IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 1, January 2015.
www.ijiset.com
ISSN 2348 – 7968

Indoor Location in WLAN Based on Competitive Agglomeration Algorithm

Qing Jiang, Kunpeng Li*, Mu Zhou, Zengshan Tian

Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, P. R. China

Abstract
In the area of Wireless Local Area Network (WLAN) based indoor localization, the k-nearest neighbors (KNN) fusion clustering algorithm has been studied extensively. But the number of the clustering and the value of K is set manually and fixed, so it can't adapt to the environment changes. Besides, the algorithm localization with a single Received Signal Strength (RSS), and ignored other deeper information such as the physical location information. Aiming at the shortcomings of the fusion algorithm, in this paper, we proposed a novel indoor localization algorithm based on competitive agglomeration (CA). The algorithm soft partition radio map based on RSS and physical location information in succession, and select the clustering number based on real time information in the environment to estimate user's position coordinates. Finally, based on the extensive experiments conducted in a real WLAN indoor environment, our proposed algorithm is proved to outperform traditional positioning algorithm.

Keywords: *Indoor Localization, Location Fingerprints, Competitive algorithm, Soft partition.*

1. Introduction

With the rapid development of economy and mobile internet, people have become increasingly demanding on context aware service, especially the location based service (LBS). The vigorous development of O2O (Online to Offline) business model makes people are increasingly concerned with positioning accuracy. High-precision indoor location becomes the core of LBS. In the outdoor, the Global Positioning System (GPS) is commonly used for navigating and positioning. Nowadays, people spend average 80% of the time in the indoor, where the GPS signal is difficult to be received due to the building barrier. In the meantime, people entering in a mobile internet era, average 80% of the intelligent equipment connect network in the indoor. WLAN has been widely deployed in commercial areas, office building, campus, airport, residential areas, hotels and other kinds of environmental. The widely deployment of the WLAN and the popularity of the intelligent equipment makes the indoor positioning technology based WLAN achieve positioning function without the hardware upgrades compared with other positioning technologies. Indoor positioning technology based on WLAN has become the mainstream of the indoor positioning technology.

WLAN indoor localization approaches are generally classified into two categories, the model-based localization approach and location fingerprint approach. The model-based localization approach is based on the RSS characteristics from the Access Points (APs) to the target. The propagation model, Angel of Arrival (AOA), Time of Arrival (TOA), and Time Difference of Arrival (TDOA) are the four types of techniques in model-based localization. Since the propagation channels are site dependent, the radio propagation model is compromised by multipath fading, the law of signal attenuation is very complex in the indoor, so the precision of the model-based indoor location is very unreliable [2]. The location fingerprint approach contains two phases, namely the off-line phase and on-line phase. In off-line phase, the RSS data from hearable APs are used to construct a radio map contains RSS and their corresponding physical locations. In on-line phase, the target locations are estimated based on the matching from the newly recorded RSS measurements against the radio map. The location fingerprint approach could effectively reduce the influence of the multipath propagation and improve localization accuracy[3]. At the same time, the location fingerprint approach is relatively simple operation.

The k-nearest neighbors (KNN) fusion clustering algorithm is a widely applied indoor localization algorithm. While the cluster number and the value of K in the fusion algorithm is set manually and fixed, and the matching with single RSS ignored other information [5]. Aiming at the shortcomings of the traditional fusion indoor localization algorithm, in this paper we present an algorithm of indoor localization based on competitive agglomeration. The indoor localization based on competitive agglomeration could select dynamically the cluster number to generate competitive classes, and choice the value of K based on the real environment information both RSS and the physical location information of RPs'. We proposed algorithm improve the localization accuracy with low complexity and low cost.

The rest of this paper is structured as follows. In Section 2, we introduce the overall localization system in detail including competitive agglomeration algorithm. Experimental results are provided in Section 3. Finally, we conclude the paper in Section 4.

2. CA indoor localization system

In fingerprint localization system, first select Reference points (RPs) in the target environment and collect RSS fingerprinting in the RPs establish radio map in the off-line phase. Then, in the on-line phase the target location is estimated based on the matching from the newly collected RSS data against the pre-constructed radio map. The indoor localization system is shown in Fig.1.

