Digital Communications and Networks 9 (2023) 175–182

Contents lists available at ScienceDirect

Digital Communications and Networks

journal homepage: www.keaipublishing.com/dcan

A WKNN-based approach for NB-IoT sensors localization

Ennio Gambi [a], Linda Senigagliesi [a,*], Andrea Barbaresi [b], Matteo Mellini [b], Adelmo De Santis [a]

[a] Department of Information Engineering, Università Politecnica delle Marche, 60131, Ancona, Italy
[b] TIM S.p.A. - Business - Pre Sales Engineer, 60100, Ancona, Italy

A R T I C L E I N F O

Keywords:
Narrowband Internet of Things (NB-IoT)
Sensor localization
Sensor networks
RF fingerprint Positioning
LTE

A B S T R A C T

With the recent introduction of NarrowBand Internet of Things (NB-IoT) technology in the 4th and 5th generations of mobile radio networks, the mobile communications context opens up significantly to the world of sensors. By means of NB-IoT, the mobile systems within 3GPP standardization introduce the peculiar functions of sensor networks, thus making it possible to satisfy very specific requirements with respect to those which characterize traditional mobile telecommunications. Among the functions of interest for sensor networks, the possibility of locating the positions of the sensors without an increase in costs and energy consumption of the sensor nodes is of utmost interest. The present work describes a procedure for locating the NB-IoT nodes based on the quality of radio signals received by the mobile terminals, which therefore does not require further hardware implementations on board the nodes. This procedure, based on the RF fingerprinting technique and on machine learning processing, has been tested experimentally and has achieved interesting performances.

1. Introduction

Nowadays the Internet of Things (IoT), which consists in using the Internet to collect data from a multiplicity of sensors scattered over a territory, is gaining more and more attention. Wireless sensor networks for smart metering have been implemented [1], together with solutions which are able to bring the content closer to the end users in wireless networks, such as caching files in densely deployed small-cell Base Stations (BSs) with a large storage capacity [2]. The NarrowBand IoT (NB-IoT) is a radio technology introduced within the 4G Long Term Evolution (LTE) and 5G mobile networks, specifically designed to allow sensors to transmit data to a server in the cloud over the Internet by relying on the extensive coverage offered by those networks. The NB-IoT technology (aka LTE Cat-NB) has recently been made operational by various mobile network operators across the world and, thanks to its specifications, it allows the radio signal to be received even in environments where other currently widespread wireless technologies would not be able to provide any data connection service.

The development of geolocation techniques alternative to the use of the Global Navigation Satellite System (GNSS) systems is a very interesting and practical research area in the IoT context. In fact, the GNSS chipset is not often present in the IoT modules commercially available on the market, primarily because these modules must have both an extremely low cost and the lowest possible energy consumption. A further reason is that the location of devices makes using GNSS systems not possible in all conditions due to the attenuation of satellite signals (e.g., in typical indoor environments).

The main objective of our work is to assess the level of accuracy that can be achieved by an NB-IoT network in operation as far as a geolocation algorithm based on the RF fingerprinting technique is concerned. The considered location method is based on the K-Nearest Neighbor (KNN) algorithm and the outdoor measurement environment consists of a portion of the commercial NB-IoT network of Telecom Italia in the 800 MHz band.

At the moment, there are only few studies in the literature regarding the level of accuracy of methods for geolocating NB-IoT devices in real scenarios, as NB-IoT is a relatively recent technology. Nevertheless, there are several papers concerning this topic for the LTE technology. This leads to the opportunity to perform a comparative analysis on the results achievable by the RF fingerprinting methods specifically within the IoT context.

The rest of the paper is organized as follows. In Section 2 the state-of-the-art studies are analyzed. Section 3 describes the NB-IoT technology and its parameters of interest for the present research. Section 4 introduces the different approaches of geolocating a sensor device. Section 5 explains the experimental setup that was implemented to obtain the

* Corresponding author.
E-mail addresses: e.gambi@univpm.it (E. Gambi), l.senigagliesi@staff.univpm.it (L. Senigagliesi), andrea.barbaresi@telecomitalia.it (A. Barbaresi), matteo.mellini@telecomitalia.it (M. Mellini), adelmo.desantis@staff.univpm.it (A. De Santis).

https://doi.org/10.1016/j.dcan.2022.04.033
Received 19 October 2020; Received in revised form 26 April 2022; Accepted 28 April 2022
Available online 13 May 2022

position data and the RF levels processed by the location algorithm are described in Section 6. Section 7 shows the results of the measurement campaign and the localization performance of the considered algorithm. Finally, conclusions are drawn in Section 8.

2. Related works

NB-IoT is a relatively recent technology, and advances in the field of sensor location is a new research branch in which little has been written so far. The LTE location is well consolidated and the literature in this field can be used as a reference for our present work. In Ref. [3], the authors proposed a study for the geolocation of LTE devices based on the fingerprinting method, which exploits the Channel State Information (CSI) and considers only simulations done in an indoor scenario. In Ref. [4], an RF fingerprinting method for LTE is described, using the Minimization of Drive Testing (MDT) technique and the simulation on 3 artificial scenarios: rural, urban and hetnet. Mahalanobis Distance (MaD) method and the Kullback-Leibler Divergence (KLD) method are used. In the urban case, the best accuracy is obtained by using the KLD algorithm with an error of up to 38 m in 68% of the cases.

A precise radio map and application of indoor positioning with the dual-frequency Wi-Fi fingerprinting method is proposed in Ref. [5]. In Ref. [6], an RF fingerprinting algorithm is evaluated on a commercial LTE network in Finland, using the MDT method. Similar to the scenario considered in this paper, measurements related to an outdoor environment are considered, the range is divided into smaller areas, and a commercial LTE network is used. In particular, outdoor measurements are made on the 800 MHz LTE band using a mobile phone. The results obtained show that, by dividing the entire measurement area into 20 m $\times$ 20 m smaller areas, the accuracy is within 182 m for 68% of the cases. A development of the work of [6] is considered in Ref. [7], where the knowledge of Timing Advancing (TA) information is used together with the radio fingerprinting to improve the position estimation, and a set of measures in a real scenario are made.

