2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 30 Sept. - 3 Oct. 2019, Pisa, Italy

E-Loc: Enhanced CSI Fingerprinting Localization for massive Machine-Type Communications in Wi-Fi Ambient Connectivity

Brieuc Berruet*†‡, Oumaya Baala*, Alexandre Caminada† and Valery Guillet‡
*FEMTO-ST Institute, Univ. Bourgogne Franche-Comte, CNRS, UTBM, Belfort, France
brieuc.berruet@utbm.fr and oumaya.baala@utbm.fr
†Univ. Côte dAzur, CNRS, Inria, I3S, COATI, Nice, France
alexandre.caminada@unice.fr
‡ Orange Labs Networks, RNM, WEP, Belfort, France
valery.guillet@orange.com

***Abstract*—A location-based service in the massive machine-type wireless communications (mMTC) must respect different requirements such as a minimal energy consumption at the target device or estimating the location in ambient connectivity. The solutions in mMTC must then consider localization approaches that provide target locations with few transmitted signals and with the support of only one anchor gateway. It is also major to use a relevant input data that manages the complex radio propagation mediums. In indoor environments, a solution builds on fingerprinting approach based on the channel state information (CSI) between the target device and a single anchor gateway. This paper presents a novel CSI fingerprinting localization method, E-Loc for mMTC dedicated to indoor systems in the Wi-Fi ambient connectivity context. E-Loc architecture is based on a convolutional neural network implementing inception models with an innovative design. CSI has been collected in a complex indoor environment, post-processed to handle the transmit power diversity, phase and timing offsets and fed to E-Loc. In various spatial distributions of training locations, E-Loc outperforms other tested solutions with a 99% confidence level for localization errors around 5 meters.**

I. Introduction

The concept of massive machine-type communications (mMTC) has emerged from the Internet of Things (IoT) paradigm and brought new challenges for the next decades [1]–[3]. A few of those deal with massive and heterogeneous communications of low-cost and low-energy devices. This condition is easily met indoors with the multiplication of devices in the different existing networks such as the Wi-Fi, ZigBee or Bluetooth.

These technologies are major candidates for indoor localization but the design of a solution must consider at first the battery saving of target devices. This implies to support low-traffic communications and limited information about the data link and signal propagation. For instance, the idea is to send only a minimal amount of signals to the gateways that will process and estimate the location of the target device. A network-centric approach such as a cloud computing is also a good solution to limit the energy wasting but it is beyond the scope of the paper. Furthermore, the exponential growth of location-based services in different use cases tends to have quickly deployable solutions. This means a location-based service must be operated with any kind of gateways distribution as an environment covered by an unique gateway. This scenario known as ambient connectivity penalizes a received signal strength (RSS) based solution that must verify the gateways distribution ensures an unique location estimation [4]. Hence, all these conditions limit the solutions oriented with ZigBee or Bluetooth technologies.

978-1-7281-1788-1/19/$31.00

Precisely, the Wi-Fi technology is more suitable for respecting this specific context. Indeed, the Wi-Fi technology allows to collect the channel state information (CSI) [5], [6], a set of complex values that leverages the multipath effects. This information is available thanks to the implementation of orthogonal frequency division multiplexing (OFDM) scheme that is often extended by multiple inputs multiple outputs (MIMO) technologies to improve data throughputs. The MIMO-OFDM Wi-Fi system takes advantage of the multipath effects to improve the data communication quality and allows the location estimation with a single gateway. For instance, it is possible to estimate the location directions of arrivals (DoA) of signal paths in the propagation medium combined with the time of arrival of the line-of-sight (LOS) path or RSS [7], [8]. However, this approach is time consuming because of complex processing schemes that increase the probability of network congestion at the gateway. Furthermore, the system must consider specific antenna array geometry as defined in the processing scheme and these approaches have limited performances in non line-of-sight (NLOS) conditions [9]–[11].

Another solution is build on fingerprinting approach that aims to estimate the location of target devices by knowing the variations of signals in the experiment area. The first step consists in collecting Wi-Fi CSI data in the experiment area that are stored in a database. Different methods exist for the data collection such as the simultaneous localization and mapping (SLAM) [12]–[15] that takes advantage of the heterogeneity of equipment in the experiment area to collect, updates and refines the database. In the second step, the locations of future target devices are estimated by the best correlation of the incoming data in the database. This procedure is often performed with a machine learning technique that is able to provide a fast response. This method has been widely experimented with the RSS metric such as RADAR, HORUS or other systems [16]–[20]. CSI proved to be an efficient metric where FIFS [21] or CSI-MIMO [22] enhanced the performances compared to RADAR and HORUS.

Nowadays, the growth of computation power accelerates the training of the deep learning techniques in different field of applications. In indoor localization, ConFi [23] and CiFi [24] performed location estimations with a deep convolutional neural networks (CNN) and improved the mean localization error by 40% compared to HORUS. The major drawback of these solutions is the input tensor requires multiple samples. This leads to a repetitive transmission of Wi-Fi signals for a long period of time that does not fit with a low-energy communication. DelFin [25], another CNN-based solution has been designed to estimate location with one CSI sample and improved the accuracy by 40% compared to FIFS and other standard machine learning techniques. However, this solution did not consider the phase of CSI data and the diversity of devices. The solutions based on deep belief network (DBN) and restricted Boltzmann machine (RBM) [26]–[29] requires only one CSI sample to provide a location and improved the performances compared to HORUS or RADAR in line-of-sight (LOS) conditions. However, these solutions require to have a deep learning structure for every training location in the experiment area. These solutions limit their deployment in vast areas and require a prior study to handle the heavy data storage. Nevertheless, these solutions proved that deep learning is a promising approach for fingerprinting localization that outperforms the standard machine learning techniques.

