2011 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 21-23 September 2011, Guimarães, Portugal

Rank Based Fingerprinting Algorithm for Indoor Positioning

J. Machaj*, R. Piché** and P. Brida*

* University of Zilina, Department of Telecommunications and Multimedia, Zilina, Slovakia. Email: Machaj@fel.uniza.sk, Brida@fel.uniza.sk

** Tampere University of Technology, Tampere, Finland. Email: Robert.Piche@tut.fi

***Abstract*—A novel Received Signal Strength (RSS) rank based fingerprinting algorithm for indoor positioning is presented. Because RSS rank is invariant to bias and scaling, the algorithm provides the same accuracy for any receiver device, without the need for RSS calibration. Similarity measures to compare ranked vectors are introduced and their impact on positioning accuracy is investigated in experiments. Experimental results shown that the algorithm can achieve better accuracy than some commonly used fingerprinting algorithms.**

***Keywords*—Indoor positioning; Fingerprinting; Localization**

I. Introduction

The basic requirement for Location Based Services (LBS) [1] is knowledge of the mobile device position. This can be achieved in many different ways. Global Navigation Satellite Systems (GNSS), like GPS (Global Positioning System) or GLONASS (Global Navigation Satellite System), are widely used, and these systems work very well for outdoors, especially in areas with a clear view to the satellites. In dense urban environments GNSS can suffer from high signal attenuations and reflections, which can seriously degrade position estimate accuracy. In indoor environment the situation is even worse, as GNSS signals are mostly too weak to be received at all.

These drawbacks of GNSS have motivated the development of positioning algorithms that use signals from existing radio networks. These algorithms use different properties of radio signals. Most common in indoor environment are measurements of RSS (Received Signal Strength) and ToA (Time of Arrival). The work presented in this paper deals with RSS measurements, which have the advantage that they are available on almost every device.

Indoor positioning systems can be based on different wireless technologies, for example Bluetooth [2], UWB (Ultra Wide Band) [3] and WiFi (IEEE 802.11) [4-8]. This work deals with WiFi signals, because WiFi is the most common technology and it is supported by a wide range of devices.

Most indoor positioning systems based on WiFi use some kind of fingerprinting algorithm. In fingerprinting algorithms, measured RSS values stored in a database (known as a radio map) are compared to RSS values measured by the mobile device. A basic difficulty here is that because of hardware and software differences between different devices (even devices of the same make and model), the RSS reported by the mobile device may differ from the RSS in the database, and this can degrade the positioning accuracy [9].

One approach to dealing with this issue is to calibrate the RSS scale and bias for the device, for example using a self-calibration learning algorithm as proposed in [4].

In this paper we propose a novel fingerprint positioning algorithm that uses only the rankings of the RSS values. Because rank information is invariant to any monotonic increasing transformation (bias and scale), the algorithm's performance should be unaffected by the calibration of the mobile device.

The rest of the paper is organized as follows. In the next section related work in indoor positioning algorithms is introduced. Section III describes the proposed rank based algorithm. Similarity measures used in the algorithm are described in Section IV. Results of tests in a real environment are given in Section V and Section VI concludes the paper.

II. Related Work

A. Rank based localization

Rank based localization in wireless networks was introduced by Yedavalli et al. in [10]. Their Ecolocation algorithm uses a set of constraints to estimate the position of a mobile device. Measured RSS values from the APs (Access Points) that are within range are sorted and compared with constraints. Position is estimated as the centroid of points with the highest number of satisfied constraints.

An improved version of the Ecolocation algorithm was introduced in [11]. Bisector lines were introduced as lines connecting points with the same RSS values from two different APs. Position is estimated as a weighted mean of positions with the highest number of satisfied constraints. Another modification of the algorithm [12] uses the centroids of nearest three regions with the highest number of satisfied constraints estimated by the previous algorithm.

A drawback of these methods is their use of bisector lines, because signal propagation in indoor environment, where there are many obstacles, is not accurate and equal values of RSS from two APs are not in a straight line.

B. Fingerprinting localization

In fingerprinting localization, the position of a mobile device is estimated by comparison of measured RSS values and RSS values stored in a radio map database.

