	37.4	49.1	29.0	46.8
(1m,3m)	39.1	44.6	38.5	45.8	31.8	43.3
(2m,3m)	41.2	44.0	40.7	44.8	35.9	41.4
(3m,3m)	44.1	41.5	43.7	41.8	40.1	36.9
(4m,3m)	46.8	39.5	45.8	38.5	42.5	32.6
(5m,3m)	49.7	39.2	48.4	36.8	44.2	28.8
(0m,4m)	42.4	50.5	32.3	50.4	33.9	48.0
(1m,4m)	43.3	48.1	35.3	45.5	33.9	44.5
(2m,4m)	43.9	45.1	37.2	42.2	38.1	42.5
(3m,4m)	45.4	44.4	40.7	39.0	40.2	39.0

Table 4.5 Fingerprint database for 6 reference nodes (cont.)

Fingerprint Location	Ref. 1 (-dBm)	Ref. 2 (-dBm)	Ref. 3 (-dBm)	Ref. 4 (-dBm)	Ref. 5 (-dBm)	Ref. 6 (-dBm)
(4m,4m)	48.1	44.0	44.5	36.2	42.4	34.7
(5m,4m)	50.9	44.2	49.0	31.9	47.7	33.6
(0m,5m)	47.2	50.8	26.0	49.0	37.8	49.1
(1m,5m)	47.3	49.9	31.3	44.2	36.6	47.4
(2m,5m)	48.3	49.7	36.9	42.1	40.1	43.1
(3m,5m)	47.5	45.0	41.2	38.8	42.2	41.7
(4m,5m)	49.3	47.1	42.9	33.8	44.7	37.1
(5m,5m)	50.7	46.7	45.8	27.0	47.1	37.5

4.2.1.2 Clustering Result for 2D Indoor Localization

In this subsection, the clustering result for 2D indoor localization from FCM algorithm is presented. The group has been clustered by FCM based on the similarity between the signal information (RSSI) value from target node. In this result, not all of the testing position is given, but selected positions which have the interesting result are emphasized. The left, right diagonal, vertical and horizontal position is briefly explained before for both of 4 set and 6 set of reference nodes. In the table will be stated that the desired cluster and actual cluster should be similar. This means that desired cluster is justified from the logic of position of the target node and the fingerprint node. The actual cluster is the cluster from result of FCM processing. As previously mentioned, for 2D environment, the 36 as {1}-{36} fingerprint nodes and 1 the target node as node number {37} will be clustered by FCM.

4.2.1.2.1 Clustering Result for Left Diagonal Position

In the left diagonal testing, FCM can cluster successfully in all of 16 positions as same with the desired cluster. In some positions located in the same location of fingerprint the actual cluster shows only two members in one cluster as desired cluster does.

Table 4.6 Clustering result for left diagonal position for 4 set reference node

No.	Desired Cluster	Actual Cluster
1	{1},{37}	{1},{37}
3	{1},{2},{7},{8},{37}	{1},{2},{7},{8},{37}
8	{15},{37}	{15},{22},{37}

10	{22},{37}	{22},{37}
14	{29},{30},{35},{36},{37}	{29},{30},{35},{36},{37}
16	{36},{37}	{36},{37}

Table 4.7 Clustering result for left diagonal position for 6 set reference node

No.	Desired Cluster	Actual Cluster
1	{1},{37}	{1},{37}
3	{1},{2},{7},{8},{37}	{1},{2},{7},{8},{37}
8	{15},{16},{21},{22},{37}	{21},{37}
10	{22},{37}	{16},{22},{23},{37}
14	{29},{30},{35},{36},{37}	{29},{30},{35},{36},{37}
16	{36},{37}	{36},{37}

4.2.1.2.2 Clustering Result for Right Diagonal Position

The yellow marker shows the unsatisfying result for clustering since FCM cannot cluster the target node in the specific fingerprint database member as desired or expected. This result could be affected by bad database training in certain location. Since the accuracy of this technique is depend on the quality of the database [6], [19], [20], [21].

Table 4.8 Clustering result for right diagonal position for 4 set reference node

No.	Desired Cluster	Actual Cluster
1	{6},{37}	{5},{6},{11},{12},{37}
3	{5},{6},{11},{12},{37}	{5},{6},{11},{12},{37}
8	{15},{16},{21},{22},{37}	{15},{22},{28},{37}
10	{20},{37}	{9},{15},{37}
14	{25},{26},{31},{32},{37}	{25},{31},{32},{37}
16	{31},{37}	{31},{37}

Table 4.9 Clustering result for right diagonal position for 6 set reference node

No.	Desired Cluster	Actual Cluster
1	{6},{37}	{5},{6},{37}
3	{5},{6},{11},{12},{37}	{5},{6},{11},{12},{37}

8	{15},{16},{21},{22},{37}	{15},{21},{37}
10	{20},{37}	{21},{37}
14	{25},{26},{31},{32},{37}	{25},{31},{32},{37}
16	{31},{37}	{31},{32},{37}

4.2.1.2.3 Clustering Result for Horizontal Position

Table 4.10 Clustering result for horizontal position for 4 set reference node

No.	Desired Cluster	Actual Cluster
1	{13},{14},{19},{20},{37}	{13},{14},{19},{20},{37}
3	{13},{14},{19},{20},{37}	{13},{14},{19},{20},{37}
8	{15},{16},{21},{22},{37}	{15},{37}
10	{15},{16},{21},{22},{37}	{15},{22},{37}
14	{17},{18},{23},{24},{37}	{17},{23},{24},{37}
16	{17},{18},{23},{24},{37}	{17},{23},{24},{37}

Table 4.11 Clustering result for horizontal position for 6 set reference node

No.	Desired Cluster	Actual Cluster
1	{13},{14},{19},{20},{37}	{14},{19},{20},{37}
3	{13},{14},{19},{20},{37}	{14},{19},{20},{25},{37}
8	{15},{16},{21},{22},{37}	{16},{21},{22},{23},{37}
10	{15},{16},{21},{22},{37}	{16},{21},{22},{37}
14	{17},{18},{23},{24},{37}	{17},{18},{ 24},{37}
16	{17},{18},{23},{24},{37}	{17},{18},{ 24},{37}

From Table 4.10 and 4.11, the reduction member of cluster compared with the desire cluster gives us benefit factors for the calculating process in the pattern matching algorithm that will be impacted in the power consumption of the node. For instance, in almost of horizontal testing position, the reduction of member in desired cluster is performed by FCM. This means that FCM is minimized the number of data that will be processed in the pattern matching algorithm as its function as pre-processing method for fingerprint technique.

4.2.1.2.4 Clustering Result for Vertical Position

Table 4.12 Clustering result for vertical position for 4 set reference node

No.	Desired Cluster	Actual Cluster
1	{3},{4},{9},{10},{37}	{3},{9},{37}
3	{3},{4},{9},{10},{37}	{4},{9},{10},{16},{37}
8	{15},{16},{21},{22},{37}	{15},{22},{37}
10	{15},{16},{21},{22},{37}	{15},{22},{37}
14	{27},{28},{33},{34},{37}	{27},{28},{34} ,{37}
16	{27},{28},{33},{34},{37}	{27},{28},{33},{34},{37}

Table 4.13 Clustering result for vertical position for 6 set reference node

No.	Desired Cluster	Actual Cluster
1	{3},{4},{9},{10},{37}	{3},{9},{15},{37}
3	{3},{4},{9},{10},{37}	{3},{9},{15},{37}
8	{15},{16},{21},{22} ,{37}	{16},{22} ,{37}
10	{15},{16},{21},{22},{37}	{21},{37}
14	{27},{28},{33},{34},{37}	{28},{34},{37}
16	{27},{28},{33},{34},{37}	{27},{28},{33},{34},{37}

Similar with the above explanation of horizontal testing position, in the vertical position FCM can cluster the target successfully among the fingerprint database. Furthermore, it can reduce the member of cluster that desired cluster has.

4.2.1.3 Estimated Error Location

The member of clustered group from FCM is matched with target node in order to know which fingerprint nodes has the most similarity with the target. After matching, the error distance will be occurred. If the error distance is 0 m, it means that the target node is located in the same location of fingerprint. But, if it is more than 0 m, it means that the target node is located between the fingerprint nodes. In this subsection, the comparison between the different sets of reference node is analyzed.

4.2.1.3.1 Left Diagonal Estimated Error Location

Figure 4.3 depicts the left diagonal estimated location error between four and six reference nodes. As we can see from the Figure 4.3, the error is reduced successfully in some positions. Not only reducing the error but it shows 0 m error as shown in the target node position number 13 and 16. In the other hand, in some position the additional sets of reference node cannot reduce the estimated error distance, such as in the position number 8, 11, and 12 which show the relatively high error.

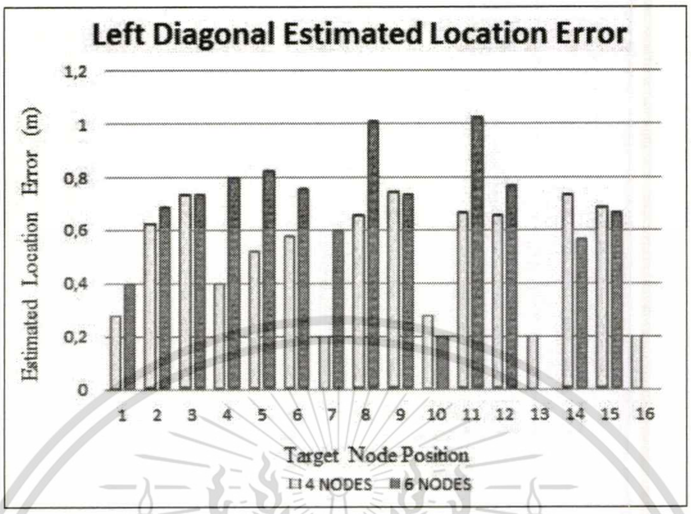


Fig. 4.3 Left diagonal estimated error location for 4 and 6 reference nodes

4.2.1.3.2 Right Diagonal Estimated Error Location

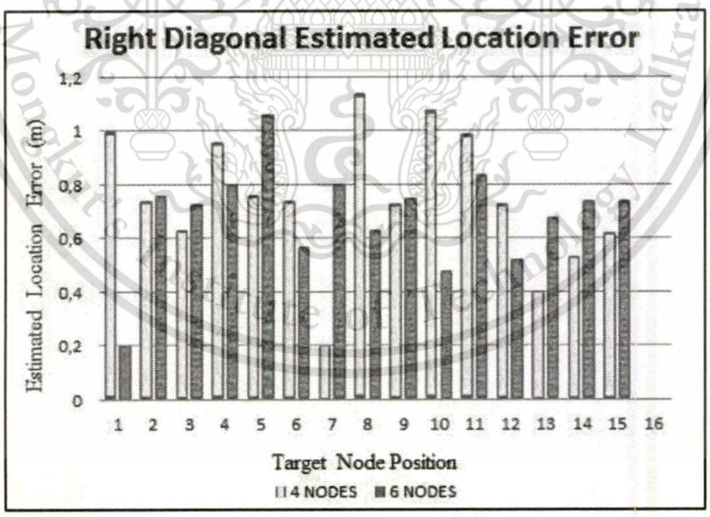


Fig. 4.4 Right diagonal estimated error location for 4 and 6 reference nodes

In Figure 4.4, when the target node approaches to reference nodes or in the border of the observed area, the error will reach relatively high rather than observed locations that are quite far from reference nodes. For instance, observed locations number 1, 5, 8 and 10, the results show that the error is relatively high rather than observed locations located in the center of the experiment area. We can see from the figure that the error for the 6 reference nodes less than the system using sets of 4 nodes in some observed points. It is occurred because of several factors, such as diffractions at edges, reflection on metallic objects, propagation media, etc [9], [10].

4.2.1.3.3 **Horizontal Estimated Error Location**

Similar case with other results, some error positions can be fixed by adding the sets of reference node but the others cannot be fixed. The effects relating with this issues can be named a few, propagation effect, the environment parameters i.e., temperature, humidity that can be different in the time of collecting data.