A different approach for the localization of an LTE mobile station has been recently presented in Ref. [8]. It is based on the fusion of one Round Trip Time (RTT) observation associated with a serving BS with the Time-Difference Of Arrival (TDOA) observations associated to the serving and neighboring BSs. The proposed approach achieves good performance in terms of Root Mean Square Error (RMSE) for both simulated networks and real-field experiments.

An NB-IoT development platform and indoor deployment coverage is analyzed in Ref. [9]. An indoor location method for NB-IoT based on CSI is proposed in Ref. [10], where a method based on the KNN algorithm is used to estimate the target position of the sensor. Different from our proposal, the authors consider an indoor environment where the CSI, which usually is not immediately available for the IoT devices, is exploited for location. Moreover, the transmitting stations are in known positions and the use of a commercial network is not considered.

In [11], a general overview of location methods is proposed by 3GPP for LTE-M and NB-IoT technologies. In particular, the location method based on the Observed Time Difference of Arrival (OTDOA) is described. Results concerning the NB-IoT technology show that, when the outdoor scenario is addressed, an error of up to 77 m is obtained in about 70% of cases.

[12] proposes a method for improving Received Signal Strength Indication (RSSI) ranging accuracy in Long Range (LoRa) systems by using a Wiener-based method, aiming at minimizing the distance logarithm error derived from the Friis path loss model equation. In Ref. [13], a RSSI-based fingerprinting localization of sensors nodes is shown for the ultra narrow band Sigfox IoT networks.

In summary, from the analysis of the state-of-the-art studies we can conclude that few references address the accuracy of localization of NB-IoT sensors by means of on-field measurements in a real scenario. Most of the proposed localization methods are evaluated through simulations with respect to ideal scenarios. Moreover, in most cases, the emphasis is given to the peculiarity of the proposed algorithms with respect to the performance obtainable in a real scenario by an algorithm belonging to a general class of localization methods, just as the one considered in the present work. A further consideration regards the fact that, in some cases, the considered algorithms make use of radio parameters that are not easily accessible from the NB-IoT chipsets used on the sensors, such as the timing advance and the channel state information. Conversely, in this paper, only the Reference Signal Received Power (RSRP) levels and the Physical Cell IDentifiers (PCI) are used, which can be easily obtained from the RF chipset.

3. Narrowband IoT

NB-IoT is a cellular technology introduced in 3GPP Release 13 to provide wide area coverage for the IoTs [14]. The normative phase of the NB-IoT work item within 3GPP started in September 2015 and the core specification ended in June 2016 [14]. A single NB-IoT carrier consists of a 180 kHz band for both the downlink and the uplink. This is equivalent to the bandwidth of one Physical Resource Block (PRB) in the LTE [15]. This band can also be taken within the GSM band, occupying a 200 kHz GSM channel [16].

The NB-IoT reuses the design of the LTE, including the downlink Orthogonal Frequency-Division Multiple-Access (OFDMA), the uplink Single-Carrier Frequency-Division Multiple-Access (SC-FDMA), channel coding, interleaving and so on [14]. The NB-IoT technology opens the way to 5G massive Machine Type Communications (mMTC) and through a standard developed specifically for using the features of the Internet of Things and savings in consumption, allows a long life of the batteries of the connected objects.

Thanks to a greater coverage capacity and a low energy consumption required from the connected devices, NB-IoT allows objects that previously could not be connected to communicate with each other. The uplink transmission uses single-tone and multi-tone transmissions. For the single-tone transmission, 3.75 kHz and 15 kHz channels are supported. For the multi-tone transmission, the SC-FDMA scheme with 15 kHz subcarriers spacing is used [16]. The spectrum usage can be triggered in three different modes: stand-alone, guard-band and in-band [17], the latter being the one used in this article and adopted by Telecom Italia - TIM in its network. It uses one or more resource blocks of 180 kHz wide, which are allocated in the same band of LTE [17].

NB-IoT uses some radio parameters to trigger actions and to give a mean to study the available coverage. RSRP is defined as the average power of Resource Elements (RE) which transport the cell-specific Reference Signals (RS) [18]. The metric chosen by 3GPP to evaluate the coverage is the Maximum Coupling Loss (MCL). It represents the maximum attenuation that the system can support on the transmission channel (Eq. (1)) and can be written as

$$MCL_{dB} = P_{tx} - (Noise_{figure} + SINR + Noise_{floor}) \quad (1)$$

One of the key features of NB-IoT lies in its ability to deliver the service coverage even toward the sensors installed behind the metallic sheets or underground. This represents a great improvement over the GSM technology: more than 20 dB increase in MCL, with a target value of 164 dB [17]. The great coverage value is obtained without the need for a highpower transmitter. The sensors' output power is 23 dBm or 20 dBm, thus allowing for the power amplifier to be integrated in a System on a Chip (SoC) [19]. To reach such a good coverage performance, it is possible to use a narrow-band channel for the uplink transmission of 3.75 kHz or 15 kHz instead of the default bandwidth of 180 kHz [19].

Moreover, the Transport Block (TB) can be retransmitted up to 2048 times in the downlink and 128 times in the uplink [19], thus allowing for the Coverage Enhancement (CE) implementation. The NB-IoT technology allows for a very high number of sensors (more than 50000) to be handled by a single sector of each eNodeB on a 180 kHz channel, with a minimum throughput of 160 bit/s for each sensor [17]. NB-IoT is energy efficient and should allow a very long battery life, which is estimated to be 6+ years [19]. Other specific NB-IoT features, such as Extended Discontinuous Reception (eDRX) and Power Saving Mode (PSM), help to extend the battery life even more [20].

