International Journal of Innovative
Computing, Information and Control
Volume **18**, Number **1**, February **2022**

ICIC International ©2022 ISSN 1349-4198
pp. 231–252

AUTOMATIC DATA ACQUISITION SYSTEM FOR WI-FI FINGERPRINT-BASED INDOOR POSITIONING SYSTEM

Alvin-Ming-Song Chong, Boon-Chin Yeo and Way-Soong Lim

Faculty of Engineering and Technology
Multimedia University
Jalan Ayer Keroh Lama, 75450 Melaka, Malaysia
1142700757@student.mmu.edu.my; { bcyeo; wslim }@mmu.edu.my

Received April 2021; revised August 2021

Abstract. *Indoor Positioning System (IPS) can be developed by using different wireless technologies with different positioning techniques. One of the popular areas of study is the Wi-Fi fingerprinting based IPS because it has high flexibility. However, the implementation of a fingerprint based IPS required obtaining a fingerprint database beforehand. Manually performing data acquisition to obtain a fingerprint database requires a significant amount of time and workforce. Thus, this paper presented a method to perform data acquisition automatically by utilizing the tiled floor characteristic. Furthermore, the fingerprint database constructed was utilized to perform studies on the impact of fingerprint patterns and errors in the location tag of fingerprints on the positioning accuracy of neural networks. Studies show that neural network models presented smaller positioning errors when fingerprints used are with environmental reading formed by RSS from more Wi-Fi anchors. Studies also show that when the location tag of fingerprints used for neural network models training consists of errors, positioning accuracy when validated with fingerprints in the validation set (errorless) depends on the type of error. The location tag of fingerprints used for training suffering large mean errors will significantly impact validation positioning accuracy compared with large standard deviation errors.*
Keywords: Indoor Positioning System (IPS), Non-Line of Sight (NLOS), Line of Sight (LOS), Fingerprint, Received Signal Strength (RSS), Tiled floor

1. **Introduction.** Navigation systems, such as Waze and Google Maps, are commonly used to locate the position of the user and to navigate the user to the desired destination. The technology is extremely convenient for the users when they are unfamiliar with the places of interest. These navigation systems generally work based on the well-known Global Positioning System (GPS), which the United States Air Force first uses as a space-based radio navigation system [1]. Even though GPS is a sophisticated positioning system in the outdoor environment, many research works have pointed out the limitations of GPS to work in the indoor environment [2]. Thus, many researchers have started researching alternative solutions by developing positioning systems specially for indoor purposes. These positioning systems are called Indoor Positioning System (IPS). IPS can be developed by using different kinds of technologies. Examples of technology are image processing, Inertial Measurement Unit (IMU), and wireless technology. This paper focuses on the IPS developed using Wi-Fi, a wireless technology that has been attracting the attention of many researchers in the past until now. When Wi-Fi is the wireless technology selected to develop an IPS, fingerprinting positioning technique with Received Signal Strength (RSS) from multiple Wi-Fi anchors as environmental reading is commonly adopted to estimate target position. However, fingerprinting positioning technique has a well-known disadvantage:

DOI: 10.24507/ijicic.18.01.231

the need of obtaining a fingerprint database beforehand. A significant amount of time and workforce is required for conducting data acquisition to obtain a fingerprint database of a study environment. Thus, this paper will be presenting a method proposed to obtain the fingerprint database (by automating the data acquisition process) for a selected study environment. In addition, two experiments will be conducted using the fingerprint database constructed by the method proposed. The first experiment will study the positioning performance of IPS developed when fingerprints used are with environmental reading formed by RSS from different types of Wi-Fi anchors or different numbers of Wi-Fi anchors. The meaning of RSS from different types of Wi-Fi anchors is referring to whether the Wi-Fi anchor is in Line of Sight (LOS) or Non-Line of Sight (NLOS) [3] condition with tag node (Automatic Surveyor). The second experiment will study the impact on positioning performance when the location tag of fingerprints in the training set contains errors.

The remainder of this paper is arranged as follows. First and foremost, this paper begins by giving a brief understanding to readers on wireless technology based IPS. Next, the advantages and disadvantages of Wi-Fi, a wireless technology frequently selected by researchers to develop an IPS, will be discussed. Furthermore, the working principle of a positioning technique (fingerprinting positioning technique) commonly used to develop Wi-Fi based IPS will be explained. In addition, different surveyor techniques used by different researchers to obtain a fingerprint database required when fingerprinting positioning technique is being adopted will be presented. Besides, a method adopted to develop a Wi-Fi fingerprinting based IPS using a neural network as the position algorithm will be explained. This includes an explanation of a method proposed to conduct automatic data acquisition to obtain a fingerprint database for a selected study environment. Furthermore, the procedure to carry out the two experiments mentioned will be described. Next, the results of each experiment will be discussed. Last but not least, a conclusion that concludes this paper is presented.

2. **Wireless Technology Based IPS.** Over the past decades, researchers had developed IPS using different kinds of wireless technologies. Examples of popular wireless technology being used are Wi-Fi, Bluetooth [4-6], and Ultra-Wide Band (UWB) [7-9]. Although IPS developed by using different wireless technologies might be using different positioning techniques, wireless technology based IPS is commonly formed with two types of nodes: tag node and anchor node. A tag node is a node that is unable to know its position at first. It is a node that position is to be estimated by a positioning algorithm. This type of node is sometimes called a dumb node, free node, unknown node, or target node. Meanwhile, the anchor node is a node that knows its position without using a positioning algorithm. It is a static node that knows its position at the very beginning during the node installation. Besides, it is a node that assists a positioning system to locate the tag node position. Sometimes this node is also known as landmark node, reference node, or beacon node.

IPS developed using different wireless technologies as positioning medium will benefit developers from different aspects and face different issues. For example, UWB based IPS will benefit the developer in terms that IPS developed can provide centimetre level accuracy. However, UWB based IPS will induce high development costs, and most of the portable user devices are not built-in with a UWB module. The wireless technology used can result in IPS developed with high accuracy. However, accuracy is not the only measuring criterion for developers of IPS. This is because different applications/services require IPS with different levels of accuracy. Another high concern criterion is development cost. IPS with low development cost and sufficient accuracy is a more practical choice, and Wi-Fi suits the statement well.

2.1. **Wi-Fi.** Wi-Fi is a popular wireless technology chosen by researchers to develop IPS in the past decade [10-12]. The reason is because of its several high-impact advantages. First and foremost, Wi-Fi has a great infrastructure. An IPS developed using Wi-Fi will enjoy a lower development cost [13]. Existing Wi-Fi anchors can be utilized, reducing the need for extra costs on hardware and electrical power. Secondly, devices that can detect Wi-Fi signals are widely available worldwide as people are now in an era that everyone could have a smartphone with them. This reduces the need for extra devices to act as tag nodes. Furthermore, Wi-Fi signals can penetrate through walls and obstacles relatively easier when compared with some other wireless technologies. This made it suitable to develop IPS without considering too much on having an LOS between tag node and anchor node. In addition, Wi-Fi signal coverage is more extensive. The signal coverage can extend up to 1 kilometre, according to the IEEE 802.11ah standard [14]. Although Wi-Fi has many advantages, it also has its disadvantages. One of the major disadvantages of Wi-Fi based IPS is that high positioning accuracy is hard to achieve due to signal fluctuation and interference [15]. When Wi-Fi is chosen to develop IPS, fingerprinting positioning technique with environmental reading formed by RSS is a popular technique used to estimate a target position [12,16,17].

2.2. **Fingerprinting positioning technique.** Fingerprinting is a positioning technique commonly used to construct different wireless technology based IPS [18-20]. It is a flexible positioning technique and promises to provide good accuracy with simple implementation and low cost. Besides, its complexity is low as LOS measurements between nodes and the exact position of installed wireless anchors are not required beforehand. In addition, it has high applicability in complicated indoor conditions [12,21]. Figure 1 illustrates a block diagram that provides a big picture of processes involved in fingerprinting positioning technique using RSS from multiple wireless anchors as environmental reading.

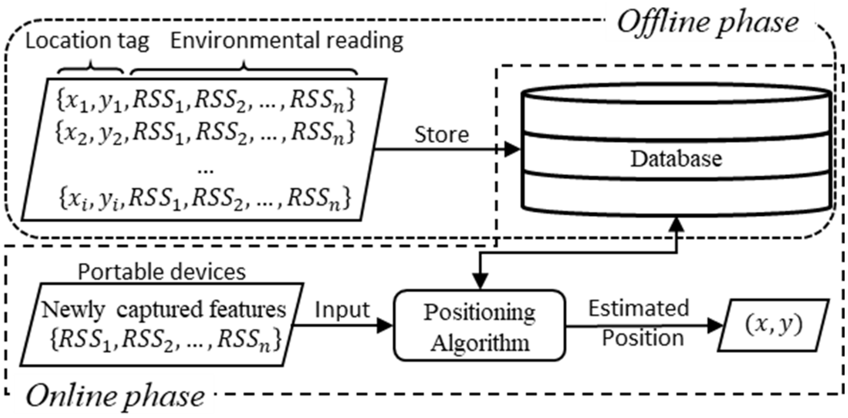


FIGURE 1. Block diagram of fingerprinting positioning technique using RSS from multiple wireless anchors as environmental reading

Fingerprinting positioning technique typically consists of two main phases: offline phase and online phase. During the offline phase, a single objective is to be achieved: obtaining a fingerprint database (signal map). Therefore, data acquisition will be performed in a study environment from one location to another location. Each time data acquisition is performed, a fingerprint will be created and saved in a fingerprint database. Once data acquisition on a study environment is completed, a fingerprint database is obtained. A fingerprint is actually a vector formed by two pieces of information: location tag and environmental reading. Location tag is an address of a location. It depends on how a location in a study environment is being defined at the very beginning. Suppose a study environment is defined using a 2D Cartesian coordinate system. In that case, the location tag will be like the location tag in fingerprints, as shown in Figure 1. The location tag will be the x-coordinate value and y-coordinate value of the ith location when data acquisition is performed. Next, environmental reading is a collection of distance-dependent parameters (features) values obtained at a specific location. It can be formed by multiple types of distance-dependent parameters such as Time of Arrival (TOA), Time Difference of Arrival (TDOA), and Angle of Arrival (AOA). A distance-dependent parameter can be any parameter that value will differ when extracting at a different location. As shown in Figure 1, the environmental reading of a fingerprint is a series of RSS values from n number of wireless anchors obtained at a specific location.