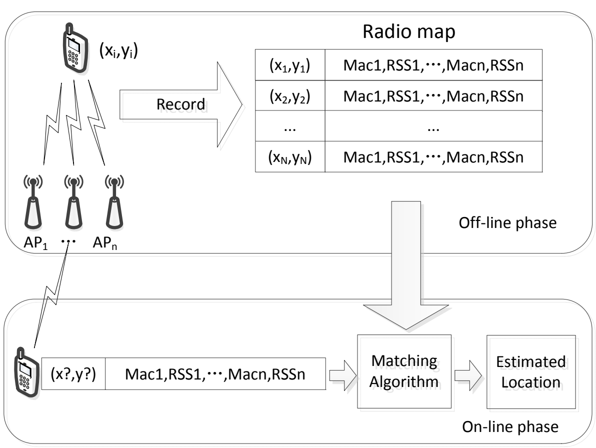


Fig. 1 Indoor localization system

KNN fusion clustering algorithm is extensively applied fingerprint localization algorithm. First, use clustering algorithm cluster radio map into several classes. Then, compute the Euclidean distance between clustering centers with the newly collected RSS. The Euclidean distances are defined as follows:

$$d(S,C_q)=\sqrt{\sum_{i=1}^{N}(s_i-c_{q,i})^2}\ ,(q=1,2,...,M) \quad (1)$$

Where $S=[s_1,s...,s_N]$ is the newly collected RSS vector in target location. The N is the number of access points (APs) in the environment. $C_q=[c_{q,1},c_{q,2},...,c_{q,N}]$ is the qth clustering center's vector. The M is the number of clustering. $d(S,C_q)$. M denotes the Euclidean distance between newly collected RSS vector in target location with qth clustering center. Choose the cluster whose clustering center's Euclidean distance is nearest to the newly collected RSS as the target cluster class among all clusters. Last, K RPs from target cluster is chosen according to the first k minimum Euclidean distance between RPs' RSS vector with the newly collected RSS. The target location is calculated as follows:

$$(\hat{x},\hat{y})=\frac{1}{K}\sum_{j}^{K}(x_j,y_j)\ ,(j=1,2,...K) \quad (2)$$

Where $(x_j,y_j),(j=1,2,...K)$ is the K RPs chosen from target cluster. $(\hat{x},\hat{y})$ is the estimated target location coordinates.

For traditional fusion algorithm, parameter M and K must be fixed. The fixed values of M and K will influence the performance of the localization system with the change of the circumstances. In order to solve the problem, we put forward to competitive agglomeration algorithm in this paper. Competitive Agglomeration could soft partition radio map and get different competitive clusters according to the dynamic environment. Besides, the traditional fusion algorithm localization with single RSS similarity and ignore RPs' physical location information. It's easy to choose some wrong RPs which RSS is similar with the RSS of target location, while physical location is far away from target location, such as shown in Fig.2.

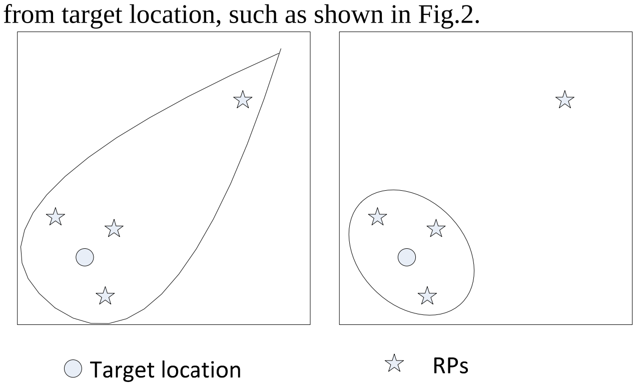


Fig. 2 Schematic diagram of the shortcomings of the traditional location algorithm

In order to solve the shortcomings and improve the accuracy of the localization, our proposed algorithm that use twice competitive agglomeration soft partition Radio map respectively based on RSS and physical location information.