4. The proposed geolocation approach

Geolocation is the process by which the geographic position of an object is determined by exploiting satellite positioning systems or mobile networks radio signals (radiolocation). This article focuses on a positioning method in the seoncd context. The User Equipment (UE) sensor location can be carried out in two different modalities, namely, "network-based" or "handset mode". The first requires to process the radio measurements and signaling data received by the BSs, which can be accessed only by the network operator and allows a great precision in urban areas. The handset mode approach requires to install a minimal software component on the UE to track the level of signals received either from the serving cell (i.e., the cell used by the UE to transmit data to the network) or from the neighboring cells (i.e., the other cells measured by the UE in order to perform a serving cell reselection, when needed). The fingerprinting-based method, which belongs to the "handset mode" approach, is the one chosen in the current work. It relies on the creation of "RF power fingerprints" of the radio signals transmitted by the BSs, which can be received by the UE. The reason for this choice is because the fingerprinting-based approach does not rely on network-based data which are typically available only to the network operator, nor on other privileged information, so this approach is considered more preferable than other ones.

The RF fingerprinting method is a database correlation method that can be used to estimate the UE position [4]. It requires a radio environment map, which is created by measuring the UE's received power levels of the signals transmitted by different BSs and by varying the UE position. In order to take into consideration the non-stationary and time-variant environment in which the UE lies or moves, this approach requires a set of measures that must be repeated periodically. Following the approach presented in Ref. [21], the fingerprinting location method requires a "training phase" and a "positioning phase". On the basis of this approach, a database is populated in the first phase of our work, which contains all the radio power measures taken on a set of known geographic locations. In the positioning phase, the current power measures received by the UE located in an unknown position are transmitted to a server, then an algorithm is applied to select the position in the database which best matches the measured data. In order to compare the measurements transmitted by the UE with the ones saved in the central database, the KNN or the maximum likelihood algorithms are often used. As a result of this approach, it is not mandatory to know the exact geographic position of the BSs involved in the location process (which in most cases cannot be considered as public available).

5. Experimental setup

Our experimental tests involve the use of a couple of development boards. The SODAQ Sara Arduino Form Factor (AFF) N211 [22] is an Arduino compatible board which supports the NB-IoT Radio Access Technology (RAT) communication thanks to the uBLOX N211 module installed, which allows the access to NB-IoT bands 8 and 20. The SODAQ board is also equipped with a variety of sensors like an accelerometer, a magnetometer and a Global Positioning System (GPS) receiver. The core MCU is Atmel SAMD21. The uBLOX module supports standard AT commands as well as manufacturer-specific extended AT commands, by which the user can also have access to several internal radio parameters of the module.

The Wislink Cellular BG96 from RAK Wireless is an Arduino shield built around the Quectel BG96 module. It supports multi-RAT 2G, 4G/LTE, LTE Cat M1 and NB-IoT operations and has a GPS receiver onboard. Also the BG96 module supports extended AT commands in addition to the standard ones, by which it is possible to have access to several internal parameters. The Wislink shield can be used either in conjunction with an Arduino UNO board or in a stand-alone configuration.

On the network side, the "All Things Talk" cloud platform, which acts as a data collector for IoT devices, is adopted. For the aims of the experimental tests, measurement data from the SODAQ board are stored on this platform.

The first developed firmware for the SODAQ board has been a plain "pass through" which enables the AT commands written on the serial line to be conveyed directly to the uBLOX module. In this way, the module has been initially configured and tested to check which kind of data can be retrieved from it. To this aim, by issuing the 'AT+NEUSTATS = "ALL"' command, the radio parameters monitored by the RF chipset for all the received cells are shown as follows:

- NUESTATS: "APPSMEM", "Num Allocs":88;
- NUESTATS: "APPSMEM", "Num Frees":21;
- NUESTATS: "CELL", 6290,234,1,-1093,-120,-1004,51;
- NUESTATS: "CELL", 6290,156,0,-1099,-108,-1464,31;
- NUESTATS: "CELL", 6290,472,0,-942,-128,-1351,40;
- OK.

Parameters identified by the "CELL" keyword in the string formatted as "CELL",p1,p2,p3,p4,p5,p6,p7 are very important within the scope of this work, which have the following meanings:

- p1: EARFCN (E-UTRA Absolute Radio Frequency Channel Number);
- p2: Physical Cell ID (Cell IDentifier);
- p3: serving cell (1) or neighbor cell (0) indication;
- p4: RSRP (Reference Signals Received Power) expressed in tenth of dBm (−1093 is equivalent to −109.3 dBm);
- p5: RSRQ (Reference Signals Received Quality) expressed in tenth of dB (−120 is equivalent to −12 dB);
- p6: RSSI (Received Signal Strength Indication) expressed in tenth of dBm (−1004 is equivalent to −100.4 dBm);
- p7: SNR (Signal-to-Noise Ratio) expressed in tenth of dB (51 is equivalent to 5.1 dB).

Also in the case of the Wislink cellular shield, radio parameters monitored by the RF chipset can be accessed by issuing the "Engineering Mode" via the AT+QENG command: RSRP, RSRQ, RSSI and SINR measurements related to the serving cell (AT+QENG = "servingcell") as well as the ones related to the neighbor cells (AT+QENG = "neighbor") can be retrieved in this way. The format of the output produced by these command is described in Ref. [23].

According to the technical specifications of the two modules, uBLOX N211 is based on Huawei HiSilicon Neul Hi2110 RF chipset whereas Quectel BG96 is based on the MDM9206 RF chipset made by Qualcomm. Within the scope of the present work, this means that the achieved results in terms of practical feasibility to obtain the radio measurements of interest for the location algorithm via the AT commands reported above can be considered valid also for many other NB-IoT modules which are based on these chipsets.