Furthermore, during the online phase, a positioning algorithm will be developed. The positioning algorithm will utilize the fingerprint database constructed to estimate a target position. Firstly, when a tag node reaches a specific position, environmental reading (with the same features as in the offline phase) will be measured. The positioning algorithm will then use the environmental reading as input for performing position estimation. Several positioning algorithms have been used in the past decade. One of the popular positioning algorithms is the k-Nearest Neighbor (k-NN) algorithm [22-24]. When k-NN is adopted as the positioning algorithm, k-NN will take environmental reading obtained in real time as input. It will then compute the distance between the input with other environmental readings in the fingerprint database. Based on distance computed, k number of fingerprints with environmental reading computed having distance closest to the input will be selected. Then, the fingerprint label with the highest votes will become the estimated target position outputted by the k-NN if the k-NN is a classification k-NN. If the k-NN is a regression k-NN, the estimated target position outputted will be the average of all the labels of k number of fingerprints selected.

Besides k-NN, neural network [16,25,26] is another popular positioning algorithm used to develop fingerprinting based IPS. This is because neural networks can recognize patterns available in data (in this case, referring to feature patterns available in the fingerprint database). In addition, neural networks can avoid overfitting to noise available in data. Although fingerprinting positioning technique is easy to implement, less in complexity, and promises to provide high accuracy depending on the quality of the fingerprint database, there are still some disadvantages [27]. For example, fingerprinting positioning technique will cause an increase in the computing load and memory size [28]. Besides, there is another well-known disadvantage: require obtaining a fingerprint database beforehand. The data acquisition process to obtain a fingerprint database requires a large workforce, time-consuming and error-prone [29].

2.3. **Fingerprint database generation techniques.** Researchers had proposed various methods to obtain a fingerprint database (signal map) required by fingerprinting based IPS [29]. In [30], a traditional method is adopted to obtain a fingerprint database required to develop a Wi-Fi RSS based IPS. The traditional method is by manually conducting data acquisition on each grid point in a study environment by a surveyor carrying a device installed with an Android application developed to log Wi-Fi RSS values. This method is time-consuming and requires a large workforce if the study environment selected is large. Besides, the surveyor must remember its current location as the location tag is generated based on the current location inputted by the surveyor.

In order to mitigate issues caused by the traditional method, the authors of [31] proposed a method by developing a sampling device that can automatically generate a location tag for each environmental reading. Location tag is generated automatically according to an estimated location based on a starting point and motion sensor reading. In addition, the sampling device will display the current location of the surveyor on a map request from a venue service. This function made it easier for the user to identify errors in estimated location. If the estimated location is incorrect, an interface is provided for the user to correct the estimated location. This method proposed had reduced the need for users to remember their current location. Besides, this method can experience a shorter data acquisition time compared to the traditional method. The reason is that the need for users to input their current location every time is excluded. However, this method still required the surveyor to conduct data acquisition manually from a location to another location in a study environment to obtain the fingerprint database.

Furthermore, another group of authors in [32] proposed an alternative method to obtain a fingerprint database. The method is by the use of a ray-tracing algorithm to predict wireless signal field intensity information. After that, a few points will be selected to conduct physical manual field measurement to record actual wireless signal strength and use it to adjust parameters in the algorithm. The advantage of this method is rapid signal map generation and requires less workforce. However, theoretical reading can have a significant difference when compared with the actual reading.

Last but not least, many researchers proposed to use an SLAM robot to automate the data acquisition process for obtaining the fingerprint database. However, the location tag generated by the SLAM algorithm faces the problem of cumulative error [33]. Location tags generated with cumulative errors may reduce the quality of the fingerprint database. At the same time, the performance of the IPS developed may be impacted as well. Using SLAM robots to automate the data acquisition process is a good idea, reducing the human workforce. However, the development cost of an SLAM robot is high. The reason is that high accuracy sensors are required to minimize error in position estimated. Thus, this paper will present a method proposed to obtain a fingerprint database by automating the data acquisition process with a robot that estimates its position based on grout lines/joints on a tiled floor.

3. Methodology.

Development of a Wi-Fi RSS fingerprinting based IPS requires starting by obtaining a fingerprint database for a study environment. Thus, hardware preparation for study environment setup is the first step. This section is divided into five subsections. The first subsection is the design synthesis of hardware. All the hardware developed for constructing a Wi-Fi RSS fingerprinting based IPS will be introduced in the subsection. However, only hardware developed for environment setup will be explained in detail. The second subsection will explain hardware developed that made automatic data acquisition with high accuracy location tags automatically generated become possible. The third subsection will present the structure of the study environment will with a floor plan. The fourth subsection will explain the process of automatic data acquisition. Finally, the fifth subsection will explain the procedure to conduct the two experiments mentioned in the introduction.

3.1. Design synthesis of hardware.

Hardware for developing a Wi-Fi RSS fingerprinting based IPS is categorized into hardware developed to serve as data acquisitor and hardware developed for environment setup. Hardware developed to serve as data acquisitors is called Automatic Surveyor. It is a robot developed capable of automatically conducting data acquisition with environmental reading labelled with high accuracy location tag when generating a fingerprint. Although robots had been used to automate data acquisition, the method to generate accurate location tags is still under investigation. Detailed explanation on Automatic Surveyor will be discussed in the following subsection. This subsection will focus on describing hardware developed for environment setup only.

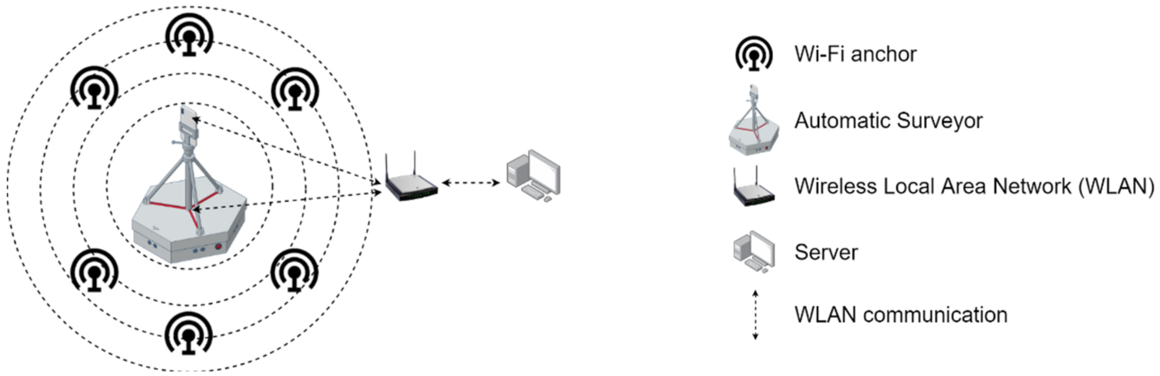


FIGURE 2. Overall environment setup

However, the relationship between all the hardware is illustrated as shown in Figure 2. Excluding the Automatic Surveyor, the remaining hardware shown in Figure 2 is hardware developed for environment setup.

Hardware developed for environment setup is separated into three types: hardware that acts as Wi-Fi anchors, hardware that enables long-range communication between devices (a WLAN), and hardware that provides database service and venue service (a server). First and foremost, a Wi-Fi anchor is hardware that can transmit Wi-Fi signals. Since the IPS to be developed uses fingerprinting positioning technique with RSS from multiple Wi-Fi sources to estimate the position of a tag node, hardware that can transmit Wi-Fi signals is required. The hardware used to serve as a Wi-Fi anchor is NodeMCU. NodeMCU consists of an ESP8266 as the main component providing Wi-Fi service. Reasons for selecting NodeMCU to serve as a Wi-Fi anchor are that it is open-source and Arduino IDE compatible, which will significantly speed up the developing process because libraries are available that make firmware development an easy job. Besides, NodeMCU consists of a build-in PCB antenna to extend its operating range, which as a result is sufficient to cover the whole study environment. Details about the study environment selected will be discussed in a later subsection. In addition, NodeMCU power consumption is low and can easily be powered by a power bank. This made the placement of Wi-Fi anchors in the study environment easier as the Wi-Fi anchors can be set up anywhere without considering the presence of a wall electrical outlet. A total of six NodeMCU will be prepared and separately placed in the study environment. Each NodeMCU will be uploaded with a firmware prepared that configures them to serve as a Wi-Fi anchor. NodeMCU can act in three modes: access point mode, station mode or both modes simultaneously. Since NodeMCU prepared aims to serve as Wi-Fi anchors, firmware that configures them to function in access point mode will be uploaded. While a NodeMCU is configured to function in access point mode, a Wi-Fi scanner can discover its Wi-Fi RSS value. Last but not least, each firmware uploaded to each of the six NodeMCU will also configure them with a unique SSID. The reason is that SSID is used by Wi-Fi tag to identify the source of RSS instead of MAC address for a better explanation in this paper.

Next, a long-range communication platform is required so that each device involved can easily communicate with each other or exchange information. As a result, a Wi-Fi modem, hardware that combines modem and router functionality, was used to establish a WLAN for the study environment. Devices such as hardware that serve as a server, smartphone of Automatic Surveyor and mobile robot of Automatic Surveyor will all be connected to the WLAN established. For communication purposes, each device connected to the WLAN is assigned a static IP address.

