

Research Article

A New Three-Dimensional Indoor Positioning Mechanism Based on Wireless LAN

Jiujun Cheng,^{1,2} Yueqiao Cai,² Qingyang Zhang,² Junlu Cheng,² and Chendan Yan²

¹ Department of Computer Science and Technology, Tongji University, Shanghai 201804, China

² Key Laboratory of Embedded System and Service Computing of Ministry of Education, Tongji University, Shanghai 201804, China

Correspondence should be addressed to Yueqiao Cai; caiyueqiao@163.com

Received 9 December 2013; Accepted 15 January 2014; Published 23 February 2014

Academic Editor: Huaicheng Yan

Copyright © 2014 Jiujun Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The researches on two-dimensional indoor positioning based on wireless LAN and the location fingerprint methods have become mature, but in the actual indoor positioning situation, users are also concerned about the height where they stand. Due to the expansion of the range of three-dimensional indoor positioning, more features must be needed to describe the location fingerprint. Directly using a machine learning algorithm will result in the reduced ability of classification. To solve this problem, in this paper, a “divide and conquer” strategy is adopted; that is, first through k-medoids algorithm the three-dimensional location space is clustered into a number of service areas, and then a multicategory SVM with less features is created for each service area for further positioning. Our experiment shows that the error distance resolution of the approach with k-medoids algorithm and multicategory SVM is higher than that of the approach only with SVM, and the former can effectively decrease the “crazy prediction.”

1. Introduction

With the development of mobile communication technology and the growing demand for new services, location-aware computing especially the Location Based Service (LBS) gradually attracts people’s attention, and how to determine the user’s location is the core issue in LBS. The global positioning system (GPS) [1] and the cellular network positioning system are location service systems widely used in open outdoor environment [2]. However, these positioning systems in an indoor environment are out of people’s satisfaction. And there are already many researches about the indoor positioning technology [3].

Nowadays, various radio techniques have been applied to the indoor positioning, such as UWB (ultra-wide band), RFID (radio-frequency identification), and VHF (ultra-high frequency) technology [4–7]. However, these positioning systems require redeployment of the network and additional signal measurement hardware, and the cost is relatively high, so their application is limited.

In recent years, with the increasing popularity of the application of WLAN, WLAN-based indoor positioning technology has been rapidly developed. The technology can

leverage existing wireless LAN resources without additional network deployments or facilities, which leads to the most attractive advantage, low cost.

WLAN-based indoor positioning technology research mostly uses the location fingerprint method and usually is based on the nearest neighbor search, naive Bayesian statistics, BP neural network or support vector machines [8–13], and other machines learning ideas. The positioning technology based on location fingerprint focuses more on the application of two-dimensional case, and relevant experiments are also conducted in a two-dimensional plane [14]. However, in practice the location systems in large shopping malls, libraries, offices, hospitals, airports, museums, and other places need to provide location information that includes not only the latitude and longitude coordinates, or other two-dimensional representations, but also the height or the room number.

When it comes to three-dimensional indoor positioning environments and the problem begins to be tougher, those methods previously mentioned may not be simply adaptive or scalable, and often a “divide and conquer” strategy acts as a good direction guide.

2. Related Work

Cotroneo et al. [15] proposed a naive partition positioning method that the communication range covered by an AP can be considered as a subregion, and which subregion the mobile terminal belongs to is determined by the signal strength it receives from each AP.

Xu et al. [16] divided the location space into multiple zones and used the distance-loss model choosing different parameters for each region. In the positioning phase, they used the maximum likelihood estimation to determine the location of the mobile terminal.

Nel Samama et al. [17] pointed out that many indoor positioning scenarios often do not require high accuracy and just tell the user some symbolic information, such as which corridor or room the current place is at. They proposed a 3D symbol positioning algorithm that divides the location space into different positioning symbolic subspace and designed a simple rule of symbolic subspace resolution to give the user positioning information.

Gansemmer et al. [18] pointed out that the 3D indoor positioning is more realized by UWB, RFID, and other technologies. They stressed the need for 3D WLAN indoor positioning and proposed a method that extends isolines algorithm [19], used in 2D WLAN indoor positioning to 3D space.

