Wireless Networks (2019) 25:3019–3027
https://doi.org/10.1007/s11276-018-1700-9

Wi-Fi fingerprint using radio map model based on MDLP and euclidean distance based on the Chi squared test

Ju-Hyeon Seong[1] · Dong-Hoan Seo[2]

Published online: 8 March 2018

Abstract
The Wi-Fi fingerprint, which can be used on existing wireless networks, is one of the main indoor positioning techniques that utilizes the received signal strength (RSS). In smartphones, the positioning performance of the fingerprint has been significantly improved through fusion algorithms along with terrestrial magnetism and acceleration sensors. However, the positioning accuracy and speed of the fingerprint is based on radio maps. Although these maps are separate databases obtained without using these sensors, they are important reference elements for initial position estimation and sensor error compensation. In order to minimize the DB of fingerprint and to improve the speed of system construction according to the area of positioning is expanded, this paper proposes a Wi-Fi fingerprint using a radio map construction model based on the minimum description length principle, which can automatically optimize radio maps and the Euclidean distance algorithm based on the Chi squared test. Unlike the existing RSS-classification-based radio map construction method, the proposed access point (AP) classification-based radio construction model not only automatically distinguishes the continuity of all the RSSs acquired from the APs but also optimizes the radio map by eliminating unnecessary APs, based on the information gain. In the positioning phase, based on the proposed radio map, the accuracy of the signals is distinguished using a Chi squared test for the AP RSSs measured in real-time. Therefore, the Euclidean distance, based on the Chi squared test, improves the positioning performance by determining the position accuracy using weighted values of the RSSs, with high reliability.

Keywords Fingerprint · MDLP · Chi squared test · Euclidean distance · Radio map

1 Introduction

With the global propagation of smartphones and the advancements in wireless device technologies such as the Wi-Fi and Bluetooth low energy (BLE), user-location accuracy has gradually improved [1, 2]. The fingerprint method, which is a Wi-Fi positioning technology with long transmission distances and the highest versatility between indoor devices, uses radio maps; hence, it is robust in non-line of sight (NLOS) environments compared to the time of arrival (TOA) and Time Difference of Arrival (TDOA) method, which measure the time of arrival of the transmission signals between devices. The fingerprint method is advantageous because it can utilize existing wireless networks [3–6].

However, the size of the radio map increases in accordance with the area, whose location is to be estimated and the density of the Wi-Fi APs; hence, its creation for fingerprinting is time-consuming. Moreover, the positioning resolution of the fingerprint, which is of the order of meters based on the interval of the Wi-Fi signal that is set for creating the radio map, is lower than that of TOA, which shows results up to centimeters.

To overcome these disadvantages, several converged algorithms such as the ultra-wide band (UWB), acceleration sensors have been developed [7, 8]. Despite the application of these fusion techniques, the starting point

✉ Dong-Hoan Seo
dhseo@kmou.ac.kr

[1] Department of Electrical and Electronics Engineering, Korea Maritime and Ocean University, #727 Taejong-Ro, Youngdo-Gu, Busan 606-791, Korea

[2] Division of Electronics and Electrical Information Engineering, Korea Maritime and Ocean University, #727 Taejong-Ro, Youngdo-Gu, Busan 606-791, Korea

relies on the Wi-Fi fingerprint because it is extremely difficult to automatically determine the location using sensors with methods that estimate the starting point of the positioning technique indoors. Therefore, several studies endeavor to reduce the database (DB) size and improve the accuracy in order to minimize the calculation speed of the Wi-Fi fingerprint.

Jung et al. [9] analyzed the positioning performances of four types of radio map models (point-by-point manual calibration, walking survey, semi-supervised learning-based method, and the unsupervised learning-based method) using a fingerprint model called the signal fluctuation matrix (SFM). The positioning performance of their proposed model had the advantage of selecting models based on the environment, considering the size of the positioning area and the number of APs; however, the positioning performance was approximately 2 m, similar to the existing one. Thus, the positioning-performance improvement was not considered, while comparing the performances of these fusion algorithms.

Carlos et al. [10] improved the position accuracy using two types of information and learning algorithms generated by modifying the support vector machine (SVM). As the SVM-based mechanical learning algorithm used in their study had a slow processing speed compared to the other classification algorithms, it rapidly increased the computation time, in accordance with the size of the radio map. Therefore, additional research is required to resolve this issue.

Gary Chan et al. [11] present some recent progress in reducing site survey and online adaptation to signal/fingerprint change based on crowdsourcing. This approach has great advantages in reducing the time to create the database, but it is difficult to reduce the amount of the database definitely.

As described above, in order to expand and activate the recognition area in the fingerprint-based positioning technology, the speed and the accuracy of the construction of the radio map have recently become increasingly important.

