

# Multiple simultaneous Wi-Fi measurements in fingerprinting indoor positioning

Adriano Moreira<sup>1</sup>, Ivo Silva<sup>1</sup>, Filipe Meneses<sup>1,2</sup>, Maria João Nicolau<sup>1</sup>, Cristiano Pendão<sup>1</sup> and Joaquín Torres-Sospedra<sup>3</sup>

<sup>1</sup> Algoritmi Research Centre, University of Minho, Guimarães, Portugal

<sup>2</sup> Centro de Computação Gráfica (CCG), Guimarães, Portugal

<sup>3</sup> Institute of New Imaging Technologies, Universitat Jaume I, Castellón, Spain

adriano.moreira@algoritmi.uminho.pt, ivo.silva@eng.uminho.pt, filipe.meneses@ccg.pt,

maria.j.nicolau@algoritmi.uminho.pt, cristiano@dsi.uminho.pt, jtorres@uji.es

**Abstract**—The accuracy of fingerprinting-based positioning methods accuracy is limited by the fluctuations in the radio signal intensity mainly due to reflections, refractions, and multipath interference, among other factors. We consider that the fluctuations (often modelled as a Gaussian process for simplification purposes) can be minimized by exploiting the richness of multiple signals collected simultaneously through independent network interfaces. This paper introduces an analysis of Wi-Fi signals' statistics using simultaneous measurements which shows that RSSI values obtained from independent devices are not highly correlated. The low correlation between Wi-Fi interfaces might be exploited to improve the positioning accuracy. The validation of the proposed fingerprinting approach in a real scenario shows that the mean and maximum error in positioning can be reduced by more than 40% when five Wi-Fi interfaces are simultaneously used for fingerprinting.

**Keywords**—Indoor positioning; Wi-Fi fingerprinting; multiple interfaces; indoor vehicles

## I. INTRODUCTION

In contrast to outdoor positioning, where global navigation satellite systems (GNSS) are well established, it is well-known that providing an indoor equivalent to GNSS, that fits in any indoor environment, is a challenge. There is no gold solution and the best choice depends on many factors that have to be carefully balanced: complexity of the environment, economic costs, user requirements and technical parameters, among many others.

For tracking autonomous vehicles indoor there is not a leading technology and existing solutions are based on a large variety of ad hoc technologies [1]. The approach proposed in [2] for mobile robot localization is based in UWB range data and odometry. Adding complementary proprioceptive sensors is a common approach in mobile robotics according to the authors. The wheeled robotic assistant c-Walker used odometry and sensor fusion for positioning [3]. Sporadic RFID and visual marks matching were also used to enrich positioning. The authors of [4] proposed an integrated algorithm based on available technologies: vehicular communications, RFID, dead reckoning and GPS. Wi-Fi based positioning was explored for vehicle tracking in [5,6] as a GPS alternative for outdoors and low speed vehicles. In [7] a robust Wi-Fi fingerprinting method was also proposed for vehicle positioning.

Among all technologies for indoor positioning, Wi-Fi fingerprinting is a well-established solution. A Wi-Fi fingerprint

solution consists of measuring the Received Signal Strength Indicator (RSSI) of the available Wi-Fi Access Points (APs) in the environment. The location of a device is determined by computing the similarity between a fingerprint collected by the device in an unknown position and the radio map, a set of fingerprints previously collected at well-known positions. Wi-Fi fingerprinting is attracting the attention of many researchers because Wi-Fi connectivity is already widespread in many indoor environments, so the infrastructure that was deployed for communication can also be used for positioning. Although it is well-suited for smartphone-based positioning, it is not commonly applied for vehicle tracking due to its low expected accuracy, which is in the range of a few meters.

Locating yourself in a map or calculating the route to a place of interest are the natural positioning applications in smartphones, but taking a photo, checking the weather, looking for a restaurant in the neighbourhood, or checking the public transport status are also common activities that may require positioning. This positioning can be improved by using the Wi-Fi information, especially indoors where the GNSS might not be available. In fact, disabling Wi-Fi in smartphones may decrease location accuracy as result [8]. Despite Wi-Fi fingerprinting is attached to smartphones [8], we consider that it can also be suitable for assisting vehicle tracking, if the positioning errors are reduced.

This paper introduces a novel method for indoor positioning based on Wi-Fi fingerprinting. In contrast to traditional methods, where the fingerprints are collected by just one single network interface, the proposed method fuses the fingerprints simultaneously captured by multiple interfaces in order to reduce the noise present in the RSSI values and obtain a more robust positioning system. Combining multiple fingerprints with the aim of stabilizing the RSSI values has already been attempted [9-11], although through the average of the RSSI values collected over time at the same position. As far as we known, the basic idea of combining synchronized fingerprints collected using multiple interfaces has not been explored yet.

Section II describes Wi-Fi fingerprinting in detail. Section III presents the empirical statistics associated to the correlation of fingerprints simultaneously collected by multiple interfaces. Section IV describes the proposed multi-interface fingerprinting method and introduces a simulation study. Section V describes the experimental setup and results. Main conclusions resulting from this work are summarized in Section VI.

## II. WI-FI FINGERPRINTING

Wi-Fi fingerprinting is a positioning solution which relies on a fundamental assumption: the strength of the Wi-Fi signals measured in the environment have a unique signature - the Wi-Fi fingerprint - at a given location. Wi-Fi fingerprinting has two differentiated phases: calibration and operation. In the calibration phase, which occurs before the operational phase, the environment is surveyed in order to create a radio map. This process is empirically done by collecting Wi-Fi fingerprints at different reference points whose positions are known. Later, during the operational phase, a new fingerprint is collected with the signal strength of all detected APs in the unknown position. This fingerprint is compared with the fingerprints in the radio map and the unknown position is computed using the positions of the most similar fingerprints.

