A New WKNN Localization Approach

Amin Gholoobi
Faculty of Pure and Applied Sciences
Open University of Cyprus
Nicosia, Cyprus
Email: amin.gholoobi@st.ouc.ac.cy

Stavros Stavrou
Faculty of Pure and Applied Sciences
Open University of Cyprus
Nicosia, Cyprus
Email: stavros.stavrou@ouc.ac.cy

Abstract **— Location aware applications are constantly under development. By using an Indoor Location System (ILS), a user can localize himself, plan an indoor route and destination, or receive useful information and services in malls, airports, shopping centers, etc. To this purpose, various signal strength or timing based localization methods exist, each one having their own advantages and disadvantages. Localization based on the Received Signal Strength (RSS) methods, typically require only off-the-shelf equipment to operate. Moreover, they can utilize the existing infrastructure. This paper presents a new method on how to use the captured RSS values for localization purposes. The proposed method analyses a captured signal over a short distance and stores it into the fingerprint database for later comparison, rather than using an average value obtained from static measurements. Performance of proposed method is compared with a typical RSS-based localization method. Real-world measurements are used in order to validate our approach.**

Keywords - Indoor localization; location fingerprinting; RSS; faded signals; wireless LANs; WKNN

I. INTRODUCTION

Location based services (LBSs) based on Global Positioning System (GPS) [1], [2] created a revolutionary world. Today, localization plays a vital role in everyday activities. As the demand for localization applications rise, the localization methods improve and get customized to be applied to various environments and scenarios [3], [4]. GPS, GLONASS and Galileo might be the most commonly used technologies to achieve localization in outdoor environments; however they cannot typically provide localization in indoor environments.

Localization based on Received Signal Strength (RSS) proves to be more applicable to indoor environments [5], [6]. One reason is the implementation simplicity of RSS-based localization methods compared to other methods such as TOA [7], [8], TDOA [9], [10] and AOA [11], [12].

Accuracy, cost and complexity are some of the most important factors of an indoor localization system based on 802.11 wireless networks. RSS-based methods can dramatically decrease the cost and complexity of a localization system, since RSS measuring capabilities are available in almost all wireless systems. Moreover, the achievable accuracy when using RSS-based methods, are within the range of meters [13]-[16], which makes them a great candidate for typical indoor localization systems.

RADAR [13] was the first RF based technique for location determination in indoor environments. Some other RSS-based methods include QRSS [14], SpotOn [15] and Nibble [17]. RSS is used to characterize and map the indoor environment into distinguishable areas. To achieve that, a training phase (offline phase) is necessary in order to collect RSS values and save them to a Fingerprint Database (FDB). Localizing the target happens during the online phase. The Mobile Terminal (MT) gathers RSS values from surrounding Access Points (APs) and compares them with values stored in FDB to find the best match. Like many other localization techniques RSS-based methods typically need multiple APs to localize a target.

The proposed method in this paper suggests a novel way to perform the offline and online phases. The method characterizes the environment based on dynamic paths and not static points as the traditional methods do. Collecting RSS values over a path allows for the collection of faded signals. The collected faded signals help to provide a better statistical description of the localization environment aiming to lead to improved localization procedures. This paper is organized as follows: The proposed method is explained in section II. This includes offline phase, online phase and related algorithms. Section III presents the experiments that have been carried out to investigate the accuracy and reliability of the proposed method. Section IV summarizes the outcome and presents the conclusions of this study.

II. METHODOLOGY

Most of the localization methods based on RSS share the same fundamentals. First a database needs to be created which contains all the collected data from the offline/training phase. These data aims to assign unique characteristics to each point/path in the localization environment. Quantity and density of the collected points/paths in a localization environment depends on the environment's geometry, the expected localization accuracy and etc.

Depending on the scenario and with regards to the localization system implementation, either the infrastructure or the MT collects the RSS values related to the locations of interest.

The proposed method constructs a Radio Map (RM) of the localization environment based on the signal fading statistics of multiple short paths. Each entry in the FDB represents a short path in the localization environment. The length of a path for this method was arbitrary selected to be 1.8 m, which contains 3 × 60 cm floor tiles. To create the FDB, a MT moves along the paths in the area of interest and collects RSS values from the surrounding Access Points (APs). Collected data includes the fast faded signals received from the APs. The collection of the fast faded signals was evident by the collection of signal strengths that varied up to 40 dB on the same path.

