RSS-based Fingerprinting Localization with Artificial Neural Network

Kenny Fong Peng Wye, Syed Muhammad Mamduh Syed Zakaria, Latifah Munirah Kamarudin, Ammar Zakaria, Norhawati Binti Ahmad and Kamarulzaman Kamarudin

1 School of Microelectronic Engineering, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, Arau, Perlis, 02600, Malaysia
2 Centre of Excellence for Advanced Sensor Technology (CEASTech), Universiti Malaysia Perlis (UniMAP), Kompleks Pengajian Jejawi II, Arau, Perlis, 02600, Malaysia
3 School of Computer Engineering, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, Arau, Perlis, 02600, Malaysia
4 School of Mechatronic Engineering, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, Arau, Perlis, 02600, Malaysia

smmamduh@unimap.edu.my

Abstract. Radio Frequency (RF) based indoor is challenging due to the multipath effect in indoor signal propagation such as reflection, absorption, diffraction due to obstacles, interference and moving objects within the environments. The multipath effect phenomenon will be worsened if Received Signal Strength (RSS) is used as the localization measurement parameter. The advancement of Artificial Intelligence (AI) may hold the key for the improvement of RSS based localization. The Artificial Neural Network (ANN) in AI outperforms the traditional algorithms in indoor localization due to its capability to learn the unique features given in the training datasets. This paper discusses indoor fingerprinting localization with different architectures of ANN network to localize object of interest with the given indoor environment. The size of the experimental testbed is 14m x 8m and the testbed consists of four identical receivers that receive the signal from an active transmitter. All the tested ANN architectures have RSS as inputs and the Cartesian coordinates as outputs with different hidden layers and hidden nodes. The relationship between hidden nodes and layers in ANN and the regression losses is studied in this paper. The RSS-based fingerprinting with ANN in this paper is considered as a multi-output regression problem. The result shows that the ANN architecture with four layers with a total number of 800 hidden nodes has achieved an average of 6.01098 regression losses and a mean Euclidean Distance error of 2.54m.

1. Introduction
The researches in indoor localization have grown rapidly due to the increasing need for location-based applications in recent years. In the Radio Frequency (RF) based localization, there are four main measurement approaches which are the Time of Arrival (TOA), Angle of Arrival (AOA), Time
Difference of Arrival (TDOA) and Received Signal Strength (RSS) [1]. In particular, RSS based indoor localization has the advantages of minimum hardware support requirement and easily obtained from various devices such as WiFi Access Points [2,3], Bluetooth devices [4,5], Long Range devices (LORA) [6], Radio Frequency Identification (RFID) [7]. However, the RSS based indoor localization suffer from complex signal propagation in the indoor environment due to the obstacles, furniture, environmental changes and other signal propagation interference in the environment [8]. This results in fluctuation in the RSS signal data. In order to improve the localization accuracy in RSS measurement, RSS based fingerprinting localization is presented as one of the solutions to record the unique RSS presentation for every position in the testbed [9]. In general, the RSS based fingerprinting localization is performed by comparing the current RSS with the prebuilt RSS database in the offline phase. The location will be predicted based on the closest fit to the predefined data in the database.

2. Related work
A wide variety of researches and reviews have been done in the field of indoor localization. Compared to various methods introduced for indoor localization, fingerprinting localization is still gaining much attention in the field for its wide applicability in most unique indoor environment and sufficient localization accuracy. RADAR [10], the pioneer RF based fingerprinting indoor localization done by Microsoft Research that used K Nearest Neighbour (KNN) and Euclidean Distance as the location prediction measurement between the prediction location and the actual location. A typical fingerprinting technique, LANDMARC [11] works in a similar localization method like RADAR, the RF map of LANDMARC is prebuilt with the placement of active RFID tags in the testbed and the RSS from active tags are recorded as fingerprints for localization purposes. Both of these techniques subjected to environmental changes and difficult to design if the indoor environment is large.