Wi-Fi technology was initially designed to support communications, not for positioning. The propagation of radio signals in an indoor environment is not easy to predict due to the presence of people [12,13] and existing obstacles, which create reflections, refractions and multipath interference [14] that impair very accurate positioning. Close locations may share similar fingerprints and, therefore, the typical accuracy of Wi-Fi fingerprint is within a few meters (1-10 m according to [15]). Advanced fingerprint methods report an accuracy of about 1 to 4 m [16,17], but the heterogeneity of testing environments hinders a fair comparison.

Montgomery established that the nuisance factors in an experimental setup can be controlled, uncontrolled or, even, unknown (noise) [18]. Although there are many factors that have a direct impact on the Wi-Fi fingerprinting accuracy (size of environment, number and location of reference points, number and distribution of APs, among others), their discussion is out of this paper's scope. This paper focuses on reducing the impact of the noise in RSSI values, which might originate from many diverse sources, into the positioning accuracy.

### A. Assumptions on the modelling of RSSI values

The addition of white Gaussian noise is the usual starting point for understanding basic performance relationships in the study of communication systems [19].

Elnahrawy et al., in [20], explored the fundamental limits of localization using signal strength in indoor environments. In order to better explore the limits of localization performance, they developed some area-based solutions. In one of them, they assumed that the RSSI values (and, therefore, the noise in the RSSI readings in a given position) follows a Gaussian distribution. Although this assumption was not always true, it significantly simplified their computations with little performance loss, according to the authors. For the Bayesian network solution that they proposed, a t-distribution was added to the net. They selected the t-distribution, rather than a Gaussian one, in order to better model the outliers of real data.

In [9] it was stated that the RSSI distribution model changed according to the mean RSSI value. On the one hand, larger tails to the left appeared as stronger the mean RSSI was. On the other hand, more symmetric distributions were obtained as the mean RSSI became weaker. According to their study, most of RSSI

distributions (70% approx.) were often left-skewed, despite the log-normal distribution being usually used.

In [21], Luo and Zhan presented a study about the characterization of RSSI values measured with smartphones. For 150 histograms, only 34.67% fitted in the Gaussian distribution according to the kurtosis coefficient. To improve the Gaussian distribution model, they proposed to adjust the standard deviation with the kurtosis coefficient. However, the bi-modal phenomena, that rarely appears, could not be modelled.

The authors in [22] also studied the distribution of the RSSI values by using both the RSSI values reported by the driver and also the FTT data for every OFDM sub-carrier obtained directly from the baseband (NICS with the Qualcomm/Atheros AR92xx and AR93xx chipsets). They stated that a clear decision about the RSSI modelling cannot be drawn just from histograms because of signal strength quantization, as also concluded in [23]. On the other hand, the spectral scan mode showed more precise intensity readings, and the normal distribution almost perfectly fitted the RSSI values. However, the tests were done outdoors, in free space, without the interference typically present in indoor environments.

Berkvens et al. [24] proposed an extension of a Gaussian-based sensor model that explicitly allowed an AP to be undetected and that also incorporated the uncertainty on the antenna gain, because of its pose, by using a convolution with a uniform distribution.

As stated in [21], there are some conflicting and ambiguous conclusions about the best model to fit the RSSI data distributions, and many other distributions are also considered in the literature. However, most authors agree that the RSSI histograms resemble a Gaussian distribution in most of the cases.

### B. Other common assumptions

One may think that devices from the same brand or model should perform equally. Lui et al. performed some analyses to determine how different devices behave in a practical empirical evaluation [25]. First, Wi-Fi devices from the same vendor might not have the same performance according to their tests. Similarly, different generations of the same device might also not have the same performance. Furthermore, an empirical test showed that even three identical Wi-Fi cards, that provided similar averaged RSSI values at most of the testing points, differ significantly at a few testing points. In the empirical tests reported in [21], it is also shown that data collected by different smartphones vary significantly, even for samples taken simultaneously in the same place. Therefore, the assumption of device equivalence must be checked beforehand and a calibration step is necessary if multiple Wi-Fi devices are used.

Another common assumption in Wi-Fi fingerprinting is that any device is suitable for positioning. Lui et al. also demonstrated that this assumption is not valid [25]. There are devices that may report non-valid or useless RSSI values, which makes them completely incompatible with positioning. Also, there are devices that might report unusual patterns, which turn them into undesirable devices for positioning.

### III. THE STATISTICS OF WI-FI SIGNALS

Solid conclusions about the statistics of RSSI values collected using real Wi-Fi network interfaces in real world settings are difficult to draw from the results previously published by several research teams around the world. Aiming to contribute to clarify this issue, we report here some results on our own analysis of the RSSI statistics.

#### A. Our own analysis setup

Data to support this study was collected in the Ubicomp lab, at the Engineering building, University of Minho, Portugal, and also in the Geotec lab, at the University Jaume I, Spain. All data were collected during working hours, with several people standing seated at their working tables and also moving around the labs without any constraints.

The setup used to collect data in the Ubicomp lab is shown in Fig. 1. It is built around a Raspberry Pi 3 Model B to which 4 USB Wi-Fi interfaces (Edimax EW-7811un) are connected, some through a USB hub (D-Link, DUB-H4). A specially crafted Java program, using 5 threads, was used to collect RSSI samples from the 5 Wi-Fi interfaces (4 USB + the internal interface of the Raspberry Pi 3) simultaneously. Each sample includes a timestamp, the identification of the Wi-Fi interface used to collect the sample, and a list of all detected Access Points (APs) represented by their MAC addresses, RSSI value, SSID and radio channel number.

Several configurations for the USB Wi-Fi interfaces were considered, as illustrated in Fig. 2. These were used to assess the impact of the relative position between interfaces on the values of the collected samples. In all configurations, the maximum distance between interfaces was less than 15 cm which, at 2.4 GHz, is a little over one wavelength (12.5 cm). Data was collected for consecutive periods of one hour, at the maximum possible sampling rate. For each one of the configurations, the total number of samples collected per interface was around 2500 (1/1.44 samples/second), with small variations due to the variable sampling periods imposed by the processing time (parsing of data and writing them into a memory card).

