Journal of Engineering and Applied Sciences 12 (5): 1192-1199, 2017
ISSN: 1816-949X

An Improvement of Indoor Navigation System Based on Fingerprinting Technique Using K-Means Clustering Algorithm

Jirapat Sangthong and Sathaporn Promwong
Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang,
10520 Bangkok, Thailand

Abstract: Indoor navigation system is the one of interesting application among the researchers in an indoor environment due to the meter-level accuracy requirement in complex structure. This research proposed an improvement of the indoor navigation system based on fingerprinting technique by using K-Means (KM) clustering algorithm. The unknown positions are estimated by using Least Square (LS) and K-Nearest Neighbor (KNN) algorithms. The experimental results show the performance comparison between no-clustering case and KM-clustering case. Finally, we found that the KM clustering algorithm can be improved the accuracy of indoor navigation system both LS and KNN algorithm.

Key words: Indoor navigation system, fingerprinting technique, K-Means (KM), clustering algorithm, Least Square (LS), K-Nearest Neighbor (KNN)

INTRODUCTION

Over the last few years, the indoor Location-Based Service (LBS) has been extensively applied in many applications namely in guided tours of museums; indoor robotics tracking; store navigation; emergency services; and so on (He and Chan, 2016; Pahlavan *et al*., 2015). Navigation system is the one of interesting application among the researchers in an indoor environment due to the meter-level accuracy requirement in complex structure (Montanes *et al*., 2013).

For estimate the position, the wide variety of signal parameters are considered depend on the accuracy requirement and the limitation of the transmitter or receiver devices. Commonly, the measurement of signal parameters are based on Angle of Arrival (AOA), Time of Arrival (TOA), Time Difference of Arrival (TDOA) and Received Signal Strength (RSS) (Yang and Shao, 2015; Dardari*et al*., 2015). Triangulation technique is based on angle of the received signal. It requires antenna arrays in order to AOA estimation (Elsamnty *et al*., 2016; Malajner *et al*., 2015; Chuang *et al*., 2015; Liu and Cui, 2015). Besides, TDOA is proposed to find the unknown positions by using hyperbolas technique (Jafari *et al*., 2015). For TDOA estimation, it requires the synchronization clock on all Access Points (APs) (Makki *et al*., 2016). In addition, trilateration technique is the one of most popular technique. This technique required at least three APs to find the unknown position by using the intersection line between the circles (Yang and Liu, 2010). It can be used TOA or RSS to estimate the positions by calculate distances from considered signal parameter. Like TDOA, TOA estimation also needs synchronization clock to get precision timing. RSS is quite fascinating in practically transceivers for indoor navigation system. Although, RSS estimation does not require specialized hardware (Huang and Manh, 2016; Pagano *et al*., 2015; Wang *et al*., 2015; Talvitie *et al.,* 2015) but it is affected from the shadowing, interference and multipath fading. Fortunately, fingerprinting technique can be solved the indoor environmental effects (Mao *et al*., 2016).

The fingerprinting technique is approached in many literatures (Zhuang *et al*., 2016; Shu *et al*., 2016; Li *et al*., 2016; Kim *et al*., 2016). The main concept is to collect the considered signal parameter at reference positions in the interesting area and keep them to the database. Then, the unknown position can be estimated by using the matching algorithm in real-time process. This technique does not require the multiple APs which can reduce the cost significantly compared with the triangulation, hyperbolas and trilatelation techniques. Several algorithms have been presented for estimate the unknown position in real-time process. K-Nearest Neighbor (KNN) is the one of attractive algorithm because it is easy computation (Mao *et al*., 2016; Hossain *et al*., 2013). KNN considers all nearest K positions to calculate the unknown position. However, the value of K must be found in training process in order to more accurate estimation.