This paper presents E-Loc, a CNN based CSI fingerprinting solution for mMTC dedicated to indoor localization in Wi-Fi ambient connectivity. The solution is able to estimate locations with only one CSI sample transmitted to an unique anchor gateway. E-Loc has been tested on data collected with a channel sounder [30] in an experiment area composed of five rooms and an external corridor mixing LOS and NLOS conditions. The transmitter was a single antenna element and the receiver was an eight-element uniform linear array antenna. 20 MHz of bandwidth has been selected to reflect the most low-energy communication between an unique Wi-Fi anchor gateway and a low-cost low-energy target device. The transmitter was set at different training locations spatially distributed according to the fingerprinting protocol and at different testing locations to analyze the robustness and accuracy of E-Loc. The number of collected CSI samples per training location has been limited to respect a low-traffic communication and to have a solution that supports fast deployment. Then, every CSI sample has been processed to eliminate the transmit power and path loss effects as well as the phase and timing offsets to be robust to the devices diversity. At the same time, the resulting tensors bring richer information in this new structure. Finally, E-Loc has learned different spatial distributions of training locations and provided estimations of testing locations to evaluate its localization performances.

E-Loc is an innovative and unique solution for mMTC indoor localization for the Wi-Fi ambient connectivity that outperforms the previous solutions. To our knowledge, E-Loc is the first indoor localization solution to integrate the inception model (IM), an efficient deep features extraction in image processing [31] that reduces the depth of CNN architecture. Hereafter, we propose a novel CSI processing scheme to extract more informative data tensors and to handle from the transmission power diversity as well as the timing and phase offsets.

This paper is organized as follows: Section II details CSI data and the CSI processing scheme. Section III highlights the E-Loc architecture, the different layers and the training procedure. Section IV presents the analyses and discussions and Section V concludes on this work.

II. CSI Wi-Fi data

A. Received Signal Strength Drawbacks

In many wireless technologies, the medium access control (MAC) layer affords many information such as the receiving signal strength (RSS) to calibrate the data rate according to the quality of communication medium. A receiver is able to estimate its distance from a transmitter thanks to a propagation model and RSS measurements. However, its accuracy is extremely disturbed by the multiple paths of the transmitted signal and the robustness of RSS-based localizations drops severely in complex topology environments. An estimation of RSS statistical distributions [19], [32] or a time average of RSS values [16], [20] are solutions to decrease the location estimation errors but it requires to send a large number of messages in the mMTC requirements. Furthermore, a specific network must be designed to ensure that every RSS fingerprint is mathematically unique in the studied area by positioning multiple gateways [4] or to ensure a specific area topology. This last condition does not allow a fast deployment based on an ambient connectivity. The list of conditions shows Bluetooth and ZigBee technologies do not stick with our context. In other hand, these solution properties are more respected with the channel state information (CSI) in Wi-Fi technologies combined with the fingerprinting approach [21]–[28]. Finally, a database refinement technique such as the simultaneous localization and mapping (SLAM) allows the fingerprinting approach to stick with the ambient connectivity by updating the database [12]–[15], [33].

B. CSI Data

On the physical (PHY) layer, Wi-Fi implements the orthogonal frequency division multiplexing (OFDM), a digital multi-carrier modulation method. The signal is transmitted on multiple elements named subcarriers that are regularly spaced in the bandwidth and centered around the carrier frequency. Here, the orthogonality requires to have a specific spacing in frequency domain between subcarriers for respecting inter-carrier interferences cancellations. This method provides a set of complex values with a length corresponding to the number of subcarriers supported by the selected Wi-Fi technology. Finally, the scheme is often associated with multiple inputs multiple outputs (MIMO) technology for improving the data throughputs with the spatial multiplexing and beamforming techniques. Then, the set of complex values is estimated for every spatial stream i.e. every antenna element of the receiver. The result is called the channel state information (CSI) that can be expressed in the frequency or time domain. Mathematically, if R, S and T are respectively the number of receiving antenna elements, subcarriers and transmitting antenna elements, CSI is a three-dimensional complex tensor that can be defined as follows:

$$h_{r,s,t} = |h_{r,s,t}| e^{j\angle h_{r,s,t}} \tag{1}$$

where $r \in [1, \dots, R]$, $s \in [1, \dots, S]$ and $t \in [1, \dots, T]$. This three-dimensional complex tensor denoted H^{sample} represents the transfer function of the propagation medium at a specific time. The Wi-Fi technology is not fundamentally designed to the mMTC context and thus the system must set up a Wi-Fi communication that must be in accordance with the mMTC context. This work proposes to do location estimations of Wi-Fi systems with 20 MHz bandwidth at 5.2 GHz, a target device with one antenna element and an unique anchor gateway with eight antenna elements. This leads to have a dimensionality of H^{sample} represented by the tuple (R, S, T) equal to $(8, 56, 1)$.

