CrossMark

CSI-based fingerprinting for indoor localization using LTE Signals

Giovanni Pecoraro[*] (ID), Simone Di Domenico, Ernestina Cianca and Mauro De Sanctis

Abstract

This paper addresses the use of channel state information (CSI) for Long Term Evolution (LTE) signal fingerprinting localization. In particular, the paper proposes a novel CSI-based signal fingerprinting approach, where fingerprints are descriptors of the "shape" of the channel frequency response (CFR) calculated on CSI vectors, rather than direct CSI vectors. Experiments have been carried out to prove the feasibility and the effectiveness of the proposed method and to study the impact on the localization performance of (i) the bandwidth of the available LTE signal and (ii) the availability of more LTE signals transmitted by different eNodeB (cell diversity). Comparisons with other signal fingerprinting approaches, such as the ones based on received signal strength indicator or reference signal received power, clearly show that using LTE CSI, and in particular, descriptors as fingerprints, can bring relevant performance improvement.

Keywords: CSI, Localization, Fingerprinting, LTE

1 Introduction

The range of applications requiring ubiquitous (indoor and outdoor) high-accuracy localization is rapidly increasing, and it is well known that the accuracy and availability of the Global Navigation Satellite System (GNSS), which remains the most common positioning technology, drop in indoor environments and urban canyons. This aspect has motivated extensive work on alternative localization solutions [1], which are based on radio signals that are either transmitted by dedicated sensors or by opportunistic transmitters (e.g., WiFi routers or towers of a cellular system). In both cases, fingerprinting may play an important role [2] in meeting the high accuracy and low complexity requirements of indoor positioning [3].

The basic idea of a signal fingerprinting approach is to find the location of a mobile device by comparing its signal pattern received from multiple transmitters, such as WiFi Access Point (AP), cellular Base Station (BS) or dedicated transceivers, to a predefined database of signal patterns. Most of the proposed signal fingerprinting approaches for indoor positioning use WiFi signals [3, 4], which are widely available indoor without deploying a dedicated

infrastructure. In most of the mentioned works, Received Signal Strength Indicator (RSSI) is used as a fingerprint due to its simplicity and low complexity requirements. However, the use of Channel State Information (CSI), which can be easily made available by commodity hardware, is attracting more and more interest in the context of WiFi [3, 5, 6]. Radio signals transmitted by BSs of a cellular system would have the advantage of a wider availability (also indoor), and in any case, they could complement the coverage of WiFi APs. Measurements on the radio signal transmitted by the BSs of a cellular system have been considered for localization, but mainly for outdoor applications, using Global System for Mobile Communications (GSM) [7] or Universal Mobile Telecommunications System (UMTS) [8] and considering the RSSI as fingerprint. Few recent works have considered Long Term Evolution (LTE) signal for radio fingerprinting localization [9–11]. However, in all the mentioned works, signal strength measurements are used as fingerprints.

This work aims to provide better insights on the use of LTE signal for radio frequency (RF) fingerprinting and in particular in the use of measurements that are not only related to the signal strength but also to a finer-grained knowledge, at subcarrier level, of the channel gain, such as the one provided by the CSI. As a matter of fact, the term CSI usually refers to WiFi and indicates the vectors of

---
*Correspondence: giovanni.pecoraro@uniroma2.it
Department of Electronic Engineering, University of Rome "Tor Vergata", Via del Politecnico, 1, 00133 Rome, Italy

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 2 of 18

channel gains per subcarrier that can be extracted by commodity hardware. In this work, more generally, we call CSI a vector of channel gains per subcarrier that represents an estimate of the channel frequency response of the propagation channel. Therefore, this paper proposes an LTE signal fingerprinting localization method that uses CSI as fingerprint. Moreover, the proposed approach is different from other CSI-related approaches that can be found in the literature, mainly employed for indoor localization and based on WiFi signals, where fingerprints are vectors containing the values of measured CSI. We propose to use as fingerprints not only the vectors of CSI but also some "descriptors" of the "shape" of the CSI calculated on these vectors. This would greatly reduce the requirements in terms of memory for the database and also the computational complexity of the matching phase. It is also worth outlining that the proposed method extracts the CSI from signalling messages meaning that the mobile device does not need to have any subscription with any mobile operator. As a matter of fact, the device can receive the signalling messages of different eNodeBs regardless the specific operator.

Experimental results, in rooms of different size and different furniture, prove the feasibility of the proposed approach, which has the further advantage of reducing the computational complexity of the matching phase and the fingerprint database size. Moreover, the paper shows how the use of CSI measurements can improve the performance with respect to signal strength measurements in fingerprinting localization using LTE signals. Being one of the first works that investigates the use of LTE signal and CSI measurements for indoor localization, the paper also studies the impact of the available bandwidth and the possibility to get cell diversity gain by using LTE signals transmitted by different eNodeBs.

The paper is organized as follows: Section 2 provides an overview about the state-of-the-art on RF fingerprinting techniques; Section 3 describes the theoretical background on LTE and signal fingerprinting; Section 4 presents the proposed localization method; and Sections 5 and 6 show the experimental setup and results, while the conclusions are drawn in Section 7.

2 Related works

The literature on wireless localization based on RF fingerprinting is wide; hence, we will only discuss the major results on the topic. Several wireless systems have been used to apply the concept of RF fingerprinting for localization, including the most common WiFi networks [12], but also RFID technologies [13], GSM/UMTS/LTE, and LoRa. A combination of wireless technologies can be used to improve localization accuracy [14].

Fingerprinting localization is an indirect localization method which is based on the collection of RF received

signal characteristics using a two-step process: an offline training phase which builds a database of measurements for each known position and an online matching/classification phase which estimates the current position using similarity scores between the received signal characteristics and the measurements stored in the training database. Fingerprinting methods are prone to environmental changes of the area of interest; hence, the offline training phase has to be carried out periodically. Fingerprinting uses measurements made by the access network (e.g., the base stations), the user equipment (UE), or a combination of both. In this work, we focus on RF measurements made by the UE (i.e., using downlink signals). In principle, features extracted from the received signals at each location should differ from all the others to avoid ambiguity.

RF fingerprinting can be carried out using received signal features, in particular signal strength, such as RSSI, Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), and Channel Frequency Response (CFR)/CSI or equivalently the Channel Impulse Response (CIR) [3, 5, 15, 16]. Time of Arrival (TOA), Timing Advancing (TA), Time Difference of Arrival (TDOA), and Angle of Arrival (AOA) are less frequently used in a fingerprinting-based approaches [17, 18].

Most of the fingerprinting localization systems are based on the use of signal strength measures. Methods based on signal strength have two disadvantages: (1) signal strength measures have a high variability over time for a fixed location, due to the multipath effects; (2) signal strength measures are coarse information which do not exploit all the available information on the wireless channel. However, the multipath effect can even be very constructive when employed in fingerprinting localization systems. In particular, CSI-based fingerprinting exploits the uniqueness of the multipath delay profile of the wireless channel at each receiver location [6, 16].

In DeepFi [3], CSI information is collected through a dedicated device driver and analyzed with a deep network with four hidden layers. During the training phase, CSI values collected at each location are utilized to learn features, which are then stored as fingerprints. In the test phase, online data is used to match the closest spot with the similar features stored in the training phase. The distance between two spots was set to 50 cm, which can maintain the balance between localization accuracy and preprocess cost. The achieved mean error ranges from 1.8 to 1.9 m.

For what concerns an LTE network, a UE must detect and monitor the signals transmitted by multiple cells and perform cell reselection. Cell mobility procedures and physical layer adaptation schemes are based on the following measurements on the signal received from each eNodeB: RSRP, RSSI, RSRQ, and CSI.

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 3 of 18

A fingerprinting localization method in LTE cellular systems is presented in [10]. In this work, the altitude is added to the adaptive enhanced cell identity (AECID) using observed time difference of arrival (OTDOA) or uplink-time difference of arrival (U-TDOA) measurements.

The authors of [19] presented a multi-radio positioning concept which can be used to enhance the positioning performance in the WCDMA and LTE networks.

Several other works on localization in LTE networks exploit TOA or TDOA measurements [20–22]. The following works are based on signal strength measurements.

In [23], LTE measurements at the UE side such as RSRP, RSRQ, and RSSI have been used for indoor localization. Three Samsung Galaxy S5 phones running an application were used to gather data from around 60 different spots in two indoor environments, where a location represents a 2×2-m spot within the indoor space. The achieved average prediction accuracy ranges between 59 and 84% using only LTE signals. However, it was demonstrated that the combination of 2G and 4G signals increases the prediction accuracy.

In [24], a feature extraction algorithm is applied to select channel parameters with non-redundant information that are calculated from the LTE downlink signals. A feedforward neural network with the input of fingerprint vectors and the output of UEs' known locations is trained and used by UEs to estimate their positions. Preliminary experimental results obtained through signal measurements in LTE networks demonstrated that the proposed localization technique yielded median positioning error distances of 6 m for indoor and of 75 m for outdoor scenarios, respectively.

The feasibility of using TA information together with RSSI/RSRP for improving position estimation in LTE networks was studied in [18]. Performance results show that RF fingerprint containing both RSRP and TA measurements provides a higher positioning accuracy with respect to the fingerprint containing only RSRP information.

RSSI pattern matching and dynamic time warping are exploited in [11] for localization in LTE and LoRa networks.

In the past research works, fingerprinting localization using CSI was carried out only in WiFi networks. In [25], it was preliminarily investigated the possibility to use the CSI extracted from LTE signals for signal fingerprinting localization. Our work represents the first study on the use of CSI for fingerprint-based localization in a LTE network.

3 Theoretical background
3.1 LTE CSI
In orthogonal frequency division multiplexing (OFDM) systems, CSI is a vector of channel gains per subcarrier which are estimated by the receiver and used for channel equalization. There has been an increased interest in using WiFi CSI both for localization and activity recognition, as it can be easily extracted using a customized firmware and an open source Linux wireless driver for the Intel 5300 WiFi card. In the case of LTE, with the term CSI, we still mean a vector of channel gains per subcarriers which represents an estimate of the channel frequency response. However, we need to specify how to extract these channel gains from the receiver. In the following, we consider only the LTE frequency division duplexing (FDD) mode, in which uplink and downlink channels are separated in frequency. In LTE, the information data is transmitted over a time-frequency grid. The time-frequency resources are organized as follows: *radio frame* has a duration of 10 ms and is the largest unit of time in the LTE resource grid. Each radio frame is subdivided into ten *subframes* having a duration of 1 ms, each of which is split into two *slots* of duration 0.5 ms. Each slot consists of six or seven OFDM symbols (depending on the cyclic prefix length). In the frequency domain, the finest granularity is provided by OFDM subcarriers which are spaced by 15 kHz from each other. The minimum resource unit is called resource element (RE) and consists of one OFDM subcarrier in the frequency domain and OFDM symbol in the time domain. The smallest unit of resource that can be allocated to a UE is called resource block (RB), which is a group of 12 contiguous OFDM subcarriers (180 kHz) in a time slot of seven (or six) OFDM symbols.