Corresponding Author: Jirapat Sangthong, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, 10520 Bangkok, Thailand

Clustering algorithm is the method to group the similar data into the same cluster (Lam *et al.*, 2015; Tsai *et al.,* 2014). It can be used in many fields such as image processing, data mining, pattern recognition, machine learning, and so on. K-Means (KM) is the one basic algorithm because it is simplicity and low complexity in computation process (Alsheikh *et al.*, 2014). The data are grouped by considered the Euclidean distance between the centroid of each cluster and considered data.

This research proposed an improvement of the indoor navigation system based on fingerprinting technique by using K-Means clustering algorithm. The experimental test bed area can be covered 334.5 m^2. The RSS values are collected by using Samsung Galaxy S3 smart phone. The ZyXel P-660HN-T1A Access Points (APs) at the frequency of 2.4 GHz are used as the transmitters. The unknown positions are estimated by using Least Square (LS) and K-Nearest Neighbor (KNN) algorithms. The experimental results show the performance comparison between no-clustering case and KM-clustering case. Moreover, the multiple APs are considered in KM-clustering case.

MATERIALS AND METHODS

Fingerprinting technique with k-means clustering: The fingerprinting technique can be divided into two processes: a training process that the RSS values at the reference positions are surveyed. Then, the similarly RSS values have been grouped by using KM. After that, the RSS samples are generated and collected them into the database. For real-time process, the positions are estimated by using the matching algorithm such as LS and KNN algorithm.

For training process, the RSS values are collected at the reference positions with t sample times. The average of RSS values ($RSS_{i,j}$) from access point (i) at reference positions (j)can be calculated by:

$$RSS_{i,j} = \frac{1}{\tau}\sum_{a=1}^{\tau} RSS_{i,j}(a) \tag{1}$$

After that, all $RSS_{i,j}$ are stored into the database (FP) which can be represented by:

$$FP = \begin{Bmatrix} RSS_{1,1} & RSS_{2,1} & \cdots & RSS_{m,1} \\ RSS_{1,2} & RSS_{2,2} & \cdots & RSS_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ RSS_{1,n} & RSS_{2,n} & \cdots & RSS_{m,n} \end{Bmatrix} \tag{2}$$

Where:
m = The number of access point
n = The number of reference positions

The KM clustering algorithm is used to group the $RSS_{i,j}$ in the database. First, the number of cluster k is considered and initialed the centroid of each cluster (C_b) for represents the cluster group. Then, the Euclidean distance between the RSS values ($RSS_{i,j}$) and the centroid of each cluster can be calculated by:

$$d_{i,b} = \left\| RSS_{i,j} - c_b \right\| \tag{3}$$

where, b = 1, 2, ..., k. The cluster is considered from the minimum Euclidean distance ($d_{i,b}$) which can be represented by:

$$c_e = \arg\min d_{i,b} \tag{4}$$

Then, the new centroid of each cluster can be calculated by considered the average value in each cluster:

$$c_{b,new} = \frac{1}{|c_e|}\sum_{x\in c_e} RSS_{i,j} \tag{5}$$

After that, the K-Means clustering objective function value is given by:

$$KM = \sum_{j=1}^{k}\sum_{x\in c_b} \left\| x_i - c_b \right\|^2 \tag{6}$$

The process is repeated again from Eq. 3-6 until each cluster is no change:

$$\left| KM_{old} - KM_{new} \right| \le \varepsilon \tag{7}$$

The final centroid of each cluster is denoted as flag (f_n). Then, the equation information in the database can be represent by:

$$FP = \begin{Bmatrix} RSS_{1,1} & RSS_{2,1} & \cdots & RSS_{m,1} & f_1 \\ RSS_{1,2} & RSS_{2,2} & \cdots & RSS_{m,2} & f_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ RSS_{1,n} & RSS_{2,n} & \cdots & RSS_{m,n} & f_n \end{Bmatrix} \tag{8}$$

For real-time process, the RSS value is collected from unknown position (RSS_u) and sent it to the server.

