IOP Conference Series: Earth and Environmental Science

PAPER • OPEN ACCESS

WLAN Based Position Estimation System Using Classification Fuzzy K-Nearest Neighbor (FK-NN)

WLAN Based Position Estimation System Using Classification Fuzzy K-Nearest Neighbor (FK-NN)

R F Malik*[1], Mardiah[1], A Farissi[1], D Stiawan[1], R Zulfahmi[1], M R Ahmad[2] and A S Khirbeet[2]

[1]Communication Network and Information Security Research Lab, Faculty of Computer Science, Universitas Sriwijaya Jalan Srijaya Negara Bukit Besar, Bukit Lama, Ilir Bar. I, Kota Palembang, Sumatera Selatan 30128 - Indonesia
[2]Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Jalan Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

rezafm@unsri.ac.id*

Abstract. Increasing the number of public hotspots using Wi-Fi technology is one of opportunity to gain advantage for proposing many new technologies. One of emerging technology is an estimation system to locate the object/person position using Wi-Fi. The object estimation position is the technology to estimate object position accuracy, using signal Received Signal Strength (RSS) from Wi-Fi Access Point. The RSS is an information about the strength of the signal indicates the distance between the access point device. Through the Indoor Positioning System (IPS), RSS value information from multiple access points are processed in order to provide position information. In this study, the IPS using Fuzzy K-Nearest Neighbour (FK-NN) classification method which is a combination of Fuzzy algorithm and K-NN to increase the accuracy of the object estimation position based on the learning data as reference point. Through hybridization from the algorithm is expected to calculate the position estimation more effectively and accurately and minimize errors in estimation. The results show that the algorithm FK-NN obtain the average location error of 2.4 meters with an accuracy percentage of 76%.
Keywords: Indoor Positioning System, Fuzzy K-Nearest Neighbour

1. Introduction

The Indoor Positioning System (IPS) utilize wireless technology to find objects in the building via RSS values from access point devices. Currently, several IPS applications have implement in the the hospital for navigation system, employee monitoring, guiding the blind, tracking small children and the elderly [1]. Existing technology such as the Global Positioning System (GPS) for navigation system by using signals from satellites. However, GPS is less effective when used in the indoor environment. The drawback because the weak signals received from the satellite by the GPS device is hard to penetrate the building structure[2]. Because of limitation, alternative solution using Indoor Positioning System (IPS) able to utilize Wi-Fi signal in each building. Wi-Fi signal sources can received from various wireless devices in the indoor environment such as offices, hospitals, schools, colleges and other buildings [1]. Signals broadcast from access points have an information as known as the Receive Signal Strength (RSS) in dBm. Thus, the RSS as an information base to provide the signal strength values. The RSS values from several access points are processed to provide position information of the objects.

In a previous study by Houria and Chami Chabbar Mouhcine [2] using a wireless signal as a reference for position estimation using fingerprint method to calculate the euclidean distance and determine the user's position. They can achieved 2 meters accuracy. Dong Li, Baoxian Zhang and Cheng Li [3] proposed the estimated position using the wireless signal strength with FS-KNN classification method (Scaling Feature-based K-Nearest Neighbor). The accuracy improvement can achived average 1.70 meters accuracy. Mathematical modeling of Bayesian methods has been carried out by N. Wiranata and R. F. Malik [4] for the IPS system.

In this paper, we use the fingerprint algorithm using Fuzzy classification and K-Nearest Neighbor (FK-NN). FK-NN classification method is a combination of Fuzzy Logic algorithm and K-NN to increase the accuracy of the estimate of the object position based on RSSI values.

2. The Indoor Position System

The Indoor Positioning System (IPS) is a technology to determine the location of an object or a person inside a building by using radio waves, magnetic fields, acoustic signals or other sensor information collected from wireless devices. IPS is an alternative solution to solve the drawback of GPS location service. IPS utilizing different wireless technologies such as such as Wi-Fi Access Point (AP), magnetic positioning and dead reckoning, to measure a current object position from the nearest reference points. [5]