In the LTE standard, the concept of antenna port has been introduced. An antenna port is a generic term used for signal transmission under identical channel conditions and is defined for each independent channel in the downlink direction. LTE symbols that are transmitted via identical antenna ports are subjected to the same channel conditions, and thus, a UE must carry out a separate channel estimation for each antenna port through separate reference symbols. Any eNodeB serving the cell c, in order to exploit spatial diversity, can map these logical antenna ports to T_c physical transmitting antennas. Let us denote with $\mathbf{x}_c$ the vector of complex symbols transmitted over N subcarriers by the eNodeB and with $\mathbf{y}_c$ the complex vector received from one of the possible transmitting antennas of the eNodeB after the N-point Fast Fourier Transform (FFT) at the receiver. The latter can be written as:

$$\mathbf{y}_c = \mathbf{X}_c \mathbf{h}_c + \mathbf{w} \tag{1}$$

where $\mathbf{X}_c$ is the transmitted diagonal complex matrix, $\mathbf{h}_c$ is the vector containing the channel complex gains per subcarrier, and $\mathbf{w}$ is a complex white Gaussian random process representing the noise and the inter-cell interference. The channel gains are estimated by the receiver using the Cell-specific Reference Signal (CRS) inserted in specific OFDM symbols within every slot.

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 4 of 18

Figure 1 shows the CRSs in two consecutive RBs, and CRSs are represented by different colors corresponding to different antennas (four antennas are assumed). In one slot, there is a total of four CRSs per antenna, located over different subcarriers and OFDM symbols. In this work, we extract the complex channel gain per subcarrier for two different signal bandwidths, i.e., 1.4 and 10 MHz. For both cases, we assume that the channel stays rather stationary over a slot (0.5 ms), i.e., the coherence time is equal or greater than the time slot duration. Under this assumption, the CRSs in different positions in the same slot can be aligned at the same instant of time for each antenna, doubling the size of the CRS resulting vector. In the case of 1.4 MHz signal bandwidth, 6 RBs are used, which allow to extract from the received signal $N = 6 \times 2 \times 2 = 24$ (6 RBs $\times$ 2 CRSs $\times$ 2 positions) complex channel gains for 24 different OFDM subcarriers. While receiving the 10-MHz signal bandwidth, 50 RBs are employed for transmission; therefore, $N = 50 \times 2 \times 2 = 200$ (50 RBs $\times$ 2 CRSs $\times$ 2 positions) CRSs are used to estimate the complex channel gains for 200 OFDM subcarriers. The complex vector that represents the channel estimation is what we call CSI vector in the following. It is worth outlining that an LTE receiver can extract a CSI vector $\hat{\mathbf{h}}_{c,t}$ for each eNodeB that is received in a specific position and for each transmitting antenna port of that eNodeB:

$$\hat{\mathbf{h}}_{c,t} = \left[\hat{h}_{c,t}(0), ..., \hat{h}_{c,t}(i), ..., \hat{h}_{c,t}(N-1)\right], \quad \forall c, \forall t \in T_c \quad (2)$$

In this paper, we also consider, for comparison, other measurements of the LTE signal for fingerprinting purposes, such as RSRP and RSSI. The RSRP corresponds to the average power of REs that carry CRSs over the entire bandwidth, so RSRP for each transmitting antenna port can be calculated on the vectors $\hat{\mathbf{h}}_{c,t}$ as follows:

Fig. 1 LTE Cell-specific Reference Symbols. Cell-specific Reference Symbols for each antenna port in two consecutive Resource Blocks

$$\mathrm{RSRP}_{c,t} = \frac{1}{N} \sum_{i=0}^{N-1} \left|\hat{h}_{c,t}(i)\right|^2, \qquad \forall c, \forall t \in T_c \quad (3)$$

The RSSI, instead, is the measure of the average received power observed only in OFDM symbols containing CRSs for antenna port 0 (symbols 0 and 4 in a slot) over all the resource blocks in the measurement bandwidth. The measured RSSI value includes the power from co-channel serving and non-serving cells, adjacent channel interference, and the thermal noise:

$$\mathrm{RSSI}_c = \frac{1}{L} \sum_{i=0}^{L-1} |\mathbf{X}_c(i,i)|^2 \quad (4)$$

where L is the overall number of available subcarriers, that in case of 1.4 MHz signal bandwidth (six RBs) is equal to $L = 6 \times 12 = 72$, while for the 10 MHz case (50 RBs) is equal to $L = 50 \times 12 = 600$.

3.2 Signal fingerprint-based localization

Signal fingerprint-based localization techniques find the location of a device by comparing its signal pattern received from multiple transmitters (e.g., WiFi APs or cellular BSs) to a predefined database of signal patterns. There is a variety of signal fingerprint-based localization systems in the literature both for outdoor and indoor localizations. The "signatures" are usually extracted from the signal received from an AP or a BS of a cellular system. Moreover, in most of the previous works, the fingerprint is a tuple comprising the AP or BS identifiers and the corresponding measured RSSI values. In this work, we propose a signal fingerprinting localization method, which uses as fingerprint-specific features calculated on the CSI estimated by an LTE receiver. In general, a fingerprint-based localization consists of the following phases:

- Fingerprint database building phase—The purpose of this phase is to build up the offline fingerprint database, which stores for each reference point (RP) a fingerprint. The fingerprint is obtained by properly processing some measurements of the signal received by an AP or a BS. Let us denote with $\mathbf{RF}_r$ the reference fingerprint in the RP r.
- Fingerprint acquisition phase—For each test point (TP), whose position is unknown, the fingerprint is calculated using the same measurements on the received signal. Let us denote with $\mathbf{TF}$ the fingerprint in a TP.
- Fingerprint matching phase—This phase consists of associating to the fingerprint measured in the TP, the fingerprint stored in the built database which is closest to the measured one according to a predefined matching rule. The user location is then calculated as the location of the RP corresponding to the found fingerprint.

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 5 of 18
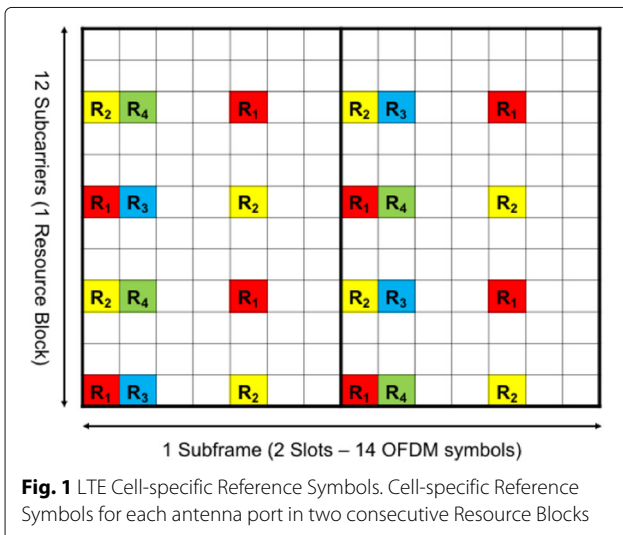
The RP that is associated to the TP is identified by minimizing a function of the fingerprints. In the case of nearest neighbor (NN) matching, the RP $\bar{r}$, that is associated to the TP with fingerprint **TF**, is the RP that minimizes a function called Fingerprint Distance (FD):

$$\bar{r} : \mathrm{FD}\left(\mathbf{RF}_{\bar{r}}, \mathbf{TF}\right) \leq \mathrm{FD}\left(\mathbf{RF}_r, \mathbf{TF}\right), \quad \forall r \neq \bar{r} \tag{5}$$

Then, the location (x, y) of the TP is calculated through the following association:

$$(x, y) \implies (x_{\bar{r}}, y_{\bar{r}}) \tag{6}$$

In case of K-nearest neighbor (KNN), the coordinates of the K best RPs are arithmetically averaged:

$$(x, y) = \frac{1}{K} \sum_{r=1}^{K} (x_r, y_r) \tag{7}$$

In case of weighted K-nearest neighbor (WKNN), the weighted average of the coordinates of RPs with the shortest distance to the TP are computed, where the weights are the inverse of the fingerprint distances D_r:

$$(x, y) = \frac{\sum_{r=1}^{K} \frac{1}{D_r}(x_r, y_r)}{\sum_{r=1}^{K} \frac{1}{D_r}}, \quad D_r = \mathrm{FD}(\mathbf{RF}_r, \mathbf{TF}) \tag{8}$$

3.3 CSI spatial correlation and temporal stability

The first hypothesis of this work, which motivates to leverage CSI as fingerprint, is that the CSI vectors of points within a certain distance from a reference point are similar to the CSI vector of the reference point.

As a matter of fact, in the matching phase, the test point is associated to the position of the reference point with a more "similar" CSI. In other words, CSI must be characterized by some spatial correlation. On the other hand, this spatial correlation should decrease with the distance, and points that are "far" from the reference point should be characterized by low correlation.

Figure 2 shows the spatial correlation of RSSI and CSI vectors in one of the rooms of the considered apartment, in terms of normalized absolute distance with respect to the values measured in the center of the area (red cross): deep blue means that CSI vectors, or RSSI of that point is very "similar" to the values of the reference point in the center of the area. From Fig. 2, it is evident that CSI has a good spatial correlation around the reference point up to a distance that is around 50 cm. After that distance, the correlation decreases more quickly, and after 1 m, it is almost completely uncorrelated. The trend of the spatial correlation is better for CSI with respect to RSSI.

The second hypothesis encouraging the employment of CSI as fingerprint is that it is expected to show greater temporal stability at a fixed location. This hypothesis is confirmed by Fig. 3 that shows the standard deviation

normalized to the mean value for RSSI, RSRP, and CSI. The standard deviation is evaluated over a time interval of 30 min. At a fixed location, CSI is much more stable than RSSI and also RSRP.

4 CSI fingerprinting localization

In this section, we explain more in detail how the signal fingerprint-based localization approach can be applied to LTE CSI (Fig. 4).

The LTE receiver provides for each antenna port of each decoded cell a sequence of CSI vectors $\mathbf{h}_{c,t}$ over a time interval corresponding to N_{SLOT}. CSI amplitudes are considered and time-averaged to remove noise. The receiver also provides the vectors of RSRP and RSSI over the same time interval equivalent to N_{SLOT}:

$$\mathbf{RSRP}_{c,t} = \begin{bmatrix} \mathrm{RSRP}_{c,t}(0) \\ \vdots \\ \mathrm{RSRP}_{c,t}(s) \\ \vdots \\ \mathrm{RSRP}_{c,t}(N_{\mathrm{SLOT}} - 1) \end{bmatrix} \tag{9}$$

$$\mathbf{RSSI}_c = \begin{bmatrix} \mathrm{RSSI}_c(0) \\ \vdots \\ \mathrm{RSSI}_c(s) \\ \vdots \\ \mathrm{RSSI}_c(N_{\mathrm{SLOT}} - 1) \end{bmatrix} \tag{10}$$

where $\mathrm{RSSI}_c(s)$ represents the average RSSI value between the symbols 0 and 4 in a slot. From all these vectors relative to the LTE signals transmitted by different antennas of multiple eNodeBs, a database like the one shown in Table 1 is built.