C. CSI Processing

Different transmission powers of target devices deteriorate also the accuracy of CSI and RSS based localization systems. In fingerprinting, this occurs when the transmission power of training data collection differs from the power in use. Furthermore, the phase of CSI data may show differences because of phase and timing offsets between communicating wireless systems for every spatial stream [26], [28], [34]. The CSI phase may be written as follow:

$$\angle h_{r,s,t} = \angle \hat{h}_{r,s,t} - 2\pi \frac{k_s}{S} \delta_s + \beta + Z_{r,s,t} \tag{2}$$

where $\hat{h}_{r,s,t}$, k_s, δ_s, β and Z denote respectively the true CSI phase, the subcarrier index, the timing offset, the phase offset and the thermal noise. Algorithm 1 proposes to make the CSI amplitude division and CSI phase difference in two-by-two comparisons between a reference receiving antenna element and other receiving ones. Assuming an invariant orientation and height of the gateway during the whole data collection in the experiment area, the amplitude division removes the unknown transmitted power of hardware and the phase difference results in a new value without δ_s and β. This operation is close to existing solutions but has the advantage to provide also an enriched structure where the resulting tensors are $H^{|\cdot|}$, $H^{\Re}$ and $H^{\Im}$. The Frobenius norm is also applied to ease the learning procedure and the resulting tensors have been concatenated for forming H^{input}, the input tensor of the deep learning architecture of shape $(28, 56, 3)$.

Algorithm 1: CSI Processing

Data: R, S and T the number of receiving antenna elements, subcarriers and transmitting antenna elements. $\mathcal{F}$ the Frobenius norm. $\Re$ and $\Im$ the real and imaginary part of a complex value.
Input: H^{sample} the CSI complex tensor.
Output: $H^{|\cdot|}$, $H^{\Re}$ and $H^{\Im}$ three real 3D tensors.

```
for t ∈ [1, ..., T] do
    for s ∈ [1, ..., S] do
        loop = 1;
        for i ∈ [1, ..., R − 1] do
            for j ∈ [i + 1, ..., R] do
                h^{|·|}_{loop,s,t} ⟵ |h_{i,s,t}| / |h_{j,s,t}|;
                h^{ℜ}_{loop,s,t} ⟵ ℜ(e^{j(∠h_{i,s,t} − ∠h_{j,s,t})});
                h^{ℑ}_{loop,s,t} ⟵ ℑ(e^{j(∠h_{i,s,t} − ∠h_{j,s,t})});
                loop+ = 1;
    h^{|·|}_{:,:,t}, h^{ℜ}_{:,:,t}, h^{ℑ}_{:,:,t} ⟵ F(h^{|·|}_{:,:,t}), F(h^{ℜ}_{:,:,t}), F(h^{ℑ}_{:,:,t})
```

III. Deep Learning Solution

In this paper, the E-loc solution has an architecture that relies on some major features of convolutional neural networks and recent developments in image processing. This section describes the deep convolutional neural network (CNN), the E-Loc inception model and the procedure to select the best configuration of a deep learning architecture i.e. the optimal solution.

A. Deep Convolutional Neural Network

CNN is one of the most famous deep learning techniques in image recognition and has been efficient in different applications in indoor localization [24], [25], [35], [36]. CNN is a stacking of multiple hidden layers to extract complex features with convolution operations [37]. Here, a 2D CNN is generally composed of the following layers:

- ***Convolutional layers*** (CONV) that apply two-dimensional convolution operations to an input tensor. This operation depends on the number of convolution kernels and their size and the stride of kernels along the two first axes of the tensor. Then, each element of the tensor is processed by an activation ??? El-Loc implements the sELU activation function because it provides the best performances for a defined architecture.
- ***Pooling layers*** (POOL) that consist in extracting information such as the maximum or mean value in a sliding window. The goal of this operation is to extract relevant features in input tensors and may reduce the data dimensionality depending on the user specification. E-Loc achieves this with several maximum (MPOOL) and average (APOOL) pooling layers.
- ***Dropout layers*** (DOUT) that prevent the overfitting in the training procedure by setting randomly some values

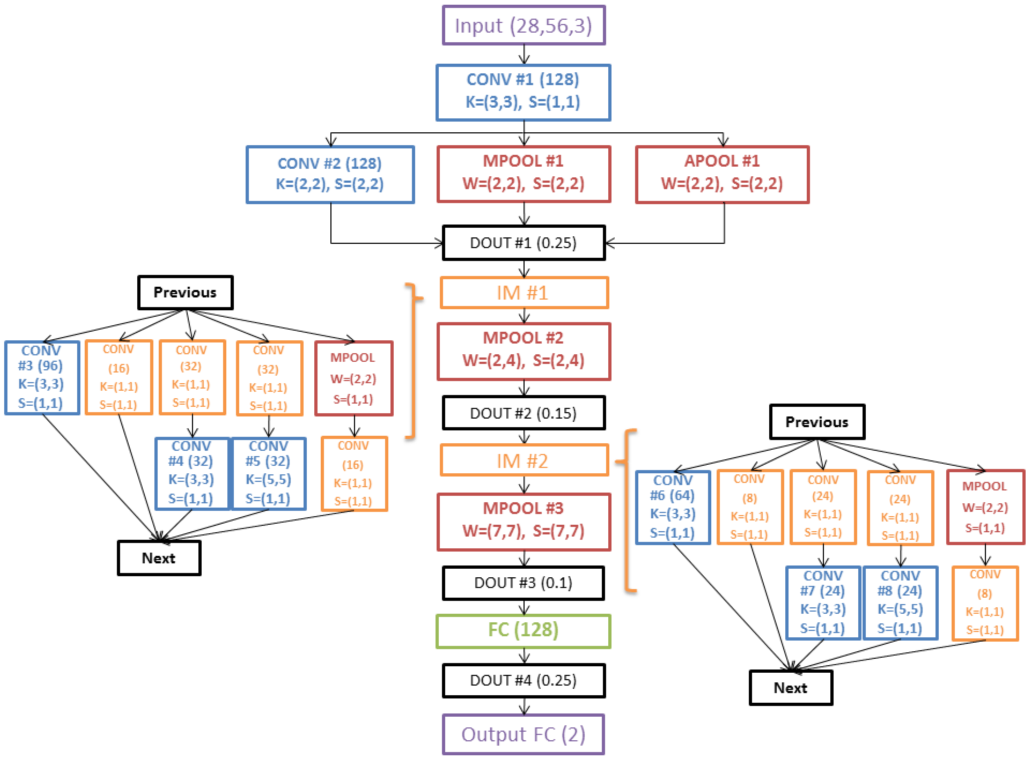


Fig. 1. E-Loc Deep Learning Architecture

to zero. The DOUT layers in E-Loc follow uniform distributions.