Similar data was also collected with two of the same USB interfaces connected to a laptop computer (Asus P452LJ with Linux Mint). In this case, the internal Wi-Fi interface (Intel Corporation Wireless 7260) of the laptop was also used to collect data. This experiment was conducted to assess if different device drivers would impact the collected data. Data were

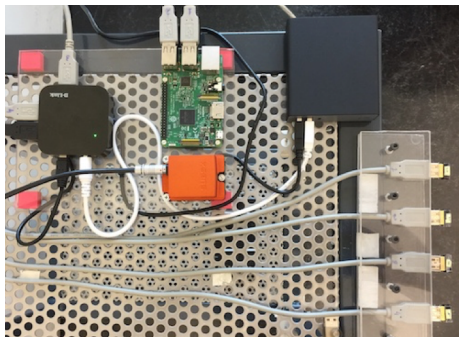


Fig. 1 Main data collection setup (Raspberry Pi 3 Model B + 4 USB Wi-Fi interfaces).

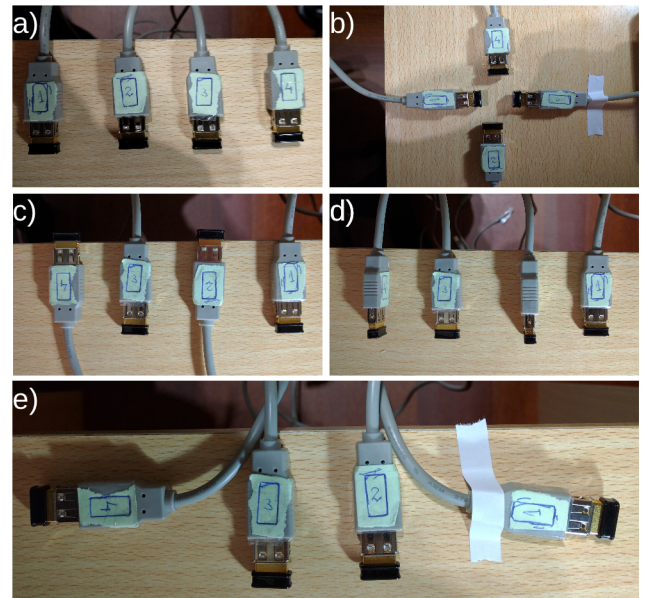


Fig. 2 Physical configurations for the USB Wi-Fi interfaces.

collected for 30 minutes at the maximum possible sampling rate. A total of 763 samples were collected from the internal interface (1/2.4 samples/second), and around 2500 from each one of the USB interfaces (1/0.72 samples/second). This sampling rate of the USB interfaces connected to the laptop was twice the one obtained with the Raspberry Pi. For some still unknown reason, the sampling rate of the internal interface was very low.

Four datasets were also collected using two Android smartphones (Samsung Galaxy S3, with Android 4.3, and Samsung A5 model 2015, with Android 6.0.1). In this case, data were collected only from the embedded Wi-Fi interface. Data were collected for two hours at a sampling rate of around 1/4.6 samples/second, on a total of 1560 samples per dataset.

Throughout this and the following sections, the random nature of the RSSI values is referred to as noise, even knowing that the observed variations are the result of noise introduced by the receiver, multipath propagation, interference from other radio sources, and other phenomena. Each sequence of RSSI values referring to the same AP observed from a particular interface is treated as a time series.

All these datasets were analysed and a summary of the results are presented next. Firstly, the distributions of the RSSI signals are analysed, and the results are compared to those obtained by other research teams. A model for the RSSI noise is also discussed. Secondly, the correlation between time series obtained from different Wi-Fi interfaces is investigated.

#### B. "Noise" distribution and modelling

Fig. 3 shows the sequence of RSSI values observed by the 4 USB Wi-Fi interfaces simultaneously, always at the same position, referring to the same AP.

This example shows that the same AP is observed quite differently by the 4 interfaces, despite the fact that they are very close to each other, and that the used hardware and software is similar. In this case, interfaces 1 and 3 measured similar RSSI values, while interfaces 2 and 4 measured significantly different

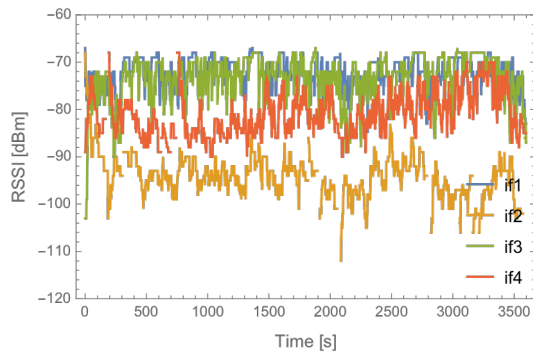


Fig. 3 RSSI time series obtained from similar Wi-Fi interfaces simultaneously (data from configuration represented in Fig. 2a).

values from the other three. Very similar results were obtained when observing other APs. A similar behaviour was also observed while using the other configurations illustrated in Fig. 2, showing that the relative physical orientation of the antennas does not play a major role in the differences among interfaces.

In several previous works on indoor positioning based on Wi-Fi fingerprinting, many research teams identified the difficulties arising from the use of different hardware for creating the radio maps and for collecting samples during the online phase [25, 26]. In order to minimize these difficulties, several authors proposed techniques based on conversion tables tuned for each type of Wi-Fi interface, and other techniques [9, 26]. The results in Fig. 3 show that even with similar hardware and software, the observed RSSI values can vary significantly.

An analysis of the distribution of RSSI values suggests a Gaussian distribution, although significantly skewed to the left (Fig. 4), as previously described by other authors (see Section II). In this case, the skewness of the distribution is very clear for the signal from interface if1 (the strongest one). The example shown in Fig. 4 also suggests that the range of RSSI values is much larger for interface if2 than for if1. While we did not find any evidence that stronger signals are always associated to more skewed distributions, we found that stronger signals tend to be more stable, i.e., have lower values of standard deviation. This trend is depicted in Fig. 5, where the standard deviation is plotted against the average RSSI for all the pairs (AP, Wi-Fi interface).