In recent years, the advancement of AI developed a trend of indoor localization improvement with AI especially in Machine Learning (ML). The ML is a subfield of AI that learns through data and previous experiences make the indoor localization efficient and cost effective. Many indoor localization researches introduce ML models such as Support Vector Machines (SVM) [12,13], Random Forest [14], Decision Trees [15] and ANN to solve the intrinsic problems in indoor localization. In the particular of ANN fingerprinting localization, the RSS with Cartesian coordinates are trained in the neural network to adjust the weights and biases accordingly. The capability of ANN to localize in harsh manufacturing environments is also discussed in [16]. The uncertainty of RSS fingerprinting location measurements by the ANN and Gaussian Distribution approach is discussed in [17] and it is able to reduce the errors in a range of 23.3% to 32.3%. Meanwhile, Riya et al. [18] achieve 98% accuracy with ANN and a range of error estimation between 1.37 to 15.17 units in an area of 100 unit x 100 unit testbed. Although several researches have shown significant improvement by reducing the positioning error between the prediction and the true location of the object of interest, this paper would like to emphasize the effect of hidden layers and hidden nodes toward the accuracy of the location prediction ANN fingerprinting localization.

As most of the researches discussed, the RSS signal fluctuations are an intrinsic problem especially the presence of occupants and obstacles. In the previous work, the nature of indoor propagation has been studied in [19] and different algorithms proposed to ensure the reliability of signal transmission in both indoor and outdoor environments in [20,21]. While the results of proposed methods shown similar accuracies to methods from other researchers, further probing needs to be done to explore other modalities of RF-based localization. As part of continuing work, this paper presents the tracking capability of ANN approaches for fingerprinting localization in the indoor environment. The contribution of the work presented in this paper is the exploration of ANN performance in feature extraction in the collected RSS data in fingerprinting localization and the study of the relationship between the hidden layers and nodes towards the neural network regression losses and location prediction.
3. Experimental setup

In this paper, the RSS signal is collected with an active Radio Frequency Identification (RFID) tag throughout the experimental testbed and being trained in the ANN as a regression problem. In a regression problem, the ANN is being trained to minimize the error between the predicted output and the actual given output data. The RSS fingerprinting localization with ANN in this paper is considered as a multiple output regression where the output of the ANN gives the Cartesian coordinates. The predicted location by the ANN is then being compared to the actual active transmitter location and determined the error of the actual position and the prediction position.

In order to evaluate the performance of ANN architecture, the experiment is in UniMAP Solutions Laboratory in Universiti Malaysia Perlis (UniMAP). The size of 14m x 8m within the laboratory is chosen as the testbed for the experiment. Four identical 15dBi omnidirectional antennas connected with each of the 2.45GHz active RFID readers. The antennas are mounted on the walls at 1.4m height from ground and the height placement of the transmitter (RFID tag) stands at approximately 1.35m height. There are total of 135 reference positions located at 1 meters apart. Each position is labelled based on the Cartesian coordinates for the ANN training input. The floor plan of the laboratory is shown in Figure 1.

![Figure 1. Experimental testbed of 14m x 8m](image)

Although the signal reflected by the surrounding walls are unavoidable, the testbed is maintained at the optimum level for the RSS fingerprinting localization. All the obstacles such as tables, chairs are removed to reduce the multipath effect that may be generated from the obstacles. The active transmitter is placed at each reference position to collect 2 minutes data with sample rate of 1Hz. There is a total of 16200 training data for the ANN architecture from the 135 reference positions.

The mean of the collected sample data for every reference position is used as the radio map RSS for each of the receivers respectively. The radio map RSS to observe the signal pattern is plotted in the contour pattern with the arrangement as shown in Equation (Eq.) 1.

\[
\text{RSS}_{\text{radio map mean}} = \begin{bmatrix}
\text{RSS}_{0,0} & \cdots & \text{RSS}_{x,0} \\
\vdots & \ddots & \vdots \\
\text{RSS}_{0,y} & \cdots & \text{RSS}_{x,y}
\end{bmatrix} \quad \text{Eq.1}
\]
3.1. Artificial Neural Network (ANN) model optimization

Artificial Neural Network (ANN) consists of perceptron that inspired by human brains to learn and adapt any given input by adjusting its weights and biases by gradient descent and achieve its desired output classification with different loss functions. The designed ANN architecture is shown in Figure 2. The RSS signals at every location collected from four receivers are applied to the input of the ANN architecture. The output layer of the ANN architecture has two nodes that represent the Cartesian coordinates of the location (x,y) respectively. The hidden layers and hidden nodes of the ANN architecture is then increased to observe the effect of different hidden layers and hidden nodes toward the model losses for location prediction.

![Figure 2. Multi output neural network architecture](image)

The prediction of location coordinates (x,y) is calculated by the optimum weights adjusted by the neural network and the RSS signal inputs with activation function at every layer. The estimated position by the neural network is obtained by Eq. 2.

\[ \text{Pred. Location}_{x,y} = f_{output}\left( \sum (w_{hidden})(f_{hidden}) \sum (RSS_i)(w_{input}) \right) \]  

**Eq. 2**

Where the RSS_i refers to the signal strength from each receiver respectively, f indicates with activation function in the different layers respectively, w represents the weights and RSS as the input for the neural network.