- ***Fully-connected layers*** (FC) that connect every neuron of a layer to every neuron of the previous one. Each connection has a weight and the result of every neuron in a FC layer is a weighted sum of values of every neuron in the previous layer. An activation step as described in convolutional layers is applied to resulting value of every neuron. In E-Loc, the output layer implemented the rectified linear unit (ReLU) as an activation function to avoid negative values as the estimated coordinates are all positive.

B. Inception Layer

In the above part, we highlighted the different components of CNN where the depth depends on the number of stacked layers to extract more complex and hidden relevant features. This approach has been often used in image recognition because of the complexity of features into images [35], [36]. Deeper is the architecture, slower is the training because of the exponential growth of the number of parameters. This requires large training dataset and heavy computation powers. Furthermore, this architecture increases the risk of overfitting that penalizes fingerprinting localization.

To deal with it, the inception model (IM) has been designed to dodge this approach [31]. It consists in building a layer with multiple branches where each branch has its own design composed of CONV and MPOOL layers. Then, the resulting tensors of each branch are concatenated along the last axis. In E-Loc, the basic IM architecture is slightly modified from the model in image processing applications to fit well with the indoor localization.

C. E-Loc Solution

Fig. 1 presents the full architecture of E-Loc that is composed of 4 major hidden layers and requires one CSI data sample as input to predict a location. The output layer estimates the label of CSI input data that are 2D Cartesian coordinates. According to the input shape, this architecture is designed for reducing the tensor dimensionality to have a features vector after all CONV and MPOOL layers. In other words, this means the two first dimensions of an input tensor are reduced to 1 before applying the fully-connected layers that is equivalent to a feature extraction procedure. Finally, the Adam algorithm, one of the fastest optimizer [38] has optimized the weights based on the mean squared errors between estimated and genuine coordinates.

D. Optimal Solution

Based on fingerprinting, CSI data is usually collected at different training and testing locations in an experiment area with the support of a technical team or with the ambient connected devices. Thereafter, collected data are divided into two datasets for learning and testing the localization solution. Then, it is necessary to determine how to select the best configuration of a deep learning architecture. To do this, E-Loc exploits the metric defined in [25] that is:

$$M_{test} = \frac{C_1 P_{50\%} + C_2 P_{90\%} + C_3 P_{99\%} + C_4 P_{loss}}{C_1 + C_2 + C_3 + C_4} \quad (3)$$

where $P_{50\%}, P_{90\%}, P_{99\%}$ are respectively the median, 90% and 99% confidence levels of localization errors calculated with the testing dataset, P_{loss} is the mean training localization error and C_1, C_2, C_3, C_4 are user-defined coefficients. The defined metric helps the user to balance between accuracy and robustness to outliers of a deep learning architecture in localization. Here, the user-defined coefficients are set to with $C_1 = 0.75, C_2 = 1, C_3 = 0.5$ and $C_4 = 0.5$ forcing the deep learning architecture to manage correctly outliers and to provide a good accuracy. Finally, the best parameters configuration corresponds to the lowest found value of M_{test} during the learning phase.

In this work, the context supposes a limited number of collected data at every training location for fast deployment. Then, the training data is learned in multiple epochs to find an optimal solution and the procedure is automatically stopped after 1,024 epochs from the lowest calculated M_{test}.

IV. Results and Discussions

This section presents the experiment area and the collected data, some insights about the training procedure of E-Loc and comparisons of E-Loc with some existing solutions in different spatial distributions of training locations.

A. Data and Experiment Area

CSI has been collected with a channel sounder [30] according to the data collection procedure presented in the DelFin solution [25] that sticks with IoT and ambient connectivity contexts. Fig. 3(a) shows the original map of data collection locations where the blue dots and red squares are training and testing locations in the experiment area, a 5-room apartment with an external corridor. The yellow stellar is the location of the anchor gateway. In this experiment setup, some testing locations are in Line-of-Sight (LOS) or non-LOS that allows to analyze the robustness of solutions in daily life applications. Moreover, the training data has been collected in two environment scenarios designed for on-the-field measurement or with a radio propagation simulator. The testing dataset was collected in a dynamic environment with 3 moving people and topographical modifications. The measurement equipment acquired 40 samples per training location and 80 samples per testing location. Finally, the number of training locations was relatively low to limit the gateway resources occupation and to fasten solutions deployment that induces a specific learning approach for deep learning methods.

B. M_{test} Metric

As introduced in Section III-D, M_{test} is designed to validate a localization solution according to user-defined weights. Here, this is also useful to configure a deep learning architecture as mentioned in [25]. However, this metric does not take into account the dependence on some learning parameters such as the learning rate, regularization with dropouts, the batch shuffle of training dataset or even the initialization of convolution kernels and weights in FC layers.