Finally, hardware that can provide database service and venue service is required. A computer is configured to function as a server connected to the WLAN established to provide the two services. Database service will store data (fingerprints) received from a client, and venue service will return a survey path to the client requesting a path of survey. A survey path is a file that consists of a series of commands defining the path to be followed by Automatic Surveyor when performing automatic data acquisition. Thus, the survey path needs prepared beforehand and kept on the computer so that the venue service can reply to clients when a request is made.

3.2. **Automatic Surveyor.** Conducting data acquisition to obtain a fingerprint database can be done manually by humans or automated with the help of robots. Suppose data acquisition is conducted manually. The generation of a fingerprint database for a study environment will be time-consuming, a large workforce is required, and environmental reading might be mislabelled. Thus, many researchers had adopted robots to conduct data acquisition automatically to obtain a fingerprint database required. However, the method to automatically generate accurate location tags (for labelling environmental reading obtained to generate a fingerprint) by robots adopted is yet a topic to be investigated. This section will present a robot developed and adopted for conducting data acquisition automatically in a study environment selected. The robot is capable of automatically self-generating accurate location tags using the characteristic of a tiled floor.

Nowadays, the floor of many indoor environments is tiled. Two square tiles placing side by side will form a grout line with a wide of few millimetres. Another two square tiles placing below the previous two tiles will result in two grout lines crossing each other perpendicularly and form a grout joint. Suppose a floor is tiled in a grid pattern. In that case, the tiled floor will eventually look like a large two-dimensional Cartesian plan. By using this characteristic, the robot for automatically conducting data acquisition in a selected study environment is developed. The robot is called Automatic Surveyor, as shown in Figure 3(a).

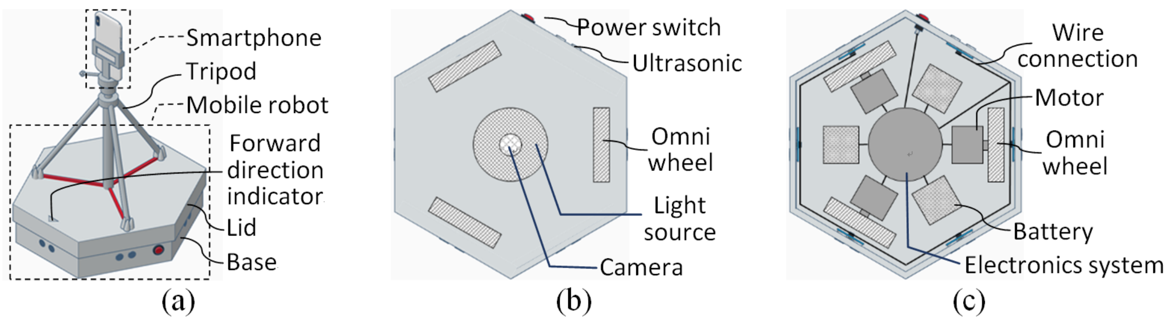


FIGURE 3. (a) Automatic Surveyor; (b) the base of the mobile robot (bottom view); (c) the base of the mobile robot (top view)

Automatic Surveyor is made of two main components: a component that can extract environmental reading and a component that can provide mobility ability to the previous component. The component that can provide mobility ability will be a mobile robot framed as shown in Figure 3(a). Based on Figure 3(a), the mobile robot consists of a base and a lid. The base is installed with various electronic components. The lid is to cover the electronic components installed on the base. Besides, the lid is also served as a platform for placing other devices on top of the mobile robot. The lid is allowed to be opened for the user to access the electronic components installed on the base. Electronic components installed on the base are shown in Figure 3(b) and Figure 3(c). Figure 3(b) illustrates the base view from the bottom, and Figure 3(c) illustrates the base view from the top.

According to Figure 3(b), a camera and a light source are installed at the bottom centre of the mobile robot base. The purpose of the camera is to capture grout lines/joints of a tiled floor, and the light source aims to provide constant light intensity while the camera is operating. By processing the images of grout captured by the camera, the mobile robot will be able to move accurately from one coordinate to another coordinate. According to Figure 3(c), there are three motors installed on the mobile robot base. The shaft of each motor is connected with an Omni wheel. Each of the motors (with respective Omni wheels connected) is arranged with a central angle of 120 degrees apart. In this configuration, the mobile robot developed is capable of moving in omnidirectional. In addition, three batteries are also installed with 120 degrees central angle apart from each other on the mobile robot base. However, they are having an offset of 60 degrees from the three motors. Besides that, on each side of the base, an ultrasonic sensor is installed to detect obstacles present in the movement direction of the mobile robot. On one side of the base, a power switch is installed to connect and cut off the power from the power source (the three batteries). Furthermore, at the centre of the mobile robot base, the electronic system of the mobile robot is installed. The electronic system is composed of a power regulating and distributing board, a motor controller, and two processing units. The relationship between all electronic components used for developing the mobile robot of Automatic Surveyor is illustrated using a block diagram, as shown in Figure 4.

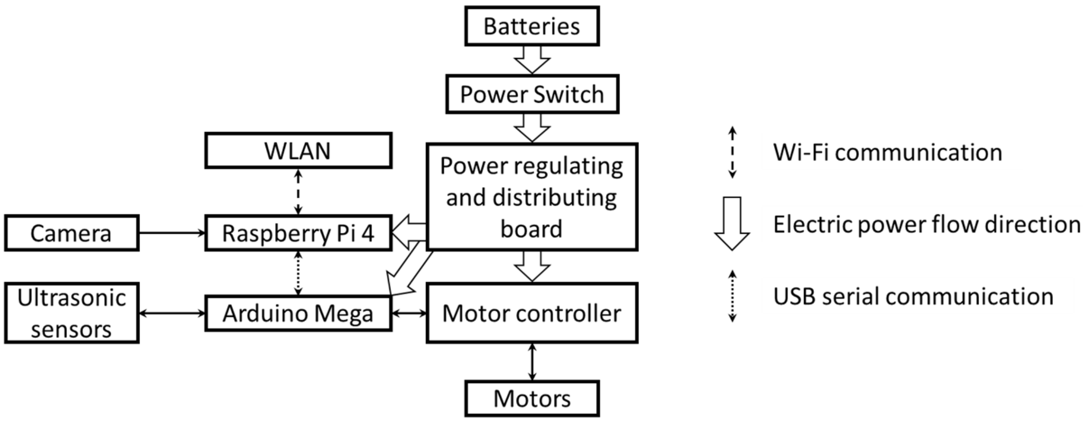


Figure 4. Overall relationship between electronic components of mobile robot

Based on Figure 4, batteries installed on the base are the electrical power source of the mobile robot. The electrical power from these batteries will flow through a power switch and finally being regulated and distributed to other electronic components by a power regulating and distributing board. Besides, the two processing units used on the mobile robot can be observed in Figure 4. The first processing unit is Arduino Mega. It is a microcontroller board based on ATmega2560. It will focus on processing reading from the ultrasonic sensors and controlling the motion of the motors. The second processing unit is Raspberry Pi 4. Raspberry Pi 4 is configured to function as a server that focuses on processing motion requests made by a client (through the WLAN established). When a motion request is received, image processing will be conducted on images captured by the camera to detect grout lines/joints on the tiled floor. Based on the position and orientation of the grout detected and the direction of motion request made by a client, motion commands will be generated by the Raspberry Pi 4 and sent to the Arduino Mega through USB serial communication to carry out the motion.

Lastly, regarding the component that can extract environmental reading, a smartphone is proposed to be used. In this modern era, most people have at least a smartphone device usually equipped with various sensors such as gyroscope, accelerometer, and Wi-Fi module. The reading of these sensors has commonly been collected by researchers while conducting data acquisition to obtain a fingerprint database for a fingerprinting based IPS. Therefore, a smartphone is chosen as the component to extract environmental reading. An automatic data acquisition application had been developed and installed on the smartphone. The application layout consists of three number input spinners, two start buttons, and a data table. Two number input spinners are for the user to input the starting/current x-coordinate and y-coordinate of the smartphone (also referring to the current coordinates of the mobile robot since the smartphone will be equipped on a tripod that is fixed on the lid of the mobile robot, as shown in Figure 3(a)). Another number input spinner is for the user to input the number of times data acquisition to be conducted on a coordinate. Once the user provides these three inputs, the user can click either the manual mode button to start with data acquisition on the current location only or click the auto mode button to start with automatic data acquisition algorithm execution. The process of automatic data acquisition will be explained in detail in a later section. Lastly, there is a table at the bottom of those input spinner and start buttons. The table is scrollable in both vertical and horizontal directions. It will display on time the data acquired.

3.3. **Environment layout.** After hardware preparation, the next stage will be deciding the study environment. The selected study environment needs to fulfil two basic requirements. The first requirement is that the selected study environment needs a floor tiled in grid patterns. The reason is that the developed Automatic Surveyor is navigating in a study environment based on groutlines/joints formed by tiled floor in grid patterns. The second requirement is that the study environment selected needs to be enclosed since an IPS will be developed and examined. After a short discussion, a classroom on the fifth floor of Faculty of Business is finally chosen out of a few proposed areas as the study environment. The layout of the classroom is shown in Figure 5.

