

An experimental study of indoor RSS-based RF fingerprinting localization using GSM and Wi-Fi signals

Gökhan ÇELİK^{1,*}, Hasari ÇELEBİ¹, Hakan Paşa PARTAL², Engin ZEYDAN³,

İlyas Alper KARATEPE³, Ahmet Salih ER³

¹Department of Computer Engineering, Gebze Technical University, Kocaeli, Turkey

²Department of Electronics and Communications Engineering, Yıldız Technical University, İstanbul, Turkey

³Turk Telekom Labs, İstanbul, Turkey

Received: 08.01.2016

Accepted/Published Online: 20.10.2016

Final Version: 30.07.2017

Abstract: Localization of mobile users in indoor environments has many practical applications in daily life. In this paper, we study the performance of the received signal strength (RSS)-based radio frequency (RF) fingerprinting localization method in a shopping mall environment considering both calibration and practical measurement cases. In the calibration case, the test data for the RSS fingerprinting database are built offline by receiving signals from Global System for Mobile Communications (GSM) base stations, which are collected by a dedicated measurement tool, i.e. the Test Mobile System. In order to see the localization performance, the k-nearest neighbors (K-NN) and random decision forest (RDF) algorithms are implemented. The RDF algorithm provides a better localization performance than K-NN in this case. For the practical implementations, the RSS values of both GSM and Wi-Fi signals are collected by ordinary smartphones. Localization is performed using different classification algorithms, i.e. BayesNet, support vector machines, K-NN, RDF, and J48. Moreover, the effects of the received signal type, phone type, and number of reference points on localization performance are investigated.

Key words: Indoor localization, radio frequency fingerprinting, received signal strength, Wi-Fi, Global System for Mobile Communications

1. Introduction

The localization of wireless nodes by employing radio signals is now very attractive due to the increased demand for location-based services [1,2]. In the localization context, the Global Positioning System (GPS) is commonly used for the localization of the wireless nodes in outdoor environments [3]. A wireless node can estimate its own position by employing the GPS signal from at least four satellites. However, a GPS signal strength is too weak and faces great attenuation when the wireless node is in indoor environments. Therefore, GPS-based localization in indoor environments produces inaccurate location information. Instead of GPS signals, Wi-Fi or Global System for Mobile Communications (GSM) signals are used to estimate the location of a node in indoor environments [4,5]. The location of a node can be estimated by different methods that use the different parameters of the received signal, such as time of arrival, angle of arrival, time difference of arrival, and received signal strength (RSS) [4]. In this paper, we consider the RSS-based localization of nodes using the RF fingerprinting method.

*Correspondence: gcelik@gtu.edu.tr

Several theoretical and experimental studies have shown the challenges of RSS-based localization techniques in the literature [6–14]. Cost-effective solutions utilize readily available commercial RF signals, such as cellular GSM and Wi-Fi, and employ software algorithms without any additional hardware requirements [5]. Cellular base stations (BSs) or Wi-Fi access points (APs) already transmit RF signals and the localization of the mobile users can be performed by their mobile devices, which are collecting RSS values [11–16]. In this line, the RF multilateration method requires the modeling of transmitter and receiver antenna performances, which is an easy task for antenna designers. However, due to nondeterministic effects, such as fading, multipath propagation, and shadowing, the modeling requires an additional offline training phase for different scenarios. Note that the RF multilateration method provides an acceptable positioning accuracy with the following conditions [9,12,14]: 1) smaller test area, 2) large number of reference nodes (APs and BSs), 3) known path loss exponent for each reference node, and 4) line-of-sight environment. The RF multilateration method is not preferred for the localization of mobile devices in large indoor environments, such as shopping malls and airports, since they cannot satisfy these requirements. In the literature, RF fingerprinting method has better accuracy and a more reliable performance in indoor environments [9]. Therefore, in this paper, the RF fingerprinting method is considered and experimentally analyzed.