In Table 1, for each RP r, and for each eNodeB c (denoted as *Cell*), whose signal is received in the RP, a reference fingerprint $\mathbf{R}_{r,c}$ is calculated. We distinguish the following approaches for calculating $\mathbf{R}_{r,c}$:

Reference Symbols Received Power $\mathbf{R}_{r,c}$ is linked to the vector of $\mathbf{RSRP}_{c,t}$, which represents the sequence of the values of the reference symbol received power from the cell c and the Tx antenna t. Therefore, considering the cth eNodeB with $T_c = 4$ antennas, $\mathbf{R}_{r,c}$ is given by:
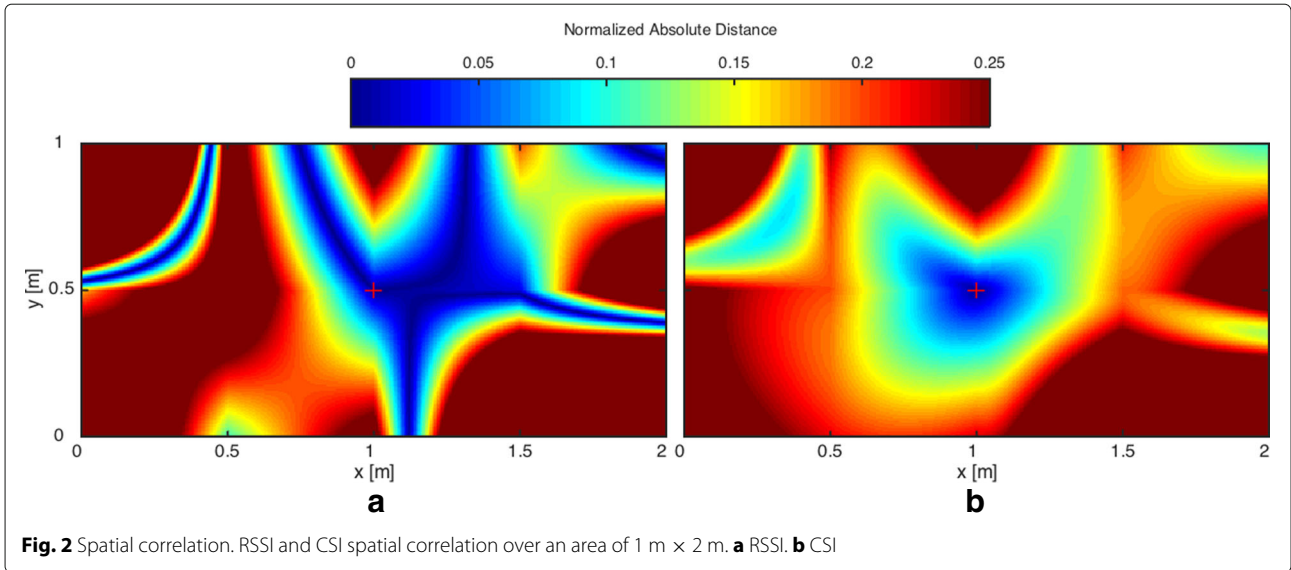
$$\mathbf{R}_{r,c} = \begin{bmatrix} \mathrm{R}_{r,c,1}, \mathrm{R}_{r,c,2}, \mathrm{R}_{r,c,3}, \mathrm{R}_{r,c,4} \end{bmatrix} \tag{11}$$

where

$$\mathbf{R}_{r,c,t} = \overline{\mathbf{RSRP}}_{c,t} = \frac{1}{N_{\mathrm{SLOT}}} \sum_{s=0}^{N_{\mathrm{SLOT}}-1} \mathrm{RSRP}_{c,t}(s) \tag{12}$$

and then the whole reference fingerprint is

$$\mathbf{RF}_r = \begin{bmatrix} \mathbf{R}_{r,1}, \ldots, \mathbf{R}_{r,c}, \ldots, \mathbf{R}_{r,C} \end{bmatrix} \tag{13}$$

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 6 of 18

Fig. 2 Spatial correlation. RSSI and CSI spatial correlation over an area of 1 m × 2 m. **a** RSSI. **b** CSI

Obviously, the same procedure is followed for every test fingerprint **TF**, and according to Eq. 5, it is also necessary to define a fingerprint distance metric FD, which in the case of RSRP, comparison has been chosen equal to:

$$\text{FD}(\mathbf{RF}_r, \mathbf{TF}) = \frac{1}{C} \sum_{c=1}^{C} \left[d(\mathbf{RF}_{r,c}, \mathbf{TF}_c) \right] \quad (14)$$
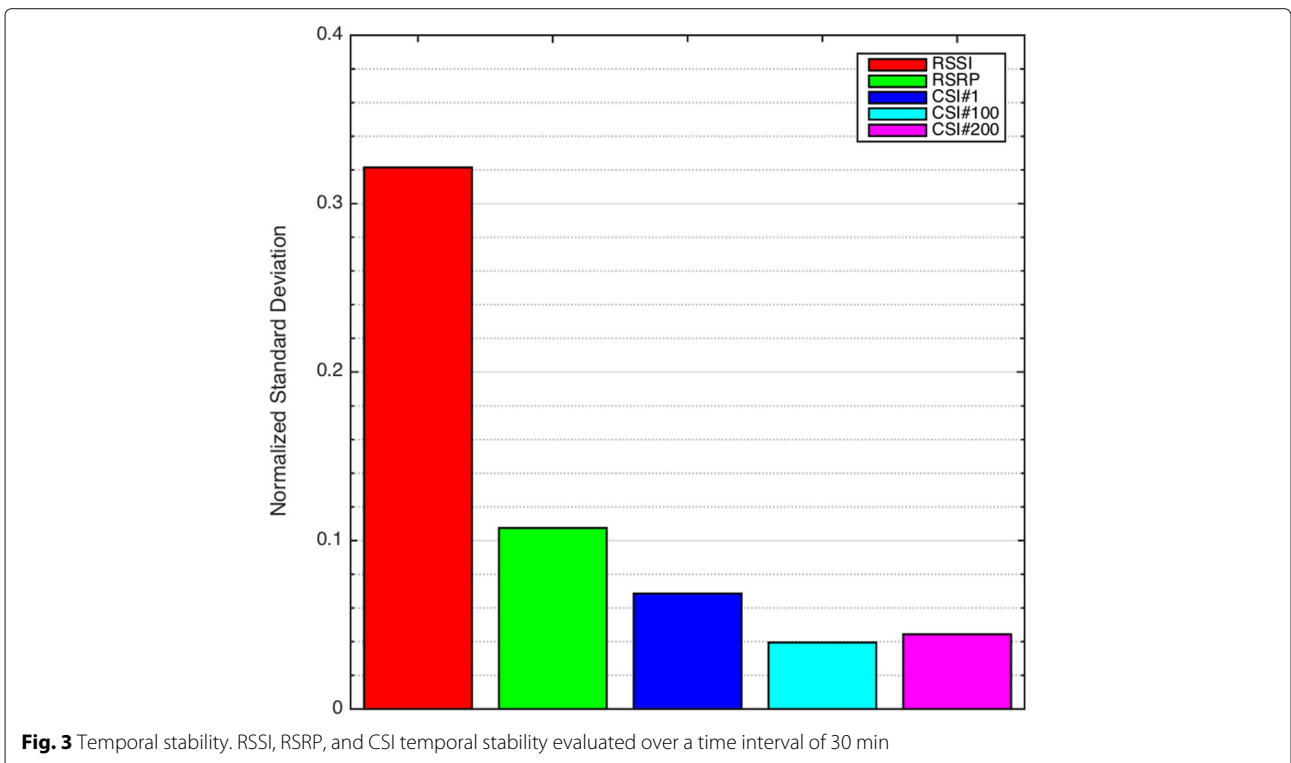
where d is the Euclidean distance between two vectors, and FD is its average over all the available cells.

Direct Channel State Information $\mathbf{R}_{r,c}$ is a vector containing directly all vectors $\mathbf{h}_{c,t}$, which are the CSIs estimated on the signal received from the cell c and the Tx antenna t. Therefore, considering again the cth eNodeB with $T_c = 4$ antennas, $\mathbf{R}_{r,c}$ is given by

$$\mathbf{R}_{r,c} = \left[\mathbf{R}_{r,c,1}, \mathbf{R}_{r,c,2}, \mathbf{R}_{r,c,3}, \mathbf{R}_{r,c,4} \right] \quad (15)$$

where

$$\mathbf{R}_{r,c,t} = \mathbf{h}_{c,t} = [h_0, ..., h_n, ..., h_{N-1}] \quad (16)$$

Fig. 3 Temporal stability. RSSI, RSRP, and CSI temporal stability evaluated over a time interval of 30 min

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing*   (2018) 2018:49
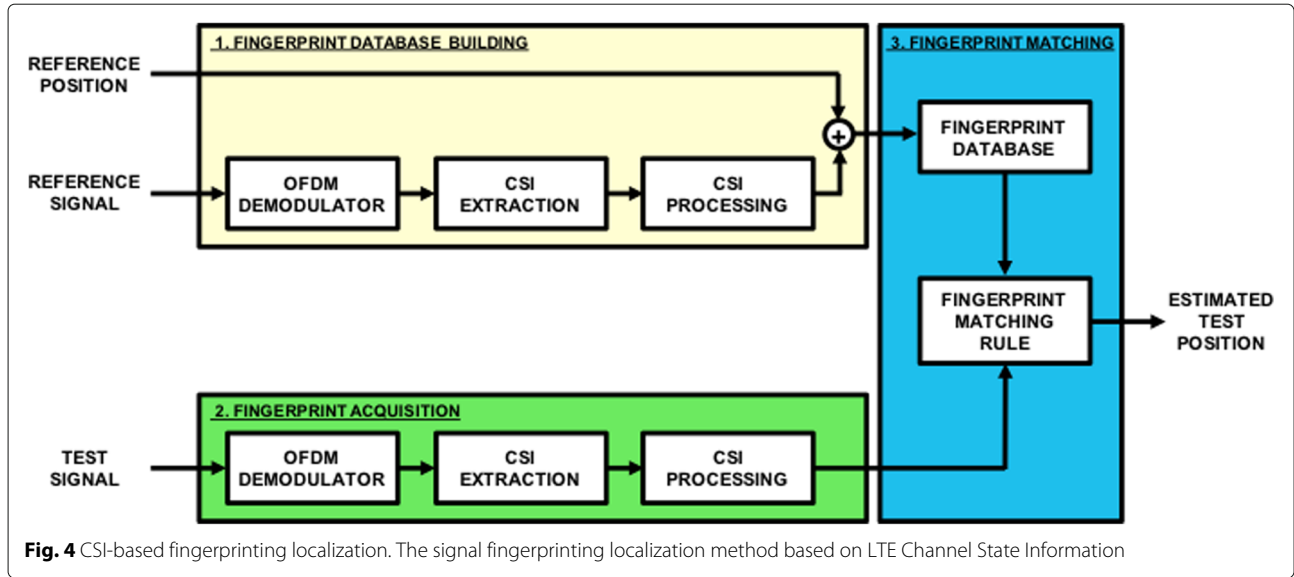
Page 7 of 18

Fig. 4 CSI-based fingerprinting localization. The signal fingerprinting localization method based on LTE Channel State Information

This is how traditionally CSI fingerprinting approaches work. In case of direct CSI comparison, the matching fingerprint distance metric FD has been chosen equal to:

$$
\text{FD}(\mathbf{RF}_r, \mathbf{TF}) = \frac{1}{C} \sum_{c=1}^{C} \left[\frac{1}{T_c} \sum_{t=1}^{T_c} d\left(\mathbf{RF}_{r,c,t}, \mathbf{TF}_{c,t} \right) \right]
$$
(17)

where d is again the Euclidean distance between two vectors, that is now computed between vectors relative to the same transmitting antenna of the same eNodeB.