Equation (1) presents the data that has to be collected for each path and saved in FDB:

$$P_{x,y} = (a1, a2, \ldots, ak) \tag{1}$$

Where *x, y*, represents the coordinates for center of the path on the map and a_k represents a RSS value for access point *k* derived from (2).

Equation (2) allows to calculate *a*, for each visible AP. Its components are *V*, the average of the squared differences from the sample mean (Variance), and *s* the sample standard deviation.

$$a_k = \frac{V^2}{2.(s^2)} \tag{2}$$

To address device diversity, all a_k values have to be normalized [18], [19]. Equation (3) calculates *N* which is the normalized vector for $A = (a_1,a_2,\ldots,a_k)$:

$$N = (a_{n=1}^{k} - Max(A)) \tag{3}$$

Prior to the localization process the radio map of the environment is created which concludes the offline/training phase.

During the online/localization phase, the MT collects RSS values from all visible APs. Next, all a_k values are calculated and normalized based on (2) and (3). Finally, the extracted normalized vector for the MT is compared to the radio map using a k-nearest neighbor (kNN) pattern matching technique [5], [20]. To find the nearest neighbors, first the Euclidean distance between the MT and each *P* in the FDB is calculated by using (4):

$$D_{RL} = \sqrt{\sum_{i=1}^{K} (TN_i - N_i)^2} \tag{4}$$

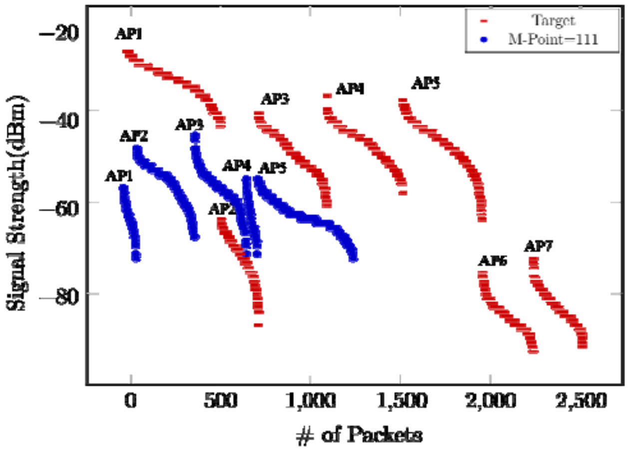

Figure 1. A false positive introduced as a nearest neighbor in traditional *k*NN algorithm.

Where *D* represents the Euclidean distance between the MT and an entry in the fingerprinting database, *RL* is the index of the reference position, *TN* is the normalized linear value of the MT signal strength vector, *N* represents the linear vector of the signal strength values as stored in FDB, *i* represents the index of the AP and *K* is the number of available APs.

After finding the nearest neighbors to the target, location of the MT is estimated by using a kNN algorithm and a weighted term given by (5) and (6).

Equation (5) shows the location estimation based on multiple nearest neighbors by using the weighted factor:

$$x = \frac{\sum_{m=1}^{M} \frac{1}{D_m} . x_m}{C} \qquad y = \frac{\sum_{m=1}^{M} \frac{1}{D_m} . y_m}{C} \tag{5}$$

Where *x, y* are the location coordinates, *M* is the number of nearest neighbors and *C* is the sum of calculated coefficients for all the selected neighbors. Equation (6) presents the calculation of *C*:

$$C = \sum_{m=0}^{M} \frac{1}{D_m} \tag{6}$$

Applying the weighted factor for the proposed method is a crucial step, as in some scenarios, values of *D* for selected nearest neighbors can have considerable differences. In such cases, estimating MT's location purely based on a simple average coordinates of nearest neighbors, might not produce accurate results.