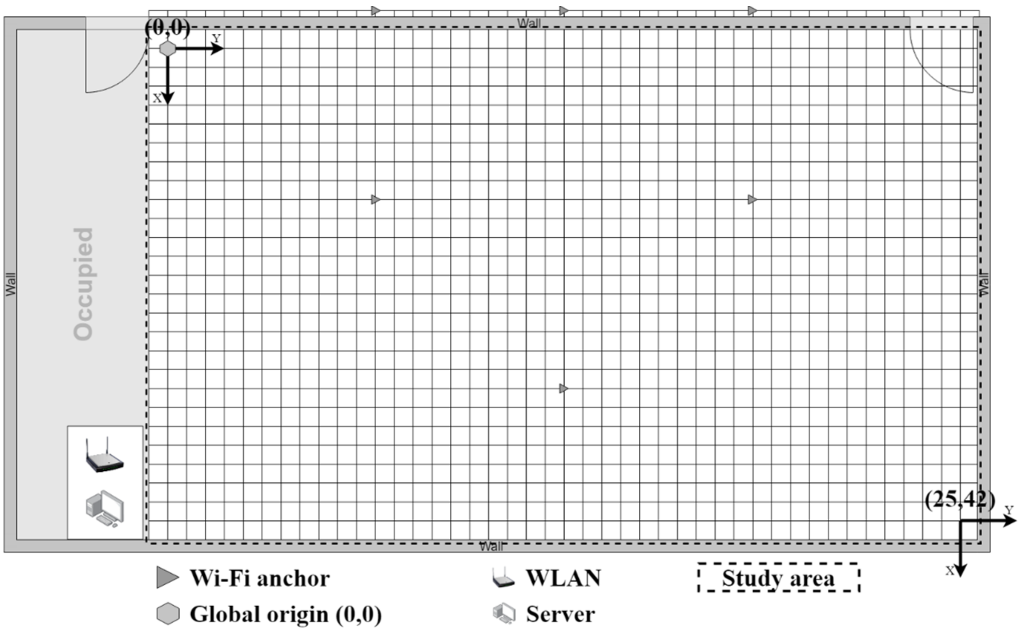


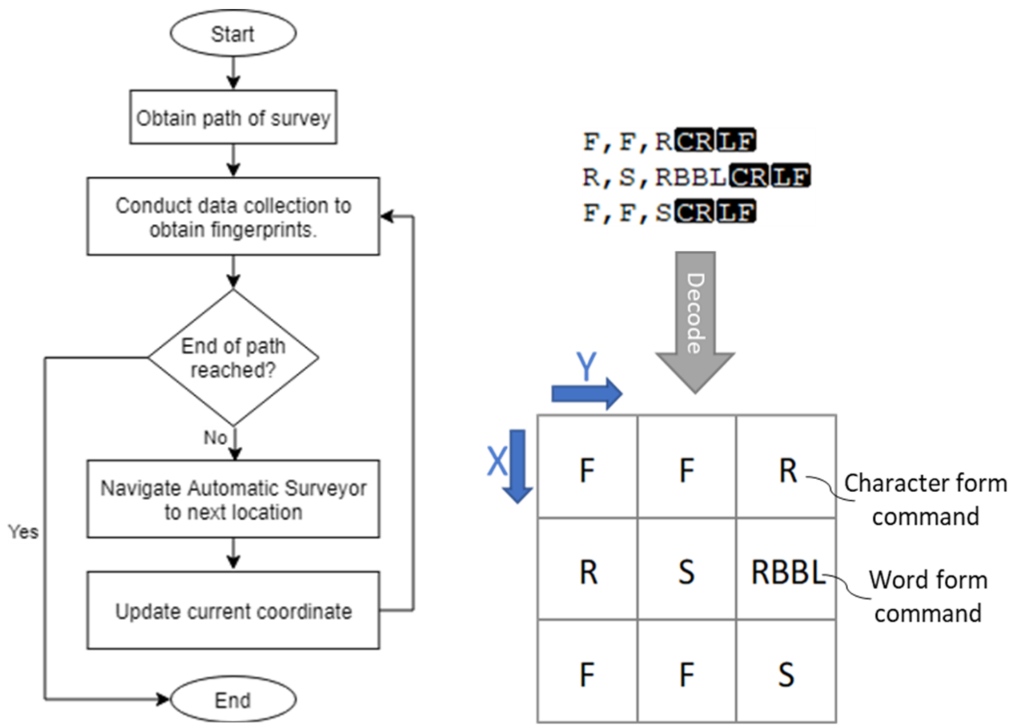
FIGURE 5. Environment layout

Based on the floor plan shown in Figure 5, the study environment is set up with hardware prepared. First and foremost, the six NodeMCU developed to function as Wi-Fi anchors are separately placed in the study environment at locations marked with a grey colour triangle that indicates Wi-Fi anchor position. Three NodeMCU preprogrammed to function as Wi-Fi anchors with AP1, AP2, and AP3 as SSID respectively were separately placed inside the classroom. This ensures at least three Wi-Fi anchors in LOS with tag node (Automatic Surveyor) when data acquisition is conducted in the study area. Moreover, the remaining three NodeMCU preprogrammed to function as Wi-Fi anchors with AP4, AP5, and AP6 as SSID respectively had been set up at the interior corridor outside the classroom. This ensures at least three Wi-Fi anchors in NLOS (with wall act as the obstacle) with tag node when data acquisition is conducted. With Wi-Fi anchors arrangement as mentioned, the fingerprint database obtained after data acquisition not only can be used for the study of the performance of the developed IPS using fingerprints formed by RSS from different numbers of Wi-Fi anchors but also used to study the performance of IPS that uses fingerprinting formed by RSS from Wi-Fi anchors in LOS and NLOS condition with tag node.

Next, the computer configured as a server and the modem served as WLAN were also set up in the study environment. Furthermore, from the layout illustrated in Figure 5, it can be seen that the study environment consists of two doors. Both of the doors are connecting to the same interior corridor. However, these doors will maintain in close condition when data acquisition is being conducted. Next, the area framed with a short-dashed square, as shown in Figure 5, represents the study area. This area is tiled with a total of 27 tiles in the x-direction and 44 tiles in the y-direction. Each tile is having a dimension of 0.3m $\times$ 0.3m. The grout lines and grout joints formed between these tiles will be the path to be followed by the Automatic Surveyor developed and the position of data acquisition being conducted. Lastly, the grout lines/joints formed between tiles result in a large cartesian plan. Thus, a global origin had been declared on it, as shown in Figure 5. The global origin is used as a reference point to define other locations in the study area.

3.4. **Data acquisition.** Data acquisition is a process required for obtaining a fingerprint database of a study environment. The previous few subsections provided explanations on hardware prepared to set up the selected study environment and conduct data acquisition. Thus, this section will explain the use of the Automatic Surveyor developed in helping to obtain a fingerprint database required by conducting automatic data acquisition for the selected study environment. First and foremost, a survey path for the study environment will be prepared and saved on the server to be used by the venue service. The survey path will cover all the coordinates/locations within the study area except the three coordinates placed with Wi-Fi anchors. This means that when the Automatic Surveyor performs automatic data acquisition according to the survey path, data collection will be conducted on 1115 different coordinates. Next, the developed Automatic Surveyor will be placed on the global origin facing toward positive y-axis (as forwarding direction) with the power of mobile robot switched on and smartphone installed with data acquisition application being executed. Upon this point, the computer configured to serve as a server providing database service and venue service, the Raspberry Pi 4 of the mobile robot, and the smartphone fixed on the tripod on top of the mobile robot required to ensure are all connected to the WLAN established by a modem set up in the study environment. Before starting with executing the automatic data acquisition algorithm by clicking on the auto mode button, the three number input spinners display on the application needs to be set. Two of the number input spinners will be set with the current coordinate of the Automatic Surveyor, which at this moment is the global origin $(0, 0)$. The remaining number input spinner will be set with the number of times to conduct scanning for environmental reading at each location which is 10 for this study. This means ten samples of environmental readings will be collected at each location.

Once the three number input spinners had been set, the auto mode button will be clicked to execute the automatic data acquisition algorithm. Figure 6(a) shows a block diagram with major processes that happens during the automatic data acquisition. The automatic data acquisition algorithm will start by creating three variables to store: the current x-coordinate, the current y-coordinate, and the number of times to perform scanning for environmental reading at each coordinate inputted by the user using the three number input spinners. Next, the automatic data acquisition algorithm will continue by sending a request through WLAN to the server providing venue service for obtaining a survey path. The server providing venue service will process the request and generate a response. The response is containing the survey path initially prepared and saved on the server. The survey path is a list of commands separated with commas and Carriage Return (CR) and Line Feed (LF) at the end of each row. Next, the automatic data acquisition algorithm will decode the survey path and save it into a two-dimensional array indexed using the current coordinate, as shown in Figure 6(b).



(a) (b)

FIGURE 6. (a) Automatic data acquisition workflow; (b) sample survey path before and after decode into array form

After the survey path is saved into a two-dimensional array, the automatic data acquisition algorithm will conduct data collection to obtain a fingerprint. First and foremost, scanning for environmental reading will be conducted using the built-in Wi-Fi module on the smartphone. The environmental reading obtained after the scanning process is formed by six distance-dependent parameters: Wi-Fi RSS from the six Wi-Fi anchors set up. Secondly, the environmental reading obtained will be labelled with a location tag (automatically created based on the current coordinate value) to generate a fingerprint.

Thirdly, the fingerprint generated will be submitted to the server providing database service through WLAN. The automatic data acquisition algorithm will repeat conducting data collection for N times. N is the number of times to perform scanning for environmental reading at each coordinate, inputted by the user before starting with the execution of the automatic data acquisition algorithm. Once data collection at a coordinate is completed, the array containing the decoded survey path will be indexed using the current coordinate to read the command that indicates the path to the next coordinate. Each command can be in the form of a character (formed by an English alphabet) or a word (formed by English alphabets). Only five of the English alphabets (repetitive allowed) are used to form a command: F (forwards), B (backwards), L (left), R (right), and S (End of Path (EoP)).

Suppose the command is in the form of a character. The algorithm will be terminated if the character indicates the EoP reached. If not EoP, based on the character, a motion request will be sent by the automatic data acquisition algorithm to the server on the mobile robot (hosted by Raspberry Pi 4) through WLAN. The server on the mobile robot will then process the request and execute image processing algorithm to detect the grout lines/joints on each image captured by the camera installed at the bottom centre of the mobile robot. Based on the position and orientation of grout lines being detected on each image captured, motion commands will be generated by the server on the mobile robot. These motion commands will be sent to the Arduino Mega installed on the mobile robot. The Arduino Mega will then navigate the mobile robot based on the motion command received. Once the image processing algorithm detects a grout joint (indicate next coordinate reached) on images captured by the camera, the server will generate a stop command to Arduino Mega. The server will also follow by generating a response and sending it through WLAN to the client (automatic data acquisition algorithm) to inform the complete execution of the motion request. Once the automatic data acquisition algorithm receives the response, the current coordinate will be updated based on the previously executed command. For example, suppose it is a forward command. In that case, the current x-coordinate will maintain, but the current y-coordinate will increase by one as the Automatic Surveyor is set up facing toward positive y-axis.