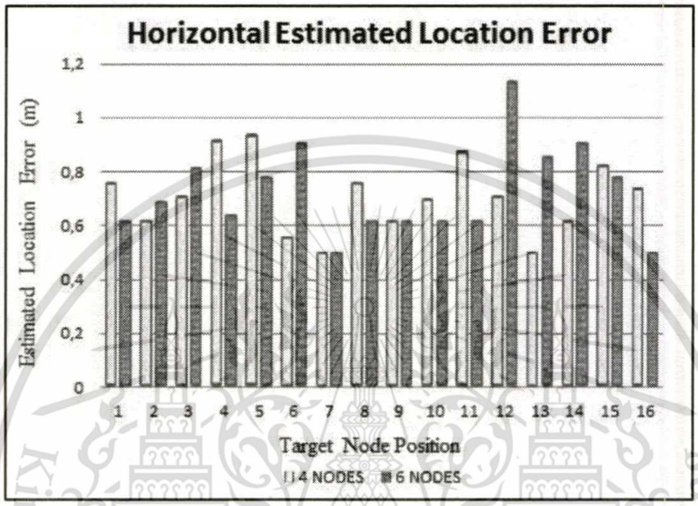


Fig. 4.5 Horizontal estimated error location for 4 and 6 reference nodes

4.2.1.3.4 Vertical Estimated Error Location

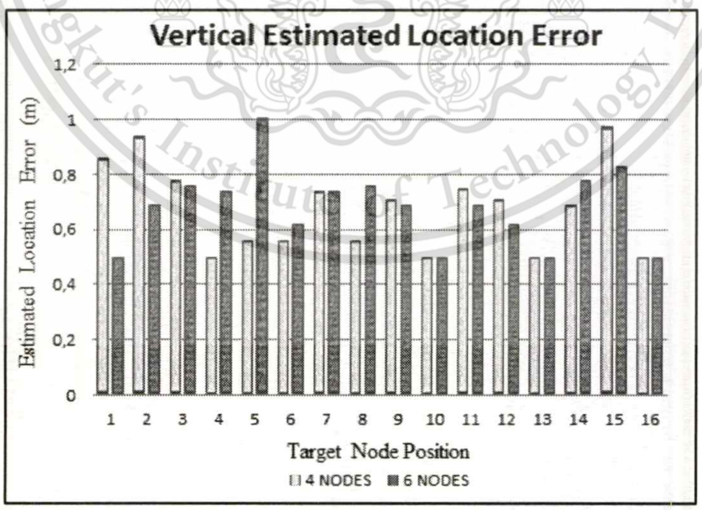


Fig. 4.6 Vertical estimated error location for 4 and 6 reference nodes

From the figures of estimated error location, the improvement in the accuracy of position estimation of each observed location as a function of the sets of reference nodes is seen in some positions. However, few locations give a relatively high error in the additional sets of reference node. The experiment is conducted many times, and the results are similar

4.2.2 3D Indoor Localization

In the 3D indoor localization, the system using a bookshelf which estimates position using fingerprint technique is applied. FCM algorithm is also deployed in this system. For validating the performance of the system, the different scenario of target estimation is utilized. The clean environment means minimizing the furniture and obstacle in the area of interest, the human body and furniture effects are considered as the 3D indoor localization testing as scenarios. Figure 4.7 shows the real measurement setup using a bookshelf. Figure 4.8 shows the Xbee that deployed as the reference nodes and the target node.

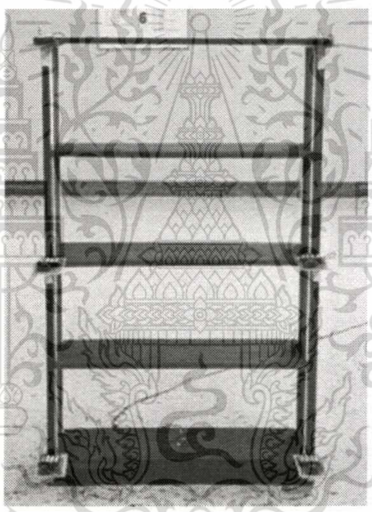


Fig. 4.7 A bookshelf as area of interest

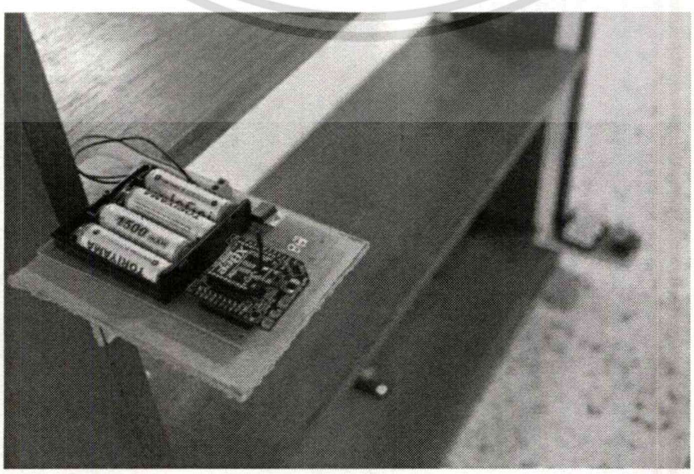


Fig. 4.8 Xbee 24-ZB as reference node

4.2.2.1 Fingerprint Database

In 3D indoor scenario, 12 fingerprint nodes are deployed. Since the 4-level of bookshelf is uniform (same high, same wide) the 3 fingerprint nodes in each level is expected to be dense enough. The gap between fingerprints is 22 cm. First experiment is conducted by placing the target node in the same position of fingerprint. This scenario has purpose to validate the performance of the fingerprint database. For this experiment, the result shows 0 cm error in all of target positions (12 positions). FCM can cluster the target node successfully in all the target node positions. From here, we can conclude that the quality of the fingerprint database is good.

The other scenarios i.e., between fingerprints in a clean environment, between fingerprints in furniture's effect, and between the fingerprints in human body's effect are presented in the next subsection.

Table 4.14 Fingerprint database

Fingerprint Location	Ref. 1 (-dBm)	Ref. 2 (-dBm)	Ref. 3 (-dBm)	Ref. 4 (-dBm)	Ref. 5 (-dBm)	Ref. 6 (-dBm)
(0m,0m)	30.4	42.9	34	34	48	53
(1m,0m)	30	38	31	34.9	45	53.9
(2m,0m)	34	40	36.8	32	52	55
(3m,0m)	56.7	40	28	31	54	40
(4m,0m)	34	69.3	31	32	41	43
(5m,0m)	34	62.5	35	29	40	52
(0m,1m)	52	53	34	46	33	34
(1m,1m)	45	53.6	32	40	31	32
(2m,1m)	44.1	56	42	45.2	35	32
(3m,1m)	60.4	56	51	40	28	38
(4m,1m)	52	56	40	40	33	34
(5m,1m)	48.7	58	42.5	43.1	50.5	28

The fingerprint database is obtained by conducting ten times experiment and ten times acquiring data for each fingerprint. The interval time between 1 data to another data is approximately 10-15 minutes. The data is stored to the PC directly via Xbee supported software. Table 4.14 shows the final fingerprint database (average of ten times RSSI experiment data from 6 reference nodes).

4.2.2.2 Clustering Result for 3D Indoor Localization

In this subsection, the clustering result from FCM algorithm is presented for 3D indoor localization. The groups have been clustered by FCM based on the similarity between the signal information (RSSI) value from target node and fingerprint nodes. The 16-target positions is deployed to validate the quality of the database and to observe the ability of FCM to cluster the target node among fingerprint database. Table 4.15 and Table 4.16 show the coordinate of fingerprint (FP) database and the target in the area of interest (bookshelf), respectively. The illustration of these coordinates can be seen in the previous chapter.

Table 4.15 Coordinates of fingerprint database

No.	FP Number	Coordinate (x, y, z)	No.	FP Number	Coordinate (x, y, z)
1.	FP 1	22 cm, 12.5 cm, 105 cm	7.	FP 7	22 cm, 12.5 cm, 35 cm
2.	FP 2	44 cm, 12.5 cm, 105 cm	8.	FP 8	44 cm, 12.5 cm, 35 cm
3.	FP 3	66 cm, 12.5 cm, 105 cm	9.	FP 9	66 cm, 12.5 cm, 35 cm
4.	FP 4	22 cm, 12.5 cm, 70 cm	10.	FP 10	22 cm, 12.5 cm, 0 cm
5.	FP 5	44 cm, 12.5 cm, 70 cm	11.	FP 11	44 cm, 12.5 cm, 0 cm
6.	FP 6	66 cm, 12.5 cm, 70 cm	12.	FP 12	66 cm, 12.5 cm, 0 cm

Table 4.16 Coordinates of the target

No.	Target Number	Coordinate (x, y, z)	No.	Target Number	Coordinate (x, y, z)
1.	T 1	11 cm, 12.5 cm, 105 cm	9.	T 9	11 cm, 12.5 cm, 35 cm
2.	T 2	33 cm, 12.5 cm, 105 cm	10.	T 10	33 cm, 12.5 cm, 35 cm
3.	T 3	55 cm, 12.5 cm, 105 cm	11.	T 11	55 cm, 12.5 cm, 35 cm
4.	T 4	77 cm, 12.5 cm, 105 cm	12.	T 12	77 cm, 12.5 cm, 35 cm
5.	T 5	11 cm, 12.5 cm, 70 cm	13.	T 13	11 cm, 12.5 cm, 0 cm
6.	T 6	33 cm, 12.5 cm, 70 cm	14.	T 14	33 cm, 12.5 cm, 0 cm
7.	T 7	55 cm, 12.5 cm, 70 cm	15.	T 15	55 cm, 12.5 cm, 0 cm
8.	T 8	77 cm, 12.5 cm, 70 cm	16.	T 16	77 cm, 12.5 cm, 0 cm

The target as mention in Table 4.16 is stationary placed in the bookshelf. The consideration of testing is emphasized on 3 types of environments; clean, furniture effect (book's effect), and human body effect. The clustering result for FCM as pre-processing method will be presented in the tables in the next subsection.

4.2.2.2.1 Clustering Result for Clean Environment Testing

In the clean environment testing, the obstacles inside or near the area of interest is minimized. There has no significant obstacles such as table, chair, and furniture inside the bookshelf. FCM can cluster successfully in almost all of 16 positions as same with the desired cluster. In some positions, the actual cluster shows only two members in one cluster. For 3D environment, 12 fingerprint nodes and 1 the target node {13} will be clustered by FCM.

Table 4.17 Clustering result for clean environment

No.	Desired Cluster	FCM Cluster	No.	Desired Cluster	FCM Cluster
T 1	{1},{2},{3},{13}	{1},{2},{3},{13}	T 9	{7},{8},{9},{13}	{9},{12},{13}
T 2	{1},{2},{3},{13}	{1},{2},{3},{13}	T 10	{7},{8},{9},{13}	{4},{7},{8},{13}
T 3	{1},{2},{3},{13}	{1},{2},{3},{13}	T 11	{7},{8},{9},{13}	{7},{8},{9},{13}
T 4	{1},{2},{3},{13}	{1},{2},{3},{13}	T 12	{7},{8},{9},{13}	{9},{12},{13}
T 5	{4},{5},{6},{13}	{4},{13}	T 13	{10},{11},{12},{13}	{10},{13}
T 6	{4},{5},{6},{13}	{1},{2},{3},{13}	T 14	{10},{11},{12},{13}	{10},{13}
T 7	{4},{5},{6},{13}	{5},{6},{13}	T 15	{10},{11},{12},{13}	{4},{12},{13}
T 8	{4},{5},{6},{13}	{5},{6},{13}	T 16	{10},{11},{12},{13}	{12},{13}

4.2.2.2.2 Clustering Result for Furniture Effect

In this experiment testing, books are placed surround the target to see the effect of propagation effect from books.