In order to minimize the DB of fingerprint and to improve the speed of system construction according to the indoor space of positioning is expanded, this study proposes a Wi-Fi fingerprint using a new radio map construction model based on MDLP, for optimizing radio maps and the Euclidean distance algorithm based on the Chi squared test, for improving the positioning performance. Unlike existing position-based classification algorithms such as the SVM, the proposed novel radio map construction model based on MDLP applies RSS-based classification to optimize radio maps; it realizes this through AP elimination by automatically classifying the continuity of the RSS data sets acquired from each AP and determining whether the RSS data set is needed in the radio map. Unpredictable RSS values from the AP RSSs measured in real-time during the positioning phase, based on this optimized radio map, are automatically eliminated from the calculations; the Euclidean distance, based on the Chi squared test, improves the positioning performance by determining the position accuracy using weighted values of the RSSs, with high reliability. Using the proposed fingerprint, there is an improvement of approximately 5% in the position accuracy and a superior performance in optimizing the radio map by decreasing the radio map size by 38.56%.

2 Relevant theories

2.1 SVM-based fingerprint

In general, fingerprinting is divided into two phases: training and positioning.

In the training phase, a reference point is established for collecting the RSSs at set intervals in the given space for which the position must be estimated, based on the internal structure, number of APs, etc. of the indoor environment. Subsequently, a radio map is created based on the RSSs of the Wi-Fi APs measured at each point. Because the reference point estimates the user location, the positioning performance improves when the intervals are narrow. However, a reference point interval of 2–3 is generally applied owing to the deviation of the measured RSSs by reflection and refraction of the radio wave or the occurrence of inherent measurement errors.

The RSS is measured and collected by sensor-based methods, walking surveys, learning algorithms, etc. In general, using these collected RSSs, radio maps are created by deterministic or probabilistic models [9, 12–14], crowdsourcing [15, 16], classification models (e.g., Bayes classifier, artificial neural networks [17], SVMs [10, 18–20]), etc.

Among them, the SVM that extends the concept of the perceptron is the most popular classification algorithm with logistic regression. Unlike the perceptron, which minimizes classification errors, it is an algorithm that maximizes margins. Here, the margin is the distance between the boundary for classification and the nearest trading data to this boundary, and the trading data closest to this boundary is called the support vector. This boundary line is called hyperplane of multidimensional space. The general hyperplane formula is as follows.

$$\omega^T x + k = 0 \tag{1}$$

where x is an vector in the vector space, ω is the weighted vector. k is the bias term. As shown in Fig. 1, when a

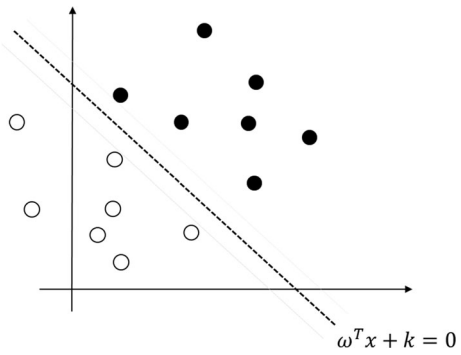


Fig. 1 Classify data by SVM on hyperplane

hyperplane that can obtain the maximum margin is derived, the whole data can be divided into two parts data. Therefore the data can be divided according to the reference (axis information) by iterative calculation as necessary.

It is applied to fingerprint radio maps, in particular because it can reduce the computational complexity at a higher dimension, rendering it suitable for the automatic classification of complex multidimensional fingerprints. However, because this method requires considerable time compared to several other classification algorithms, real-time processing is difficult. Thus, it is used in the training phase, which does not require real-time processing. In the training phase, the initial radio map, which includes a set of RSSs measured at the reference points, is defined as follows:

$$\mathrm{D}_t = \begin{bmatrix} P_{AP1(1)} & P_{AP1(2)} & P_{AP1(3)} & \cdots & P_{AP1(n)} \\ P_{AP2(1)} & P_{AP2(2)} & P_{AP2(3)} & \cdots & P_{AP2(n)} \\ P_{AP3(1)} & P_{AP3(2)} & P_{AP3(3)} & \cdots & P_{AP3(n)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ P_{APk(1)} & P_{APk(2)} & P_{APk(3)} & \cdots & P_{APk(n)} \end{bmatrix} \quad (2)$$

where D_t is the set of all collected RSS data, $P_{APk(n)}$ are the stored AP RSSs, k is the number of APs measured in the area, and n is the number of reference points. SVM is applied to automatically classify the RSSs of the initially collected radio map, according to the time axis (reference points). On applying SVM to a radio map, the RSSs are separated based on the reference points and the positions are classified in accordance with the signal intensity, which varies according to the position of the RSSs of the Wi-Fi AP signals measured at each reference point. SVM maps to the input space to find the optimal hyperplane to separate the dataset of AP. A margin is introduced by getting the distance between a RSSI vector x_i and the boundary, which can be written as

$$\frac{|\omega^T x_i + k|}{\omega} \quad (3)$$

SVM is to optimize the maximum boundary by minimizing (3), which is expressed as

$$min_i \left|\omega^T x_i + k\right| = 1 \quad (4)$$

Through this expression, it can be reduced to a maximization

$$y_i\left(\omega^T x_i + k\right) \geq 1 \quad (5)$$

where y_i states the classification category. If it is 1, $P_{APk(n)}$ belongs to positive region, otherwise, $P_{APk(n)}$ is in negative region [15].