2.1 Fingerprint Method

Applications and Location-based Service make location as basic information. Localization technique being important element at positioning application [6]. Currently there are many localization and navigatin system techniques designed and implemented in the previous study. One of the approach is collecting the Received Signal Strength (RSS) value for estimating the object location at indoor or outdoor environment. Information brings by Signal Strength contain the degree of signal receive by receiver devices from transmitter. Utilization of RSS can be used to estimate the distance between the transmitter and receiver through the fingerprint method. In fingerprint method, RSS values from several places will be stored at database and the points were collected called reference points [7]. Thus, the unknown location can be estimated by finding a pattern match between the existing fingerprint and the previous fingerprint from RSS measurement [8]. Basically, fingerprint method consists of 2 phases, Online and Offline phases [9]:

2.1.1 Offline Phase

In Offline phase a collection of RSS values from the access point in every reference points are collected into a database. [10]. The collection of fingerprints for all known locations as known as reference points are called radio map [11]. The value from radio map is fluctuative taken from measurements and set as the radio fingerprint for every reference points.

2.1.2 Online Phase

If offline phase has purpose to build an empirical training database at every reference points. Incontrary, data were compared with the existing RSS and matched in order to estimate the object position [9] [12]. The process is named as online phase. This online phase will receive real time data from access points. The data will compare to get a pattern and predict the object location. The Pattern-Matching Algorithm using Fuzzy K-Nearest Neighbor (FK-NN) to compare and find a match among the RSS values between offline and online phases[13]. Overall procedure of indoor position system is carried out by linear or non-linear mapping F : $R^N \rightarrow R^2$ [10].

2.2 Fuzzy K-Nearest Neighbor (FK-NN)

Fuzzy K-Nearest Neighbor (FK-NN) is a combination of Fuzzy and K-Nearest Neighbor[14]. In the fuzzy theory, the data has a membership value. Thus, the K-NN will use for predicting based on the data testing and put the k-nearest neighbors ratio as an input to calculate the membership value [5].

The data are collected at offline phase will use as training data. In the online phase, object position request by user to get an information accurately. To obtain an accurate position, the RSS values at offline phase will compare and find the pattern based on the closest RSS value among data in offline and online phases. The shortest distance is obtained by comparing the testing data and training data using equation euclidean distance as follows:

$$\mathbf{d}\ (x_i, x_j) = \sqrt{\sum_{r=1}^{n} (x_i - x_j))^2} \tag{1}$$

Where :

d (x_i, x_j) = Distance Euclidean,

x_i = Test data

x_j = Data training

In the FK-NN, euclidean distance value using the fuzzy as an input for position estimation. The Fuzzy logic procedure is described as follows:

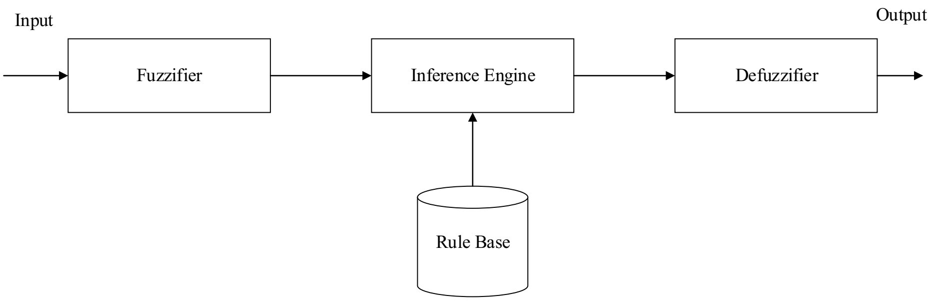


Figure 2.1 Fuzzy Logic Systems [15]

2.2.1 Fuzzifier

The triangle curve, as shown in Figure 2.2, is basically a combination of the two lines.The range of variable input value for distance is divided into five linguistic values. The linguistic values used in this research are: VeryClose, Close, Medium, Far and VeryFar.

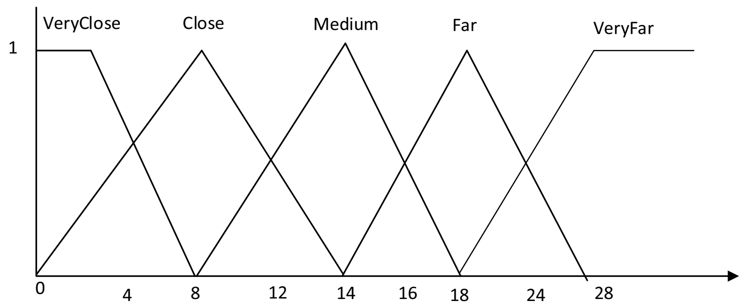


Figure 2.2 Variable Input for Distance[15]

The Variable Output value for weight is divided into five linguistic values. VeryLarge (VL), Large (L), Medium (M), Small (S) and VerySmall (VS) are the linguistic values in term of weight description. The weight value graph for the decision variables are shown in Figure 2.3.