Previous studies in the literature considered RF fingerprinting methods and tested them in small areas with no environmental effects and obstacles, such as laboratory environments [13]. However, in this study, the localization performance of the RF fingerprinting method in a shopping mall is tested in two different measurement cases. For calibration, we used Test Mobile System (TEMS) investigation 13.0.3 model from Ascom (Baar, Switzerland), which consists of both hardware and software commonly used by telecom operators for wireless network test driving, benchmarking, monitoring, and analysis. Since the TEMS device can only support cellular type signals (i.e. GSM), only this type of signal is used for localization purposes in the calibration case. Since special purpose measurement equipment is used in the TEMS, the errors experienced during the measurement are minimized. As a result, the TEMS is used in the calibration case in order to collect accurate RSS values from multiple BSs simultaneously. Additionally, the test and training data are collected offline and the localization algorithms are performed on these data in this case. Machine learning algorithms (k-nearest neighbors (K-NN) and random decision forest (RDF)) are employed for localization. Various types of phone are used. In this context, the localization accuracy of a mobile user depends on the receiver sensitivity of the phone used. Therefore, the effects of the different phones on the localization performance are also investigated. Note that the aim of this study is to apply the method for practical scenarios and discuss the improvement approaches for both measurements and algorithms [17,18]. Therefore, a practical measurement case with an ordinary smartphone is considered in order to see the localization performance in a real measurement scenario. The RSS values of GSM and Wi-Fi signals are used to obtain the fingerprints of the points, respectively. In order to realize the practical case, an Android-based LG Nexus 4 is used to record the RSS of the received signals from the reference points. Localization performance in this case is evaluated using different machine learning algorithms [17–20]. The general system model for both measurement cases is shown in Figure 1.

The rest of this paper is organized as follows. In Section 2, RF fingerprinting-based localization in a shopping mall is presented. In Sections 2.1 and 2.2, the calibration and the practical measurement cases of the RF fingerprinting-based localization method are covered, respectively. The concluding remarks are given in Section 3.

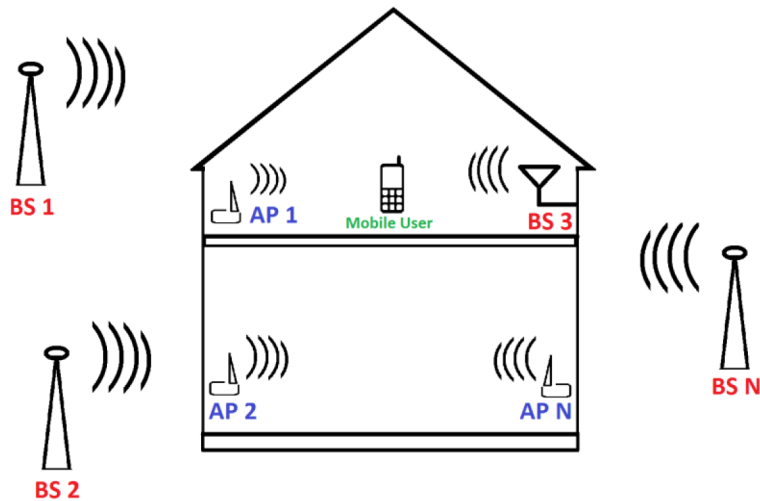


Figure 1. General system model for localization.

2. RF fingerprinting-based localization

2.1. Calibration measurement case

In this case, the RF fingerprinting localization is performed in the 4-story Palladium Shopping Mall in Ataşehir, İstanbul, Turkey. The RSS measurements are not performed with an ordinary smart phone or another device. Instead, the TEMS device and its software are used during the measurement in order to reduce measurement errors. The localization algorithms are tested with the RSS values, which are collected before as training data, not with instantaneous RSS values. Therefore, this case is the calibration case.

The RSS values of the GSM signals are measured for each floor of the shopping mall and the RSS measurements are recorded using a laptop computer. In addition, the RSS values are collected with a Sony Ericsson smartphone. The smartphone is connected to the laptop through a USB port. The measurement values are recorded to a database.

2.1.1. Data collection

In the shopping mall, 254 points are determined on each floor of the Palladium shopping mall. The entrance stairs of each floor is divided into grids. The distance between the points are selected as 1.7 m and RSS values are recorded from 7 GSM BSs, one located inside the mall and the others outside the mall. Moreover, 1000 RSS values are recorded for each BS by standing for 1.5 min at each point. A total of 7000 RSS values are obtained. Considering the polarization effects of the receiver antenna, the phone is fixed at a reference point for 45 s. After that, the phone is rotated 90° with respect to the reference point for the remaining 45 s.

