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

Indoor WLAN positioning Using Hybrid SVM Hyperplane Margin Clustering and Regression

Mu Zhou, Ming Xiang, Lingxia Li and Zengshan Tian

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

Abstract

This paper proposes a novel indoor Wireless Local Area Network (WLAN) positioning algorithm by using the Support Vector Machine (SVM) Hyperplane Margin Clustering and Regression (SVMCR). First of all, we rely on the SVM Hyperplane Margin Clustering (SVMC) to reduce the search space of the fingerprint database. Second, we use the Support Vector Regression (SVR) to characterize the relations of the Received Signal Strengths (RSSs) and physical locations for the sake of achieving the accurate positioning. Compared with the conventional indoor WLAN positioning algorithms, the proposed one significantly reduces the storage overhead, as well as guarantees the high positioning accuracy.

Keywords: *WLAN positioning, location fingerprinting, support vector machine, hyperplane margin, regression*

1. Introduction

In recent decade, the demand for indoor location-based services, like the personalized information delivery, medical services, and target discovery [1] has increased rapidly. Due to the existence of the multi-path fading and shadowing in indoor environment, the widely-used GPS and cellular network based location techniques cannot satisfy the accuracy requirement of indoor positioning. To solve this problem, the Received Signal Strength (RSS) based indoor Wireless Local Area Network (WLAN) positioning technique becomes significantly favored [2]. The indoor WLAN positioning technique is generally based on the existing indoor WLAN infrastructures and mobile terminals without any extra hardware devices, and meanwhile provides high positioning accuracy.

Most of the existed indoor WLAN positioning architectures are based on the concept of location fingerprinting [3], which consists of the offline and online stages. In offline stage, the RSSs from every hearable Access Point (AP) at each Reference Point (RP) are collected for the sake of constructing the location fingerprinting database (a.k.a. the radio map). After that, in online stage, when a location query occurs, the target locations are estimated by matching the newly collected RSSs against the pre-constructed radio map. The location fingerprinting based positioning technique is featured with the process of pattern recognition. Other pattern recognition approaches which have been applied to the indoor WLAN positioning can be found in [4, 5].

The performance of location fingerprinting based positioning technique seriously depends on the variations of RSSs, and meanwhile the growing number of individuals walking around in the target environment, as well as the multi-path fading will have significant impact on the variations of RSSs [6]. Therefore, with the purpose of achieving high positioning accuracy, a large number of location fingerprints are required to be collected in offline stage to characterize the relations between the variations of RSSs and physical locations. In this case, the offline stage always involves a large amount of time overhead. In online stage, high storage overhead is also needed to conduct the location estimation [7, 8].

The rest of this paper is structured as follows. In Section II, we introduce the overall system and the associated algorithm in detail. The experimental results are provided in Section III. Finally, we conclude the paper in Section IV.

2. System Description

Figure 1 shows the architecture of the proposed system. In offline stage, we collect the RSSs to construct the radio map, and then conduct Support Vector Machine (SVM) Hyperplane Margin Clustering and Regression (SVMCR) on the location fingerprints in radio map. In concrete terms, we first classify the RPs into different clusters by using the SVM Hyperplane Margin Clustering (SVMC). Second, the supervisory learning based Support Vector Regression (SVR) is applied for data training with the purpose of revealing the relations between the RSSs and physical locations. In online stage, we rely on the classification criteria acquired in offline stage to select the cluster which each newly collected RSS belongs to. After that, the SVR function corresponding to the selected cluster is used to estimate the target locations.

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

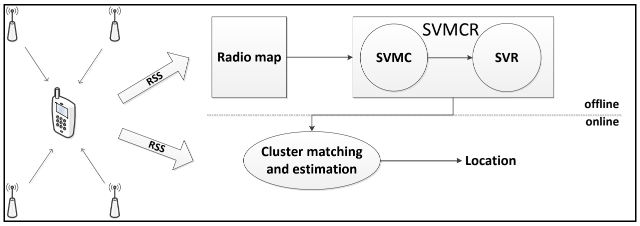


Fig. 1 System Architecture

2.1 SVM Hyperplane Margin

In SVM, if the data are linearly separable, we can select two canonical hyperplanes to separate the data. The distance between the two canonical hyperplanes is called the hyperplane margin. The purpose of SVM is to find the optimal separating hyperplane corresponding to the largest hyperplane margin. Based on the geometrical relations, the distances between the separating hyperplane and any point on the canonical hyperplanes equal to $1/\|w\|$. Therefore, the largest hyperplane margin is calculated by

$$L = \underset{w}{\arg\max} \frac{2}{\|w\|} \quad (1)$$

where w is the weighting vector [9]. The hyperplane margin is $L = 2/\|w\|$.