2.2 Minimum description length principle

As the minimum description length principle (MDLP), a typical discretization technique based on class entropy depicting the uncertainty of data, can analyze the correlation of variables and minimize information data loss, it is frequently used in machine learning and probability modeling. If the volume and complexity of the data to be applied for modeling is high, the MDLP is advantageous in terms of simplifying and optimizing this data. The entropy has a value between 0 and 1 and it is suitable for discriminating data uncertainty because it indicates that the data set is composed of only one value as it approaches zero. The entropy is denoted by,

$$Ent(S) = -\sum_{j=1}^{k} P\left(C_j, S\right)\log\left(P\left(C_j, S\right)\right) \quad (6)$$

$$P\left(C_j, S\right) = \frac{Count\left(C_j, S\right)}{|S|} \quad (7)$$

where S is the set of given data, C_j is the set of class values, $Count\left(C_j, S\right)$ is the number of classes in C_j in S, and $|S|$ is the total number of given data.

The entropy derived through the above formula is applied in the MDLP as follows:

$$E(X, T, S) = \frac{|S_1|}{|S|}Ent(|S_1|) + \frac{|S_2|}{|S|}Ent(|S_2|) \quad (8)$$

where S is a data set, X is a continuous variable, and T is a split point of S. $S_1$and S_2 are the two data sets that were separated using T. If this is applied as a conditional expression, the data classification accuracy can be confirmed, based on the reference points, by dividing the AP signals several times or until the divisible limit is achieved. As the entropy is determined only by data continuity, it is not sufficiently considered for the attribute (reference point, AP, etc.). Therefore, this entropy is often represented by the information gain (IG), which indicates the reduction in entropy, when data are classified based on certain attributes.

2.3 Chi squared test

The Chi squared test, which is widely used in statistics, is a method for judging whether the measured value is actually reliable from expected values when the variables are normally distributed. Based on this, the following equation is used to verify the correlation between two variables.

$$\chi^2 = \sum \frac{(O-E)^2}{E} \tag{9}$$

Where O is the observed frequency and E is the expected frequency. The observation frequency will represent the numerically measured number and the expected frequency can be calculated from the expected ratio from any one characteristic. The value of χ^2 obtained through this process is used to determine the suitability of the value through the significant interval. In the proposed algorithm, the Chi squared test is the observation frequency of the RSS values of the DB, and the measured values are applicable to the expected frequency. In other words, it is possible to derive the normal distribution of APs according to each reference point (position) through the measured DB values, and then to estimate the position through the actual measurement values in the positioning phase of Fingerprint.

3 Proposed fingerprint based on supervised learning

As shown in Fig. 2, the proposed algorithm estimates the real-time positions of users through the training phase, in which a radio map is created using the measured RSSs from the Wi-Fi receiver and the positioning phase, in which the user position is estimated. In the training phase, to classify and minimize the initial radio map, the set of obtained RSSs from each AP is automatically analyzed and unnecessary APs are removed using the MDLP and IG. In the positioning phase, the radio map created through training phase is used to estimate the final position using the Euclidean distance algorithm, based on weighted values that applies the Chi squared test. As Wi-Fi RSSs measured in real-time are not regular and have considerable deviations, a Chi squared test is used to determine their reliability; only appropriate RSSs are applied to the Euclidean distance. The algorithm is explained in detail below.

1. Radio map construction method

In the fingerprint, because radio maps are used as references for estimating the user location in the positioning phase, their creation phase is critical in determining the accuracy of the system. Therefore, this phase requires the maximum time in the entire fingerprint process. The proposed radio map algorithm based on MDLP has a higher processing speed and is suitable for the classification of independent data, compared to the radio map algorithm based on SVM.

In the indoor area for which the radio map is to be created, a set of consecutively measured RSSs of any AP for each reference point can be expressed as follows:

$$P(re) = \left[P^1_{APi(1)} \quad P^2_{APi(1)} \quad P^r_{APi(1)} \quad \cdots \quad P^1_{APi(2)} \quad \cdots \quad P^r_{APi(n)} \right] \tag{10}$$

where APi is the identifier of the $i-$ th AP, n is the reference point, and r is the number of RSS measurements of APi at n's location. Hence, $P(re)$ is the set of all RSSs continuously measured from an AP and these sets are combined to create a radio map. This radio map is discretized by applying the MDLP in Eq. (8) to determine whether the dataset is suitably partitioned over time. The SVM shown in Fig. 3(a) is the most commonly used radio map classification method; it classifies based on the reference points, for categorizing the positions of the continuously-measured data sets. On the other hand, the proposed MDLP method shown in Fig. 3(b) classifies by determining the continuity and stability of the RSSs measured at each AP. The SVM and Proposed methods are similar algorithms for classifying consecutive data, there are different criteria to apply. SVM is a technique to classify reference points according to measuring time. The proposed algorithm classifies APs according to the characteristics of APs based on data measured according to reference points.