Note that the TEMS module is used in active mode, because data collection is much faster in active mode compared to the idle mode on this device. In the idle mode, only a handful of data per minute can be collected. Meanwhile, in the active mode, the TEMS module can easily collect about 1000 RSSI data samples per minute. The amount of data is vitally important for the localization algorithm to work accurately. Otherwise, momentary RSSI changes cannot be fetched and the offline training cannot be completed. Since insufficient RSSI data can degrade the accuracy of the localization algorithm, the idle mode was not used and the active mode was preferred for the TEMS module.

2.1.2. Localization algorithms

The RF fingerprinting method consists of two steps, which are online and offline stages. In the offline stage, the measurement area (the Palladium Shopping Mall in this case) is divided into 1.7×1.7 m grids and each point is labeled manually with a unique 2D coordinate number (x, y) . At each grid point, the RSS values from the GSM BSs are recorded for 1.5 min and a database is generated, which contains the RSS fingerprints (signatures) of the points in the grids [21]. Then, in the online stage, the fingerprint of the mobile user is generated in real time using the RSS values from GSM BSs. The generated fingerprint is compared to the fingerprint values in the database formed during the offline step [21]. The instantaneous location of the mobile user is then determined using the machine learning algorithms.

In this study, the K-NN and RDF classification algorithms are used to determine the location of a node. K-NN is a supervised learning method that decides by comparing its k neighbors [17]. On the other hand, RDF is a classification method consisting of many decision trees [18]. RDF finds the results as a tree, which is the most repeated one on the decision tree.

For recording the RSS values of the different brands of phones, the ground floor of the shopping mall is reserved. This area is divided into 32 uniformly separated points, 1.7 m apart. The RSS values are recorded with a Nokia phone from 3 different GSM BSs: one in the mall (indoor BS) and the others outside the mall (outdoor BSs). Hence, 500 RSS values are recorded for each point. As a result, 500 different fingerprint values are obtained from 3 different GSM BSs for each point. Similarly, a Sony Ericsson phone is used to collect RSS values from the same GSM BSs at the same grid points. Consequently, 500 different fingerprint values are obtained for each point for this phone.

2.1.3. The effects of localization algorithms

In this section, the localization performance results are provided for the K-NN and RDF classification algorithms in the shopping mall. Note that the 80% of the measured RSS values are used for training data and the remaining 20% of the RSS values are used for test data. Besides the effects of both classification algorithms on the localization error, the effects of the number of BSs are also investigated. The localization performance of the K-NN algorithm in terms of cumulative distribution function (CDF) of the localization error is shown in Figure 2. Notice that k is selected as 3 for this case. According to the results, a localization accuracy under 2 m is obtained with the following results: 34.62% for 3 BSs, 82.16% for 4 BSs, 95.55% for 5 BSs, 98.50% for 6 BSs, and 99.00% for 7 BSs.

Similarly, RDF performance is obtained and shown in Figure 3. According to the results, a localization accuracy under 2 m is obtained with the following results: 35.77% for 3 BSs, 79.59% for 4 BSs, 95.89% for 5 BSs, 99.11% for 6 BSs, and 99.78% for 7 BSs. If we compare the results for K-NN and RDF, one can see that the RDF algorithm provides better localization accuracy than K-NN for all the cases except the case with 4 BSs, which is an outlier. This is because the RDF method is relatively insensitive to outliers compared to K-NN. The results show that the number of BSs can be optimized in order to achieve the target localization accuracy.