The final version of the firmware running on the SODAQ board has been used in the measurement campaign to store data on the basis of the equipment's current position. When the board is powered on, both the position from the GPS chipset and the radio parameters from the NB-IoT RF chipset are acquired and sent to the cloud platform. Until the board position does not change, radio parameters are aggregated by PCI and their average values are calculated. As soon as the board moves and the distance from the previous recorded position becomes greater than 20 m, new data are acquired and sent to the cloud. For each position, all the received PCIs and the related radio parameters are recorded.

A measurement campaign has been carried out, in which the SODAQ board with the developed firmware is used to collect the RF measurements in a real scenario. Tests show that collecting data through an IoT platform is a good choice when a great number of stationary sensors are present in the area of interest and when real-time processing of data is not mandatory.

Once the practical feasibility to populate the RF fingerprint database by collecting the measurements performed by the UE via the manufacturer-specific AT commands and by transmitting the data to the cloud platform using the NB-IoT network is validated, the second board is used. In fact, the Quectel BG96 module let the user access and monitor radio-parameters in real time by using a dedicated proprietary debugging interface called Qualcomm Diagnostic Interface (DIAG). This interface allows for a direct connection to the NB-IoT Qualcomm chipset using a USB interface: the diagnostic data can be configured and retrieved in real-time using Qualcomm's Qxdm proprietary software as well as open source tools like diag-parser [24], which is able to parse the 2G, 3G and 4G radio messages in DIAG format and to covert them to Osmocom GSMTAP [25] for analysis in wireshark and other utilities.

This way of accessing the internal data of the Qualcomm RF chipset can be considered as an alternative (much more complete and powerful) to the use of manufacturer specific AT commands via the serial interface. For this practical reason, within the scope of this work, the Diag interface has been exploited to collect the radio measurements while moving to the area of interest and to save the data directly to a personal computer.

The Wislink board has an external GPS input too, which allows for georeferencing of measurements samples taken during the first training phase. Tests were taken on a 4.5 km^2 wide area, as shown in Fig. 1.

The data are organized by using a spreadsheet in which the columns represent the following data (see Fig. 2):

- N – index of the i-*th* measurement point;
- Latitude and Longitude coordinates – the position where the measurement sample is taken;
- LTE-NB Physical Cell ID - PCI;
- LTE-NB RSRP RSRQ RSSI - Radio parameters.

The data are converted into the CSV format to be easily imported and processed in Matlab. Latitude and longitude data from the first row (first sampled data) are extracted and taken as a reference point. Then the next row is analyzed: if the distance from the former point is less than 20 m, the PCI value is checked. If a measure for the same PCI is already present, the average of the radio-parameter data is calculated. Once all the rows in the CSV file have been processed, the data are organized in a matrix which contains only values taken at 20 m distance. The parameter *Counter* takes into account the recorded number of samples with the same PCI. Table 1 reports the average values of RSRP calculated over different measurements which have the same coordinates and the same PCI within a distance of 20 m. As an example, for the coordinates in the first row of Table 1, measurements (see column $Count_1$) with PCI 234 are collected, 10 ($Count_2$) with PCI 156, 8 ($Count_3$) with PCI 472, etc. The same

Fig. 1. Mapped area.

operation is performed for the other radio parameters, such as the Reference Signal Received Quality (RSRQ) and the Received Signal Strength Indication (RSSI). Fig. 3 shows the plot of the processed data on the Google Earth platform, using “20 m” clusters. Fig. 4 shows a detailed image of the points in which data are gathered.

By sampling the cell radio-parameters on a 20 m basis, there is a great probability to gather data for multiple PCIs in a single measurement. Different BSs have unique power parameters in a given position. Therefore, the more PCI data are gathered, the more accurate the position estimation can be. Results achieved by our experimental setup demonstrate the practical feasibility of accessing RSRP, RSRQ, RSSI measurements of the two NB-IoT chipsets under investigation, to populate the RF fingerprinting database needed by the location algorithm. It is worth to note that, according to 3GPP specifications [26], it is mandatory for the UE to continuously perform the measurements of these radio parameters to correctly manage the mobility procedures of cell selection in the idle mode and cell reselection in the connected mode.

6. NB-IoT location algorithm

In this article, some novelties are introduced, using the RSRP parameters and PCI as a basis of using the Weighted K-Nearest Neighbor (WKNN) algorithm. KNN turns out to be the most often used algorithm for RF fingerprinting in literature [27–29], despite its simplicity and limitations that make it necessary to adjust it to increase its performance. The good performance of KNN is highly dependent on the metric used for computing the pair ware distances between data points [30]. Algorithms more complex than the KNN have been used for WiFi fingerprinting, such as the principal component analysis and the support vector machines. However, the estimation accuracy of such algorithms is comparable to that of KNN [31]. The radio parameters are acquired in a real outdoor scenario, thus giving results that are immediately useable for the deployment and provisioning of sensors.

Once the dataset containing RF-fingerprints is created, it is processed with the WKNN algorithm to estimate the current sensor position. The WKNN algorithm evaluates the Euclidean Distance (ED) between a reference value and all the other entries in a dataset. In the present work, an RF-fingerprint is a set of RSRP values measured by the sensor on the main BS and neighbor BSs in a given position. In order to geolocate a sensor, the measured RF-fingerprint needs to be classified. Each entry in the dataset is considered as a separate class, so no classes are with multiple RF-fingerprints. This requires a different approach in using KNN: once K closest points are calculated, the weighted average of their positions is determined and used to geolocate the unknown point. Weights are determined based on the metric defined according to the criterion: *the closer the distance, the bigger the associated weight.*

The localization method proposed here requires a pre-processing of the measurements values before classification, as is described in Section 5 (as shown in Table 1 and related comments). The position is determined by applying a KNN algorithm which, on the basis of the defined distance (which can be, for example, the ED), finds the K clusters with the shortest distance from the identified PCIs. Latitude and longitude values are then weighed according to the distance values. A more detailed description of the algorithm is provided in the following subsection. Thus the proposed approach is not based on a classification in the strict sense, but rather on the application of a WKNN algorithm for the comparison of the distance values.