IG is used for checking whether the data sets, which are divided after being classified by MDLP, are distinct. IG indicates the decrease in entropy, after the classification of datasets based on certain properties. The IG using M as a

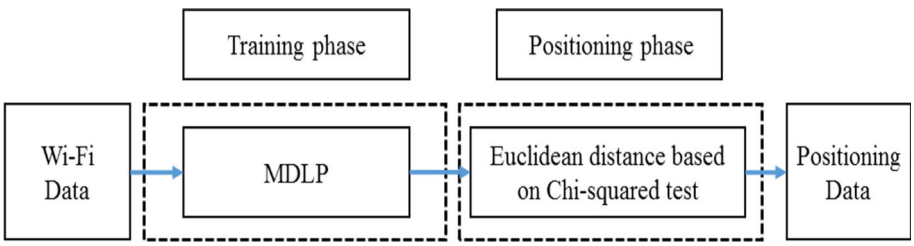


Fig. 2 Flowchart of the proposed fingerprint

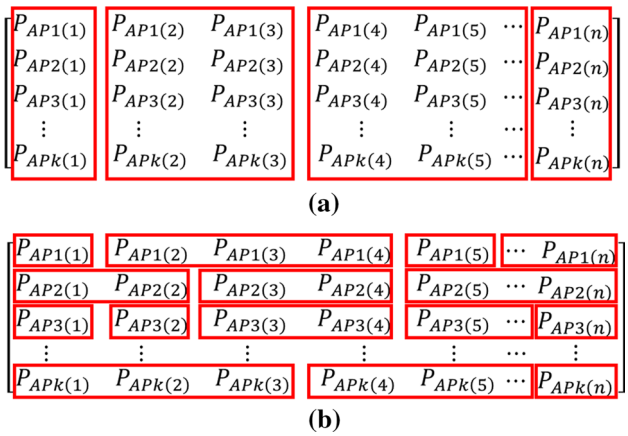


Fig. 3 Data classification method: **a** SVM, **b** The proposed MDLP

certain property of the data set, S, is IG (S, M) and can be defined as follows:

$$IG(S,M) = Ent(S) - \sum_{m \in Value\{M\}} \frac{|S_m|}{|S|} Ent(S_m) \tag{11}$$

where $Value\{M\}$ is the set of all possible values of the property M and S_m is the partial value of S, when M has a value, m. If the RSS data have no distinctions, based on certain properties, (position) through this equation, they are considered to be unnecessary AP RSSs with no distinction in the actual positioning; thus, the radio map size can be reduced by eliminating these APs. Radio maps that are minimized through the proposed MDLP and IG method can be generated and applied to the actual positioning method.

2. Positioning method

In the positioning phase, the user position is estimated based on the correlation between the radio map created using the proposed method and the Wi-Fi RSSs measured in real-time. The proposed positioning method applies the weighted values obtained by the Chi squared test, based on the Euclidean distance, to each AP; Euclidean distance is an algorithm that can determine the similarity of the AP RSSs, from the RSSs measured in real-time and the proposed radio map. This method extracts the reliability of the RSSs measured in real-time using the Chi squared test for such RSSs, in accordance with the distribution of the radio map RSS signals. The weighted values for each AP are applied, based on the Euclidean distance, for estimating the final user position. The positioning algorithm with the Euclidean distance and Chi squared test is as follows:

$$Mod\_Dist(i) = \sqrt{\sum_{j=1}^{n} \alpha^2 \left(AP_{Mj} - AP_{rj}\right)^2} \tag{12}$$

where AP_{Mj} is the RSS of the $j-$ th AP stored in the radio map and AP_{rj} is the RSS of the $j-$ th AP. α is the p value extracted using the Chi squared test; it is applied to each AP signal, according to the significance level. α, which has a value between 0 and 1, indicates the degree of nearness to the center based on the Chi squared test graph of the RSSI data sets of each AP corresponding to the RSSI measured in real time. In other words, α, which means p value, can be expressed as a measure that distinguishes whether RSSI entered in real time is a value that can be commonly measured in the RSSI dataset of the corresponding AP stored in the database. Through this process, a high weighted value is applied for RSSs with high reliabilities and those with low reliabilities are eliminated or the weighted values are reduced and applied for positioning. From the $Mod\_Dist(i)$ calculated for each AP, the final positions, P, of users with the highest similarities are determined as follows:

$$P = argmin(Mod\_Dist(i)) \tag{13}$$

where $argmin$ is the argument of the minimum. Thus, the position value that is most similar, among the values calculated by applying the weighted values, is used to assess the final user position, which is then displayed.

4 Experiment and results

4.1 Experimental environment

To verify the performance of the proposed algorithm, an approximately 150-m long corridor on the fourth floor of the first engineering college building at the Korea Maritime and Ocean University was used as the experimental area, as shown in Fig. 4. As this experimental area included fixed APs in addition to active private Wi-Fi networks, various Wi-Fi AP signals could be collected. As shown Fig. 5, A