2.1.4. The effects of phone types

In order to investigate the effects of the different phone brands on localization accuracy, the measurements are obtained with both a Sony Ericsson and a Nokia under the same conditions. The K-NN algorithm, with $k = 3$, is considered for obtaining results. During this study, 80% of the data is used for training data and the remaining

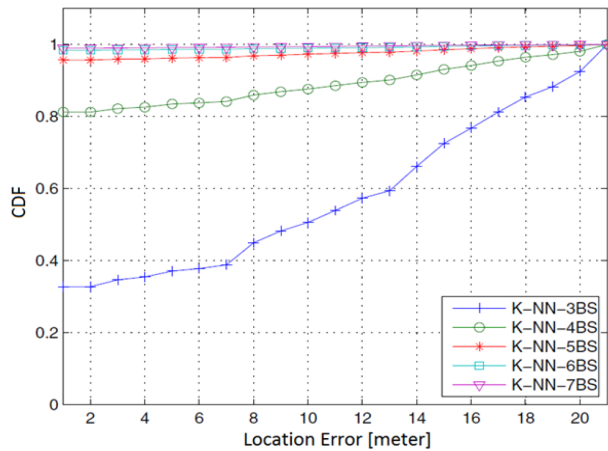


Figure 2. Localization performance of K-NN algorithm with different numbers of GSM BSs.

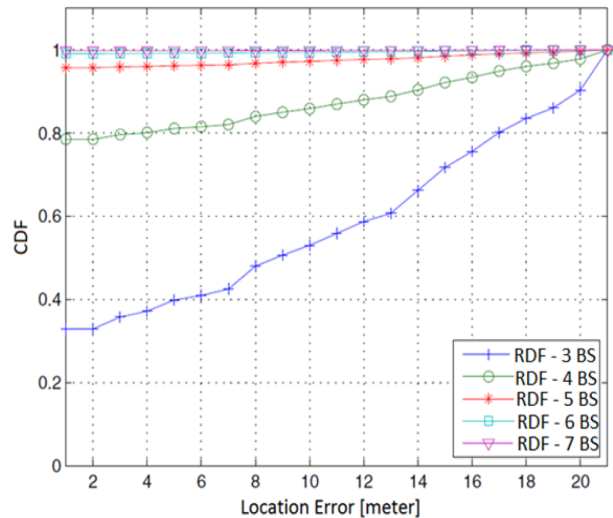


Figure 3. Localization performance of RDF algorithm with different numbers of GSM BSs.

20% is used for test data. Under these conditions, the CDF of the location estimation error with Sony Ericsson and Nokia phones for the localization error is shown in Figure 4. The Sony Ericsson provides better localization accuracy than the Nokia phone. For example, the Sony Ericsson phone gives a 4-m localization error with 90% probability, whereas the Nokia phone gives a 9-m localization error with the same probability. The main reason for this difference is that the Sony Ericsson phone has better receiver sensitivity than the Nokia phone. Therefore, the training data for the Nokia phone were not as good as the Sony Ericsson phone data due to the receiver sensitivity levels. It should also be noted that the receiver sensitivities of different brand phones vary, and then the RSS data reception capabilities also vary, which may cause different training schemes.

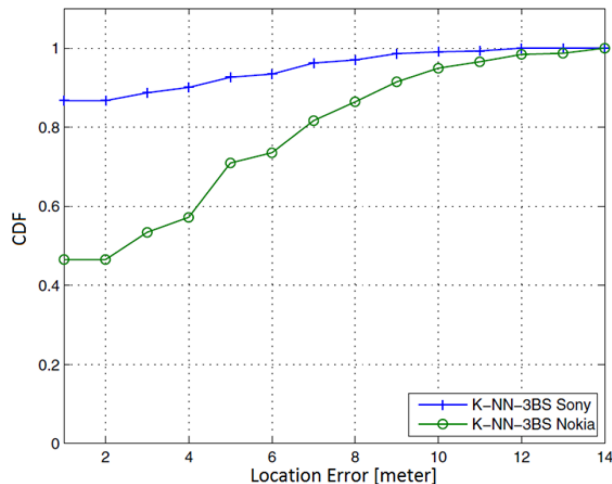


Figure 4. The effects of phone types on localization accuracy.

The calibration case of indoor RF fingerprinting-based localization is performed in the Palladium Shopping Mall of İstanbul. The location of the mobile user is determined using the GSM signals obtained from both indoor and outdoor GSM BSs. The localization performance evaluations of the K-NN and RDF algorithms are represented in terms of CDF. According to the results, the RDF algorithm gives better localization accuracy

than the K-NN algorithm. In addition, the effects of the smartphones on localization accuracy are investigated. According to the results, the Sony Ericsson phone gives more accurate 5-m results than the Nokia phone with 90% probabilities.