The cluster of unknown position can be found by considered closest centroid of cluster which can be represented by:

$$d_b = \|RSS_u - c_b\| \tag{9}$$

After that, the position is estimated by using the matching algorithm. For LS algorithm, the unknown position is derived from the position that has minimum difference between its RSS value and the RSS information in the database. The difference of fingerprint (d_f) can be calculated by:

$$d_f(j) = \sqrt{\sum_{i=i}^{m}(RSS_{i,j}(f_u) - RSS_{u,i})^2} \tag{10}$$

For LS algorithm, the estimated position (x_e, y_e) is considered as the position with minimum difference of fingerprint and can be written as:

$$(x_e, y_e) = \arg\min_{x,y} d_f(j) \tag{11}$$

For KNN algorithm, the set of k reference positions (x_i, y_i) are selected according to the minimum difference of fingerprint. Then, the unknown position can be estimated by:

$$(x_e, y_e) = \frac{1}{k}\sum_{i=1}^{k}(x_i, y_i) \tag{12}$$

Experimental setup: The experiments are conducted at the 3rd floor, E-Building, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. The test bed area is covered 334.5 m^2. The obstacle object is the non-measurable zone with covered area is 4.84 m^2. The RSS values are collected by using Samsung Galaxy S3 smart phone. The five ZyXel P-660HN-T1A Access Points (APs) are used as the transmitters. There are operated at the frequency of 2.4 GHz based on the iEEE802.11b/g standard. The RSS samples from 5 APs are collected with 10 sample times over 1,289 reference positions. The layout of the test bed area and the experimental setup.

RESULTS AND DISCUSSION

This study provides the experimental results (Fig. 1). For building the fingerprints, the average of RSS values of each AP with 0.5 m space are collected. The fingerprints were built from RSS values of AP1-AP5 can be shown in Fig. 2-6, respectively.

Table 1: The summary of all fingerprints information

Access point	RSS values (dBm)	Max. RSS value position (x,y)	No. of covered positions (points)
AP1	-48.60 to -97	(4,41)	1,172
AP2	-44.80 to -98	(9.5,48)	1,219
AP3	-51.20 to -95	(13,36)	1,288
AP4	-42.40 to -99	(14.5,28)	1,253
AP5	-44.40 to -99	(14.5,3)	924

The considered fingerprints will be used in training process. Table 1 shows the summary of all fingerprints information. To investigate LS and KNN algorithm, the AP4 fingerprints are selected to be the information in the database. The RSS values are between -42.40 to -99 dBm. The maximum RSS value is obtained at the position (14.5, 28). The AP4 can be covered 1,253 positions.

However, the fingerprinting technique can be used many information to improve the accuracy of indoor navigation. The rest of the results are shown the accuracy of fingerprinting technique using KM clustering algorithm with multiple APs.

Accuracy of LS algorithm: The unknown positions are estimated by using Eq. 12. The histogram of distance errors of no-clustering case can be shown in Fig. 7. The 1 m accuracy is 41 positions with 3.27%. Figure 8 shows the histogram of distance errors of KM clustering case. The 1 m accuracy is 47 positions with 3.75%. Figure 9 shows the Cumulative Distribution Function (CDF) of distance errors of LS algorithm. For no-clustering case, the average distance error is 12.73 m. While, the average distance error of KM clustering case is 9.81 m.

Accuracy of KNN algorithm: The unknown positions are estimated by using Eq. 3. The highest accuracy are found in k = 2. Then, the histogram of distance errors of no-clustering case is shown in Fig. 10. The 1 m accuracy is 33 positions with 2.63%. From Fig. 11, the histogram of distance errors of KM-clustering case is shown. The highest accuracy are found in k = 5. The 1 m accuracy is 33 positions with 2.63%. After that, the CDF of distance errors of KNN algorithm is shown in Fig. 12. For no-clustering case, the average distance error is 12.48 m. While, the average distance error of KM-clustering case is 9.76 m.

Accuracy of fingerprinting technique using KM clustering algorithm with multiple APs: For fingerprinting technique, many information can be used for building the fingerprints that can be improved the accuracy of indoor navigation. The accuracy of KM-clustering case using 2 APs to 5 APs can be shown in Fig. 13-16, respectively.