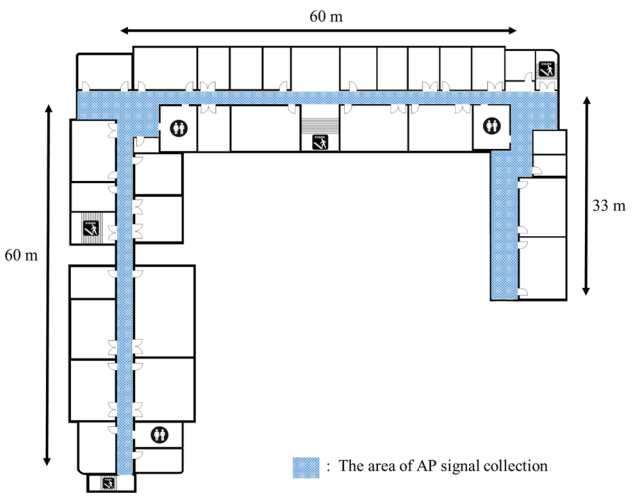
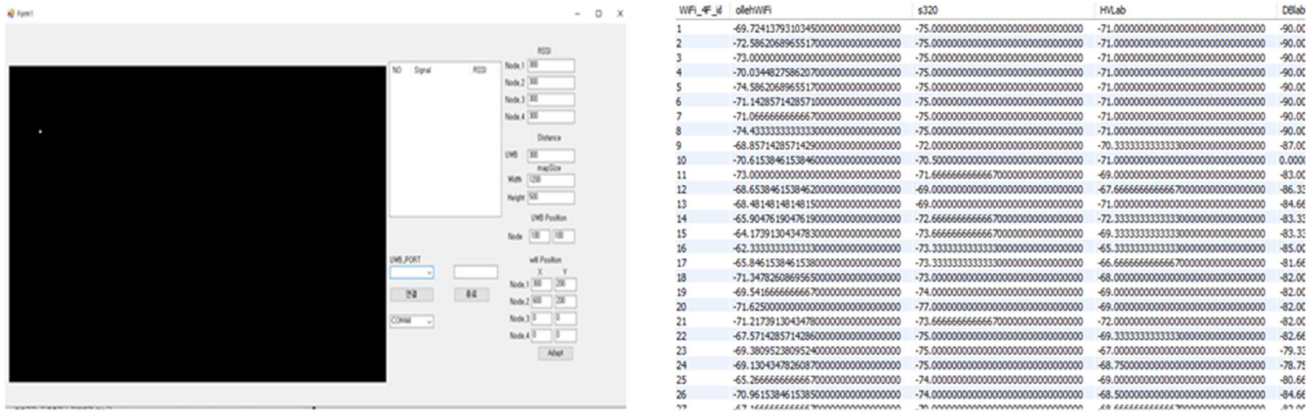


Fig. 4 Experimental environment for the positioning



(a) (b)

Fig. 5 The radio map used in the experiment. **a** Signal receive program(C#), **b** the initial radio map

program was created in C#, which could store the service set identifier (SSID) and RSS or Wi-Fi in real-time for collecting the AP signals measured in this area; 53 reference points were set at 3 m intervals and the Wi-Fi signals were measured at each point. A total of 139 APs were measured in the experimental area and the initial radio map (measurement count $\times$ 139) was created by the RSSs measured from these APs. 3-m is the most commonly used interval in Wi-Fi radio map together with 2, and 3-m is mainly applied in indoor space where there are many disturbances. Using the proposed algorithm, the radio map was optimized and the positioning was estimated, based on this initial radio map.

4.2 Performance evaluation of the proposed radio map algorithm

Figure 6 shows the IG values of data divided using the MDLP. The *X*-axis represents the measured APs and the Y axis the derived IG values. If the IG values obtained by MDLP are closer to zero, the classification of the AP RSS is not clear, according to the reference points. Therefore, the RSS data sets of APs with IG = 0 cause position errors or increase the computation in the positioning phase. In the proposed algorithm, because there is no precise thresholding standard using the IG values, we eliminate unnecessary APs with zero values. To evaluate the performance of the proposed MDLP method, the collected Wi-Fi RSSs were evaluated by the class-attribute contingency coefficient (CACC) and class-attribute interdependence maximization (CAIM) algorithms, which are division method