2.2. Practical measurement case

In this case, the practical method of RF fingerprinting-based localization is performed in the Gebze Center Shopping Mall in Kocaeli, Turkey. The practical implication of the RF fingerprinting method is selected in this case. The main difference of the practical case from the calibration case is as follows: the RSS values are collected and recorded with an ordinary smartphone, i.e. an LG Nexus 4, with no special software or hardware used. The database is generated based on the actual recorded RSS values. In addition, not only the GSM signals but also Wi-Fi signals are used to localize a mobile user in the shopping mall. Furthermore, the test values are instantaneous values selected from a mobile user, not from a database.

An Android mobile application is developed in order to collect RSS values of different GSM BSs and Wi-Fi APs automatically and this application is run on an LG Nexus 4 smartphone. Using these GSM and Wi-Fi RSS values, a database is generated on the phone. After this step, the reference point and the RSS value received at the reference point are sent to the server via short message service (SMS). The reference and the RSS values at the server are read by the server computer from a subscriber identity module (SIM) card using the SMSLIB Java library. The SIM card of the server is in a USB 3G modem.

The location of the mobile user is determined by reading the RSS values from the SIM card and comparing these values with the RSS values stored in the predefined database. Using the Abstract Window Toolkit (AWT) in Java, the location of the mobile users, their routes, and the tracks are shown on the server screen. The route that mobile users generated is shown on the map of the shopping mall using the AWT. The steps of the localization system are shown in Algorithm 1.

Algorithm 1 Summary of the localization algorithm.

Collect the RSS values from GSM and Wi-Fi reference points

Send the RSS data to the server machine via SMS

Read the SMS from USB modems on the server machine

Send the data to the localization application

Estimate the location of the mobile user by employing different machine learning algorithms

Demonstrate the estimated location of the mobile user on the map of the shopping mall

2.2.1. Data collection

An LG Nexus 4 is used to collect the RSS values of GSM BSs and Wi-Fi APs. The RSS values are collected from predetermined points for a specific time. While collecting RSS values, the coordinates of the point are also attached to the stored data. The coordinate and the RSS values of GSM BSs and Wi-Fi APs are stored in two different files on the LG Nexus 4. The recorded data are then transferred to the PostgreSQL database. The RSS values are collected from two floors of the mall. For each floor, 100 points are selected, 2 m apart. The data are collected over four different days and the measurement is performed twice a day. The measurement time is 4 min for each point.

The received signal power of the strongest GSM BS over 24 h is shown in Figure 5.

The average signal power is -62.35 dBm between 16:00 and 17:00 hours, whereas this value is -71.88 dBm between 17:00 and 18:00 hours [14]. There are remarkable average power differences within this 1-h time

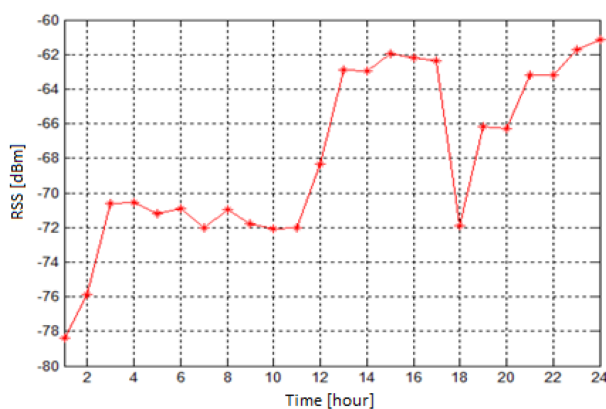


Figure 5. The RSS values of the most powerful GSM BS near the shopping mall over 24 h on Friday, 18 October 2013.

period.

The minimization of the location estimation errors is critical for designing high-precision localization systems. In order to reduce the localization error of the RF fingerprinting method and obtain more accurate results, data collection is performed twice for each point. In addition, the data are collected on different days in order to obtain more realistic results. It is worth noting that the test data are collected for 30 s independently from the training data.