Fingerprinting algorithms have two phases – an offline learning phase and an online operating phase.

In the offline phase, the radio map database is created. The localization area is divided into small cells [6], and each cell is represented by a reference point. RSS values from all APs within range are measured and stored in the radio map, which is a collection of data vectors that can be described as:

$$P_j = (\alpha_1, \dots, \alpha_{N_j}, \theta_j) \qquad j = 1,2,\dots,M , \qquad (1)$$

where N_j is the number of APs heard at the j-th reference point, M is the number of reference points, α_i are RSS values, and parameter vector θ_j contains additional information that can be used in the localization phase.

In the online phase the mobile device measures RSS values from all APs within range. These values are compared to data stored in the radio map database. Algorithms used for comparison between RSS data from the two phases and estimation of position of mobile device can be divided into two main frameworks – deterministic and probabilistic.

In the probabilistic (or statistical) framework the mobile device's position is modeled as a random vector [13]. The location candidate γ is chosen if its posterior probability is the highest. The decision rule uses Bayes' theorem:

$$P(\gamma_i | S) = \frac{P(S|\gamma_i)P(\gamma_i)}{P(S)} , \qquad (2)$$

where posteriori probability $P(\gamma_i|S)$ is a function of likelihood $P(S|\gamma_i)$, prior probability $P(\gamma_i)$ and observed evidence $P(S) = \sum_i P(S|\gamma_i)P(\gamma_i)$, vector S represents the observed RSS values during online phase and γ_i stands for i-th location candidate.

The deterministic framework is based on optimizing the similarity between the measurement and the fingerprints. The position estimate is computed using the weighted average:

$$\hat{x} = \frac{\sum_{i=1}^{M} \omega_i \cdot \gamma_i}{\sum_{i=1}^{M} \omega_i} , \qquad (3)$$

where ω_i is a non-negative weighting factor. Weights can be calculated as the reciprocal of the distance between RSS vectors from online and offline phase. Usually the Euclidian distance is used but different distance metrics are also possible [14].

The estimator (3) which keeps the K largest weights and sets the others to zero is called the WKNN (Weighted K-Nearest Neighbor) method [7]. WKNN with all weights $\omega_i = 1$ is called the KNN (K-Nearest Neighbor) method. The simplest method, where $K = 1$, is called the NN (Nearest Neighbor) method. In [6] it was found that WKNN and KNN methods perform better than the NN method, particularly when values of parameter K are 3 or 4.

III. RANK BASED FINGERPRINTING

The main difference between conventional fingerprinting algorithms and the proposed Rank Based Fingerprinting (RBF) localization algorithm is the way in which measured data in offline and online phases are compared and used to estimate position. In classical fingerprinting algorithms, vectors of RSS values measured in online and offline phase are directly compared to each other.

In the proposed algorithm (Fig. 1) the RSS values measured in the online phase from different APs are first sorted from strongest to weakest. Then ranks (1, 2, 3, …) are assigned to APs based on their position in the sorted vector. The sorted vector of APs detected in the online phase is then compared to vectors stored in the radio map. Rank vectors are created for vectors stored in the database. Ranks are assigned based on the MAC (Media Access Control) address of AP and the rank of the AP in online phase. In case that one (or more) of the APs from the online phase is not found in the database, the rank vector created from the radio map is padded with 0, to achieve the same length as the rank vector from the online phase.

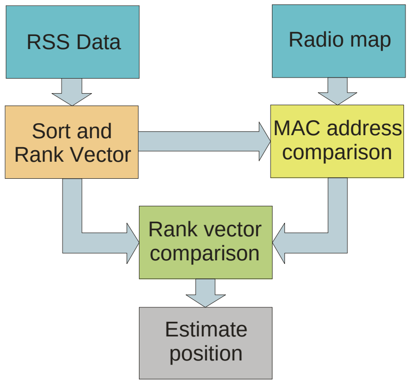

Figure 1. Block diagram of proposed RBF algorithm

These rank vectors are then compared to the online phase vector using one of the similarity measures introduced in the next section. The K reference points with smallest difference are used to calculate the estimated position $\hat{x}$ using the weighted average formula (3).