A. *Neighbors selection*

The proposed method selects the nearest neighbors based on path analysis. Compared to traditional kNN methods where selection of the nearest neighbors is based on static single point measurements, the proposed method uses the data collected over a path. During our measurement campaign, we have found that by using traditional kNN methods, some of the selected neighbors were the result of

DOI 10.5013/IJSSST.a.16.06.07

ISSN: 1473-804x online, 1473-8031 print

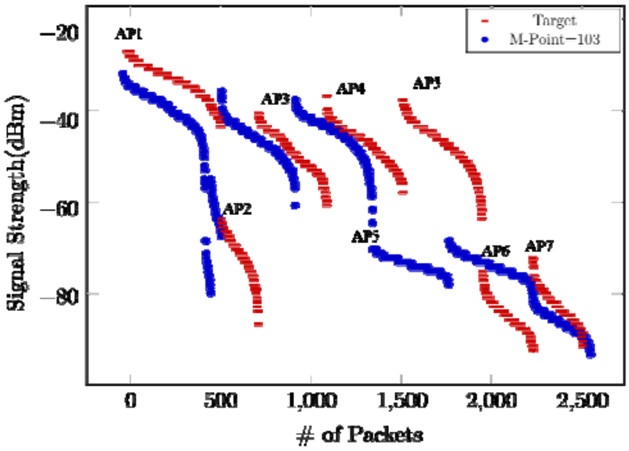

Figure 2. Sorted RSS signals of the target and a nearest neighbor found by the proposed method.

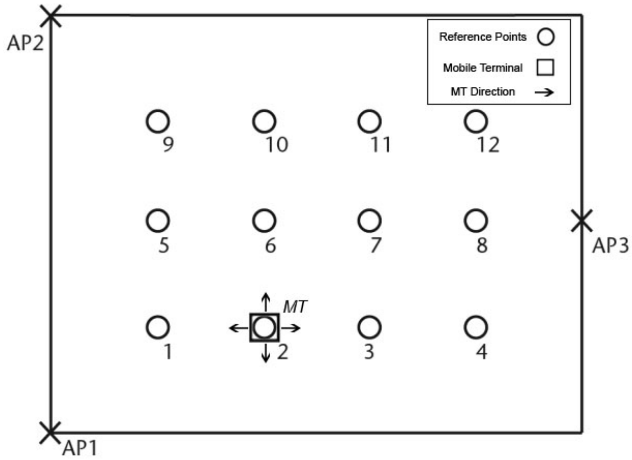

Figure 4. Typical RSS based localization setup.

false positives. Traditional methods which select neighbors based on the smallest distance (*D*) did not perform well specially in the areas where the radio signal could not attenuate well due to the characteristics of the environment (Fig. 1). As a result, fewer packets have been collected leading to a smaller *D* calculation for that point.

Fig. 1 shows the target signal (Red) and a nearest neighbor to it (Blue), as suggested by a traditional kNN algorithm.

By using the proposed algorithm, such errors can be eliminated since it does not select the nearest neighbors purely based on the *D* value. Fig. 2 presents the selected nearest neighbor for the same target signal by using the proposed method. The comparison of the visual presentation of RSS values for both methods, suggest that a selected neighbor chosen by the proposed method, shares more similarities with the target signals, as demonstrated in Fig. 2.

To select *k* nearest neighbors for an online measurement, one must first calculate distance *D* for all of FDB entries and carry out a sorting process in an ascending order. The raw data gets sorted based on AP names and RSS values (strongest to weakest) in an effort to better distinguish the closest neighbors and eliminate false positives. Based on the sorted data, a list is generated which includes strongest to weakest AP names by averaging the RSS values for each AP. By comparing the target's sorted AP list to FDB entries' list, the closest neighbors to the target's signal are selected. The proposed method's algorithm is outlined in Table I.

III. EXPERIMENTAL EVALUATION

A proof-of-concept localization system was developed to investigate the accuracy and reliability of the proposed method. The experimental setup consisted of eight APs (Linksys WRT54GL), one MT (Sony Vaio VPCYB2M1E) with three different Wireless LAN (WLAN) cards (To investigate device diversity), and an SQL Database which is located on MT itself. The APs have been distributed throughout the localization environment. Fig. 3 shows the Localization environment and AP locations marked by dark triangles.

Over a 100 reference points have been collected during the offline/training phase and recorded into the FDB. A reference point in the proposed method refers to a path with a 1.8 m length. The MT moves along a straight line (a reference point) with a steady pace, collecting signal strengths from all visible APs. The localization platform is implemented in C#.