The application of ANN in fingerprinting localization that consists of a considerable amount of data may cause overfitting issue in machine learning. One of the solutions to avoid overfitting issue is to split the dataset into training data and validation data. It is common to split the dataset into 80% of training data and 20% of validation data to evaluate the designed model is being overfitted. The dropout method can be applied in the ANN architecture that randomly dropout of the hidden nodes to avoid overfitting. The dropout layers act as an optimization method to reduce the overfitting where random nodes are temporarily removed in order to improve the “regularize” a fixed sized model [22]. In this work, the dropout of each hidden layer in the network is set at 0.5 (50%).

This paper introduces the Min-Max normalization to normalize the raw RSS data to a zero mean, unit variance new normalized RSS value in each of the receivers respectively. The normalization performed by Min-Max normalization can be calculated as shown in Eq. 3.
Result and discussion

The experiment starts with the collection of RSS value throughout the 135 labelled reference points in the testbed. Since the RSS of the active transmitter is sampled at every second, multiple RSS data at a reference position are collected to increase the data robustness during the neural network training session. The boxplots graph that represents the RSS distribution throughout the testbed in each receiver is plotted in Figure 3.

![Boxplots graph for collected RSS in each receiver](image)

**Figure 3.** Boxplots graph for collected RSS in each receiver

Based on the boxplots in Figure 3, the RSS fluctuates within the range of -50dB to -100dB for all the reference locations within the testbed with the furthest outliers of -128dB. Since the RSS is sampled at every second, missing sampled RSS data during the experiment is unavoidable. The missing RSS data is considered as -128dB which is out of the RSS range of the received signal manual. This is due to the multipath effects discussed in the literature and signals reflection by the surrounding walls are unavoidable in the environment. Although collected data shows some unexpected RSS peaks during the experiment, the outliers are not removed as they may carry important RSS unique features to represent the location.

The mean of the collected RSS at every reference position of each receiver is used to demonstrate the RSS signal propagation in the indoor environment, the radio maps for each of the receivers with Eq. 1 are plotted as shown in Figure 4(a)-(d) respectively.
Figure 4. Contour map for each receiver (a) RSS radio map of receiver 1 located at (0,0) (b) RSS radio map of receiver 2 located at (0,8) (c) RSS radio map of receiver 3 located at (14,0) (d) RSS radio map of receiver 4 in (14,8)

Although identical receivers are used in the different positions in the same testbed, the RSS signal propagation for each of the receiver show the uniqueness to represent the location as shown in Figure 4. These results shown that the signal propagation model to determine the transmitter location is not suitable as the receivers would have unique RSS representation at different locations. The offline database is established after all the 135 reference positions’ RSS have been recorded. The raw collected RSS data are then normalized with Eq. 3 to obtain the new normalized RSS value with zero mean, unit variance between 0 and 1. In order to train the neural network, the single hidden layer and 20 hidden nodes ANN architecture is used as the reference for the experiment.

The nodes of ANN architecture are then increased to 50, 100, 150, 200 respectively. By performing these experiments, the relationship of hidden nodes and the regression losses of the ANN is obtained. The experiments continue with increasing the hidden layer from a single layer until five hidden layers for every different number of nodes mentioned to obtain the relationship of hidden regression losses and hidden layers. The results of regression losses against different hidden nodes and different hidden layers in ANN is plotted in Figure 5.
Based on Figure 5, both of the graphs show different relationships between the losses, hidden layers, and nodes. In Figure 5(a), the results have shown improvement in the decrement of regression losses at the first few layers but the losses increased after the optimum layers for hidden nodes are reached. In contrast with the work in [23], the result obtained in this paper indicated that deeper ANN architecture does not indicate a guaranteed improvement in localization accuracy. As the neural network goes deeper, the more complicated calculation is performed by the neural network to adjust the weights and biases to extract the feature within the data supplied. Furthermore, the regression losses decreased with increment of hidden nodes in a constant layer experiment as shown in Figure 5(b). As more hidden nodes in a hidden layer are introduced, more feature extracted can be obtained from the data supplied without complicated calculation pass through from layers to layers in a deeper network.