2.2.2. The effects of localization algorithms

In this step, localization is performed using different machine learning algorithms from the Weka Java library on the Android OS. The localization performance of the algorithms is evaluated using the Weka library. Different algorithms are tested by using 30 s of test data and 3.5 min of training data from 17 different Wi-Fi APs for each point. The performance of the algorithms is obtained in terms of CDF, as shown in Figure 6. The probability of localization error of 4 m and below of the algorithms is 49.63% for J48, 55.35% for K-NN, 60.16% for RDF, 76.79% for SVM, and 77.27% for BayesNet.

J48 is a decision tree-based classification algorithm similar to the RDF algorithm [22]. The SVM algorithm classifies objects by maximizing the orthogonal distance of the closest points of the classes. On the other hand, BayesNet is a probabilistic classification algorithm that considers the conditional probabilities of two random events and their relations. Since the BayesNet algorithm considers the relations between the RSS values received from different BSs, it performs better than the other algorithms. According to the results, BayesNet and SVM are the most suitable algorithms for the practical case. If we compare the performance of K-NN (or RDF) under calibration and practical cases, it can be easily seen that K-NN (or RDF) is not suitable for practical cases.

2.2.3. The effects of number of reference points

In order to see the effects of the number of reference points on localization accuracy, the BayesNet algorithm is used with 10, 12, 14, and 17 Wi-Fi reference APs and 30 s of test data and 3.5 min of training data. The results are shown in Figure 7. Localization accuracy is improved when the number of reference points is increased. At least 12 reference points are needed to achieve error of 4 m and below with 70% probability, and 17 reference points are needed to achieve this error limit with 80% probability.

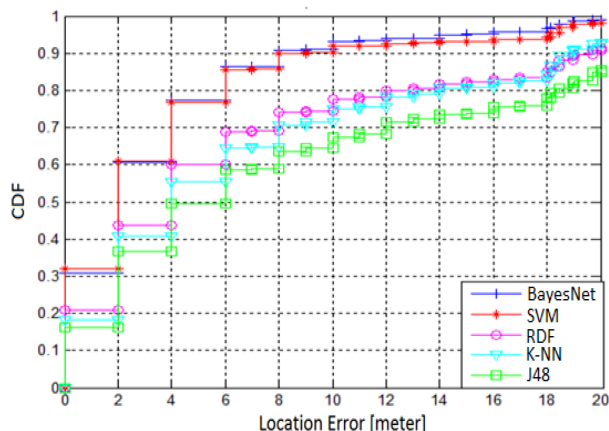


Figure 6. The effects of different algorithms on localization accuracy in the practical case.

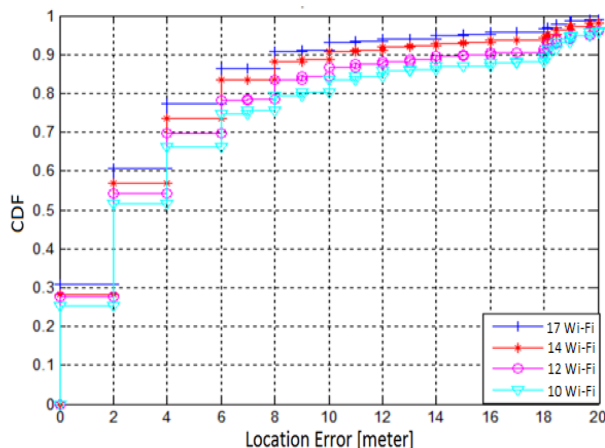


Figure 7. The effects of number of reference nodes on localization accuracy.

2.2.4. The effects of reference signal types

In the practical case, only Wi-Fi reference signals are used for locating the mobile device. This is because the CRLB expression for the RSS-based localization was derived in [4] and the localization error fundamentally depends on the standard deviation of shadowing, the distance between the transmitter and receiver, and the path loss exponent [4]. According to these fundamental limits, the estimation error of localization is proportional to the standard deviation of shadowing and the distance between the BS (or AP) and the mobile device, and it is inversely proportional to the path loss exponent, which is an environment-dependent parameter. We can infer that if the distance between the BS (or AP) and the mobile device is short, the estimation error of the localization system can be small. Therefore, since the typical value of this distance parameter for the Wi-Fi system (typical range) is less than for the GSM system (typical cell radius), we prefer to use a Wi-Fi system for the practical case in order to minimize localization errors.