Suppose the command is in the form of a word. The word will be separated into a series of character commands. Each character command will be executed in order with the same process as a command in the form of a character until all characters in the series of character commands are completed. After that, the automatic data acquisition algorithm will loop back to conduct data collection for obtaining fingerprints. The automatic data acquisition algorithm will continue executing until the EoP is reached. While the automatic data acquisition algorithm is executing, the orientation of the Automatic Surveyor will always remain the same from a coordinate to another coordinate. Once the Automatic Surveyor completed automatic data acquisition on the study environment selected, the database server will consist of a total of 11150 fingerprints. For a better understanding of the movement of Automatic Surveyor while executing automatic data acquisition algorithm, we can have a look at Figure 7.

Figure 7 illustrates the movement of the Automatic Surveyor in a sample environment with four tiles in the x-direction and four tiles in the y-direction. The movements are based on a sample survey path starting at a coordinate value of $(0, 0)$. The hexagon icon in solid colour is the current location of the Automatic Surveyor before executing the command in the sample survey path. Besides, the translucent arrow is the direction of movement, and the translucent hexagon is the future location of the Automatic Surveyor after executing the command in the sample survey path.

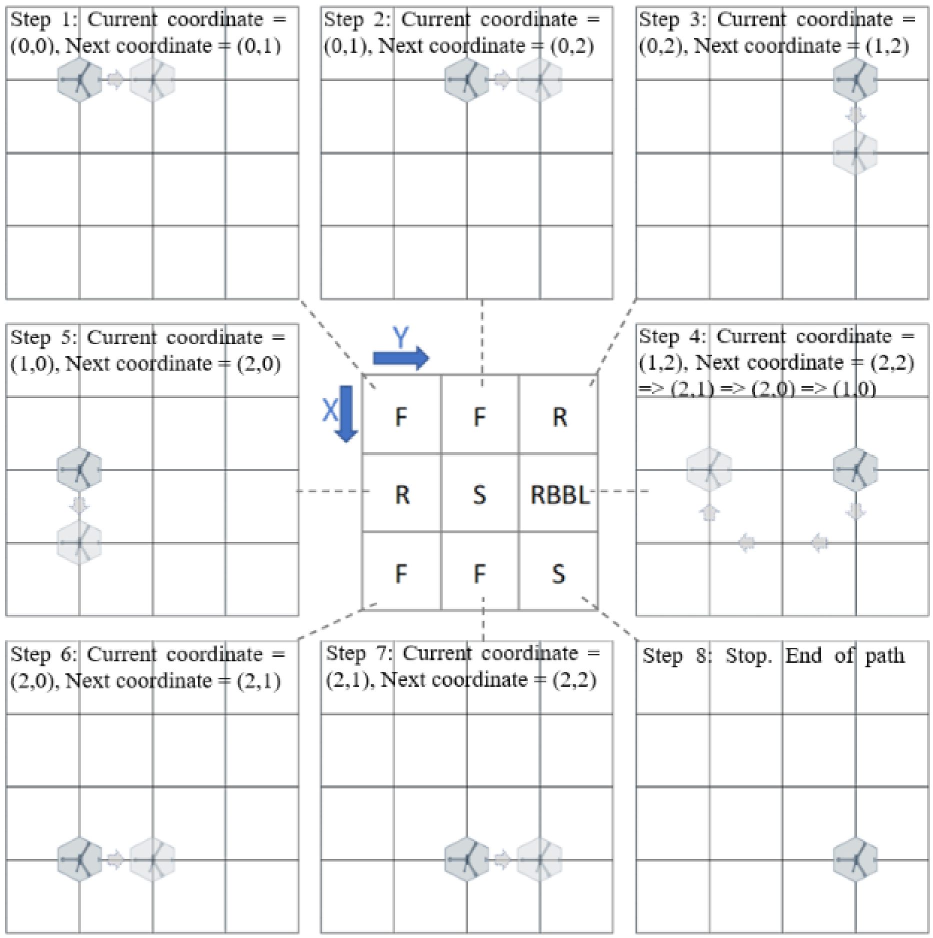


Figure 7. Automatic data acquisition process following a sample survey path

In a nutshell, developing a fingerprinting based IPS required obtaining a fingerprint database for a study environment beforehand. This section had presented the use of the developed Automatic Surveyor in helping to obtain a fingerprint database for a selected study environment (described with a floor plan in the previous subsection). The use of the developed Automatic Surveyor for obtaining a fingerprint database has several advantages. Firstly, the workforce required is reduced as data acquisition for the selected study environment is done automatically by the Automatic Surveyor. Secondly, data acquired by Automatic Surveyor while performing automatic data acquisition can be observed in real time as each fingerprint generated is submitted to the server. Besides, the time required to complete data acquisition for a study environment can be reduced by multiplying the Automatic Surveyor and defining a different survey path for each Automatic Surveyor. Last but not least, compared to performing data acquisition using an SLAM robot where location tags generated are facing cumulative errors, the location tags generated by Automatic Surveyor not only will not experience cumulative errors as SLAM robot, the location tags generated also have high accuracy as the method to generate location tag is different. SLAM robots generate location tags using current position estimated by using multiple sensors values. In contrast, Automatic Surveyor developed generates location tags using current position update when Automatic Surveyor travels from one position to another by tracking grout lines/joints formed by floor tiled in grid patterns.

3.5. **Experiment procedure.** This subsection will present the procedure to carry out two experiments using the fingerprint database constructed for the selected study environment. The first experiment will focus on studying the positioning performance of IPS developed when fingerprints used are with environmental reading formed by RSS from different types of Wi-Fi anchors or different numbers of Wi-Fi anchors. The second experiment will focus on studying the impact on positioning performance of IPS when the location tag of fingerprints used for training of positioning algorithm adopted contains errors. Throughout both experiments, a neural network is selected as the positioning algorithm adopted to develop an IPS. It is a mathematical model that is well known for recognizing patterns in data. Thus, a neural network used as a positioning algorithm throughout both experiments is expected to recognize features patterns and follow by estimating target position. One thing that needs special attention is that neural network architecture is also a factor affecting positioning accuracy. Thus, only one neural network architecture will be used throughout both experiments. This is to exclude neural network architecture becomes a factor affecting positioning accuracy while performing both experiments. The programming environment used to develop neural networks to conduct studies for both experiments is Google Colaboratory. Python is the programming language, and TensorFlow is the AI framework used to develop neural networks.

Before starting with both experiments, few tasks must be done in advance. First and foremost, all fingerprints in the fingerprint database are exported into a CSV file. Each row of the data kept inside the file is a fingerprint. Thus, the file consists of 11150 rows, as Automatic Surveyor developed performed automatic data acquisition on 1115 locations in the selected study environment, and data collection was performed ten times at each location. Besides, each row of the data consists of eight columns. The first two columns are the x-coordinate and y-coordinate, together known as a location tag. The remaining columns are RSS values from the six Wi-Fi anchors (with SSID from AP1 to AP6 in order), together known as an environmental reading. A python program is prepared to load all fingerprints in the CSV file and standardize each distance-dependent parameter (feature) of environmental readings. Standardizing is a necessary preprocessing to speed up the training process of a neural network later by reaching convergence faster. All fingerprints with standardized environmental reading will be exported by the python program into a new file to be used in both experiments later. In both experiments, during the training and validation processes, distance-dependent parameters of each environmental reading will be fed into neural networks as inputs. Location tag will be the expected output of neural networks. However, which distance-dependent parameters are selected to feed into neural networks as inputs depends on the study conducted.

The first experiment will start by loading the file containing fingerprints with standardized environmental reading into the programming platform. Next, the 11150 fingerprints loaded will be separated into two sets: a training set and a validation set with a ratio of $9 : 1$ to be used as data for training and validation of the neural network. In order to make sure all the coordinate is being validated, nine fingerprints obtained on each location will be randomly picked and put into the training set. The remaining one fingerprint will be put into the validation set. As a result, the training set will consist of 10035 fingerprints and the validation set consisting of 1115 fingerprints. After obtaining the training set and validation set, the next step will be identifying a high performing neural network architecture. This also means identifying the hyperparameter of a high performing neural network as hyperparameters are parameters that define the neural network architecture. In this study, the hyperparameters of a high performing neural network are obtained through grid search and fine-tuning process. While performing grid search and fine-tuning, inputs of neural networks will be RSS from all Wi-Fi anchors. The best positioning accuracy is expected when RSS from more number Wi-Fi anchors is used as input. The resulting best performance neural network architecture after the grid search and the fine-tuning process is with a minimum of three up to a maximum of six neurons on the input layer (depend on the study conducted), three hidden layers with 512 neurons and hyperbolic tangent function as activation function, an output layer with two neurons and linear as activation function, optimiser algorithm is Adam, the loss function is mean square error, the learning rate is 0.001 and batch size is 64. As mentioned early in this subsection, only one neural network architecture will be used throughout both experiments. Thus, the resulting neural network architecture after the grid search and fine-tuning process will also be used for the second experiment.

The first experiment is separated into two parts. The first part will study the positioning performance of neural networks developed when fingerprints used are with environmental reading formed by RSS from different types of Wi-Fi anchors. The second part will study developed neural networks positioning performance when fingerprints used are with environmental reading formed by RSS from different numbers of Wi-Fi anchors. There are a total of six Wi-Fi anchors categorized into two types: Wi-Fi anchors that are LOS with the Automatic Surveyor (which are Wi-Fi anchors with AP1, AP2, and AP3 as SSID) and NLOS with the Automatic Surveyor (which are Wi-Fi anchors with AP4, AP5, and AP6 as SSID).