Table 4.18 Clustering result for furniture's effect

No.	Desired Cluster	FCM Cluster	No.	Desired Cluster	FCM Cluster
T 1	{1},{2},{3},{13}	{1},{2},{3},{13}	T 9	{7},{8},{9},{13}	{9},{12},{13}
T 2	{1},{2},{3},{13}	{1},{2},{3},{13}	T 10	{7},{8},{9},{13}	{7},{8},{9},{12},{13}
T 3	{1},{2},{3},{13}	{1},{2},{3},{13}	T 11	{7},{8},{9},{13}	{10},{13}
T 4	{1},{2},{3},{13}	{1},{2},{3},{13}	T 12	{7},{8},{9},{13}	{8},{9},{12},{13}
T 5	{4},{5},{6},{13}	{4},{13}	T 13	{10},{11},{12},{13}	{10},{13}
T 6	{4},{5},{6},{13}	{1},{2},{3},{13}	T 14	{10},{11},{12},{13}	{10},{13}
T 7	{4},{5},{6},{13}	{1},{2},{3},{13}	T 15	{10},{11},{12},{13}	{4},{12},{13}
T 8	{4},{5},{6},{13}	{5},{6},{13}	T 16	{10},{11},{12},{13}	{12},{13}

In the furniture's effect testing, it is found that some of target positions have different cluster members as desired cluster as shown in yellow marker. It could be occurred as a result of propagation effect from books. Moreover, the position between target and books can result the shadowing of antenna from reference node since the books behave like a wall between the target node and references nodes. It means the RSSI values from reference nodes could have errors and unreliable.

4.2.2.2.3 Clustering Result for Human Body Effect

In this experiment testing, the author stands 50 cm in the front of bookshelf in order to observe the effect of human body effect to the FCM clustering result.

Table 4.19 Clustering result for human body effect

No.	Desired Cluster	FCM Cluster	No.	Desired Cluster	FCM Cluster
T 1	{1},{2},{3},{13}	{1},{2},{3},{13}	T 9	{7},{8},{9},{13}	{4},{12},{13}
T 2	{1},{2},{3},{13}	{1},{2},{3},{13}	T 10	{7},{8},{9},{13}	{7},{8},{9},{13}
T 3	{1},{2},{3},{13}	{1},{2},{3},{13}	T 11	{7},{8},{9},{13}	{7},{8},{9},{13}
T 4	{1},{2},{3},{13}	{1},{2},{3},{13}	T 12	{7},{8},{9},{13}	{9},{12},{13}
T 5	{4},{5},{6},{13}	{5},{6},{13}	T 13	{10},{11},{12},{13}	{10},{13}
T 6	{4},{5},{6},{13}	{1},{2},{3},{13}	T 14	{10},{11},{12},{13}	{10},{13}
T 7	{4},{5},{6},{13}	{1},{2},{3},{13}	T 15	{10},{11},{12},{13}	{4},{10},{12},{13}
T 8	{4},{5},{6},{13}	{1},{2},{3},{13}	T 16	{10},{11},{12},{13}	{4},{12},{13}

In the Table 4.19 for human body effect, it seems that human body effect in the front of area of interest give the propagation effect to the system. As the yellow marker show the clustering error (do not follow the trend of desired cluster).

4.2.2.3 Estimated Error Location

The member of clustered group from FCM is matched with target node in order to know which fingerprint nodes has most similarity with the target. After matching, the error distance will occur. The target is placed between the fingerprint nodes (11 cm). The evaluation of fingerprint database will be validated by three types of environment: clean environment, furniture's effect, and human body effect.

4.2.2.3.1 Clean Environment

Figure 4.9 depicts the graph of estimated error location for 16 target locations in between fingerprint nodes in clean environment. The average error for all 16 positions is 19.5 cm. The target position number 6 gives the error of 37 cm, in the FCM result, it is located in the fingerprint number 1 (first upper level). In fact, the target position number 6 is located in 2nd upper level of bookshelf). The other position which give error (1, 2, and 9), these positions show error but FCM can cluster them in the same level of bookshelf.

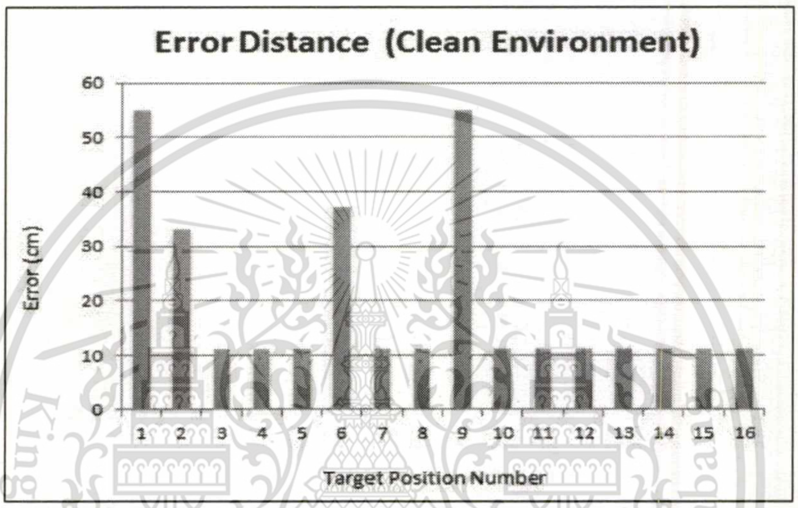


Fig. 4.9 Estimated error distance for clean environment

4.2.2.3.2 Furniture Effect

In this furniture effect experiment, books are placed in the center of each level of bookshelf. Figure 4.10 depicts the graph of error distance for furniture's effect.

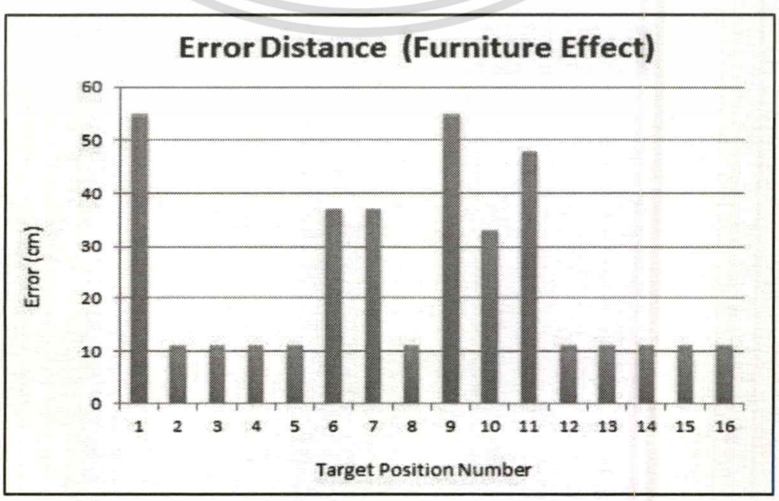


Fig. 4.10 Estimated error distance for furniture effect

In this experiment, the average error for al 16-position is 23.4 cm. The position number 6, 7 and 11 give the unsatisfying result since FCM cluster them into the different level of bookshelf. This can be happened as the issues of the edge, propagation effect, also the effect of furniture's itself (propagation effect from books).

4.2.2.3.3 Human Body Effect

Human body's effect is the last experiment to see the system performance. Here, author is 50 cm standing in the front of bookshelf. In this experiment, bookshelf is clean (books have been removed). Figure 4.11 depicts the error distance graph from this experiment.

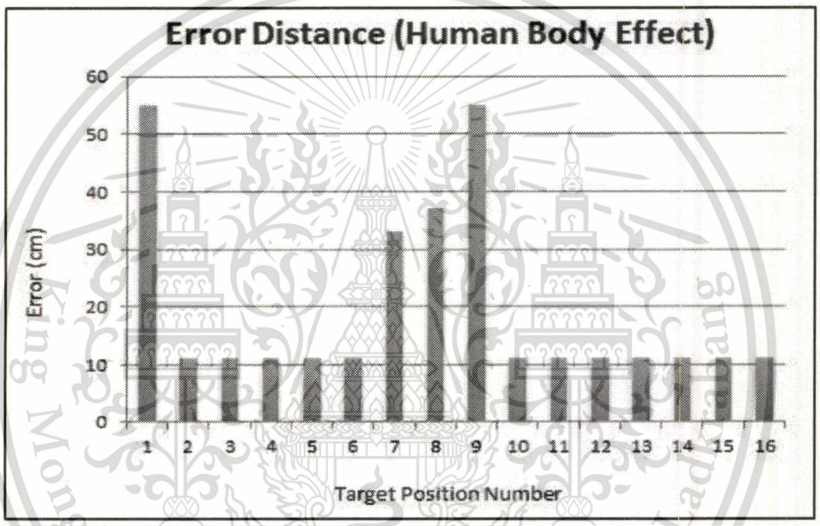


Fig. 4.11 Estimated error distance for human body effect

The average of error distance for human body's effect experiment is 19.5 cm. The results show good result since the FCM can cluster the target node in the same level of desired fingerprint nodes. In other words, the errors come from the nodes which have been clustered by FCM in the same level of bookshelf. The results also show that the human body's effect does not affect much in the result of proposed localization system.

Chapter 5

Conclusions

5.1 Conclusions

The objective of this thesis is to implement the clustering technique as pre-processing phase in the fingerprint-based indoor localization technique. Fuzzy C-Means (FCM) as commonly method for data clustering is deployed. This implementation is expected to improve the performance of the localization system in its accuracy and its efficiency. Two different methods are used to investigate the performance of FCM to cluster the data in groups as the pre-processing before the pattern matching algorithm deciding the best match between fingerprint and the target. The indoor experiments are conducted as the area of interest in 2 Dimensional (2D) and 3 Dimensional (3D), respectively. The accuracy of FCM to cluster the target along with the fingerprint database is carried out. The accuracy as the ability of indoor localization system is emphasized.

As the number of fingerprint database in wireless sensor networks (WSNs)-based localization system increased, the demand of developing the effective system including the power effective, positioning accuracy, are also increased. Since the common technique fingerprint technique is used the whole fingerprint database for matching in order to find the location of the target, the system will consume power and much time in the localization process. As clustering technique can group the fingerprint database and the target in one group (the member of group is minimized), the effective system can be achieved by employing the pattern matching algorithm in that group, the work is carried out under this assumption. In this thesis, the clustering technique and FCM described in Chapter 2 are presented to be implemented to the conventional fingerprint technique. In doing so, the large number of fingerprint database implied the higher accuracy but still low power and effective localization system can be accommodated.

The Scenarios and research methodology are described in Chapter 3. The explanation on how the validation of our system is also given in Chapter 3.

The performance of FCM for clustering fingerprint database and the target node is presented in Chapter 4. From the simulation results, it can be shown that FCM can mostly cluster the target node in the fingerprint database and minimize the member for calculating in the pattern matching process. But, in some position testing FCM cannot cluster successfully. Further investigation is needed to conclude that is not only the quality of fingerprint database but also how we maintain the reference nodes and the target node performance in acquiring RSSI values. In some position, especially in 3D localization, it is found that if there is one node is not in a good performance, the whole data will be unsatisfying. Our work can be summarized in this chapter, Chapter 5 as conclusion and future work.

The use of FCM as pre-processing phase in Fingerprint Technique for localization gives a new idea for minimizing the data for the next process (pattern matching process). One advantage is the minimization of complexity in order to achieve the simple, accurate, and effective system.

5.2 Future Research Works

The details theoretical study of Fuzzy-C Mean clustering technique and its implementation in fingerprint-based indoor localization technique as pre-processing phase are presented in this thesis. For the future research, the practical implementation including the real-time acquiring data and the target position is needed.

To enhance the performance for minimizing power and complexity, the FCM algorithm can be installed in the micro-processor system to maintain the power of clustered group and non-clustered group. For which, the power effective will be obtained as well as an efficient localization system. Furthermore, the implementation of FCM in fingerprint technique will be a good topic to work in comparison with the performance of other techniques.

References

[1] Caffery, J. J. and Stuber, G. L. *Overview of Radiolocation in CDMA Cellular Systems*, IEEE Communications Magazine, pp. 38-45, April 1998.

[2] Rappaport, T. S., Red, J. H., and Woerner, B. D. *Position Location Using Wireless Communications on Highways of the Future*, IEEE Communications Magazine, pp. 33-41, October 1996.

[3] Sun, G., Chen, J., Guo, W., and Liu, K. J. R. *Signal Processing Techniques in Network Aided Positioining*, IEEE Signal Processing Magazine, pp. 12-23, July 2005.

[4] Kikuchi, S., Shano, A., Tsuji, H., and Miura, R. *Mobile Localization Using Local Scattering Model in Multipath Environments*, Proc. 60th IEEE Veh. Technol. Conf., vol. 1, pp. 339-343, Sept. 2004.