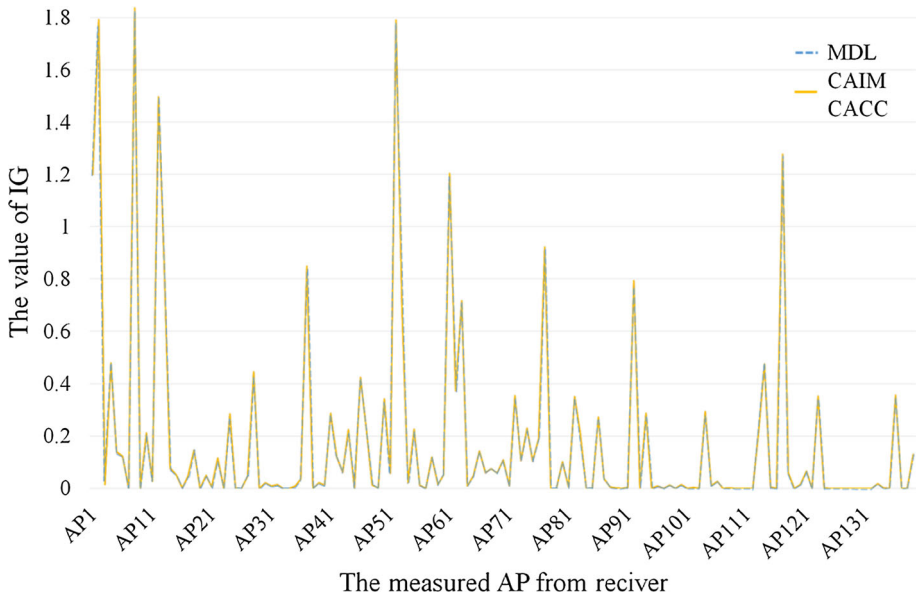


Fig. 6 IG values of the radio map using MDLP

discretization algorithms equivalent to the MDLP, for performance comparison, as shown in Fig. 6. The *X*-axis represents the measured APs and the *Y*- axis the differences in the IG values of the proposed algorithm, the CACC, and the CAIM, respectively.

The IG results using MDLP, as depicted in Fig. 7, and the CACC and CAIM results were 0.004, demonstrating that there were nearly no differences in the values. However, because the CACC and CAIM have similar operation processes and are classifiers that do not use IG values, the proposed algorithm based on MDLP is more advantageous for eliminating APs, based on the IG and can reduce unnecessary operations.

Table 1 presents a comparison of the radio map sizes optimized by the proposed algorithm, the CACC, and the CAIM, respectively. In the initial radio map 139 APs were used; with the proposed algorithm, the CACC, and the CAIM, it was optimized to 94, 97, and 97 APs, respectively. Compared to CACC and CAIM, MDLP has better data resolution and performance because it can acquire information gain value and use it for classification operation.

To estimate the user's actual position based on the radio map optimized by the proposed MDLP method, the Euclidean distance, based on a Chi squared test, was applied in the positioning phase.

The availabilities of RSSs measured in real-time were determined using the RSSs of the proposed radio map, with a Chi squared test and α (importance values) between 0 and 1 were deduced according to the results, as shown in Fig. 8. The *X*-axis represents the APs of the proposed radio map and the *Y*-axis the importance values based a Chi squared test, using the radio map. The importance values were deduced using a Chi squared test for the RSSs measured in real-time, after calculating the radio maps using a Chi squared test at each AP. It can be observed that the importance value differ in accordance with the RSS measured in real-time and that the importance of each AP differs accordingly. The importance values are applied such that the importance changes, when the RSS measured in real-time is not regular and has a significant deviation from the surrounding environment or receiver performance (obstacles, wall material, structures, etc.) in order to increase the position accuracy using highly reliable signals. If the importance value is zero, the RSS is not within the range of the radio map and the relevant AP can be eliminated during actual calculation.

Figure 9 depicts the comparison between the positioning performances of the Euclidean distance and the proposed algorithm based on radio maps with MDLP. The *X*-axis represents the reference points and the *Y*-axis the number of measurements. A total of 300 positioning results, which are the values that accurately measure the position, were measured for each position. The experimental results demonstrate that while there were positions where the performance was weak compared to the existing Euclidean distance method, the positioning accuracy at 54 reference points improved by an average of 5% compared to the Euclidean distance method.

The proposed Euclidean distance method based on a Chi squared test used an average of 84 APs for calculation, which is a 10.6% decrease from the existing Euclidean distance method that used 94 APs, as shown in Table 2. According to the reference points of 139 AP signals

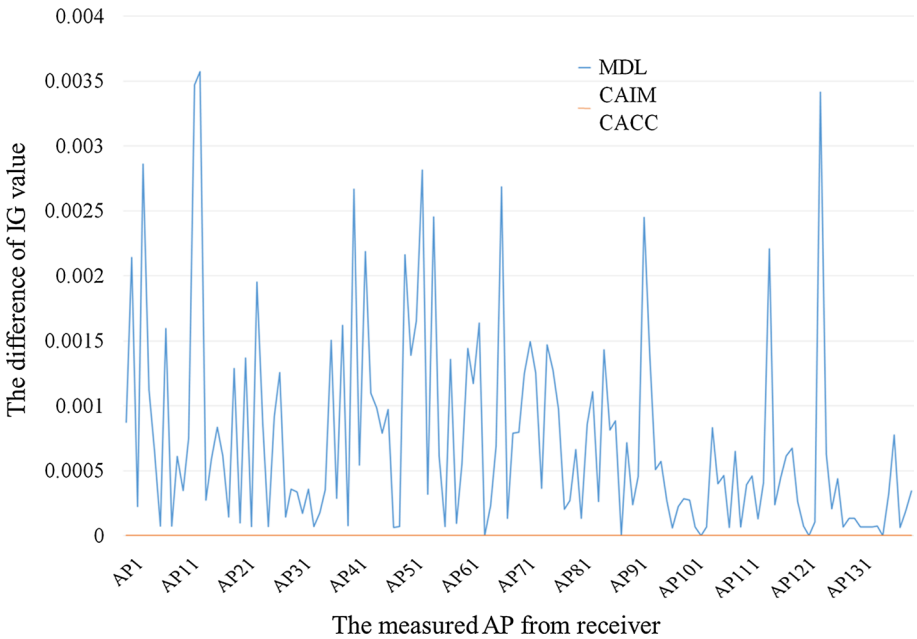


Fig. 7 Comparative differences between the IG values using MDLP, CACC, and CAIM

Table 1 Sizes of radio maps optimized with the MDLP, CAIM, and CACC

The adapted classifier	Initial radio map	Radio map based on the proposed algorithm	Radio map based on CAIM	Radio map based on CACC
Number of APs	139 APs	94 APs	97 APs	97 APs
Size of the radio map	100%	67.6%	69.8%	69.8%

