Increasing Interference Robustness of WiFi Fingerprinting by Leveraging Spectrum Information

Filip Lemic, Arash Behboodi, Vlado Handziski, Adam Wolisz
Telecommunication Networks Group (TKN)
Technische Universität Berlin (TUB), Germany
{lemic, behboodi, handziski, wolisz}@tkn.tu-berlin.de

***Abstract*—RF-based indoor localization is constantly gaining popularity, and WiFi-based fingerprinting algorithms belong to the most promising candidates due to their well-known advantages. While it is known that RF interference can adversely influence the accuracy of those algorithms, it is still unclear if this effect could be efficiently mitigated. To this end, we demonstrate the impact and propose a procedure for reducing the influence of RF interference on WiFi beacon packets RSSI-based fingerprinting algorithms. The proposed procedure adjusts the RSSI measurements based on estimates of their variability caused by RF interference. For estimating the variability in RSSI measurements, the procedure leverages information about the spectrum power levels in the frequency band on which the fingerprinting algorithm performs. The proposed procedure can be inserted in the usual workflow of fingerprinting algorithms. We experimentally compared the performance of two well-known WiFi-based fingerprinting algorithms without and with the proposed interference mitigation procedure. Our experimental evaluation in different interference scenarios shows that the proposed procedure for mitigating the influence of RF interference significantly improves the localization accuracy, while it does not notably increase the latency of evaluated fingerprinting algorithms.**

I. Introduction

In recent years, indoor localization is gaining high popularity as an enabler of different location-based services and applications [1]. The usage of Radio Frequency (RF) signals is a promising candidate for an accurate indoor localization, as witnessed by the amount of work on this topic, with just some examples being [2]–[5]. Among different approaches in RF-based indoor localization, Wireless Fidelity (WiFi)-based fingerprinting using beacon packets Received Signal Strength Indicator (RSSI) measurements is one of the most promising candidates, due to its various well-known advantages in details described in [6].

The number of wirelessly connected devices is constantly increasing, as well as the wireless traffic generated by each device [7]. Naturally, with the increase in the number of connected devices and in the overall traffic among those devices, the probability of wireless interference is also increasing. In other words, it is well known that RF interference can degrade the performance of wireless systems, which has been repeatedly reported in the literature [8], [9]. More specifically, some recent results indicate that RF interference can significantly degrade the performance of RF-based indoor localization solutions. The authors in [10] experimentally show that background IEEE 802.11 traffic as interference source can increase the 80^{th} percentile localization error of an IEEE 802.15.4-based indoor localization solution in the worst case scenario. In [11], the authors give an analytical model and show by simulations the effect of several types of interference on the performance of an Radio-Frequency Identification (RFID)-based indoor localization solution. In our previous work [12], we have shown that various RF interference patterns generally degrade the performance of a large number of RF-based indoor localization solutions in the 2.4 GHz Industrial, Scientific and Medical (ISM) frequency band. This effect is mostly visible through the increased localization errors in the evaluation scenarios with RF interference, in comparison to the evaluation scenarios with minimized interference.

This work builds on our recent findings [13], where we have theoretically characterized and experimentally verified the effect that RF interference has on the RSSI measurements obtained from WiFi beacon packets. This effect, depending on the interference power, ranges from no practical influence, to a linear correlation with RSSI variability, and finally to a full packet-loss. To this end, we propose a procedure for reducing the effect of RF interference on the WiFi RSSI-based fingerprinting algorithms. The proposed procedure, to which we refer as *interference mitigation procedure*, can be applied after the RSSI measurements collection procedure of fingerprinting algorithms in order to adjust the RSSI measurements before their processing for location estimation. Our work is based on the assumption that a fingerprinting system leverages a certain level of environmental awareness, i.e. spectrum power levels in the 2.4 GHz ISM frequency band, in order to estimate the average interference power levels at different frequencies. In this work, we also propose two methods for estimation of the average interference power using spectrum power levels. By using the estimated average interference power levels it is possible to calculate the additive variability of the RSSI measurements due to RF interference. Additive variability of RSSI measurements is here defined simply as the change in RSSIs due to RF interference and this term should not be confused with statistical variance-like terms. Once this variability is estimated, the proposed procedure for reducing the influence of interference can leverage this information to adjust RSSI measurements in both training and runtime phases of fingerprinting, which results in improved performance of fingerprinting algorithms based on RSSI measurements from WiFi beacon packets.

We compared the performance of two such fingerprinting algorithms with the performance of the same algorithms with inserted interference mitigation procedure. The evaluation was carried in four interference scenarios where different types of controlled RF interference were generated. The obtained results firstly show that RF interference can severely degrade the performance of WiFi-based fingerprinting algorithms. Secondly, the results show that the proposed procedure for reducing the influence of interference can significantly improve the interference robustness of evaluated algorithms, without practically affecting their processing time.

The rest of the paper is structured as follows. In Section II, the most important theoretical findings from our previous work are highlighted and extended in order to obtain the estimate of the additive variability of the RSSI measurements due to interference. In Section III, we describe a general workflow of RSSI-based WiFi fingerprinting algorithms, while in Section IV we show how this workflow can be extended with a procedure for reducing the effect of RF interference on fingerprinting algorithms. In the same section, we describe in details the proposed procedure, as well as two methods for estimation of average interference power. Section V describes the methodology, scenarios, experimental setup, and the two fingerprinting algorithms used in the evaluation. In Section VI, we present and discuss the results of the performance evaluation. Finally, Section VII concludes the work and gives directions for future improvements.

