Binary Fingerprinting-Based Positioning Systems with Uncertainty Regions

Miguel Omeñaca Muro[1], Marouan Mizmizi[2], and Luca Reggiani[2]

[1] Universidad Politécnica de Madrid, Madrid, Spain, m.omenaca@alumnos.upm.es
[2] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, 20133 Italy, {marouan.mizmizi,luca.reggiani}@polimi.it

Abstract. A great variety of Indoor Positioning Systems (IPSs) have been developed with different technologies and algorithms. In the context of fingerprinting (FP) applications, the reduction of quantization levels in the Received Signal Strength Indicator (RSSI) till to its binary representation has been recently proposed for addressing one of the common drawbacks of FP, i.e. the large data size and consequently the large search space and computational load.
In this paper, the binary fingerprinting technique is completed by the use of guard or uncertainty regions on the edges between the zones in which the area of interest is partitioned. This simple solution is shown to be effective for recovering the performance gap between binary fingerprinting and fingerprinting with full complexity. In addition, we introduce also an alternative design of the cells in the area of interest in order to address better the errors due to fluctuations of the RSSI measure in a real deployment.

Keywords: Fingerprinting, Indoor Positioning System, Localization, Binary Fingerprinting.

1 Introduction

For more than a decade we have been witnessing a steady growth of mobile terminals and Internet of Things (IoT) systems. This expansion has made the role of location information more and more central in the digital age. Position information, in terms of coordinates, is an added value to many services and it has even enabled new Location Based Services (LBSs).

The Fingerprinting (FP) technique [1, 2] takes advantage of Location Dependent (LD) features of the received signals, exploited as unique signatures associated to the target locations. Firstly, a radio map containing stored LD parameters measured over predetermined points (grid points) is built during an off-line or training phase and then target position is estimated via pattern matching between measured LD parameters and those previously recorded. The search space during the online step can be computationally burdensome, either because the deployment area is wide (e.g. smart city, hospitals or, large factories) or because it is based on devices such as Bluetooth Low Energy (BLE) tags with strong limitation on power consumption and limited hardware capabilities. To overcome this restraint, the works in [3, 4] and reference therein, have proposed clustering and spatial filtering techniques that limit the positioning algorithm to a subset of reference points (RPs) in order to narrow down the search space and focus on the relevant subset of RPs. As a consequence, the resulting performance suffers from a reduced map resolution and therefore a loss in position accuracy.

In [5], the role of quantization in the RSSI-based FB has been studied and the numerical results shows that the computational complexity can be limited by adapting the RSSI quantization w.r.t. the variance of the measured RSSI error at the target. In [6], we have focused on the simplest quantization, with two levels, exploring its relations with binary encoding and presenting a design of a specific binary representation of the RSSI signatures and measures (binary fingerprinting or BFP). This novel design is appropriate when the beacons are characterized by very limited size, cost and computational capability, like in the BLE or in the future technologies such as in pervasive Internet of Things or molecular communication.

In the context of low complexity FP systems, this paper considers the design presented in [6] with the aim of improving some of the weaknesses of the system, in particular, the performance degradation w.r.t. RSSI representation at the edge of two or more cells. The main contributions are the following:

- the use of guard spaces on the edges of the sub-areas labeled by the binary fingerprinting (cells) in order to address performance loss in the zones where the error rate is higher;
- a novel online positioning algorithm that takes into account the measures inside these guard spaces and, more generally, for treating ambiguous measurements, *i.e* two or more labels having the same hamming distance to the measurement;
- the definition of an additional cell design step, in which the actual shapes of the FP cells are redefined to fit better the real RSSI measures in the area.

The remainder of this work is organized as follow: Sect. 2 describes the network scenario and, in Sect. 3, the binary version of FP is briefly revised. Then Sect. 4 introduces the ternary representation with the guard spaces and Sect. 5 the optimized cell design. Finally, Sect. 6 is dedicated to the numerical results and their analysis.

2 System model

We consider an asynchronous sensor network containing a number of target devices over a limited rectangular area on a single floor, with two-dimensional coordinates (x, y).

2.1 Layout

In the area, there are N_B fixed nodes called beacons with known positions $\mathbf{P}_i^{BS} = \{x_i, y_i\}, \{i = 1, 2, ..., N_B\}$. A uniform grid is defined over the two-dimensional area and any estimate of a target location is limited to the points on this grid. The grid of points has resolution δ, over which the RSSI measures are stored for recovering the target locations.