It is also possible that we can use GSM signals alone, or both GSM and Wi-Fi signals at the same time. For each point, the RSS values from 17 Wi-Fi, 17 GSM, and 17 GSM–Wi-Fi (15 Wi-Fi + 2 GSM) BSs and APs are recorded with 30 s of test and 3.5 min of training data. The BayesNet algorithm is employed for these three scenarios in order to see the effects of reference signal type on the localization error. The results of these measurements are illustrated in Figure 8.

It is clear that using Wi-Fi signals instead of GSM signals provides better localization accuracy. The main reason for this result is that the distance between the AP and mobile device in the Wi-Fi case is shorter than that in the GSM case. However, using GSM signals alongside Wi-Fi signals improves localization accuracy, because the fingerprint of each grid is diversified by using additional different features, i.e. GSM and Wi-Fi signals.

3. Concluding remarks

In this paper, the calibration and practical cases of an RSS-based RF fingerprinting localization method are investigated. In the calibration case, indoor RF fingerprinting-based localization is performed in the Palladium Shopping Mall of İstanbul, Turkey. The location of the mobile user is determined using GSM signals. According to the results, the RDF algorithm provides better localization accuracy than K-NN in the same scenario for the calibration case. In the practical case, the BayesNet and SVM algorithms provide better localization accuracy

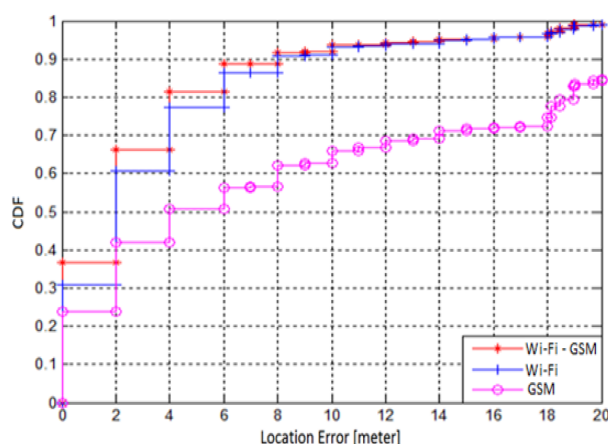


Figure 8. The effects of different reference signals on localization accuracy.

than the other algorithms employed in this study. The Wi-Fi reference signals yield better estimation accuracy than the GSM signals due to the distance between the APs and mobile devices in Wi-Fi networks being shorter than in GSM networks. In addition, increasing the number of reference nodes used in the localization can further decrease the localization error. According to the results, in order to improve the performance of RF fingerprinting methods, the number of transmitter stations should be increased, the amount of offline (training) data must be increased, offline training data should be collected on different days and at different times, and an effective machine learning algorithm must be used.

Acknowledgment

We appreciate the help and comments from Hasan Buyruk, Ahmet Bacak, Ömer İleri, Salih Ergüt, A Kenan Keskin, Yakup Genç, Furkan Tektaş, and Şeyma Şendil during the measurement campaigns.

References

- [1] Schiller J, Voisard A. Location-Based Services. 1st ed. San Francisco, CA, USA: Morgan Kaufmann Publishers, 2014.
- [2] Patwari N, Ash J, Kyperountas S, Hero AO, Moses RM, Correal NS. Locating the nodes: cooperative localization in wireless sensor networks. *IEEE Signal Proc Mag* 2005; 22: 54-69.
- [3] Pratap M, Enge P. Global Positioning System: Signals, Measurements and Performance. 2nd ed. Lincoln, MA, USA: Ganga-Jamuna Press, 2006.
- [4] Gezici S. A survey on wireless position estimation. *Wireless Pers Commun* 2008; 44: 263-282.
- [5] Liu H, Darabi H, Banerjee P, Liu J. Survey of wireless indoor positioning techniques and systems. *IEEE T Syst Man Cy C* 2007; 37: 1067-1080.
- [6] Pahlavan K, Li X, Makela JP. Indoor geolocation science and technology. *IEEE Commun Mag* 2002; 40: 112-118.
- [7] Nerguizian C, Despins C, Affes S. Geolocation in mines with an impulse response fingerprinting technique and neural networks. *IEEE T Wirel Commun* 2006; 5: 603-611.
- [8] Lin TN, Lin PC. Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks. In: *IEEE 2005 International Conference on Wireless Networks, Communications, and Mobile Computing*; 13–16 June 2005; Maui, HI, USA. New York, NY, USA: IEEE. pp. 1569-1574.