RSRP and CSI descriptors $\mathbf{R}_{r,c}$ is a vector containing F features calculated on the vector $\mathbf{RSRP}_{c,t}$ or $\mathbf{h}_{c,t}$. Each feature is a number which is somehow related to the statistics of the RSRP (Table 2) or to the "shape" or statistics of the CSI (Table 3). A similar approach has already been successfully employed in [26] to perform device-free crowd counting and occupancy estimation by WiFi.

Table 1 Reference fingerprint database

RP	Coordinates	Cell 1	...	Cell c	...	Cell C
1	(x_1, y_1)	$\mathbf{R}_{1,1}$	...	$\mathbf{R}_{1,c}$	...	$\mathbf{R}_{1,C}$
2	(x_2, y_2)	$\mathbf{R}_{2,1}$	...	$\mathbf{R}_{2,c}$	...	$\mathbf{R}_{2,C}$
⋮	⋮	⋮		⋮		⋮
r	(x_r, y_r)	$\mathbf{R}_{r,1}$	...	$\mathbf{R}_{r,c}$	...	$\mathbf{R}_{r,C}$
⋮	⋮	⋮		⋮		⋮
R	(x_R, y_R)	$\mathbf{R}_{R,1}$	...	$\mathbf{R}_{R,c}$	...	$\mathbf{R}_{R,C}$

The offline fingerprint database storing fingerprints for RPs

Since descriptors are heterogeneous quantities, which can vary in very different intervals, before performing the deterministic classification, it is fundamental to normalize the fingerprints in order to balance all the terms for distance calculation. A *min-max normalization* approach is applied to both reference and test fingerprints:

$$
\hat{\mathbf{R}}_{r,c,t} = \frac{\mathbf{R}_{r,c,t} - \min_{r,c} \left[\mathbf{R}_{r,c,t} \right]}{\max_{r,c} \left[\mathbf{R}_{r,c,t} \right] - \min_{r,c} \left[\mathbf{R}_{r,c,t} \right]}, \quad \forall t
$$
(18)

$$
\hat{\mathbf{T}}_{c,t} = \frac{\mathbf{T}_{c,t} - \min_{r,c} \left[\mathbf{R}_{r,c,t} \right]}{\max_{r,c} \left[\mathbf{R}_{r,c,t} \right] - \min_{r,c} \left[\mathbf{R}_{r,c,t} \right]}, \quad \forall t
$$
(19)

and in this case, the fingerprint distance is calculated as the vector distance between the normalized fingerprints and then averaged over the available cells:

$$
\text{FD}\left(\hat{\mathbf{RF}}_r, \hat{\mathbf{TF}} \right) = \frac{1}{C} \sum_{c=1}^{C} \left[d\left(\hat{\mathbf{RF}}_{r,c}, \hat{\mathbf{TF}}_c \right) \right]
$$
(20)

The descriptor approach, which is novel, has two fundamental advantages:

- It reduces the amount of data that must be stored in the database.
- It reduces the computational complexity associated to the matching phase.

Received Signal Strength Indicator In the traditional RSSI fingerprint-based localization approach, the reference fingerprint $\mathbf{R}_r$ is simply given by the average RSSI measured from all the available eNodeBs:

$$
\mathbf{R}_{r,c} = \overline{\mathbf{RSSI}}_c = \frac{1}{N_{\text{SLOT}}} \sum_{s=0}^{N_{\text{SLOT}}-1} \text{RSSI}_c(s)
$$
(21)

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 8 of 18

Table 2 RSRP descriptors

Descriptor	Formula	Description
Mean	$\mu = \frac{1}{N_{\text{SLOT}}} \sum_{s=0}^{N_{\text{SLOT}}-1} \text{RSRP}[s]$	The arithmetic mean of the RSRP.
Standard deviation	$\sigma = \sqrt{\frac{1}{N_{\text{SLOT}}-1} \sum_{s=0}^{N_{\text{SLOT}}-1} (\text{RSRP}[s] - \mu)^2}$	The standard deviation of the RSRP.
Fano factor	$\text{FF} = \frac{\sigma^2}{\mu}$	The ratio between the variance of the RSRP and its arithmetic mean.

The list of RSRP statistical descriptors

$$
\begin{aligned}
\mathbf{RF}_r &= \left[\text{R}_{r,1}, \ldots, \text{R}_{r,c}, \ldots, \text{R}_{r,C} \right] \\
&= \left[\overline{\text{RSSI}}_1, \ldots, \overline{\text{RSSI}}_c, \ldots, \overline{\text{RSSI}}_C \right]
\end{aligned}
\tag{22}
$$

The fingerprint distance metric FD in case of RSSI comparison has been chosen equal to:

$$
\begin{aligned}
\text{FD}\,(\mathbf{RF}_r, \mathbf{TF}) &= \frac{1}{C} \sum_{c=1}^{C} \left[d\left(\text{R}_{r,c}, \text{TF}_c \right) \right] \\
&= \frac{1}{C} \sum_{c=1}^{C} |\text{R}_{r,c} - \text{TF}_c|
\end{aligned}
\tag{23}
$$

where d is simply the magnitude of the difference between the reference and the test RSSI.

Table 3 CSI descriptors

Descriptor	Formula	Description
Mean	$\mu = \frac{1}{N} \sum_{n=0}^{N-1} h_n$	The arithmetic mean of the CSI.
Standard deviation	$\sigma = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} (h_n - \mu)^2}$	The standard deviation of the CSI.
Fano factor	$\text{FF} = \frac{\sigma^2}{\mu}$	The ratio between the variance of the CSI and its arithmetic mean.
Spectral centroid	$f_n = (3n - 6\text{RB})15 \text{ kHz}$ $\quad SC = \frac{\sum_{n=0}^{N-1} h_n f_n}{\sum_{i=0}^{N-1} h_i}$	The "center of mass" calculated as the weighted mean of the frequency values with CSI normalized magnitudes as weights.
Spectral lambda	$\lambda = -\frac{1}{N-1} \sum_{n=1}^{N-1} \frac{h_n - h_{n-1}}{f_n - f_{n-1}} \frac{2}{h_n + h_{n-1}}$	The mean of the derivative function for the CSI.
Spectral entropy	$SE = -\sum_{n=0}^{N-1} \frac{h_n}{\sum_{j=0}^{N-1} h_j} \log_2 \frac{h_n}{\sum_{i=0}^{N-1} h_i}$	The amount of information contained in the CSI.
Spectral flatness	$SF = \frac{\sqrt[N]{\prod_{n=0}^{N-1} h_n}}{\frac{1}{N} \sum_{n=0}^{N-1} h_n}$	A measure used in digital signal processing to quantify how noise-like the CSI is.
Spectral slope	$SSL = \frac{\sum_{n=0}^{N-1} (f_n - \bar{f}_n)(h_n - \mu)}{\sum_{n=0}^{N-1} (f_n - \bar{f})^2}$	A measure of the slope of the spectral shape of CSI.
Spectral spread	$SSP = \sqrt{\frac{\sum_{n=0}^{N-1} h_n (f_n - SC)^2}{\sum_{i=0}^{N-1} h_i}}$	A measure of how the spectrum is distributed around its centroid.
Spectral moment	$\eta_j = \frac{\sum_{n=0}^{N-1} h_n f_n^j}{\sum_{i=0}^{N-1} h_i}$	The jth order spectral moment of the CSI.
Spectral central moment	$\xi_j = \frac{\sum_{n=0}^{N-1} h_n (f_n - SC)^j}{\sum_{i=0}^{N-1} h_i}$	The jth order spectral central moment of the CSI.
Spectral kurtosis	$SKU = \frac{\sum_{n=0}^{N-1} h_n T_n^4}{\sum_{i=0}^{N-1} h_i}, \quad T_n = \frac{f_n - SC}{\sqrt{\xi_2}}$	A measure of the "tailedness" of the CSI.
Spectral skewness	$SSK = \frac{\sum_{n=0}^{N-1} h_n T_n^3}{\sum_{i=0}^{N-1} h_i}$	A measure of the asymmetry of the CSI about its spectral centroid.

The list of CSI statistical and shape descriptors

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 9 of 18

It should also be considered that in small and static environments, there is a fixed set of base stations received at every reference point during database building phase and at every test point during matching phase. However, one should not neglect that in larger areas, it is easy to observe different eNodeBs between reference and test points, and furthermore, in dynamically changing environment, these sets of eNodeBs can also temporally vary due to signal fluctuations. This issue is managed by comparing two fingerprints only in relation to the cells in common and excluding the others from the averaging process.

5 Methods
5.1 Experimental test-bed
Experimental results have been carried out indoor in a 90-m^2 apartment with six rooms and three balconies (Fig. 5). The apartment has rooms of different sizes and differently equipped. Firstly, we have assessed the performance of the proposed method over this complex environment. Then, we have also evaluated the performance within single rooms, as it has been most commonly done in other literature works. The experimental area has been mapped by reference points in a mesh grid of 0.5 m. Since the apartment is not empty and there is furniture, it was not possible to completely grid the rooms. We have acquired two different datasets, one as reference and one for testing. A total of 113 RPs were placed in the apartment, and 30 TPs were randomly chosen all over the area, as shown in Fig. 6. During the experiments, the furniture was not moved, all internal doors were closed, and all the balconies were open.

5.2 Data acquisition and processing
The receiver employed for data acquisition is the Great Scott Gadget HackRF equipped by an ANT500 omni-directional antenna, which is able to tune its operating frequency from 1 MHz to 6 GHz and to capture up to
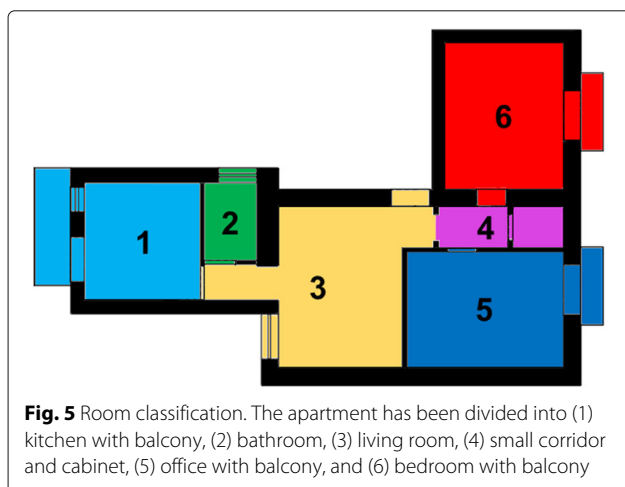

Fig. 5 Room classification. The apartment has been divided into (1) kitchen with balcony, (2) bathroom, (3) living room, (4) small corridor and cabinet, (5) office with balcony, and (6) bedroom with balcony

20 MS/s with 8-bit resolution. The receiver can receive LTE signals both at 800 and 1800 MHz. The HackRF SDR platform has acquired the raw samples of the received signal in each RP and TP and has stored them in binary files. The raw files have then been post-processed by an LTE software receiver, which performs the channel estimation and outputs RSSI, RSRP, and CSI vectors to be imported into MATLAB. A time interval of 1 s, corresponding to $N_{SLOT} = 2000$ slots, has been considered to incoherently average the captured RSRP and CSI or to calculate descriptors.