Assuming that grid spacing results in K_x points along the x coordinate and K_y along the y coordinate, we have $K_T = K_x \times K_y$ positions in the area. Any position can be represented by a 2-tuple with label (x, y) where x and y represent the $2D$ coordinates on the floor plane. Therefore, the coordinates $(x, y)_i$ denote the *i-th* FP signature location.

2.2 Signal

The most common model adopted for the real RSSIs, recorded and stored during the off-line phase, responds to log-normal shadowing, i.e.

$$RSSI(d) = A - L_0 - 10\alpha log_{10}(d/d_0) + L_{SH} \quad (1)$$

where A is a constant, given by the transmitted power and the antenna gains, L_0 is the average propagation loss at the reference distance d_0 (usually 1 m), α is the Path Loss Exponent (PLE), d is the distance between transmitter and receiver and $L_{SH} \sim \mathcal{N}(0, \sigma_{SH}^2)$ is the log-normal fluctuation mainly due to obstacles in the environment. In this work, the values of the shadowing L_{SH} in a given area are assumed correlated according to the parameter d_{SH}, which separates independent shadowing samples in the space. Then, according to the relative position between a grid point and a beacon, the values of L_0 and α depend also on the Line of Sight (LoS) or Non LoS (NLoS) conditions. In fact, each RSSI measured by the target during the online phase is affected by an additional variance measure, which depends on the environment changes, device orientation and different device responses. The RSSI variance is probably the most relevant issue in RSSI-based IPSs and some countermeasures are suggested in [7] [8]. Different measurement campaigns in indoor environments report a typical standard deviation within $2 - 2.83$ dB [9]. From this perspective, the online RSSI measurements are modeled as affected by an additional random log-normal component, uncorrelated to the channel shadowing component in (1) and we assume, again in dB,

$$RSSI_{MEAS}(d) = RSSI(d) + W, \quad (2)$$

where $W \sim \mathcal{N}(0, \sigma_W^2)$.

3 Overview of binary fingerprinting

The key point of BFP is the reduction of the quantization levels of the RSSI up to the binary state. Considering beacons with low cost and small size, N_B can be large and σ_W^2 higher. As discussed in [6], BFP allows us to consider the system from a different point of view, related to a *binary code interpretation* of the fingerprinting scheme: the $log_2(K_T)$ bits that enumerate the K_T grid positions and the corresponding fingerprinting signatures are now transformed, or *encoded*, in the $N_B \cdot log_2(L)$ bits of each signature, where L is the number of levels used for each RSSI measure. The resulting code rate would be defined as

$$R = N_B \cdot \frac{log_2(L)}{log_2(K_T)} = \frac{N_B}{log_2(K_T)} \quad if \quad L = 2. \quad (3)$$

Through the quantization of the RSSI with one bit, the RSSI will have 2 levels and the beacons will have a coverage obviously divided into two zones, according to a threshold $RSSI_{REF}$, as in

$$r_i = \begin{cases} 1 & \text{if } RSSI_{measured} \geq RSSI_{REF} \\ 0 & Otherwise. \end{cases} \quad (4)$$

Therefore, the binary design of quantized fingerprinting can be driven by concepts taken from error correcting codes theory an the optimal schemes for 2×2 and 4×4 cells have been presented in [5, 6].

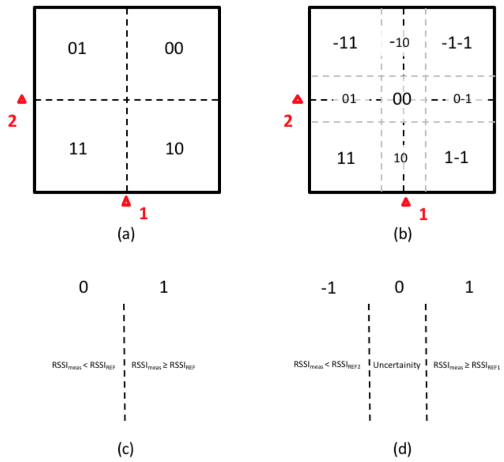

Fig. 1: Binary labeling in (a,c) vs Ternary labeling in (b,d)

4 Uncertainty Region

Performance of BFP is competitive with traditional FP techniques but there is an accuracy degradation that occurs when the target approaches the boundaries of a cell. This loss is due to the uncertainty in the bit assigned because of the hard threshold on the RSSI.