In Fig. 5 we also included data collected through the Raspberry Pi 3 internal interface (if0) to show that, in this case, the standard deviation is typically lower compared to that of the USB interfaces, and also less dependent on the average RSSI.

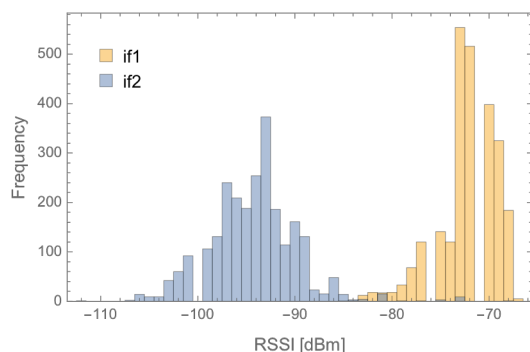


Fig. 4 Distribution of the RSSI values (for simplicity, only data from two interfaces is shown; data is the same as shown in Fig. 3).

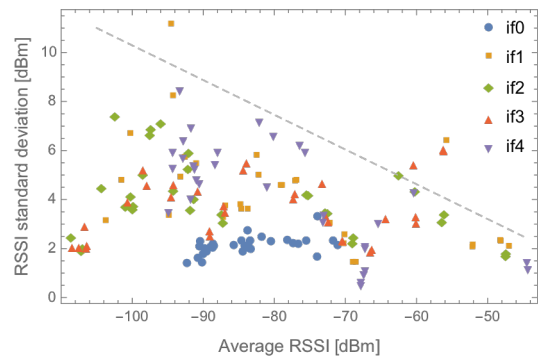


Fig. 5 Standard deviation vs. average RSSI for all pairs (AP, Wi-Fi interface) (configuration represented in Fig. 2a plus the internal interface, if0, of the Raspberry Pi).

On the other hand, when we analysed the data collected using the laptop computer, the RSSI values collected through the USB interfaces were observed to be a lot more stable, with standard deviation values always lower than 5. These results show that the same hardware combined with different software (device drivers and operating systems) leads to different noise statistics.

Given that the RSSI value is known to be affected by the dynamics of the space, such as the movement of people around the receiver, we investigated the stationarity of the corresponding random process. Fig. 6 shows (grey lines) the values of the standard deviation computed within a window of 2 minutes along the entire observation period, for a particular AP observed through 4 Wi-Fi interfaces. Similar calculations, and also regarding the variation of the average RSSI, were performed for every pair of (AP, Wi-Fi interface). These results, illustrated in Fig. 6, show that we are facing a non-stationary random process as both the mean and standard deviation vary along the time (see section IV for the meaning of the blue and orange lines in Fig. 6).

Modelling the RSSI noise has been addressed previously by many authors (see section II). We performed some analysis of the distributions that best fit the noise observed in our datasets, and concluded that stronger signals often resemble Gaussian distributions clearly skewed to the left, while weaker signals exhibit a more symmetric distribution around the mean value. This behaviour was observed in all our datasets, however data collected with the smartphones was found to exhibit less skewed distributions. Further details are not included here due to lack of space.

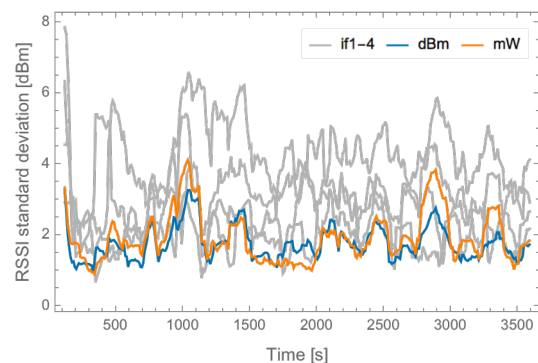


Fig. 6 Standard deviation computed for a sliding window of 2 minutes, over the time (shifted every 10 seconds).



### C. Correlation

In this section, we further investigate the properties of the RSSI signals by assessing the level of correlation between time series collected simultaneously through independent Wi-Fi interfaces. For this assessment, the Pearson correlation between synchronized time series was used. For each pair of time series, only the samples that were collected with a very short time difference between them were considered (which includes most of the samples for very visible APs, but might include just a subset of all the samples for APs with low visibility). An AP is considered to be very visible if it is detected in the large majority of the samples. An AP is usually very visible if near the receiver.

Fig. 7 shows the values of the Pearson correlation between all pair of interfaces, for all APs that were visible in at least 25% of the fingerprints. APs are sorted from the most visible, on the left, to the least visible, on the right. The values in Fig. 7 show that the time series are weakly correlated. The higher values on the right correspond to the least visible APs, for which the correlation values are not relevant because the time series have large gaps without samples.

The values of the correlation shown in Fig. 7 were computed for the entire duration of the time series, i.e., around one hour for this dataset. In order to evaluate the correlation on the short time scale, we also computed the Pearson correlation within short time periods of two minutes using a sliding window (10 seconds shifts). One illustrative result is shown in Fig. 8, where the correlation between a few pairs of time series, associated to the same AP, is plotted. It shows that, as with the standard deviation, the correlation between a pair of time series varies widely over time, with periods of considerably high correlation (values close to 1 and -1) and periods of low correlation (values close to 0). In general, there is not a clear overall correlation among all the interfaces. A scatter plot of RSSI values from different pairs of interfaces (not shown here) also reveals a low correlation between time series.

How this characteristic of the RSSI signals can be exploited to improve fingerprinting-based positioning systems is introduced in the next section.

## IV. USING MULTIPLE INTERFACES

The rationale for using multiple interfaces simultaneously is that their signals can be combined to, eventually, reduce the positioning error. We hypothesize that merged time series are less noisy and more stable in time, leading to improved performance of the positioning system.

