Localising crowds through Wi-Fi probes[☆,☆☆]

Francesco Potortì[a,*], Antonino Crivello[a,b], Michele Girolami[a], Paolo Barsocchi[a], Emilia Traficante[c]

[a] *ISTI Institute of CNR, Via G. Moruzzi 1, 56124, Pisa, Italy.*
[b] *DIISM Department of the University of Siena, Italy.*
[c] *Cloud4Wi, Pisa, Italy.*

Abstract

Most of us carry mobile devices that routinely disseminate radio messages, as is the case with Wi-Fi scanning and Bluetooth beaconing. We investigate whether it is possible to examine these digital crumbs and have them reveal useful insight on the presence of people in indoor locations, as the literature lacks any answers on this topic. Wi-Fi probes are generated sparsely and often anonymised, which hinders the possibility of using them for targeted localisation or tracking. However, by experimenting in three different indoor environments, we demonstrate for the first time that it is possible to extract from them some positioning information. Possible applications include identifying frequented regions where many people are gathered together. In the described experimentation with sniffing devices we adopted fingerprinting interpolation, which requires no survey phase and automatically adapts to changes in the environment. The same process can be carried out using the Wi-Fi access points already installed in the environment, thus allowing for operation free of installation, surveying and maintenance.

Keywords: Passive indoor localisation; survey-free fingerprinting; Wi-Fi probe eavesdropping

1. Introduction

Most mobile device we bring with us regularly send Wi-Fi packets called *probe requests* to detect nearby Wi-Fi networks. The great majority of such messages are discarded by networking devices that receive them, yet these digital

☆☆DOI: 10.1016/j.adhoc.2018.03.011
*Corresponding author
Email addresses: potorti@isti.cnr.it (Francesco Potortì), antonino.crivello@isti.cnr.it (Antonino Crivello), michele.girolami@isti.cnr.it (Michele Girolami), paolo.barsocchi@isti.cnr.it (Paolo Barsocchi), etraficante@cloud4wi.com (Emilia Traficante)

crumbs hide potential for revealing some aspects of human behaviour, such as people mobility.

In this paper we analyse Wi-Fi probe request messages sent by mobile devices, with the goal of estimating the device positions in indoor environments. The purpose of this experiment is to demonstrate the feasibility of identifying crowded regions by sniffing probe request, a task that can be performed by Wi-Fi access points already installed in the region of interest, without any additional hardware installation and without the need for a surveying phase or maintenance.

Wi-Fi devices emit probes to discover the existence of access points (APs). This procedure is called active scan, in contrast with passive scans during which devices passively listen for Wi-Fi beacons sent by APs. We use sniffing devices designed by Cloud4Wi® [1], called FogSenses, which passively collect probes sent by stationary and mobile devices and their received signal strength (RSS).

We study some meaningful statistics of the gathered data, such as the received signal strength (RSS) distribution of the probes and how the number of probes vary according to the periodic activities of people.

Using RSS, we then evaluate the accuracy of some range-free indoor localisation techniques, which are techniques that do not rely on the radio propagation properties of the environment. Most of our experiments are concerned with *fingerprinting*, a method based on RSS measurements made at a series of known significant points in the environment.

Usually fingerprinting requires a survey phase, which is a time-consuming measurement campaign devoted to populate a *fingerprinting database* which is then used to perform localisation; instead, we adopt an survey-free procedure to build the database, with a twofold advantage. First, we can build and periodically rebuild the fingerprint database without any human intervention, and without any extra cost; second, in case the number of Wi-Fi APs is too low, we can extend our sensing architecture by adding FogSenses without any manual reconfiguration other than registering the location coordinates of new FogSenses. Our work extends the preliminary study proposed in [8], by considering additional data collection campaigns in environments with different features, and by greatly expanding the span of location estimating algorithms, which we then combine into an *ensemble estimator*.

The goal of this analysis is to assess whether Wi-Fi probes are suitable to identify sample locations of human presence in indoor locations. We are not striving to provide the best possible localisation accuracy, but to demonstrate for the first time that Wi-Fi probes can be used to sample human presence locations.

We speak about sample locations because Wi-Fi devices emit probes only occasionally, so probes cannot be used to track a device or to reliably identify its presence. Such samples can be used for multiple purposes, for example

[1] "*the industry's leader in service platforms for advanced guest Wi-Fi*", `http://cloud4wi.com/`

as an aid to intrusion detection system, which would work by spotting the presence of unauthorised Wi-Fi devices in a given area. Another usage is as an activity heartbeat, by assessing the presence and regular activity of a stationary device in a given area. What we think is the most interesting is performing *crowd localisation*, that is estimating crowded regions, areas where several Wi-Fi enabled devices are located for some time. Our work does not perform any real-life experiments in any of these scenarios; rather, its purpose is to assess the feasibility of using Wi-Fi probes to such aims.

The experimental results we obtain show a median localisation error below 5.5 m, which is in line with the state of the art in indoor localisation algorithms based on Wi-Fi only, such as those participating to the EvAAL-ETRI off-site competition at IPIN 2015 [9]. It is worth noting that results obtained during the IPIN competitions are generally worse than those claimed by authors of indoor localisation papers, because they are measured in realistic and controlled conditions, rather than the same environment where a given system has been developed [10].

This is the first time that accurate measurements of probe-based localisation performance are done; they give a strong indication that it is indeed possible to implement and deploy practical systems that use only Wi-Fi probe detection to perform crowd localisation in areas such as a shop inside a mall or an open space office.

The remainder of this paper is organised as follows. Section II covers related work in the field of Wi-Fi probes used for localisation purposes. Section III describes our sensing architecture based on the FogSenses. Section IV describes the data gathering campaigns with an analysis of the quality of the obtained data set. Finally Section V introduces our localisation framework with a comparative analysis of the different techniques used. Section VI draws some conclusions.

2. Related work

Few works concentrate on sniffing Wi-Fi probes. Even less do so with the aim of localising people.

The main technical difficulty is that probes are sent only occasionally, as discussed in [3], with an experimental study of several factors that influence the number and the frequency of the probes sent by the popular smartphones. There are two major factors determining the behaviour of devices, namely the Operating System (OS) and the existence of known Wi-Fi networks. As an example, devices based on Android 5.0.1 are observed to emit about 1500 probes per hour in general, while for iOS devices (iOS 8.1.3) the number drops to 120 per hour. Devices usually send bursts of probes, the frequency of bursts strongly depends on the existence of known networks. The observed frequency of bursts ranges from one every 66 s (Android 5.0.1) to one every 330 s (iOS 8.1.3).