To deal with this weakness, in this paper, we propose an additional region denominated *uncertainty region* around the boundaries separating the cells (as shown in Fig. 1). In the proposed approach, whenever an on-line measurement falls in the uncertainty region, the algorithm assigns a *neutral* value to differentiate from the other two cases. In the interpretation of BFP as an error correcting code [6], this neutral value corresponds to the concept of *erasure*.

The novel ternary labeling is now defined, for the measure coming from the i-th beacon, as

$$r_i = \begin{cases} 1 & \text{if } RSSI_{measured} \geq RSSI_{REF,i} + \Delta/2 \\ -1 & \text{if } RSSI_{measured} < RSSI_{REF,i} - \Delta/2 \\ 0 & \text{Otherwise} \end{cases} \tag{5}$$

where Δ is the size of the uncertainty region, in dB, around the threshold $RSSI_{REF,i}$ and the neutral value is conventionally identified by the 0. In Fig. 1, the differences between binary labeling and ternary labeling are shown.

The target position is estimated as follow

- in presence of neutral values, the corresponding positions in the binary label would not be considered in the search of the candidates in the binary radio map.
- When there are N_m multiple Hamming distances (this event clearly becomes more likely in presence of one or more neutral values), the position estimated is computed as

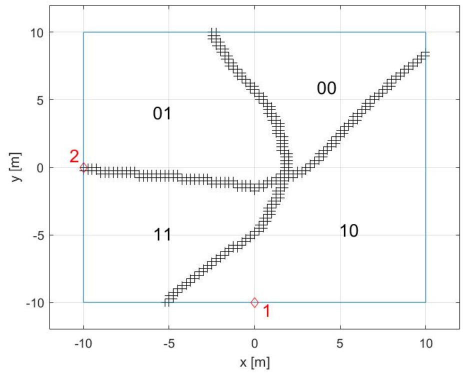

Fig. 2: Binary labeling 2×2 for a typical, realistic RSSI scenario.

$$P_{Target} = \frac{1}{N_m} \sum_{i \in \mathcal{A}_m} C_i = \frac{1}{N_m} \sum_{i \in \mathcal{A}_m} (x_i, y_i) \tag{6}$$

where $\mathcal{A}_m$ is the set of FP signatures with minimum Hamming distance from the measured binary one and $C_i = (x_i, y_i)$ are the coordinates of the corresponding cells centers.

5 Cells design

Another weakness in the spatial labeling proposed and analyzed in [5,6] regards the relation between the ideal lines that partition the space, i.e. the square cells in Fig. 1a, and the real lines determined by the field measure used for the FP, i.e. the RSSI values in this case. As shown in Fig. 2, a real environment will return an RSSI map that is affected by the real propagation environment in terms of pathloss, shadowing and fast fading; on the other hand, even in presence of ideal pathloss, the straight lines of the squared cells cannot correspond to lines with equivalent RSSI since the ideal pathloss depends just on the distance and the real lines would be circular. This misalignment between the lines or borders in the design and the real ones increases the overall rate error, in terms of estimated cell and estimated position of the target. The procedure followed for deriving the threshold $RSSI_{REF,i}$ that divides the area for each beacon i is simply given by these two steps:

- measure the RSSI values along the border that divides the area according to the value of the bit assigned to each beacon i;
- compute $RSSI_{REF,i}$ as the average (in [dB] here) of the values collected on this border.

If we apply the $RSSI_{REF,i}$ in the area, the actual partitioning of the space according to (4) will be different from the ideal one, as in the example reported in Fig. 2. Therefore, the final performance incorporates the error induced by the misalignment between the ideal and the real binary map of the area. In this paper, we will call this design strategy, used also in [5, 6] *Forced Cell Design*

Table 1: System model parameters.