Algorithm details

The sensor's position is extracted from a collection of radio-parameters measured at an unknown geographical point. KNN requires to evaluate the ED between the sensor data and each entry in the dataset. Hence, three cases are possible:

N	Latitude	Longitude	PCI	RSRP	RSRQ	RSSI
1	43,5671698	13,5133874	234	-94.0	-14.3	-79.7
2	43,5671698	13,5133874	234	-88.4	-15.6	-72.8
3	43,5671698	13,5133873	156	-80.1	-9.6	-70.5
4	43,5671697	13,5133873	234	-91.9	-14.3	-77.6
5	43,5671697	13,5133872	234	-91.9	-14.3	-77.6
6	43,5671697	13,5133871	156	-80.3	-1.9	-78.4
7	43,5671696	13,513387	234	-88.1	-15.7	-72.4
8	43,5671696	13,5133869	156	-81.2	-3.6	-77.6
9	43,5671695	13,5133868	234	-88.9	-17.0	-71.9
...	...	...	...	...	...	...

Fig. 2. Excerpt of data organized in a spreadsheet.

Table 1
Excerpt of data processing results.

Latitude	Longitude	$Count_1$	PCI_1	$RSRP_1$	$Count_2$	PCI_2	$RSRP_2$	$Count_3$	PCI_3	$RSRP_3$	…
43.5672	13.5134	14	234	−90.4643	10	156	−82.5400	8	472	−91.3375	…
43.5670	13.5135	4	156	−85.1250	1	33	−93.9000	3	254	−90.4667	…
43.5666	13.5138	3	156	−77.0333	2	33	−90.8000	2	234	−84.8000	…
43.5664	13.5139	3	156	−82.7000	1	33	−85.5000	1	252	−90.2000	…
43.5662	13.5140	4	472	−92.0750	2	156	−76.2500	1	33	−75.3000	…
43.5661	13.5141	4	156	−80.8750	2	234	−88.3500	1	33	−79.8000	…
43.5659	13.5143	4	234	−86.4000	2	156	−84.4500	1	33	−91.9000	…
43.5658	13.5145	3	156	−79.3333	3	33	−91.4333	3	234	−80.8000	…

Fig. 3. 20 m clusters.

Fig. 4. Detailed image of some 20 m clusters.

- Sensor and dataset RF-fingerprints contain parameters from the same PCIs;
- Sensor RF-fingerprints data contains PCIs that are not present in the dataset;
- Dataset RF-fingerprints contain PCI that are not present in sensor data.

In the first case, the ED is easily obtained with the following equation:

$$D = \sqrt{(RSRP_{1,r} - RSRP_{1,u})^2 + (RSRP_{2,r} - RSRP_{2,u})^2} \quad (2)$$

where subscript u marks the RSRP value at the sensor and subscript r denotes the RSRP at the reference entry.

In the other two cases, the missing PCI has a power level below a sensitivity threshold (THR). As an example, with reference to Table 2, let us assume a case in which the reference entry and the sensor entry do not share any PCI:

$$D_{direct} = (RSRP_{1,r} - THR)^2 + (RSRP_{2,r} - THR)^2 \quad (3)$$

$$D_{reverse} = (RSRP_{3,u} - THR)^2 + (RSRP_{4,u} - THR)^2 \quad (4)$$

The total ED is calculated from Eq. (3) and Eq. (4) as

$$D_{tot} = \sqrt{D_{direct} + D_{reverse}} \quad (5)$$

Once the ED is calculated, we choose the K nearest point in the dataset and calculate a weighted average of their geographic positions. Weights are evaluated based on the ED

Table 2
Euclidean Distance for different PCIs.

Data	RSRP	PCI	RSRP	PCI
Dataset value	$RSRP_{1,r}$	$PCI_{1,r}$	$RSRP_{2,r}$	$PCI_{2,r}$
Sensor value	$RSRP_{3,u}$	$PCI_{3,u}$	$RSRP_{4,u}$	$PCI_{4,u}$

$$p_i = \frac{1 - \frac{D_i}{\sum_{j=1}^{K} D_j}}{K-1} \quad (6)$$

Latitude and longitude estimates are obtained as

$$Latitude_{es} = \sum_{i=1}^{K} p_i \times latitude_i \quad (7)$$

$$Longitude_{es} = \sum_{i=1}^{K} p_i \times longitude_i \quad (8)$$

Once the sensor position is determined, the estimated position error is calculated. Let "minlat" and "maxlat", "minlon" and "maxlon" respectively be the smallest and the greatest latitude and longitude values of the sensor's closest points in the dataset. By using the "spherical law cosine" formula, we can determine the estimate error as

$$\begin{aligned} D_{error} = \ & \arccos(\sin(minlat) \times \sin(maxlat) + \\ & + \cos(minlat) \times \cos(maxlat) \times \\ & \times \cos(minlat - minlon)) \times 6371 \end{aligned} \quad (9)$$

7. Experimental results and performance evaluation

In order to process a large number of measurements, a Matlab script is developed, which allows for many parameters to be changed in order to evaluate their effect on the accuracy of the algorithm.

The dataset which collects the measurements made during the experiments has a dimension of 2374. Almost all such measurements refer to different positions, as already shown in Fig. 2. The data is pre-processed before classification according to the procedure described in Section 5. In such a way, the number of "20 m" clusters is 353 (while, for example, the dataset dimension is 483 when the measurements within a distance of 10 m are grouped, and becomes 215 for 40 m).