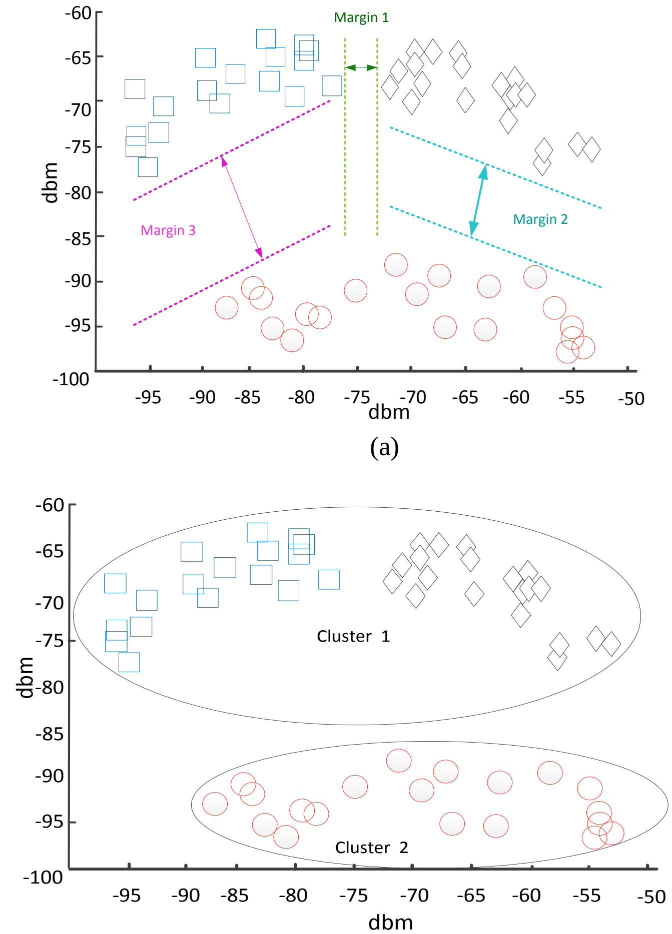
If the data are not linearly separable, an effective approach to guarantee the linear separability of the data is by using the kernel function [11]. The concept of kernel function is by first mapping the raw low-dimensional data into the one in a high-dimensional feature space, and then conducting the linear SVM on the mapped data. In this paper, the Gaussian kernel [12] is selected as the kernel function, as shown in (2).

$$k(x_i, x_j) = e^{-\gamma\|x_i - x_j\|} \quad (2)$$

where γ is the kernel width, which determines the similarity of x_i and x_j.

2.2 SVMC

Many conventional indoor WLAN RSS clustering algorithms are based on the Euclidean distances of RSS centroids at RPs. Without using the RSS centroids, Fig. 2 gives an example of clustering results by using the proposed algorithm, which takes the RSS distributions at



(b)

Fig.2 An example of clustering results by using the proposed algorithm. (a) The largest SVM hyperplane margins with respect to the RSSs at three RPs. (b) Results of clustering.

RPs into account, as well as relies on the SVM hyperplane margin to conduct the RSS clustering. The smaller SVM hyperplane margin between the RSS distributions of two RPs generally indicates the higher similarity of these two RPs.

The purpose of SVMC is to classify the M RPs into K clusters. We set $R = [r_1, r_2, \cdots r_M]$ as the set of RSS samples, where M is the number of RPs and $r_i\,(i = 1, \cdots, M)$ is the set of RSS samples collected at the i-th RP. At each RP, we collect N RSS samples with the dimensions of D which equals to the number of APs. The steps of SVMC are described as follows.