II. Interference Effect on RSSI Measurements

In this section, a theoretical framework is adopted to study the effect of interference on Received Signal Strength (RSS) values. We assume an Additive White Gaussian Noise (AWGN) channel, a signal source with transmission power P'_X, a destination with noise variance N, and an interferer. The interferer is a continuously transmitting source with fixed transmission power P'_I. The transmission rate of the source is r and it is also assumed that RSS values can only be obtained if the destination can correctly decode the source messages. This assumption is used to model the packet-based nature of RSS values, where RSS values, reported by some technologies, can only be obtained if the packet is correctly decoded. If all powers are measured in mW, it is shown in [13] that the effect of interference on the RSS value is additive, until the moment that the destination Signal to Interference plus Noise Ratio (SINR) is below the necessary level for decoding source messages. To observe the change of δ dB in RSS value at the receiver, the interference power should increase above certain threshold which is dependent on δ. As shown in [13], the change of RSS value in dB due to RF interference can be described as follows:

$$\delta = 10 \log \left(\frac{P_I + P_X + N}{P_X + N} \right), \tag{1}$$

where P_X and P_I are signal and interference power *at the receiver*, respectively. However, increasing interference power P_I at the receiver leads to decreasing SINR, which eventually falls below the threshold γ_{snr} needed for correct decoding of the message and hence no RSS value can be reported. In this situation, the maximum observable change of RSS value at the receiver is as follows:

$$\delta_{max} = 10 \log \left(1 + \frac{1}{\gamma_{snr}} \right). \tag{2}$$

The interesting point is that if δ_{max} is too small, with respect to the previously set threshold, then no change in RSS value is observed as long as the source message is correctly decoded. The visibility threshold δ represents the change in RSS value at the receiver that can be observed and interpreted as an increase by interference. Based on this threshold, we can distinguish three different operational regimes. In the noise-limited regime, RSS values at the receiver are not affected by RF interference. In the interference-limited regime, RSS values at the receiver are changed with interference power according to Equation 1. Finally, in the collision regime, RSS values at the receiver cannot be reported, since the source messages can not be decoded correctly, which corresponds to the case when the full packet-loss occurs. A graphical presentation of different regimes detected in our previous work is given in Figure 1. In figure, red and blue lines show the increase in variability of received RSS values and decrease in the Packet Reception Rate (PRR) with the increase in the interference power at the receiver, respectively.

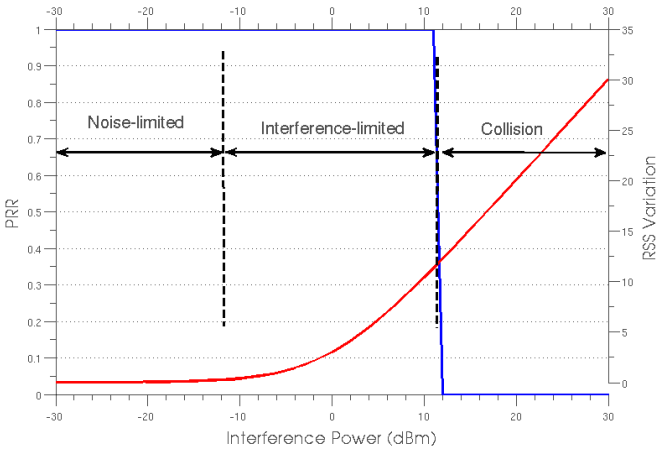


Fig. 1: Operational regimes for RSS-PRR variation

As it can be seen in Equation 1, the amount of change in RSS value can be measured if one has access to P_I and $P_X + N$. The destination can measure total received power $P_{RX} = P_I + P_X + N$, which corresponds to the numerator of the fraction inside the logarithm in Equation 1. If the interferer transmits continuously and if the destination can measure the power before the start of the transmission, the destination can also find $P_I + N$. But these two values are not enough to find the change δ in RSS value. In general, when the noise power N is comparable with received interference power P_I, nothing can be said about δ based on two measurements done by the destination. However, if P_I is much bigger than N, the value

$P_I + N$ can be roughly approximated by P_I and hence δ is given as follows:

$$\delta = 10 \log \left(\frac{P_{RX}}{P_{RX} - P_I} \right), \quad (3)$$

where $P_X = P_{RX} - P_I$. The previous equation provides the change of RSS in dB in terms of two measurements done by the destination. Let us assume that the RSSI measurements, reported by different technologies and specifically WiFi, are the quantized versions of RSS values discussed here. In this case, Equation 3 can be used to "correct" the RSSI measurements in interference-limited regime, under the assumption that the WiFi receiver provides the P_{RX} power level and the P_I is provided by leveraging spectrum information at the receiver.

III. General Workflow of Fingerprinting

The overview of a general workflow of WiFi fingerprinting using RSSI measurements from beacon packets sent by different Access Points (APs) in the environment of interest is given in Figure 2 and in details described in [14]. In the first step of the workflow, i.e. in the procedure for collection of raw RSSI measurements, the user generates a WiFi scan of an indoor environment at an unknown location. This scan is then sent to the fingerprinting server, and using some method in the fingerprint creation procedure a fingerprint is created out of the raw RSSI measurements. The workflow continues with a pattern matching procedure, in which the generated fingerprint is compared with fingerprints from a training database, which is comprised of a set of fingerprints previously surveyed in the environment of interest. Using some method in the pattern matching procedure the similarity between a fingerprint generated by the user and the training fingerprints is calculated. A training fingerprint with the highest similarity to the user's generated one is reported as the estimated location. Optionally, a set of training fingerprints with the highest similarities to the user's generated one can be post-processed using some method in the post-processing procedure, and the result of the post-processing procedure is reported as the estimated location.

IV. Interference Mitigation Procedure

We extend this general workflow of fingerprinting algorithms by inserting a procedure for reducing the effect of RF interference on the RSSI measurements. This procedure is applied on both the RSSI measurements collected in the training survey of the environment of interest and on user's generated RSSI measurements at an unknown location. The reason lies in the fact that, in both training and runtime phase of fingerprinting algorithms, collected RSSI measurements can be influenced by RF interference. An overview of the extended workflow of fingerprinting is given in Figure 2.