As a consequence, it is only possible to get sparse samples of people's positions. Our goal is to study whether Wi-Fi probes are usable to identify the presence of unspecified people in a given indoor area, without any attempt at

tracking or identifying specific devices. Specifically, our work focuses on the accuracy of the position samples through experimentation in a static environment.

In [7], Wi-Fi probes are used to estimate the trajectory of devices, which is a tracking task. This is made possible by instrumenting an arterial road 2.8 km long with 7 Wi-Fi monitors. The authors manage to track some individual devices with a median error of about 50 m with monitors placed on average 460 m apart. They use a hidden Markov model of possible trajectories and make the final estimate using the Viterbi algorithm. They do not only sniff for Wi-Fi probes spontaneously sent by mobile devices, but use several additional techniques to elicit response packets from devices and increase the length of packet bursts sent by each device, thus improving tracking performance once the device radio is on, but can do nothing when the device turns off the radio or anyway decides to not transmit anything. Accuracy performance of this approach is very good, considered how much far apart are the monitors, but it is only achievable in an environment with where two requirements are simultaneously satisfied: few well-defined possible trajectories and device tracking. Our work instead is aimed at being applicable in wide unstructured indoor areas such as a mall where a great number of trajectories are possible and configuring a Markov model would be a complex and long task, which contradicts our aim of simple setup. Moreover, many modern devices' operating system use some form of probe anonymisation which prevents tracking, unless the device is associated with a Wi-Fi network, which is generally not true.

The only other paper we found that exploits Wi-Fi probe messages for localisation purposes is [13], where the authors track pedestrians in an outdoor environment using triangulation. The final figures indicate a 1 m mean positioning error on a single experiment without any details on the number of samples taken. While it is possible to observe this high accuracy in a small outdoor environment with few obstacles, the adopted triangulation method is not generally usable in an indoor environment, where reflections from ceiling and floors are strong and no line of sight from the device to the monitor is a common situation, leading to a generally weak relationship between received signal strength and distance, which makes triangulation unreliable.

In summary, to the best of our knowledge there is no direct comparison for our work in the literature and no measurement campaigns, whether extensive or not, have been published on the positioning accuracy that one can obtain by eavesdropping Wi-Fi probe request packets using APs or sniffing devices.

3. The probe sensing architecture

Devices with an enabled Wi-Fi network periodically emit Wi-Fi probe requests. Their purpose is to actively scan the network searching for available Wi-Fi access points or for a previously accessed access point. This discovery phase usually prepares an association phase through which a device establishes a connection to a specific network. Devices send probes with frequency depending on several factors, including the Wi-Fi device driver and decisions made by

Figure 1: A FogSense Wi-Fi sensor used in the measurement campaign.

the operating system. For example, some devices do not perform any Wi-Fi scanning when they are connected to a wired network, while other devices still emit Wi-Fi probes even if they are connected. Probes are sensed by all APs in the area as part of their normal activity, as the IEEE 802.11 standard mandates. Using them for different purposes can be done internally to the APS or externally by a server to which APs send the collected probes. For simplicity of experimental set-up, we collect the probes emitted by Wi-Fi-enabled devices by means of a network of sniffing devices, namely FogSense devices distributed by Cloud4Wi. FogSenses are plug-and-play Wi-Fi sensors provisioned with a USB port as well as a mini-USB port for configuration. (figure 1). The Wi-Fi module is a Broadcom WICEDTM from USI, supporting IEEE 802.11 b/g/n Wi-Fi standards. A FogSense logs Wi-Fi probes emitted by nearby Wi-Fi-enabled devices and sends the logs to a server at intervals of 15 s. The data stored by the server includes information extracted from the captured probes, including reception time, MAC address of the sending device, ID of the receiving FogSense and RSS measured by the FogSense.

4. Experimental setting

We perform our experiments in three scenarios characterised by different layouts, sizes and number of sniffers needed to cover the area. Analysis of probes presented in this work are based on anonymised data. Maps of the three scenarios are shown in figure 2.

In a real deployment scenario probes are normally gathered by already installed APs, and FogSenses are only deployed if the number and positions of APs is not sufficient to obtain good accuracy performance. In our experiments, however, we only work with FogSenses, for simplicity of set-up.

The CNR area in Pisa (from now on CNR) covers about 350 m^2 and it is characterised by a straight corridor with offices located on both sides. The sensing region includes 12 offices where we deployed 4 FogSenses as shown in figure 2a. The Cloud4Wi Italian office (from now on C4WIT) covers about 250 m^2 and is located in an old historical building with 9 offices of irregular shape, where we deployed 8 FogSenses (figure 2b). Finally, the Cloud4Wi San Francisco headquarter (from now on C4WUS) is an open office covering about 500 m^2 with 3 small offices and a meeting area (on the right and on the top left side of figure 2c) where we deployed 5 FogSenses.

It is apparent from the maps that the three scenarios are quite different. C4WUS is an open space, with no obstructions. This is not very dissimilar from CNR, where the walls are part gasbeton and part drywalls, both of which are not a serious obstacle to Wi-Fi signals. On the other hand, C4WIT is quite different: this is an ancient building with many brick and stone walls up to 60 cm thick. We expect this scenario to produce more accurate results, because different FogSenses generally receive very well differentiated signal strengths from devices.

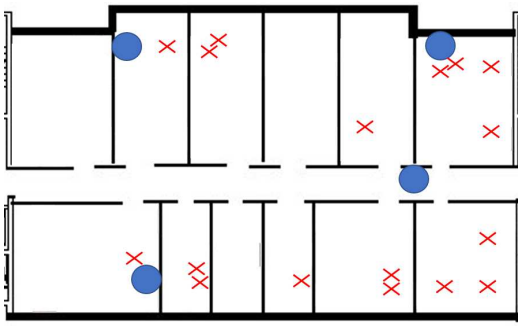
We installed different number of FogSenses in the three areas, specifically a higher number is needed in the C4WIT location, because the effect of the walls is similar to significantly increasing the distances.

