An Accurate, Robust and Low Dimensionality Deep Learning Localization Approach in DM-MIMO Systems Based on RSS

Seyedeh Samira Moosavi (seyedeh-samira.moosavi.1@ulaval.ca)
Universite Laval Faculte des sciences et de genie
Paul Fortier
Universite Laval Faculte des sciences et de genie

Research Article

Keywords: Distributed massive MIMO (DM-MIMO), Fingerprinting (FP), Localization, Affinity propagation clustering (APC), Deep neural networks (DNNs)

Posted Date: March 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-178416/v1

Noname manuscript No.
(will be inserted by the editor)

An Accurate, Robust and Low Dimensionality Deep Learning Localization Approach in DM-MIMO Systems Based on RSS

Seyedeh Samira Moosavi · Paul Fortier

Received: date / Accepted: date

Abstract Currently, localization in distributed massive MIMO (DM-MIMO) systems based on the fingerprinting (FP) approach has attracted great interest. However, this method suffers from severe multipath and signal degradation such that its accuracy is deteriorated in complex propagation environments, which results in variable received signal strength (RSS). Therefore, providing robust and accurate localization is the goal of this work. In this paper, we propose an FP-based approach to improve the accuracy of localization by reducing the noise and the dimensions of the RSS data. In the proposed approach, the fingerprints rely solely on the RSS from the single-antenna MT collected at each of the receive antenna elements of the massive MIMO base station. After creating a radio map, principal component analysis (PCA) is performed to reduce the noise and redundancy. PCA reduces the data dimension which leads to the selection of the appropriate antennas and reduces complexity. A clustering algorithm based on K-means and affinity propagation clustering (APC) is employed to divide the whole area into several regions which improves positioning precision and reduces complexity and latency. Finally, in order to have high precise localization estimation, all similar data in each cluster are modeled using a well-designed deep neural network (DNN) regression. Simulation results show that the proposed scheme improves positioning accuracy significantly. This approach has high coverage and improves average root-mean-squared error (RMSE) performance to a few meters, which

Seyedeh Samira Moosavi
Department of Electrical and Computer Engineering, Laval University, Quebec, QC G1V 0A6, Canada
Tel.: +1 418-573-3316
E-mail: seyedeh-samira.moosavi.1@ulaval.ca

Paul Fortier
Department of Electrical and Computer Engineering, Laval University, Quebec, QC G1V 0A6, Canada
Tel.: +1 418- 656-2131
E-mail: paul.fortier@gel.ulaval.ca

is expected in 5G and beyond networks. Consequently, it also proves the superiority of the proposed method over the previous location estimation schemes.

Keywords Distributed massive MIMO (DM-MIMO) · Fingerprinting (FP) · Localization · Affinity propagation clustering (APC) · Deep neural networks (DNNs).

1 Introduction

Location-based services (LBSs) have recently attracted significant attention in wireless network applications. Today, several consumer goods are equipped with user location features which provide immediate and accurate localization of lost, delayed, or damaged assets [1]. Location information has a great application potential in industry, medicine, emergency management, surveillance, controlling autonomous vehicles, and many other various fields [2]. At the same time, the demand for recognition of a mobile terminal's (MT) location has greatly increased so that numerous researches have been conducted on MTs' location [3]. Therefore, developping localization technology is becoming more and more important. The source of most localization systems are based in urban settings [3]. Today, the global positioning system (GPS) is the most used technology for outdoor localization due to its availability [4]. However, its accuracy deteriorates in shadowed locations and in the vicinity of high-rise buildings due to diminished satellite signals in the absence of line-of-sight (LoS) propagation [5]. Also, it requires too much power on an MT. Therefore, LBSs significantly need accurate and real-time localization to obtain notable performance improvement over existing cellular localization networks.

Currently, with the advent of 5G, the use of massive multiple-input multiple-output (M-MIMO) systems is drawing attention from the localization research community [2]. M-MIMO systems have been introduced as an enabling technology for 5G networks to improve localization accuracy in addition to enhancing communication performance [1]. In our work, we consider positioning multiple users simultaneously in a distributed massive MIMO (DM-MIMO) system, wherein the users are served on the same time-frequency resource by a large number of spatially-separated remote radio heads (RRHs) distributed over the whole area [6].