Parameter	Value
Cells Layout	$K_T = 2 \times 2,\ 4 \times 4$
Grid Side L	20 m
Beacons number N_B	2, 6
Attenuation constant A	0 dBm
Propagation Loss L_0	-35 dB
Path Loss Exponent α	2
Reference distance d_0	1 m
Log-Normal Shadowing σ_{SH}	4 dB
Correlation Distance d_{SH}	5 m
Measurements Noise σ_W	$\{0.1, \cdots, 10\}$ dB

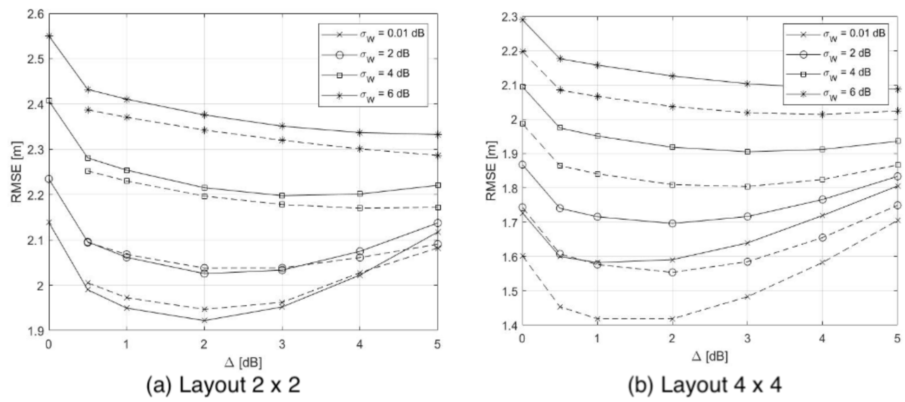

Fig. 3: RMSE in the FCD (continuous lines) and OCD (dashed lines) as function of the uncertainty region size Δ

(FCD) and we will consider also an alternative, referred to as *Optimized Cell Design* (OCD), in which the new real borders are used to redefine the cells and their centers, to be used in the online stage of the BFP. The real border are derived from measurements in the considered area, for example by interpolation of the collected radio map. This modification does not increase the computational load since it changes just the centers of the cells used for estimating the final target position by means of the algorithm, including the application of (6), after the reduction of the fingerprint to the binary level according to (4) or (5).

6 Numerical results

In this section, numerical evaluation of the proposed technique are carried out using MATLAB simulations. The performance results are compared between Nearest-Neighbors (NN) based FP and BFP techniques with the use of the uncertainty region and for the two design strategies, FCD or OCD. The system parameters are summarized in Table 1, unless otherwise stated. The performance is measured in a uniform grid of points with resolution 0.25 m. In the 2×2 cells

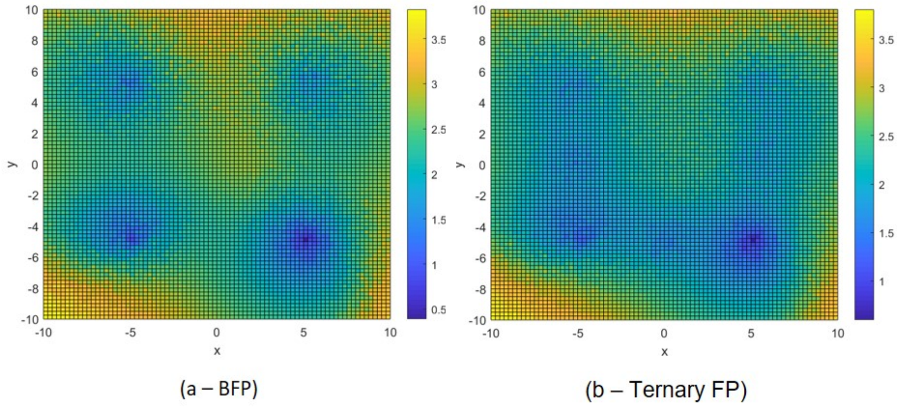

Fig. 4: Mean Error [m] as a function of the position for the 2×2 FCD layout, with $\Delta = 0$ dB (a - BFP) and $\Delta = 2$ dB (b - Ternary), $\sigma_W = 4$ dB.

layout, Fig. 3a shows how the Root Mean Square Error (RMSE) changes as a function of the uncertainty region size Δ in the FCD and OCD cases; the best value of Δ becomes larger for high values of the measures error ($\sigma_W \geq 4$ dB), where aso OCD and FCD show the highest performance gap. For the sake of completeness, we report also the spatial distribution of the error for the OCD case, corresponding to Fig. 2: this picture clearly shows that the border regions are the main sources of error. It is also necessary to notice that the error for conventional FP with infinite number of bits in the measure is close to the values of BFP with $\Delta = 0$ since with just 2 beacons the gain effect guaranteed by the Euclidean distance, instead of the Hamming one of the BFP, is negligible. More interestingly, moving to the 4×4 cells layout as described in [6], in Fig.