To evaluate the performance of the proposed methods, we noted the position of some *known* stationary devices, e.g. workstations, laptops, smartphones and other Wi-Fi-equipped devices present in each location. The positions of known devices is the ground truth of our experiment: accuracy is measured by comparing their real position with the one estimated by different localisation methods. All devices are stationary: this is strictly true for workstations and laptops, and almost always true for the smartphones. Given office working habits, we estimate that each smartphone, during the whole experiment, is located into its known position for about 90% of the time it is found inside the measurement area.

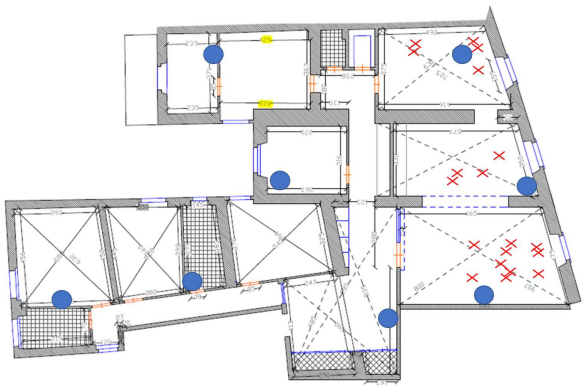
Note that experimenting with stationary devices, as we did, implies no generality loss with respect to experimenting with moving devices. The localisation procedure, in fact, relies on fixed sniffers to receive a packet sent by the device, at moving speeds that have no influence on radio propagation. Additionally, the prospected usage of the methods described in this paper is to gather samples of people's position, rather than tracking them, so the movement pattern of probe-emitting devices is largely irrelevant in this scope.

Among the known devices we could not include any that uses MAC address randomisation techniques, such as those based on recent iOS operating systems, because randomisation makes it impossible to identify which device is sending the Wi-Fi probe. While this is a limitation as far as our experiment is concerned, it does not impose any constraints for the intended usage of our technique which, again, does not involve tracking.

The data gathering campaigns have different duration, ranging from 30 days at C4WUS to 70 days at CNR, and different number of FogSenses installed in each scenario. We show the number of distinct MAC addresses that we observe



(a) Map of site CNR. Map width is 22 m.



(b) Map of site C4WIT. Map width is 25 m.

(c) Map of site C4WUS. Map width is 24 m.

Figure 2: Maps of the scenarios selected for the experiments. Blue dots show the FogSense positions, red crosses indicate the reference devices.

Table 1: Scenario characteristics

Scenario	FogSenses	MAC addrs	Duration	Known devices	Probes	Size
CNR	4	24000	70 days	16	2.2e6	350 m^2
C4WIT	8	130000	60 days	18	2.3e6	250 m^2
C4WUS	5	34000	30 days	12	1.6e6	500 m^2

in each scenario. The different numbers of observed MAC addresses are due to the proximity of offices to roads. Table 1 summarises the features of the three scenarios.

Since this paper is concerned with assessing whether Wi-Fi probes can be used for the purpose of localisation, and since no other measurement campaign of this kind is available, in appendix Appendix A we try to give an idea of the numbers we are working with.

5. Performance of localisation algorithms with Wi-Fi probes

We experiment with some localisation algorithms, in order to find the ones with best performance in terms of accuracy and robustness to changing environmental conditions. Our purpose is investigating whether we can find an algorithm with performance sufficient to be used as a basis for crowd localisation.

Generally speaking, RSS-based localisation techniques can be divided into *range-based* and *range-free* methods. Range-based techniques estimate a user's position by considering the received signal strength of that user's device and exploiting a Wi-Fi signal propagation model. They are prone to errors due to reflection of waves over the walls, floor and ceiling, especially in the presence of obstacles obstructing line of sight between transmitter and receiver. On the other hand, range-free techniques do not rely on the radio propagation properties of the environment. We only consider range-free algorithms.

Each *algorithm* we use has several parameters to be tuned. Choosing an algorithm and a set of parameters gives rise to a different localisation *method*.

All algorithms are based on k-NN classification, so each algorithm gives rise to different methods based on the value of k, which in our experiments varies from 1 to 3. Given the target application, we expect that each device is seen by a low number of FogSenses, so we have not experimented with higher values of k. The final estimate is the centroid of the k estimates.

The simplest algorithm, which we call *strongest*, estimates that the observed device is in the same location as the FogSense which has observed the highest RSS (Received Signal Strength), among those that have received the probe. When k is greater than 1, estimates are additionally considered for the second strongest up to the k-th strongest. Since we use k from 1 to 3, the *strongest* algorithm gives rise to 3 methods.

All the algorithms apart from *strongest* are based on *fingerprinting*. Fingerprinting is a technique commonly used for indoor localisation, which is composed of an installation *off-line* phase followed by a run-time *on-line* phase. During the *off-line phase*, one takes measurements of the RSS of Wi-Fi packets received from the APs (Wi-Fi Access Points), as observed at a number of reference points. These reference observations are collected into a *fingerprint database*. During the *on-line phase*, when an agent needs localisation, it makes a new observation, by measuring the RSS it gets from the visible APs at the location. This new observation is compared with those present in the fingerprint database. The entry in the database that is closest to the new observation is selected, and the agent's estimated position is set to that of the closest entry in the database, or the centroid of the k closest entries when k-NN is used. Fingerprint methods have been first proposed many years ago [1] and are still being actively investigated [12], since they are at the base of most indoor localisation systems. For example, all competitors in the EvAAL-ETRI off-site competition at IPIN 2015 used some form of fingerprinting [9].

Fingerprints observed during the on-line phase are variable in length, because the number of FogSenses receiving a given probe from a device is variable: in fact probes are lost for a variety of reasons, including collisions, interference and insufficient transmitting power. Generally speaking, the higher the number of FogSenses receiving a probe, the higher the accuracy of estimation we can possibly obtain, but the lower the number of probes we can consider as valid samples. The trade-off between accuracy and number of usable probes depends on the FogSense positioning, the number of devices expected in the area, the presence of other Wi-Fi networks, the expected accuracy of the results obtained and should be decided for each scenario, on a case-by-case basis. In this work, we use a threshold of 3 for all scenarios; in other words, we only consider probes which have been received by at least 3 FogSenses.