Zhong-liang et al. [20] adopted k-means clustering algorithm to partition a three-dimensional indoor space into multiple regions; namely, location fingerprints with similar Euclidean distance are clustered into one region and the central fingerprint of every region is saved. When in the positioning phase, the fact that the fingerprint received by the mobile terminal is closest to which central fingerprint helps estimate in which area the mobile terminal is located. But they just used this principle to determine which floor the mobile terminal is located on.

Various research works suggest and prove the effectiveness of the partitioning thought which will also be used in this paper. We will first use k-medoids algorithm to cluster the location space into different service areas to achieve coarse positioning and then build a multicategory SVM for each service area to achieve further fine positioning.

3. The Hybrid Indoor Positioning Approach

3.1. Cluster by k-Medoids

Definition 1 (location fingerprint). A location fingerprint is a vector bound to a specified location, which consists of a series of Wi-Fi signal strength values received by the mobile terminal from different APs (access point, Wi-Fi hotspot, or WLAN wireless router):

$$LF = [AP_1, AP_2, AP_3, \dots, AP_n]. \quad (1)$$

Definition 2 (record). A record here is a structure consisting of a location representation and a location fingerprint:

$$R = \{\text{location representation, location fingerprint}\}. \quad (2)$$

We adopt k-medoids [21] algorithm to partition the location space into different services areas. Compared to k-means

algorithm, k-medoids algorithm is not sensitive to outliers, which allows us to get a better partition center.

As the complexity of the original implementation of k-medoids (Partitioning around Medoids, PAM [22, 23]) is too high, we need only a method to partition the location space, especially to solve the problem that whether partitioning those locations in stairs into up-floor or down-floor service area is better. Hence, we design our implementation of k-medoids algorithm, as shown in Algorithm 1.

Note. $m_i \cdot \{\}$ is a set of records in i th service area with m_i as its medoid; the initial value of variable *changed* is true; the initial value of variable *count* is 0.

3.2. Classify by Multiclass SVM. The support vector machine (SVM) is a popular classification technique. Professor Chih-Jen Lin has done lots of researches deeply upon SVM for about many years, and he and his fellows or students developed and maintained a very useful SVM library LIBSVM [24].

The standard SVM is a binary classifier, but more often we need a multicategory SVM, especially in the indoor positioning problem, where a location is one category. In the library LIBSVM, there are three methods implementing multicategory SVM, namely, one-against-all, one-against-one, and DAG-SVM [25]. We will choose the DAG-SVM because the testing time of the DAG-SVM is less than the other two and the testing time is an online and time-sensitive operation.

In [26], Hsu et al. also presented some tricks on improving the performance of an SVM, such as scaling on input data and using cross-validation to get proprietary parameters of the RBF kernel.

3.3. Indoor Positioning Approach Combining k-Medoids and Multicategory SVM. The location space that we are now faced with is not limited to the scope of a room or a floor but has been extended to the whole building.

Take our college building as an example; usually there is an AP in a room and there are several APs in a corridor. So the total number of all APs in the whole building is very considerable.

Hence, the number of APs become large, so as the dimension of the location fingerprint. If an SVM is directly applied to classify the location fingerprints with large dimension, the ability to classify will decrease.

In the preparing phase, we will first through k-medoids algorithm use large dimension location fingerprint to partition the location space into several service areas and save a medoid location fingerprint for each service area. Thus, there are a set of location fingerprints (including the medoid itself) bound to each service area. Then we reduce the dimension of fingerprints in every service area through deleting those APs that are shared by none location in that service area. We will create a multicategory SVM upon each set location fingerprints. Detailed processing is shown in Algorithm 2.

In the positioning phase, when the mobile terminal receives many Wi-Fi signals from different APs, some of them are constructed as a fingerprint, and then the medoid

Input: $\{R_1, R_2, \dots, R_n\}$, R_i , i th record, total n records;
 $\{M_1, M_2, \dots, M_k\}$, M_i , i th initial medoid, total k medoids; $maxcount$, maximum iteration time.