IV. SIMILARITY MEASURES

In this section similarity measures used to compare ranking vectors in RBF algorithm are introduced. In all cases the online and offline RSS vectors are assumed to have the same length.

A. Spearman distance

Spearman distance [15] is the square of Euclidean distance between two rank vectors:

$$D_S = \sum_{k=1}^{n} (x_k - y_k)^2 , \qquad (4)$$

where x_k is the rank of k-th element in vector X, y_k is the rank of k-th element in vector Y and n is the number of elements in vectors X and Y.

B. Spearman's footrule

Spearman's footrule distance measures total element-wise displacement between two permutations [16]. It is similar to the Manhattan distance for quantitative variables. Spearman's footrule distance can be computed as:

$$D_F = \sum_{k=1}^{n} |x_k - y_k| . \tag{5}$$

C. Jaccard coefficient

The Jaccard coefficient is used to measure the similarity of two sets of data. It is defined as the size (cardinality) of the intersection of the data sets divided by the size of the data sets [17]. It is a special case of the normalized Hamming distance and can be computed using:

$$C_J = \frac{\sum_{k=1}^{n} [x_k \neq y_k]}{n} . \tag{6}$$

D. Hamming distance

Hamming distance is the number of disagreements between two vectors. Hamming distance can also be used for ordinal variables to measure disorder of elements in two vectors [18]. In the RBF algorithm a weighted Hamming distance was used to compute distance between two rank vectors:

$$D_H = \sum_{k=1}^{n} \omega_k \cdot [x_k \neq y_k], \tag{7}$$

where ω_k denotes the weight assigned to the k-th element of the rank vector.

E. Canberra distance

The Canberra distance is the sum of fraction differences between two vectors. Each fraction difference is a value between 0 and 1 [19]. If one of coordinates is zero, the term become unity regardless the other values, thus the distance will not be affected. A weighted version of Canberra distance was used in the RBF algorithm:

$$D_C = \sum_{k=1}^{n} \frac{|x_k - y_k|}{|x_k| + |y_k|} \cdot \omega_k . \tag{8}$$

V. Experimental results

The experiment was carried out in Tietotalo building at Tampere University of Technology. The area was covered with 96 reference points. The average number of heard APs per fingerprint was 29 and altogether 206 APs were detected during data collection.

Measurements in the offline phase of the fingerprinting algorithm were done with Nokia N900 mobile phone and the data collecting software was implemented with Qt Developer. The area where the test was performed, with the positions of reference points, is shown in Fig. 2.

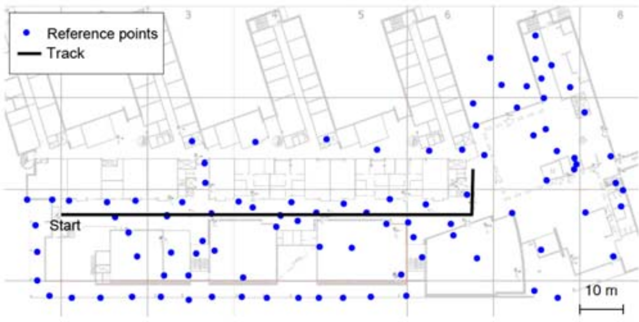

Figure 2. Localization area

Experimental measurements in the online phase were done a month later with the same mobile phone Nokia N900 and then a couple of weeks later with an Asus N63 laptop using WirelessMon software. Measured RSS data were used to estimate the position of the mobile device using fingerprinting algorithms. Measurements were done at 43 points. Track of mobile devices in online phase is shown on Fig. 2 as a black line.

In this scenario differences in localization accuracy are caused by the change of device used in the online phase and also by changes in environment. Results achieved in this scenario using proposed RBF algorithm with different similarity measures are shown in Fig. 3.