First, tests are performed by varying the K values from 1 to 4. Let us recall that varying K influences the number of dataset measurements the algorithm takes to estimate the sensors' positions. By comparing the results for different values of K, we can see that the best choice is $K = 2$ as shown in Fig. 5, which leads to an average error of 91 m. Taking into account the Cumulative Distribution Function (CDF) in Fig. 6, the error for the best case ($K = 2$) is below 109 m for 70% of the total cases. Increasing the K value leads to worst performances as the algorithm has to cope with data that has a bigger ED from the reference point. Being an NB-IoT narrowband signal, it is greatly affected by frequency fading, which may cause variations of the reference signal's amplitude even for short displacement of the sensor. The LTE nodes are less affected by this problem for this technology is based on wideband signals. Nevertheless, the accuracy values obtained in our analysis are overall comparable with results that can be found in literature (e.g. Ref. [6]). In order to have a fair comparison with the results obtained in our work, we consider only rows with LTE 800 MHz in Table 1 of [6]. It is evident that the Positioning Error (PE), even with the best %-ile (PE of 182 m in the 68%-ile), is larger than the one obtained by the NB-IoT approach based on the WKNN algorithm, which achieves a PE of 109 m for 70% of the cases.

A further test-set involves a different choice of the threshold value, by setting it at −140 dBm and varying the K values. Results are shown in Fig. 7. The algorithm performances are slightly better, with a 1% increase in performances for the average error in the cases of $K = 2$ and $K = 4$.

One more test-set is made by varying the cluster size for averaging the measurement data. The cluster size is varied from 10 m × 10 m–40 m × 40 m, and the cumulative results are shown in Fig. 8.

The effect of the weight values is assessed by executing the algorithm and setting the weight value to 1. By this setup, the sensor's position is estimated using the KNN algorithm. The RF fingerprints in the dataset

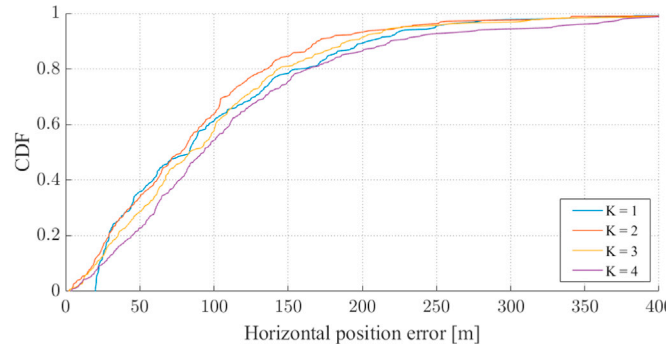

Fig. 6. CDF of error by varying K.

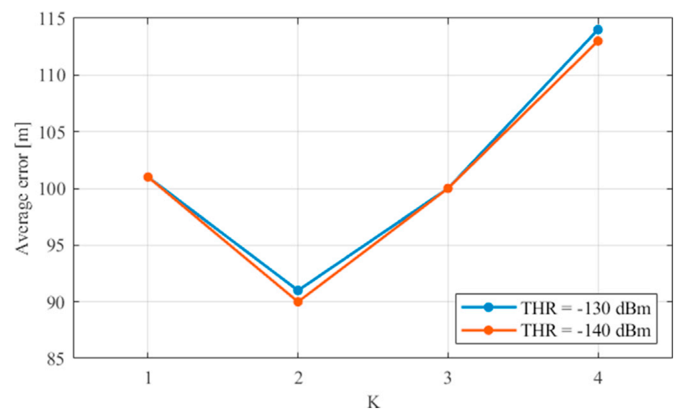

Fig. 7. Average error by varying K, with THR = −130 dBm and THR = −140 dBm and cluster 20 m × 20 m.

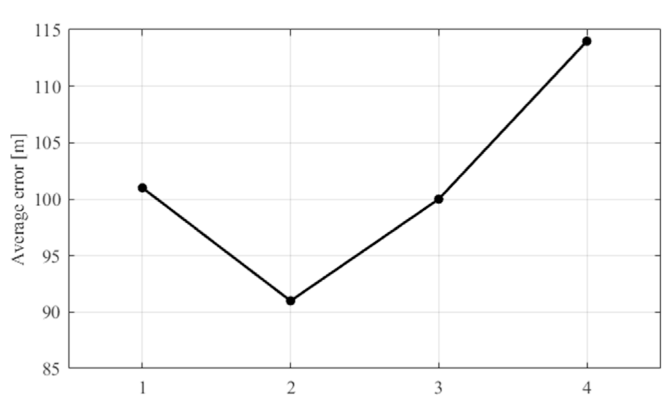

Fig. 5. Average error by varying K, with THR = −130 dBm and cluster 20 m × 20 m.

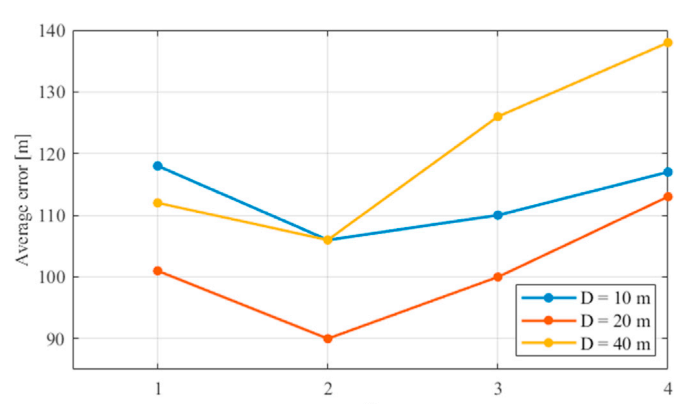

Fig. 8. Average error by varying K, with THR = −140 dBm. Cluster size variation.