- [9] Al Shayokh M, Alkasi U, Partal HP. Performance improvement techniques for RSS-based localization methods. In: IEEE 2013 International Conference on Electrical Information and Communication Technology; 13–15 February 2014; Khulna, Bangladesh. New York, NY, USA: IEEE. pp. 1-6.
- [10] Alkasi U, Al Shayokh M, Partal HP. An experimental comparison study on indoor localization: RF fingerprinting and multilateration methods. In: IEEE 2013 10th International Conference on Electronics Computer and Computation; 7–9 November 2013; Ankara, Turkey. New York, NY, USA: IEEE. pp. 255-259.
- [11] Alyafawi I, Braun T, Dimitrova DC. Robust indoor localization of narrowband signals. In: IEEE 2015 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications: 30 August–2 September 2015; Hong Kong. New York, NY, USA: IEEE. pp. 2193-2198.
- [12] Buyruk H, Keskin AK, Şendil Ş, Çelebi H, Partal HP, İleri Ö, Zeydan E, Ergüt S. RF fingerprinting-based GSM indoor localization. In: IEEE 2013 21st Signal Processing and Communications Applications Conference; 24–26 April 2013; Nicosia, North Cyprus. New York, NY, USA: IEEE. pp. 1-4.
- [13] Yang J, Chen Y. A theoretical analysis of wireless localization using RF-based fingerprint matching. In: IEEE 2008 International Symposium on Parallel and Distributed Processing; 14–18 April 2008; Miami, FL, USA. New York, NY, USA: IEEE. pp. 1-6.
- [14] Bacak A, Çelebi H. Practical considerations for RSS RF fingerprinting-based indoor localization systems. In: IEEE 2014 22nd Signal Processing and Communications Applications Conference; 23–25 April 2014; Trabzon, Turkey. New York, NY, USA: IEEE. pp. 497-500.
- [15] Birkel U, Weber M. Indoor localization with UMTS compared to WLAN. In: IEEE International Conference on Indoor Positioning and Indoor Navigation; 13–15 November 2012; Sydney, Australia. New York, NY, USA: IEEE. pp. 1-6.
- [16] Otsason V. Accurate indoor localization using wide GSM fingerprinting. MSc, Tartu University, Tartu, Estonia, 2005.
- [17] Kontkanen P, Myllymaki P, Roos T, Tirri H, Valtonen K, Wettig H. Topics in probabilistic location estimation in wireless networks. In: IEEE 2004 15th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, 5–8 September 2004. New York, NY, USA: IEEE. pp. 1052-1056.
- [18] Ho TK. Random decision forests. In: 1995 3rd International Conference on Document Analysis and Recognition; 14–16 August 1995. New York, NY, USA: IEEE. pp. 278-282.
- [19] Brunato M, Battiti R. Statistical learning theory for location fingerprinting in wireless LANs. *Comput Netw* 2005; 47: 825-845.
- [20] Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. 1st ed. New York, NY, USA: Cambridge University Press, 2000.
- [21] Mondal R, Turkka J, Ristaniemi T. An efficient grid-based RF fingerprint positioning algorithm for mobile user location estimation in heterogeneous small cell networks. In: 2014 International Conference on Localization and GNSS; 24–26 June 2014; Helsinki, Finland. New York, NY, USA: IEEE. pp. 1-5.
- [22] Patil TR, Sherekar MSS. Performance analysis of naive Bayes and J48 classification algorithm for data classification. *International Journal of Computer Science and Applications* 2013; 6: 256-261.