Similar to the general fingerprinting workflow, in the workflow with inserted interference mitigation procedure, the user collects RSSI measurements at an unknown location in the environment of interest. In addition, in the raw data collection procedure, the user at the same time samples the spectrum in the 2.4 GHz ISM frequency band and sends it to the fingerprinting server. Firstly, using the spectrum scan provided by the user, an interference power estimation method is leveraged for estimating the average interference power levels using the spectrum power levels. Using the user's provided beacon packets RSSI measurements, which also indicate the operating IEEE 802.11 channel for each AP, and using the estimated average interference power level on each IEEE 802.11 channel, according to Equation 3 the additive variability for RSSI measurements from each AP can be estimated. The estimated variability for each AP is then subtracted from original RSSI measurements, and these adjusted measurements are used as an input to the fingerprint creation procedure. The fingerprinting workflow continues further in the same way as in the general workflow of fingerprinting, with the exception that the same interference mitigation procedure has to be performed in the training step of fingerprinting to adjust RSSI measurements before storing them in a training database.

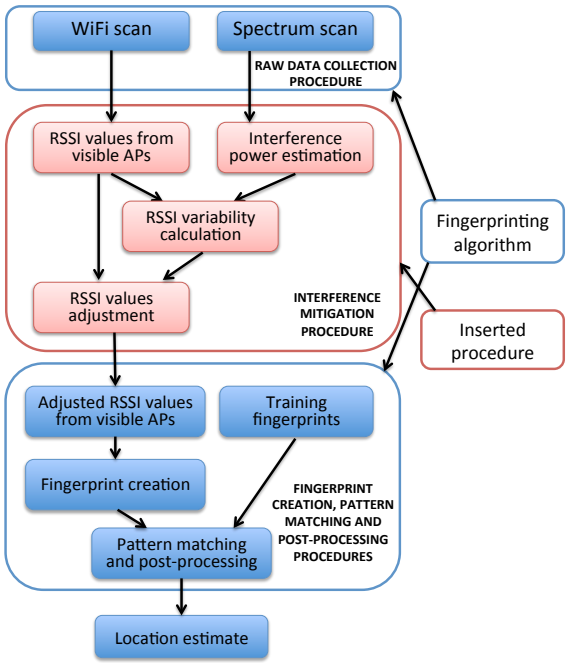


Fig. 2: Overview of the extended workflow of fingerprinting

A. Interference Power Estimation

In this work, we propose two methods for average interference power estimation using the sampled power levels on different carrier frequencies in the spectrum.

1) Interference Power Estimation Method 1: The first method uses sampled spectrum power levels on different carrier frequencies and estimates the average interference power level at the receiver on each IEEE 802.11 channel in the 2.4 GHz frequency band. The method firstly filters all samples that are considered as noise, i.e. smaller than the threshold of -99 dBm. This threshold is selected because it is a standard noise level for WiFi devices and because the power level below this threshold has no practical influence on RSSI measurements, as discussed before in the paper. Secondly,

the method for interference power estimation clusters samples sampled on different carrier frequencies into groups corresponding to IEEE 802.11 channels in the 2.4 GHz frequency band. For example, all samples taken at sampling frequencies between 2400 and 2422 MHz are arranged in a group that corresponds to IEEE 802.11 channel 1. The frequencies are clustered corresponding to IEEE 802.11 channels since the fingerprinting algorithms are WiFi-based. In other words, RSSI measurements used in fingerprinting algorithms are related to one IEEE 802.11 channel, making it reasonable to estimate the average interference power level corresponding to IEEE 802.11 channels. Finally, the method does averaging of the samples in each group and the result of averaging is the estimated interference power on each IEEE 802.11 channel in 2.4 GHz frequency band.

2) Interference Power Estimation Method 2: The second interference power estimation method, additionally to the sampled spectrum power levels, uses the probability of interference occurrence for estimating the interference power at the receiver. Instead of only averaging samples in each group, this method takes into account the ratio between the number of samples in which interference occurs and a total number of samples in one group (including noise samples). The method uses the calculated ratio as a weight with which the averaged samples in each group are multiplied. This ratio is a rough estimate of the probability of interference occurrence, and is used with the following intuition. If one IEEE 802.11 channel is highly used and the other one is less used, but they both have the same spectrum power levels when interference occurs, this methods will provide different average interference power level estimates for these two channels, while the previous method will result in the same estimate of the average interference power.

The proposed average interference power estimation methods at the receiver can reasonably be used under the condition that a time during which the signal occupies the spectrum is much lower than a time when there is interference in the spectrum. Otherwise, the methods would provide the estimated signal level from the spectrum information. Since for WiFi fingerprinting RSSI measurements from beacon packets are used, and beacon packets are by default transmitted periodically every 100 ms and have a short duration of around 50 μs depending on the transmission rate, the described methods can reasonably be used and will provide an estimate of the average interference power levels at different IEEE 802.11 channels in the 2.4 GHz frequency band.

The accuracy of the average interference power estimation based on the spectrum information depends on the spectrum sampling frequency. Naturally, if the sampling frequency is higher than Nyquist rate and therefore it is high enough to realistically capture the changes in spectrum power levels, the interference power can be obtained using the second method. If the sampling frequency is lower than Nyquist rate, it is possible to envision the usage of sparse signal processing methods, e.g. [15] to estimate precisely the average interference power. However, these methods need a specific spectrum sensing design. Current hardware, with its limitations, generally cannot support such high sampling frequencies and specific sensing design. Due to that, our average interference power estimation methods are based on two different heuristics. The first method considers a worst case scenario. It is assumed that the interference is present for the whole duration of signal transmission and therefore the average interference power level in the spectrum corresponds to the average value of powers during the measurement period. In other words, if the beacon packet is hit by interference, it endures the interference in its whole transmission duration. This provides the justification for filtering the spectrum samples below certain threshold. Worst case scenario is reasonable when an interference source is designed to transmit whenever the beacon packet is transmitted. In the second proposed average interference power estimation method, the main assumption is that interference can be present or absent during signal transmission. Particularly in this case, a beacon packet in its duration can be equally affected or not affected by interference. Therefore in the long run, the average interference effect on received power is dependent on the probability that the interference is present. The second proposed average interference power estimation method then estimates this probability by calculating the ratio of samples in which interference is transmitting to total number of samples.