In order to conduct the study for the first part of the first experiment, the high performing neural network architecture (identified previously) will be used to develop neural network models. All the neural network models developed will have three input neurons as the maximum number of Wi-Fi anchors LOS or NLOS with Automatic Surveyor is both three. The neural network models developed will be trained and validated using fingerprints with environmental reading formed by RSS from three different Wi-Fi anchors. The environmental reading of each fingerprint in both the training and validation sets is formed by RSS from six Wi-Fi anchors. When RSS from three out of six Wi-Fi anchors is selected, twenty different combinations without repetition can be formed. Equation (1) had been used to obtain the number of possible combinations. The letter n denotes the number of different Wi-Fi anchors with RSS values extracted by Automatic Surveyor during automatic data acquisition. The letter r denotes the number of different Wi-Fi anchors with RSS values being extracted is selected out of n. Thus, 20 neural network models will be developed using the same high performing neural network architecture. Each neural network model will be trained and validated using fingerprints with environmental reading formed by different combinations of RSS from three Wi-Fi anchors. The 20 different combinations are categorized into four groups according to the types of RSS selected: 3 LOS only = 1 combination, 2 LOS + 1 NLOS = 9 combinations, 1 LOS + 2 NLOS = 9 combinations, and 3 NLOS only = 1 combination. The validation result of each neural network will be summarized according to the groups and presented in the coming section.

$$ {}_nC_r = \frac{n!}{r!(n-r)!} \tag{1} $$

Next, the process of conducting the second part of the first experiment is similar to the first part. The second part of the first experiment will also start by developing neural network models using previously identified neural network architecture. However, the number of input neurons of neural network models developed will vary from three to six as there are six Wi-Fi anchors set up, and a minimum of three is required for 2D positioning. The neural network models developed will be trained and validated using fingerprints with environmental reading formed by RSS from a minimum of three up to a maximum of six different Wi-Fi anchors. Thus, this study will result in a total of 42 neural network models (${}_6C_3 = 20$, ${}_6C_4 = 15$, ${}_6C_5 = 6$, ${}_6C_6 = 1$) being developed.

Furthermore, obtaining a training set and validation is the first step in the second experiment. The method to obtain both sets is the same as the first experiment. After obtaining both sets, the location tag of fingerprints in the training set will be modified by added with error. Next, neural network models will be developed, trained using fingerprints (with error added) in the training set and validated using fingerprints (without error added) in the validation set. The neural network models developed will be with six input neurons as fingerprints used in this second experiment will be with environmental reading formed by RSS from all six Wi-Fi anchors. Two types of errors will be separately added and being investigated. The first type of error is a random error generated according to a probability density function with shorthand notation represented as in (2) where e_x and e_y indicate random x-coordinate error and random y-coordinate error generated according to a bivariate normal distribution (denoted by N) with correlation (denoted by ρ) equal to zero and with given mean (denoted by μ) and standard deviation (denoted by σ). Random error generated according to (2) will be studied from two aspects: fixed mean ($\mu = 0$) with varying standard deviation ($\sigma \in \mathbb{Z}$, $\sigma \in [1, 15]$) and varying mean ($\mu \in \mathbb{Z}$, $\mu \in [0, 15]$) with fixed standard deviation ($\sigma = 1$). Fingerprints with the exact location tag will be added with the exact magnitude of errors as these fingerprints are obtained at the same location during data collection.

$$\begin{pmatrix} e_x \\ e_y \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \rho\sigma_y\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}\right], \quad \left\{ \begin{array}{c} \mu_x = \mu_y = \mu \\ \sigma_x = \sigma_y = \sigma \\ \rho = 0 \end{array} \right\} \tag{2}$$

The second type of error is a cumulative error. The cumulative error will be studied from two aspects: cumulative error on x-direction only and y-direction only. While performing the study for cumulative error on x-direction, 43 error vectors, each with 26 elements, are generated. Forty-three error vectors are because y-coordinate is ranged from 0 to 42. Twenty-six elements are because x-coordinate is ranged from 0 to 25. Elements in each error vector are random errors generated according to a probability density function with shorthand notation, as shown in (3), where the random error (denoted by e) is distributed according to a normal distribution (denoted by N), with a mean (denoted by μ) of 0 and with different standard deviation (denoted by σ). Next, elements in each error vector will be modified, starting from the last element in each vector, using Equation (4), where e_n denoted random error at n position in a vector will be replaced with the sum of absolute random error at the first position until n position in the vector. Once modification is done, each error vector is now a cumulative error vector. The location tag of each fingerprint in the training set will be used to index cumulative error vectors to obtain cumulative error to be added to the first element in the location tag of each fingerprint in the training set. For example, a fingerprint with the location tag of $(2, 5)$, its first element will be added with error in the third position of the sixth cumulative error vector.

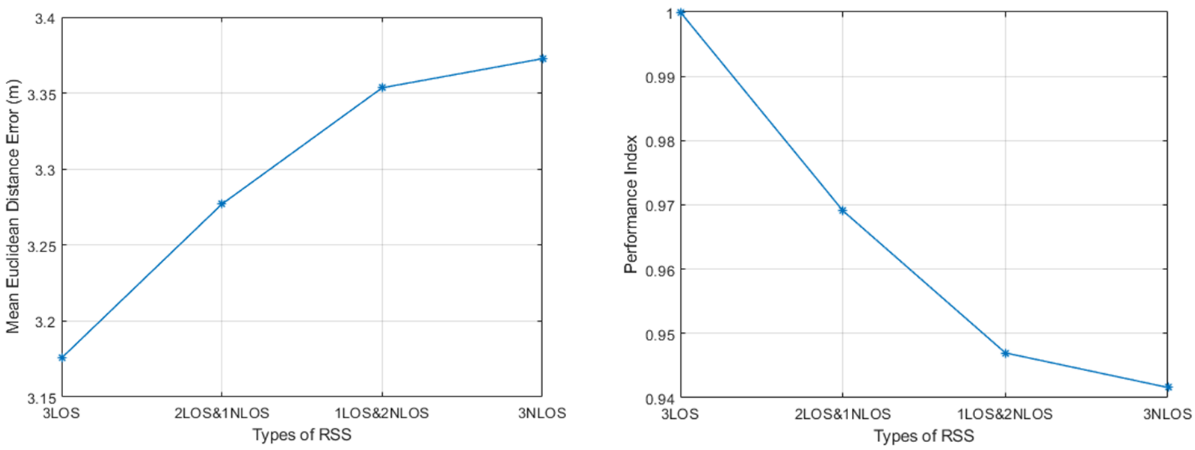
$$e \sim N[\mu, \sigma], \quad \left\{ \begin{array}{c} \mu = 0 \\ \sigma \in \mathbb{R} \mid 0.05 \leq \sigma \leq 1.00,\ \sigma \text{ is multiple of } 0.05 \end{array} \right\} \tag{3}$$

$$e_n = \sum_{i=1}^{n} |e_i| \tag{4}$$

A similar procedure will undergo when performing the study for cumulative error on y-direction. The differences are the number of error vectors generated based on the number of x-coordinate, the number of elements in each error vector is based on the number of y-coordinate, and the error generated is added to the second element in a location tag of fingerprints in the training set. Lastly, when conducting the training and validation for all the developed neural network models in both experiments, the early stopping regularization technique is used to prevent overfitting that commonly happens during neural network development. Overfitting is an issue that occurs when a neural network model developed can make good predictions with training data but make bad predictions when the neural network model is fed in with data that is not being seen before by the neural network.

4. **Experimental Results.** Throughout both experiments, the positioning error is calculated using the Euclidean distance formula. For simplicity purposes, later in this section, when RSS LOS fingerprints or RSS NLOS fingerprints exist, its mean fingerprints with environmental reading formed by RSS from all Wi-Fi anchors that are LOS with Automatic Surveyor or NLOS with Automatic Surveyor, respectively. After performing the first part of the first experiment, results show that the neural network model trained and validated using RSS LOS fingerprints performs better than when using RSS NLOS fingerprints. The mean positioning error of the neural network model validated using RSS LOS fingerprints achieved 3.176 meters smaller than 3.373 meters which is the mean positioning error when using RSS NLOS fingerprints. However, the neural network model validated using RSS LOS fingerprints presented a standard deviation of 1.853 meters, which is slightly larger when comparing with the neural network model validated using RSS NLOS fingerprints with a standard deviation of 1.823 meters. The trend of mean positioning error of neural network models validated using fingerprints with environmental reading formed RSS from different types of Wi-Fi anchors is shown in Figure 8(a). It can be observed that the mean positioning error is increasing when the amount of RSS from Wi-Fi anchors LOS with tag node is reducing. Furthermore, the performance index of the neural network models trained and validated using fingerprints with environmental reading formed by different types of RSS combinations is illustrated in Figure 8(b). Although there is a decrease in the performance index when RSS LOS is replaced with RSS NLOS, the decrease is just up to 0.0584.