DM-MIMO systems potentially provide higher spectral efficiency [7], energy efficiency [8], average throughput rate, and coverage probability compared to conventional co-located massive MIMO (CM-MIMO) systems, where the base station (BS) is equipped with an array of co-located antennas [7]. Although DM-MIMO systems have significant benefits, user localization has not yet been as well established as with CM-MIMO systems [9]. A variety of wireless signal properties have been considered in M-MIMO systems for MT's localization. Among them, we concentrate on the received signal strength (RSS) since it has lower cost and complexity [9]. Machine learning (ML) approaches are then employed on the extracted signal properties, which are considered

as to be a unique fingerprint of a specific location. Therefore, the localization problem can be solved using pattern recognition, which consists of fingerprint extraction, fingerprint matching, and ultimately location estimation [9].

In this study, we propose a robust and precise localization method using a dimension reduction technique, clustering, and regression, which are accomplished via two modes: the offline mode and the online mode. The extracted RSS samples from the whole area are analyzed during the offline mode. First, the dimensions of RSS samples are reduced using principal component analysis (PCA). The whole area is then split into several sub-areas using a combination of clustering algorithms based on the K-means and affinity propagation clustering (APC) algorithms. Ultimately, a deep neural network (DNN) regression is applied to the RSS samples of each cluster. The accuracy of the model is estimated using a validation dataset. When a new fingerprint is given in the online mode, it is first preprocessed, then its cluster is specified, and finally, its location is estimated. We now summarize our major contributions in four aspects.

1. In the preprocessing step, we apply the PCA technique on all data samples to denoise the RSS sample and extract effective features and reduced unimportant features from RSS vectors. Preprocessing leads to speed up and improve the accuracy of our proposed machine learning-based method since the training time and complexity are reduced significantly with fewer dimensions (features). Also, it helps us to select a proper set of RRHs.
2. In the clustering step, a fast convergence, and initial value independent clustering method relying on a combination of K-means and AP clustering algorithms is proposed. This method reduces latency and computational complexity and helps to improve localization accuracy.
3. We propose a DNN regression for each cluster using all the data of the corresponding cluster to estimate the location more precisely.
4. The performance of the proposed localization method is evaluated in terms of root-mean-squared-error (RMSE) via simulations and compared to the works in [9].

The rest of the paper is organized as follows. Section 4.2 overviews existing techniques and related works for localization. In Section 4.3, we present the system model. The positioning method is proposed in Section 4.4. Simulations results are presented and discussed in Section 4.5. A conclusion is presented in Section 4.6.

2 Related Work

2.1 User Positioning in Massive MIMO System

In recent years, user positioning in M-MIMO has attracted much attention and there are several research works in this area. The authors in [10], [11], and [12] use angle-of-arrival (AoA) information to estimate UE position in

M-MIMO systems. In [13] and [14], the combined information of AoA, angle-of-delay (AoD), and time delay is used for user positioning in M-MIMO, where in [13] a mm-Wave M-MIMO system including LOS scenarios is considered. In [15], a compressed sensing approach is proposed to estimate the location of a MT from time-of-arrival (ToA) data recorded at multiple M-MIMO BSs. In [16], an environment sensing method is employed in a highly directional 60 GHz mm-Wave network to estimate MT's positions. However, the localization in all of the above techniques is completed with based on the information obtained from a CM-MIMO system configuration, where the BS hosts an array of antennas. But, these methods are not applicable in DM-MIMO systems, where single-antenna remote radio heads are considered. In [17] and [18], the Gaussian process regression (GPR) ML algorithm is employed based on RSS measurements in DM-MIMO systems. In [3], the performance of several ML algorithms, which are used in conjunction with fingerprint-based MT localization for DM-MIMO wireless systems configurations, is investigated and evaluated. In [9], RSS-based positioning using a machine learning method relies on the affinity propagation clustering algorithm and the GPR algorithm. Among the relevant works, the study of [9] is the most pertinent for our investigation, wherein the focus of the analysis is based on GPR. We expand on the work presented in [9], using data compression and deep learning algorithms to provide higher localization accuracy and less computational complexity.