Output: $\{m_1, m_2, \dots, m_k\}$, m_i , a record as i th clustered medoid.

- (1) For $i \leftarrow 1$ to k do
- (2) $m_i \leftarrow M_i$
- (3) EndFor
- (4) While $changed \ \&\& \ count < maxcount$ do
- (5) For $i \leftarrow 1$ to n do
- (6) $dist \leftarrow EuclideanDistance(R_i, m_1)$
- (7) $index \leftarrow 1$
- (8) For $j \leftarrow 2$ to k do
- (9) $temp \leftarrow EuclideanDistance(R_i, m_j)$
- (10) If $temp < dist$ do
- (11) $dist \leftarrow temp$
- (12) $index \leftarrow j$
- (13) EndIf
- (14) EndFor
- (15) add R_i to $m_{index}\{\}$
- (16) EndFor
- (17) For $j \leftarrow 1$ to k do
- (18) $centroid \leftarrow Average(m_j\{\})$
- (19) $m_j = record \in m_j\{\}$ nearest centroid
- (20) EndFor
- (21) If $\{m_j\}$ don't change do
- (22) $changed \leftarrow false$
- (23) EndIf
- (24) count++
- (25) EndWhile

ALGORITHM 1: Our implementation of k-medoids.

Input: $\{m_1\{\}, m_2\{\}, \dots, m_k\{\}\}$, $m_i\{\}$, set of records in i th service area with m_i as its medoid;

Output: $\{SVM_1, SVM_2, \dots, SVM_k\}$, SVM_i , trained SVM for i th service area;
 $\{AL_1, AL_2, \dots, AL_k\}$, AL_i , AP list of the fingerprints for i th service area.

- (1) For $i \leftarrow 1$ to k do
- (2) For $j \leftarrow 1$ to $m_i\{\}.length$ do
- (3) $r \leftarrow m_i\{\}.index[j]$
- (4) For $x \leftarrow 1$ to $r.fingerprint.length$ do
- (5) If $r.fingerprint[x] \neq NaN$ do
- (6) add $AP_{r.fingerprint[x]}$ to AL_i
- (7) EndIf
- (8) EndFor
- (9) EndFor
- (10) For $j \leftarrow 1$ to $m_i\{\}.length$ do
- (11) $r \leftarrow m_i\{\}.index[j]$
- (12) delete $AP \notin AL_i$ from r
- (13) EndFor
- (14) train SVM_i on new $m_i\{\}$ with LIBSVM
- (15) EndFor

ALGORITHM 2: Train an SVM for each service area.

fingerprint which is nearest to this finger can help determine which service area the mobile terminal is located in. Finally, the fingerprint modified by reducing dimension is input to the multicategory SVM corresponding to the determined service area, and the fine location is output by the SVM. Detailed processing is shown in Algorithm 3.

4. Experiment and Analysis

4.1. The Experiment Procedure. Our experiment is conducted on the 2nd, 3rd, and 4th floors in the building of the College of Electronics and Information Engineering (CEIE), Jading campus of Tongji University.

Input: fp, fingerprint of certain place;
 $\{m_1, m_2, \dots, m_k\}, m_i$, medoid of i th service area;
 $\{AL_1, AL_2, \dots, AL_k\}, AL_i$, AP list of the fingerprints for i th service area;
 $\{SVM_1, SVM_2, \dots, SVM_k\}, SVM_i$, trained SVM for i th service area.

Output: predicted location representation.

- (1) $dist \leftarrow \text{EuclideanDistance}(fp, m_1)$
- (2) $index \leftarrow 1$
- (3) For $j \leftarrow 2$ to k do
- (4) $temp \leftarrow \text{EuclideanDistance}(fp, m_j)$
- (5) If $temp < dist$ do
- (6) $dist \leftarrow temp$
- (7) $index \leftarrow j$
- (8) EndIf
- (9) EndFor
- (10) delete AP $\notin AL_{index}$ from fp
- (11) call SVM_{index} with input \leftarrow new fp

ALGORITHM 3: Predict the location of a certain place.