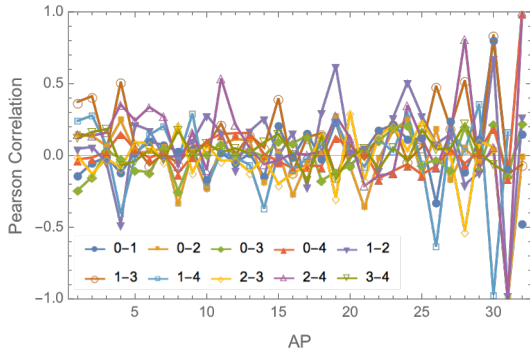


Fig. 7 Pearson correlation between all pairs of time series.

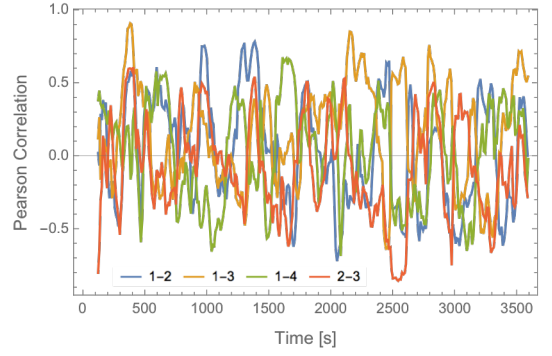


Fig. 8 Short term Pearson correlation (calculated for a sliding window of 2 minutes, shifted every 10 seconds).

Three alternative approaches have been considered to combine the time series. Let us assume that  $N$  interfaces are used, and that they all collect a sample at the same time instant (the interfaces are synchronized). We define a sample  $S$  as the list of the RSSI values measured by interface  $i$ , from each of the  $M$  Access Points, in the time instant  $t$ :

$$S_t^i = (\text{RSSI}_1^i, \dots, \text{RSSI}_M^i) \quad (1)$$

A combined sample can be obtained by simply computing the average of the RSSI values measured from each interface:

$$S_t = (\overline{\text{RSSI}}_1, \dots, \overline{\text{RSSI}}_M) \quad (2)$$

where

$$\overline{\text{RSSI}}_j = (\text{RSSI}_j^1 + \dots + \text{RSSI}_j^N) / N \quad (3)$$

Since RSSI values are usually represented in dBm, an alternative approach is to mimic a hypothetical merging in the analogue (radio) domain by computing the average of the received powers in Watts. It requires a conversion of all the dBm values to mW (eq. 4). Then, averaging is applied to the mW values. Finally, a conversion back to the dBm scale (eq. 5):

$$P_{\text{mW}} = 1\text{mW} \cdot 10^{P_{\text{dBm}}/10} \quad (4)$$

$$P_{\text{dBm}} = 10 \cdot \log_{10}(P_{\text{mW}}/1\text{mW}) \quad (5)$$

Yet another approach is to generate a combined sample by concatenating the  $M$  RSSI values measured by each one of the  $N$  interfaces, thus leading to a  $N \times M$  long sample:

$$S_t = (\text{RSSI}_1^1, \dots, \text{RSSI}_M^1, \text{RSSI}_1^2, \dots, \text{RSSI}_M^N) \quad (6)$$

We refer to these three alternatives as “dBm”, “mW”, and “concatenated”. We now investigate how the first two of these approaches affect the statistics of the RSSI signals, first in what concerns the noise standard deviation, then the stationarity, and finally the distribution. The impact of using multiple interfaces in the positioning error is then assessed through simulation and experimentally.

### A. Noise statistics

Fig. 9 shows the standard deviation for the combined signals associated to most visible APs observed by the  $N$  interfaces (same dataset used in the previous results). The “dBm” and “mW” approaches provide similar results, which point to a

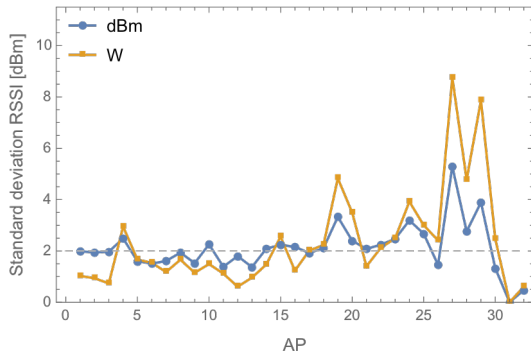


Fig. 9 Standard deviation of the combined signals.

significant reduction of the standard deviation when compared to that of independent interfaces (see Fig. 5 for comparison). While the standard deviation of independent signals is well above 2 dBm in most of the cases (Fig. 5), the corresponding values for the combined signals is around or below 2 dBm (the higher values shown in Fig. 9 are associated to APs that are barely visible).

An analysis of how the mean RSSI and standard deviation varies over time also shows a considerably more stable (more time invariant) signal when compared to that of individual interfaces (see Fig. 6 for a comparison between the independent and merged signals – blue and orange lines).

### B. Noise modelling

As we merge several time series, we would expect to get noise better described by a Gaussian distribution. While this is true for weaker signals, stronger merged signals still exhibit a skewed behaviour.

Fig. 10 shows an example of the distribution of RSSI values after merging the time series obtained through 4 identical interfaces. In this example, resulting from merging strong signals, the skewness to the left is clearly visible. An attempt to model the noise using Normal, SkewNormal, and logNormal distribution showed a poor fit: all three rejected the null hypothesis test using the Cramér-von Mises [27] method (p-values of 0,  $1.7 \times 10^{-11}$  and 0, respectively). On other cases, the noise of the merged signals was observed to be a lot more symmetrical around the mean value. Given these results, showing that a perfect model is difficult to generalize, in the next section we decided to model the noise using a Normal distribution with null mean.