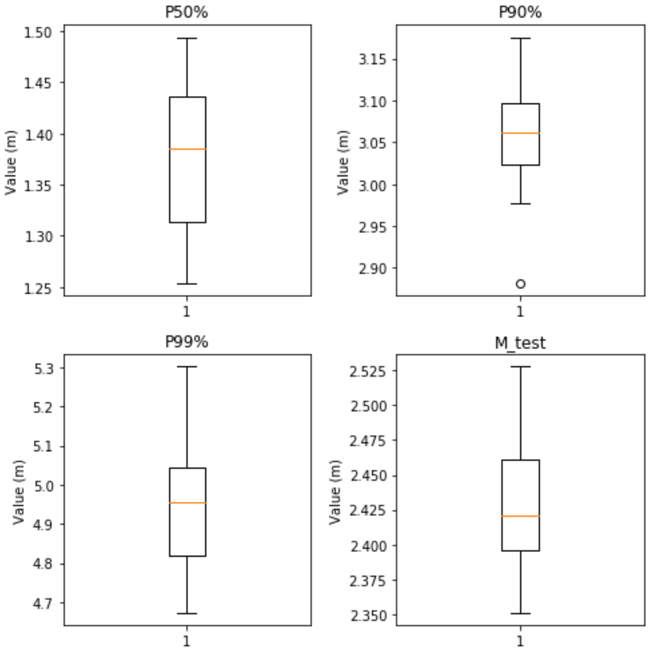
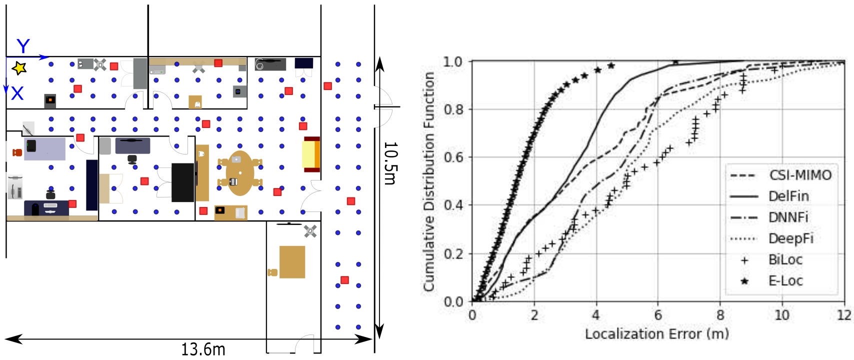


Fig. 2. Variation of $P_{50\%}, P_{90\%}, P_{99\%}$ and M_{test} with 21 generated E-Loc models

Fig. 2 presents a box-and-whisker plot of $P_{50\%}, P_{90\%}, P_{99\%}$ and M_{test} obtained with 21 models generated from E-Loc architecture. The results shows the impact of learning parameters on the metric. M_{test} has a variation from 2.352 to 2.533 meters that proves the difficulty to select the best deep learning architecture. If the user tends to select with $P_{50\%}$, $P_{90\%}$ or $P_{99\%}$, the problems would be equivalent.

Hence, the validation procedure is reinforced with a new approach to handle this problem. A parameters configuration is considered as better when M_{test} is 10% lower than the results of other solutions. If this first condition is not met, both concurrent solutions generate 21 models where every model is associated with a value of M_{test}. Then, every solution has a list of 21 M_{test} and it is possible to calculate the list average. Finally, the lowest average corresponds to the best solution. This approach is extended to further statistical moments if the first order did not yield to satisfying performances. Among all tested CNN architectures and configurations, E-Loc presented in Section III provided the best result with this procedure.

Finally, the selection of E-Loc configuration is also dependent on the testing dataset that is just a parsimonious representation of unknown locations in the experiment area. To handle this, 21 models are generated from the E-Loc architecture and the model corresponding to the median of M_{test} is saved as the optimal solution. To simplify the rest of the analysis, the selected model resulting from this second phase is named E-Loc.

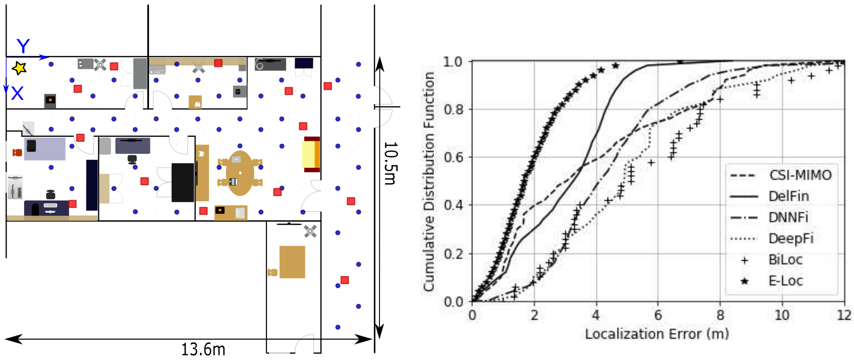


(a) Original training/testing locations

(b) Cumulative function of localization errors

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	3.24	7.54	10.60
DelFin	3.06	4.96	6.63
DNNFi	4.2	6.67	11.56
DeepFi	4.92	8.31	12.22
BiLoc	4.99	8.76	9.99
E-Loc	1.41	3.06	4.81