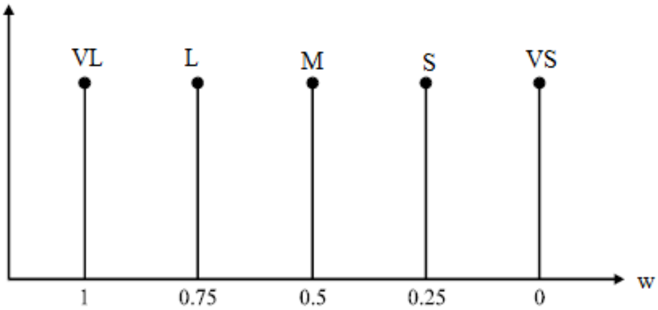


Figure 2.3 Output Membership Function[15]

2.2.2 Fuzzy Rules

Fuzzy Rules are applied and have five rules to give a decision framework. The fuzzy rules will determine the object position based the data input. These are the fuzzy rules as follows:

1. If (distance VeryClose) then (weight value VeryLarge)
2. If (distance Close) then (weight value Large)
3. If (Medium distance) then (Medium weight value)
4. If (distance Far) then (Small weight value)
5. If (distance VeryFar) then (weight VerySmall)

2.2.3 Defuzzyfication

Defuzzification process is taking an weight average using equation 2.

$$Z^* = \frac{\sum y\, \mu_r(y)}{\sum \mu_r(y)}$$

$$Z^* = \frac{(1*\mu_{VeryClose})+(0{,}75*\mu_{Close})+(0{,}5*\mu_{Medium})+(0{,}25*\mu_{Far})+(0*\mu_{VeryFar})}{(\mu_{VeryClose})+(\mu_{Close})+(\mu_{Medium})+(\mu_{Far})+(\mu_{VeryFar})} \quad (2)$$

Where :

Z * = output inference engine

The Figure 2.4 is floor plan with access points surrounding the building. For example, the person/object is in the room 1 and will detect one nearest access point as a nearest neighbor (1-NN). Thus, the FK-NN method will estimate the object/person location and will identify the room 1 is corrected location of object/person. The Figure 2.5 shown three access points are involved for object/person detection and act as a nearest neighbor (3-NN). Thus, the result can be classified the object/person positioning as the nearest neighbour (1-NN).

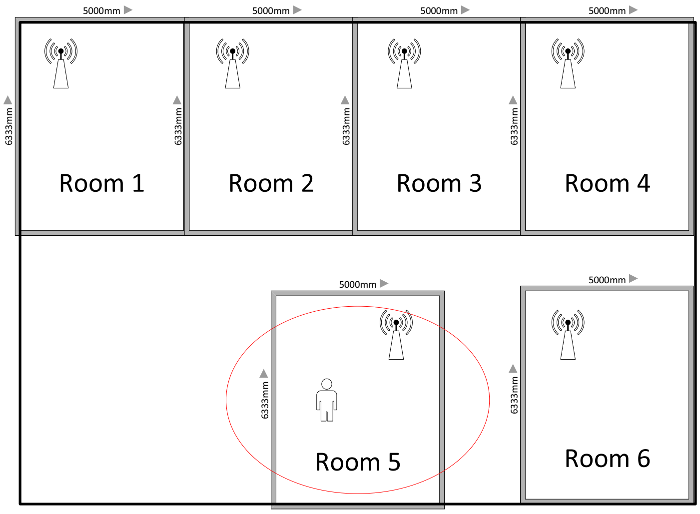


Figure 2.4 Example 1 for FK-NN with Nearest Neighbor (1-NN) [5]