[5] Parameswaran, A.T, Husain, M. I, and Upadhyaya, S. *Is RSSI a Reliable Parameter in Sensor Localization Algorithms – an Experimental Study,* Field Failure Data Analysis Workshop (F2DA 2009), New York, USA, 2009.

[6] Cherntanomwong, P., Takada, J., and Tsuji H., *Signal Subspace Interpolation from Discrete Measurement Samples in Constructing a Database for Location Fingerprint Technique*, *IEICE Trans. Commun.*, Vol.E92-B, no.9, pp. 2922-2930. September 2009.

[7] Bezdek, J. C., *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York. 1981.

[8] Hoang, D. C., Kumar, R., and Panda, S. K., "Fuzzy C-Means Clustering Protocol for Wireless Sensor Networks," *IEEE International Symposium on Industrial Electronics (ISIE-2010),* Bari, Italy, 2010.

[9] Akyldiz, I. F., Su, W., Sankarasubranamniam, Y., and Cayirci, E., *A Survey on Sensor Networks*, IEEE Communications Magazine, pp. 102–114, August 2002.

[10] Figueiras, J., and Frattasi, S., *Mobile Positioning and Tracking-From Conventioanl to Cooperative Techniques*," John Willey and Sons, June 2010.

[11] Nakarach, N., Bhokam., T., and Jomprom, K., *Localization and Tracking in Wireless Sensor Networks*," Bachelor Thesis, KMITL 2011.

[12] Cherntanomwong, P., *Radio Localization using Array Antenna for Outdoor Applications*, Dissertation, Tokyo Institute of Technology, Tokyo, Japan , 2008.

[13] Egan, M. A., "Locating Clusters in Noisy Data: A Genetic Fuzz C-Means Clustering Algorithm." *Proceeding of the 1998 Conference of the North American Fuzzy Information Processing Society. IEEE*, pp. 178-182, 1998.

[14] Yager, R.R., *Fuzzy Sets and Applications: Selected Papers by L.A. Zadeh*, John Wiley, New York, 1987.

[15] Valente de Oliveira, J., and Pedrycz, W., *Advance in Fuzzy Clustering and Its Application*, John Wiley and Sons, New York, 2007

[16] Faharani, S., *ZigBee Wireless Networks and Transceivers*, Newnes, 2008.

[17] The XBee/XBee-PRO/S2B ZB Manual, Digi International Cooperation, 2010.

[18] Tadakamadla, S., *Indoor Local Positioning System for ZigBee, Based on RSSI*, Master Thesis, Mid Sweden Universty, 2006.

[19] Cherntanomwong, P., and Suroso, D. J., "Indoor localization using Wireless Sensor Networks for Stationary and Moving Target," *The 8th ICICS 2011*, Singapore, Dec. 2011.

[20] Suroso, D. J., Cherntanomwong, P., Sooraksa, P., and Takada, J., "Location Fingerprint Technique using Fuzzy C-Means Clustering Algorithm for Indoor Localization," *IEEE TENCON 2011*. November 2011.

[21] Suroso, D. J., Chuenurajit, T., Cherntanomwong, P., and Takada, J., "Fingerprint Technique for 3D Indoor Localization using Fuzzy C-Means Clustering Algorithm," *ITC-CSCC 2012*. Japan. July 2012.

[22] Sayadi, M., Tlig, L., and Fnaiech, F., "A New Texture Segmentation Method Based on the Fuzzy C-Means Algorithm and statistical Features," *Applied Mathematical Sciences*, vol.1, no.60, pp. 2999–3007, 2007.

[23] Roullier, V., Cavaro-Menard, C., Calmon, G., and Aube, C., "Fuzzy algorithms: Application to adipose tissue quantification on MR images," *Biomedical Signal Processing and Control*, vol. 2, issues 3, pp 239-247, July 2007.

[24] Jain, N. and Shukla, S., "Fuzzy Databases Using Extended Fuzzy C-Means Clustering," *International Journal of Engineering Research and Applications (IJERA)*, vol. 2, issues 3, pp 1444-1451, June 2012.

[25] Rezzae, R.M., Lelieveldt, B.P.F., and Reiber, J.H.C., "A New Cluster Validity Index for the Fuzzy *C*-Mean," *Pattern Recognition Letters*, vol. 19, pp 237-246, 1998.

[26] http://www.stg.com/wireless/ZigBee_netw.html

[27] http://www.ladyada.net/make/xbee/wiring.html

Author Biography

Personal Information

Name	DWI JOKO SUROSO
Nationality	Indonesian
Date of birth	August 23, 1988
Place of birth	Sleman, Yogyakarta, Indonesia

Education

Bachelor degree

Field	Engineering Physics
Duration	2006-2010
Department	Department of Engineering Physics
University	Gadjah Mada University (UGM), Indonesia

Master degree

Field	Computer Engineering
Duration	2010-2012
Department	School of Computer Engineering
Faculty	International College
University	King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand

Research Interests

Wireless Communication Systems: wireless sensor networks, target finding algorithm, wireless module and hardware applications, fingerprinting techniques.

List of Publication

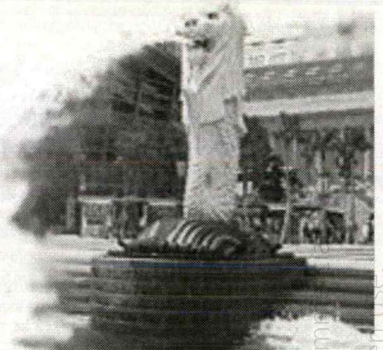
1. Nuttapon Nakarach, Thanasate Bhokam, Kanokwan Jomprom, **Dwi Joko Suroso**, Panarat Cherntanomwong, "Implementation of RSSI-based Localization System Using Wireless Sensor Networks Based on ZigBee Standard," Regional Conference of Information and Communication Technology, AUN/SEED-Net JICA, Lao PDR, March 2011.
2. **Dwi Joko Suroso**, Achmad bayhaqy, Muhfizaturrahmah, Panarat Cherntanomwong, and Sunarno, "Design of Remote Early Warning Cold Lava Avalanche System using Wireless Sensor Networks, Case Study of Merapi Volcano, Yogyakarta, Indonesia," Conference of Information Technology and Electrical Engineering (CITEE 2011), Gadjah Mada University, Indonesia, July 2011.
3. **Dwi Joko Suroso**, Panarat Cherntanomwong, and Jun-ichi Takada, "Fingerprint-based Localization Technique in Wireless Sensor Networks using Fuzzy C-Means Clustering Algorithm, Emphasizing in the Quality of the Database", Thailand-Japan MicroWave 2011 (TJMW 2011), King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, August 2011.
4. **Dwi Joko Suroso**, Panarat Cherntanomwong, Pitikhate Sooraksa, and Jun-ichi Takada, "Location Fingerprint Technique using Fuzzy C-Means Clustering Algorithm for Indoor Localization", TENCON 2011, Bali, Indonesia, November 2011.
5. **Dwi Joko Suroso**, Panarat Cherntanomwong, and Jun-ichi Takada, "Fingerprint-based Localization Technique in Wireless Sensor Networks using Fuzzy C-Means Clustering Algorithm", ISPACS 2011, Chiang Mai, Thailand, December 2011.
6. Panarat Cherntanomwong and **Dwi Joko Suroso**, "Indoor Localization System using Wireless Sensor Networks for Stationary and Moving Target", ICICS 2011, Singapore, December 2011. (*Won High Quality Paper Award*)
7. Thanapong Chuenurajit, **Dwi Joko Suroso**, and Panarat Cherntanomwong, "Implementation of RSSI-based 3D Indoor Localization System using Wireless Sensor Networks Based on ZigBee Standard", NC ICT 2012, Thailand, April 2012.
8. Sisongkham Phimmasean, **Dwi Joko Suroso**, and Panarat Cherntanomwong "Development of User-friendly GUI for Indoor Localization using RFID Passive Tag", AIT 2012, Thailand, 23 June 2012.
9. **Dwi Joko Suroso**, Thanapong Chuenurajit, and Panarat Cherntanomwong "Fingerprint Technique for 3D Indoor Localization using Fuzzy C-Means Clustering Algorithm", ITC-CSCC 2012, Japan, July 2012.

10. Sisongkham Phimmasean, **Dwi Joko Suroso**, and Panarat Cherntanomwong "Fingerprint Technique for Indoor Localization in RFID Passive Tag Using Dot Product", TJMW 2012, Thailand, August 2012. (*Won Young Research Best Paper Award*)
11. **Dwi Joko Suroso,** Panarat Cherntanomwong, and Sunarno "Mitigation Disaster: An Application of Wireless sensor Networks", TEKNOFISIKA 2012, Indonesia, April 2012. (*Journal Paper*)
12. Thanapong Chuenurajit, **Dwi Joko Suroso**, and Panarat Cherntanomwong "Implementation of RSSI-based 3D Indoor Localization System using Wireless Sensor Networks Based on ZigBee Standard", JIST 2012, Thailand. (*Journal Paper*)

ICICS 2011

8th International Conference on Information, Communications and Signal Processing

December 13 - 16, 2011; SINGAPORE



Networking

Multimedia Systems

Communications

Signal Processing

Information Systems

IEEE Catalog Number: CFP11435-CDR
ISBN: 978-1-4577-0030-9

Introduction | Sessions | Author Index | Search

Indoor Localization System using Wireless Sensor Networks for Stationary and Moving Target

Panarat Cherntanomwong
Department of Computer Engineering
Faculty of Engineering, KMITL
Bangkok, Thailand
krpanara@kmitl.ac.th

Dwi Joko Suroso
International College, KMITL
Bangkok, Thailand
S3601152@kmitl.ac.th

***Abstract*— The recent research in localization technique has been supported by the emerging of wireless sensor network (WSN) technology. The issues of the estimated location accuracy have becoming the main research topics in WSN-based localization technique. In this paper, the application of WSN for localization technique using IEEE 802.15.4 standard is proposed. As IEEE 802.15.4 standard, ZigBee is widely used because of its advantages. It specially provides useful parameters for location estimation, i.e. received signal strength indicator (RSSI) and link quality indicator (LQI). In this paper, the simple but effective localization system is proposed. Range-based and location fingerprint-based are implemented. The different set of the reference nodes is applied. The effectiveness of this method is verified by data obtained from an indoor experiment in the 5m × 5m observed area. The error from estimated locations using both techniques is analyzed. The estimated location results from different amount of reference nodes are compared. The results show that the error from the observed area is less than 1.2 m for both localization techniques, respectively. The GUI is created to help the experiments in the visualization and for showing the estimated location. From this result, the simple but effective system is really beneficial for the real application.**

***Keywords*—Indoor Localization; Wireless Sensor Network; ZigBee; Range-based; Location Fingerprint Technique**

I. INTRODUCTION

Radio localization using smart sensor nodes in WSNs has become popular in recent years. It is supported by the advancement of micro-mechanical systems and development of digital electronics technology. Radio localization is realized to be an important application for daily life and WSNs is popularly selected to be used for a localization system. ZigBee as IEEE 802.15.4 standard is widely used because of its various advantages, i.e. cost-effective, low-power consumption, security, robustness, reliability and supporting low data rates. ZigBee provides useful parameters for location estimation, i.e. RSSI and LQI [1], [2].

Localization technique can be classified into several categories based on the parameter that is used. Range-based and location fingerprint technique are most common technique in WSN applications. In this paper, an indoor scenario is applied. In the indoor scenario, many issues such as diffraction at edges, refraction by media with different propagation velocity and reflection in metallic objects. These issues affect to the localization system performance [3].

In this paper, we implement the user-friendly GUI for sending commands, receiving and calculating data based on the initial algorithm. For the experiment, we use the range-based and fingerprint-based techniques. The command data is sent to the reference nodes then it will initialize the position of reference nodes. Then, the reference nodes will transmit parameter that will be received by the target node as received power. The received signal power that is received by the target node will be transferred to personal computer (PC) through serial communication. The GUI for this experiment has the serial communication connection to provide the communication between PC and the target node. The GUI will display the position estimation results and has ability to store the experiment data result. The experiment is evaluated using two different types of the target node, i.e., stationary and moving target node.