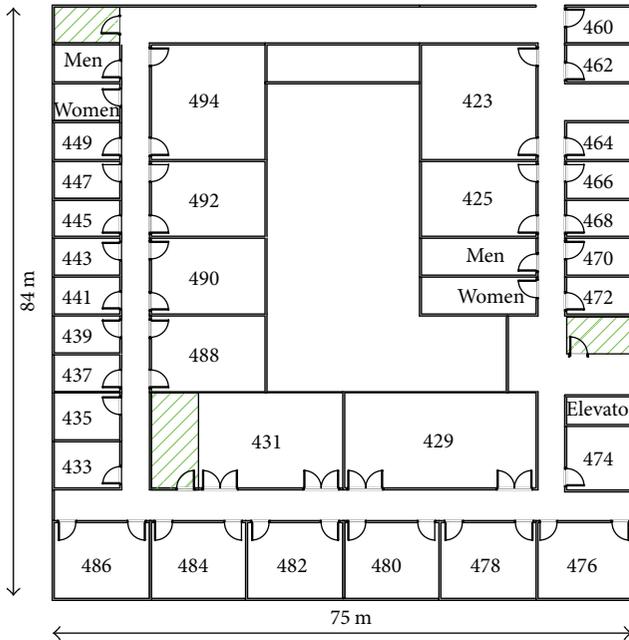


FIGURE 1: Experiment on the 4th floor of CEIE.

Figure 1 shows one of the experiment deployment floor plans, and the other two are similar. 100 sample spots are chosen from the corridor or stairs. These spots are labeled from number 1 to number 100 and every two neighbor spots are 5 m or 8 m or 12 m apart.

Table 1 lists two Wi-Fi info items received from nearby APs by an Android application (Wi-Fi scanner) developed by ourselves when the mobile phone is set in one of sample spots. The application can detect a list of info items (frequently there are more than 10 APs around) every time in the building of CEIE, and the situation can also be found in the library and dormitory buildings of the campus. Although we do not know the location of these APs, the information we can

TABLE 1: Two information items received from near APs.

Address of AP	00:21:29:7a:b2:b6
Network name	tongji419
Channel frequency	2437 MHz
Signal strength	-51 dBm
Address of AP	38:83:45:41:22:24
Network name	ciclab4
Channel frequency	2412 MHz
Signal strength	-58 dBm

get is enough for our experiment, and the “address of AP” and “signal strength” are used to construct a fingerprint of a location.

The main steps of our experiment are shown below.

4.1.1. The Preparing Phase

- (1) At every sample spot, with the mobile phone held in the hand of our tester, run Wi-Fi scanner 5 times for north, west, south, east, and a random direction, respectively; store the Wi-Fi info items for every time.
- (2) Delete those items with the “address of AP” that do not appear 5 times from the list of Wi-Fi info items, because their signal may be weak or unstable at that sample spot.
- (3) Count the frequency of every “address of AP,” sort those addresses by their frequency, reserve the first 2/3 of the sorted addresses, and discard the rest, 1/3, because we only need the address shared by more sample spots.
- (4) Shuffle the 2/3 of sorted addresses, and store them as a set of index indicators of a fingerprint, called Index Set.
- (5) Construct 5 fingerprints for every sample spot received from every 5 lists of Wi-Fi info items using

the principle that if the item has the “address of AP” in Index Set, the “signal strength” is assigned to the corresponding element with the same index of a fingerprint (the initial value of all elements is -100.0dBm), and these 5 fingerprints are bounded to the sample spot as one location or category for below SVM.

- (6) We get a set of fingerprints, and the size of this set is large. Because these data are collected from 9 corridors of 3 floors, we choose a fingerprint in every middle corridor to get 9 initial medoids; then we use Algorithm 1 with the input (the set of fingerprints, the 9 initial medoids, $maxcount = 100$) to partition them into 9 subsets for every subset as a service area and store the output (9 clustered medoids).
- (7) As every service area does not need so many “addresses of AP,” we can discard those addresses shared by none sample spot in every subset; in other words, we can reduce the dimension of the fingerprint in that service area.
- (8) Use the multicategory SVM of LIBSVM to train the data on every service area; then finally we get 9 SVMs for 9 service areas, respectively.