5.3 Experimental methodology
The following experiments have been carried out in the same environment.

Experiment 1 The HackRF platform has been configured to capture spectrum raw samples at a sampling rate of 1.92 MS/s and with a receiving filter bandwidth of 2.5 MHz. The receiver gains were set to fixed values in order to avoid distortions introduced by the Automatic Gain Control (AGC). The RF amplifier was disabled and the intermediate frequency gain was set to 40 dB and the baseband gain to 30 dB. The ANT500 omni-directional receiving antenna has then been placed on the ground in each RP and TP, and signal samples relative to the LTE cells 49 (796 MHz), 255 (806 MHz), 116 (816 MHz), 285 (1817.5 MHz), and 21 and 75 (1870 MHz) have been captured for a time interval of 10 s at each frequency. The raw samples have been processed and, at such sampling rate, have allowed to observe 6 RBs and thus to get $N = 24$ complex channel gains. Since the just-mentioned LTE cells are all configured to work with two transmitting antennas, each fingerprint has included data from both of them. Figure 7 shows the percentage of points, with respect to all reference and test points, where it was possible to decode the signal corresponding to one specific eNodeB.

Experiment 2 The HackRF receiver has been configured to work at a sampling frequency of 15.36 MS/s and with a receiving bandwidth of 10 MHz. The RF amplifier and AGC were disabled and the intermediate frequency gain was set to 40 dB and the baseband gain to 30 dB as in the previous experiment. Since the scope of the experiment was just to exploit the effect of a bandwidth increase to the localization performance, only the samples relative to the best cell 255 (806 MHz), which, as shown in Fig. 7, was available in all RPs and TPs, were captured for the same time interval of 10 s. In this case, the higher rate samples allowed to observe up to 50 RBs and provided $N = 200$ complex channel gains for both antennas.

5.4 Performance evaluation criteria
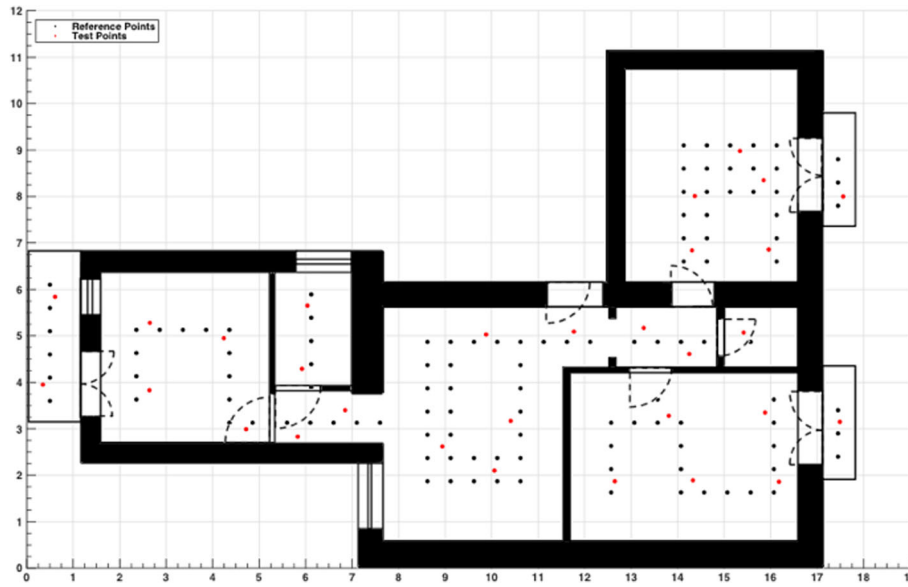We have considered the following performance metrics:

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing*   (2018) 2018:49

Page 10 of 18

Fig. 6 Indoor experimental layout. The indoor experimental area is an apartment of 90 m^2, which has been mapped by 113 reference points (black points) and 30 test points (red points)

Mean Distance Error (MDE) is calculated by averaging the Euclidean distance between all the TPs' true positions $(x_{t,\text{true}}; y_{t,\text{true}})$ and the estimated positions $(x_{t,\text{est}}; y_{t,\text{est}})$ to which they have been matched:

$$\text{MDE} = \frac{1}{N_{\text{TF}}} \sum_{t=1}^{N_{\text{TF}}} \sqrt{\left(x_{t,\text{true}} - x_{t,\text{est}}\right)^2 + \left(y_{t,\text{true}} - y_{t,\text{est}}\right)^2}$$

(24)

where N_{TF} represents the overall number of test fingerprints extracted from all the test points.

Cumulative Distribution Function (CDF) of the localization error is used to provide a better characterization of the localization performance. When two techniques are compared, if their accuracies are the same, it should be preferred the system with the CDF graph, which reaches higher probability values faster. In this work, the CDF has been described through the percentile format and, in
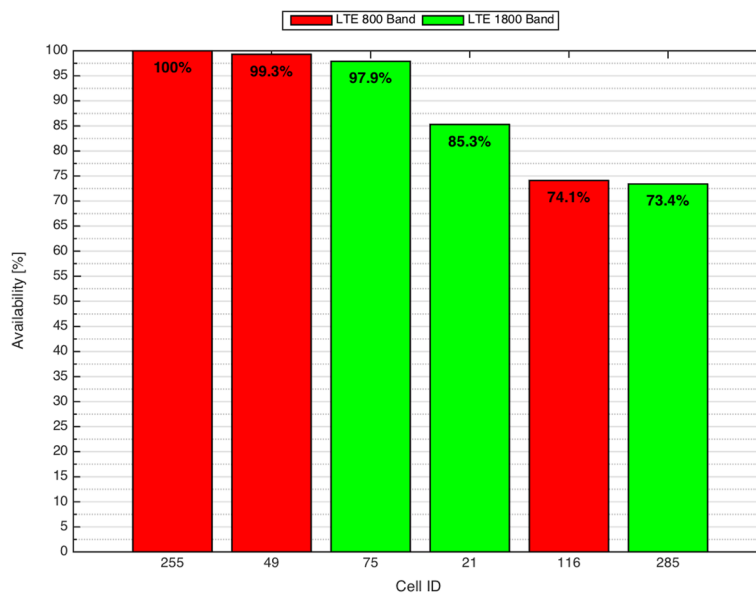

Fig. 7 LTE cell availability. Percentage of points over the whole test area correctly decoding the LTE cells taken under analysis

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 11 of 18

particular, by providing the 50th and 75th percentile of the localization error. MDE is related to the accuracy, but its value could be strongly affected by an outlier. Therefore, the use of both MDE and the CDF gives a more complete understanding of the performance.

Room Classification Accuracy (RCA) is the percentage of TPs over the total number of TPs for which the identified position is within the correct room, as show in Fig. 5. In other words, we wonder, regardless of the accuracy in terms of MDE, which is the technique in which rooms are confused as less as possible.

6 Experimental results and discussion
6.1 CSI vs RSSI vs RSRP in a narrowband scenario
First of all, we wonder whether (and when) the use of CSI for LTE signal fingerprinting improves the performance with respect to the use of more classical approaches based on RSSI and RSRP. Moreover, we wonder whether using RSRP/CSI descriptors as fingerprint rather than direct CSI vectors is feasible and, in case, convenient. Results shown in this section are achieved using the signal received only from cell 255 (cell diversity is not considered) using 6 RBs (experiment 1), i.e., the LTE signal is observed over a small bandwidth of 1.4 MHz (narrowband scenario).

Table 4 summarizes the achieved results in terms of MDE, RCA, and 50th and 75th percentile of the achieved accuracy. Table 4 compares the results when different fingerprints are used: RSSI, RSRP, direct CSI vectors (CSI in the table), and RSRP/CSI descriptors (as explained in Section 4). In case of RSRP/CSI descriptors, three cases are considered: two, three, or four descriptors. We have not considered a higher number of descriptors as we have observed that increasing the number of descriptors does not contribute to further improve the performance.

The selected RSRP/CSI descriptors providing the best performance are shown in the last column of Table 4. First of all, Table 4 and Fig. 8 clearly show that using CSI greatly improves the performance with respect to the use of RSSI. The MDE goes from 4.32 m with RSSI to 2.91 m, while RCA increases from 36 to 57% using CSI vectors as fingerprint. CSI-based approaches (CSI vectors and descriptors) also show better performance with respect to the RSRP approach even if, in this narrowband scenario, only a slight improvement of performance is observed.

Furthermore, the very interesting result is that the reduction of the database size and the computational complexity of the matching phase that can be achieved using descriptors rather than CSI vectors is not paid in terms of performance; as a matter of fact, a signal fingerprinting method based on the use of descriptors outperforms a method based on CSI vectors. Moreover, with respect to the RSRP, using two descriptors reduces the MDE of 6%, the 50th percentile of 15%, the 75th percentile of 10%, and only slightly reduces the RCA (1%). It is worth outlining once more that the results shown in this section do not consider cell diversity, and the bandwidth of the LTE signal is rather small (1.4 MHz). Moreover, these results are achieved using an NN classifier. Therefore, in next sections, we first investigate the impact of the selected classifier, considering also the KNN and the WKNN classifiers. Then, the impact of cell diversity and a larger bandwidth of the acquired LTE signal is investigated.

6.2 Selection of classifier in a narrowband scenario
In Section 3.2, it was stated that in case of deterministic matching, the estimated location could be found not only by considering the most similar fingerprint (NN) but also extracting the first *K*-nearest neighbor fingerprints

Table 4 Comparison of signal fingerprinting approaches based on RSSI, RSRP, CSI vectors, and RSRP/CSI descriptors (1 cell, 6 RBs)

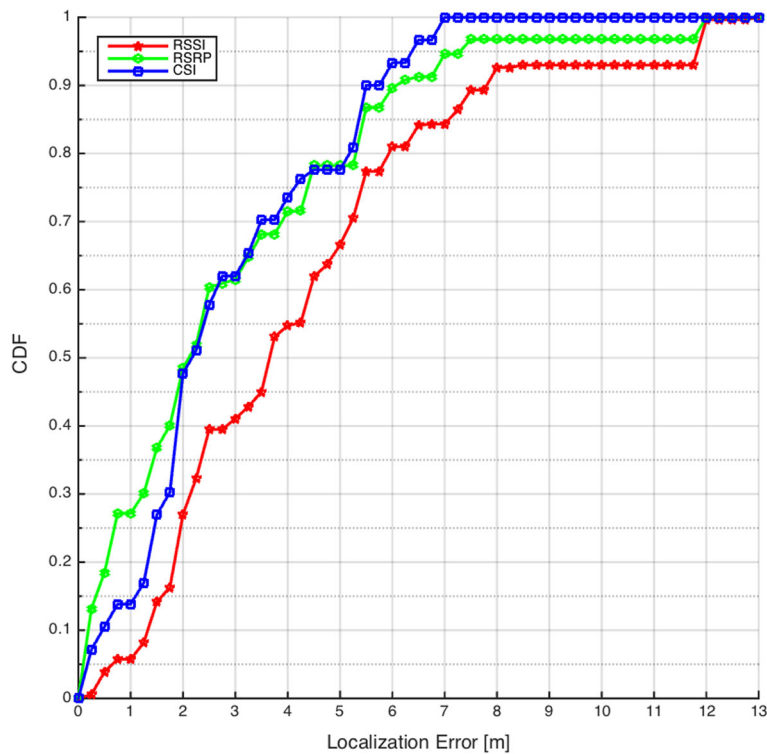| Fingerprint	MDE	CDF 50%	CDF 75%	RCA	Descriptors
	[m]	[m]	[m]	[%]	
RSSI	4.32	3.76	5.40	36	–
RSRP	3.00	2.28	4.63	58	–
CSI	2.91	2.28	4.17	57	–
2 RSRP/CSI descriptors	2.82	1.94	4.17	57	CSI mean
					RSRP standard deviation
3 RSRP/CSI descriptors	2.81	2.07	4.17	57	CSI mean
					RSRP mean
					RSRP standard deviation
4 RSRP/CSI descriptors	2.86	2.07	4.63	57	CSI mean
					RSRP mean
					RSRP standard deviation
					RSRP Fano factor