2.1 Competitive r Agglomeration Algorithm

Competitive agglomeration is efficient algorithm in clustering algorithm theory. The CA algorithm doesn't need set an accurate cluster number prior to clustering. In CA, the initial partition has an over-specified number of clusters, which is progressively reduced until an optimum is reached. Also, the CA algorithms can be rapidly converged in a few iterations regardless of the initial number of clusters, and can also converge to the same optimal and most competitive partition regardless of its initialization. Its principle is described as follows:

Set $X = \{x_1, x_2, \cdots, x_n\}$ is the set of all the data points, and n is the number of data points. At the same time set $V = \{v_i \mid 1 \le i \le C\}$ as the set of clustering centers, where C is the number of clustering. Let u_{ij} denote the membership degree of x_j belonging to ith class, then we can define the fuzzy c-partition U of the given data set. i.e., $U = \{u_{ij} \mid 1 \le i \le C, 1 \le j \le n\}$. U is a $C \times n$ fuzzy classified matrix. U satisfies the formula of (3) :

$$1 = \sum_{i=1}^{C} u_{ij} \quad j \in \{1, 2, \dots, n\} \quad (3)$$

The CA minimizes objective function of CA is generally formulated as:

$$J = \sum_{i=1}^{C}\sum_{j=1}^{n}(u_{ij})^2 d^2(x_j, v_i) - \alpha \sum_{i=1}^{C}[\sum_{j=1}^{n}(u_{ij})]^2 \quad (4)$$

Where $d(x_j, v_i)$ denotes the Euclidean distances between data point x_j and the ith clustering center. α is a balancing item. The first term given in the equation (4) controls the size and the shape of the clusters. The second term controls the number of clusters. Lagrange coefficient is used in equation (4) to minimize the objective function. The steps of CA are described as follows:

Step 1: Set an initial value of the number of clustering $C_{\max}$ which is a relatively large number. i.e., $C = C_{\max}$. And initialize the iterative times $k = 0$. Initialize the fuzzy c-partition matrix U . Compute the Clustering cardinality $N_i (i = 1, 2, \dots, C)$ as equation (5)

$$N_i = \sum_{i=1}^{n} u_{ij} \quad (5)$$

Step 2: Compute the cluster centers by

$$v_i = \frac{1}{\sum_{j=1}^{n}(u_{ij})^2} \sum_{j=1}^{n}(u_{ij})^2 x_j \quad (6)$$

Step 3: Compute the membership degree and modify fuzzy classified matrix U as follows.

$$u_{ij} = u_{ij}^{FCM} + u_{ij}^{Bias} \quad (7)$$

Where the first component in equation (7) u_{ij}^{FCM} considers only the relative distances of a data point to all clusters and is computed as the equation [8]. And the second component in equation (7) is the deviation of the membership degree which is computed as the equation (9).

$$u_{ij}^{FCM} = \frac{[1/d^2(x_j, v_i)]}{\sum_{i=1}^{C}[1/d^2(x_j, v_i)]} \quad (8)$$

$$u_{ij}^{Bias} = \frac{\alpha(k)}{d^2(x_j, v_i)}(N_i - \bar{N}_j) \quad (9)$$

In the equation (9) $\bar{N}_j$ is a simple weighted average of the cardinality and is computed as equation (10)

$$\bar{N}_j = \frac{\sum_{i=1}^{C}[1/d^2(x_j, v_i)]N_i}{\sum_{i=1}^{C}[1/d^2(x_j, v_i)]} \quad (10)$$

$\alpha(k)$ is the value of the balancing item in the kth iterative, k is the iteration index. $\alpha(k)$ is computed as equation (11)

$$a(k) = \eta(k)\frac{\sum_{i=1}^{C}\sum_{j=1}^{n}(u_{ij})^2 d(x_j, v_i)}{\sum_{i=1}^{C}(\sum_{j=1}^{n}u_{ij})^2} \quad (11)$$

$\eta(k)$ in the equation (11) is computed as equation (12)

$$\eta(k) = \eta_0 exp(-k / \tau) \quad (12)$$

Where η_0 is the initial value and τ is the time constant. η_0 and τ could be valued as some constant according to particular cases. In our experiments, the initial value η_0, the time constant τ are set 5 and 10,respectively[8]. So, $\eta(k)$ in our experiments could be expressed as equation (13)

$$\eta(k) = 5exp(-k / 10) \quad (13)$$

Step 4: Compute the iteration stop threshold ε . Discard the spurious clusters whose cardinalities drop below a threshold clusters. However the clustering is sensitive to the threshold parameter. In this paper, we investigate an alternative process to improve the agglomeration. The details of the process are given as