2.2 Machine Learning and Deep Learning for User Positioning

Localization techniques are classified into four main categories: proximity-based, angle-based, range-based, and fingerprinting-based. The proximity-based is the most straightforward technique where the location is provided approximately in a particular radio coverage area based on the locations of the BSs. Therefore, BSs are required, which is not suitable for large areas [19]. The angle-based technique, which is based on the AoA of the received signal, is not efficient in non-line-of-sight (NLoS) situations because it produces a coarse error for positioning [20],[21]. In range-based techniques, one must compute the distance between the MT and at least three BSs. Then the MT location is estimated using trilateration. This can be accomplished through radio signal information received from MTs such as ToA and received signal strength (RSS). The ToA method is known for its complexity because it requires very expensive hardware at the BS, such as high accuracy clocks for time synchronization [22]. In addition, it has low performance in NLoS environments. It has been demonstrated that the RSS method is appropriate in non-urban environments because by increasing the distances the path loss is expected to decrease steadily [3]. This issue can be mitigated when the RSS method is employed in conjunction with a fingerprinting (FP) based method [17].

In a FP-based method, the location of MTs is estimated based on a pre-recorded data, called fingerprint, using ML and deep learning (DL) algorithms [17]. Since FP-based positioning methods have a good performance in highly-

cluttered multipath environments [17], [23], they can be used in many systems such as WiFi networks [24, 25, 26]. In addition to received signal information, channel state information (CSI) [27], [28] is used as the position fingerprint. In recent years, the FP-based localization method has attracted significant interest by combining mobile positioning requirements into 5G wireless communication systems due to its broad applicability and high cost-efficiency without any hardware requirement on the MTs [29].

Several machine learning methods including GP methods [17], and more recently, deep learning methods [30], [27] have been applied and investigated for wireless user positioning. However, the proposed methods for WiFi systems [27, 31, 32] are not applicable for M-MIMO systems because they do not consider the associated inter-user interference. In addition, they concentrate on the downlink, where the MTs estimate their positions by managing the computational cost while in M-MIMO systems, positioning is performed on the uplink, where the BS estimates the MTs' position.

3 System Description

In the considered single-cell DM-MIMO system (Fig. 1), there are K single-antenna users that transmit signals to M single-antenna RRHs on the same time-frequency resource. The high-speed front-haul links connect RRHs to a central processor unit (CU). When the RRHs receive signals transmitted by the users on the uplink, individually record their own multi-user RSS values and send them to the CU. The CU handles the multi-user interference and extracts the per-user RSS values from the multi-user RSS values. Then the CU from each user forms an $M \times 1$ RSS vector to perform localization [33], [34]. Details are as follows.

3.1 Propagation Model

To explain the uplink of a multi-user DM-MIMO system in more detail, let $\mathbf{w}_k$ be the symbol vector transmitted by user k with transmission power ρ. If g_{mk} is the flat-fading channel gain between user k and RRH m, the sum symbol vector $\mathbf{y}_m$ received at RRH m is given by

$$\mathbf{y}_m = \sqrt{\rho} \sum_{k=1}^{K} g_{mk} \mathbf{w}_k + \mathbf{n}_m \tag{1}$$

In (1), $g_{mk} = q_{mk}\sqrt{h_{mk}}$ is a flat-fading channel where $q_{m,k}$ denotes small-scale fading represented by an independent and identically distributed (i.i.d.) zero mean complex Gaussian random variable with unit variance, i.e., $q_{mk} \sim \mathcal{CN}(0,1)$, h_{mk} is the large-scale fading coefficient, and $\mathbf{n}_m \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I})$ is the additive white Gaussian noise. Note that the large-scale fading coefficient h_{mk} can be modeled [35] as