5.1. Interpolating the fingerprint database

Usually, building a fingerprint database starts with a survey phase during which several calibration points are selected. The purpose is to measure, at each point, what is the RSS observed from each of a number of APs in the area. In our case, we need the opposite: we should measure the RSS observed by the FogSense when a probe is sent by a device located at the calibration points. From a conceptual and practical point of view, this change of perspective is unimportant, and all the procedures commonly used for fingerprinting stay the same.

During the survey phase, the RSS values associated with each access point are collected at the calibration points over a certain period of time and then stored in fingerprint database together with the location coordinates. During the on-line phase, the person or object of interest is localised by comparing its observed fingerprint to those stored in the database, looking for the most similar ones. Building a fingerprint database is a time-consuming task, especially for large areas that may contain thousands of calibration samples.

In order to be commercially viable, the proposed method should require very little or no installation and maintenance measurements. To this aim, we completely remove the survey phase by profiting from the probes sent by the FogSenses themselves, which are connected to a server via Wi-Fi, and therefore occasionally send a probe request which is collected by the other FogSenses. This is enough to build a self-updating fingerprint database composed of fingerprints relative to the positions of the FogSenses. When using APs instead of FogSenses, we can profit from the probes sent by APs during routine neighbourhood scanning.

A fingerprint database obtained with this survey-free procedure, however, is too sparse for obtaining a satisfying positioning accuracy, because the typical density of FogSenses in the environment should be low. In order to get a denser fingerprint database, we resort to interpolation on a square grid, an idea already proposed in the indoor localisation literature [5, 6]. In particular, we further refine the solution proposed in [8] by exploiting several 2-D interpolation strategies.

5.1.1. Off-line phase: building the fingerprint database

The fingerprint database is automatically built with a survey-free procedure requiring no human intervention thanks to interpolation, which in real deployment scenarios allows for installation-free systems when the number and position of APs allows it, and for automatic fingerprint update when additional FogSenses are needed to improve positioning accuracy.

The first interpolation strategy we use is based on *linear interpolation over Delaunay triangulation* whose vertices are the known points, that is the FogSense positions. Note that this strategy does not provide extrapolation, which means that it provides no estimates for unknown points that lie outside of the convex hull of the known points.

The second interpolation strategy is *inverse distance*. At each unknown point, the estimate is the mean of the values at the known points, each weighted by the inverse of their distance to the unknown.

The third interpolation strategy is based on Kriging [5]. Kriging is an interpolation strategy originally adopted in the mining industry. Suppose that one can draw scalar samples from an unknown function of points belonging to a given domain. In our case, the samples are RSS measurement and the points in domain are the locations in the area where we take measurements. Kriging is based on the assumption that the variance of the difference of the samples taken at two different points is only dependent on the distance of the two points. The function that relates the variance to the distance is called *variogram*. We speak of *simple Kriging* when the mean of the samples is a known constant. *Ordinary Kriging* can work with an unknown constant mean. If we need to drop the constraint that the mean is constant, we resort to *universal Kriging*, where one can impose a trend on the mean of samples as a function of distance. This is our case, because the RSS expressed in dB can be modelled, as a first approximation, as a linearly decreasing function of distance. We assumed the same parameters adopted by [6]: *spherical model* with a *range* of 6 m, *sill* set

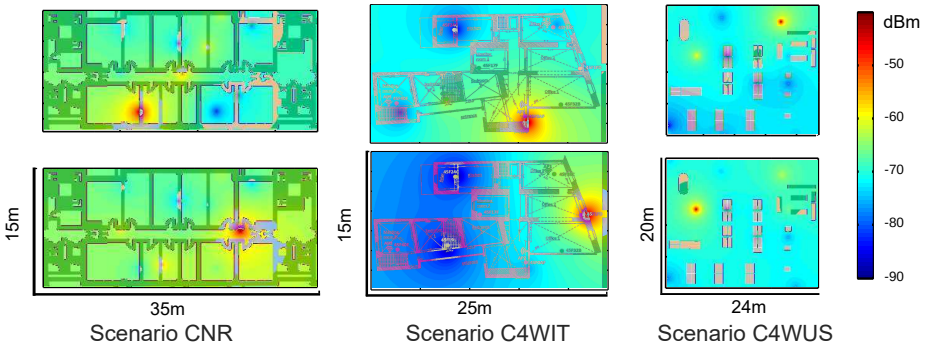


Figure 3: Examples of fingerprint maps generated with *inverse distance* interpolation.

to 31 dBm^2 and nugget set to 9 dBm^2 and linear trend. Our experiments have shown that these choices are in fact good enough in our scenarios.

By interpolating the measured cross-FogSense fingerprints over a regular grid, we obtain an interpolated set of fingerprints, that is, our final fingerprint database. Figure 3 shows some interpolated RSS radio maps. For illustration purposes, the maps are computed on a very small grid width of 10 cm. Each map is seen from the point of view of one FogSense, whose position on the map is the point where the RSS value is highest (the red point).

5.1.2. On-line phase: using the fingerprint database

During the on-line localisation phase, the fingerprint of the probe request sent by a mobile phone is compared with the RSS fingerprints stored in the database, an operation which requires a *measure of distance* to be defined. Fingerprints are N-D arrays, where N is the number of probes that are received. As stated above, we worked with $N \geq 3$. We experimented with several measures of distance: 1- and 2-norm distances, differential 1- and 2-norm distances, cosine distance and FreeLoc distance.