The knowledge about the spectrum power levels at an unknown location is expected to be available for the users' end-devices. In other words, in the era of cognitive radios and seamless cooperation between heterogeneous devices this knowledge will become essential in enabling all the envisioned capabilities. Either in a form of a connectivity brokerage as a central entity for providing the information [16], [17], or by embedding the spectrum sensing capabilities in the end-devices [18], the information about the spectrum power levels is expected to be available. It is already possible to leverage the functionalities of various WiFi chipsets to measure 2.4 GHz spectrum information, e.g. for Atheros chipsets[1]. Under this assumption we use the information about the spectrum as given, in order to show the feasibility of our system as a proof-of-concept.

V. Experimental Evaluation

In this section, we provide a description of the evaluation procedure and an overview of four different artificially generated interference scenarios used for the performance evaluation. Finally, in this section, two WiFi RSSI-based fingerprinting algorithms used in the experimental evaluation are presented.

A. Evaluation Methodology

The evaluation methodology in this work follows guidelines established in the EVARILOS Benchmarking Handbook (EBH) [19], which is aligned with the upcoming ISO/IEC 18305 standard "Test and Evaluation of Localization

[1]http://wireless.wiki.kernel.org/en/users/drivers/ath9k/spectral_scan

and Tracking Systems". The EBH promotes the usage of a well defined evaluation procedure and evaluation scenarios, with the particular aim on scenarios with artificially generated RF interference. The result of the evaluation, according to the EBH, is a set of performance metrics, from which in this work we select point accuracy, room-level accuracy and processing time as the most relevant ones.

All the evaluation experiments presented in this work were performed using a testbed infrastructure specifically designed for the evaluation and benchmarking of RF-based indoor localization algorithms in the scenarios with artificially generated RF interference context [20]. The leveraged testbed infrastructure provides the possibility of collecting highly accurate measurements with minimized external influences, such as uncontrolled interference, influence of experimenter's body, etc. The experiments were performed during weekend afternoons, so the influence of uncontrolled interference, people walking and slight movements of objects (chairs, tables) in the testbed premises have been minimized.

As described before, fingerprinting algorithms generally require a training step in which the localization environment is surveyed for a set of distinguishable features (training fingerprints). The locations of training points for both fingerprinting algorithms used in the evaluation are given in Figure 3. The testbed infrastructure was used for collecting the raw RSSI measurements in different interference scenarios. For each interference scenario two sets of measurements were collected at 20 evaluation points with locations indicated in Figure 4.

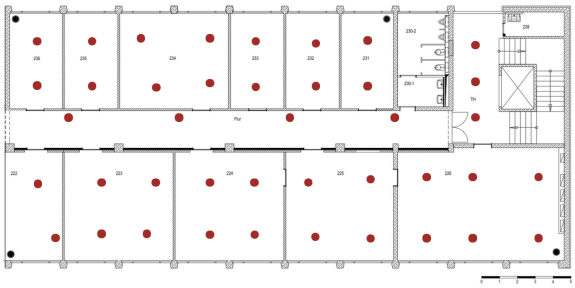


Fig. 3: Locations of training points

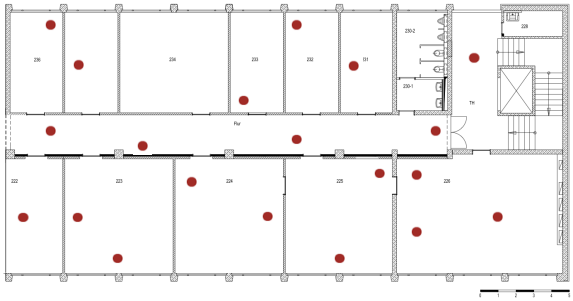


Fig. 4: Locations of evaluation points

Two repetitions of the same experiment provide additional insight into the temporal stability of the obtained results, which strengthens the reliability of our observations. The collected raw RSSI measurements were stored in a web-based platform for streamlined experimental evaluation of RF-based indoor localization algorithms using previously collected raw datasets [21]. This web-based platform provides a simple way of reusing datasets for evaluation of different RF-based indoor localization algorithms. By leveraging the platform, we were able to use the same raw datasets for the evaluation of the two described fingerprint algorithms without and with the procedure of mitigating the influence of RF interference and with two different methods for estimating interference power levels in the interference mitigation procedure.

B. Interference Scenarios

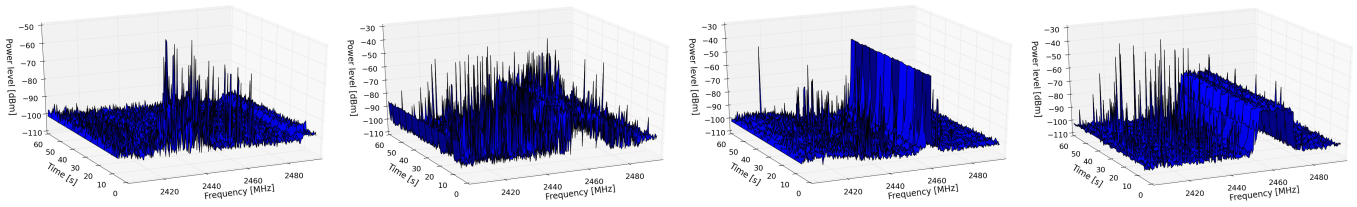
The evaluation was performed in four interference scenarios which are adopted from the EVARILOS project and are in details described in [12]. The spectrum power levels at an example location in the environment for each interference scenario is depicted in Figure 5. The leveraged interference scenarios do not represent a real-life interference context in which multiple sources of different types of interference are expected. On the contrary, the generated interference scenarios are simplistic on purpose, which gives an opportunity to evaluate if and to what extent interference of a particular type can influence the performance of evaluated algorithms.

a) Reference scenario (no interference): The name reference scenario reflects the fact that in this scenario no artificial interference was generated and the presence of uncontrolled interference was minimized, making the performance of fingerprinting algorithms achieved in this scenario a "reference" for evaluating the impact of interference generated in the interference scenarios.