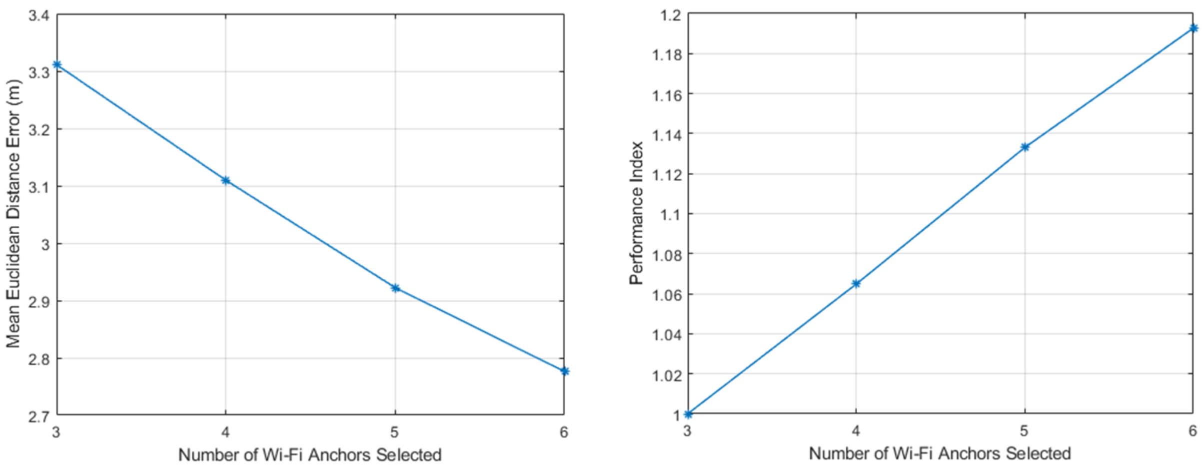
Next, the results obtained after performing the second part of the first experiment show that the positioning accuracy of neural network models increases when fingerprints used for training and validation are with environmental reading formed by RSS from



(a) (b)

FIGURE 8. (a) Trend of mean Euclidean distance error against fingerprints with environmental reading formed by different types of RSS; (b) performance index against different types of RSS

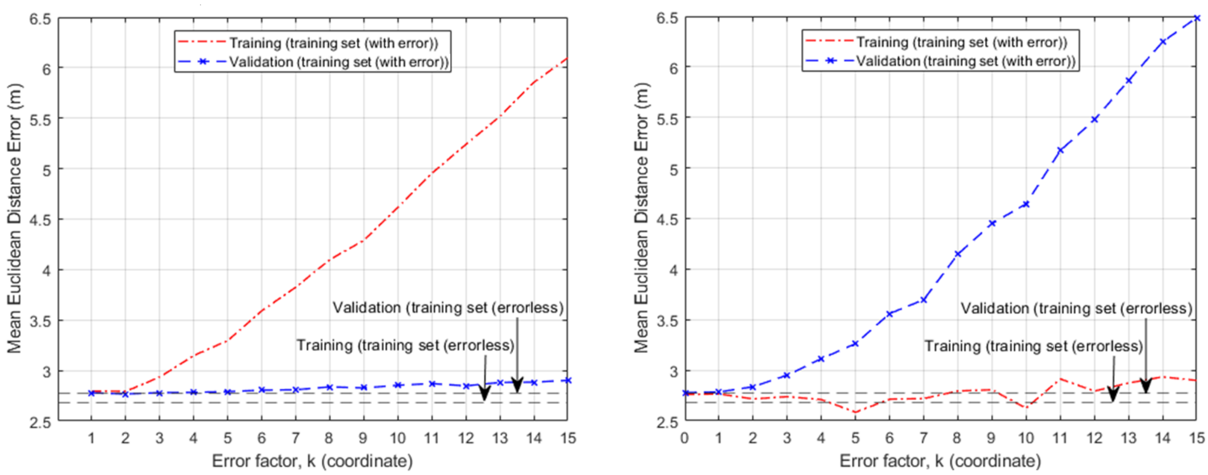
more number of Wi-Fi anchors. Figure 9(a) illustrates the trend of mean positioning error of neural network models after being trained and validated using fingerprints with environmental reading formed by RSS from different numbers of Wi-Fi anchors. It can be observed that the positioning error minimized from 3.311 meters down to 2.776 meters when the number of Wi-Fi anchors with RSS selected as distance-dependent parameters of environmental reading increased from three to six. Besides, the reduction in positioning error is quite linear to the number of Wi-Fi anchors with RSS used to form environmental reading. Although the number of Wi-Fi anchors with RSS used to form environmental reading increased until the highest number of six, the positioning error still reduces with a large negative gradient. This means further increasing the number of Wi-Fi anchors with RSS selected to form environmental reading can obtain a smaller positioning error. Next, the performance indexes of neural network models trained and validated using fingerprints with environmental reading formed by RSS from different Wi-Fi anchors are illustrated in Figure 9(b). The figure shows that the performance index increases until 1.1927 when all the Wi-Fi anchors set up in the study environment are with RSS selected to form environmental reading of fingerprints used for training and validation.



(a) (b)

Figure 9. (a) Trend of mean positioning error against fingerprints with environmental reading formed by different numbers of RSS; (b) performance index against different numbers of RSS

Furthermore, the result obtained after performing the second experiment is illustrated in Figure 10 and Figure 11. In both figures, two dashed horizontal lines mark the training and validation mean positioning errors of a neural network model (trained using fingerprints in the training set (errorless)) when examined with fingerprints in the training set (errorless) and validation set (errorless), respectively. Figure 10 illustrates the training and validation mean positioning errors of neural network models trained using a training set containing fingerprints with errors on the location tag. The errors were generated according to a bivariate normal distribution with a given mean and standard deviation. When a neural network model is trained using a training set with errors having a large standard deviation or large mean, it can be observed that the training and validation mean positioning errors both increase when examined with the training set (with errors) and validation set (errorless), respectively. However, the impact on the training mean positioning error and validation mean positioning error is different. When the location tag of fingerprints in the training set has errors with a larger standard deviation, a larger

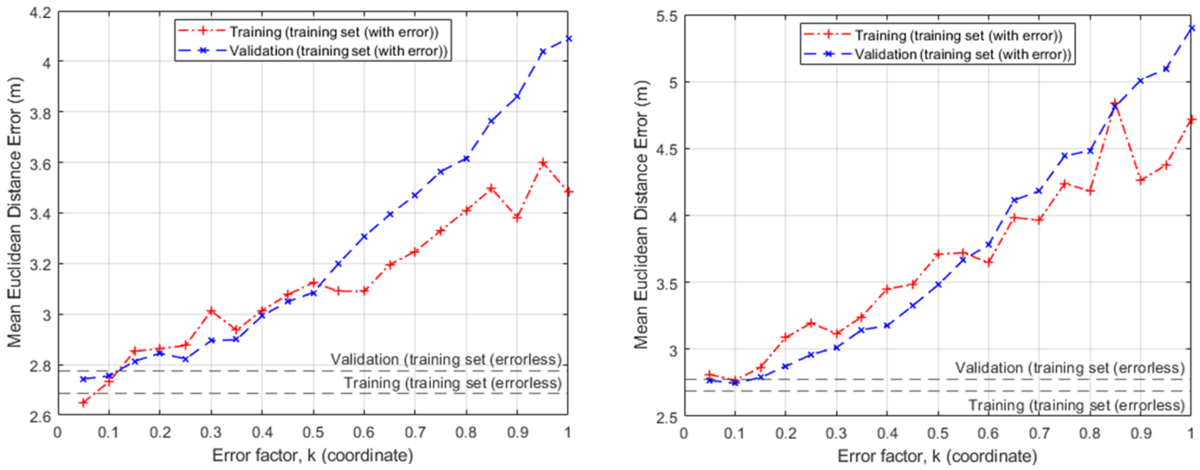


(a) (b)

FIGURE 10. Mean positioning error for neural network models when the location tag of fingerprints used for training is with errors generated according to a bivariate normal distribution: (a) With $\mu = 0$ and $\sigma = k$; (b) with $\mu = k$ and $\sigma = 1$

training mean positioning error can be observed than validation mean positioning error. Inverted circumstances can be observed when the location tag of fingerprints in the training set has errors with a larger mean. This result shows that suppose fingerprints collected for a neural network model training have a large error on the location tag of each fingerprint. The mean position error of the trained neural network when examined with a validation set (errorless) may not be heavily impacted. It depends on whether the errors on the location tag of fingerprints used for neural network training have a large standard deviation or a large mean. Suppose it is a large standard deviation. The impact on mean positioning error when examined with a validation set (errorless) is negligible.

Last but not least, Figure 11 illustrates the training and validation mean positioning errors of neural network models trained using a training set containing fingerprints with cumulative errors on the location tag generated according to a normal distribution with a mean of zero and different standard deviation. Based on the figure, when neural network models trained with the training set containing fingerprints with x-coordinate or y-coordinate of the location tag consist of cumulative error with mean $= 0$ and increasing standard deviation, the neural network models trained will present an increasing trend in training and validation mean positioning errors when examined with training set (with cumulative error) and validation set (errorless), respectively. This is because although errors on the location tag of each fingerprint in the training set are generated according to a normal distribution function with a mean of zero and small standard deviation $[0.05, 1]$, when these errors accumulated along each axis direction, the overall mean and standard deviation of error existing in the location tag of fingerprints used for training are both increasing. This study of errors cumulative on x-direction and y-direction simulates when data collection is performed using an SLAM robot moving straight in the x-direction only and y-direction only starting from $x = 0$ and $y = 0$, respectively. This study result shows that for an SLAM robot with a slight error while moving from a coordinate to another coordinate when the error is accumulated, the quality of fingerprinting database constructed will reduce. Thus, this also reflects the importance of having a robot that can automatically conduct data acquisition and generate high accuracy location tags.



(a) Cumulative error on x-direction (b) Cumulative error on y-direction

FIGURE 11. Mean positioning error for neural network models when the location tag of fingerprints used for training has cumulative errors generated according to a normal distribution with $\mu = 0$ and $\sigma = k$

5. **Conclusions.** In a nutshell, this paper presents a method to conduct automatic data acquisition by utilizing the characteristic of floor tiled in grid patterns. After adopting the method presented in a selected study environment, the fingerprint database constructed is used to conduct two experiments. Both experiments are studying the performance of neural network models when trained and validated using different kinds of fingerprints. Only one neural network architecture is used to avoid neural network architecture becoming a factor affecting positioning performance. The first experiment result shows that neural network models trained and validated using fingerprints with environmental reading formed by RSS from more Wi-Fi anchors LOS with tag node (Automatic Surveyor) presented a smaller positioning error. Besides, the first experiment also shows that neural network models trained and validated using fingerprints with environmental reading formed by RSS from a greater number of Wi-Fi anchors can minimize positioning error. Until the maximum amount of six Wi-Fi anchors placed with RSS being extracted to form environmental reading of fingerprints, the gradient of decrease in positioning error is still significant. Next, the second experiment shows that when the location tag of fingerprints used for neural network models training suffers from errors, the impact on positioning accuracy when validated with a validation set (errorless) might not always be high. The location tag of fingerprints used for training suffering from error with small mean and large standard deviation will have less impact on positioning accuracy. However, the location tag of fingerprints, when consisting of errors with a large mean, although with a small standard deviation, will significantly impact positioning accuracy when validated with a validation set (errorless). Lastly, the cumulative error study done in the second experiment shows that suppose the location tag of fingerprints of training set contains errors generated based on a normal distribution with a small standard deviation. When these errors are absolute and accumulated (simulate the cumulative error on the location tag generated by the SLAM robot when performing data collection by moving straight in one direction), the overall mean and standard deviation will increase. Results in neural network model trained using these fingerprints present a large mean positioning error when validated with a validation set (errorless). For future work, the positioning performance of neural networks using fingerprints with environmental reading formed by a greater

number of features can be investigated. A hybrid wireless technology (Wi-Fi and UWB) fingerprinting based IPS can be constructed to study the trend of positioning accuracy when fingerprints used are with environmental reading formed by Wi-Fi RSS replaced or combined with UWB RSS.