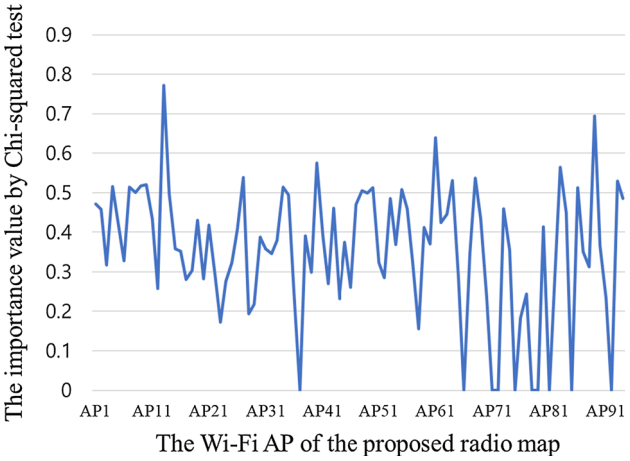


Fig. 8 RSSs of APs measured in real-time and the importance values based on a Chi squared test, using radio maps

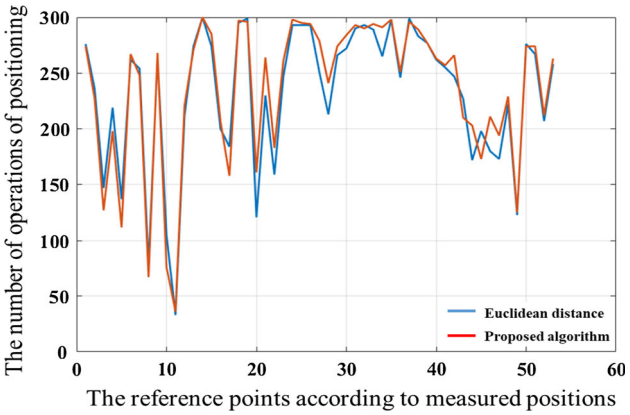


Fig. 9 Position accuracy according to the measurement count with respect to the reference point

acquired by the initial Wi-Fi reception, APs with irregular RSSs have been removed by MDLP and additionally Euclidean distance method based on a Chi squared test to minimize the radio map. Therefore, when compared with the initial radio map, the radio map was reduced by 39.6% through the proposed algorithm.

5 Conclusion

A Wi-Fi-based fingerprint, which combines a novel MDL-based radio map automatic classification and optimization algorithm, and a positioning algorithm based on the Chi squared test, was proposed in this paper. The proposed MDLP-based radio-map automatic classification algorithm uses an AP-based MDLP discretization method, which is entirely different from the extensively used radio-map-classification SVM method based on reference points, for classification; it uses the IG, based on entropy, for Wi-Fi RSSs that are continuously measured from the APs to automatically determine the suitability of data. The discretization performance was compared and analyzed using CAIM and CACC, discretization methods equivalent to the MDLP; the proposed MDLP algorithm was found to be marginally superior in terms of the AP elimination performance, compared to the other methods. The proposed Euclidean distance algorithm, uses the Chi square distributions estimated for each AP based on radio maps, for extracting the importance of RSSs measured in real-time. This importance was applied to the Euclidean distance. RSSs that were not measured by probability in this process were excluded from the calculation; the weighted values for accurate signals, in accordance with the importance, improved the position accuracy. Compared to the Euclidean distance algorithm in which weighted values were not applied, the accuracy of the proposed algorithm improved by an average of 5% and the number of APs were reduced by an average of 10.6%, on calculating 300 times for each reference point on the proposed radio map. Therefore, the size of the initial radio map decreased by 38.56% with the proposed fingerprint, while the positioning performance improved by approximately 5%. The proposed fingerprint

Table 2 The size of the radio map according to algorithms

The adapted algorithm	Original radio map	Proposed radio map (only MDLP)	Proposed Fingerprint
Number of APs	139 APs	94 APs	84 APs
Size of the radio map	100%	67.6%	60.4%

is a supervised learning method; after RSS collection, it automatically generates radio maps for estimating the position, with excellent AP elimination performance. However, because RSS elimination is vulnerable, there is a limit on decreasing the radio-map dimension based on the RSS. Therefore, methods that can decrease the dimension using data compression techniques must be researched in future.

Acknowledgements This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant No. 2016R1D1A1B03934812).