(c) Values of $P_{50\%}, P_{90\%}, P_{99\%}$

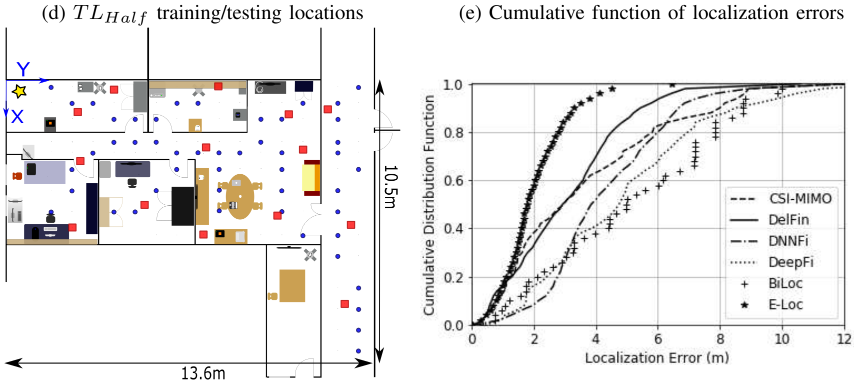


(d) TL_{Half} training/testing locations

(e) Cumulative function of localization errors

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	2.84	7.32	11.18
DelFin	3.28	4.78	6.04
DNNFi	4.13	7.02	11.51
DeepFi	4.91	8.73	11.62
BiLoc	5.12	9.18	11.91
E-Loc	1.68	3.45	5.01

(f) Values of $P_{50\%}, P_{90\%}, P_{99\%}$

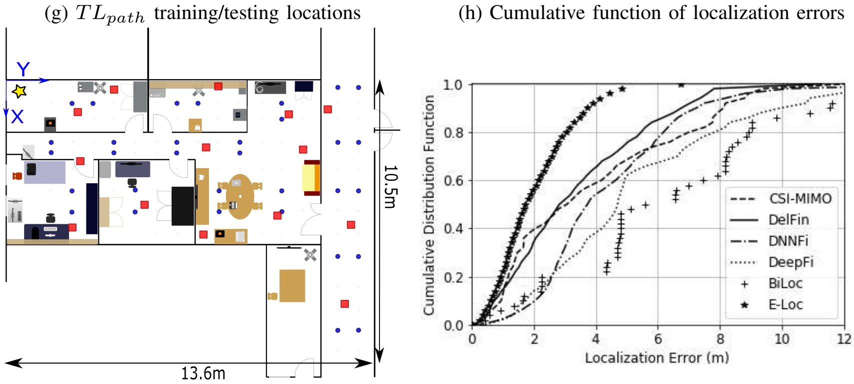


(g) TL_{path} training/testing locations

(h) Cumulative function of localization errors

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	2.98	8.00	10.64
DelFin	2.93	5.44	7.22
DNNFi	3.85	6.67	10.24
DeepFi	4.75	8.49	12.57
BiLoc	4.99	8.76	9.99
E-Loc	1.79	3.31	4.86

(i) Values of $P_{50\%}, P_{90\%}, P_{99\%}$



(j) $TL_{cluster}$ training/testing locations

(k) Cumulative function of localization errors

Solution	$P_{50\%}$	$P_{90\%}$	$P_{99\%}$
CSI-MIMO	3.07	8.13	10.82
DelFin	2.82	6.77	8.07
DNNFi	3.73	7.24	11.6
DeepFi	4.75	9.98	12.82
BiLoc	5.58	11.53	12.93
E-Loc	1.73	3.72	5.20

(l) Values of $P_{50\%}, P_{90\%}, P_{99\%}$

Fig. 3. Performance of localization solutions in multiple spatial distributions of training locations

C. Comparison with Existing Solutions

The performances of E-Loc is compared to 5 other methods: DelFin [25], DNNFi [29], DeepFi [27], BiLoc [26] and CSI-MIMO [22]. All these solutions provide a location estimation of a target device with only one data sample and a CSI processing scheme that removes the phase and timing offsets. We have also normalized the amplitude with the Frobenius norm to stick with the context of transmit power diversity.

Fig. 3(b) and 3(c) present respectively the cumulative distribution function of localization errors and the values of $P_{50\%}, P_{90\%}, P_{99\%}$ in the original training distribution presented in Fig. 3(a). Without doubts, E-Loc performs the best results with a median localization error of 1.41 meters, and $P_{90\%}$ and $P_{99\%}$ of 3.06 meters and 4.81 meters respectively. E-Loc decreases $P_{50\%}$ by 64% and $P_{90\%}$ by 38.7% compared to DelFin, the second best solution. Here, almost 99% of testing dataset is localized with a localization error less than 5 meters that highlights a good robustness of the solution in this experiment setup. BiLoc provides the worst results with a median localization error of 4.99 meters. In this experiment, the DBNs provide a coarse localization compared to CNNs. This phenomenon is mainly caused by the lack of CSI training samples per location compared to the number of samples used in BiLoc, DNNFi and DeepFi.