The paper is organized as follow: Section 2 presents the description of the algorithm model. The experiment setup is explained in Section 3. Result and discussion are described in Section 4. Finally, in Section 5, we conclude this paper with plans for future work.

II. DESCRIPTION OF ALGORITHM MODEL

A. RSSI Definition

RSSI in dBm is defined as ten times the logarithm of the ratio of power (P) at the receiving end and the reference power ($Pref$). Power at the receiving end is inversely proportional to the square of distance. The received signal strength depends on the transmitted power and the distance between the transmitter and the receiver. In the embedded devices, the received signal strength is converted to RSSI. The relationship between RSSI and distance can be determined as equation [4].

$$RSSI[dBm] = A - [10 \cdot n \cdot log_{10}(d/d_0)], \tag{1}$$

where n is the path loss exponent or the signal propagation constant, d is the distance from transmitter in meter, d_0 is a reference distance, typically 1 meter, and A is the received signal strength at 1 meter distance, in dBm.

B. Range-based Localization

In this paper, the maximum likelihood algorithm is applied as range-based localization. In Range-based, the node distance is acquired through measurement of RSSI without additional node hardware design. Fig. 1 depicts the illustration of measurement system.

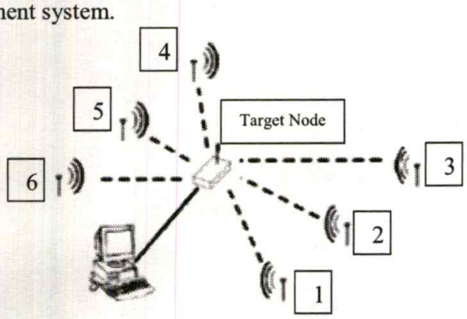


Figure 1. The illustration of measurement system

Suppose coordinate of reference nodes are $B1(x_1, y_1), B2(x_2, y_2)$,... and $Bn(x_n, y_n)$. And coordinate of the unknown node is determined as $O(x, y)$, the distance between the unknown node and reference nodes are $d_1, \ldots, d_n$, respectively. A group of non-linear equation can be obtained according to the following formula in a two-dimensional space as shown [5]:

$$\begin{cases} (x-x_1)^2 + (y-y_1)^2 = d_1^2 \\ (x-x_2)^2 + (y-y_2)^2 = d_2^2 \\ \vdots \\ (x-x_n)^2 + (y-y_n)^2 = d_n^2 \end{cases} \tag{2}$$

When the last equation is subtracted from the other equations in turn beginning from the first equation, hence

$$\begin{cases} x^2 - x_n^2 - 2(x_1 - x_n)x + y_1^2 \\ -y_n^2 - 2(y_1 - y_n)y = d_1^2 - d_n^2 \\ \vdots \\ x_{n-1}^2 - x_n^2 - 2(x_{n-1} - x_n)x + y_{n-1}^2 \\ -y_n^2 - 2(y_{n-1} - y_n)y = d_{n-1}^2 - d_n^2 \end{cases} \tag{3}$$

Linear equation in equation (3) can be represented as

$$\boldsymbol{Xb} = \boldsymbol{a}$$

where

$$\boldsymbol{X} = \begin{bmatrix} 2(x_1 - x_n) & 2(y_1 - y_n) \\ \vdots & \vdots \\ 2(x_{n-1} - x_n) & 2(y_{n-1} - y_n) \end{bmatrix},$$

$$\boldsymbol{a} = \begin{bmatrix} x_1^2 - x_n^2 + y_1^2 - y_n^2 - d_1^2 + d_n^2 \\ \vdots \\ x_{n-1}^2 - x_n^2 + y_{n-1}^2 - y_n^2 - d_{n1}^2 + d_n^2 \end{bmatrix},$$

$$\boldsymbol{b} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

The error of the node estimation location can be obtained by applying the standard minimum mean square error as

$$\hat{\boldsymbol{b}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{a}\,. \tag{4}$$

C. Location Fingerprint Technique

Location fingerprinting is a location sensing technique which involves a two-phase process. First, during the off-line calibration phase the received signals at selected locations are recorded in a database. Then, the second phase, called the on-line, pattern matching algorithms are used to infer a target's location by comparing the current observed signal features to the pre-recorded values in the database. The key points of this technique are the selection of the spatial signature and the method in constructing database. The commonly used of pattern matching algorithm as shown in Figure 1 [6], [7].

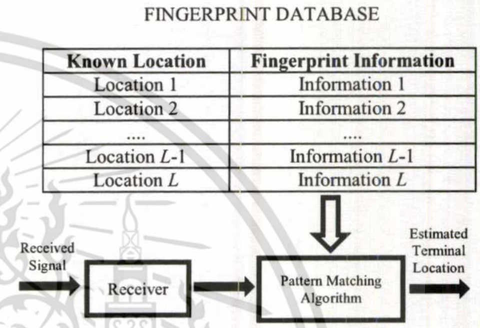


Figure 2. Location fingerprint technique

For simplicity, the fingerprint is defined as sets of measured RSSIs from all sensor nodes. Hence, let the number of sensor nodes that placed at fix locations as N, number of fingerprint locations as M and D_i as fingerprints at position $D_i (i = 1,2, \ldots, M)$. Table I shows the illustration of fingerprint database for the experiment.

TABLE I. FINGERPRINT DATABASE

Fingerprint Location	Fingerprint Information
$F_1(x_1, y_1)$	$D_1 = \{rssi_{1,F_1}, rssi_{2,F_1}, \ldots . rssi_{N,F_1}\}$
$F_2(x_2, y_2)$	$D_2 = \{rssi_{1,F_2}, rssi_{2,F_2}, \ldots . rssi_{N,F_2}\}$
......	......
$F_M(x_M, y_M)$	$D_M = \{rssi_{1,F_M}, rssi_{2,F_M}, \ldots . rssi_{N,F_M}\}$

When the target node is in the fingerprint area, it collects the RSSIs from all sensor nodes as

$$T = \{rssi_{1,T}, rssi_{2,T}, \ldots . rssi_{N,T}\}.$$

Euclidean distance is used to find the location of target node by comparing it with the fingerprint database, D_i. D_i and T are two points in Euclidean N-space, then the distance is obtained by the equation

$$dist_i(x,y) = \sqrt{\sum_{n=1}^{N}\left(rssi_{n,T} - rssi_{n,D_i}\right)^2}. \quad (5)$$

III. Experiment System And Setup

A. Experiment System

The authors reported the localization system based on ZigBee standard for the stationary target [8]. In this paper, we improve the system which is able to estimate the location of moving target. The ZigBee module, XBee-24ZB is used as reference nodes and target node. The ZigBee provides the RSSI parameter for experiments. The different amount of reference nodes is deployed for the experiments. There are 4 and 6 reference nodes. Fig. 2 depicts the illustration for experiment system. The 5m × 5m observed area is used as the experiments location for both maximum likelihood as range-based and location fingerprint techniques, respectively.

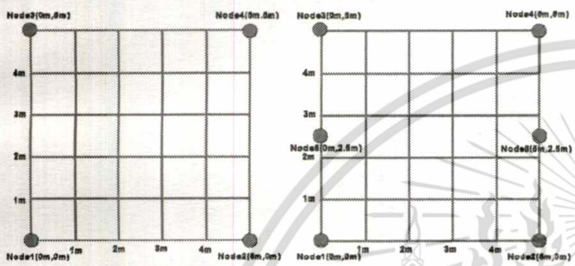


Figure 2. The illustration of experiment system using 4 reference nodes (left) and 6 reference nodes (right)

In this paper, the system estimate the location based on RSSI in the observed area. The target node is a wireless device which receives a packet from reference nodes, which each measures the received power. After receiving a packet, the target node measures RSSI and sends the results to the sink node. In this research, we use personal computer (PC) as sink node which graphical user interface (GUI) is already installed inside. The GUI for the experiment is shown in Fig. 3. The GUI is very useful for the users to see the location of the target node in the real time.

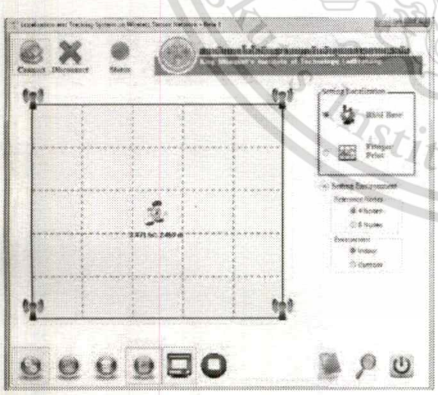


Figure 3. The GUI for The Experiment

As shown in Fig. 3, the GUI displays the appearance of the observed area. The 4 reference nodes are shown as the antenna signs which are located at the corners of the observed area. The target node is represented as the small robot with the letter that represents the location coordinate. Fig. 4 shows the corridor where the experiments are conducted.

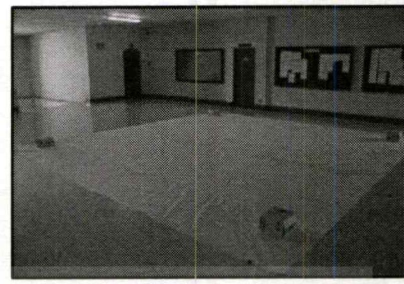


Figure 4. The indoor location for data collection

B. Experiment Setup

In order to verify the accuracy between range-based and fingerprint-based localization, the location of the target node is divided into two types, stationary and moving. Fig. 5 shows the target position for experiments. The 16 locations represented by a small circle along the diagonal line of the target node are used for validating the range-based and fingerprint-based localization techniques. Fig. 6 depicts the 36 database locations of sensor nodes as database information, represented as circles for fingerprint technique experiments.

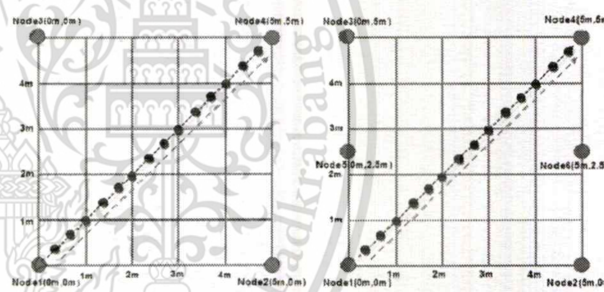


Figure 5. The left diagonal position of the target node for the experiments

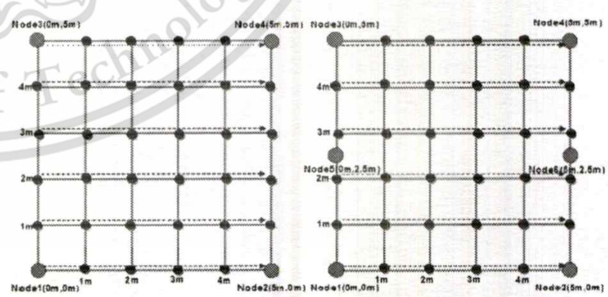


Figure 6. The 36 database locations for the fingerprint technique

IV. Experiment Result And Discussion

The objective of this paper is to compare the accuracy between the range-based and fingerprint-based localization techniques for the stationary and moving target. The effectiveness of the both techniques is verified by an indoor experiment and the estimated location errors are analyzed. The GUI is created to help the experiments in the visualization and for showing the estimated location.

A. Estimated Location Error for Range-based Technique

1. Stationary Target Node

For the stationary target node, we start placing the target node at (0 m, 0 m) and increase continuously on 0.33 m. For instance, the target node location for the diagonal left bottom to right up is change from (0 m, 0 m) and the experiment is continued until reach at (5 m, 5 m). Fig. 7 shows the estimated location error in meter of the range-based technique versus the position of the target node. It depicts the results for the 4 and 6 reference nodes, respectively.

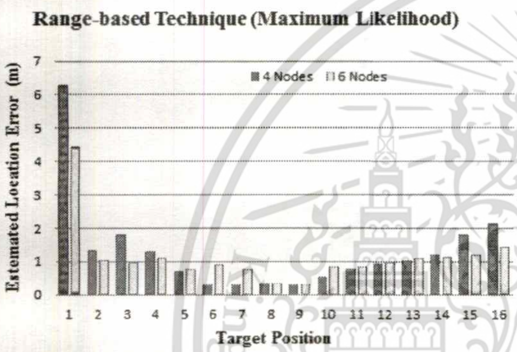


Figure 7. The estimated location error of range-based technique

Fig. 7 shows that the most of the estimated location errors is reduced as a function of the sets of reference nodes. The highest error is occurred in the 1st of the target position. From Fig. 7, we can see that the relatively high of the estimated location errors are in the target position where near or exactly in the position of the reference node. This result could be caused by the issue of diffractions at edges.

