

Impact of Radio Fingerprints Processing on Localization Accuracy of Fingerprinting Algorithms

J. Machaj, P. Brida

*Department of Telecommunications and Multimedia, University of Zilina,
Univerzitna 1,010 26, Zilina, Slovakia, phone: +421 41 5132243, e-mail: juraj.machaj@fel.uniza.sk*

crossref <http://dx.doi.org/10.5755/j01.eee.123.7.2391>

Introduction

In past few years lot of attention was attracted to Location Based Services (LBS) [1, 2]. These services are used not only in transport systems, but have also importance in applications in indoor environment. Since the basic requirement for LBS is knowledge of user position, novel localization algorithms, which can be used in indoor environment, have to be developed. These algorithms are mostly based on radio networks [4], e.g. Wi-Fi, GSM, Bluetooth, because satellite navigation systems e.g. GPS can not reliable work in indoor environment due to high signal attenuations.

In this work algorithms based on Wi-Fi networks were used. Main advantage is almost ubiquitous Wi-Fi coverage in indoor environment and implementation of Wi-Fi receivers into almost all devices. So there is no need to develop any additional hardware. In this work fingerprinting algorithms are used to estimate position of mobile device.

Fingerprinting algorithms utilize information about Received Signal Strength (RSS) and do not need to know position of Access Points (APs). Main advantage of fingerprinting algorithms is that they seem to be immune to multipath propagation, which is very strong in indoor environment. On the other hand, drawback of this method seems to be calibration (offline) phase, where time consuming measurements of radio map takes place.

It is clear that RSS fluctuations have impact on localization accuracy. For this reason at least 20 samples of RSS should be measured on each position [5] (in both online and offline phase) and fingerprint is calculated from these samples. Since there are many ways how to process measured RSS data into fingerprint, optimal solution will be found using simulations.

The rest of the paper is organized as follows. In next section fingerprinting localization algorithms used in simulation will be described in detail. In section three different methods of fingerprint processing will be

introduced. Section four describes the simulation model and simulation scenario. Results of the simulations will be shown in section five. Section six concludes the paper.

Fingerprinting algorithms

In this section fingerprinting algorithms used in simulations will be described. Fingerprinting algorithms, in general consists from two phases. First phase is the offline phase (also called calibration phase). In this phase of radio map is created and stored in the database. Second phase is called online phase, in this phase position of mobile device is estimated using one of fingerprinting algorithms.

Radio map

Radio map is built during offline stage. Area where localization services will be offered is divided into small cells in this stage. Each cell is represented by one reference point (RP). In these points RSS values from all transmitters in range – fingerprint is measured for certain period of time and stored in database [3, 4]. Element of radio map has the form

$$P_j = (N_j, \alpha_{ji}, \theta_j), \quad j = 1, 2, \dots, m, \quad (1)$$

where N_j is number of j -th reference point, m is the number of all RPs, α_{ji} is the vector of RSS values and parameter θ_j obtains additional information used in localization phase. Radio map can be modified or preprocessed before the online stage to reduce memory requirements or computational cost of used localization algorithm.

Localization algorithms

Deterministic framework is based on assumption that RSS values on each position represents non-random vector. Estimate of mobile node position \hat{z} can be calculated using

$$\hat{z} = \frac{\sum_{i=1}^m \omega_i \cdot P_i}{\sum_{i=1}^m \omega_i}, \quad (2)$$

where P_i is position of i -th reference point, ω_i represents weight of i -th reference point and m is number of RPs in radio map.

Weights can be calculated as inverted value of Euclidean distance between RSS vectors from online and offline phase. The estimator (2), which keeps the K biggest weights and sets the others to zero, is called the WKNN (Weighted K-Nearest Neighbor) method. 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).

WKNN and KNN methods performs better than the NN method, particularly when values of parameter K are $K = 3$ or $K = 4$ [3]. In the other hand NN algorithm can achieve almost the same results as KNN and WKNN algorithm in case that radio map density is high enough.