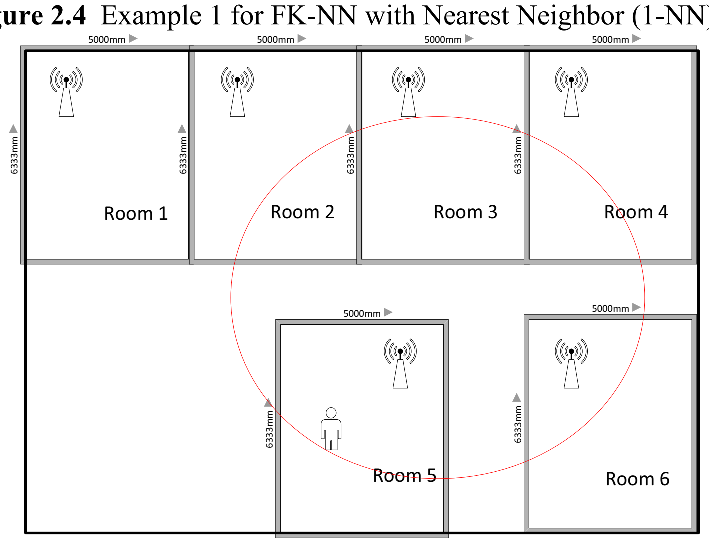


Figure 2.5 Example 3 for FK-NN with Nearest Neighbor (3-NN) [5]

3. Research Methodology

The laptop and access points are required for running the experimental scenarios. The laptop act as client and server then access points are intermediate devices. The signal from the access points as RSS values will send to the server. While NetSurveyor is software to record data in RSSI values.

In this paper, the floor plan use the building D, Faculty of Computer Science, Universitas Sriwijaya as shown in Figure 3.1. The building D consists of four rooms: D.1.1, D.1.2, D.1.3 and D.1.4. There are 67 reference points set as point to collect data for certain period. The spacing between reference points are 2 meters. Total of access point surrounding the building are 6 units and will obtain for training data. Based on MAC Address, the dataset will collect from access points are 402 data.

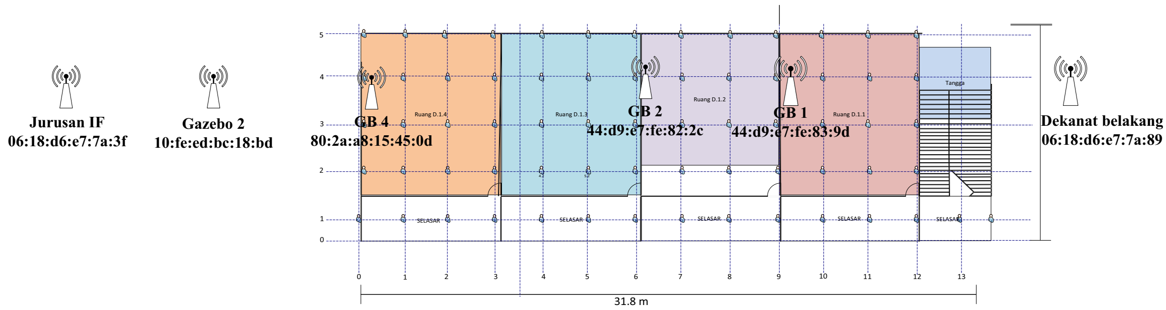


Figure 3.1 Floor plan for Building D Faculty of Computer Science

As mention in section 2, the method is divided into two sphases: offline and online. For offline phase, training data is done by fingeprint access point at any point of reference. The data property is the signal strength in dBM from access points. Training data is obtained and will input into a csv format then the data processed manually. Each data has a label for identification.

In online phase, request location will perform using the testing data from access points and compare it with the training data. The matching process using FK-NN classification and the result is object/person position.

Figure 3.2 shown the FK-NN flowchart to depict the the indoor positioning system. The system starts from the user retrieve the data from access points. The scan results are used as a testing data and proceed calculate to get classification from training data using FK-NN. Thus, the results predict the client positions based on the classification.

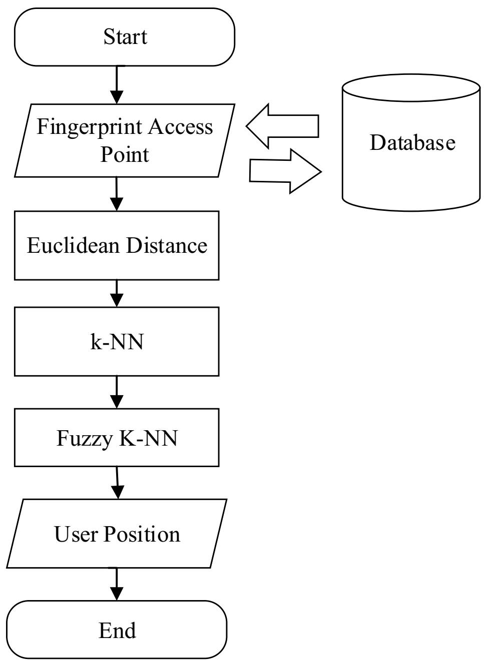


Figure 3.2 FK-NN Classification Flowchart

During classification process, the data are calculated the distance using the euclidean distance equation. Thus, the results are determining the euclidean distance from K-NN and proceed as input in a fuzzy system. The K-NN membership value will obtain as the linguistic variables. Defuzzification is next process after get the linguistic variables. The biggest weight value will serve as the estimated position of the object in the defuzzification process.