References

1. Arain, Q. A., Memon, H., Memon, I., Memon, M. H., Shaikh, R. A., & Mangi, F. A. (2017). Intelligent travel information platform based on location base services to predict user travel behavior from user-generated GPS traces. *International Journal of Computers and Applications, 39,* 1–14.
2. Arain, Q. A., Deng, Z., Memon, I., Zubedi, A., & Mangi, F. A. (2017). Map services based on multiple mix-zones with location privacy protection over road network. *Wireless Personal Communications, 97*(2), 2617–2632.
3. So, J. M., Lee, J. Y., Yoon, C. H., & Park, H. J. (2013). An improved location estimation method for Wi-Fi fingerprint-based indoor localization. *International Journal of Software Engineering and Its Applications, 7*(3), 77–86.
4. Arai, K., & Tolle, H. (2013). Color radio map interpolation for efficient fingerprint WiFi-based indoor location estimation. *International Journal of Advanced Research in Artificial Intelligence, 2*(3), 10–15.
5. He, J., Geng, Y., Liu, F., & Xu, C. (2014). CC-KF: Enhanced TOA performance in multipath and NLOS indoor extreme environment. *IEEE Sensors Journal, 14*(11), 3766–3774.
6. Ahmed, H. I., Wei, P., Memon, I., Du, Y., & Xie, W. (2013). Estimation of time difference of arrival (TDoA) for the source radiates BPSK signal. *IJCSI International Journal of Computer Science Issues, 10*(3), 1694–1784.
7. Seong, J.-H., Choi, E.-C., Lee, J.-S., & Seo, D.-H. (2017). High-speed positioning and automatic updating technique using Wi-Fi and UWB in a ship. *Wireless Personal Communications, 94*(3), 1105–1121.
8. Yiu, S., Dashti, M., Claussen, H., & Perez-Cruz, F. (2017). Wireless RSSI fingerprinting localization. *Signal Processing, 131,* 235–244.
9. Jung, S.H., Moon, B.-C., & Han, D. (2017). Performance evaluation of radio map construction methods for Wi-Fi positioning systems. *IEEE Transactions on Intelligent Transportation Systems, 18*(4), 880–889.
10. Figuera, C., Rojo-Álvarez, J. L., Wilby, M., Mora-Jiménez, I., & Caamaño, A. J. (2012). Advanced support vector machines for 802.11 indoor location. *Signal Processing, 92*(9), 2126–2136.
11. Jiang, Q., Ma, Y., Liu, K., & Dou, Z. (2016). A probabilistic radio map construction scheme for crowdsourcing-based fingerprinting localization. *IEEE Sensors Journal, 16*(10), 3764–3774.
12. Kjærgaard, M. B. (2007). A taxonomy for radio location fingerprinting. In *International Symposium on Location-and Context-Awareness*. Berlin: Springer.
13. Mirowski, P., Ho, T. K., & Whiting, P. (2014). Building optimal radio-frequency signal maps. In *2014 22nd International Conference on Pattern Recognition (ICPR)*. IEEE.
14. Jiang, Q., Ma, Y., Liu, K., & Dou, Z. (2016). A probabilistic radio map construction scheme for crowdsourcing-based fingerprinting localization. *IEEE Sensors Journal, 16*(10), 3764–3774.
15. Kim, Y., Shin, H., Chon, Y., & Cha, H. (2015). Crowdsensing-based Wi-Fi radio map management using a lightweight site survey. *Computer Communications, 60,* 86–96.
16. Chang, K., & Han, D. (2014). Crowdsourcing-based radio map update automation for wi-fi positioning systems. In *Proceedings of 3rd ACM SIGSPATIAL International Workshop Crowdsourcing Volunteered Geographic Information* (pp. 24–31).
17. Laoudias, C., Eliades, D. G., Kemppi, P., Panayiotou, C. G., & Polycarpou, M. M. (2009). Indoor localization using neural networks with location fingerprints. In *Proceedings of the 19th International Conference on Artificial Neural Networks: Part II, ser. ICANN'09* (pp. 954–963). Berlin: Springer.
18. Brunato, M., & Battiti, R. (2005). Statistical learning theory for location fingerprinting in wireless lans. *Computer Networks, 47*(6), 825–845.
19. Wu, Z., Fu, K., Jedari, E., Shuvra, S. R., Rashidzadeh, R., & Saif, M. (2016). A fast and resource efficient method for indoor positioning using received signal strength. *IEEE Transactions on Vehicular Technology, 65*(12), 9747–9758.
20. Tran, D. A., & Pham, C. (2014). Fast and accurate indoor localization based on spatially hierarchical classification. In *2014 IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)* (pp. 118–126). IEEE.

Ju-Hyeon Seong He received his B.S. and M.S. degrees in electrical and electronics engineering from Korea Maritime and Ocean University, South Korea, in 2012 and 2014, respectively. He is currently pursuing the Ph.D. degree in electronics engineering at Korea Maritime and Ocean University. His research interests include positioning system, sense network and embedded signal processing.

Dong-Hoan Seo He received his B.S., M.S., and Ph.D. degrees in electronic engineering from Kyungpook National University, South Korea, in 1996, 1999, and 2003, respectively. Since 2004, he has been with Korea Maritime and Ocean University, where he is currently a professor in the Division of Electronics and Electrical Information Engineering. His research interests include sense network, signal processing, and computer vision.