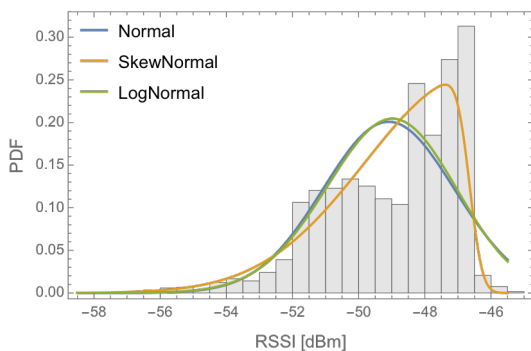


Fig. 10 Distribution of merged RSSI values, and corresponding models.

### C. Simulation

In order to quantify the potential benefits of using multiple interfaces in the reduction of the error in the position estimates, as hypothesised above, we first resorted to simulation. The considered scenario is a rectangular area of  $50 \times 20$  meters with 8 Access Points installed near the longer walls: four of them (APs 1 to 4) placed at 3.9 m above the floor, and the other four placed at 5.4 meters. Up to 25 simultaneous Wi-Fi interfaces were considered to collect samples.

A radio map was created by generating 10 samples per reference point for each one of the interfaces. A total of 931 reference points were placed on a grid with a one meter separation, starting at point (1,1) (no ref. points in the walls).

For evaluation, three different cases were considered. In the first case, a total of 9310 testing samples, with one fingerprint per interface, were generated at random positions uniformly distributed around the reference points. In the second case, a total of 1000 testing samples, with one fingerprint per interface, were generated over a grid with a one meter separation starting at point (0.5, 0.5). In the third case, a trajectory with 3601 points was randomly simulated through the environment. The simulations were repeated 100 times to obtain a representative average of the mean and maximum positioning errors. Each fingerprint was generated by computing the RSSI expected at the receiver position from each AP using the path loss radio propagation model with  $RSSI_0 = -40$  dBm and attenuation factor of 2. The noise was modelled as a Gaussian random process with null mean and  $\sigma = 4$  dBm.

For each one of the testing samples, the corresponding position was estimated using a simple KNN method, and computing the centroid of the  $k=3$  best matches. The Manhattan distance function was used to measure the similarity between fingerprints. The positioning errors were calculated as the Euclidean distance between the estimated and the true positions.

The comparison of the results obtained by using 1, 3, 5, 10 and 25 interfaces is shown in Table I, where the average mean and the average maximum positioning errors, over the 100 runs, are listed. Also, the standard deviation values are shown in the table. We considered 10 and 25 simultaneous interfaces in the simulation to explore the limits of using multiple interfaces, although it might be not practical in real implementations.

TABLE I. ERROR STATISTICS: MEAN AND MAXIMUM POSITIONING ERRORS (AVERAGED VALUES OVER 100 SIMULATIONS).

Conf.	Random Grid		Fixed Grid		Random Route	
	Mean [m]	Max [m]	Mean [m]	Max [m]	Mean [m]	Max [m]
Single IF	4.84±0.08	20.9±4.5	4.90±0.09	21.8±4.8	4.86±0.05	25.1±4.8
3 IFs dBm	2.98±0.05	10.4±1.3	3.04±0.05	10.7±1.2	3.00±0.03	11.9±1.2
5 IFs dBm	2.39±0.04	8.2±0.9	2.44±0.05	8.4±0.7	2.41±0.02	9.3±1.0
10 IFs dBm	1.75±0.03	6.2±0.8	1.80±0.03	6.2±0.6	1.78±0.03	6.9±0.8
25 IFs dBm	1.17±0.02	4.0±0.4	1.20±0.02	4.1±0.5	1.18±0.01	4.6±0.5
3 IFs mW	3.28±0.06	11.8±1.6	3.33±0.06	12.2±1.8	3.29±0.03	13.8±2.0
5 IFs mW	2.68±0.05	9.4±1.0	2.74±0.04	9.6±1.0	2.71±0.03	10.6±1.2
10 IFs mW	2.03±0.04	7.0±0.7	2.08±0.03	7.3±0.9	2.06±0.03	8.0±0.8
25 IFs mW	1.39±0.02	4.9±0.6	1.42±0.02	4.8±0.5	1.40±0.01	5.5±0.6
3 IFs conc	3.27±0.06	11.0±1.1	3.36±0.06	11.6±1.3	3.29±0.04	12.6±1.2
5 IFs conc	2.78±0.04	9.2±0.8	2.84±0.05	9.4±1.0	2.80±0.03	10.3±0.8
10 IFs conc	2.23±0.04	7.3±0.7	2.30±0.04	7.4±0.6	2.25±0.04	8.1±0.7
25 IFs conc	1.70±0.03	5.5±0.5	1.76±0.03	5.7±0.5	1.72±0.02	6.2±0.6

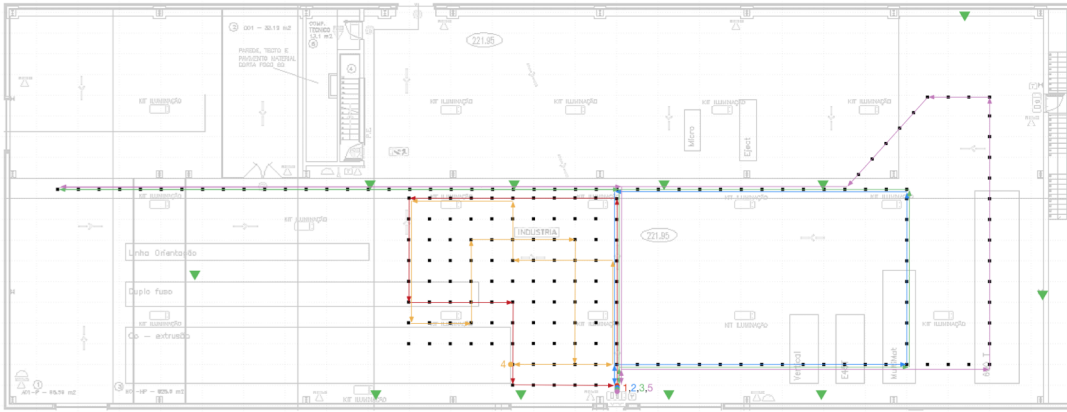


Fig. 11 Experimental environment: APs are represented by green triangles; ref. points by black dots; and trajectories by coloured lines.