ISSN 2348 – 7968

equation (14).

$$\varepsilon = \frac{8\cdot[\sum_{i=1}^{C}\sum_{j=1}^{n}u_{ij}]}{10\cdot C} \quad (14)$$

Step 5: Modify the number of clustering C based step 4. If the number of the clustering unchanged then the iteration is finished, else modify iteration index k=k+1 and turn to step 2.

2.2 Localization system design

We proposed localization algorithm in this paper consists of two main modules: the i) soft partition the radio map based on RSS through competitive agglomeration and chose the most similarity cluster with the RSS vector collected in the target location. ii) soft partition the cluster chosen in step i) base on the physical location information and estimate the physical location coordinates of the target location. The overview of the localization algorithm based competitive agglomeration is described by a flow chart shown in Fig. 3.

The localization algorithm is generally divided into four steps. Firstly, partition the radio map based on the RSS of the RPs and new collected in the target location with CA. Choose the RPs which RSS is in the cluster which contains the RSS of new collected in the target location as the RSS similarity cluster.

Second, compute the sum of each RP to all other RPs Euclidean distance respectively, and choose the global nearest point as equation (15) and equation (16)

$$D_i = \sum_{j=1}^{m} d(\mathrm{RP}_i, \mathrm{RP}_j) \quad (15)$$

$$\min(D_1, D_2, ..., D_m) \quad (16)$$

Where in the equation (15), m is the number of the RPS in the RSS similarity cluster. The global nearest point is the RP which index is the same as the index calculated as equation (16).

Third, partition the RSS similarity cluster based on the physical location coordinates with CA. Choose the RPs which location coordinates is in the cluster which contains the global nearest point as the coordinates similarity cluster.

Forth, the coordinates of target location is estimated by compute the weighted mean value of coordinates of the RPS in the coordinates similarity cluster. The calculation formula of the target location is shown as equation (17) and.

$$(\bar{x}, \bar{y}) = \sum_{i=1}^{p} w_i \cdot (x_i, y_i) \quad (17)$$

$$w_i = u_{Ri} + u_{Ci} \quad (18)$$

Where in the equation (17) P is the number of the RPs of the coordinates similarity cluster. In the equation (18) u_{Ri} and u_{Ci} is membership degree of ith RP belong to the RSS similarity cluster and the coordinates similarity cluster respectively.

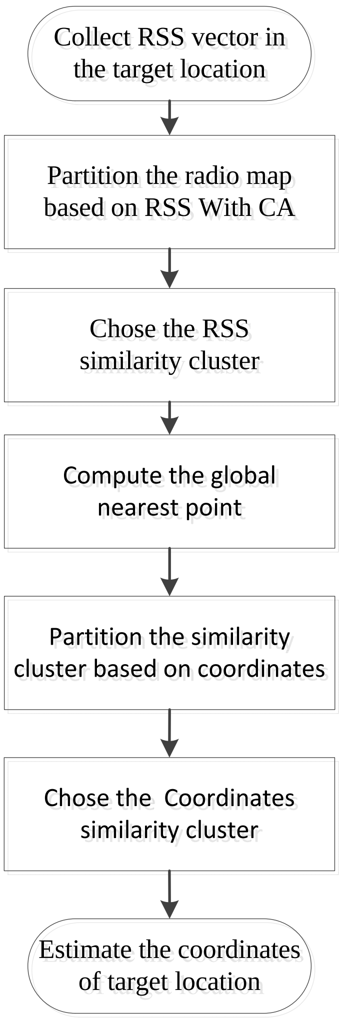


Fig. 3 The overview of the localization algorithm

3. Experimental Results

We conduct experiments in two typical environments, a simulated regular LOS environment and a real irregular Non-line-of-sight (NLOS) environment.