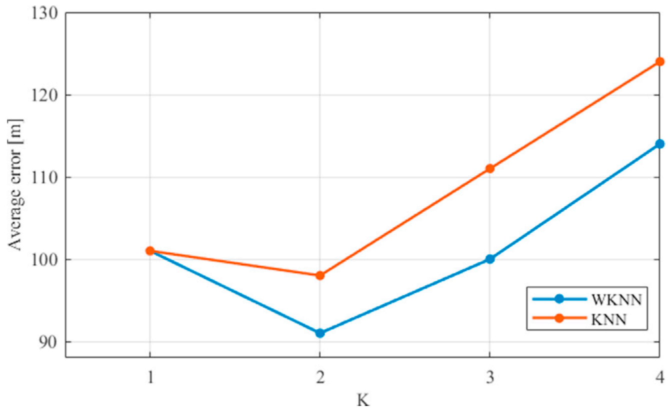

Fig. 9. Average error by varying K, with THR = −130 dBm, cluster 20 m × 20 m. KNN vs WKNN.

that are closer to the sensor's reported values are not given a big weight during the position estimation. Results show that the KNN approach is more error prone than WKNN, as is shown in Fig. 9.

The last test set focuses on the analysis of sensor measures that contain only one PCI versus other PCIs. The idea behind the test is to verify that the more PCIs are present in the sensor RF-fingerprint, the bigger is the accuracy. The tests do not confirm this assumption, as is shown in Fig. 10. An explanation for this is that, in the 3GPP standard [26], it is stated that when a BS signal is strong and exceeds a certain threshold, the UE can stop looking for neighbor BSs. While the signal being so strong, the interference and noise levels are low, thus leading to a great accuracy in position estimate.

Test results are also analyzed using Google Earth, by plotting the position of the reference row in the dataset, the nearest points and the sensor position estimate. An example is given in Fig. 11.

8. Conclusions and future work

This article focuses on the possibility to use a RF-fingerprinting approach in the estimation of NB-IoT sensor position. A dataset is built by using an NB-IoT board equipped with a GPS receiver. Gathered data are used to implement a location algorithm based on the RSRP values measured in a real-world scenario. Data are processed with a WKNN algorithm which shows an average error of 91 m with $K = 2$. Taking into account the CDF, the error for the best case ($K = 2$) is below 109 m for 70% of the total cases. The accuracy values obtained in our analysis are overall comparable with results that can be found in literature for the LTE technology which, under similar conditions, produces a positioning error of 182 m in 68% of the cases (as reported in Ref. [6]). This work realizes a Proof of Concept (PoC) of a location technique that can be used on the NB-IoT sensor networks, without using GPS receivers, thus extending the life of battery and sensors.

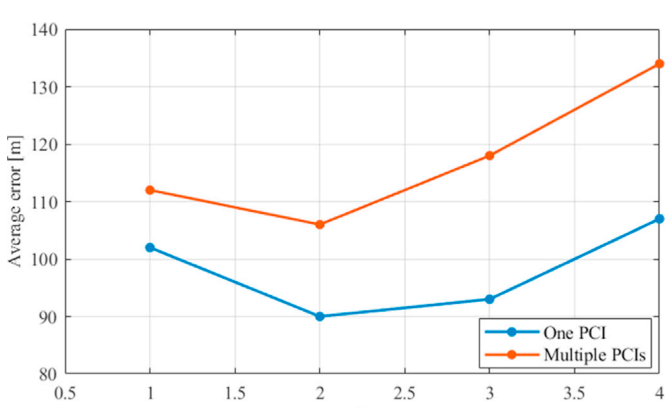

Fig. 10. Average error by varying K, with THR = −130 dBm, cluster 20 m × 20 m. Single vs multiple PCIs.

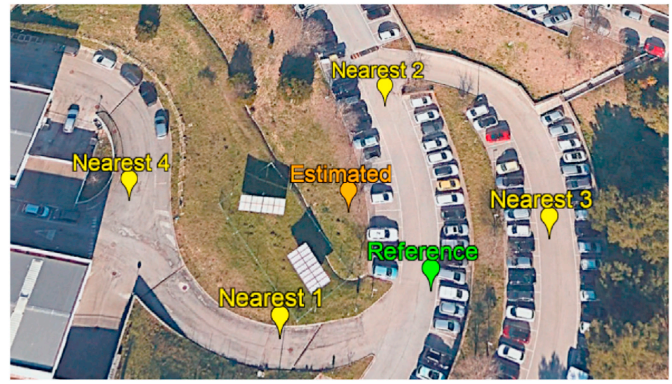

Fig. 11. Google Earth plot.

Future work may extend the preliminary outcomes to evaluate the algorithm performance with different NB-IoT SoCs, which may provide different radio-parameter values. Field tests have been performed in dense urban scenarios, and the same could be done in an indoor or rural environment.

Declaration of competing interest

We declare that we have no potential competing interests.

We confirm that all authors have approved the manuscript for submission.

We confirm that the content of the manuscript has not been published, or submitted for publication elsewhere.

Acknowledgements

The authors would like to thank Telecom Italia SpA for its support on the development of the here proposed research.