From the original spatial distribution of training locations, the comparison have been extended to three other data collections based on different spatial distributions of training locations. The first TL_{Half} that is composed of 55 training locations as presented in Fig. 3(d) kept the regular distribution but increased by a half the space between training locations that led to a diamond grid-like distribution. The second TL_{Path} picked 49 training locations as drawn in Fig. 3(g) to build a path-like distribution that can be provided with dead-reckoning approaches. The last $TL_{cluster}$ was a sparse distribution of small training locations clusters composed of 32 training locations as sketched in Fig. 3(j). The clusters have been manually chosen to cover completely and uniformly the experiment area. A cluster is composed of 2 training locations to have a rich-enough representation of CSI data per region.
The new training dataset in TL_{Half} deteriorates slightly the performances of E-Loc where its indicators increase approximatively by 5 to 10%. DelFin has a slight degradation of $P_{50\%}$ whereas CSI-MIMO decreases $P_{50\%}$ by 12.4%. Here, the median localization error with DBN-based solutions are not affected by the new distributions. This second experimentation shows the initial training location distribution may be divided to reduce the time of data collection i.e. the network occupation and human interventions without impinging the localization accuracy. This implies to store a smaller database that reduces energy costs and cloud storage infrastructure. E-Loc has still estimated 99% of the testing dataset with a localization error less than 5 meters.
In the distribution of TL_{Path}, E-Loc performs equivalent localization while CSI-MIMO gets some difficulties compared to TL_{Half}. DelFin and DBN-based solutions decreases globally the median localization error $P_{50\%}$. This last distribution of training locations highlights E-Loc is also really efficient where CSI data has been collected with dead-reckoning approaches.
A distribution such as $TL_{cluster}$ decreases $P_{50\%}$ of BiLoc or E-Loc compared to the first spatial distribution of training locations. In another hand, DNNFi and DelFin has a lower median localization error in non-dense distributions of training locations. However, all the solutions except for CSI-MIMO lose some robustness to extreme values where $P_{90\%}$ is decreased until 40% for DelFin. Nevertheless, E-Loc keeps at least 99% of the testing dataset around 5 meters of localization errors as per the Fig. 3(i).
In this experiment, E-Loc performs generally the best results and handles TL_{Path} and $TL_{cluster}$ that is crucial for providing fast and low-cost deployment for mMTC solutions in ambient connectivity. Furthermore, 99% of testing dataset is estimated with a localization error around 5 meters in all the spatial distributions of training locations that is really promising in this limiting context.

V. Conclusion

In the context of massive machine-type communications (mMTC) and the Wi-Fi ambient connectivity, this paper proposes E-Loc, a new deep learning architecture for indoor fingerprinting localization. This has been achieved in an experiment area composed of 5 rooms and an external corridor that is submitted to multiple signal perturbations such as shadowing and multipath fading. The relevant input data, the channel state information (CSI) has been collected with a channel sounder representing the communication of 1-antenna element target devices with a single 8-antenna elements anchor gateway. After data collection in the communication context, the channel state information has been processed with a new and unique framework to eliminate the transmission power and the timing and phase offsets. Then, the resulting tensors have been learned by E-Loc, a deep convolutional neural network with an unique inception model to reduce the architecture depth. During the training session, a metric evaluated the localization performance of E-Loc to find the best parameters configuration. This approach has been improved by additional models generation to handle the performance variations implied by the regularization of dropout layers. Finally, E-Loc has been tested in the experiment area with different spatial distributions of training locations against DelFin, DNNFi, DeepFi, BiLoc and CSI-MIMO. The methods have been evaluated in a dense grid-like distribution, a diamond grid-like distribution, a path-like distribution and a sparse distribution of small clusters of training locations. In every tested distribution, E-Loc performed the best results with a 99% confidence level of localization errors of around 5 meters. Hence, E-Loc proved to be efficient for mMTC solutions in ambient connectivity and reliable in fast or unsupervised data collections that is essential to ease the solutions deployment.

However, the evaluation of E-Loc was conducted in a single environment that assumes the same central frequency carrier in the training and operation phases and an anchor gateway composed of eight antenna elements. Then, future works will study the reliability of E-Loc where the Wi-Fi channel of testing datasets differs from the one used for training locations. The robustness of E-Loc will be also evaluated with different antenna elements configurations of the anchor gateway. Parallel to this, the evaluation of E-Loc will be conducted in other experiment areas. E-Loc will also investigate other CSI processing schemes to improve the localization performances. Finally, the deep learning architecture of E-Loc will be tested with other wireless communication technologies such as the long-term evolution for machines (LTE-M).

REFERENCES

[1] (2017) Battery life in connected wireless iot devices. [Online]. Available: http://www.silabs.com/whitepapers/battery-life-in-connected-wireless-iot-devices

[2] (2016) Paving the path to narrowband 5g with lte internet of things (iot). [Online]. Available: https://www.qualcomm.com/media/documents/files/paving-the-path-to-narrowband-5g-with-lte-iot.pdf

[3] K. Takeda, W. A. Hapsari, H. Takahashi, D. Fujishima, and Z. Miao, "New technologies for achieving iot in lte release 13," *NTT Docomo Technical Report*, vol. 18, pp. 39–51, Oct. 2016.

[4] O. Baala, Y. Zheng, and A. Caminada, "The impact of ap placement in wlan-based indoor positioning system," in *Proc. IEEE 2009 Eighth International Conference on Networks (ICN)*, Gosier, Guadeloupe, Mar. 2009, pp. 12–17.

[5] D. Halperin, W. J. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," in *Proc. ACM SIGCOMM 2010 conference (MobiCom15)*, New Delhi, India, Sep. 2010, pp. 159–170.

[6] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wifi," in *Proc. ACM 21st Annual International Conference on Mobile Computing and Networking (MobiCom15)*, Paris, France, Sep. 2015, pp. 53–64.

[7] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, pp. 276–280, Mar. 1986.

[8] R. Roy and T. Kailath, "Esprit - estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 984–995, Jul. 1989.

[9] A. Gaber and A. Omar, "A study of wireless indoor positioning based on joint tdoa and doa estimation using 2-d matrix pencil algorithms and ieee 802.11ac," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 2440–2454, May 2015.

[10] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single wifi access point," in *Proc. of the 13th USENIX Symposium on Networked Systems Design and Implementation*, Santa Clara, CA, USA, Mar. 2016, pp. 165–178.

[11] C. Yang and H.-R. Shao, "Wifi-based indoor positioning," *IEEE Communications Magazine*, vol. 53, pp. 150–157, Mar. 2015.