4. Results and Discussions

4.1 *Modeling Classification of Fuzzy K-Nearest Neighbor (FK-NN)*
Here is an position estimation have been done using FK-NN. The first step is the training data as input will test the data. The total consists of 15 training data and 1 testing data from 6 from access points. Table 4.1 is a sample of training data and Table 4.2 is a sample testing data.

Table 4.1 Sample Training Data

No.	Reference point [X, y]	06:18: d6:e7: 7a: 89 (dBm)	44:d9: e7: fe: 83: 9d (dBm)	44:d9: e7: fe: 82: 2c (dBm)	80:2a: a8:15: 45: 0s (dBm)	10: fe: ed:bc: 18: bd (dBm)	06:18: d6:e7: 7a: 3f (dBm)	Room
1	0.1	-95	-100	-100	-62	-66	-75	Terrace GB
2	4.1	-66	-100	-100	-73	-100	-76	Terrace GB
3	6.1	-69	-95	-55	-88	-82	-100	Terrace GB
4	3.2	-74	-89	-68	-55	-73	-70	D.1.4
5	3.3	-77	-87	-71	-54	-68	-75	D.1.4
6	3.4	-95	-100	-70	-67	-95	-78	D.1.4
7	6.2	-77	-95	-62	-68	-80	-86	D.1.3
8	6.3	-81	-76	-62	-73	-80	-100	D.1.3
9	6.4	-74	-69	-58	-95	-76	-72	D.1.3
10	9.2	-77	-59	-58	-84	-83	-84	D.1.2
11	9.3	-69	-73	-52	-100	-100	-82	D.1.2
12	9.4	-75	-67	-53	-78	-87	-78	D.1.2
13	12.2	-77	-67	-68	-100	-100	-100	D.1.1
14	12.3	-81	-60	-72	-100	-91	-100	D.1.1
15	12.4	-73	-60	-72	-100	-100	-100	D.1.1

Table 4.2 Sample Testing Data

06:18:d6: e7:7a:89(dBm)	44:d9:e7: fe:83:9d (dBm)	44:d9: e7:fe:82:2 c(dBm)	80:2a:a8: 15:45:0s (dBm)	10:fe:ed: bc:18:bd (dBm)	06:18: d6:e7:7a: 3f (dBm)	Room
-67	-55	-62	-100	-89	-100	unknown

After determining the training data and test data, the second step is calculate the Euclidean distance between the testing data and training data using equation 1. This stage, euclidean distance process is carried out on all training data records on database using equation 1. In Table 4.3 shown as the results from the shortest distance calculation between the testing data and all the training data.

Table 4.3 The Closest Distance Calculation Results

No.	Reference point [x, y]	Distance Data Testing Data into Training	Target class
1	12.4	16.793	D.1.1
2	12.3	18.028	D.1.1
3	12.2	20.025	D.1.1
4	9.2	26.077	D.1.2
5	9.3	29.547	D.1.2
6	6.1	32.281	Terrace GB
7	6.4	35.2	D.1.3
8	9.4	35.511	D.1.2
9	6.3	38.04	D.1.3
10	6.2	54.782	D.1.3
11	3.3	66.235	D.1.4
12	3.2	66.499	D.1.4
13	3.4	66.948	D.1.4
14	4.1	69.972	Terrace GB
15	0.1	82.771	Terrace GB

The next step is to determine the value of the data k nearest euclidean distance as shown in Table 4.3. The data were taken with the smallest value of k. For k = 3, table 4.4 shown distance calculation based on Equation 1.