2. Moving Target Node

For the result of the moving target, we will show the result in the GUI appearance. For both algorithms, we use automatic robot which carries the target node and moves it alongside the diagonal line. Fig. 8 and Fig. 9 show the results for both sets of reference nodes for the maximum likelihood algorithm. According to the result of Fig.7, the set of 6 reference nodes gives the better result for estimating the location of target node. The GUI shows the distribution of the target node, represented as the small robot is spread along the diagonal line.

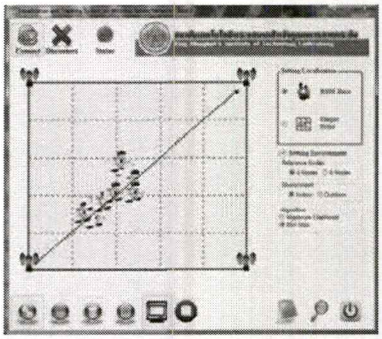


Figure 8. Result of Moving Target for 4 Reference Nodes in GUI using range-based technique

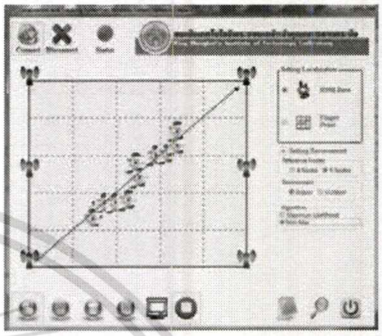


Figure 9. Result of Moving Target for 6 Reference Nodes in GUI using range-based technique

B. Estimated Location Error for Fingerprint Technique

1. Stationary Target Node

The fingerprint information for the experiment is a collection of the RSSI data which is obtained by placing the nodes in the 36 database locations as shown in Fig. 6. As mentioned previously, the main idea of this technique is to compare the target data with the database. Fig. 10 depicts estimated location errors in meter of the fingerprint technique versus the position of the target node.

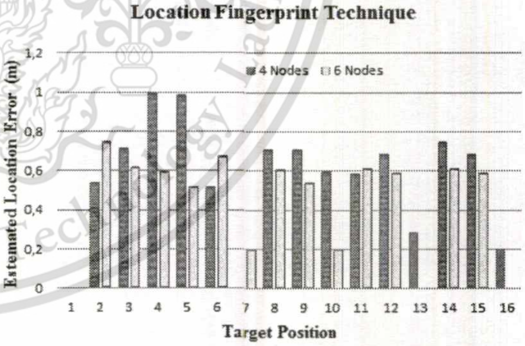


Figure 10. The estimated location error of fingerprint technique

Fig. 10 depicts that the estimated location errors in fingerprint technique relatively smaller than the range-based error. From Fig. 10, we can see that the highest error is occurred in the 4th and 5th of target position. Technically speaking, the sets of the reference node can significantly reduce the estimated location error. As in range-based, the fingerprint technique gives better result when there are 6 nodes as the sets of the references nodes.

2. Moving Target Node

Fig. 11 and Fig. 12 show the result for both sets of reference nodes for location fingerprint algorithms. According to the result of Fig. 10, the set of 6 reference nodes gives the better result for estimating the location of the target node. The GUI shows the distribution of the target node, represented as the small robot is spread along the diagonal line.

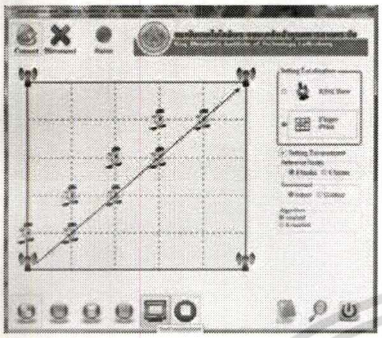


Figure 11. Result of Moving Target for 4 Reference Nodes in GUI using fingerprint-based technique

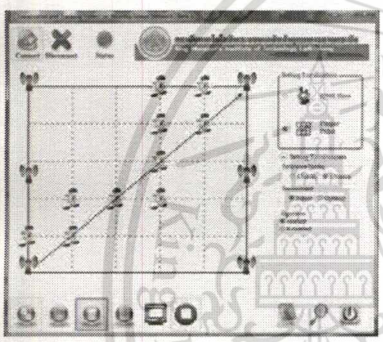


Figure 12. . Result of Moving Target for 6 Reference Nodes in GUI using fingerprint-based technique

C. The Average Error Between Range-based and Location Fingerprint Techniques for the Stationary Target

The different amount of reference nodes is used as space diversity, it affects to the accuracy of the localization system. For the average estimated error, we only use the data from the result of stationary target measurement. From Table 2, we can see that the location fingerprint gives the better result for the experiment in the left diagonal position testing.

As we can see, the error from both techniques is acceptable, since the error is 1.31 m in the range based using 4 reference nodes and 1.12 m for using 6 reference nodes. For the location fingerprint, the average estimated error reach 0.56 m for 4 reference nodes and 0.44 for 6 reference nodes. We can see from Table 2, the additional amount of reference nodes gives better result for both techniques, respectively. From this result, we can get a conclusion that for the diagonal position of the target node, fingerprint technique gives the better accuracy.

TABLE II. THE AVERAGE ERROR TABLE FOR STATIONARY TARGET

Localization Technique	Estimated Location Error (m)	
	4 Nodes	6 Nodes
Maximum Likelihood	1.31	1.12
Location Fingerprint	0.56	0.44

V. CONCLUSION

This paper proposes the method for comparing the range-based and location fingerprint techniques in indoor scenario. The target node is moved along the left diagonal position. The results show the average error reached less than 1.2 m in the range-based technique and less than 0.5 m in the location fingerprint technique. From the result of the estimated locations error, the location fingerprint technique gives the better accuracy than the range-based localization. Since the target node is moving, the error can be triggered by many factors, such as the accuracy of RSSI measurement in the devices, also the power and electromagnetic effects during the experiment. The result of both stationary and moving targets, show that the simple but effective system is obtained. The GUI for the experiments shows the estimated location for both techniques effectively in real time. This result gives the explanation that the proposed method can be applied in the real application.

VI. ACKNOWLEDGMENT

This research work is financial supported by the Thailand Research Fund, Thai Network Information Center Foundation and AUN/SEED-Net.

REFERENCES

[1] I.F. Akyldiz, W. Su, Y. Sankarasubranamniam and E. Cayirci, "A Survey on Sensor Networks," *IEEE Communications Magazine*, pp. 102–114, August 2002.

[2] A.T. Parameswaran, M. I. Husain, and S.Upadhyaya, "Is RSSI a Reliable Parameter in Sensor Localization Algorithms – an Experimental Study," *Field Failure Data Analysis Workshop (F2DA 2009)*, New York, USA, 2009.

[3] R. Grossmann, J. Blumenthal, F. Golatowski and D. Timmermann, "Localization in ZigBee-based Sensor Networks," *Technical Report (German)*, University of Rostock, Institute MD, April 2007.

[4] F. Sottile, R. Giannantonio, M. A. Spirito and F. L. Bellifemine, "Design, Deployment and Performance of a Complete Real-time ZigBee Localization System," *Wireless Day 2008 (WD'08)*, 1st IFIP, pp. 1–5, November 2008.

[5] R. Priwgharm, K. Srivilas, and P. Cherntanomwong, "Indoor Localization System Using RSSI measurement in Wireless Sensor Networks Based on ZigBee Standard," *The 3rd National Conference on Information Technology (NCICT 2010), Bangkok,Thailand*, pp. 225–228, October 2010.

[6] A. Taheri, A. Singh, and E. Agu, "Location Fingerprinting on Infrastructure 802.11 Wireless Local Area Networks," *Proceedings of the 29th Annual IEEE International Conference on Local Computer Network (LCN'04)* , vol.00, pp. 676–683, 2004.

[7] P. Cherntanomwong, J. Takada, and H. Tsuji, "Signal Subspace Interpolation from Discrete Measurement Samples in Constructing a Database for Location Fingerprint Technique ," *IEICE Trans. Commun.*, vol.E92-B, no.9, pp. 2922–2930, September 2009.

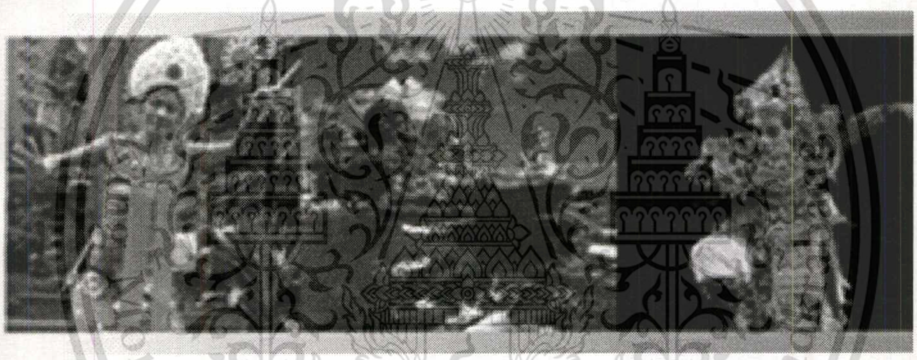
[8] N. Nakarach, T. Bokam, K. Jomprom, D. J. Suroso and P. Cherntanomwong, "Implementation of RSSI-based Localization System Using Wireless Sensor Networks Based on ZigBee Standard," *The 3rd RC ICT 2011*, Lao PDR, March 2011.

INDONESIA SECTION

TENCON 2011

2011 IEEE REGION 10 CONFERENCE

Trends and Development in Converging Technology towards 2020



November 21-24, 2011
Bali, Indonesia

http://www.tencon2011.com/

TENCON 2011

TENCON 2011

2011 IEEE REGION 10 CONFERENCE

21-24 November 2011

Bali, Indonesia

COPYRIGHT AND REPRINT PERMISSION :

Papers are printed as received from the authors.

All opinions expressed in the Proceedings are those of the authors and are not binding on the Institute of electrical and Electronics Engineers, Inc.

IEEE Catalog Number : CFP11TEN-ART

ISBN : 978-1-4577-0255-6

Editors : Wahidin Wahab, Abdul Muis and Mia Rizkinia

Publisher : Electrical Engineering Department, Universitas Indonesia
Depok, West Java 16424 INDONESIA

Location Fingerprint Technique using Fuzzy C-Means Clustering Algorithm for Indoor Localization

Dwi Joko Suroso
International College, KMITL
Bangkok, Thailand
s3601152@kmitl.ac.th

Panarat Cherntanomwong and Pitikhate Sooraksa
Dept. of Computer Engineering
Faculty of Engineering, KMITL
Bangkok, Thailand
{krpanara, kspitikh}@kmitl.ac.th

Jun-ichi Takada
Graduate School of Science and Engineering
Tokyo Institute of Technology
Tokyo, Japan
takada@ide.titech.ac.jp

***Abstract*— The recent researches in localization technique have been supported by the emerging of wireless sensor network (WSN) technology. The issues of power and time consumption have become the main research topics in WSN-based localization technique. ZigBee as IEEE 802.15.4 is commonly used as supporting device because of its advantages for low-power, small and smart sensor nodes. This paper proposes the new technique in radio frequency (RF) fingerprint technique-based localization using Fuzzy C-Means (FCM) clustering algorithm. This technique provides an efficient localization system that gives benefit in the time-efficient and low power consumption. In this paper, received signal strength indicator (RSSI) is used as the fingerprint information which indicates the location of sensor nodes. The different amount of the reference nodes is applied. The effectiveness of this method is verified by an indoor experiment. The estimated location results from different sets of reference nodes are compared. The time consumption in experiment is compared with those using the common fingerprint technique.**

Keywords-localization; wireless sensor network; RSSI; fingerprint technique; fuzzy c-means

I. INTRODUCTION

Localization technique is an important challenge in wireless sensor networks (WSNs). Localization usually refers to the process of dynamically determining the location(s) of one or more target node(s). The various techniques are implemented during the development of localization system. Location fingerprint technique as pattern matching-based localization is commonly used as non-real time localization technique. This technique allows us to have the database training in order to estimate the target location by comparing it with the database training. The disadvantages of this technique are time consuming and requiring a lot of effort during the process of collecting data [1]. This paper proposes a solution to the problems using clustering data technique. The Fuzzy c-means (FCM) algorithm is applied. By applying the FCM method, a partition of the feature vectors into different regions can be performed. The improvement of accuracy can be achieved while the RSSI values as database are clustered into vectors that have similarity among others. The clustering technique is also well-known technique with the spatial advantages related to scalability and efficient communication [2]. In this research, we focus on a localization system using ZigBee standard as the sensor nodes and the target node. The location fingerprint technique is used to estimate the target node location. We present the approach in the clustering technique as the new method for measuring the target location in the fingerprint technique. This paper is organized as follows. Section 2 explains the description of the algorithm model. Section 3 describes the experiment system and setup. Section 4 shows the results and gives discussion. Finally, the conclusion is given in Section 5.