4.1.2. The Positioning Phase

- (1) Run Wi-Fi scanner at a testing spot; get a list of Wi-Fi info items; use the same method mentioned in the preparing phase to construct a new fingerprint.
- (2) Compare this new fingerprint with every clustered medoid (totally 9 medoids); choose the medoid nearest to the new fingerprint with Euclidean distance, and the testing spot is located corresponding to service area.
- (3) According to the subset of fingerprints of the service area, reduce the dimension of the new fingerprint and input the dimension-reduced fingerprint to the corresponding SVM; we finally get the category or location of the testing spot.

4.2. Testing and Analysis. We consider two testing scenarios to evaluate the performance of the hybrid indoor positioning approach.

One is in-place testing scenario and the other is middle-place testing scenario, as is shown Figure 2. The left and right circles represent two sample spots. In the in-place testing scenario, we choose every testing spot that is almost the same with certain sample spot; while in the middle-place testing scenario, we choose every testing spot that is in the middle of two neighbor sample spots.

In these two scenarios, the hybrid k-medoids + SVM approach is compared with the only SVM approach which does not have partitioning and reducing dimension steps and directly trains all fingerprints with one multicategory SVM.

Figure 3 shows the cumulative distribution function (CDF) of the error distance for the hybrid k-medoids + SVM and only SVM approaches methods in the in-place scenario,

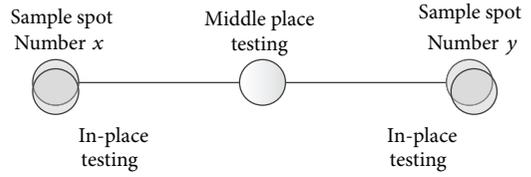


FIGURE 2: In-place and middle-place testing scenarios.

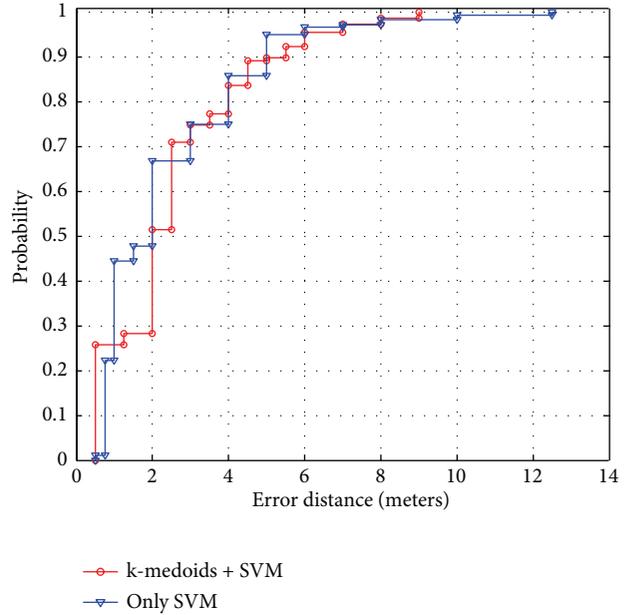


FIGURE 3: CDF of error distance/in-place scenario.

TABLE 2: The performance of two approaches.

Approach	50th percentile	75th percentile	Crazy prediction
k-medoids + SVM	6 m	8 m	0
Only SVM	7.5 m	9.3 m	3 times

while Figure 4 shows the counterpart in the middle-place scenario.

From Figure 3, in the in-place scenario, taking the 75th percentile for example, the error distance is less than 3.25 m with k-medoids + SVM approach and is less than 3.5 m with only SVM. We find that the difference between the performances of the two approaches is not very obvious. This is not hard to expect, because the in-place scenario represents an ideal situation, in which two approaches both achieve their best performance.

From Figure 4, we can easily find that the k-medoids + SVM approach performs better than only SVM approach in the middle-place scenario. The comparison in this scenario is listed in Table 2.