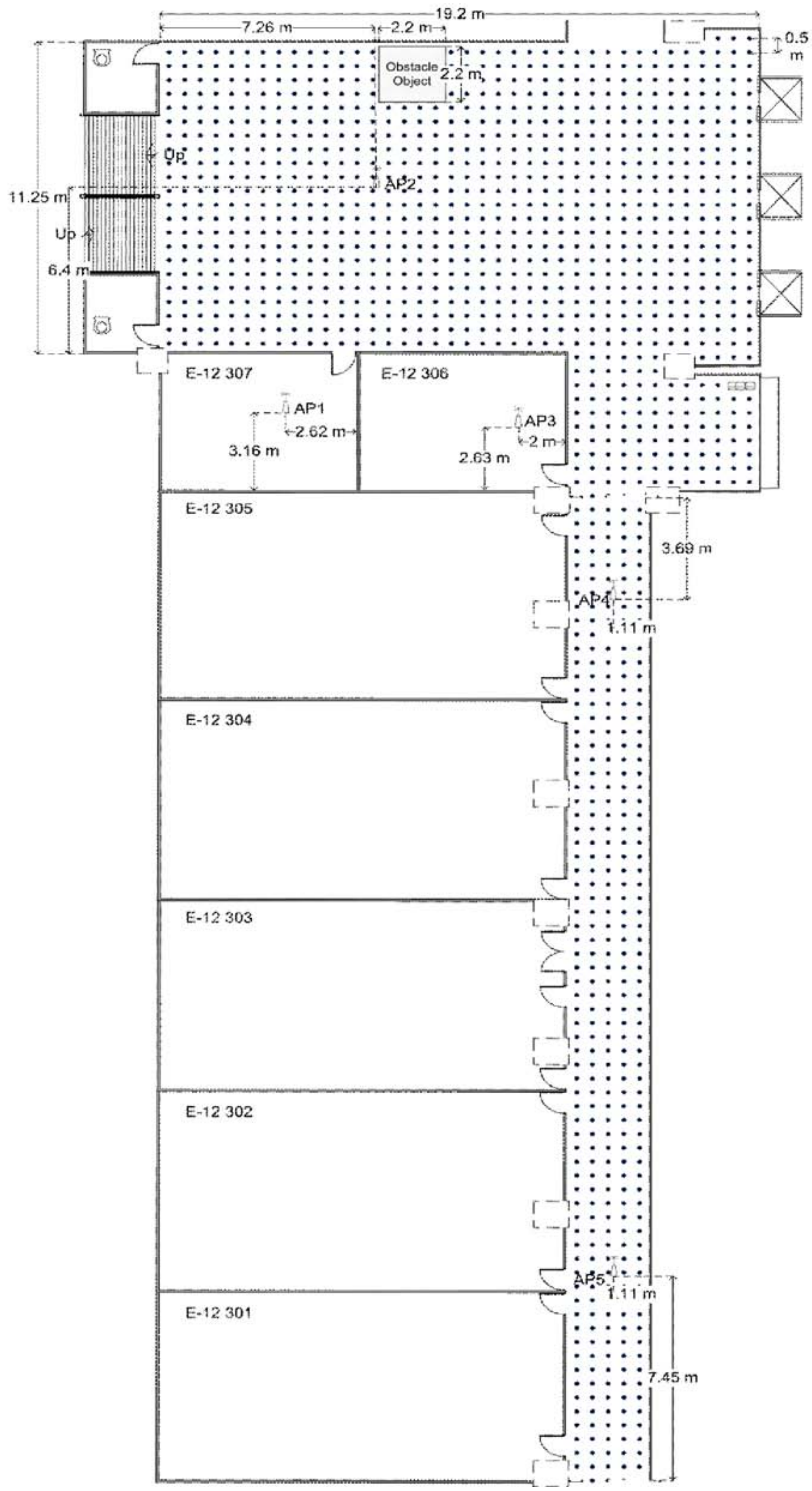

Fig. 1: The layout of the test bed area and the experimental setup

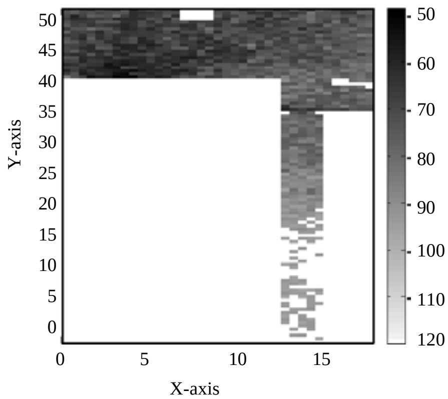

Fig. 2: The fingerprints built from RSS values of AP1

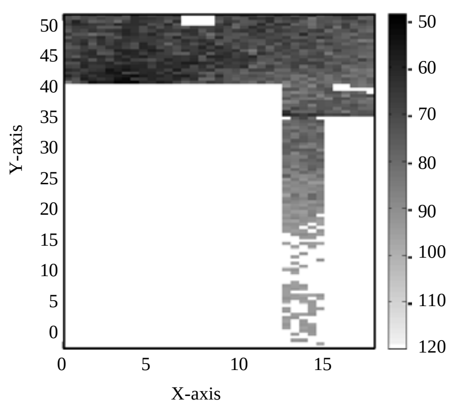

Fig. 3: The fingerprints built from RSS values of AP2

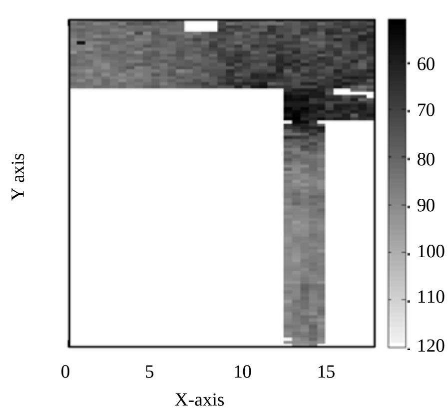

Fig. 4: The fingerprints built from RSS values of AP3

Table 2 shows the summary of fingerprinting technique using KM clustering algorithm with multiple Aps. For example, 5 Aps case, the