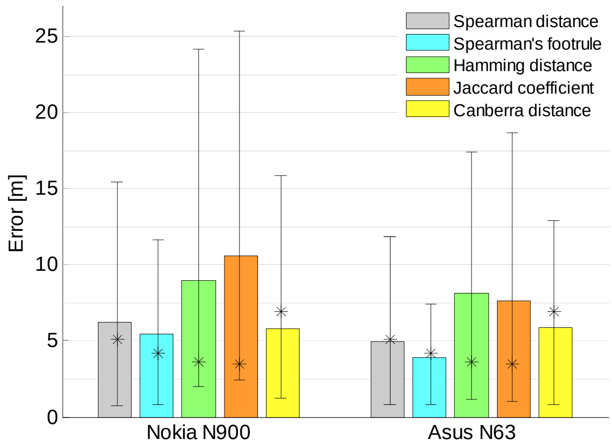

Figure 3. Bars show mean error achieved using RBF and error bars show the 5% and 95% quantiles Asterix show median error.

From results shown in Fig. 3 it can be seen that Spearman's footrule performs best in this real world scenario. When Spearman's footrule was used, median error does not change, and mean error decreased by 1.5 meter when different devices were used in online and offline phases. It is interesting to see that Asus does better than Nokia, even though the Nokia was used to create the radio map. This may be caused by changes of the environment and also by hardware and software equipment of used devices.

When best similarity measure in RBF algorithm was found, performance of this algorithm can be compared to NN and WKNN algorithms. For this comparison the same data were used; in these algorithms the distance between RSS vectors and weights were calculated using Euclidean distance.

From results shown on Fig. 4 it can be seen that mean error of proposed RBF algorithm outperforms commonly used NN and WKNN algorithms. It is clear that position error is less affected by change of the mobile device and environment. From these results RBF algorithm seems to be a great improvement, compared to NN and WKNN algorithms.

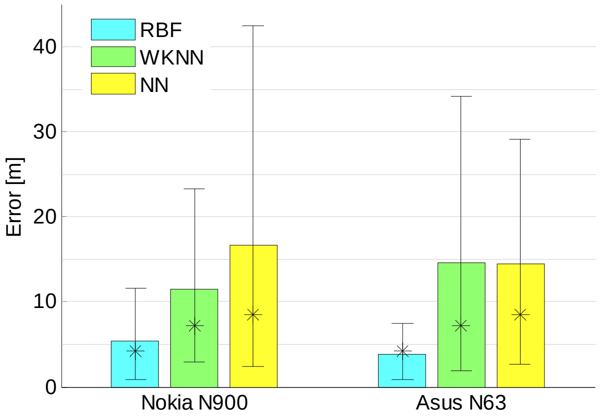

Figure 4. Comparison of RBF with NN and WKNN

Compared to WKNN, the mean error (bars on Fig. 4.) of RBF algorithm is 50% lower when the same device was used and 65% lower in case the devices in online and offline phases were different. Note that RBF performs better than NN and WKNN with every implemented similarity measure.

VI. Conclusion and Future Work

We have described a novel RBF algorithm for indoor localization. The main advantage of this algorithm is that its performance is about the same for any receiver, without the need of calibration of RSS values. Experimental results show that proposed algorithm achieves better accuracy than algorithms NN and WKNN.

The impact of different similarity measures used in RBF algorithm was also investigated. Spearman's footrule seems to perform best among all implemented measures in a real indoor environment. RBF in combination with any of the described similarity measures performs better than NN and WKNN algorithms.

In future more experimental tests will be done, and the impact of AP placement and the number of APs on localization accuracy will be investigated. The impact of density of reference points used in offline phase of algorithm to accuracy of proposed algorithm will be another part of future research. Other similarity measures, such as in Webber et al. [20], will also be studied.

Acknowledgment

This work was partially supported by the Slovak Research and Development Agency under contract No. LPP-0126-09 and by the Slovak VEGA grant agency, Project No. 1/0392/10. We thank Laura Wirola, née Koski, and Toni Fadjukoff for assistance in the data collection.