REFERENCES

[1] T. Mai, *Global Positioning System History*, 2017, https://www.nasa.gov/directorates/heo/scan/communications/policy/GPS_History.html, Accessed on 30-Jan-2021.

[2] B. Peterson, D. Bruckner and S. Heye, Measuring GPS signals indoors, *Proc. of the 10th Int. Tech. Meet. Satell. Div. Inst. Navig. (ION GPS 1997)*, Kansas City, MO, pp.615-624, 1997.

[3] J. Liao, B. Guo and X. Wu, An indoor location tracking optimization algorithm based on NLOS model, *ICIC Express Letters, Part B: Applications*, vol.9, no.1, pp.77-84, 2018.

[4] F. S. Danis, A. T. Cemgil and C. Ersoy, Adaptive sequential Monte Carlo filter for indoor positioning and tracking with Bluetooth low energy beacons, *IEEE Access*, vol.9, pp.37022-37038, 2021.

[5] H. Yao, H. Shu, X. Liang, H. Yan and H. Sun, Integrity monitoring for Bluetooth low energy beacons RSSI based indoor positioning, *IEEE Access*, vol.8, pp.215173-215191, 2020.

[6] H. C. Lee and T. T. Lin, Low-cost indoor human tracking by utilizing fluctuation of received radio signal strength, *IEEE Sens. J.*, vol.20, no.21, pp.13029-13036, 2020.

[7] Z. Wei, X. Chen, L. Fang, N. Zhao, S. Guo and X. Li, Joint positioning technique based on TOF and TDOA, *2018 IEEE Int. Instrum. Meas. Technol. Conf.*, pp.1-6, 2018.

[8] J. M. Hong, S. H. Kim, K. J. Kim and C. G. Kim, Multi-cell based UWB indoor positioning system, in *Intelligent Information and Database Systems. ACIIDS 2019. Lecture Notes in Computer Science*, N. Nguyen, F. Gaol, T. P. Hong and B. Trawiński (eds.), Cham, Springer, 2019.

[9] A. Schneider and H. Feussner, Tracking and navigation systems, *Biomed. Eng. Gastrointest. Surg.*, pp.443-472, 2017.

[10] T. A. Z. Rahman, N. Ramli, M. A. A. Rahman and H. Mohamad, Hybrid nearest neighbour and feed-forward neural networks algorithm for indoor positioning system, *2020 IEEE 8th Conference on Systems, Process and Control (ICSPC)*, pp.37-40, 2020.

[11] W. Njima, I. Ahriz, R. Zayani, M. Terre and R. Bouallegue, Smart probabilistic approach with RSSI fingerprinting for indoor localization, *2017 25th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp.1-6, 2017.

[12] S. He and S. H. G. Chan, Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons, *IEEE Commun. Surv. Tutorials*, vol.18, no.1, pp.466-490, 2016.

[13] N. A. M. Maung and W. Zaw, Comparative study of RSS-based indoor positioning techniques on two different Wi-Fi frequency bands, *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pp.185-188, 2020.

[14] I. Lee et al., Wi-Fi HaLow for long-range and low-power Internet of Things: System on chip development and performance evaluation, *IEEE Commun. Mag.*, vol.59, no.7, pp.101-107, 2021.

[15] W. Shao, H. Luo, F. Zhao, Y. Ma, Z. Zhao and A. Crivello, Indoor positioning based on fingerprint-image and deep learning, *IEEE Access*, vol.6, pp.74699-74712, 2018.

[16] M. Ibrahim, M. Torki and M. Elnainay, CNN based indoor localization using RSS time-series, *2018 IEEE Symposium on Computers and Communications (ISCC)*, pp.1044-1049, 2018.

[17] S. Dai, L. He and X. Zhang, Autonomous WiFi fingerprinting for indoor localization, *Proc. of 2020 ACM/IEEE 11th Int. Conf. Cyber-Physical Syst. (ICCPS2020)*, pp.141-150, 2020.

[18] D. Sun, E. Wei, L. Yang and S. Xu, Improving fingerprint indoor localization using convolutional neural networks, *IEEE Access*, vol.8, pp.193396-193411, 2020.

[19] T. Arsan and M. M. N. Hameez, A clustering-based approach for improving the accuracy of UWB sensor-based indoor positioning system, *Mob. Inf. Syst.*, vol.2019, no.2, pp.1-13, 2019.

[20] K. R. Athira and A. Babu, Zigbee based indoor location tracking and monitoring system, *Int. J. Recent Technol. Eng.*, vol.7, no.6, pp.1651-1656, 2019.

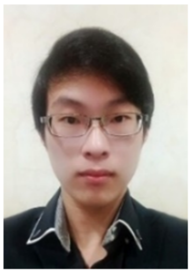
[21] X. Wang, L. Gao, S. Mao and S. Pandey, CSI-based fingerprinting for indoor localization: A deep learning approach, *IEEE Trans. Veh. Technol.*, vol.66, no.1, pp.763-776, 2017.

[22] J. Yoo, Change detection of RSSI fingerprint pattern for indoor positioning system, *IEEE Sens. J.*, vol.20, no.5, pp.2608-2615, 2020.

[23] K. Wang et al., Learning to improve WLAN indoor positioning accuracy based on DBSCAN-KRF algorithm from RSS fingerprint data, *IEEE Access*, vol.7, pp.72308-72315, 2019.

[24] D. Zhe, M. Chen and W. Liu, Implementation of indoor fingerprint positioning based on ZigBee, *Proc. of the 29th Chinese Control and Decision Conference (CCDC2017)*, pp.2654-2659, 2017.
[25] D. I. Nastac, F. A. Iftimie, O. Arsene and C. Cherciu, A statistical estimation analysis of indoor positioning WLAN based fingerprinting, *2017 IEEE 23rd International Symposium for Design and Technology in Electronic Packaging (SIITME)*, pp.239-242, 2017.
[26] R. F. Malik et al., The indoor positioning system using fingerprint method based deep neural network, *IOP Conf. Ser. Earth Environ. Sci.*, vol.248, no.1, 2019.
[27] S. Xia, Y. Liu, G. Yuan, M. Zhu and Z. Wang, Indoor fingerprint positioning based on Wi-Fi: An overview, *ISPRS Int. J. Geo-Information*, vol.6, no.5, p.135, 2017.
[28] R. Xu, W. Chen, Y. Xu and S. Ji, A new indoor positioning system architecture using GPS signals, *Sensors*, vol.15, pp.10074-10087, 2015.
[29] S. Yiu, M. Dashti, H. Claussen and F. Perez-Cruz, Wireless RSSI fingerprinting localization, *Signal Processing*, vol.131, pp.235-244, 2017.
[30] W. K. Zegeye, S. B. Amsalu, Y. Astatke and F. Moazzami, WiFi RSS fingerprinting indoor localization for mobile devices, *2016 IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp.1-6, 2016.
[31] A. R. Patel et al., *Indoor Location Survey*, US20150350845A1, 2014.
[32] J. Zhang, J. Wu, C. Qin and Z. Lai, *Rapid Indoor Wireless Signal Fingerprint Database Creation through Calibration of Ray-Tracing Propagation Model*, US9998928B2, 2014.
[33] D. I. Nastac, E. S. Lehan, F. A. Iftimie, O. Arsene and B. Cramariuc, Automatic data acquisition with robots for indoor fingerprinting, *2018 12th Int. Conf. Commun.*, pp.321-326, 2018.

Author Biography



Alvin-Ming-Song Chong obtained his B.Eng. (Hons.) Electronics majoring in Robotics and Automation from Multimedia University, Malaysia, in 2019.

Mr. Chong is currently pursuing his Master of Engineering Science (by Research) study at Multimedia University. His main research interest includes robotics, automation, and machine learning.

Boon-Chin Yeo obtained his Ph.D. in Engineering from Multimedia University, Malaysia, in 2017. The research topic involved background model, lane detection, traffic monitoring and control, fuzzy logic, and machine learning.

Dr. Yeo is a senior lecturer in the Faculty of Engineering and Technology, Multimedia University. Currently, he is also the Chairperson of the Centre for Advanced Robotics. He has been supervising numerous award-winning research projects related to robotics, condition monitoring, machine vision, and soft computing, especially for the applications in health care, automation, and robot positioning.

Way-Soong Lim received his Ph.D. degree (Engineering) from Multimedia University of Malaysia, in 2007. He joined the university in 1999 as a lecturer of Faculty of Engineering & Technology. He served as a Deputy Dean from 2008 to 2020, and was later appointed as the Director of International Relations of the University.

His research interest covers areas in artificial intelligence, neural network, pattern recognition, road traffic monitoring, machining-tools condition monitoring and robotics navigation. Along his career, he has published numerous international journals and conference papers, book chapters and engineering text books. Dr. Lim also maintains strong collaboration with many industrial organizations and international universities.