Last fingerprinting method used in the simulation was the Rank Based Fingerprinting (RBF) introduced in [4]. This algorithm does not use directly measured RSS data, but uses rank of APs instead. This modification seems to made algorithm more immune to change of device. Same as the NN family algorithms, it uses the estimator (2). Difference is that weights are computed from ranked vectors using Hamming distance

$$d_H = \sum_{k=1}^n \omega_k \cdot [x_k \neq y_k], \quad (3)$$

where $[.] = 1$ if the argument is true, 0 otherwise, d_H represents the distance between ranked vectors, x_k and y_k represents ranked vectors from the online and the offline phase respectively, ω_k denotes the weight assigned to the k -th element of the rank vector and n represents number of the APs heard during the online phase. Weights ω_k are higher for APs with higher ranks, because of fact that the APs with higher RSS are less affected by fluctuations and are more stable.

Processing of measured RSS data

Measured RSS data can be stored in fingerprints in different ways. Most common way is to compute mean value of RSS measured in the certain position, using

$$\overline{RSS} = \frac{1}{N_s} \sum_{i=1}^{N_s} RSS_i, \quad (4)$$

where N_s is number of samples and RSS_i is i -th measured RSS value from the AP.

It seems that this way does not represent the ideal solution since mean RSS value is affected also by samples with low probability, which represents the highest outliers in measured RSS data.

Another common way how to process the fingerprint from measured RSS data is to compute median value of RSS. Median m must to satisfy inequalities

$$P(RSS \leq m) \geq \frac{1}{2}, \quad P(RSS \geq m) \geq \frac{1}{2}, \quad (5)$$

where $P(RSS)$ is probability of measured RSS value and m stands for the median.

Above described ways are most commonly implemented in various positioning systems, but they have serious lacks, since mean and median values are affected by outlier RSS samples. It is assumed that this fact may have impact on localization accuracy. Hence, we decide to propose another solution how to process measured RSS. It consists in selection of RSS samples with the highest probability – mode value. In this case measured RSS data are divided into groups with step of 1 dB, and the group with the highest number of samples – highest probability was chosen.

In the simulations we decided to try to filter out the outlier RSS samples from measured data using histogram functions. Mean values from RSS data without outliers were then computed. Two different ways how to filter the measured RSS data were used. In both ways the outliers were filtered out from the measured RSS data by removing samples with the lowest probability. Samples were divided into groups same way as in case that mode value was calculated. Then the outlier samples were filtered out based on the defined threshold.

In the first case, threshold was set based on the number of groups, in which the measured samples were divided. Groups were sorted based on probability of samples from highest to lowest. Than mean of most probable samples was computed using

$$\overline{RSS} = \frac{\sum_{i=1}^{p \cdot M} P(RSS_i) \cdot RSS_i}{\sum_{i=1}^{p \cdot M} P(RSS_i)}, \quad (6)$$

where M is the number of the groups, p is percentage of used samples, RSS_i is value of RSS in the i -th most probable group, $P(RSS_i)$ is probability of sample from the same group.

The second way how to filter out outliers from set of measured RSS data is to create threshold of probability. In this case threshold is given by percentage of maximum probability. Mean value of RSS than can be computed using

$$\overline{RSS} = \frac{\sum_{i=1}^M P(RSS_i) \cdot RSS_i \cdot [P(RSS_i) \geq p \cdot P_{\max}(RSS)]}{\sum_{i=1}^M P(RSS_i) \cdot [P(RSS_i) \geq p \cdot P_{\max}(RSS)]}, \quad (7)$$

where $P_{\max}(RSS)$ is probability of RSS from the most frequent group and $[.] = 1$ if is true, 0 otherwise.

Simulation model and scenario

Simulations were performed in simulation model developed in the Matlab environment. Multi Wall and Floor (MWF) propagation model is implemented into simulation model in order to calculate RSS values, these values are then affected by random variable, which simulates the real world RSS fluctuations. Simulation model was described in detail in [6].

The simulations were performed at the area of 516 m², observed area was covered by 9 APs and there were 154 RPs placed in the grid at the area. Shape of simulation environment can be seen in Fig. 1.

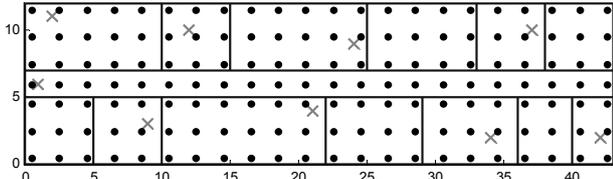


Fig. 1. Simulation environment

In the figure, black dots represent positions of the reference points, grey X shows position of APs and lines represent the walls of a building.