Table 4.4 Three Nearest Neighbor (3-NN)

No.	Reference points [x, y]	Distance (k)	Target
1	12.4	16.793	D.1.1
2	12.3	18.028	D.1.1
3	12.2	20.025	D.1.1

The fourth step is the 3-NN Fuzzification membership using the equation 2. The following is fuzzification calculation:

$$\mu_{medium}\,[16{,}793] = \frac{18 - 16{,}793}{18 - 14} = \frac{1{,}207}{4} = 0{,}3$$

$$\mu_{far}\,[16{,}793] = \frac{16{,}793 - 14}{18 - 14} = \frac{2{,}793}{4} = 0{,}7$$

$$\mu_{far}\,[18{,}028] = \frac{28 - 18{,}028}{28 - 18} = \frac{9{,}972}{10} = 0{,}9$$

$$\mu_{Veryfar}\,[18{,}028] = \frac{18{,}028 - 18}{28 - 18} = \frac{0{,}208}{10} = 0{,}02$$

$$\mu_{far}\,[20{,}025] = \frac{28 - 20{,}025}{28 - 18} = \frac{7{,}975}{10} = 0{,}8$$

$$\mu_{Veryfar}\,[20{,}025] = \frac{20{,}025 - 18}{28 - 18} = \frac{2{,}025}{10} = 0{,}2$$

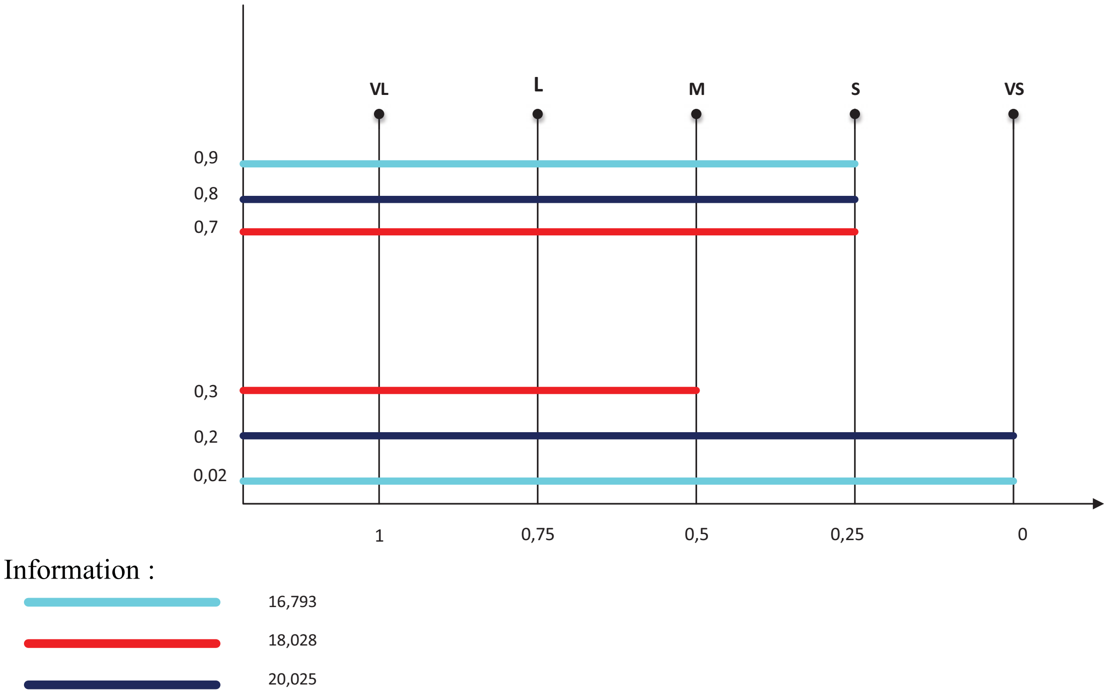


Figure 4.1. Fuzzification Membership

Next stage is the defuzzification using equation 2. Weight value varies from 0 to 1. The further the reference point and impact to the smaller the weight value in accordance with the rules. The Fuzzy K-Nearest Neighbor (FK-NN) rules as follows:

1. If (distance VeryClose) then (weight value is VeryLarge)
2. If (distance Close) then (weight value is Large)
3. If (Medium distance) then (weight value is Medium)
4. If (distance Far) then (weight value is Small)
5. If (distance VeryFar) then (weight value is VerySmall)

Table 4.5 Inference Output 1

Distance (k)	**[16.793]**
Euclidean	Weight
Very Close (VC)	Very Large (VL)
Close (C)	Large (L)
0.3	0.5
0.7	0.25
Very Far (VF)	Very Small (VS)

$$Z^* = \frac{0{,}3 \text{ x } 0{,}5 + 0{,}7 \text{ x } 0{,}25}{0{,}3 + 0{,}7} \frac{0{,}15+0{,}175}{1} = \frac{0{,}325}{1} = 0{,}325$$