Given two fingerprints A and B of dimension N, the most usual distance is the Euclidean distance:

$$\|x\|_2 = (\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}}. \quad (1)$$

Generalising the Euclidean distance brings us to the p-norm distance:

$$||x||_p = (\sum_{i=1}^{n} x_i^p)^{\frac{1}{p}}. \quad (2)$$

Setting $p = 2$ produces the Euclidean distance, $p = 1$ is the Manhattan distance. A variation on the p-dist is the differential p-dist, where only the differences between the measured values of each vector are considered. Specifically, ND vector A is converted into an$(N-1)$D vector A_d:

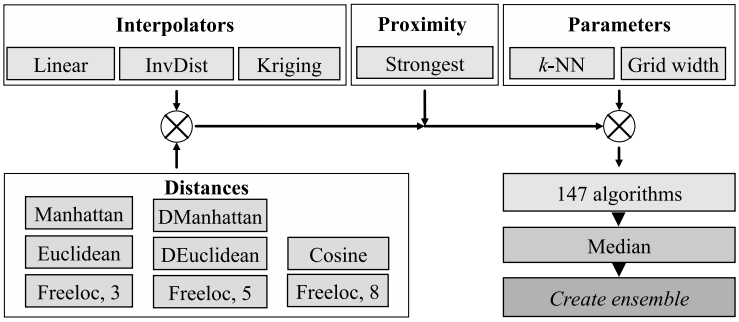


Figure 4: Generation of localisation algorithms. The ensemble creation box is expanded in figure 5

$$A = x_1, x_2, \cdots, x_N, A_d = x_2 - x_1, x_3 - x_2, \cdots, x_N - x_{N-1} \tag{3}$$

We call *differential* p-norm distance of vectors A and B the p-norm distance of vectors A_d and B_d. The purpose of differential p-norm distances is to remove the bias given by different devices possibly sending probes with different transmitting power.

The *cosine similarity* between two vectors A and B is a value in the interval $[-1, 1]$ defined as:

$$\frac{A \cdot B}{||A|| \times ||B||}. \tag{4}$$

Since we need a measure of dissimilarity, we (improperly) call *cosine distance* the complement to 1 of the cosine similarity.

The FreeLoc distance is inspired by [2]. The idea is that one should not rely on exact RSS values when comparing two fingerprints, but the only significant information comes from deciding whether the signal received by one FogSense is significantly higher, significantly lower or about the same as the signal received by another FogSense. This information is ternary, and coded as -1, 0 and +1 values. A threshold p is used to decide whether two signals are nearly equal ($|x - y| < p$), giving rise to a 0 or not, giving rise to a +1 or a -1. In our computations, we used for p one of the three values 3 dB, 5 dB, 8 dB (the latter being the value used in [2]).

Each fingerprinting vector A of length N is thus converted into a new A_f ternary vector of length $N \times (N-1)/2$, that is the number of pairs of the N dimensions. Comparing the ternary vectors is just a matter of obtaining their scalar product. As in the previous case, we obtain the complement to 1 of the similarity:

$$1 - (A_f \cdot B_f)/\frac{N \times (N-1)}{2}. \tag{5}$$

5.2. Creating ensemble estimators

Using the above-described building blocks, we define parametric algorithms for localisation, and for each we evaluate its performance in terms of accuracy. Once this is done, we turn our attention to performance in terms of robustness across varying scenarios. We start with definitions, we proceed to illustrating accuracy performance, and then we the consider trading some accuracy for robustness.

An *algorithm* is either the *strongest* algorithm or a fingeprinting algorithm. Fingerprinting algorithms are defined by the choice of an interpolator and a measure of distance. The choice of the interpolator, used in the off-line phase, affects the creation of the fingerprint database, while the distance is used in the on-line phase to identify the k fingerprints in the database which are closest to the measured fingerprint. Each algorithm is associated with several parameters to produce a set of *methods*.

For each algorithm, the parameters we consider are the interpolation grid size (not significant for *strongest*, which is not based on interpolating) and the k value. By varying the parameters, as shown in figure 4, we produce a spectrum of alternative methods. In summary, we have used 3 different interpolators and 8 different distances, which give rise to 144 methods based on fingerprinting, to be added to 3 more methods based on the *strongest* algorithm.

In order to compare the 147 methods, we choose the error median as an accuracy performance measure. We obtain an error median for each method applied to each of the three scenarios. Given a method and a scenario, the error median is computed by first obtaining the error distribution for each device of that scenario, and then merging together those distributions. This way the results are not dependent on the number of samples per device, which in fact are quite different, as shown in figure A.7.

Table 2 shows the performance of the 25 best methods for each scenario. Some methods that are used in the following discussion are marked with a letter id, whose meaning is listed in table 3.

We don't want to choose the best method for each scenario. Rather, we want to find a way to select methods that have good performance overall. To this end, we resort to the concept of *ensemble estimator*, which is employed in [11] for a similar case. Ensemble estimators (ensembles for short) are useful when dealing with optimisation on many discrete parameters. For example, in our case varying the parameters creates a total of 147 methods. Just choosing the method having the best performance would lead to overfitting. Overfitting, which means tuning the parameters to the specific case that is being analysed, can produce brittle methods, that is, methods that perform well only in a specific situation. In order to increase the robustness of the choice, and possibly the performance too, we select a set (an ensemble) of methods. Once the set is chosen, the position estimated by the ensemble estimator is defined as the centroid of the positions estimated by each method in the ensemble. To define an ensemble estimator, a criterion is needed to select the methods composing the ensemble. For example, a simple criterion would be to just choose the N best accuracy

Table 2: Median errors [m] of the best 25 methods for each scenario. Methods are indicated with keys listed in table 3.

	CNR scenario	Key	C4WIT scenario	Key	C4WUS scenario	Key
Methods	5.1	(a)	2.9	(e)	4.8	(C)
used for	5.1	(b)	3.0	(a)	4.9	(B)
reference	5.1	(c)	3.2	(f)	5.0	(h)
ensembles	5.2	(d)	3.2	(g)	5.0	(i)
	5.3		3.3		5.1	
	5.3		3.4	(B)	5.2	
	5.3		3.5		5.2	
	5.4		3.5		5.2	
	5.4		3.5		5.3	
	5.4		3.6	(A)	5.4	(D)
	5.4	(e)	3.6		5.4	
	5.5	(f)	3.6		5.4	
	5.5		3.7		5.5	
	5.5	(A)	3.7		5.6	
	5.5	(h)	3.7	(C)	5.6	
	5.6		3.7		5.6	
	5.6		3.7		5.6	
	5.6		3.8		5.7	
	5.6		3.8	(c)	5.7	
	5.6	(C)	3.8		5.7	
	5.6	(B)	3.9		5.7	
	5.6		3.9		5.7	
	5.6		3.9		5.7	(A)
	5.8	(D)	3.9	(D)	5.7	
	5.8		3.9		5.8	