b) Interference scenario 1 (IEEE 802.11b traffic): The second interference scenario was comprised of several interference sources that are typical for office or home environments. Interference was emulated using 4 WiFi embedded Personal Computers (PCs) having the roles of a server, access point, data client, and video client. During this scenario, the server acted as a gateway for the emulated services. A data client was emulated as a TCP client continuously sending data over the AP to the server. Similarly, a video client was emulated as a continuous UDP stream source of 500 kbps with bandwidth of 50 Mbps. The AP was working on a WiFi channel 11 (2462 MHz) and with the transmission power set to 20 dBm.

c) Interference scenario 2 (IEEE 802.15.4 jamming): In the first interference scenario, the interference was created using the IEEE 802.15.4 Tmote Sky nodes. The interference type was jamming on one IEEE 802.15.4 channel with a constant transmit power equal to 0 dBm. Five of these jamming nodes were present in the testbed environment. The wireless channel on which jamming was performed was IEEE 802.15.4 channel 22 (2460 MHz).

d) Interference scenario 3 (IEEE 802.11b jamming): For the third interference scenario, a signal generator was used to generate synthetic interference with an envelope that reflects WiFi modulated signals, but without Carrier Sensing (CS). The transmission power was set to 20 dBm, while the wireless channel was set on WiFi channel 11.



(a) No interference (b) 802.11 traffic (c) 802.15.4 jamming (d) 802.11 jamming

Fig. 5: Spectrum information provided by a WiSpy device in four interference scenarios at an example evaluation point

C. Experimental Setup

The environment used for evaluation is given in Figure 3, in which also the locations of training points are indicated with red dots. Fingerprinting algorithms leverage the information from all visible WiFi APs in the 2.4 GHz ISM frequency band. In addition to that, we set-up four additional WiFi APs in the corners of the environment, with their locations marked in figure with black dots. Those APs are configured to operate on IEEE 802.11 channel 11 (2462 MHz) with the transmission power of 20 dBm (100 mW). The used traffic model is IEEE 802.11b. These APs are working constantly and generally provide the highest RSSI measurements in the environment of interest, meaning that they have a high relevance for the fingerprinting procedure.

As a localization device for collecting the raw RSSI measurements a MacBook Pro laptop with an AirPort Network Interface Card (NIC) was used. We leveraged the RSSI measurements provided by the AirPort driver's method *airport – scan*, which provided us a scan of all visible WiFi APs every 0.5 sec. Twenty of these scans were used in the training phase, and eight in the runtime phase of fingerprinting.

For the collection of the spectrum measurements we used a WiSpy 2.4x device, which provides power levels in 2.4 GHz spectrum. The WiSpy device samples the spectrum by sweeping through the frequencies from 2400 MHz to 2495 MHz with a frequency step of 333 kHz, with the sweep time of 507 ms. The example spectrum information provided by a WiSpy device is graphically presented in Figure 5. This type of information was used for estimation the interference power levels at each evaluation point. While a WiSpy device is limited in the granularity of spectrum and sweeping frequency, the obtained data is sufficient for use as a proof-of-concept.

D. Fingerprinting Algorithms

For the evaluation purposes, we selected two widely used WiFi RSSI-based fingerprinting algorithms and extended them with interference mitigation procedure.

a) Euclidean distance of averaged RSSI vectors: This simple, yet popular fingerprinting algorithm [22] is computing an average value of the RSSI measurements obtained from each AP used for localization. The fingerprint is a vector of average values of the RSSI measurements obtained from all APs used for localization in both training and runtime steps, where K is the length of the vector. Let $\overline{\boldsymbol{X}}_{t,m} = [\overline{RSSI}_{t,1}, ..., \overline{RSSI}_{t,k}, ..., \overline{RSSI}_{t,K}]$ be the vector of averaged RSSI values $\overline{RSSI}_{t,i}$ from each AP i obtained in training step at point $m \in 1, ..., M_t$, i.e. training fingerprint. In the same manner, let $\overline{\boldsymbol{X}}_r = [\overline{RSSI}_{r,1}, ..., \overline{RSSI}_{r,k}, ..., \overline{RSSI}_{r,K}]$ be the vector of averaged RSSI values $\overline{RSSI}_{r,i}$ from each AP i obtained in runtime step, i.e. runtime fingerprint. The pattern matching procedure uses the Euclidean Distance (ED) between a training fingerprint at the cell m and the runtime fingerprint and it is given as:

$$D_E(\overline{\boldsymbol{X}}_{t,m}, \overline{\boldsymbol{X}}_r) = |\overline{\boldsymbol{X}}_{t,m} - \overline{\boldsymbol{X}}_r|. \tag{4}$$

Training fingerprints with the smallest distance (also called smallest weight) are then used in the post-processing procedure. In the post-processing procedure we used the non-weighed k-Nearest Neighbors (kNN) method with the parameter k set to 3, since it is shown in [14] that this method achieves the best performance results for this environment, in comparison to a set of other evaluated post-processing methods.

b) Pompeiu-Hausdorff distance of RSSI quantiles: A recently proposed procedure [23] uses a vector of q quantiles of the RSSI values from each AP as fingerprints, which are calculated in two steps. First the Cumulative Distribution Function (CDF) of the RSSI measurements from each AP is computed. Second, the quantiles, i.e. RSSI values with probabilities $k/(q-1)$, where $k = 0, 1, ..., q-1$, are calculated. The result of the quantile calculation in both training and runtime steps is a quantile matrix $Q_{K,q}$, where K is the number of APs visible at the given location and q is a number of quantiles. The pattern matching procedure of this algorithm uses the Pompeiu-Hausdorff (PH) metric for capturing similarities between fingerprints in a training dataset and the runtime fingerprint [23]. The Pompeiu-Hausdorff (PH) distance between two sets is given as follows:

$$D_{PH}(\boldsymbol{X}_{t,m}, \boldsymbol{X}_r) = \max_{x_{t,k} \in \boldsymbol{X}_{t,m}} \min_{x_{r,k} \in \boldsymbol{X}_r} d(x_{t,k}, x_{r,k}) \tag{5}$$

Here $d(x_{t,k}, x_{r,k})$ is the Euclidean Distance measurement between elements of the runtime fingerprint $\boldsymbol{X}_r$ and training fingerprint $\boldsymbol{X}_{t,m}$ at point m. The training point with the smallest PH distance with the runtime fingerprint is reported as an estimated location. In this paper we use PH as pattern matching procedure for RSSI quantile fingerprints. Same as in the previous algorithm, here we use the non-weighed 3NN method in the post-processing procedure.