[12] J.-B. Prost, "Golocalisation indoor grande chelle: Des techniques d'auto-apprentissage prparent une rvolution de la localisation indoor," in *Revue de l'lectricit et de l'lectronique*.

[13] J.-G. Park *et al.*, "Growing an organic indoor location system," in *Proc. IEEE 8th annual international conference on Mobile systems, applications and services (MobiSys '10)*, Sep. 2010, pp. 271–284.

[14] Z. Yang, C. Wuan, and Y. Liu, "Locating in fingerprinting space: Wireless indoor localization with little human intervention," in *Proc. ACM 18th annual international conference on Mobile computing and networking (MobiCom 12)*, Istanbul, Turkey, Aug. 2012, pp. 269–280.

[15] A. Goswami, L. E. Ortiz, and S. Das, "Wigem: A learning-based approach for indoor localization," in *Proc. of the 17th ACM Conference on emerging Networking Experiments and Technologies*, Tokyo, Japan, Dec. 2011.

[16] P. Bahl and V. N. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *Proc. IEEE 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'2000)*, Tel Aviv, Israel, Mar. 2000.

[17] M. Youssef and A. Agrawala, "The horus wlan location determination system," in *Proc. ACM 3rd international conference on Mobile systems, applications, and services (MobiSys '05)*, Seattle, Washington, Jun. 2005, pp. 205–218.

[18] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless lans," *Elsevier The International Journal of Computer and Telecommunications Networking*, vol. 47, pp. 825–845, Apr. 2005.

[19] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao, "A wireless lan-based indoor positioning technology," *IBM Journal of Research and Development*, vol. 48, pp. 617–626, Sep. 2004.

[20] C. Laoudias, P. Kemppi, and C. Panayiotou, "Localization using radial basis function networks and signal strength fingerprints in wlan," in *IEEE Global Telecommunications Conference (GLOBECOM 2009)*, Honolulu, HI, Dec. 2009.

[21] J. Xiao, K. Wu, Y. Yi, and L. M. Ni, "Fifs: Fine-grained indoor fingerprinting system," in *Proc. IEEE 21st International Conference on Computer Communications and Networks (ICCCN)*, Munich, Germany, Aug. 2012.

[22] Y. Chapre, A. Ignjatovic, A. Seneviratne, and S. Jha, "Csi-mimo: Indoor wi-fi fingerprinting system," in *Proc. IEEE 39th Conference on Local Computer Networks (LCN)*, 2014, pp. 202–209.

[23] H. Chen, W. L. Y. Zhang, X. Tao, and P. Zhang, "Confi: Convolutional neural networks based indoor wi-fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18 066–18 074, 2017.

[24] X. Wang, X. Wang, and S. Mao, "Cifi: Deep convolutional neural network for indoor localization with 5ghz wi-fi," in *IEEE ICC 2017 International Conference on Communications*, Paris, France, May 2017.

[25] B. Berruet, O. Baala, A. Caminada, and V. Guillet, "Delfin: A deep learning based csi fingerprinting indoor localization in iot context," in *the 9th IEEE International Conference on Indoor Positioning and Indoor Navigation*, Nantes, France, 2018.

[26] X. Wang, L. Gao, and S. Mao, "Biloc: Bi-modal deep learning for indoor localization with commodity 5ghz wifi," *IEEE Access: Cooperative and Intelligent Sensing*, vol. 5, pp. 4209–4220, 2017.

[27] X. Wang, L. Gao, S. Mao, and S. Pandey, "Csi-based fingerprinting for indoor localization: A deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 66, pp. 763–776, Jan. 2017.

[28] X. Wang, L. Gao, and S. Mao, "Csi phase fingerprinting for indoor localization with a deep learning approach," *IEEE Internet of Things Journal*, vol. 3, pp. 1113–1123, Dec. 2016.

[29] G.-S. Wu and P.-H. Tseng, "A deep neural network-based indoor positioning method using channel state information," in *2018 Workshop on Computing, Networking and Communications*, Maui, Hawaii, USA, 2018.

[30] J.-M. Conrat, P. Pajusco, and J.-Y. Thiriet, "A multibands wideband propagation channel sounder from 2 to 60 ghz," in *Proc. IEEE Instrumentation and Measurement Technology Conference (ITMC)*, Yangzhou, Jiangsu, China, Apr. 2006.

[31] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *the 28th IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, Jun. 2015.

[32] O. Baala, Y. Zheng, and A. Caminada, "Toward environment indicators to evaluate wlan-based indoor positioning system," in *Proc. IEEE/ACS 2009 International Conference on Computer Systems and Applications (AICCSA)*, Rabat, Marocco, May 2009, pp. 243–250.

[33] C. Wu, Z. Yand, Y. Liu, and W. Xi, "Will: Wireless indoor localization without site survey," *IEEE Transactions on Parallel and Distributed systems*, vol. 24, pp. 839–848, Apr. 2013.

[34] M. Kotaru, K. Josh, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using wifi," in *Proc. of the 2015 ACM Conference on Special Interest Group on Data Communication*, London, United Kingdom, Aug. 2015, pp. 269–282.

[35] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *The 3rd International Conference on Learning Representations*, San Diego, USA, May 2015.

[36] Y. Lecun, Y. B. L. Bottou, and P. Haffner, "Gradien-based learning applied to document recognition," *IEEE Access*, vol. 5, pp. 18 066–18 074, 2017.

[37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[38] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *The 3rd International Conference on Learning Representations*, San Diego, USA, May 2015.