3.1 Simulated Regular LOS Environment

Fig. 4 shows a simulated regular LOS environment with the size of 20m by 20m. The 396 RPs with the same interval of 1m are uniformly distributed in this environment.

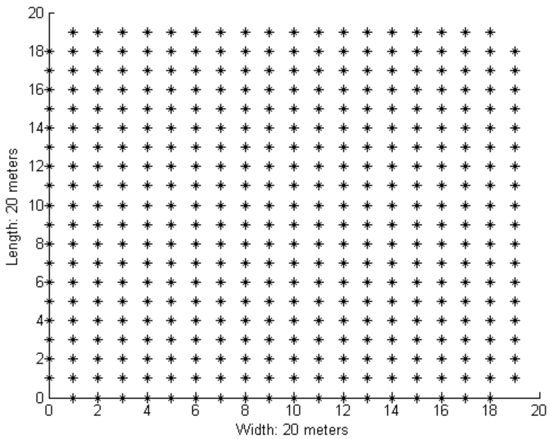


Fig. 4 Simulated regular LOS environment

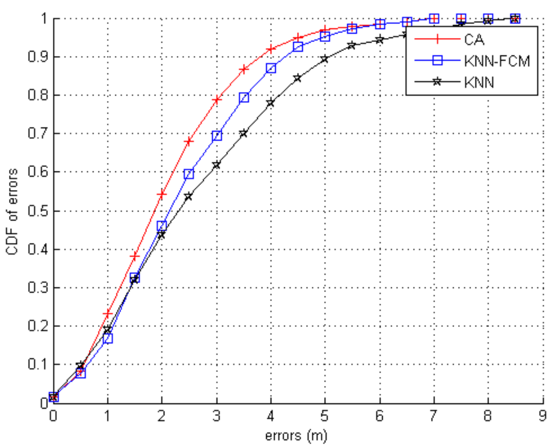


Fig. 5 CDFs of errors in the simulated regular LOS environment

Fig.5 shows the CDFs of errors with CA localization algorithm, KNN and KNN-FCM hybrid algorithm in the simulated regular LOS environment. From Fig. 5 , we can find that the proposed CA localization algorithm performs well in localization precision compared with KNN and KNN-FCM fusion algorithm. The average localization error of CA algorithm is 2.01m which is more accurate than both KNN-FCM fusion algorithm and KNN algorithm which average localization error is 2.28m and 2.4 3m respectively. The positioning accuracy improved 17%. The degree of confidence of CA localization algorithm in3m is 79% which is better than KNN and KNN-FCM fusion algorithm.

3.2 Experiments in Real Irregular NLOS Environment

Fig. 6 shows the layout of the experiment environment as well as the localization where we establish radio map. The size of the experiment environment is 36m by 21m. There are 5 Aps are placed in the environment to ensure the wireless signal coverage the entire area.

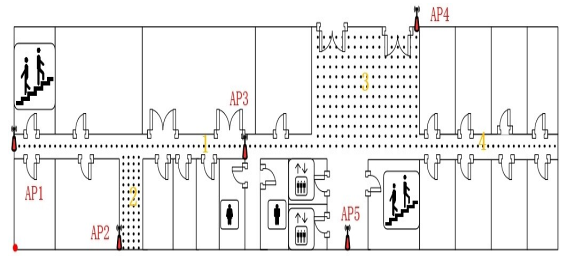


Fig. 6 The experiment environment

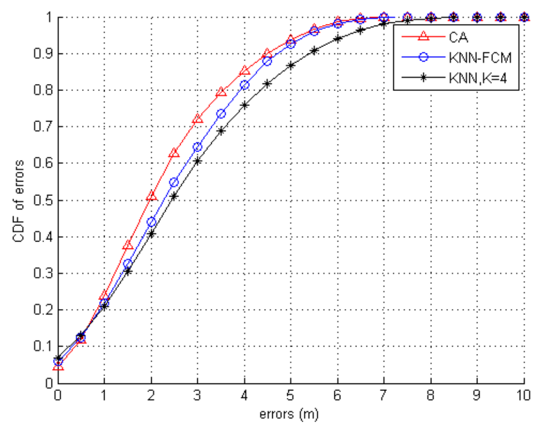


Fig. 7 CDFs of errors in the real irregular NLOS environment

degree of confidence of CA localization algorithm in3m is 72% which is better than KNN algorithm which degree of confidence in 3m is 64 and KNN-FCM fusion algorithm is 60%.