fingerprints are built from RSS values of AP1-AP5. Figure 18 shows the histogram of distance errors of KM

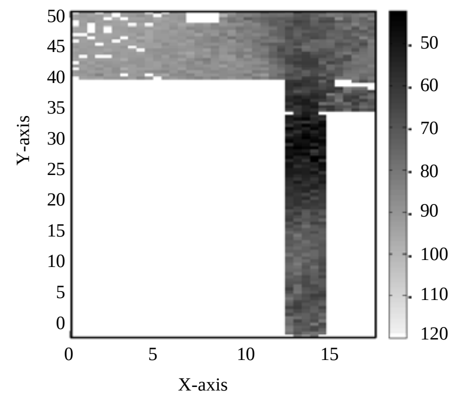

Fig. 5: The fingerprints built from RSS values of AP4

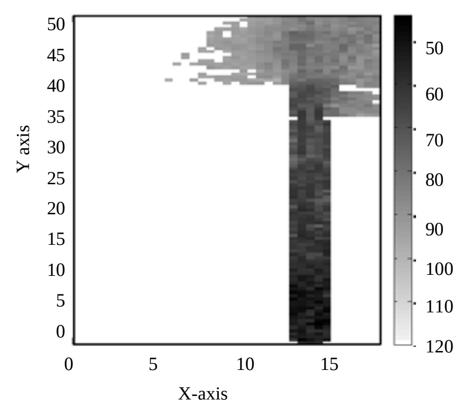

Fig. 6: The fingerprints built from RSS values of AP5

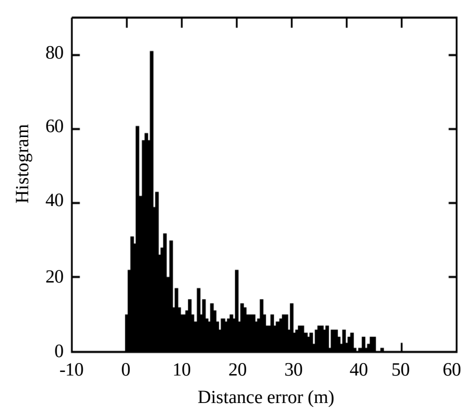

Fig. 7: The accuracy of LS algorithm with no-clustering case

clustering case using 5 Aps. The 1 m accuracy is 973 positions with 75.48%. While the average distance error is 1.04 m.