Table 4.6 Inference Output 2

Distance (k)	**[18.028]**
Euclidean	Weight
Very Close (VC)	Very Large (VL)
Close (C)	Large (L)
Medium (M)	Medium (M)
0.9	0.25
0.02	0

$$Z^* = \frac{0{,}9 \text{ x } 0{,}25 + 0{,}02 \text{ x } 0}{0{,}9+0{,}02} = \frac{0{,}225}{0{,}92} = 0{,}24$$

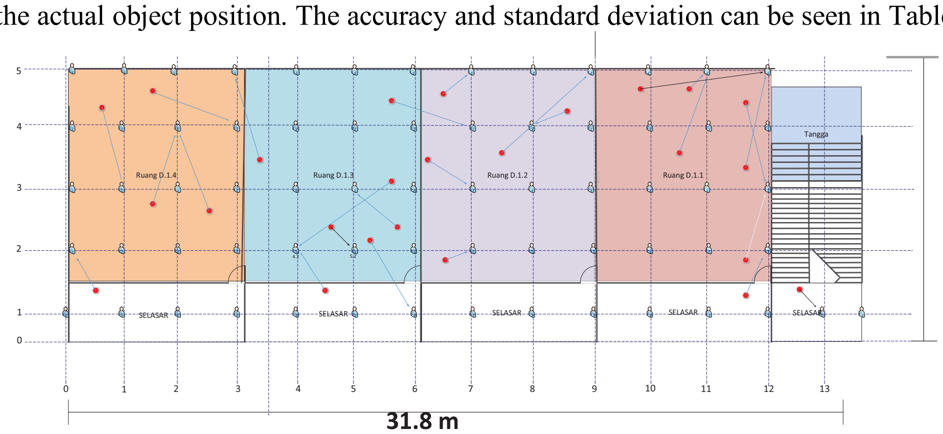
Table 4.7 Inference Output 3

Distance (k)	**[20.025]**
Euclidean	Weight
Very Close (VC)	Very Large (VL)
Close (C)	Large (L)
Medium (M)	Medium (M)
0.8	0.25
0.2	0

$$Z^* = \frac{0{,}8 \text{ x } 0{,}25+0{,}2 \text{ x } 0}{0{,}8 + 0{,}2} \frac{0{,}2}{1} = 0{,}2$$

From the above calculation results obtained three estimates the weight value for determining the object position. The K-NN data with the biggest weight is 0.325 and point to target is D.1.1 room with reference point coordinate is (12,4).

In Figure 4.2 shown the result of testing data from WLAN based position estimation system using Fuzzy K-Nearest Neighbor classification algorithm. Testing was performed using 402 training data and 25 testing data using a value of k = 3. The results obtained with the estimated object position information is accurate or not with the actual object position. The accuracy and standard deviation can be seen in Table 4.8.



: Reference Point
: Estimated Position

Figure 4.2 The Object Estimation Position Results

Table 4.8 Estimation Position and Accuracy

Test Point	Estimated Location	Accuracy (Meters)	Information
1	12.5	4.2	Corresponding
2	11.5	0.95	Corresponding
3	12.5	2.7	Corresponding
4	11.5	2.2	Corresponding
5	12.3	2.7	Corresponding
6	12.3	1.5	Corresponding
7	12.2	1.3	Not Corresponding
8	13.1	1	Corresponding
9	7.5	1.7	Corresponding
10	9.5	4.4	Corresponding
11	8.4	1.7	Corresponding
12	7.3	1.6	Corresponding
13	7.2	1	Corresponding
14	4.2	1.2	Not Corresponding
15	4.2	4.6	Corresponding
16	5.2	0.95	Corresponding
17	5.3	2	Corresponding
18	6.1	4	Not Corresponding
19	7.4	3.6	Not Corresponding
20	0.2	1	Not Corresponding
21	1.3	3.3	Corresponding
22	2.4	3.1	Corresponding
23	2.4	3.2	Corresponding
24	3.4	3.3	Corresponding
25	3.5	3.7	Not Corresponding
Average	2.436		
Standard Deviation	1.225299283		

Based on Table 4.8, the results show that the algorithm FK-NN obtain the average location error of 2.4 meters with an accuracy percentage of 76%.