VI. Evaluation Results

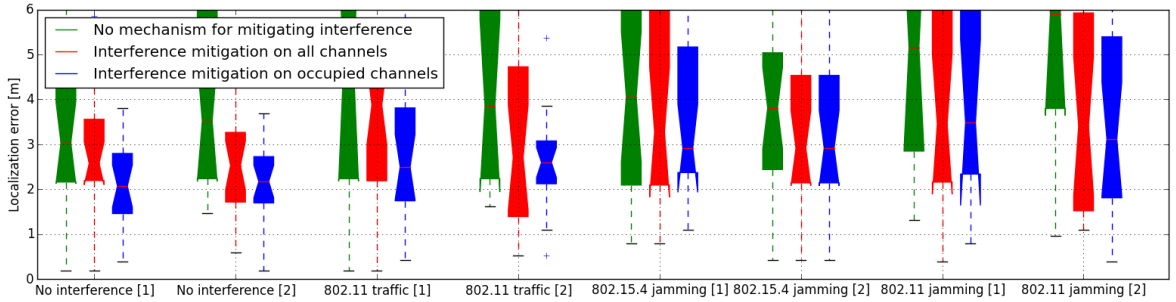
The distributions of localization errors achieved by two fingerprint algorithms in four artificially generated interference scenarios are given in Figure 6, in a regular box-plot fashion. Furthermore, statistical information about the point and room-level accuracy achieved by two algorithms is given in Table I. In the following we discuss three major observations.

The first observation concerns the stability of the obtained results. While it is impossible to have an entirely stable evaluation results, due to the intrinsic randomness of wireless environments, our results show that two repetitions (in two sequential days) of the same experiment result in highly comparable localization errors and their relevance and reliability is for that reason increased. This can be seen by comparing the localization errors obtained in two different runs of the same experiment, as shown in Figure 6 and Table I. The reason for high repeatability of the obtained results is also related to using the specifically designed automated testbed infrastructure for data collection. Namely, the infrastructure removes the need of a test-person for carrying the localization device to different evaluation points, which decreases the shadowing and fast-fading caused by the person's body and shaking of the person. The infrastructure further allows highly accurate and repeatable positioning of a localization device, with the average error in positioning being less than 15 cm [20]. Finally, the height of a localization device is exactly the same for all evaluation points and equals 55 cm above the floor level. While we cannot draw statistical claims on the stability based on only two repetitions, it does straighten our subsequent conclusions, i.e. that the interference can degrade the performance of fingerprinting algorithms and the interference mitigation methods are useful in improving the robustness of fingerprinting algorithms to RF interference.

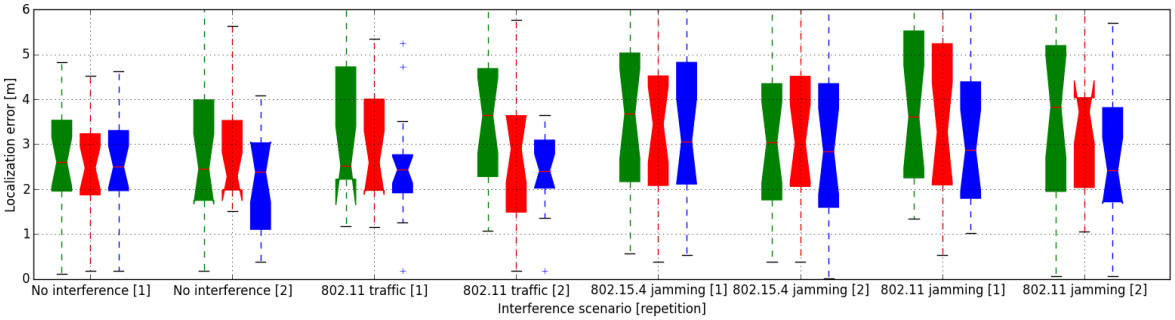
As for the second observation, the results show that RF interference can increase the achieved localization errors and, thus, decrease the accuracy of the evaluated fingerprinting algorithms. This can be observed by comparing the localization errors obtained by the algorithms without applying the interference mitigation procedure. In other words, the localization errors obtained in the evaluation in different interference scenarios are generally higher than the errors obtained in the scenario where no artificial interference is generated. It can also be observed that the smallest influence of interference on the accuracy of the algorithms occurs when the interference source is IEEE 802.11 traffic. The reason for this effect is the standard Carrier Sense Multiple Access (CSMA) mechanism in the IEEE 802.11 APs, which reduces the possibility of interfering with the beacon packets used for the localization purposes. The influence of interference is generally higher when interference source is jamming, since the CSMA mechanism does not exist. Finally, in the jamming scenarios, higher degradation of the accuracy of fingerprinting algorithms is observed when jamming is performed on a IEEE 802.11 channel. The reason for that is the higher transmission power (20 dBm) and wider channel (20 MHz), in comparison to the scenario where jamming is done on a IEEE 802.15.4 channel with transmit power of 0 dBm and 2 MHz wide channels.

The third observation is related to the improvements in the accuracy of evaluated fingerprinting algorithms by leveraging the spectrum information and applying the interference mitigation procedure. As the obtained results show, the accuracy of fingerprinting algorithms is improved in the interference scenarios when the interference mitigation procedure is used. At the same time, in the scenarios where the influence of interference on the accuracy of fingerprinting algorithms is not high, the interference mitigation procedure has no influence, i.e. it does not degrade the performance. For some scenarios, even in case no interference is generated, some improvements are visible. The reason is the uncontrolled interference in the environment, which, although minimized in our experiments, can still degrade the performance of fingerprinting algorithms and the interference mitigation procedure in that case helps. It is also visible that, in general, comparable or better performance is achieved when the average interference power estimation method used in the interference mitigation procedure takes into account the probability of the occurrence of interference, in comparison to the other method for interference power estimation. The reason lies in the fact that the first method only performs averaging of a set of spectrum samples on a given IEEE 802.11 channel, which, for the case when interference does not occur with high probability, results in too high estimates. For that reason, the difference in performance of methods 1 and 2 is clearly visible in the interference scenario 1, where the interference is IEEE 802.11 traffic. For the jamming scenarios the probability of interference is high, so both methods achieve comparable results. The improvements in the interference robustness due to leveraging an interference mitigation method are the smallest for the case when the source of interference is different technology (interference scenario 2). The reason is the difference in signal bandwidth for the two technologies, which indicates that a more accurate estimation of the interference power would result in further improvements in the interference robustness of WiFi-based fingerprinting algorithms.