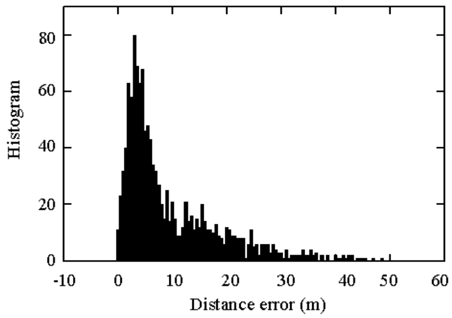

Fig. 8: The accuracy of LS algorithm with KM-clustering case

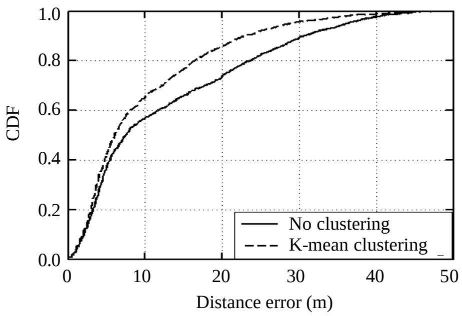

Fig. 9: The CDF of distance errors of LS algorithm

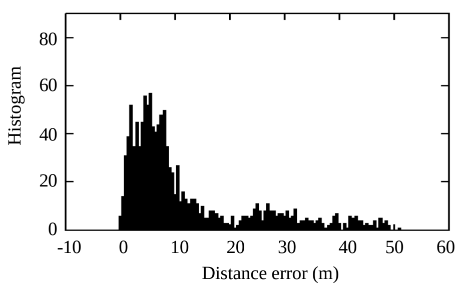

Fig. 10: The accuracy of KNN algorithm for no-clustering case

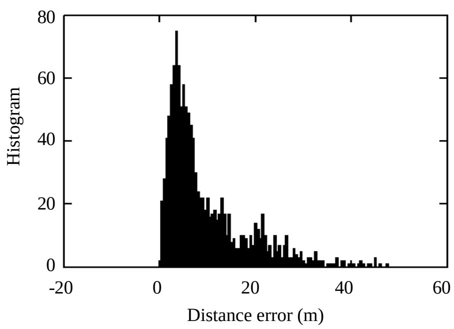

Fig. 11: The accuracy of KNN algorithm for no-clustering case

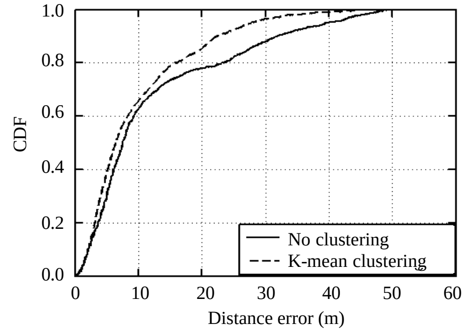

Fig. 12: The CDF of distance errors of KNN algorithm

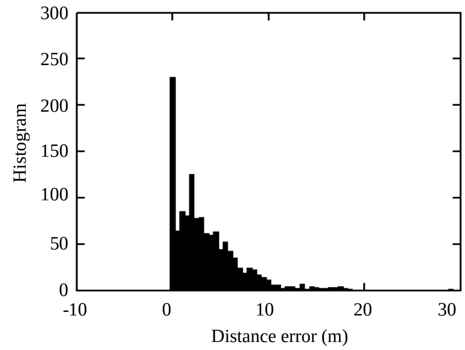

Fig. 13: The accuracy of KM-clustering case using 2 APs

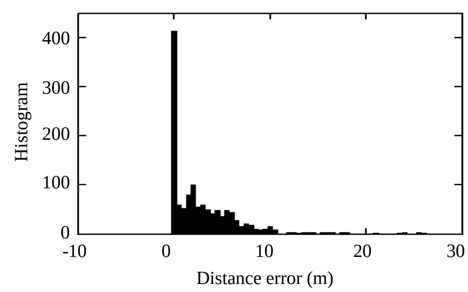

Fig. 14: The accuracy of KM-clustering case using 3 APs

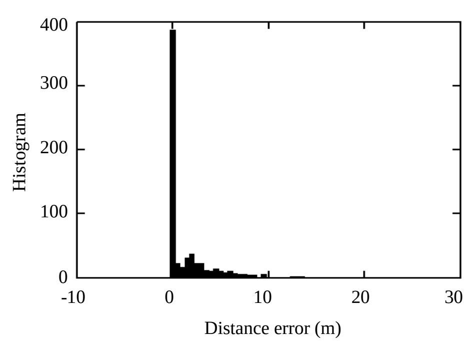

Fig. 15: The accuracy of KM-clustering case using 4 APs

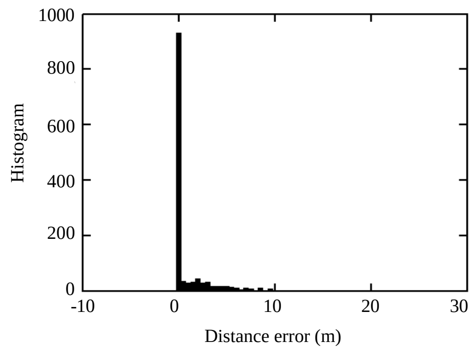

Fig. 16: The accuracy of KM-clustering case using 5 APs

Table 2: The summary of all fingerprints information

No. of APs	The selected Aps for build the fingerprints	1 m accuracy (positions)	1 m accuracy (%)	Average distance error (m)
2	AP2, AP4	324	25.14	3.57
3	AP2, AP4, AP5	573	44.45	2.43
4	AP1, AP2, AP4, AP5	830	64.39	1.45
5	AP1, AP2, ,AP3, AP4, AP5	973	75.48	1.04

CONCLUSION