The three lowest regression losses obtained from the experiments shown in Figure 5 fall under 200 hidden nodes in two, three and four layers ANN architecture respectively. After the neural network models are being trained with the offline database, the active transmitter is then placed at random available fingerprinted locations to evaluate the error of the ANN predicted location and the actual location. There are total of 10 random locations with the collected RSS is predicted with the best regression loss ANN architecture. The best ANN architecture with four hidden layers and total of 800 hidden nodes achieved the best averages losses of 6.01098.

The best pretrained ANN model is then loaded and the location prediction based on collected RSS for 10 random locations are performed. The Euclidean Distance (ED) error between the predicted location and the true location of the active transmitter is then tabulated in Table I that calculated with Eq. 4.

\[ ED \text{ Error (m)} = \sqrt{(X_{pred} - X_{true})^2 + (Y_{pred} - Y_{true})^2} \]  
\text{Eq.4}
Table 1. Tabulated error in localization prediction

| Actual X | Actual Y | Predicted X | Predicted Y | Error in X | Error in Y | ED Error (m) |
|----------|----------|-------------|-------------|------------|------------|--------------|
| 0        | 3        | 1.24        | 2.58        | 1.24       | -0.42      | 1.309198     |
| 2        | 6        | 2.58        | -0.42       | -3.2       | 0.22       | 3.227445     |
| 4        | 4        | 4.22        | -0.98       | 0.22       | 0.340588   |
| 6        | 2        | 2.26        | 0.22        | -0.26      | 0.340588   |
| 7        | 4        | 3.39        | -0.79       | -0.61      | 0.998098   |
| 8        | 7        | 3.19        | -4.81       | 2.56       | 5.448826   |
| 10       | 7        | 9.52        | -0.48       | -2         | 2.056794   |
| 12       | 1        | 10.69       | -1.31       | 1.01       | 1.654146   |
| 13       | 2        | 11.57       | -1.43       | 0.76       | 1.619413   |
| 13       | 8        | 10.37       | -6.13       | -4.76      | 7.761089   |

Based on Table 1, the ED error of localization with four hidden layers of 200 hidden units in each layer ANN architecture falls between 0.34m to 7.76m. The mean error of the ANN architecture is obtained by averaging the error in 10 locations available in Table 1. The findings in this paper demonstrated that the potential of ANN to be used reliably with fingerprinting localization with an average of 6.01098 regression losses and a mean error of 2.54m. In contrast with the work in [24], author shown improvement in ANN performance when the anchor locations are increased. The results have shown a 2.35m average error with 5 anchor location with WLAN localization with single hidden layers neural network. In this paper, we studied the relationship between the hidden layers and nodes towards the neural network regression losses and location prediction. Furthermore, it was observed that the ANN managed to capture the unique feature of RSS in every position without extensive programming required.

4. Conclusion
The work presented in this paper has managed to demonstrate the feasibility of ANN in fingerprinting localization. Although the RSS fluctuates in the indoor environment is observed as expected, the capability of ANN to extract features from the RSS and learn the uniqueness of each fingerprinting reference position. Various results are obtained with different ANN architecture and shown the fingerprinting localization with ANN is able to achieve the desired performance with the RSS signal. The optimum result obtained is from the four layers and a total number of 800 hidden nodes ANN that achieved an average of 6.01098 regression losses and a range between 0.34m to 7.76m localization error. In future planning, the main focus will be toward the minimization of the regression losses with different ANN architecture that would be further explored.

Acknowledgement
The authors would like to acknowledge the support from Malaysian Technical University Network (MTUN) grant (9002-00094 & 9028-00002) under the project of Development and Deployment of Smart Community and Innovation Centre for Smart City 4.0.

References
[1] Alarifi A, Al-Salman A, Alsaleh M, Alnafessah A, Al-Hadhrami S, Al-Ammar M A and Al-Khalifa H S 2016 Ultra wideband indoor positioning technologies: Analysis and recent advances Sensors (Switzerland) 16 1–36
[2] Xue W, Qiu W, Hua X and Yu K 2017 Improved Wi-Fi RSSI Measurement for Indoor Localization IEEE Sens. J. 17 2224–30
[3] Jang J W and Hong S N 2018 Indoor Localization with WiFi Fingerprinting Using Convolutional Neural Network Int. Conf. Ubiquitous Futur. Networks, ICUFN 2018-July 753–8
[4] Chai S, An R and Du Z 2016 An Indoor Positioning Algorithm Using Bluetooth Low Energy RSSI
276–8