The results from the comprehensive simulation show that using multiple interfaces reduces the mean and maximum positioning errors. The error is reduced to around 50%, 63% and 75% by using 5, 10 and 25 simultaneous interfaces, respectively. Moreover, using multiple interfaces also reduces the standard deviation in both metrics (mean and maximum error), which denotes that the variability in the 100 independent runs is much lower. i.e. positioning is better, more stable and less noise dependent when multiple interfaces are used.

Among the three alternatives to fuse data from multiple interfaces, the “dBm” approach is reporting the best accuracy in all the scenarios and cases (number of multiple interfaces). Moreover, it also reports the lowest deviation in almost all cases.

## V. EXPERIMENTS

Validation of the proposed method and of the simulation results presented above was performed through a real-world experiment. The experiment took place in a large laboratory used for polymer research work (PIEP) at the University of Minho. The space is around  $50 \times 20$  meters wide and more than 8 meters high and, in many aspects, is very similar to a factory plant, with large machinery, plenty of metal structures and tools, and some quite large open spaces. For the experiment, 11 Wi-Fi Access Points, operating in the 2.4 GHz band, were installed in the laboratory as shown in Fig. 11 (heights ranging from 3.9 to 5.4 meters from the floor, as for the simulation). A set of numbered paper tags were affixed to the floor to create a set of reference points (black dots in Fig. 11). Distance between adjacent reference points is one meter.

A radio map was created by collecting 20 samples at each reference point (around 4000 total samples). A set of 810 testing points was also created by collecting samples along a few trajectories (coloured lines in Fig. 11). These samples were collected using a small trolley manually moved over the trajectories at normal pedestrian speed (around 1 m/s). Each one of the samples is made of five fingerprints, each one of them collected through a different Wi-Fi interface: four USB Wi-Fi interfaces, and the internal Wi-Fi interface of a Raspberry Pi (see Section III.A). Ground truth data was manually collected, with the help of a time-synchronized video camera.

These data were processed to estimate the positions associated to each one of the testing samples. As in the simulation, a simple KNN estimation method, with  $k=3$ , was

used, and four approaches were considered: using fingerprints from one single Wi-Fi interface, and; using fingerprints from the multiple Wi-Fi interfaces merged using the “dBm”, “mW” and “concatenated” methods. For the single interface approach, positions were estimated independently for each one of the five Wi-Fi interfaces. For the multiple interface approach, positions were estimated using 3 interfaces (3,4 and 5; all USB), and all the 5 interfaces (USB and internal). In all cases, the positioning errors were computed as the Euclidean distance between the estimated positions and the true positions obtained from the ground truth data.

Fig. 12a shows the CDF of errors measured for the single interface, 3 interfaces and 5 interfaces approaches (“dBm” method). The results in Fig. 12b compare the results for the “dBm”, “mW” and “concatenated” methods (5 interfaces). These experimental results confirm the results obtained through simulation. Firstly (Fig. 12a), a considerable gain is achieved when data from multiple interfaces is combined to estimate the positions, with the simultaneous use of more interfaces leading to lower errors. In this particular case, the mean error obtained with 5 interfaces (2.39 m) is 43% lower than the mean error obtained by using a single interface (4.20 m). Secondly (Fig. 12b), among the three approaches considered to combine the fingerprints from the multiple interfaces, the “dBm” approach performs better, although with a marginal advantage over the “concatenated” approach (mean error of 2.39 versus 2.51 m). These results also confirm a considerable reduction on the maximum error: 8.47 m for the 5 interfaces “dBm” approach versus 17.33 m for the single interface approach (51% lower).

## VI. CONCLUSIONS

In this paper, we proposed to use data collected from multiple synchronized Wi-Fi interfaces to improve fingerprinting-based indoor positioning systems. An extensive analysis of the RSSI values collected from multiple independent Wi-Fi interfaces revealed that correlation among them is not strong. This weak correlation can be exploited to improve the accuracy of fingerprinting-based methods by enabling a considerable noise reduction. Validation of the proposed method was performed through a real-world experiment in a large factory-plant-like indoor space. The obtained results, with mean errors of less than 2.5 m, show that the accuracy can be improved by more than 40% using just 5 Wi-Fi interfaces.

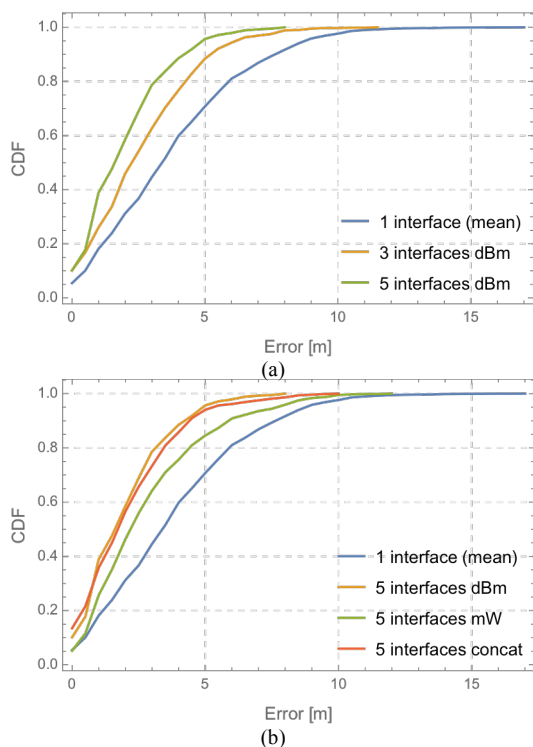


Fig. 12 Experimental results – error distribution: (a) comparison between 1, 3 and 5 interfaces, using the “dBm” merging approach; (b) comparison between the “dBm”, “mW”, and “concatenation” merging approaches. Results in both plots correspond to real experiments carried out in the scenario show in Fig. 11

#### ACKNOWLEDGMENT

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT—Fundação para a Ciência e Tecnologia within the scope of project UID/CEC/00319/2013, by the Portugal Incentive System for Research and Technological Development in the scope of the projects in co-promotion n° 002814/2015 (iFACTORY 2015-2018), and by the José Castillejo mobility grant (CAS16/00072).