References

[1] S. Spinsante, S. Squartini, L. Gabrielli, M. Pizzichini, E. Gambi, F. Piazza, Wireless m-bus sensor networks for smart water grids: analysis and results, Int. J. Distributed Sens. Netw. 10 (6) (2014), 579271.
[2] S. Kumar, A. Graell i Amat, E. Rosnes, L. Senigagliesi, Private information retrieval from a cellular network with caching at the edge, IEEE Trans. Commun. 67 (7) (2019) 4900–4912.
[3] G. Pecoraro, S. Di Domenico, E. Cianca, M. De Sanctis, Csi-based fingerprinting for indoor localization using LTE signals, EURASIP J. Appl. Signal Process. 2018 (1) (2018) 1–18.
[4] R. Mondal, J. Turkka, T. Ristaniemi, T. Henttonen, Performance evaluation of MDT assisted LTE RF fingerprint framework, in: 2014 Seventh International Conference on Mobile Computing and Ubiquitous Networking (ICMU), IEEE, 2014, pp. 33–37.
[5] B.N. Ozdemir, A. Ceylan, Constructing a precise radio map and application of indoor positioning with dual-frequency Wi-Fi fingerprinting method, Measurement (2020), 107997.
[6] J. Turkka, T. Hiltunen, R.U. Mondal, T. Ristaniemi, Performance evaluation of LTE radio fingerprinting using field measurements, in: 2015 International Symposium on Wireless Communication Systems (ISWCS), IEEE, 2015, pp. 466–470.
[7] T. Hiltunen, J. Turkka, R. Mondal, T. Ristaniemi, Performance evaluation of LTE radio fingerprint positioning with timing advancing, in: 2015 10th International Conference on Information, Communications and Signal Processing (ICICS), IEEE, 2015, pp. 1–5.
[8] K. Radnosrati, C. Fritsche, F. Gunnarsson, F. Gustafsson, G. Hendeby, Localization in 3GPP LTE based on one RTT and one TDOA observation, IEEE Trans. Veh. Technol. 69 (3) (2020) 3399–3411.
[9] K. Muteba, K. Djouani, T. Olwal, Deep reinforcement learning based resource allocation for narrowband cognitive radio-IoT systems, Procedia Comput. Sci. 175 (2020) 315–324.
[10] Q. Song, S. Guo, X. Liu, Y. Yang, CSI amplitude fingerprinting-based NB-IoT indoor localization, IEEE Internet Things J. 5 (3) (2017) 1494–1504.
[11] X. Lin, J. Bergman, F. Gunnarsson, O. Liberg, S.M. Razavi, H.S. Razaghi, H. Rydn, Y. Sui, Positioning for the internet of things: a 3GPP perspective, IEEE Commun. 55 (12) (2017) 179–185.

[12] P. Savazzi, E. Goldoni, A. Vizziello, L. Favalli, P. Gamba, A wiener-based RSSI localization algorithm exploiting modulation diversity in LoRa networks, IEEE Sensor. J. 19 (24) (2019) 12381–12388.
[13] H. Sallouha, A. Chiumento, S. Rajendran, S. Pollin, Localization in ultra narrow band IoT networks: design guidelines and tradeoffs, IEEE Internet Things J. 6 (6) (2019) 9375–9385.
[14] Y.-P.E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, H.S. Razaghi, A primer on 3GPP narrowband internet of things, IEEE Commun. 55 (3) (2017) 117–123.
[15] Y.D. Beyene, R. Jantti, O. Tirkkonen, K. Ruttik, S. Iraji, A. Larmo, T. Tirronen, J. Torsner, NB-IoT technology overview and experience from cloud-RAN implementation, IEEE Wireless Commun. 24 (3) (2017) 26–32.
[16] N. Mangalvedhe, R. Ratasuk, A. Ghosh, NB-IoT deployment study for low power wide area cellular IoT, in: 2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), IEEE, 2016, pp. 1–6.
[17] I. Collotta, B. Melis, Notiziario tecnico tim - anticipare il 5G: NB-IoT. https://www.telecomitalia.com/, 2016. (Accessed 26 April 2022).
[18] 3GPP, 3GPP technical specification 36.214 v12.2.0. https://www.3gpp.org/, 2015. (Accessed 26 April 2022).
[19] R. Fantini, F. Mondello, A. Rigallo, D. Sorbara, Notiziario tecnico tim - le tecnologie abilitanti per l'iot. https://www.telecomitalia.com/, 2016. (Accessed 26 April 2022).
[20] M. Chen, Y. Miao, Y. Hao, K. Hwang, Narrow band internet of things, IEEE Access 5 (2017) 20557–20577.
[21] Y.-H. Wang, H. Li, X.-L. Luo, Q.-M. Sun, J.-N. Liu, A 3D fingerprinting positioning method based on cellular networks, Int. J. Distributed Sens. Netw. 10 (7) (2014), 248981.
[22] SODAQ, SODAQ SARA AFF (Arduino Form factor). https://support.sodaq.com/Boards/Sara_AFF/, 2019. (Accessed 26 April 2022).
[23] QUECTEL, BG96_AT_Commands_Manual_V2.1. https://www.quectel.com/, 2018. (Accessed 26 April 2022).
[24] Zecke, DIAG parser and to GSMTAP converter. https://github.com/moiji-mobile/diag-parser/, 2017. (Accessed 26 April 2022).
[25] OsmocomBB, What is GSMTAP?. https://osmocom.org/projects/baseband/wiki/GSMTAP/, 2018. (Accessed 26 April 2022).
[26] 3GPP, 3GPP technical specification 25.304. https://www.3gpp.org/, 2017. (Accessed 26 April 2022).
[27] N. Al Khanbashi, N. Al Sindi, S. Al-Araji, N. Ali, Z. Chaloupka, V. Yenamandra, J. Aweya, Real time evaluation of RF fingerprints in wireless LAN localization systems, in: 2013 10th Workshop on Positioning, Navigation and Communication (WPNC), IEEE, 2013, pp. 1–6.
[28] S.U. Rehman, K. Sowerby, C. Coghill, Analysis of receiver front end on the performance of rf fingerprinting, in: 2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications-(PIMRC), IEEE, 2012, pp. 2494–2499.
[29] Y. Li, Y. Lin, Z. Dou, Y. Chen, Research on RF fingerprint feature selection method, in: 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), IEEE, 2020, pp. 1–5.
[30] X. Liang, X. Gou, Y. Liu, Fingerprint-based location positioning using improved KNN, in: 2012 3rd IEEE International Conference on Network Infrastructure and Digital Content, IEEE, 2012, pp. 57–61.
[31] J. Oh, J. Kim, Adaptive K-nearest neighbour algorithm for WiFi fingerprint positioning, Ict Express 4 (2) (2018) 91–94.