II. DESCRIPTION OF THE ALGORITHM MODEL

A. Fingerprint Localization Technique

Location fingerprinting is a location sensing technique which involves a two-phase process. First, during the off-line calibration phase the received signals at selected locations are recorded in a database. Then, the second phase, called the on-line, pattern matching algorithms are used to infer a target's location by comparing the current observed signal features to the pre-recorded values in the database. The key points of this technique are the selection of the spatial signature and the method in constructing database. The commonly used of pattern matching algorithm as shown in Fig. 1 [1], [3].

FINGERPRINT DATABASE

Known Location	**Fingerprint Information**
Location 1	Information 1
Location 2	Information 2
....	....
Location L-1	Information L-1
Location L	Information L

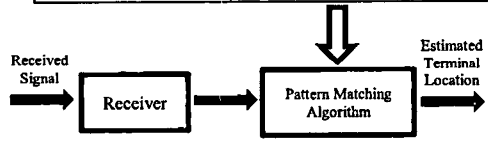


Figure 1. Location Fingerprint Technique

For instance, let the number of reference nodes as N, the number of fingerprint locations as M and D_k at position k $(1,2,\ldots,M)$ as fingerprint (a set of RSSI values for N nodes). Table I shows the illustration of fingerprint database for the experiment.

TABLE I. FINGERPRINT DATABASE

Fingerprint Location	Fingerprint Information
$F_1(x_1,y_1)$	$D_1 = \{rssi_{1,F_1}, rssi_{2,F_1}, \ldots. rssi_{N,F_1}\}$
$F_2(x_2,y_2)$	$D_2 = \{rssi_{1,F_2}, rssi_{2,F_2}, \ldots. rssi_{N,F_2}\}$
......	......
$F_M(x_M,y_M)$	$D_M = \{rssi_{1,F_M}, rssi_{2,F_M}, \ldots. rssi_{N,F_M}\}$

B. Fuzzy C-Means Clustering Algorithm

FCM is the method for data clustering which allows one piece of data belongs to two or more clusters (groups). The advantages of the FCM are its ability of uncertainty data modeling, its fairly robust behavior, its straight forward implementation, and its applicability to multi-channel data. FCM can be applied to several problems involving feature analysis, clustering and classifier design in fields such as astronomy, chemistry, geology, image analysis, target recognition, image segmentation, medical diagnosis and shape analysis [4].

The FCM algorithm minimizes the objective function for the partition of the data set, $P = [D_1, D_2, D_3, \ldots, D_M]^T$ given by:

$$J_m(u,v) = \sum_{i=1}^{c}\sum_{k=1}^{M} u_{i,k}^{m}\|D_k - v_i\|^2 \quad (1)$$

where, J_m is the objective function, M is the number of samples in the vector P, c is the number of clusters $(1 \le c \le M)$, u_{ik} is the element of partition matrix U of size $(c \times M)$ containing the membership function, v_i is the center of the i^{th} cluster, and m is a weighting factor that controls fuzziness of the membership function. The matrix U is constrained to contain elements in the range of $[0,1]$ such that $\sum_{i=1}^{c} u_{ik} = 1$ for each $u_{ik}(1 \le k \le M)$. The norm $\|D_k - v_i\|$ is the distance between the sample D_k and the clusters center v_i [5].

C. Nearest Neighbour Algorithm

The final step to estimate the location is to find the minimum error distance between the target data compared with those in the database. FCM clusters the group of fingerprint information that has similarity, including the target node. When the target node is in the fingerprint area, it collects the RSSIs from all reference nodes. N is the number of reference nodes. The RSSI value can be expressed as

$$T = \{rssi_{1,T}, rssi_{2,T}, \ldots. rssi_{N,T}\}. \quad (2)$$

The cluster or group of the sensor nodes, G, as a result of the FCM algorithm is located in the fingerprint database. Euclidean distance is used to find the location of the target node by comparing the set of RSSIs of each G member, g_i and the target node. Then, the similarity between the RSSIs of the target node and those of the C member can be calculated by

$$dist_i(x,y) = \sqrt{\sum_{n=1}^{N}\left(rssi_{n,T} - rssi_{n,g_i}\right)^2}. \quad (3)$$

By using the nearest neighbor algorithm, the fingerprint location in which its fingerprint provides the smallest distance, $dist_i$ is returned as the target node location.

III. EXPERIMENT SYSTEM AND SETUP

A. Sensor Node Placement

Fig. 2 shows the layout of references nodes and fingerprint nodes placement in the system. Fingerprint nodes, shown as the rectangular are assumed to know their own position and reference nodes, shown as the triangle symbols are located at the corners of the experiment area. The database for fingerprint technique is obtained on 36 observed location in area of 5m × 5m. in the experiment, the 1st sensor nodes is located in the left bottom of observer area. The arrangement is left to right arrangement. For example, the 1st row is consist of sensor nodes number 1 to 6, in 2nd row 7 to 12 and continues until 6th row, senor node number 31 to 36.

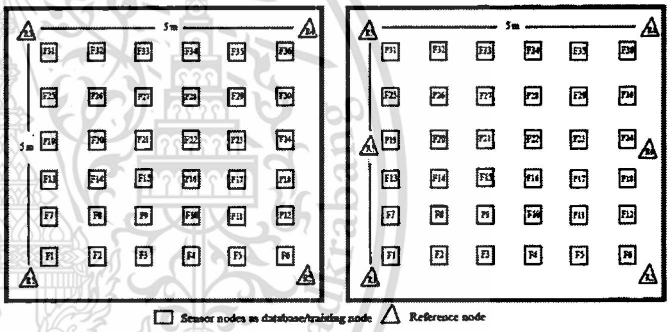


Figure 2. Reference and sensor nodes placement for experiments

B. Implementation of Fuzzy C-Means Algorithm

FCM algorithm is used to cluster the 36 sensor nodes and 1 target node as test object for identifying the target location. In this paper, the data set, P is formed as 36×4 matrix for 4 reference nodes and 36×6 matrix for 6 reference nodes. FCM clustering technique can be summarized by the following steps [5]:

Step 1: Initialization (Iteration 0)

Scan the RSSIs data row by row to construct the vector P containing all the certain level in the data set. Randomly initialize the centers of the classes vector $V^{(0)}$. From the iteration $t = 1$ to the end of algorithm:

Step 2: Calculate the membership matrix $U^{(t)}$ of element u_{ik} using:

$$u_{ik} = \left(\sum_{j=1}^{c}\left(\frac{\|D_k - v_i\|}{\|D_k - v_j\|}\right)^{\frac{2}{m-1}}\right)^{-1} \quad (4)$$

v_i is the cluster center in the iteration, $t = 0$ (the super script, t is the iteration index). Then, we can get the v_1, v_2 until v_c for the next step.

Step 3: Calculate the vector $V^{(t)} = [v_1, v_2, \dots v_c]$ using:

$$v_i = \frac{\sum_{k=1}^{d} u_{ik}^{m} p_k}{\sum_{k=1}^{d} u_{ik}^{m}} \tag{5}$$

Step 4: Convergence test: if $\|V^{(t)} - V^{(t-1)}\| > \epsilon$, then increment the iteration t, and return to the **Step 2**, otherwise, stop the algorithm. ϵ is a chosen positive threshold.

C. Experiment Setup

1.) Diagonal Position Testing

The first experiment to obtain the validation of our proposed method, the diagonal position testing is applied. Fig. 3 shows the fingerprint nodes and the target node at diagonal positions on the observed area for 4 and 6 reference nodes. The target node is stationary for each experiment.

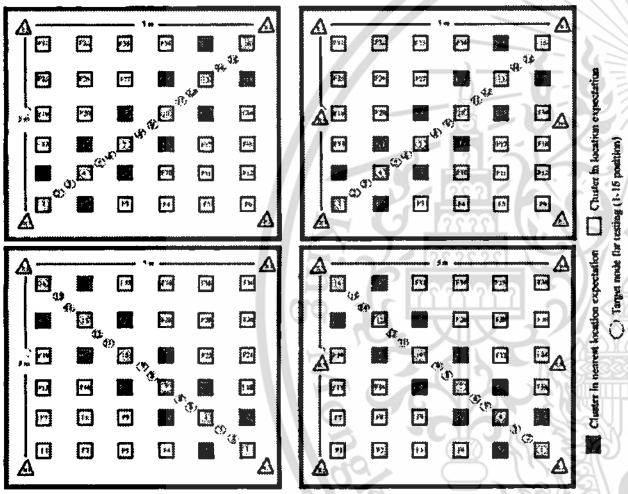


Figure 3. The 16-target node position for diagonal testing

For the left diagonal testing, we start placing the target node at (0 m, 0 m) and increase continuously on 0.33 m. For instance, the target node location for the diagonal left bottom to right up is change from (0 m, 0 m) and the experiment is continued until reach at (5 m, 5 m). For the right to the left diagonal position we change the target node location from coordinate (5 m, 0 m) to (0 m, 5 m).

2.) Horizontal and Vertical Position Testing

The other types of position are horizontal and vertical position. Fig. 4 shows the measurement setup for horizontal and vertical of target node testing position. For the horizontal and vertical position testing, we start placing the target node from (0 m, 2.5 m) for the horizontal position; with the increasing distance are 0.25 m and 0.5 m. For example, position number 1 is (0 m, 2.5 m), number 2 (0.25 m, 2.5 m), number 3 (0.75 m, 2.5 m), and number 4 (1 m, 2.5 m). The experiment continues until (5 m, 2.5 m). For the vertical position, it is began at (2.5 m, 0 m) and ended at (2.5 m, 5 m).

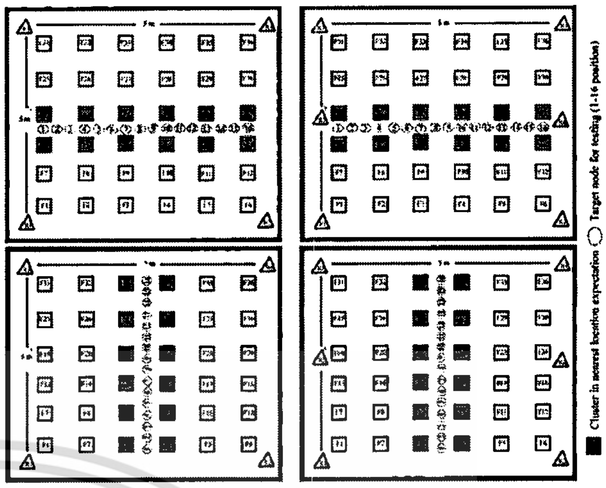


Figure 4. The 16-target node position for horizontal and vertical testing

IV. Result and Discussion

We conduct the experiment to analyze the accuracy of target location estimation for the proposed method. The accuracy of the proposed system with FCM algorithm is found by comparing the estimated position with the training database. All of these measurements were performed in the corridor at Department of Computer Engineering, KMITL. The clustering result will be explained both sets of 4 and 6 reference nodes, respectively. The experiment in each position is done using repeated five times measurement.