(1) Randomly select K RPs to construct the set of cluster

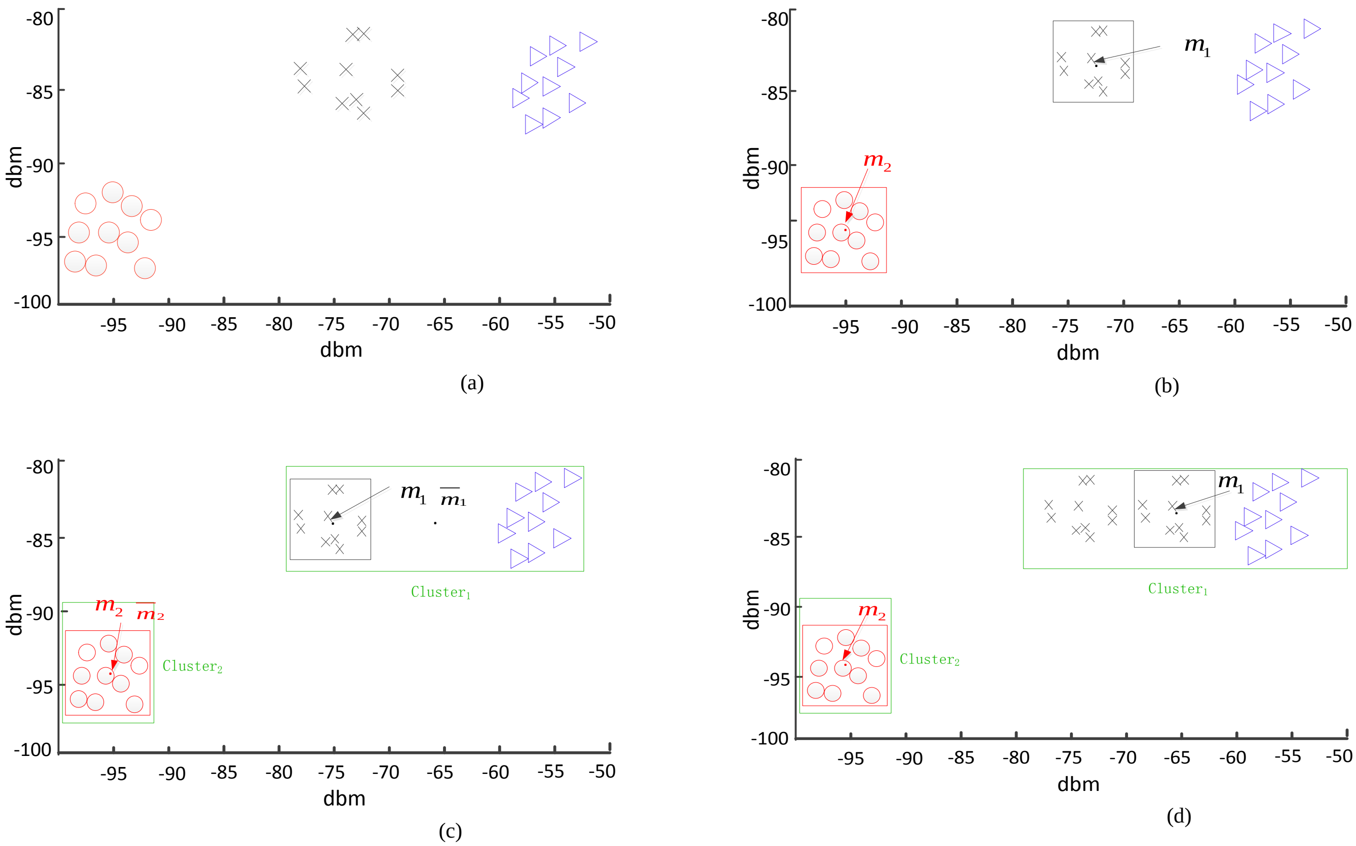


Fig.3 An example of the process of SVMC. (a) Distributions of RSSs. (b) Two cluster exemplars. (c) Classification of RPs . (d) Updating of the exemplars and centers of exemplars.

exemplars $E=[e_1,\cdots e_K]$, where $e_i=[e_{i1},e_{i2},\cdots e_{iN}]$ is the set of the N RSS samples at the i-th selected RP, RP_i, and $e_{ij}\,(j=1,\cdots N)$ is the j-th RSS sample at RP_i. The set of centers of cluster exemplars are notated as $m=[m_1,\cdots m_K]$, where $m_i=[m_{i1},\cdots m_{iD}]^{\mathrm{T}}\left(i=1,2,\cdots,K\right)$ is the center of the i-th cluster exemplar.