A. Localization Performance

To evaluate the performance of the proposed method various measurements have been performed to estimate the location of the MT. The first location estimation was carried out using the proposed method. For comparison purposes a typical RSS based localization method [21] was implemented and another fingerprinting database was created for the latter method since this was based on static measurements. Similar to the proposed method, collected reference points for the comparison method cover the same areas and the same number of reference points. The terminal collects four sets of data (one for each direction), combines all measurements and saves the results in a fingerprinting database. Fig. 4 presents a typical RSS based localization technique. This was used for comparison purposes with the proposed method. Table II presents the achieved localization accuracy of both methods:

TABLE II. METHODS PERFORMANCE

	Avg.(m)	Max.(m)	Min.(m)	σ(m)	Var(m)
Proposed	1.9	4.8	0	1.4	2.1
RSS [21]	3	8.0	0.6	1.9	3.6

DOI 10.5013/IJSSST.a.16.06.07

ISSN: 1473-804x online, 1473-8031 print

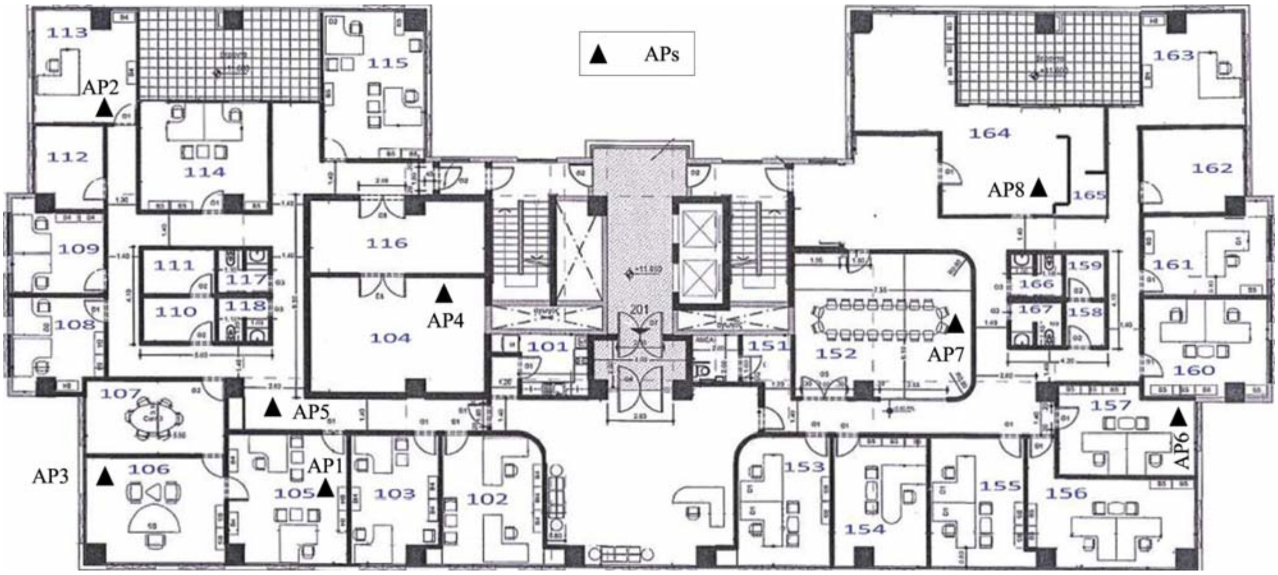

Figure 3. Floor plan and APs Distribution.

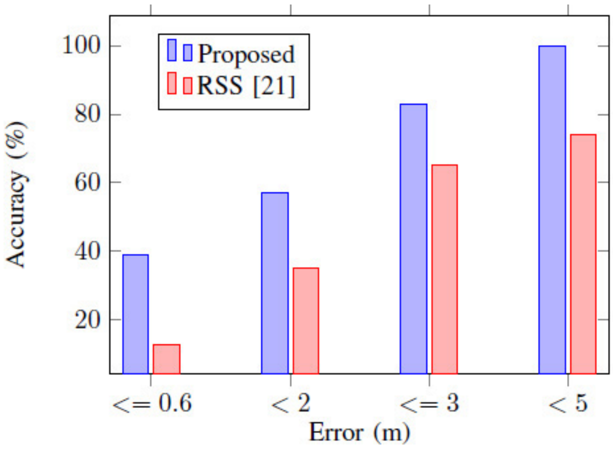

Figure 5. Methods Accuracy.