#### REFERENCES

- [1] D. Dardari, P. Closas, and P. M. Djurić, “Indoor tracking: Theory, methods, and technologies,” *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1263–1278, April 2015.
- [2] J. González, J. L. Blanco, C. Galindo, A. Ortiz-de Galisteo, J. A. Fernández-Madrigo, F. A. Moreno, and J. L. Martínez, “Mobile robot localization based on ultra-wide-band ranging: A particle filter approach,” *Robot. Auton. Syst.*, vol. 57, no. 5, pp. 496–507, May 2009.
- [3] P. Nazemzadeh, F. Moro, D. Fontanelli, D. Macii, and L. Palopoli, “Indoor positioning of a robotic walking assistant for large public environments,” *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 11, pp. 2965–2976, Nov 2015.
- [4] A. Amini, R. M. Vaghefi, J. M. de la Garza, and R. M. Buehrer, “Improving gps-based vehicle positioning for intelligent transportation systems,” in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, June 2014, pp. 1023–1029.
- [5] D. V. Nguyen, M. E. V. Recalde, and F. Nashashibi, “Low speed vehicle localization using wifi fingerprinting,” in *14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Nov 2016.
- [6] D. V. Nguyen, F. Nashashibi, N. Thanh-Huong, and E. Castelli, “Indoor intelligent vehicle localization using wifi received signal strength indicator,” in *2017 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, March 2017, pp. 33–36.

- [7] L. Wang, R. Umer, H. Ng, C. Ang, and W. Xu, “A Robust WiFi Fingerprinting-Based Vehicle Positioning System For Heterogeneous Devices,” in *22nd ITS World Congress*, 2015, pp. 5–9.
- [8] S. Nadler, “Mobile Location Tracking: Indoor and Outdoor Location Tracking,” pp. 194–209, 2017.
- [9] K. Kaemarungsi and P. Krishnamurthy, “Analysis of wlan’s received signal strength indication for indoor location fingerprinting,” *Pervasive Mob. Comput.*, vol. 8, no. 2, pp. 292–316, Apr. 2012.
- [10] C. Feng, W. S. A. Au, S. Valace, and Z. Tan, “Compressive sensing based positioning using rss of wlan access points,” in *2010 Proceedings IEEE INFOCOM*, March 2010, pp. 1–9.
- [11] S. Knauth, M. Storz, H. Dastageeri, A. Koukofikis, and N. A. Ma hser-Hipp, “Fingerprint calibrated centroid and scalar product correlation rssi positioning in large environments,” in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct 2015, pp. 1–6.
- [12] K. Kaemarungsi and P. Krishnamurthy, “Properties of indoor received signal strength for wlan location fingerprinting,” in *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, 2004. MOBIQUITOUS 2004., Aug 2004, pp. 14–23.
- [13] S. Garcia-Villalonga and A. Perez-Navarro, “Influence of human absorption of wi-fi signal in indoor positioning with wi-fi fingerprinting,” in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct 2015, pp. 1–10.
- [14] N. Marques, F. Meneses, and A. Moreira, “Combining similarity functions and majority rules for multi-building, multi-floor, wifi positioning,” in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Nov 2012, pp. 1–9.
- [15] R. Mautz, *Indoor Positioning Technologies*. Habilitation Dissertation. Institute of Geodesy and Photogrammetry, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, 2012.
- [16] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of wireless indoor positioning techniques and systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.
- [17] A. Khalajmehrabadi, N. Gatsis, and D. Akopian, “Modern wlan fingerprinting indoor positioning methods and deployment challenges,” *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2017.
- [18] D. C. Montgomery, *Design and Analysis of Experiments*. John Wiley & Sons, 2006.
- [19] B. Sklar, “Rayleigh fading channels in mobile digital communication systems Part I: characterization,” *IEEE Communications Magazine*, vol. 35, no. 9, pp. 136–146, Sep 1997.
- [20] E. Elnahrawy, X. Li, and R. P. Martin, “The limits of localization using signal strength: a comparative study,” in *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, 2004. IEEE SECON 2004., Oct 2004, pp. 406–414.
- [21] J. Luo and X. Zhan, “Characterization of smart phone received signal strength indication for wlan indoor positioning accuracy improvement,” *Journal of Networks*, vol. 9, no. 3, pp. 739–746, 2014.
- [22] M. Rademacher, K. Jonas, and M. Kessel, “An empirical evaluation of the received signal strength indicator for fixed outdoor 802.11 links,” in *Mobilkommunikation: Technologien & Anwendungen*, 2015, pp. 62–66.
- [23] S. Robitzsch, L. Murphy, and J. Fitzpatrick, “An analysis of the received signal strength accuracy in 802.11a networks using atheros chipsets: A solution towards self configuration,” in *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, Dec 2011, pp. 1429–1434.
- [24] R. Berkvens, H. Peremans, and M. Weyn, “Conditional entropy and location error in indoor localization using probabilistic wi-fi fingerprinting,” *Sensors*, vol. 16, no. 10, 2016.
- [25] G. Lui, T. Gallagher, B. Li, A. G. Dempster, and C. Rizos, “Differences in rssi readings made by different wi-fi chipsets: A limitation of wlan localization,” in *2011 International Conference on Localization and GNSS (ICL-GNSS)*, June 2011, pp. 53–57.
- [26] C. Laoudias, R. Piché, and C. G. Panayiotou, “Device self-calibration in location systems using signal strength histograms,” *J. Locat. Based Serv.*, vol. 7, no. 3, pp. 165–181, Sep. 2013.
- [27] H. Cramér, “On the composition of elementary errors,” *Scandinavian Actuarial Journal*, vol. 1928, no. 1, pp. 13–74, 1928.