Table 3: Keys used in table 2

Ensemble	Key	Interpolator	Distance	k	Grid width
CNR	(a)	invdist	cosine	2	2 m
(reference)	(b)	strongest		1	1 m
	(c)	linear	pnorm,1	1	2 m
	(d)	invdist	freeloc,8	1	2 m
C4WIT	(e)	invdist	cosine	3	2 m
(reference)	(a)	invdist	cosine	2	2 m
	(f)	invdist	cosine	2	3 m
	(g)	invdist	cosine	1	2 m
C4WUS	(C)	linear	cosine	3	2 m
(reference)	(h)	linear	cosine	2	2 m
	(B)	linear	freeloc,5	1	2 m
	(i)	linear	cosine	2	3 m
Intersect	(A)	linear	cosine	1	2 m
(ensemble	(B)	linear	cosine	2	2 m
of choice)	(C)	linear	cosine	3	2 m
	(D)	linear	pnorm,2	3	2 (m)

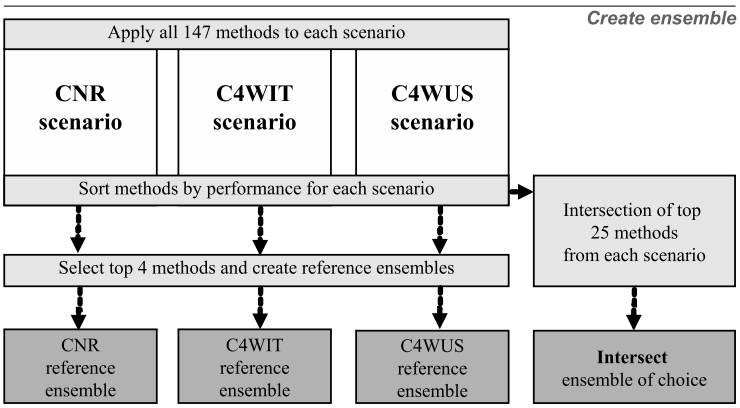


Figure 5: Generation of ensemble estimators.

performers among all the considered methods and use those as elements of the ensemble. More complex criteria are possible to select the methods that are part of an ensemble, see [4, 11] for more in-depth discussion.

The criterion we choose is depicted in figure 5 and is quite simple: we select the methods that appear among the best performers in all three scenarios, that is, a set of methods which is the intersection of the three sets whose accuracy performance is listed in table 2. The selected methods compose the *intersection ensemble*; they are marked with upper-case letters, defined in table 3.

In order to better evaluate the performance of the intersection ensemble, we compare it against three additional *reference* ensembles, each tuned on a different scenario. We create the CNR ensemble using the 4 methods having the best accuracy performance in the CNR scenario, and similarly for C4WIT and C4WUS. The methods composing these 3 scenarios are marked with lower-case letters in table 2. Note that the top performer methods are different in each scenario. For example method a is the best for scenario CNR and the second best for the C4WIT, but it is not even among the top 25 methods for C4WUS. Similar considerations apply for method b, which is the best in scenario CNR, but not in the top 25 methods for C4WIT and C4WUS.

5.3. Experimental results

Table 4 shows the accuracy performance of the three *reference* ensemble methods, each specialised for a different scenario; on the diagonal we show the median localisation error of ensemble CNR, ensemble C4WIT and ensemble C4WUS applied respectively to CNR, C4WIT and C4WUS scenarios. As expected, the results shown on the diagonal are not worse than the top method for each scenario that are listed in table 2, which confirms the effectiveness of the ensemble approach in terms of accuracy performance. For example, the error of ensemble CNR applied to scenario CNR is 4.3 m, while the best method in scenario CNR has error 5.1 m, and similarly for C4WIT and C4WUS.

Table 4: Median errors for the 4 ensembles in the 3 scenarios

	Ensemble	Scenario CNR	Scenario C4WIT	Scenario C4WUS
Reference ensembles	CNR	4.3	3.7	5.6
	C4WIT	5.3	2.9	7.2
	C4WUS	5.6	3.9	4.2
Ensemble of choice	**Intersect**	**5.5**	**3.7**	**4.7**

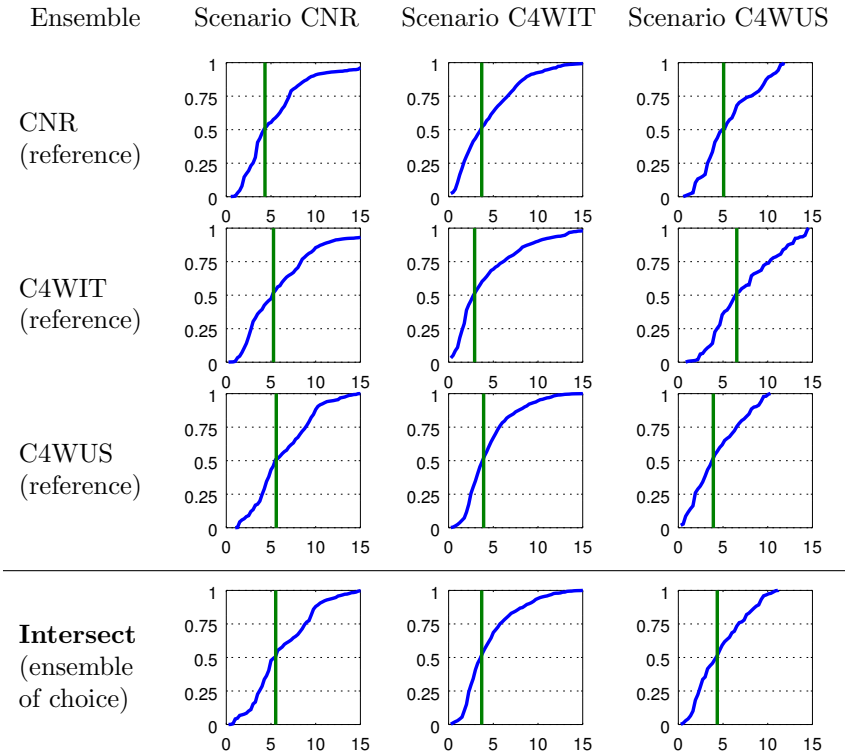


Figure 6: Cumulative density distribution of errors for the four ensembles and the three scenarios.