Fig. 8 CDF of the localization error in a narrowband scenario. Comparison of cumulative distribution functions of the localization error for signal fingerprinting approaches based on RSSI, RSRP, and CSI

and then performing an arithmetic mean (KNN) or a weighted mean (WKNN) of their relative reference coordinates. In this section, the impact of the parameter K on the localization performance for all the proposed methods is analyzed. The performance comparison between KNN and WKNN as a function of K is shown in Fig. 9, where $K \leq 5$ since it was found that larger values does not contribute to decrease the MDE. As a result, it has been found out that for both KNN and WKNN, the minimum error is achieved with $K = 2$ and $K = 3$, respectively, for CSI and descriptors. While larger values of K tend to increase the error in case of using direct CSI, the localization error using RSRP/CSI descriptors is more stable after $K = 3$; this is particularly true when WKNN is used. The best performance are achieved with KNN classifier with direct CSI for $K = 3$.
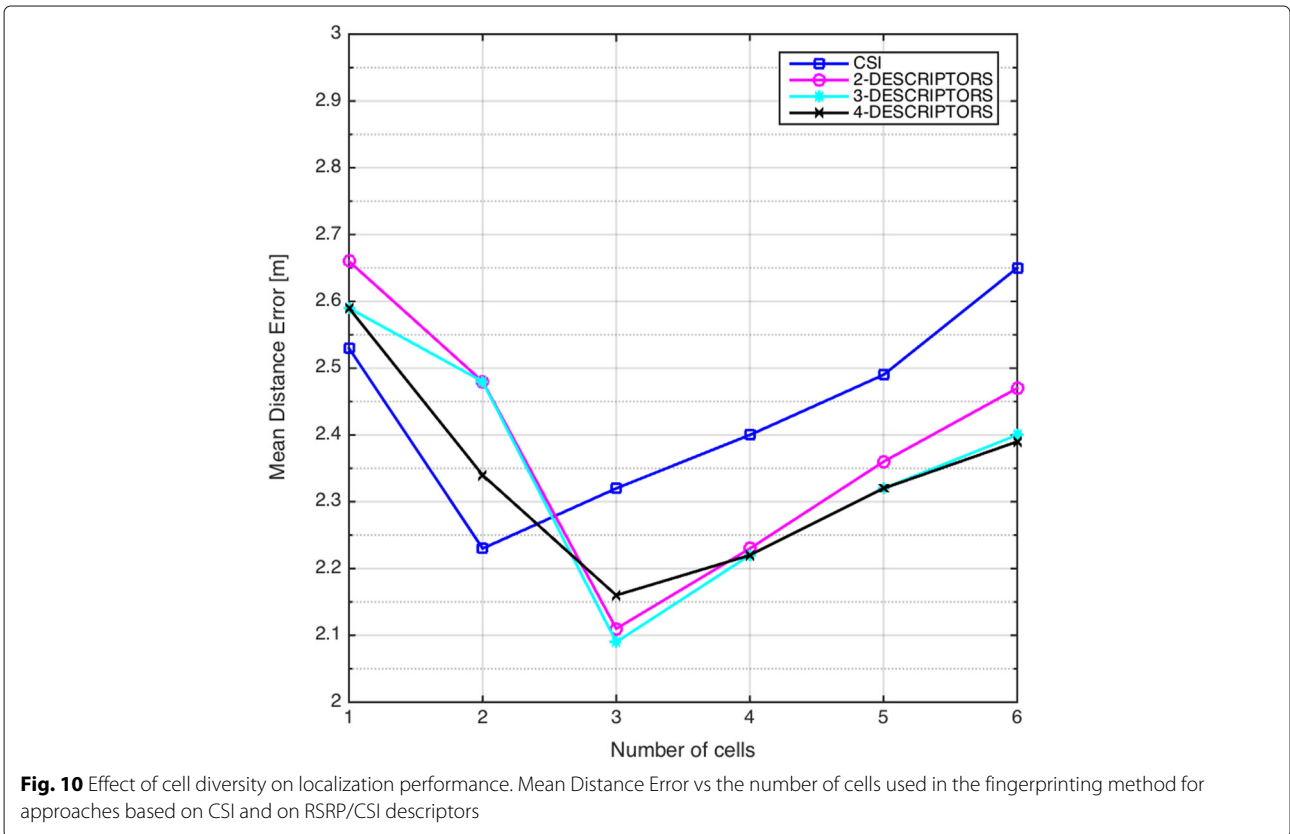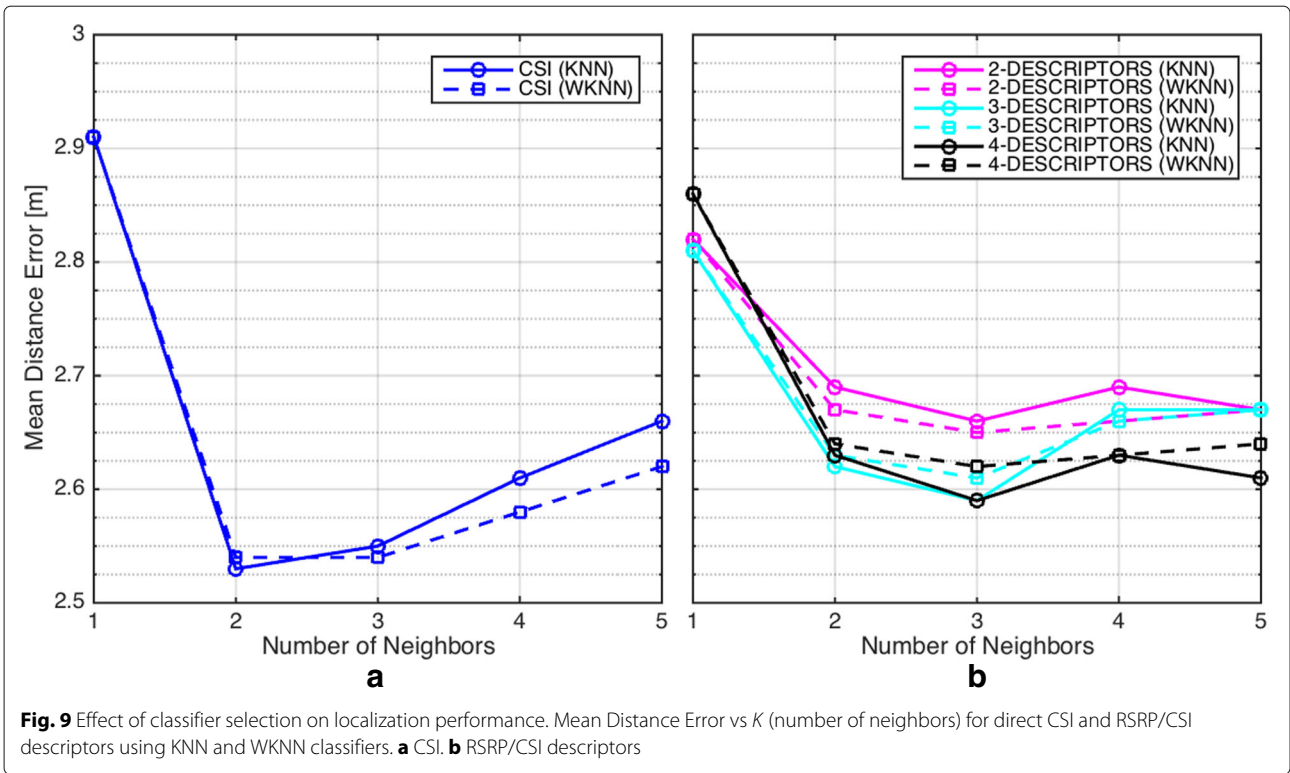
6.3 Effect of cell diversity

This section investigates the performance of different LTE signal fingerprinting localization approaches when more LTE signals, received from different eNodeBs, are available and used by the fingerprinting method. We present the results only for direct CSI and RSRP/CSI descriptors. As shown in Fig. 7, in most of the RPs, it is possible to receive up to six signals from six different eNodeBs. Therefore, the different fingerprinting approaches have

been compared assuming the use of signals from up to six cells. For each method, applying the best classifier chosen in Section 6.2, all possible combinations of available cells have been considered.

In Fig. 10, for every fingerprinting approach and for different number of cells, the MDE achieved by the best combination of cells has been plotted. All methods show a noticeable improvement in the performance when two/three cells are considered. In the case of CSI vectors, the MDE goes from 2.53 m with one signal to 2.23 m with two signals. With the descriptors, the best results are achieved using up to three signals coming from three eNodeBs, where the MDE goes below 2.10 m. However, it is also evident that further increasing in the number of cells (e.g., LTE signals from different eNodeB) does not improve the performance.

This can be explained considering that the different eNodeB signals are received with different quality; some of them have high signal quality while some others have low signal quality. In any reference point, it is not expected to receive more than two/three good quality signals. As a matter of fact, since the fingerprint is built using both "useful information" and "noise" from the signal received from each cell, with each signal having a different portion of each, then there exists a trade-off between "useful information" and "noise" bringing to an optimum number

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 13 of 18

Fig. 9 Effect of classifier selection on localization performance. Mean Distance Error vs *K* (number of neighbors) for direct CSI and RSRP/CSI descriptors using KNN and WKNN classifiers. **a** CSI. **b** RSRP/CSI descriptors

Fig. 10 Effect of cell diversity on localization performance. Mean Distance Error vs the number of cells used in the fingerprinting method for approaches based on CSI and on RSRP/CSI descriptors

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing*   (2018) 2018:49

Page 14 of 18

of signals (i.e., cells) which minimizes the localization error.

6.4   Effect of bandwidth

In the previous sections, the LTE signal is observed over a small bandwidth of 1.4 MHz. This section shows the same comparisons as in Section 6.1 using an LTE signal of larger bandwidth. In particular, the signal from cell 255 has been acquired by the highest available sampling rate of 15.36 MS/s (complex samples). This allows to extract CSI over a bandwidth of 10 MHz (wideband scenario).

Table 5 summarizes the results. Similarly to the narrowband case, CSI approaches outperform the RSSI approach. However, the most relevant result in Table 5 is that when larger bandwidth LTE signals are used, the proposed descriptor approach largely outperforms the RSRP approach and also the CSI approach based on CSI vectors. Moreover, by looking at the selected descriptors, in this case, all of them are CSI descriptors related to the *shape* of the CFR, while in the narrowband case, almost all of them are related to RSRP. It is also worth outlining that, in this wideband case, the database compression achieved by using descriptors rather than CSI vectors is even larger with respect to the narrowband case; as a matter of fact, instead of storing the CSI vector of size 200 (total number of subcarriers), we just need to store a vector of size 4 (number of descriptors). To better understand the impact of a larger bandwidth LTE signal, we have also investigated the performance considering two intermediate bandwidths between 1.4 and 10 MHz, corresponding to the use of 15 RBs and 25 RBs, which respectively correspond to the feasible LTE sampling rates of 3.84 and 7.68 MS/s. Figure 11 shows the MDE as a function of the available RBs (and hence, available bandwidth) for the CSI-based approaches. In the latter cases, the relative

CSI vectors have been obtained by decimating the ones extracted from the measurements taken at 15.36 MS/s during experiment 2.