In this research, we proposed an improvement of the indoor navigation system based on fingerprinting technique by using K-Means clustering algorithm. The unknown positions are estimated based on measurement data by using Least Square (LS) and K-Nearest Neighbor (KNN) algorithms. From the experimental results, the accuracy of fingerprinting technique can be improved by using KM clustering algorithm both LS and KNN cases. Moreover, the accuracy of fingerprinting technique using KM clustering algorithm with multiple APs is closest the requirement of indoor navigation system. This meant that this work can be used and applied in many indoor navigation systems with no specialized hardware requirement practically.

REFERENCES

Alsheikh, M.A., S. Lin, D. Niyato and H.P. Tan, 2014. Machine learning in wireless sensor networks: Algorithms, strategies and applications. IEEE. Commun. Surv. Tutorials, 16: 1996-2018.

Chuang, S.F., W.R. Wu and Y.T. Liu, 2015. High-resolution AoA estimation for hybrid antenna arrays. IEEE. Trans. Antennas Propag., 63: 2955-2968.

Dardari, D., P. Closas and P.M. Djuric, 2015. Indoor tracking: Theory, methods and technologies. IEEE. Trans. Veh. Technol., 64: 1263-1278.

Elsamnty, R., A.M. Attiya, E.A.F. Abdallah and A.H. Ammar, 2016. Beamforming and angle-of-arrival estimation of square planar antenna array. Open J. Antennas Propag., 4: 13-23.

He, S. and S.H.G. Chan, 2016. Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. IEEE. Commun. Surv. Tutorials, 18: 466-490.