A. Cluster Result

At the experiment setup, we can see the location of target node in the observed area. We also see the expect cluster in each target node position (1-16 position). For the experiments, we use value c on the FCM clustering algorithm is equal to 10 clusters or groups. As mentioned previously, the c value must not exceed the number of sample, M in the vector P. So, $c=10$ is acceptable since the number of sample is 37 (36 sensor nodes for database and 1 for target node). Numbers in the bracket, {1}-{36} are the fingerprint database and {37} is the target node.

1) The sets of four reference nodes

a) Cluster result for diagonal left

TABLE II. Cluster Result for Diagonal Left Position

No.	Desired Cluster	Actual Cluster
1	{1},{37}	{1},{37}
3	{1},{2},{7},{8},{37}	{1},{2},{7},{8},{37}
8	{15},{37}	{15},{22},{37}
10	{22},{37}	{22},{37}
14	{29},{30},{35},{36},{37}	{29},{30},{35},{36},{37}
16	{36},{37}	{36},{37}

The result for the left diagonal position for 4 reference nodes shows that the FCM can cluster the target node successfully. The target node belongs to the certain cluster in the database.

b) Cluster result for diagonal right

TABLE III. CLUSTER RESULT FOR DIAGONAL RIGHT POSITION

No.	Desired Cluster	Actual Cluster
1	{6},{37}	{5},{6},{11},{12},{37}
3	{5},{6},{11},{12},{37}	{5},{6},{11},{12},{37}
8	{15},{16},{21},{22},{37}	{15},{22},{28},{37}
10	{20},{37}	{9},{15},{37}
14	{25},{26},{31},{32},{37}	{25},{31},{32},{37}
16	{31},{37}	{31},{37}

c) Cluster result for horizontal

TABLE IV. CLUSTER RESULT FOR HORIZONTAL POSITION

No.	Desired Cluster	Actual Cluster
1	{13},{14},{19},{20},{37}	{13},{14},{19},{20},{37}
3	{13},{14},{19},{20},{37}	{13},{14},{19},{20},{37}
8	{15},{16},{21},{22},{37}	{15},{37}
10	{15},{16},{21},{22},{37}	{15},{22},{37}
14	{17},{18},{23},{24},{37}	{17},{23},{24},{37}
16	{17},{18},{23},{24},{37}	{17},{23},{24},{37}

d) Cluster result for vertical

TABLE V. CLUSTER RESULT FOR VERTICAL POSITION

No.	Desired Cluster	Actual Cluster
1	{3},{4},{9},{10},{37}	{3},{9},{37}
3	{3},{4},{9},{10},{37}	{4},{9},{10},{16},{37}
8	{15},{16},{21},{22},{37}	{15},{22},{37}
10	{15},{16},{21},{22},{37}	{15},{22},{37}
14	{27},{28},{33},{34},{37}	{27},{28},{34} ,{37}
16	{27},{28},{33},{34},{37}	{27},{28},{33},{34},{37}

2) The sets of six reference nodes

a) Cluster result for diagonal left

TABLE VI. CLUSTER RESULT FOR DIAGONAL LEFT POSITION

No.	Desired Cluster	Actual Cluster
1	{1},{37}	{1},{37}
3	{1},{2},{7},{8},{37}	{1},{2},{7},{8},{37}
8	{15},{16},{21},{22},{37}	{21},{37}
10	{22},{37}	{16},{22},{23},{37}
14	{29},{30},{35},{36},{37}	{29},{30},{35},{36},{37}
16	{36},{37}	{36},{37}

b) Cluster result for diagonal right

TABLE VII. CLUSTER RESULT FOR DIAGONAL RIGHT POSITION

No.	Desired Cluster	Actual Cluster
1	{6},{37}	{5},{6},{37}
3	{5},{6},{11},{12},{37}	{5},{6},{11},{12},{37}
8	{15},{16},{21},{22},{37}	{15},{21},{37}
10	{20},{37}	{21},{37}
14	{25},{26},{31},{32},{37}	{25},{31},{32},{37}
16	{31},{37}	{31},{32},{37}

We can see for both results, the sets of four and six reference nodes, in the right diagonal position. The yellow marker shows the bad result for clustering the data. This result is affected by bad database training in certain location. Since the accuracy of this technique is depend on the quality of the database [1].

c) Cluster result for horizontal

TABLE VIII. CLUSTER RESULT FOR HORIZONTAL POSITION

No.	Desired Cluster	Actual Cluster
1	{13},{14},{19},{20},{37}	{14},{19},{20},{37}
3	{13},{14},{19},{20},{37}	{14},{19},{20},{25},{37}
8	{15},{16},{21},{22},{37}	{16},{21},{22},{23},{37}
10	{15},{16},{21},{22},{37}	{16},{21},{22},{37}
14	{17},{18},{23},{24},{37}	{17},{18},{ 24},{37}
16	{17},{18},{23},{24},{37}	{17},{18},{ 24},{37}

d) Cluster result for vertical

TABLE IX. CLUSTER RESULT FOR VERTICAL POSITION

No.	Desired Cluster	Actual Cluster
1	{3},{4},{9},{10},{37}	{3},{9},{15},{37}
3	{3},{4},{9},{10},{37}	{3},{9},{15},{37}
8	{15},{16},{21},{22} ,{37}	{16},{22} ,{37}
10	{15},{16},{21},{22},{37}	{21},{37}
14	{27},{28},{33},{34},{37}	{28},{34},{37}
16	{27},{28},{33},{34},{37}	{27},{28},{33},{34},{37}

B. Location Estimation Error

In this subsection, the comparison between the different sets of reference node is analyzed. Fig. 5 depicts the left diagonal estimated location error between four and six reference nodes.

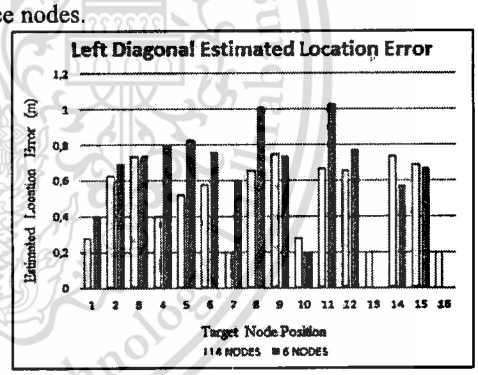


Figure 5. Left diagonal estimated location error for 4 and 6 reference nodes

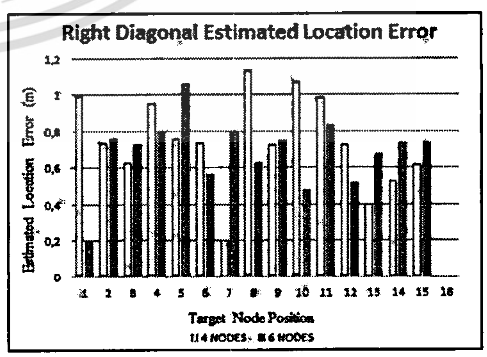


Figure 6. Right diagonal estimated location error for 4 and 6 reference nodes

In Fig. 5, the highest error occurred in the target position number 11. With the error reached 1.03 m for the set of six reference nodes. We can see that seven positions in the target node position testing, the set of six reference nodes gives a larger error than the set of four references nodes. The average error for the set of four reference nodes is 0.512 m and 0.613 m for the set of six reference nodes. In this position testing, the additional reference nodes did not give the better accuracy.

The different results occured in the right diagonal estimated location error. We can see from Fig. 6, most of the errors in the system were reduced by the set of six references nodes. For instance, the errors reduction are shown in the target node number 1, 4, 8, 10,11 and 12. In the right diagonal position testing, the average of the estimated location errors are 0.70 m for set of four references nodes and 0.64 m for the set of six reference nodes.

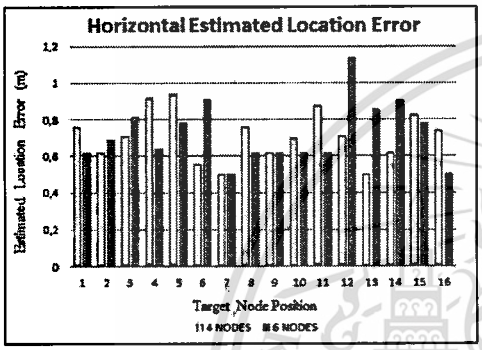


Figure 7. Horizontal estimated location error for 4 and 6 reference nodes

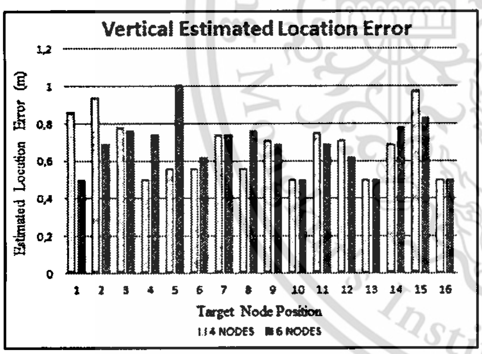


Figure 8. Vertical estimated location error for 4 and 6 reference nodes

Fig. 7 and 8 depict the horizontal and vertical estimated location error, respectively. In the horizontal and vertical position testing, the estimated location error between the set of four and six references nodes are relatively the same values in average. The average error in horizontal position is 0.71 m for four reference nodes and 0.72 m for six references nodes. For vertical position, the average estimated location for the set of four reference nodes is 0.67 m and 0.68 m for six reference nodes. As note, the errors have the relatively same pattern; we can see in Fig 5 and 6, high errors occurred in the target position number 5, 8 and 11. Fig. 7 and 8, the highest error occurred in the target node number 5.

If we see in section III.C, most of the errors occurred in the middle part of the observed area. The important thing we notice for this result, we can reduce the error by fixing or improving the quality of database in the middle part of observed area. Moreover, if we can obtain the new data for all fingerprint database. As we know, the accuracy of location fingerprint-based localization technique depends on the quality of the fingerprint database [1].

C. Comparison of Techniques

It is important to show the comparison between the common fingerprint and the proposed method. In this subsection, we will emphasize in the effectiveness of clustering algorithm to reduce the power and the time consumption. The most obvious means of power consumption is to turn transceivers off when it is not required. The common technique consumes longer time since it uses all the sensor nodes to estimate the target location. For the proposed technique, it uses only the sensor nodes in the certain cluster. It means that it can isolate the cluster that target node is belong to. So, the on-line phase to estimate the location is applied in the certain cluster; the other clusters are turned off. It shows that the proposed method can be applied in the reducing the power and the time consumption.

V. Conclusion and Future Works

This paper presents the new method for clustering data in the fingerprint localization technique using FCM algorithm. This technique is validated using 36 database locations fingerprint. The results show the average error reached less than 1 m, both for 4 and 6 reference nodes. The additional numbers of reference nodes did not give the better accuracy for estimating the location of the target node. The results for clustering show that FCM can cluster the target node in the fingerprint database. We also see some errors in the clustering result for right diagonal position, both for 4 and 6 reference nodes. The errors are caused by the not good enough RSSIs values data for the database. The future research, we will improve the quality of data for the database. From this result, we can get conclusion that new proposed method can be applied in the localization system. Furthermore, the system can be solution in issues such as high power consumption and time-efficient WSN-based localization system.

References

[1] P. Cherntanomwong, J. Takada, and H. Tsuji, "Signal Subspace Interpolation from Discrete Measurement Samples in Constructing a Database for Location Fingerprint Technique ," IEICE Trans. Commun., vol.E92-B, no.9, pp. 2922–2930, September 2009.

[2] D. C. Hoang, R. Kumar and S. K. Panda, "Fuzzy C-Means Clustering Protocol for Wireless Sensor Networks," IEEE International Symposium on Industrial Electronics (ISIE-2010), Bari, Italy, 2010.

[3] A. Taheri, A. Singh, and E. Agu, "Location Fingerprinting on Infrastructure 802.11 Wireless Local Area Networks," Proceedings of the 29th Annual IEEE International Conference on Local Computer Network (LCN'04) , vol.00, pp. 676–683, 2004.

[4] M. Sayadi, L. Tlig, and F. Fnaiech, "A New Texture Segmentation Method Based on the Fuzzy C-Means Algoritm and statistical Features ," Applied Mathematical Sciences, vol.1, no.60, pp. 2999–3007, 2007.

[5] J. C. Bezdek, Pattern recognition with fuzzy objective function algorithms, Plenum Press: New York, 1981.