Interestingly, for the signal fingerprinting approach based on CSI vectors, the performance does not improve by increasing the bandwidth since multipath is both a source of information and a disturbance. A completely different behavior is observed when descriptors are used. In the latter case, performance greatly improve by increasing the bandwidth of the LTE signals. These results confirm the intuition behind the proposed method based on CSI and shape descriptors, i.e., the fact that each point can be uniquely characterized by the way the multiple replicas of the transmitted signal are combined at the receiver, and this "unique" characteristic is effectively caught by the shape of the CFR. In the case of narrowband scenario, the CFR is almost flat, and no much gain can be achieved by its use. Over a larger bandwidth, and assuming a rather frequency selective channel, the CFR has a unique shape strictly related to the multipath propagation in that point. Therefore, a larger bandwidth increases the resolution in distinguishing different multipath profiles. However, in case of direct CSI vectors, the matching phase uses larger vectors, where each element is a noisy sample. Therefore, on one hand, these new vector elements bring information on the channel. On the other hand, they introduce more noise. This balance is evident in the case of CSI vectors. However, results show that by using some distinctive characteristic of the shape of the CFR greatly helps in effectively extracting only the useful information from the signals and take full advantage of the availability of larger bandwidths. Figure 12 shows that in the wideband case, the best performance is obtained through the WKNN with $K = 2$. In particular, the achieved localization accuracy with four descriptors has an MDE of 1.98 m;

Table 5 Comparison of signal fingerprinting approaches based on RSSI, RSRP, CSI vectors, and RSRP/CSI descriptors (1 cell, 50 RBs)

Fingerprint	MDE	CDF 50%	CDF 75%	RCA	Descriptors
	[m]	[m]	[m]	[%]	
RSSI	4.12	4.17	5.35	28	–
RSRP	3.22	2.14	4.50	55	–
CSI	3.10	2.28	3.90	61	–
2 descriptors	2.48	1.89	3.36	69	CSI mean
					CSI 2nd spectral central mom.
3 descriptors	2.24	1.89	3.18	56	CSI mean
					CSI spectral entropy
					CSI spectral spread
4 descriptors	2.11	1.82	2.93	65	CSI mean
					CSI spectral lambda
					CSI spectral kurtosis
					CSI spectral flatness

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing*   (2018) 2018:49
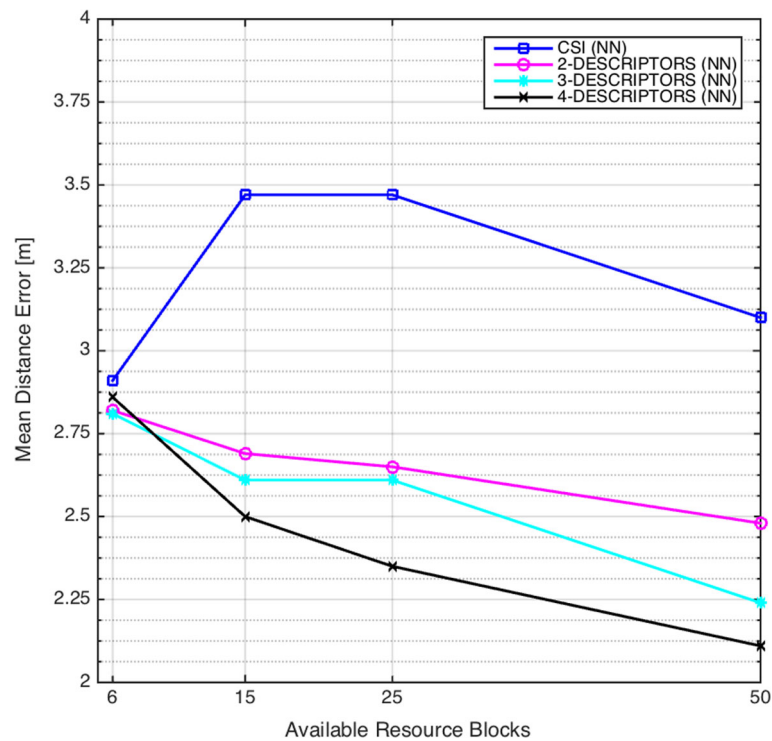
Page 15 of 18

Fig. 11 Effect of bandwidth increase on localization performance. Mean Distance Error with respect to the LTE signal bandwidth in terms of available resource blocks

50th and 75th percentiles of 1.78 and 2.58 m, respectively; and an RCA of 66%.

6.5   Effect of reference point reduction
Signal fingerprinting-based localization techniques can guarantee very high localization accuracies, but to achieve good performance, an extensive training campaign is required. In general, an increase in the number of reference points or a reduction of their relative distances provides an improvement in the localization accuracy, but in case the coverage area becomes larger, this mapping can be extremely time-consuming. For this reason, there is much work in the literature proposing interpolation-based techniques able to give a quite accurate estimate of the RSSI distribution over the area of interest by only taking a small number of reference samples [27]. In this section, it is investigated how the number of reference points can influence the localization accuracy, and for this reason, multiple reference grids have been considered in which the number of reference points has not been simply decreased, but they have also been distributed as evenly as possible all around the area of interest, and at least one reference point has been left in every environment. In detail, four different configurations have been analyzed: $\mathbf{R}_1$ with 113 RPs (100%), $\mathbf{R}_2$ with 60 (56%), $\mathbf{R}_3$ with 27 (25%), and finally $\mathbf{R}_4$ with just 1 RP for every room and balcony consisting of a total of 11 RPs (10%). The behavior

of MDE with respect to the change of the reference grid is shown in Fig. 13 for the WKNN RSRP/CSI descriptors method with $K = 2$.

As it could be expected, all methods are characterized by an MDE increase as a consequence of the gradual exclusion of fingerprints from the reference database. Nevertheless, as shown in Fig. 13, it is also worth noting that the performance obtained through the descriptors over one cell and with just one tenth of the RPs is still much better than the ones employing RSSI and all the RPs. This means that even with just one RP per environment, it is possible to find a combinations of three or four descriptors that can still guarantee good performance. This represents an important result when also the implementation issues are considered; reducing the number of "needed" reference points to get a certain accuracy allows to decrease the memory occupancy, but mainly to greatly reduce the time and the effort to take reference measurements all over the area of interest.

6.6   Single room localization performance
In the previous sections, the proposed system has been tested over a large complex indoor environment. To ease the comparison with state-of-the-art works, which usually test their system in single rooms of different sizes, Table 6 also reports the localization performance of the proposed LTE signal fingerprinting system for each single room of
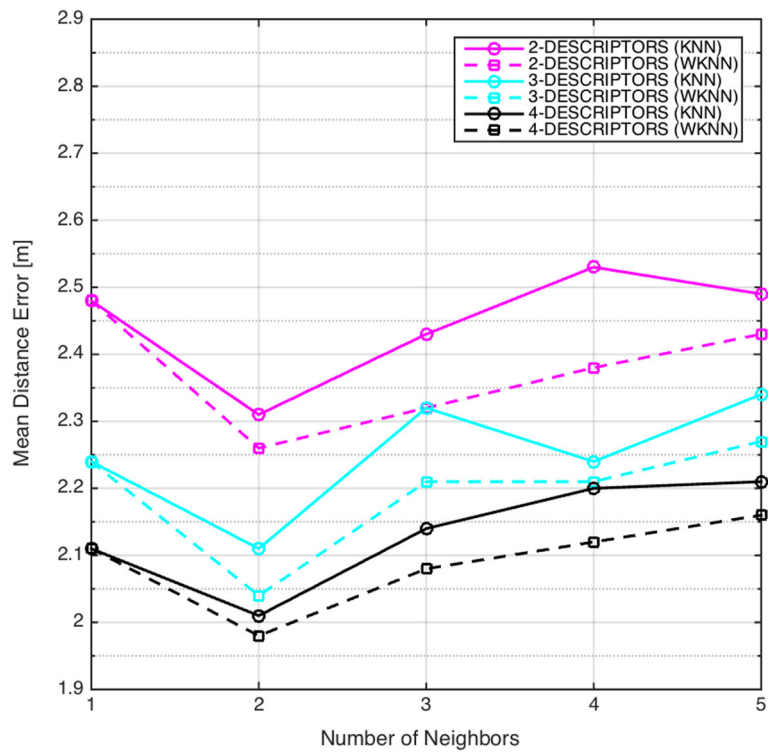
Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 16 of 18

Fig. 12 Effect of classifier selection on localization performance in a wideband scenario. Mean Distance Error vs *K* (number of neighbors) for RSRP/CSI descriptors using KNN and WKNN classifiers
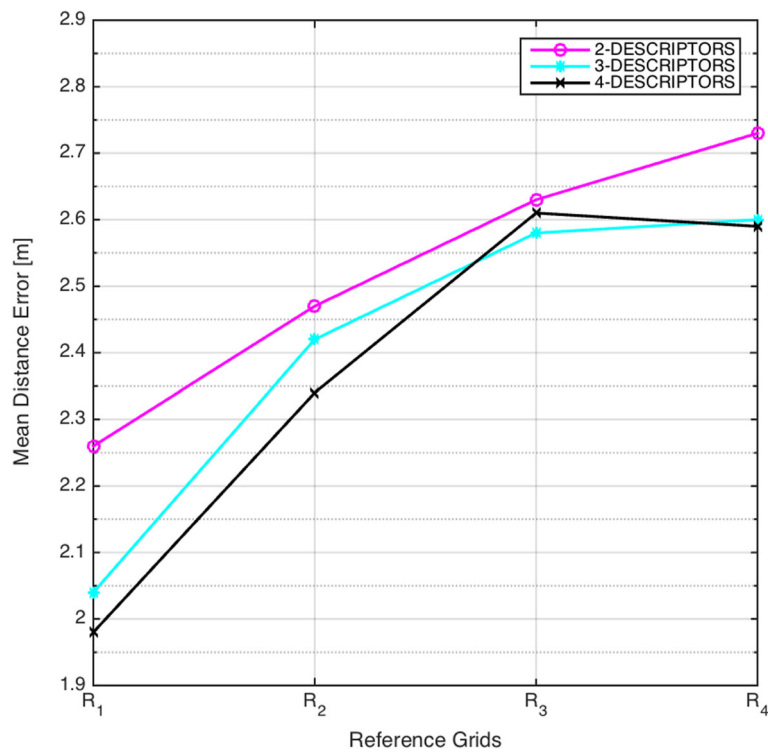

Fig. 13 Effect of reference point reduction. Mean Distance Error vs number of reference points for RSRP/CSI descriptors

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 17 of 18

Table 6 Single room localization performance

Environment	Size	Method	MDE	CDF 50%	CDF 75%
	[m]		[m]	[m]	[m]
Apartment	11 × 18	4 descriptors (LTE)	1.98	1.78	2.58
Kitchen	3.7 × 3.7		0.74	0.83	0.84
Bathroom	2.5 × 1.7		0.14	0.14	0.2
Living room	5 × 7		1.05	0.9	1.64
Office	3.6 × 4.9		0.74	0.62	0.94
Bedroom	4.6 × 3.7		0.91	0.89	1.12
Laboratory	6 × 9	DeepFi (WiFi)	1.81	–	–
Living room	4 × 7		0.94	–	–
Office	6.5 × 7	CSI-MIMO (WiFi)	0.95	–	–

Localization performance within different apartment rooms and comparison with DeepFi [3] and CSI-MIMO [28]

the apartment. In particular, the reported performance are the ones achieved employing a signal fingerprinting approach based on the use of four descriptors with 50 RBs available (10 MHz bandwidth) on a single cell. Table 6 also reports the performance achieved with recently proposed systems, namely DeepFi [3] and CSI-MIMO [28], for indoor localization with using CSI with WiFi signals.