References

[1] D. Mohapatra, S. B. Suma, "Survey of location based wireless services," *IEEE International Conference on Personal Wireless Communications, ICPWC 2005, 2005*

[2] S. S. Chawathe,. "Low-latency indoor localization using bluetooth beacons," *12th International IEEE Conference on Intelligent Transportation Systems, 2009. ITSC '09.*, pp: 1 – 7, 2009

[3] L. Zheng, W. Dehaene, G. Gielen, "A 3-Tier UWB-based indoor localization scheme for ultra-low-powersensor nodes," *IEEE International Conference on Signal Processing and Communications ICSPC 2007,* pp: 995-998, 2007

[4] L. Koski, T. Perälä, R. Piché, "Indoor positioning using WLAN coverage area estimates," *2010 International Conference on Indoor Positioning and Indoor Navigation IPIN2010*, Sept. 2010

[5] O. Krejcar, "User localization for large artifacts prebuffering and safety possibilities in mobile embedded systems," *International Conference on Advanced Computer Control*, January 2009

[6] V. Honkavirta, T. Perälä, S. Ali-Löytty, R. Piché, "A comparative survey of WLAN location fingerprinting methods," *6th Workshop on Positioning, Navigation and Communication, WPNC 2009,* pp: 243– 251, 2009

[7] B. Li, J. Salter, A. G. Dempster, C. Rizos, "Indoor positioning techniques based on wireless LAN," School of Surveying and Spatial Information Systems, UNSW, Sydney, Australia, Tech. Rep., 2006.

[8] S. Saha, K. Chauhuri, D. Sanghi, P. Bhagwat, "Location determination of a mobile device using IEEE 802.11b access point signals," *Wireless Communications and Networking, WCNC 2003.* vol. 3, pp: 1987 - 1992*,* 2003

[9] T. Vaupel; J. Seitz, F. Kiefer, S. Haimerl, J. Thielecke, "Wi-Fi positioning: System considerations and device calibration," *International Conference on Indoor Positioning and Indoor Navigation, IPIN2010*, Sept. 2010

[10] K. Yedavalli, B. Krishnamachari, S. Ravulat, B. Srinivasan, "Ecolocation: a sequence based technique for RF localization in wireless sensor networks," *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks, IPSN '05*, pp. 285–292, April 2005.

[11] D. Chen; W. Xiao; X. Zhao; "An improved localization algorithm in wireless sensor network," *IEEE International Conference on Robotics and Biomimetics ROBIO 2009*, pp.1253-1258, Dec. 2009

[12] Z. Liu; J. Chen, "A New Sequence-Based Iterative Localization in Wireless Sensor Networks," *International Conference on Information Engineering and Computer Science ICIECS 2009.*, pp.1-4, Dec. 2009

[13] Tsung-Nan Lin; Po-Chiang Lin, "Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks," *International Conference Wireless Networks, Communications and Mobile Computing 2005*, Volume 2, pp. 1569- 1574, 2005

[14] J. Machaj, P. Brida, "Performance Comparison of Similarity Measurements for Database Correlation Localization Method," *3rd Asian conference on intelligent information and database systems ACIIDS 2011*, April 2011 – "in press"

[15] I. Contreras, "Emphasizing the rank positions in a distance-based aggregation procedure," *Decision Support Systems,* Volume 51, Issue 1, pp. 240-245 April 2011

[16] R. Kumar, S. Vassilvitskii, "Generalized distances between rankings, *Proceedings of the 19th international conference on World wide web*," April 2010

[17] J. Bank, B. Cole, "Calculating the Jaccard Similarity Coefficient with Map Reduce for Entity Pairs in Wikipedia,", Wikipedia Similarity Team, December 2008

[18] A. Tarsitano, "Comparing the effectiveness of rank correlation statistics," Working Papers, Università della Calabria, Dipartimento di Economia e Statistica, 2009.

[19] S.-H. Cha, "Comprehensive survey on distance/similarity measures between probability density functions," *International Journal of Mathematical Models and Methods in Applied Science,* vol. 1, no. 4, 2007

[20] W. Webber, A. Moffat, J. Zobel, "A similarity measure for indefinite rankings," *ACM Transactions on Information Systems,* Volume 28, Issue 4, 38 pages, November 2010