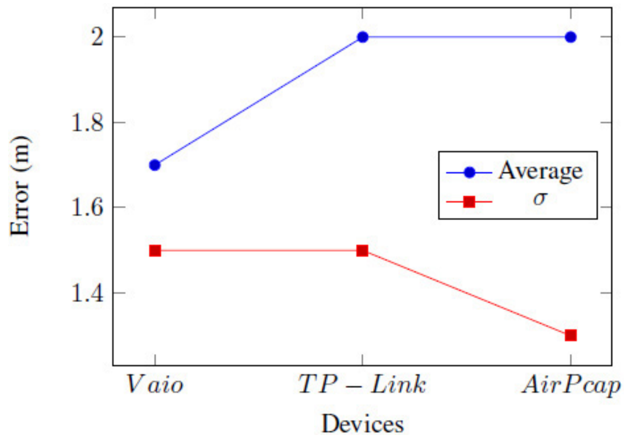

Figure 6. Device diversity and localization error.

Fig. 5 presents accuracy performance of the methods presented in Table II. The average accuracy of the proposed method was found to be less than 2 meters. The proposed method managed to localize the MT within a tile accuracy (60 cm) for almost 40% of the times. The typical localization accuracy achieved from some of the existing and well-known methods range between 2-6m [2], [13], [22], only 17% of the times, the proposed method performed with accuracy more than 3 meters.

B. k Value and Device Diversity

To test the proposed method for device diversity, localization tests have been repeated by using three different WLAN cards as follows: Atheros AR9285, TP-LINK TL-WN821N and AirPcap NX. Fig. 6 presents the average localization error for each of the devices, along with the estimated standard deviation (σ).

When considering a kNN algorithm, the number of nearest neighbors (*k*) can have an effect on localization results. To evaluate the effect of *k* on localization accuracy and stability, the collected online data have been tested by using different values for *k*. Collected samples were recorded in different environments, including office rooms, corridors and laboratories. Our experimental analysis suggests that the proposed method worked best when k=5. Averaging over *k* nearest neighbors shows that for k=5, the calculated distance error is less than 2 meters.

IV. CONCLUSION

In this paper a novel RSS-based localization method using path analysis was presented. While typical RSS methods utilize the average signal of surrounding APs recorded at stationary points, it was demonstrated that better localization results can be achieved by analyzing RSS signals

Table I. A New WKNN Approach Algorithm using Path Analysis

- Offline/Training Phase
  - For each path/point in RM (P_x, y)
    - Collect all RSS values for visible APs as (1)
      $P_{x,y} = (a_1, a_2, ..., a_k)$
    - Normalize all a values as (3)
      $N = (a_{n=1}^k - Max(A))$
    - Record it in FDB
- Online/Localization Phase
  - Collect all RSS values for visible APs $(a_1, a_2, ..., a_k)$
  - Extract the normalize vector for MT as (3)
    $N = (a_{n=1}^k - Max(A))$
  - Compare normalized vector for MT to RM using WKNN
    - For each P in FDB
      - Calculate Euclidean distance (4)
        $D_{RL} = \sqrt{\sum_{i=1}^{K} (TN_i - N_i)^2}$
      - Select the K nearest neighbors
    - Apply the Weighted term & Draw the location of MT
      - Calculate C the sum of calculated coefficients for K selected neighbors as (6)
        $C = \sum_{m=0}^{M} \frac{1}{D_m}$
      - Calculate x, y coordinates using (5)
        $x = \frac{\sum_{m=1}^{M} \frac{1}{D_m} \cdot x_m}{C}, y = \frac{\sum_{m=1}^{M} \frac{1}{D_m} \cdot y_m}{C}$

over a path. It was observed that the proposed method improves localization accuracy while decreasing the error in standard deviation. Moreover, for comparison purposes a typical RSS based localization method was implemented. The proposed method performed better in all the presented statistics. The average accuracy of the proposed method was found to be less than 2 meters. The proposed method managed to localize the MT within a tile accuracy (60 cm) for almost 40% of the times.