4. Conclusions

In this paper, we proposed a novel indoor localization algorithm based on competitive agglomeration. Through the algorithm the number of the clustering isn't fixed and changing based on the information of environment. Besides, the CA algorithm localization not only use the RSS information but also use the physical location information improve the accuracy of localization. For the future work, the computer expertise and localization in a more complex and multi-floor indoor environment forms an interesting topic.

Acknowledgments

This work was supported in part by the Program for Changjiang Scholars and Innovative Research Team in

ISSN 2348 – 7968

University (IRT1299), National Natural Science Foundation of China (61301126, 61471077), Special Fund of Chongqing Key Laboratory (CSTC), Fundamental and Frontier Research Project of Chongqing (cstc2013jcyjA40041, cstc2013jcyjA40032), Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ130528), Startup Foundation for Doctors of CQUPT (A2012-33), Science Foundation for Young Scientists of CQUPT (A2012-77), and Student Research Training Program of CQUPT (A2013-64).

References

[1] M. Zhou, A. K. Wong, Z. Tian, X. Luo, K. Xu, and R. Shi, "Personal mobility map construction for crowd-sourced Wi-Fi based indoor mapping," IEEE Communications Letters, 2014 vol. 18, no. 8, pp. 1427-1430.

[2] S. Jung, S. Hann, and C. Park, "TDOA-based optimal wireless indoor localization using LED ceiling lamps," IEEE Transactions on Comsumer Electronics, 2011, vol. 57, no. 4, pp. 1592-1597.

[3] Q. Chen, G. Huang, and S. Song, WLAN user location estimation based on receiving signal strength indicator, The 5th International Conference on Wireless Communications, Networking and Mobile Computing, 2009, pp.1-4.

[4] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," The 3rd International Conference on Mobile Systems, Applications, and Services, 2005, pp. 205-218.

[5] A. Savvides, W. Garber, S. Adlakha, R. Moses, and M. B. Strivastava, On the error characteristics of multihop node localization in ad-hoc sensor networks, Information Processing in Sensor Networks, 2003, pp. 317-332.

[6] M. Zhou, A. K. Wong, Z. Tian, V. Y. Zhang, X. Yu, and X. Luo, "Adaptive mobility mapping for people tracking using unlabelled Wi-Fi shotgun reads," IEEE Communications Letters, 2013,vol. 17, no. 1, pp.87-90.

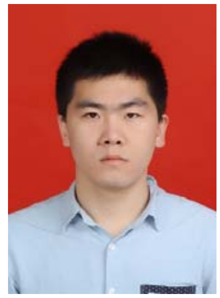
[7] P. Chai and L. Zhang, Indoor radio propagation models and wireless network planning, Computer Science and Automation Engineering, 2012, vol.2, pp. 738-741.

[8] H. FRIGUI,R. KRISHNAPURAM. Clustering by CA. Pattern Recognition, 2007, pp. 1109-1119

[9] C. Wu, Z. Yang, Y. Liu, and W. Xi, "WILL: wireless indoor localization without site survey," The 31st IEEE INFOCOM, 2012, pp. 64-72.

[10] L. Bruno, P. Robertson, "WiSLAM: improving footSLAM with WiFi," The 2nd International Conference on Indoor Positioning and Indoor Navigation, 2011, pp. 1-10.

Qing Jiang is a Professor at Chongqing University of Posts and Telecommunications, Chongqing, P.R. China. She received her B.Sc. in Communication Engineering form Chongqing University of Posts and Telecommunications, Chongqing, P.R. China. Her research interests include mobile wireless networks, mobile computing, and pervasive computing.



Kunpeng Li received the BE degree from the Department of Electronic Engineering of Huanghuai University, Henan, P.R.China. in 2013. He is currently working toward the B.Sc degree in the Chongqing University of Posts and Telecommunications His research interests include location determination technologies, and pervasive computing.