Naturally, one consequence of introducing a new interference mitigation procedure in a fingerprinting algorithm is the increase in the processing time of an algorithm, i.e. increase in time needed for reporting a location estimate. We evaluated the processing time of the two used fingerprinting algorithms by requesting the algorithms to provide 100 times the location estimates for the 20 evaluation points in the reference scenario. The time needed for providing each location estimate was measured and afterwards the statistical information about the processing time needed for providing one location estimated was calculated. The same procedure was repeated for the case when interference mitigation procedure is introduced in the fingerprinting algorithms. In Table II, the obtained results are presented. Note that the presented processing times do not include the time needed for collecting the raw data, which is constant for both cases and depends on the measuring device and used device driver. Furthermore, the processing times in



(a) Euclidean distance of averaged RSSI vectors



(b) Pompeiu-Hausdorff distance of RSSI quantiles

Fig. 6: Evaluation results for both fingerprinting algorithms in different interference scenarios

TABLE I: Summarized evaluation results - point and room-level accuracy

	No interference		IEEE 802.11 traffic		IEEE 802.15.4 jamming		IEEE 802.11 jamming	
Metrics	Repetition 1	Repetition 2	Repetition 1	Repetition 2	Repetition 1	Repetition 2	Repetition 1	Repetition 2
Euclidean distance of averaged RSSI vectors without interference mitigation procedure								
Median error [m]	3.03	3.51	4.73	3.85	4.05	3.81	5.16	5.89
Average error [m]	4.48	4.69	6.63	5.73	4.41	4.29	6.37	6.85
Minimum error [m]	0.17	1.46	0.17	1.61	0.79	0.41	1.31	0.96
Maximum error [m]	14.4	14.72	17.28	15.67	10.11	11.01	19.32	16.77
Room-level accuracy [%]	55.0	40.0	50.0	50.0	30.0	50.0	40.0	25.0
Euclidean distance of averaged RSSI vectors - method 1								
Median error [m]	2.57	2.52	3.87	2.71	3.27	2.91	3.45	3.40
Average error [m]	3.06	3.00	4.21	3.69	4.33	3.86	4.55	4.46
Minimum error [m]	0.17	0.58	0.17	0.52	0.79	0.41	0.38	1.09
Maximum error [m]	8.23	8.51	17.82	14.84	10.11	11.01	12.72	12.75
Room-level accuracy [%]	60.0	70.0	65.0	65.0	40.0	50.0	40.0	35.0
Euclidean distance of averaged RSSI vectors - method 2								
Median error [m]	2.06	2.17	2.49	2.59	2.91	2.91	3.48	3.12
Average error [m]	2.42	2.46	2.93	2.72	3.84	3.86	4.94	4.32
Minimum error [m]	0.38	0.18	0.41	0.52	1.01	0.41	0.79	0.38
Maximum error [m]	7.11	6.57	7.60	6.19	10.27	1.01	11.76	12.72
Room-level accuracy [%]	75.0	80.0	80.0	75.0	40.0	50.0	45.0	50.0
Pompeiu-Hausdorff distance of RSSI quantiles without interference mitigation procedure								
Median error [m]	2.60	2.45	2.52	3.63	3.62	3.04	3.61	3.83
Average error [m]	2.81	2.91	3.38	3.59	3.81	3.67	4.66	3.90
Minimum error [m]	0.11	0.19	1.17	1.07	0.56	0.38	1.34	0.05
Maximum error [m]	6.10	6.41	6.36	6.45	8.31	11.20	14.66	12.07
Room-level accuracy [%]	70.0	70.0	45.0	55.0	50.0	40.0	55.0	60.0
Pompeiu-Hausdorff distance of RSSI quantiles - method 1								
Median error [m]	2.47	2.28	2.59	2.89	3.44	3.03	3.26	3.71
Average error [m]	2.60	2.81	2.98	2.82	3.58	3.77	3.78	3.89
Minimum error [m]	0.19	1.50	1.16	0.18	0.38	0.38	0.53	1.06
Maximum error [m]	6.07	5.64	5.34	7.31	8.31	11.20	9.30	14.66
Room-level accuracy [%]	75.0	80.0	70.0	60.0	50.0	45.0	70.0	65.0
Pompeiu-Hausdorff distance of RSSI quantiles - method 2								
Median error [m]	2.49	2.38	2.42	2.41	3.06	2.83	2.86	2.41
Average error [m]	2.88	2.49	2.48	2.44	3.28	3.34	3.39	3.08
Minimum error [m]	0.19	0.38	0.19	0.19	0.54	0.03	1.03	0.05
Maximum error [m]	8.05	6.10	5.25	3.64	6.04	11.20	7.39	8.39
Room-level accuracy [%]	80.0	85.0	65.0	70.0	55.0	40.0	75.0	70.0

the interference scenarios are statistically the same as the processing times obtained in the reference scenario, so only results for the reference scenario are presented. The selection of method for estimation of the interference power also has a small influence on the overall processing time of an algorithm. For that reason, we only present results in which the used interference power estimation method is "method 2", i.e. the one that provides higher accuracy of indoor localization. The obtained results show that the increase in processing time of the algorithms due to the interference mitigation method is in average around 200 ms, which is an increase of around 20% in the processing time. However, the whole latency of providing location estimates consists of the time needed for RSSI collection and of the processing time of an algorithm. Depending on a hardware and a device driver, the time needed for obtaining the measurements from one WiFi scan is around 2-3 sec, making it a dominant factor in the overall latency. For that reason, the increase in the overall latency of providing location estimates due to the interference mitigation method is practically of a small importance.