(2) Classify the RPs into the clusters with the cluster exemplars having the smallest SVM hyperplane margins to the RPs. After a new RP is classified into a cluster, we update the set of centers of cluster exemplars into $\overline{m_i}=[\overline{m_{i1}},\overline{m_{i2}}\cdots\overline{m_{iD}}]^{\mathrm{T}}$.

(3) Update the cluster exemplars by (3).

$$e_{ij}=e_{ij}-(m_i-\overline{m_i}) \tag{3}$$

(4) Repeat steps 2 and 3 until the set of centers of cluster exemplars is not updated.

Fig. 3 shows an example of the process of SVMC. The RSS samples at different RPs are plotted with different color symbols in Fig. 3(a). As shown in Fig. 3(b), the RSS samples plotted with black and red symbols are selected as the initial cluster exemplars and the corresponding two centers of cluster exemplars are located at m_1 and m_2. Then, in Fig. 3(c), the two clusters of RPs based on the SVM hyperplane margins are obtained, and meanwhile the two centers of cluster exemplars are updated into $\overline{m_1}$ and $\overline{m_2}$. Finally, the cluster exemplars plotted with black symbols are updated for the sake of calculating the new centers of cluster exemplars, as shown in Fig. 3(d).

2.3 Cluster Matching

SVM is originally developed for binary classification problem, whereas in indoor WLAN positioning, the number of clusters is generally larger than two. Thus, we propose to adopt the one-versus-one (OVO) scheme [10] to apply SVM to multi-class classification. Specifically, if the number of clusters is K, $\frac{K(K-1)}{2}$ binary classifiers are required to be trained. Each binary classifier, also named as sub-SVM model, is trained based on data from two clusters. To predict the cluster which a new sample belongs to, we rely on the $\frac{K(K-1)}{2}$ classifiers to calculate the vote of each cluster, and then select the cluster with the most votes as the winning cluster which the new sample belongs to.

2.4 SVR

In SVR, we set $(x_i, y_i)(i = 1 \cdots l)$ as the training data, where $x_i \in R^D$ is the *i*-th RSS sample, D is the number of Aps, y_i is the physical location corresponding to x_i. l is the number of RPs. We use the Gaussian kernel to nonlinearly map the raw low-dimensional data into the data in a high-dimensional feature space, and then conduct convex optimization processing to construct the SVR function in (4).

$$y = \sum_{i=SV} a_i K(x, x_i) + b \qquad (4)$$

where SV is the set of support vectors, x_i. b is a constant. After the process of offline SVR training, only a small number of parameters are required to be stored, which significantly reduces the storage overhead [12].