In Table 2, compared with only SVM approach, the k-medoids + SVM approach improves more than 1 m both in terms of the 50th percentile and 75th percentile of the error distance. Because the distance between two neighbor sample

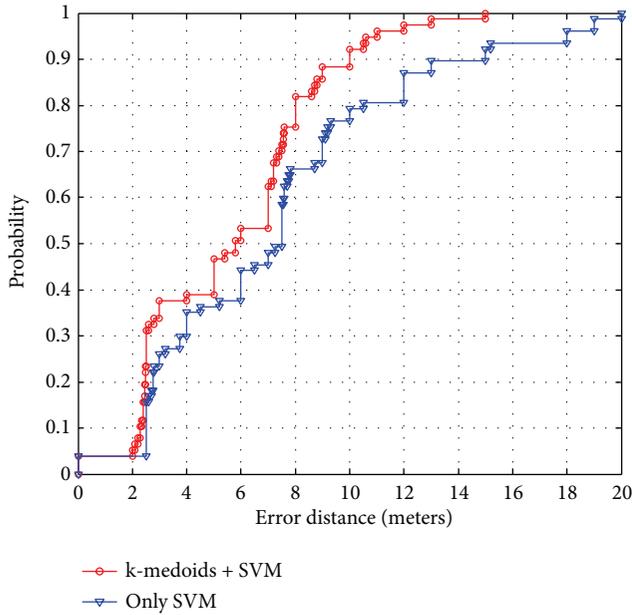


FIGURE 4: CDF of error distance/middle-place scenario.

spots in our experiment is 5 m, 8 m, or 12 m, the error distance hence becomes a little large. But indoor positioning resolution is not the key point we are concerned about in this paper. It is worth mentioning that with only SVM approach there occurs 3 times “crazy prediction,” for instance, that an actual spot on the 3rd floor is predicted on the 2nd floor or 4th floor. The k-medoids + SVM approach can well-reduce and even avoid this kind of “crazy prediction,” which also meets the requirement of indoor positioning, especially in 3D space.

5. Conclusions

A hybrid approach is proposed in this paper. The hybrid approach uses k-medoids algorithm to partition the set of fingerprints into several subsets, reduces the dimension of fingerprints of every subset, and trains a multicategory SVM on each subset data. The hybrid approach outperforms the approach just using SVM to train on all large-dimension fingerprints, in terms of error distance resolution. In addition, the hybrid approach with k-medoids algorithm and multicategory SVM can effectively reduce “crazy prediction.” Finally, we conclude that the hybrid approach can be used to solve 3D WLAN indoor positioning problem with a better performance.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the Natural Science Foundation Programs of Shanghai (Grant no. 13ZR1443100), by the 973

Program of China (under Grant 2010CB328101), by ISTCP (under Grant 2013DFM10100), by the National Science and Technology Support Plan (Grant no. 2012BAH15F03), and by NSFC (under Grants 51034003 and 51174210).