ACKNOWLEDGMENT

Research work is co-funded by the European Regional Development Fund (ERDF) and the Cyprus Research Promotion Foundation.

REFERENCES

[1] P. Enge and P. Misra, "Special issue on global positioning system," Proceedings of the IEEE, vol. 87, no. 1, pp. 3-15, Jan. 1999.

[2] Y. Liu, Z. Yang, X. Wang, and L. Jian, "Location, localization, and localizability," Journal of Computer Science and Technology, vol. 25, no. 2, pp. 274-297, 2010.

[3] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," Communications Surveys Tutorials, IEEE, vol. 11, no. 1, pp. 13-32, Jan. 2009.

[4] Z. Farid, R. Nordin, and M. Ismail, "Recent advances in wireless indoor localization techniques and system," Journal of Computer Networks and Communications, vol. 2013, pp. 1-12, 2013.

[5] T.-N. Lin and P.-C. Lin, "Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks," in Wireless Networks, Communications and Mobile Computing, 2005 International Conference on, vol. 2, pp. 1569-1574.

[6] C. Esposito and M. Ficco, "Deployment of RSS-Based Indoor Positioning Systems," Int'l J. Wireless Information Networks, vol. 18, no. 4, pp. 224-242, 2011.

[7] S. Golden and S. Bateman, "Sensor measurements for wi-fi location with emphasis on time-of-arrival ranging," Mobile Computing, IEEE Transactions on, vol. 6, no. 10, pp. 1185-1198, Oct. 2007.

[8] M. Youssef, A. Youssef, C. Rieger, U. Shankar, and A. Agrawala, "Pinpoint: An asynchronous time-based location determination system," in Proceedings of the 4th international conference on Mobile systems, applications and services. ACM, 2006, pp. 165-176.

[9] A. Günther and C. Hoene, "Measuring round trip times to determine the distance between wlan nodes," in NETWORKING 2005. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems. Springer, 2005, pp. 768-779.

[10] M. Mah, N. Gupta, and A. Agrawala, "Pinpoint time difference of arrival for unsynchronized 802.11 wireless cards," in INFOCOM, 2010 Proceedings IEEE, pp. 1-5.

[11] G. D. Stefano and A. Petricola, "A distributed aoa based localization algorithm for wireless sensor networks," Journal of computers, vol. 3, no. 4, pp. 1-8, 2008.

[12] S. Zekavat, A. Kolbus, X. Yang, Z. Wang, J. Pourrostam, and M. Pourkhaatoun, "A novel implementation of doa estimation for node localization on software defined radios: Achieving high performance with low complexity," in Signal Processing and Communications, 2007. IEEE International Conference on, pp. 983-986.

[13] P. Bahl and V. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 2, pp. 775-784.

[14] N. Patwari and A. O. Hero III, " Using proximity and quantized RSS for sensor localization in wireless networks," in Proc. ACM Int. Conf. Wirel. Sensor Net. And Appl., pp. 20-29, 2003.

[15] J. Hightower , R. Want and G. Borriello "SpotON: An indoor 3D location sensing technology based on RF signal strength", Univ. Washington, Seattle, Tech. Rep. UW CSE 2000-02-02, 2000.

[16] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," IEEE Trans. Syst., Man Cybernetics, Part C: Applications Rev., vol. 37, no. 6, pp. 1067-1080, Nov. 2007.

[17] P. Castro, P. Chiu, T. Kremenek, and R. Muntz, "A probabilistic location service for wireless networked environments," in Proc. Int. Conf. Atlanta Ubiquitous Computing, pp. 18-34, September 2001.

[18] T. Roos, P. Myllymki, H. Tirri, P. Misikangas, and J. Sievnen, "A probabilistic approach to wlan user location estimation," International Journal of Wireless Information Networks, vol. 9, no. 3, pp. 155-164, 2002.

[19] R. Zekavat and R. M. Buehrer, Handbook of position location: theory, practice and advances. John Wiley & Sons, 2011, vol. 27.

[20] V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piche', "A comparative survey of wlan location fingerprinting methods," in Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on. IEEE, pp. 243-251.

[21] A. Bensky, Wireless Positioning Technologies and Applications, Artech House, 2007.

[22] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless lans," Computer Networks, vol. 47, no. 6, pp.825-845,2005.