However, when we apply the reference ensembles to scenarios in which they are not specialised, performance drops significantly. Taking scenario CNR as an example, the error grows from 4.3 m when using the specialised ensemble CNR to 5.3 m and 5.6 m when using the other reference ensembles, as shown in table 4. We take this as indication that the reference ensembles are not robust across scenarios.

We finally analyse the performance of the ensemble of choice, the *intersection ensemble*, which is built with the purpose of being robust across scenarios, that is, of giving reasonably good results in all scenarios. The *intersection ensemble* is the intersection of the three sets of the 25 best-performing methods in each scenario. Its member methods are marked with upper-case letters A–D in table 2. The last row of table 4 shows the performance of the intersection ensemble applied to the three scenarios CNR, C4WIT and C4WUS. We observe that, as expected, while the results of the ensemble of choice (in bold) are worse than those of each reference ensemble for its own specialised scenario (underlined), they are generally good overall.

Moreover, and most importantly, the performance of the ensemble of choice can be considered satisfactory for the intended purpose of this work, meaning that it is indeed feasible to use the experimented strategy for crowd localisation. In fact, median errors ranging from 3.7 m to 5.5 m are acceptable for crowded areas such as a shop inside a mall, the space in front of a shop window, a waiting room, a bathroom area, a reception desk.

A more detailed overview of the numeric results in table 4 is given in figure 6, where the cumulative density distribution of error is depicted for all ensembles applied to all scenarios.

Results are consistent with the characteristics of the three scenarios: as expected, accuracy is higher for C4WIT. This can be explained by looking at 3: RSS varies a lot between different areas in the C4WIT map, while the picture of RSS in the other two scenarios is more homogeneous. In other words, we have more information to exploit in C4WIT than in the other scenarios, and this is reflected in a higher accuracy for C4WIT.

Another interesting observation is that the accuracy performance we observe is not so far from the state of the art in Wi-Fi indoor localisation. While a direct comparison is not possible, because we work with the little data provided by devices occasionally sending probes in small environments with a low number of FogSenses, it is interesting to note that during the EvAAL-ETRI competition at IPIN 2015 [9] one of the tracks was dedicated to off-line indoor localisation done exclusively with Wi-Fi information. The results obtained by competitors vary from a median of 4.6 m (the winner) to a median of 7 m. These results benefit from tracking techniques for error reduction, which cannot be used in our case, where tracking is impossible. We take these numbers as a hint that the methods proposed in this paper are indeed promising, since the figures measured during IPIN competitions are taken in controlled and scientifically accurate conditions, rather than by the system authors themselves in their own laboratories.

6. Conclusion

To the best of our knowledge, the literature lacks experimental investigations on using passive detection of Wi-Fi probes produced by mobile devices for indoor localisation purposes. We set up an experimental environment and we measure the performance of localisation methods that require neither installation measurements nor maintenance. Results are shown to be robust with respect to three indoor settings that exhibit quite different characteristics.

We derive some key takeaways as well as some considerations from this experimental campaign. First, the architecture we proposed is easily scalable, in the sense that in case the already-deployed APS are not enough to get satisfying positioning accuracy, it is possible to deploy additional sniffers, without any system reconfiguration. Second, our approach is survey-free, meaning that it does not require the usual configuration work needed for Wi-Fi indoor localisation systems, that is to survey the environment, to select the points where to gather the RSS values and finally to collect data with one or more sensing devices. We avoid all these steps by exploiting the probes sent by the APs and the possible additional sniffer themselves.

The results obtained with the described ensemble estimator are, in our opinion, remarkable. In fact, even if tracking cannot help with error reduction, the median errors of the intersection ensemble are directly comparable with the results of some of the best localisation algorithms based on pure Wi-Fi fingerprint, such as those that were presented and, most importantly, independently tested, during the EvAAL-ETRI 2015 competition.

Exploiting Wi-Fi probes promises to be a viable and cheap strategy for indoor localisation of devices. The method we describe can be the main building block of systems that sample the presence of people in a given area, a task that we call crowd localisation. Future work will need to build and experiment with such systems in real-life environments.

Appendix A. Measurements

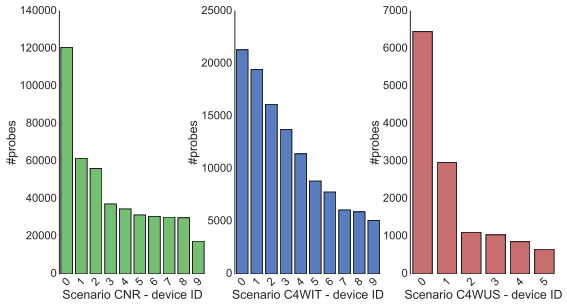


Figure A.7: Top 10 known devices by number of probes.

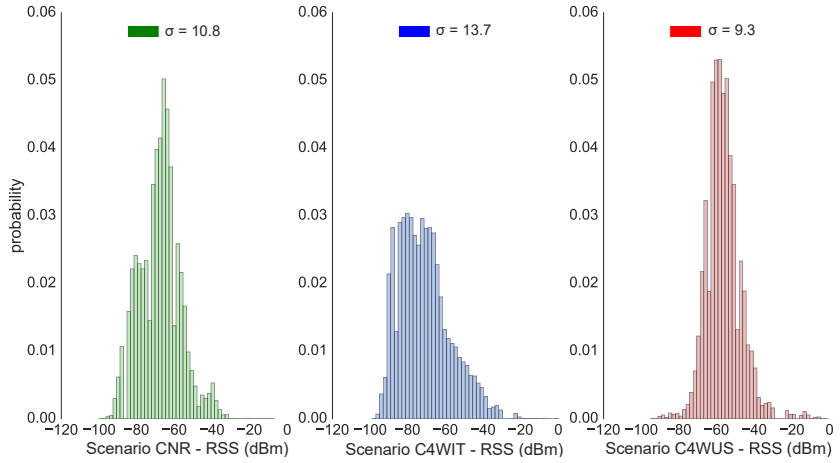


Figure A.8: Probability distribution of RSS values of known devices.