Simulations were performed with 1000 independent trials. Number of used RPs (K) to estimate position was set to 3 and 4 for the KNN and WKNN algorithms respectively [3]. For the RBF algorithm K was set to 2. These numbers result from [5]. Minimum received RSS was set to -100 dBm, when value of RSS calculated using propagation model was lower RSS was set to 'NaN' value, means that AP is out of communication range.

Simulation results

In this section results achieved in the simulations will be presented. Firstly impact of filtering was tested to find best percentage values for both methods. Results achieved for threshold given by number of groups can be seen in Fig. 2.

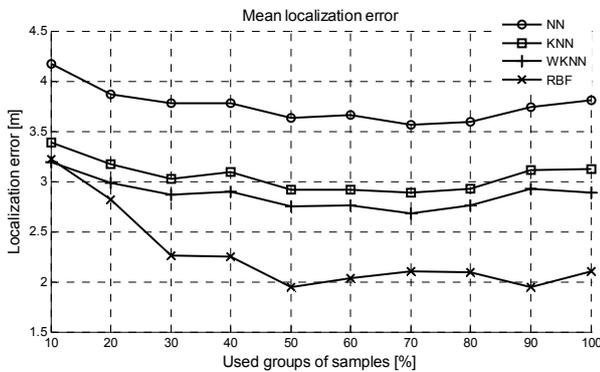


Fig. 2. Results for RSS filtering based on number of groups

From the results it can be seen that for the algorithms

from NN family best results can be achieved for p at the interval from 50% to 80%. In the other hand RBF algorithm achieved best results for $p=50%$ and $p=90%$. From the results it is also clear that RBF algorithm achieved best results in all cases.

Same process was then applied for the second way of RSS data filtering. In that case threshold was given based on probability of samples. Results achieved in these simulations can be seen in the Fig. 3.

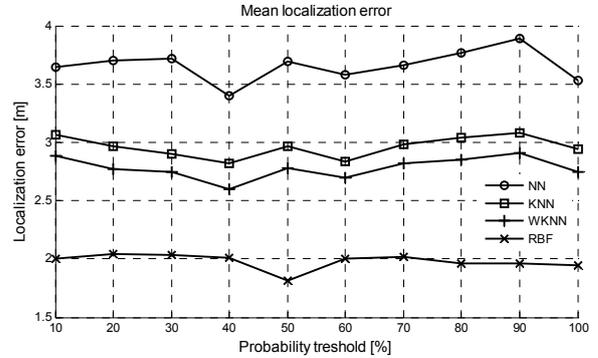


Fig. 3. Results for RSS filtering based probability

From achieved results can be seen that filtering based on probability has also impact on localization accuracy. From achieved results it is clear that algorithms from the NN family achieved best results in case that threshold p was set to 40% of maximum probability. In the other hand RBF algorithm achieved best results for threshold set to 50% of maximum probability.

Results from previous simulations leads us to conclusion that $p=50%$ achieved best results in both filtration procedures for RBF algorithm. For the algorithms from NN family 50% threshold represents optimal solution in case of filtering based on the number of RSS groups. In case that filtering based on probability is used $p=40%$ should be used for the algorithms from NN family.

In the next simulation filtration procedures were compared to results achieved by mean, median and mode values. Results of the simulation can be seen in Table 1.

From the results shown in the table it is clear that the processing of measured RSS samples can improve localization accuracy of RBF algorithm. It can be seen that for filtering based on probability the standard deviation of localization error is 20% lower while mean localization error is the same compared to results achieved using mean value to create fingerprint.

Table 1. Impact of different RSS data processing on localization accuracy

	Parameter	Algorithm	Processing of samples				
			Mean	Median	Mode	50 % of sample groups	50 % of max probability
Localization error [m]	Mean	NN	3.5	4.0	4.8	4.7	3.8
		KNN	2.8	3.3	3.9	3.9	3.1
		WKNN	2.7	3.1	3.7	3.7	2.9
		RBF	2.1	2.3	3.6	2.7	2.1
	Standard deviation	NN	2.3	2.7	3.6	3.3	2.3
		KNN	1.8	1.9	3.1	2.9	1.8
		WKNN	1.6	1.9	3.0	2.7	1.7
		RBF	2.2	2.0	3.1	2.5	1.8

Algorithms from NN family seem to be negatively affected by filtering of RSS values since achieved mean localization error is slightly higher compared to use of mean value in data processing. Standard deviation error was the same for filtering based on probability and mean value of measured RSS samples.