References

- [1] P. K. Enge, “The global positioning system: signals, measurements and performance,” *International Journal of Wireless Information Networks*, vol. 1, no. 2, pp. 83–105, 1994.
- [2] J. J. Caffery Jr. and G. L. Stüber, “Overview of radiolocation in CDMA cellular systems,” *IEEE Communications Magazine*, vol. 36, no. 4, pp. 38–45, 1998.
- [3] K. Pahlavan, X. Li, and J.-P. Mäkelä, “Indoor geolocation science and technology,” *IEEE Communications Magazine*, vol. 40, no. 2, pp. 112–118, 2002.
- [4] R. J. Fontana, E. Richley, and J. Barney, “Commercialization of an ultra wideband precision asset location system,” in *Proceedings of the IEEE Conference on Ultra Wideband Systems and Technologies*, pp. 369–373, Reston, Va, USA, November 2003.
- [5] M. N. Lionel, Y. Liu, Y. C. Lau, and A. P. Patil, “LANDMARC: indoor location sensing using active RFID,” *Wireless Networks*, vol. 10, no. 6, pp. 701–710, 2004.
- [6] J. Werb and C. Lanzl, “Designing a positioning system for finding things and people indoors,” *IEEE Spectrum*, vol. 35, no. 9, pp. 71–78, 1998.
- [7] H. Zhang, H. Yan, F. Yang, and Q. Chen, “Quantized control design for impulsive fuzzy networked systems,” *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 6, pp. 1153–1162, 2011.
- [8] P. Bahl and V. N. Padmanabhan, “RADAR: an in-building RF-based user location and tracking system,” in *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, pp. 775–784, March 2000.
- [9] M. A. Youssef, A. Agrawala, and A. U. Shankar, “WLAN location determination via clustering and probability distributions,” in *Proceedings of the 1st IEEE International Conference on Pervasive Computing and Communications (PerCom '03)*, pp. 143–150, Fort Worth, Tex, USA, March 2003.
- [10] R. Battiti, T. L. Nhat, and A. Villani, “Location-aware computing: a neural network model for determining location in wireless LANs,” Tech. Rep DIT-02-0083, 2002.
- [11] M. Brunato and R. Battiti, “Statistical learning theory for location fingerprinting in wireless LANs,” *Computer Networks*, vol. 47, no. 6, pp. 825–845, 2005.
- [12] H. Zhang, H. Yan, F. Yang, and Q. Chen, “Distributed average filtering for sensor networks with sensor saturation,” *IET Control Theory & Applications*, vol. 7, no. 6, pp. 887–893, 2013.
- [13] H. Yan, Z. Su, H. Zhang, and F. Yang, “Observer-based H_∞ control for discrete-time stochastic systems with quantisation and random communication delays,” *IET Control Theory & Applications*, vol. 7, no. 3, pp. 372–379, 2013.
- [14] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of wireless indoor positioning techniques and systems,” *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 37, no. 6, pp. 1067–1080, 2007.
- [15] D. Cotroneo, S. Russo, F. Corneville, M. Ficco, and V. Vecchio, “Implementing positioning services over an ubiquitous infrastructure,” in *Proceedings of the 2nd IEEE Workshop on Software*

- Technologies for Future Embedded and Ubiquitous Systems*, pp. 14–18, May 2004.
- [16] F.-Y. Xu, L.-B. Li, and Z.-X. Wang, “New WLAN indoor localization system based on distance-loss model with area partition,” *Journal of Electronics and Information Technology*, vol. 30, no. 6, pp. 1405–1408, 2008.
- [17] Nel Samama, “Symbolic 3D WiFi indoor positioning system: a deployment and performance evaluation tool,” in *Proceedings of the 13th World Congress of the International Association of Institutes of Navigation (IAIN '09)*, Stockholm, Sweden, October 2009.
- [18] S. Gansemer, S. Hakobyan, S. Püschel, and U. Großmann, “3D WLAN indoor positioning in multi-storey buildings,” in *Proceedings of the 5th IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS '09)*, pp. 669–672, September 2009.
- [19] U. Großmann, M. Schauch, and S. Hakobyan, “The accuracy of algorithms for WLAN indoor positioning and the standardization of signal reception for different mobile devices,” *International Journal of Computing*, vol. 6, no. 1, pp. 103–109, 2007.
- [20] D. Zhong-liang, W. Wen-jie, and X. Lian-ming, “A K-means based method to identify floor in WLAN indoor positioning system,” *Computer Engineering & Software*, vol. 33, no. 12, pp. 114–117, 2012.
- [21] L. Kaufman and P. J. Rousseeuw, “Clustering by means of medoids,” in *Statistical Data Analysis Based on the L_1 -Norm and Related Methods*, Y. Dodge, Ed., pp. 405–416, North-Holland, 1987.
- [22] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 3rd edition, 2006.
- [23] H. Yan, H. Shi, H. Zhang, and F. Yang, “Quantized H_∞ control for networked systems with communication constraints,” *Asian Journal of Control*, vol. 15, no. 5, pp. 1468–1476, 2013.
- [24] C.-C. Chang and C.-J. Lin, “LIBSVM: a Library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
- [25] 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.
- [26] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, “A practical guide to support vector classification,” Technical Report, Department of Computer Science, National Taiwan University, 2003.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