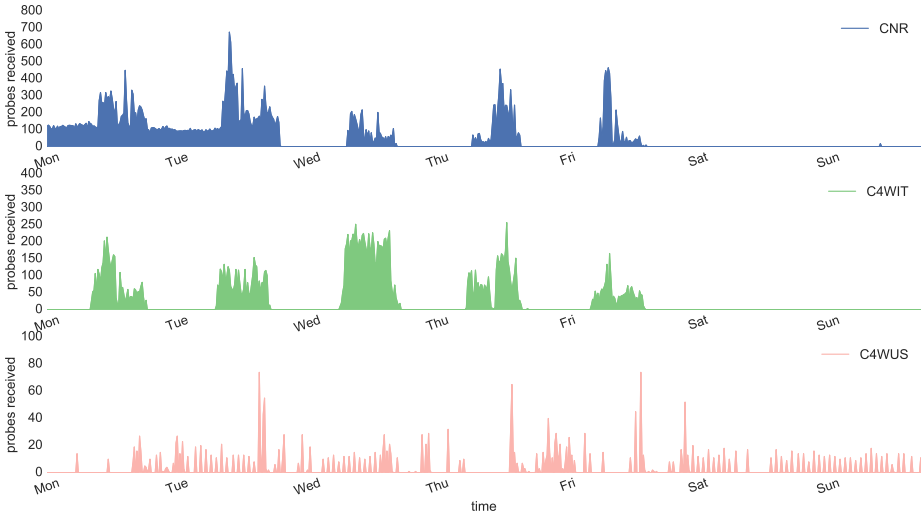


Figure A.9: Time series of captured probes in a week's time, 25-minute intervals.

Figure A.7 shows number of probes gathered by the most talkative known devices. Note how the number of probes produced can vary considerably between devices, as already discussed. We account for this difference in number of collected probes when measuring performance, in order to avoid weighting one device more than others.

Figure A.8 illustrates the RSS distribution for the known devices. The three distribution have different width, as highlighted by their standard deviation (shown in the figure). This is consistent with our previous observations on the difference of the three scenarios and is one more confirmation that C4WIT is the scenario where the FogSenses gather the most information.

Future work may investigate the usage of this information to assist the deployment in different environments, especially for deciding whether the already-installed APs are sufficient as sniffing devices to gather probe requests. In principle, adding FogSenses in the area to improve localisation accuracy could make sense unless this addition makes the standard deviation too narrow.

Finally, figure A.9 shows the number of probes received in 25-minute intervals as time series covering one week. It is apparent that, in CNR and C4WIT, the number of captured probes increases during the working hours and it drops down during off-work hours and weekends. In fact, the known devices are laptops and smartphones owned by employees at the three locations, the probes they emit well reproduce their working rhythms. At C4WUS such pattern is less clear for two reasons. First, most of the known devices are static and always connected to a local stable Wi-Fi network, which reduces the number of probes sent. Second, they are not owned by the employees, and are therefore working also during off-hours and weekends.

References

[1] P. Bahl and V. N. Padmanabhan, *RADAR: an in-building RF-based user location and tracking system*, Ieee conference on computer communications (infocom), March 2000, pp. 775–784.

[2] Pedro E. Lopez de Teruel, Oscar Canovas, and Felix J. Garcia, *Visualization of clusters for indoor positioning based on t-SNE*, Indoor positioning and indoor navigation (ipin), 2016 international conference on, 2016.

[3] Julien Freudiger, *How talkative is your mobile device?: An experimental study of Wi-Fi probe requests*, Proceedings of the 8th acm conference on security & privacy in wireless and mobile networks, 2015, pp. 8:1–8:6.

[4] Taisei Hayashi, Daisuke Taniuchi, Joseph Korpela, and Takuya Maekawa, *Spatio-temporal adaptive indoor positioning using an ensemble approach*, Pervasive and Mobile Computing (2016), –.

[5] Shau-Shiun Jan, Shuo-Ju Yeh, and Ya-Wen Liu, *Received signal strength database interpolation by Kriging for a Wi-Fi indoor positioning system*, Sensors **15** (2015), no. 9, 21377.

[6] B. Li, Y. Wang, H. K. Lee, A. Dempster, and C. Rizos, *Method for yielding a database of location fingerprints in WLAN*, IEE Proceedings - Communications **152** (2005Oct), no. 5, 580–586.

[7] A. B. M. Musa and Jakob Eriksson, *Tracking unmodified smartphones using Wi-Fi monitors*, Proceedings of the 10th acm conference on embedded network sensor systems, 2012, pp. 281–294.

[8] F. Potortì, A. Crivello, M. Girolami, E. Traficante, and P. Barsocchi, *Wi-Fi probes as digital crumbs for crowd localisation*, 2016 international conference on indoor positioning and indoor navigation (ipin), 2016Oct, pp. 1–8.

[9] Francesco Potortì, Paolo Barsocchi, Michele Girolami, Joaquín Torres-Sospedra, and Raúl Montoliu, *Evaluating indoor localization solutions in large environments through competitive benchmarking: The EvAAL-ETRI competition*, Indoor positioning and indoor navigation (ipin), 2015 international conference on, 2015, pp. 1–10.

[10] Francesco Potortì, Sangjoon Park, Antonio Ramón Jiménez Ruiz, Paolo Barsocchi, Michele Girolami, Antonino Crivello, So Yeon Lee, Jae Hyun Lim, Joaquín Torres-Sospedra, Fernando Seco, Raul Montoliu, Germán Martin Mendoza-Silva, Maria Del Carmen Pérez Rubio, Cristina Losada-Gutiérrez, Felipe Espinosa, and Javier Macias-Guarasa, *Comparing the performance of indoor localization systems through the evaal framework*, October 2017.

[11] J. Torres-Sospedra, G. M. Mendoza-Silva, R. Montoliu, O. Belmonte, F. Benitez, and J. Huerta, *Ensembles of indoor positioning systems based on fingerprinting: Simplifying parameter selection and obtaining robust systems*, 2016 international conference on indoor positioning and indoor navigation (ipin), 2016Oct, pp. 1–8.

[12] Y. Wen, X. Tian, X. Wang, and S. Lu, *Fundamental limits of RSS fingerprinting based indoor localization*, Ieee conference on computer communications (infocom), April 2015, pp. 2479–2487.

[13] Z. Xu, K. Sandrasegaran, X. Kong, X. Zhu, B. Hu, J. Zhao, and C. Lin, *Pedestrain monitoring system using Wi-Fi technology and RSSI based localization*, International Journal of Wireless and Mobile Networks **5** (2013January), no. 4, 17–34.