Conclusions

In the paper impact of different measured RSS data processing into the fingerprints on localization accuracy was investigated. From achieved results it can be seen that RSS processing has impact on localization accuracy. Two methods of measured RSS data filtering were proposed to improve localization accuracy.

Optimal solution of filtering was found using simulations in the model created in Matlab environment. From results it can be seen that optimal threshold for both filtering procedures is 50 % in case that RBF algorithm is used to estimate position of mobile device. In case that algorithm from NN family was used optimal value of threshold is in interval from 50 % to 80 % in case that procedure based on group filtering is used and in case that filtering procedure is based on probability optimal value of threshold is 40 %.

Filtering procedures shown better results compared to traditional processing of RSS data, represented by computing mean or median values. It is also important to notice that RBF algorithm achieved the best results for all types of RSS data processing.

Acknowledgements

This work has been partially supported by by the

Slovak Research and Development Agency under contract No. LPP-0126-09 and by the Slovak VEGA grant agency, Project No. VEGA 1/0704/12 and Centre of excellence for systems and services of intelligent transport II., ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF.



"Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ"

References

1. **Cerny M, Penhaker M.** Wireless Body Sensor Network in Health Maintenance Systems // *Electronics and Electrical Engineering*. – Kaunas: Technologija, 2011. – No. 9(115). – P. 113–116.
2. **Mohapatra D., Suma S.B.** Survey of Location Based Wireless Services // *Per*
3. **Honkavirta V., Perälä T., Ali-löyty S., Piche R.** A comparative survey of WLAN location fingerprinting methods // *Workshop on Positioning Navigation and Communication (WPNC'2009)*. – Hannover, Germany, 2009. – P. 243–251
4. **Machaj J., Piche R., Brida P.** Rank Based Fingerprinting Algorithm for Indoor Positioning // *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. – Guimaraes, Portugal, 2011. – P. 1–6.
5. **Xiaoyong Ch. Qiang Y.** Reducing the Calibration Effort for Location Estimation Using Unlabeled Samples // *Pervasive Computing and Communications (PerCom'2005)*, 2005. – P. 95–104.
6. **Machaj J., Duha J.** Simulation Model for Localization in Indoor Environment // *International workshop on Digital Technologies*. – Zilina, Slovakia, 2010. – P. 1–4.

Received 2012 03 09

Accepted after revision 2012 05 11

J. Machaj, P. Brida. Impact of Radio Fingerprints Processing on Localization Accuracy of Fingerprinting Algorithms // Electronics and Electrical Engineering. – Kaunas: Technologija, 2012. – No. 7(123). – P. 129–132.

In the paper impact of different processing of measured RSS data on localization accuracy is investigated using simulations. In simulations four different fingerprinting localization algorithms developed for use with Wi-Fi were used. Used fingerprinting algorithms are described in detail. Model created in Matlab environment used for simulations is also described. Achieved simulation results shown that different processing of measured RSS data has impact on localization accuracy and can improve performance of rank based fingerprinting algorithm. Ill. 3, bibl. 6, tabl. 1 (in English; abstracts in English and Lithuanian).

J. Machaj, P. Brida. Pirštų atspaudų signalų apdorojimo įtaka lokalizavimo tikslumui // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2012. – Nr. 7(123). – P. 129–132.

Naudojant modeliavimą nagrinėjama išmatuotų PSS duomenų apdorojimo įtaka lokalizavimo tikslumui. Sukurti keturi pirštų atspaudų lokalizacijos algoritmai, skirti naudoti per Wi-Fi. Aprašytas modelis sukurtas „Matlab“ aplinkoje. Gauti modeliavimo rezultatai rodo, kad skirtingas išmatuotų PSS duomenų apdorojimas turi įtakos lokalizavimo tikslumui ir gali pagerinti pirštų atspaudų atpažinimo algoritmo našumą. Il. 3, bibl. 6, lent. 1 (anglų kalba; santraukos anglų ir lietuvių k.).