Reported results in Table 6 show that the proposed LTE signal fingerprinting approach based on CSI has localization performance that are comparable with the ones of a WiFi-based system using CSI. For instance, in the case of the living room, which has almost the same size of the living room in the DeepFi system, MDE is only slightly higher (1.05 rather than 0.94).

7 Conclusions

This paper investigates the possibility to use the CSI extracted from LTE signals for signal fingerprinting indoor localization. Moreover, the paper proposes a novel way to use CSI as fingerprints; in the proposed system, fingerprints are descriptors of the "shape" of the CFR calculated on these CSI vectors, rather than direct CSI vectors. This greatly reduces the requirements in terms of memory for the database and also the computational complexity of the matching phase. It is also worth outlining that the proposed method works on the signaling messages sent by an eNodeB, and hence, the device that must be localized does not need to have any subscription with a specific operator. This also gives the possibility to use different signals coming from different eNodeBs (regardless of the specific operator) and hence eventually get a cell diversity gain. Experimental results shown in the paper prove that using CSI for LTE signal fingerprinting can effectively contribute to improve localization accuracy with respect to the RSSI and RSRP. In particular, we can conclude that:

1. When only small bandwidth LTE signals can be extracted (narrowband scenario), and with no cell diversity, CSI-based signal fingerprinting approaches improve the performance with respect to the use of RSSI and RSRP (the latter has been most commonly used in the literature with LTE signal fingerprinting).
2. The CSI-based approach can greatly take advantage of the available cell diversity up to two to three different eNodeBs included in the fingerprint.
3. When LTE signals with larger bandwidth can be extracted, the proposed CSI-based approach with descriptors shows noticeable better performance with respect to both an RSRP approach and a CSI approach with CSI vectors. These performances are comparable with the ones achieved by a state-of-the-art signal fingerprinting system based on WiFi signals and CSI.
4. The proposed approach has been proven to be robust to a reduction of the number of reference points needed for building the database.

In the future, it could be interesting to investigate the possibility to combine CSI-based approaches using both WiFi and LTE signals. Furthermore, it will be crucial to investigate implementation issues related to the database building.

Abbreviations
AECID: Adaptive enhanced cell identity; AGC: Automatic gain control; AOA: Angle of arrival; AP: Access point; BS: Base station; CDF: Cumulative distribution function; CFR: Channel frequency response; CIR: Channel impulse response; CRS: Cell-specific reference signal; CSI: Channel state information; FD: Fingerprint distance; FDD: Frequency division duplexing; FFT: Fast Fourier transform; GNSS: Global Navigation Satellite System; GSM: Global System for Mobile vommunications; KNN: *K*-nearest neighbor; LTE: Long Term Evolution; MDE: Mean distance error; NN: Nearest neighbor; OFDM: Orthogonal frequency division multiplexing; OTDOA: Observed time difference of arrival; RB: Resource block; RCA: Room classification accuracy; RE: Resource element; RF: Radio frequency; RP: Reference point; RSRP: Reference signal received power; RSRQ: Reference signal received quality; RSSI: Received signal strength indicator; TA: Timing advancing; TDOA: Time difference of arrival; TOA: Time of

Pecoraro *et al. EURASIP Journal on Advances in Signal Processing* (2018) 2018:49

Page 18 of 18

arrival; TP: Test point; U-TDOA: Uplink-time difference of arrival; UE: User equipment; UMTS: Universal Mobile Telecommunications System; WKNN: Weighted *K*-nearest neighbor

References
1. Y Gu, ACC Lo, IG Niemegeers, A survey of indoor positioning systems for wireless personal networks. IEEE Commun. Surv. Tutorials. **11**(1), 13–32 (2009). https://doi.org/10.1109/SURV.2009.090103
2. QD Vo, P De, A survey of fingerprint-based outdoor localization. IEEE Commun. Surv. Tutorials. **18**(1), 491–506 (2016). https://doi.org/10.1109/COMST.2015.2448632
3. X Wang, L Gao, S Mao, S Pandey, CSI-based fingerprinting for indoor localization: A deep learning approach. IEEE Trans. Veh. Technol. **66**(1), 763–776 (2017)
4. P Bahl, VN Padmanabhan, in *Proceedings IEEE INFOCOM 2000, The Conference on Computer Communications, Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Reaching the Promised Land of Communications, Tel Aviv, Israel, March 26-30*. RADAR: an in-building RF-based user location and tracking system (IEEE, 2000), pp. 775–784
5. K Wu, J Xiao, Y Yi, D Chen, X Luo, LM Ni, CSI-based indoor localization. IEEE Trans. Parallel Dist. Syst. **24**(7), 1300–1309 (2013)
6. J Xiao, K Wu, Y Yi, LM Ni, in *21st International Conference on Computer Communications and Networks, ICCCN 2012, Munich, Germany, July 30 - August 2, 2012*. FIFS: Fine-grained indoor fingerprinting system (IEEE, 2012), pp. 1–7
7. M Ibrahim, M Youssef, CellSense: An accurate energy-efficient GSM positioning system. IEEE Trans. Veh. Technol. **61**(1), 286–296 (2012). https://doi.org/10.1109/TVT.2011.2173771
8. J Borkowski, J Lempiäinen, in *11th European Wireless Conference 2005 - Next Generation Wireless and Mobile Communications and Services, Nicosia, Cyprus, 10-13 April*. Pilot correlation positioning method for urban UMTS networks (VDE, 2005), pp. 1–5
9. J Turkka, T Hiltunen, R Mondal, T Ristaniemi, in *2015 International Symposium on Wireless Communication Systems (ISWCS), Brussels, Belgium, August 25-28*. Performance evaluation of LTE radio fingerprinting using field measurements (IEEE, 2015), pp. 466–470
10. T Wigren, in *Proceedings of the 76th IEEE Vehicular Technology Conference, VTC Fall 2012, Quebec City, QC, Canada, September 3-6*. LTE fingerprinting localization with altitude (IEEE, 2012), pp. 1–5
11. B Yu, B Shin, J Lee, S Lee, in *Proceedings of the 30th International Technical Meeting of The Satellite Division of the Institute of Navigation ION GNSS+, Portland, Oregon*. Localization in urban canyon: Machine learning based localization using LTE or LoRa signal for 'GNSS-denied' areas, (2017), pp. 456–462
12. S He, SHG Chan, Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. IEEE Commun. Surv. Tutorials. **18**(1), 466–490 (2016)
13. Z Belhadi, L Fergani, in *2014 International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, Morocco, April 14-16*. Fingerprinting methods for RFID tag indoor localization (IEEE, 2014), pp. 717–722
14. V Nambiar, E Vattapparamban, AI Yurekli, I Guvenç, M Mozaffari, W Saad, in *International Conference on Computing, Networking and Communications,*
ICNC 2017, Silicon Valley, CA, USA, January 26-29. SDR based indoor localization using ambient WiFi and GSM signals (IEEE, 2017), pp. 952–957
15. PH Tseng, YC Chan, YJ Lin, DB Lin, N Wu, TM Wang, Ray-tracing-assisted fingerprinting based on channel impulse response measurement for indoor positioning. IEEE Trans. Instrum. Meas. **66**(5), 1032–1045 (2017)
16. ZH Wu, Y Han, Y Chen, KJR Liu, A time-reversal paradigm for indoor positioning system. IEEE Trans. Veh. Technol. **64**(4), 1331–1339 (2015)
17. B Song, S Zhang, J Long, Q Hu, Fingerprinting localization method based on TOA and particle filtering for mines. **2017**(Article ID 3215978), 1–10 (2017)
18. T Hiltunen, J Turkka, R Mondal, T Ristaniemi, in *10th International Conference on Information, Communications and Signal Processing, ICICS 2015, Singapore, December 2-4*. Performance evaluation of LTE radio fingerprint positioning with timing advancing (IEEE, 2015), pp. 1–5
19. T Wigren, Y Jading, I Siomina, A Kangas, C Tidestav, in *Proceedings of the 76th IEEE Vehicular Technology Conference, VTC Fall 2012, Quebec City, QC, Canada, September 3-6*. Enhanced WCDMA fingerprinting localization using OTDOA positioning measurements from LTE (IEEE, 2012), pp. 1–5
20. M Driusso, C Marshall, M Sabathy, F Knutti, H Mathis, F Babich, in *International Conference on Indoor Positioning and Indoor Navigation, IPIN 2016, Alcala de Henares, Spain, October 4-7*. Indoor positioning using LTE signals (IEEE, 2016), pp. 1–8
21. Q Liu, R Hu, S Liu, A wireless location system in LTE networks. Hindawi Mob. Inf. Syst. **2017**(Article ID 6160489), 1–11 (2017). https://doi.org/10.1155/2017/6160489
22. T Wigren, in *5th International Conference on Signal Processing and Communication Systems, ICSPCS 2011, Honolulu, HI, USA, December 12-14*. Clustering and polygon merging algorithms for fingerprinting positioning in LTE (IEEE, 2011), pp. 1–10
23. N Poosamani, I Rhee, in *IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops 2015, St. Louis, MO, USA, March 23-27*. Towards a practical indoor location matching system using 4G LTE PHY layer information (IEEE, 2015), pp. 284–287
24. X Ye, X Yin, X Cai, AP Yuste, H Xu, Neural-network-assisted UE localization using radio-channel fingerprints in LTE networks. IEEE Access. **5**, 12071–12087 (2017)
25. G Pecoraro, S Di Domenico, E Cianca, M De Sanctis, in *13th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob 2017, Rome, Italy, October 9-11*. LTE signal fingerprinting localization based on CSI (IEEE, 2017)
26. S Di Domenico, G Pecoraro, E Cianca, M De Sanctis, in *12th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob 2016, New York, NY, USA, October 17-19*. Trained-once device-free crowd counting and occupancy estimation using wifi: A doppler spectrum based approach (IEEE, 2016), pp. 1–8
27. S Ezpeleta, JM Claver, JJP Solano, JV Martí, RF-based location using interpolation functions to reduce fingerprint mapping. Sensors. **15**(10), 27322–27340 (2015). https://doi.org/10.3390/s151027322
28. Y Chapre, A Ignjatovic, A Seneviratne, SK Jha, in *IEEE 39th Conference on Local Computer Networks, LCN 2014, Edmonton, AB, Canada, 8-11 September*. CSI-MIMO: indoor Wi-Fi fingerprinting system (IEEE, 2014), pp. 202–209