5. Conclusion

In this paper, the experimental research on WLAN based position estimation system using fingeprinting method with Fuzzy K-Nearest Neighbor (FK-NN) algorithm is obtained. From the results, the FK-NN algorithm can estimate the position successfully. The average accuracy 2.4 meters with percentage of 76%. This accuracy level is obtained from the position estimation system is affected by several things including training data used and the noise from measurement location.

6. Reference

[1] 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.

[2] H. Chabbar and C. Mouhcine, “Indoor localization using Wi-Fi method based on fingerprinting technique,” *Ieee*, pp. 1–5, 2017.
[3] D. Li, B. Zhang, and C. Li, “A Feature-Scaling-Based k-Nearest Neighbor Algorithm for Indoor Positioning Systems,” *IEEE Internet Things J.*, vol. 3, no. 4, pp. 590–597, 2016.
[4] Nico Wiranata; Reza Firsandaya Malik, “Perancangan model wlan-location estimation system di dalam gedung menggunakan teknik fingerprint,” 2015. (Indonesian)
[5] B. F. Billyan, A. Bhawiyuga, and R. Primananda, “Implementasi Metode Klasifikasi Fuzzy K-Nearest Neighbor (FK-NN) Untuk Fingerprint Access Point Pada Indoor Positioning,” vol. 1, no. 11, 2017.(Indonesian)
[6] Y. Gu, Y. Chen, J. Liu, and X. Jiang, “Online deep intelligence for Wi-Fi indoor localization,” *UbiComp ISWC 2015 - Proc. 2015 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. Proc. 2015 ACM Int. Symp. Wearable Comput.*, pp. 29–32, 2015.
[7] T. Alhmiedat and G. Samara, “An Indoor Fingerprinting Localization Approach for ZigBee Wireless Sensor Networks,” vol. 105, no. 2, pp. 190–202, 2013.
[8] C. Laoudias, P. Kemppi, and C. G. Panayiotou, “Localization using Radial Basis Function Networks and Signal Strength Fingerprints in WLAN,” 2009.
[9] L. Jiang, “a Wlan Fingerprinting Based Indoor Localization Technique,” *Univ. Nebraska*, 2012.
[10] T. Chuenurajit, S. Phimmasean, and P. Cherntanomwong, “Robustness of 3D indoor localization based on fingerprint technique in wireless sensor networks,” *2013 10th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol.*, vol. 20, no. 3, pp. 1–6, 2013.
[11] S. Sorour, Y. Lostanlen, and S. Valaee, “RSS based indoor localization with limited deployment load,” *GLOBECOM - IEEE Glob. Telecommun. Conf.*, pp. 303–308, 2012.
[12] E. S. Pasinggi, S. Sulistyo, and B. S. Hantono, “Sistem Penentuan Posisi Di Dalam Ruangan Dengan Metode Fingerprint (Knn),” pp. 6–8, 2015. (Indonesian)
[13] G. Felix, M. Siller, and E. N. Alvarez, “A fingerprinting indoor localization algorithm based deep learning,” *2016 Eighth Int. Conf. Ubiquitous Futur. Networks*, pp. 1006–1011, 2016.
[14] E. Prasetyo *et al.*, “Fuzzy K-Nearest Neighbor In Every Class Untuk Klasifikasi Data,” *Semin. Nas. Tek. Inform.*, no. Santika, pp. 57–60, 2012.
[15] D. Y. Permana, A. Handojo, and J. Andjarwirawan, “Aplikasi Indoor Positioning System Menggunakan Android dan Wireless Local Area Network Dengan Metode Fuzzy Logic Indoor Positioning System.” (Indonesian)

Acknowledgements
This article’s publication is supported by the United States Agency for International Development (USAID) through the Sustainable Higher Education Research Alliance (SHERA) Program for Universitas Indonesia’s Scientific Modelling, Application, Research and Training for City-centered Innovation and Technology (SMART CITY) Project, grant #AID-497-A-160004, Sub Grant #IIE00000078-UI-1.
This article is presented at the International Conference on Smart City Innovation 2018 that supported by the United States Agency for International Development (USAID) through the Sustainable Higher Education Research Alliance (SHERA) Program for Universitas Indonesia’s Scientific Modeling, Application, Research and Training for City-centered Innovation and Technology (SMART CITY) Project, Grant #AID-497-A-1600004, Sub Grant #IIE-00000078-UI-1