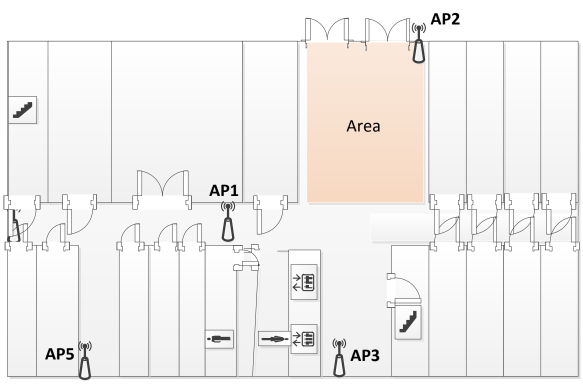


Fig.4 Experimental layout

3 Experimental Results

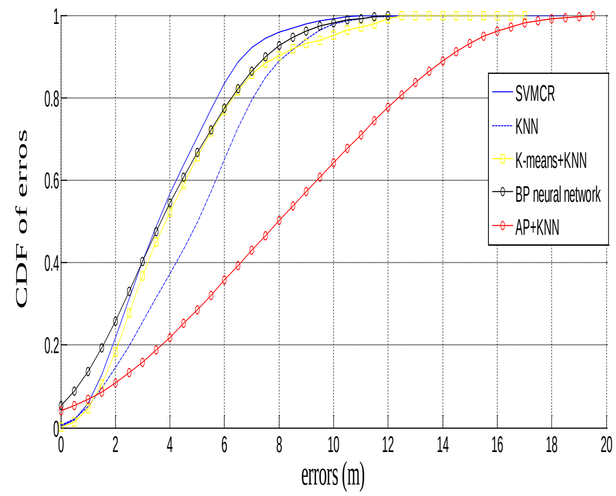


Fig. 5 Comparison of positioning accuracy

The proposed algorithm is evaluated in an actual indoor WLAN environment, as shown in Fig 4. The testing samples are collected in a $15m \times 10m$ lobby on the first floor in the Administration Building at CQUPT. There are 10 APs (TP-LINK DAP-2310) in total deployed in this environment. We select a Samsung GT-7568 smartphone to collect the RSS samples with the sampling rate of 1 Hz. The 54 RPs are uniformly calibrated with the interval of 1.6 m, and meanwhile the 40 Test Points (TPs) are randomly selected for the testing.

We compare the positioning errors achieved by the proposed algorithm and the other five conventional positioning algorithms, namely as the KNN, KNN with K-means clustering[7], KNN with Affinity Propagation (AP) clustering[8], and Back Propagation (BP) neural network [3-5], as shown in Fig.5. As can be seen from Fig. 5, we observe that the proposed algorithm performs well in positioning accuracy with the probability of errors in 5 m over 70%.

4. Conclusions

This paper proposes a novel SVMCR algorithm for indoor WLAN positioning for the sake of improving the accuracy of online positioning, as well as saving the cost for offline database construction. Furthermore, combining filter will forms an interesting work in future.

Acknowledgment

This work was supported in part by the Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), National Natural Science

ISSN 2348 – 7968

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] Munoz D, Boucherean F, and Vargas C, et a1. Position Location Techniques and Applications[M]. Burlington, USA， Academic Press．2009：17-21.

[2] Gu Y Y, Lo A, and Niemegeers I. A survey of indoor positioning systems for wireless personal networks. IEEE Communications Surveys and Tutorials, 2009, 1l(1): 13-32.

[3] Mengual L, Marb6n O, and Eibe S. Clustering-based location in wireless networks[J]. Expert Systems with Applications. 2010, 36(4): 7552-7561.

[4] Brunato M and Battiti R. Statistical learning theory for location ngerprinting in wireless LANs[J]. Computer Networks，2005，47(6): 825-845.

[5] Bahl P and Padmanabhan V N. RADAR: An in-building RF-based user location and tracking system. IEEE INFOCOM, 2000: 775-784.

[6] Carlos K F, Inmaculada M J, and Alicia G C, et at. Nonparametric model comparison and uncertainty evaluation for signal strength indoor location[J]. IEEE Transactions on Mobile Computing, 2009, 8(9) : 1250-1264.

[7] Y. Chen, Q. Yang, J. Yin, and X. Chai, "Power-Efficient Access Point Selection for Indoor Location Estimation," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 7, pp. 877－888, 2006.

[8] C. Feng, W. S. A. Au, S. Valaee, and Z. Tan, "Received Signal Strength Based Indoor Positioning Using Compressive Sensing," IEEE Transactions on Mobile Computing, vol. 11, no. 12, pp. 1983－1993, 2012.

[9] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 121－167, 1998.

[10] C.-W. Hsu and C.-J. Lin, "A Comparison of Methods for Multiclass Support Vector Machines," IEEE Transactions on Neural Networks, vol. 13, no. 2, pp. 415－425, 2002.

[11] Burnato M, Battiti R .Statistical learning theory for location fingerprinting in wireless LANs[J].Computer Networks. 2005,47:825 -845

[12] Zhang Z Z , Lin Z R..LIBSVM: a library for support vector machines. [J/OL].Taiwan.2006 http://www.cise. Ntu.edu.tw/~cjlin/libsvm.