[5] Hou X and Arslan T 2018 Monte Carlo localization algorithm for indoor positioning using Bluetooth low energy devices 2017 Int. Conf. Localization GNSS, ICL-GNSS 2017 1–6

[6] Islam B, Islam M T and Nirjon S 2017 Feasibility of LoRa for Indoor Localization on-Line, From Semant. Org 1–11

[7] Zhang J, Lyu Y, Patton J, Periaswamy S C G and Roppel T 2018 BFVP: A Probabilistic UHF RFID Tag Localization Algorithm Using Bayesian Filter and a Variable Power RFID Model IEEE Trans. Ind. Electron. 65 8250–9

[8] Hsieh C H, Chen J Y and Nien B H 2019 Deep Learning-Based Indoor Localization Using Received Signal Strength and Channel State Information IEEE Access 7 33256–67

[9] Xia S, Liu Y, Yuan G, Zhu M and Wang Z 2017 Indoor Fingerprint Positioning Based on Wi-Fi: An Overview ISPRS Int. J. Geo-Information 6 135

[10] Bahl P and Padmanabhan V N 2000 RADAR: An in-building RF-based user location and tracking system Proc. - IEEE INFOCOM 2 775–84

[11] Hu B, Peng H and Sun Z 2018 Landmarc localization algorithm based on weight optimization Chinese J. Electron. 27 1291–6

[12] Chiki A, Touati H and Snoussi H 2017 SVM-based indoor localization in Wireless Sensor Networks 2017 13th Int. Wirel. Commun. Mob. Comput. Conf. IWMC 2017 1144–9

[13] Sallouha H, Chiumento A and Pollin S 2017 Localization in long-range ultra narrow band IoT networks using RSSI IEEE Int. Conf. Commun.

[14] Guo X, Ansari N, Li L and Li H 2018 Indoor Localization by Fusing a Group of Fingerprints Based on Random Forests IEEE Internet Things J. 5 4686–98

[15] Sánchez-Rodríguez D, Hernández-Morera P, Quinteiro J M and Alonso-González I 2015 A low complexity system based on multiple weighted decision trees for indoor localization Sensors (Switzerland) 15 14809–29

[16] Gholami M, Cai N and Brennan R W 2013 An artificial neural network approach to the problem of wireless sensors network localization Robot. Comput. Integr. Manuf. 29 96–109

[17] Li Y, Gao Z, He Z, Zhuang Y, Radi A, Chen R and El-Sheimy N 2019 Wireless fingerprinting uncertainty prediction based on machine learning Sensors (Switzerland) 19 1–18

[18] Samanta R, Kumari C, Deb N, Bose S, Cortesi A and Chaki N 2018 Node localization for indoor tracking using artificial neural network 2018 3rd Int. Conf. Fog Mob. Edge Comput. FMEC 2018 229–33

[19] Turner J S C, Kamarudin L M, Ndzi D L, Harun A, Zakaria A, Shakkaf A Y M, Saad A R M and Mamduh S M 2011 Modelling indoor propagation for WSN deployment in smart building 2014 2nd Int. Conf. Electron. Des. ICED 2014 398–402

[20] Kamarudin L M, Ahmad R B, Ndzi D, Zakaria A, Ong B L, Kamarudin K, Harun A and Mamduh S M 2012 Modeling and simulation of WSNs for agriculture applications using dynamic transmit power control algorithm Proc. - 3rd Int. Conf. Intell. Syst. Model. Simulation, ISMS 2012 616–21

[21] Mamduh S M, Shakkaf A Y M, Saad S M, Kamarudin K, Kamarudin L M, Zakaria A, Kamarudin H, Ezaanuddin A M M, Sa’Ad F S A, Nooriman W M and Abdullah A H 2012 Odour and hazardous gas monitoring system for swiftlet farming using wireless sensor network (WSN) Chem. Eng. Trans. 30 331–6

[22] N Srivasta, G Hinton, A Krizhevsky, I Sutskever R S 2014 Dropout: A Simple Way to Prevent Neural Networks from Overfitting Mach. Learn. Res. 15 299 345–50

[23] Félix G, Siller M and Álvarez E N 2016 A fingerprinting indoor localization algorithm based deep learning Int. Conf. Ubiquitous Futur. Networks, ICUFN 2016-Augus 1006–11

[24] Saleem F and Wyne S 2016 Wlan-Based Indoor Localization Using Neural Networks J. Electr. Eng. 67 299–306