TABLE II: Summarized evaluation results - processing time

Metric	No mitigation	With mitigation
Euclidean distance of averaged RSSI vectors		
Average processing time [s]	0.7249	0.9649
Median processing time [s]	0.7255	0.8523
Minimum processing time [s]	0.6967	0.8138
Maximum processing time [s]	0.8024	2.2834
Pompeiu-Hausdorff distance of RSSI quantiles		
Average processing time [s]	0.7340	0.8951
Median processing time [s]	0.6931	0.8603
Minimum processing time [s]	0.6705	0.7805
Maximum processing time [s]	1.0091	1.5935

VII. Conclusion and Future Work

In this paper, we proposed a procedure for mitigating the influence of interference by leveraging knowledge of spectrum power levels in the 2.4 GHz ISM frequency band. The procedure uses estimated average interference power levels in order to remove the additive variability in WiFi beacon packets RSSI measurements due to interference. We also proposed two methods for estimating the average interference power levels from the sampled spectrum power levels. In our testbed, we evaluated the performance of two fingerprinting algorithms in four RF interference scenarios with and without applying the interference mitigation procedure. In our experimental setup, we firstly demonstrated that our evaluation results are statistically stable in time, which increases their reliability. Secondly, the results show and confirm previous findings claiming that RF interference can reduce the accuracy of WiFi fingerprinting algorithms [12]. Finally, we demonstrated that by leveraging the proposed procedure for reducing the effect of RF interference the accuracy of evaluated algorithms is improved. The processing time of evaluated algorithms increases by roughly 200 ms in average when interference mitigation procedure is introduced to the system, which is of a small practical importance. The cost of our system, in comparison to the usual fingerprinting procedure, is the necessity of having sampled spectrum power information. Although in this proof-of-concept work we used a low-power spectrum analyzer (WiSpy) as a source of spectrum information, some WiFi chipsets can also provide this information. Future work will be oriented towards building and evaluating a more realistic interference robust fingerprinting system using some of the WiFi chipsets that can provide 2.4 GHz spectrum information.

Acknowledgment

This work has been partially funded by the European Commission (FP7-ICT-FIRE) within the project EVARILOS (grant No. 317989). The author Filip Lemic was partially supported by DAAD (German Academic Exchange Service).

References

[1] I. Junglas *et al.*, "Location-based Services," *Communications of the ACM*, vol. 51, no. 3, 2008.

[2] P. Bahl *et al.*, "RADAR: An In-building RF-based User Location and Tracking System," in *INFOCOM'00*, IEEE, 2000.

[3] Z. Xiang *et al.*, "A Wireless LAN-based Indoor Positioning Technology," *IBM Journal of research and development*, 2004.

[4] M. Sugano *et al.*, "Indoor Localization System Using RSSI Measurement of Wireless Sensor Network based on ZigBee Standard," *Target*, 2006.

[5] K. Chintalapudi *et al.*, "Indoor Localization Without the Pain," in *MobiCom'10*, ACM, 2010.

[6] V. Honkavirta *et al.*, "A Comparative Survey of WLAN Location Fingerprinting Methods," in *WPNC'09*, 2009.

[7] U. Varshney and R. Vetter, "Emerging Mobile and Wireless Networks," *Communications of the ACM*, 2000.

[8] J.-H. Hauer *et al.*, "Experimental Study of the Impact of WLAN Interference on IEEE 802.15.4 Body Area Networks," in *Wireless sensor networks*, Springer, 2009.

[9] P. Fuxjager *et al.*, "The Myth of Non-Overlapping Channels: Interference Measurements in IEEE 802.11," in *WONS'07*, IEEE, 2007.

[10] S.-Y. Lau *et al.*, "A Measurement Study of ZigBee-based Indoor Localization Systems Under RF Interference," in *WiNTECH'09*, ACM, 2009.

[11] A. Papapostolou and H. Chaouchi, "RFID-assisted Indoor Localization and the Impact of Interference on its Performance," *Journal of Network and Computer Applications*, 2011.

[12] F. Lemic *et al.*, "Experimental Evaluation of RF-based Indoor Localization Algorithms Under RF Interference," in *ICL-GNSS'15*, 2015.

[13] A. Behboodi *et al.*, "Interference Effect on Localization Solutions: Signal Feature Perspective," in *VTC2015-Spring*, IEEE, 2015.

[14] F. Lemic *et al.*, "Experimental Decomposition of the Performance of Fingerprinting-based Localization Algorithms," in *IPIN'14*, 2014.

[15] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic press, 2008.

[16] T. Yucek and H. Arslan, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications," *Communications Surveys & Tutorials, IEEE*, 2009.

[17] J. Rabaey *et al.*, "Connectivity Brokerage-Enabling Seamless Cooperation in Wireless Networks," *White Paper*, 2010.

[18] D. Halperin *et al.*, "Tool Release: Gathering 802.11n Traces with Channel State Information," *ACM SIGCOMM Computer Communication Review*, 2011.

[19] T. V. Haute *et al.*, "The EVARILOS Benchmarking Handbook: Evaluation of RF-based Indoor Localization Solutions," in *MERMAT'13*, 2013.

[20] F. Lemic, J. Büsch, M. Chwalisz, V. Handziski, and A. Wolisz, "Infrastructure for Benchmarking RF-based Indoor Localization under Controlled Interference," in *UPINLBS'14*, 2014.

[21] F. Lemic *et al.*, "Web-based Platform for Evaluation of RF-based Indoor Localization Algorithms," in *Communications Workshops (ICC)*, IEEE, 2015.

[22] D. Milioris *et al.*, "Low-Dimensional Signal-Strength Fingerprint-based Positioning in Wireless LANs," *Ad Hoc Networks*, 2011.

[23] F. Lemic, "Benchmarking of Quantile based Indoor Fingerprinting Algorithm," Tech. Rep. TKN-14-001, 2014.