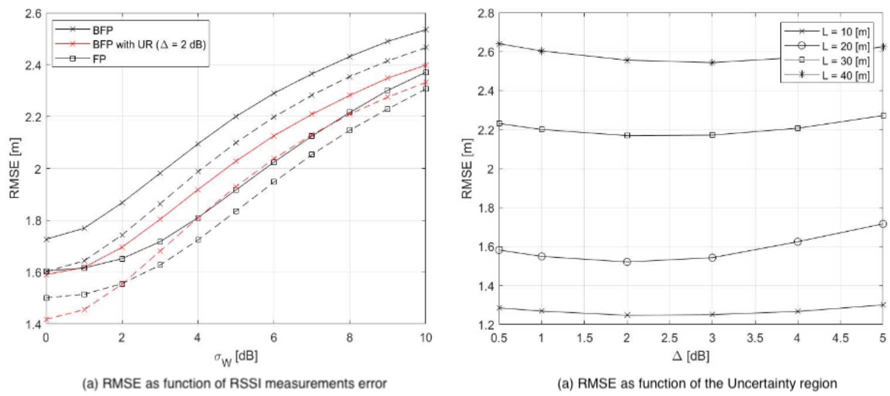

Fig. 5: RMSE for 4×4 layout in the FCD (continuous lines) and OCD (dashed lines)

5a we can show the dependance of the performance also w.r.t. the measurement error parameter σ_W; the ternary option with $\Delta = 2$ dB performs better than the original BFP and recovers most of the performance gap with FP at any value of σ_W. The optimal size of the uncertainty region around $2-4$ dB can be seen in Fig. 3b with increasing values as σ_W increases as can be expected. Finally Fig. 5b shows how the RMSE varies with the side of the square area as a function of δ.

7 Conclusions

In the last decade, the increase of demand of location based services has generated a strong interest in the field of positioning systems. Designing an IPS is a difficult task, which is generally faced by combining different technologies and methods; for techniques such as fingerprinting, there are still some issues to be addressed from a theoretical and practical point of view. This paper enhances several aspects of the Binary FP algorithm, in particular including an uncertainty region to cope with the fluctuating RSSI values on the cell borders and the design of the cells in presence of realistic RSSI distributions. Simulation results show the benefit of the proposed modifications compared to BFP without an increase of the computational or memory load.

As a future work, we aim at validating the proposed solutions with experimental data, and to prove the benefits of the BFP technique with uncertainty regions in terms of computational efficiency against the state of the art.

References

1. H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. Syst., Man, Cybern.* , vol. 37, no. 6, pp. 1067–1080, Nov. 2007.
2. Y. Gu, A. Lo, and I. Niemegeers, ""A Survey of Indoor Positioning Systems for Wireless Personal Networks"," *IEEE Communications Surveys And Tutorials*, vol. 11, no. 2, pp. 13–31, 2009.
3. A. Arya, P. Godlewski, and P. Melle, "A hierarchical clustering technique for radio map compression in location fingerprinting systems," in *2010 IEEE 71st Vehicular Technology Conference*, May 2010, pp. 1–5.
4. A. Saha and P. Sadhukhan, "A novel clustering strategy for fingerprinting-based localization system to reduce the searching time," in *2015 IEEE 2nd International Conference ReTIS*, July 2015, pp. 538–543.
5. M. Mizmizi and L. Reggiani, "Design of rssi based fingerprinting with reduced quantization measures," in *2016 International Conference IPIN*, Oct 2016, pp. 1–6.
6. ——, "Binary fingerprinting-based indoor positioning systems," in *2017 International Conference on IPIN*, Sep. 2017, pp. 1–6.
7. A. Tsui, Y. Chuang, and H. Chu, ""Unsupervised learning for solving RSS hardware variance problem in wifi localization"," *Mobile Networks and Applications*, 2009, vol.14, pag. 677–691.
8. Y. Kim, H. Shin, Y. Chon, and H. Cha, ""Smartphone-based Wi-Fi tracking system exploiting the RSS peak to overcome the RSS variance problem"," *Pervasive and Mobile Computing 9, ELSEVIER*, 2013, pag. 406–420.
9. S. Pagano, S. Peirani, and M. Valle, "Indoor ranging and localisation algorithm based on received signal strength indicator using statistic parameters for wireless sensor networks," *IET Wireless Sensor Systems*, vol. 5, no. 5, pp. 243–249, 2015.