Hossain, A.M., Y. Jin, W.S. Soh and H.N. Van, 2013. SSD: A robust RF location fingerprint addressing mobile devices' heterogeneity. IEEE. Trans. Mob. Comput., 12: 65-77.

Huang, C.C. and H.N. Manh, 2016. RSS-based indoor positioning based on multi-dimensional kernel modeling and weighted average tracking. IEEE. Sensors J., 16: 3231-3245.

Jafari, A., L. Petrillo, J. Sarrazin, D. Lautru and P.D. Doncker *et al*., 2015. TDOA estimation method using 60 GHz OFDM spectrum. Intl. J. Microwave Wirel. Technol., 7: 31-35.

Kim, Y., Y. Chon and H. Cha, 2016. Mobile crowdsensing framework for a large-scale wi-fi fingerprinting system. IEEE. Pervasive Comput., 15: 58-67.

Lam, D., M. Wei and D. Wunsch, 2015. Clustering data of mixed categorical and numerical type with unsupervised feature learning. IEEE. Access, 3: 1605-1613.

Li, Q., W. Li, W. Sun, J. Li and Z. Liu, 2016. Fingerprint and assistant nodes based wi-fi localization in complex indoor environment. IEEE. Access, 4: 2993-3004.

Liu, Y. and H. Cui, 2015. Antenna array signal direction of arrival estimation on Digital Signal Processor (DSP). Procedia Comput. Sci., 55: 782-791.

Makki, A., A. Siddig, M. Saad, J.R. Cavallaro and C.J. Bleakley, 2016. Indoor localization using 802.11 time differences of arrival. IEEE. Trans. Instrum. Meas., 65: 614-623.

Malajner, M., D. Gleich and P. Planinsic, 2015. Angle of arrival measurement using multiple static monopole antennas. IEEE. Sens. J., 15: 3328-3337.

Mao, R., P. Xiang and D. Zhang, 2016. Precise transceiver-free localization in complex indoor environment. China Commun., 13: 28-37.

Montanes, J.A.P., A.M. Rodriguez and I.S. Prieto, 2013. Smart indoor positioning-location and navigation: A lightweight approach. Intl. J. Artif. Intell. Interact. Multimedia, 2: 43-50.

Pagano, S., S. Peirani and M. Valle, 2015. Indoor ranging and localisation algorithm based on received signal strength indicator using statistic parameters for wireless sensor networks. IET. Wirel. Sens. Syst., 5: 243-249.

Pahlavan, K., P. Krishnamurthy and Y. Geng, 2015. Localization challenges for the emergence of the smart world. IEEE. Access, 3: 3058-3067.

Shu, Y., Y. Huang, J. Zhang, P. Coue and P. Cheng *et al*., 2016. Gradient-based fingerprinting for indoor localization and tracking. IEEE. Trans. Ind. Electron., 63: 2424-2433.

Talvitie, J., M. Renfors and E.S. Lohan, 2015. Distance-based interpolation and extrapolation methods for RSS-based localization with indoor wireless signals. IEEE. Trans. Veh. Technol., 64: 1340-1353.

Tsai, C.W., C.F. Lai, M.C. Chiang and L.T. Yang, 2014. Data mining for internet of things: A survey. IEEE. Commun. Surv. Tutorials, 16: 77-97.

Wang, B., S. Zhou, W. Liu and Y. Mo, 2015. Indoor localization based on curve fitting and location search using received signal strength. IEEE. Trans. Ind. Electron., 62: 572-582.

Yang, C. and H.R. Shao, 2015. WiFi-based indoor positioning. IEEE. Commun. Mag., 53: 150-157.

Yang, Z. and Y. Liu, 2010. Quality of trilateration: Confidence-based iterative localization. IEEE. Trans. Parallel Distrib. Syst., 21: 631-640.

Zhuang, Y., Z. Syed, Y. Li and N. El-Sheimy, 2016. Evaluation of two WiFi positioning systems based on autonomous crowdsourcing of handheld devices for indoor navigation. IEEE. Trans. Mob. Comput., 15: 1982-1995.