$$h_{mk} = b_0 d_{mk}^{-\alpha} 10^{z_{mk}/10} \tag{2}$$

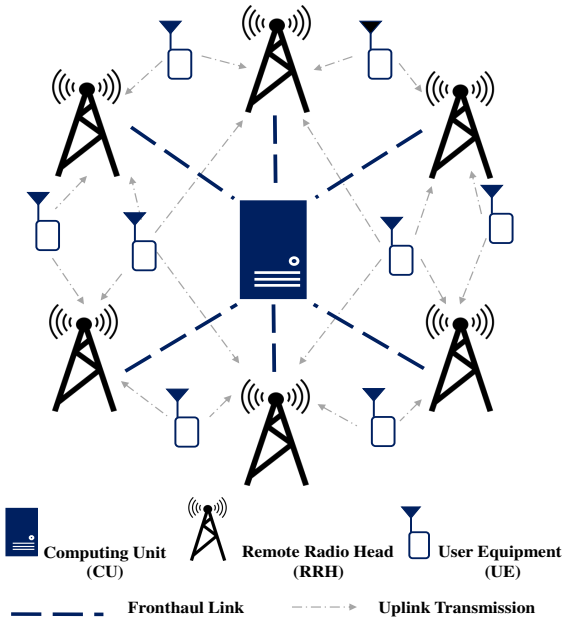


Fig. 1: Multi-user DM-MIMO system model for location estimation.

where b_0 is the path loss at reference distance d_0, d_{mk} is the distance between user k and RRH m, α is the path-loss exponent (typically dependent on the environment and the range), and z_{mk} is the log-normal shadowing noise coefficient with $10\log_{10} z_{mk} \sim \mathcal{N}(0, \sigma_z^2)$.

3.2 Mitigating Multi-user Interferences

For measuring the RSS, we consider the power of the received signal at RRH m which is given by $||\mathbf{y}_m||^2$ according to (1). But we should note that $||\mathbf{y}_m||^2$ at RRH m is in fact the multiuser RSS because the symbol vectors which are transmitted by all K users are combined at RRH m. Consequently, $||\mathbf{y}_m||^2$ cannot be directly used to estimate the position of user k. So the RSS of each user is not separately distinguishable. To overcome this, the symbol vectors $\mathbf{w}_k$ in (1) should be mutually orthogonal and should be already known at the RRH [17].

So, we need users to transmit an orthogonal set of pilot signals during channel estimation [36]. The RSS p_{mk} of user k can then be obtained from (1) [34] as

$$p_{mk} = \rho h_{mk} |q_{mk}|^2 \tag{3}$$

From (3), we can see that the RSS varies due to small-scale fading and shadowing of the wireless channel. The variation of small-scale fading can be decreased

by averaging it over multiple time-slots according to the channel hardening effect [37]. But the shadowing effect, which is position-dependent and therefore depends on the user location, cannot be averaged out [38]. Therefore, the RSS between user k and RHH m, which is obtained from (2) and (3), when converted to dB scale, is given [34] by

$$p_{mk}^{dB} = p_0^{dB} - 10\alpha \log_{10}(d_{mk}) + z_{mk} \tag{4}$$

where $p_0^{dB} = 10\log_{10}(\rho b_0)$ is the uplink RSS at reference distance d_0. Once the per-user RSS values p_{mk}, $m = 1,...,M$ and $k = 1,...,K$, are extracted as above, the CU uplink RSS vector $\mathbf{p}_k$ is given by

$$\mathbf{p}_k = [p_{1k}^{dB}, p_{2k}^{dB}, \ldots, p_{Mk}^{dB}]^T \tag{5}$$

which is considered as the fingerprint.

4 A Clustering and Deep Learning Approach-Based Fingerprinting

An overview of the structure of the proposed localization method is shown in Fig. 2, which consists of two distinct modes: the offline mode and the online mode.

During the offline mode, the system captures the RSS fingerprints from a grid of known location reference points (RPs). Then, each fingerprint is labeled with corresponding location coordinates. The labeled data is divided into training, validation, and testing datasets. In the proposed method, learning is done with the training dataset and the performance is checked with the validation dataset. The accuracy of the localization system is then presented based on the testing dataset. Then PCA is applied for dimension reduction. After that, only a subset of dimensions (features) that have the maximum variance is selected. Therefore, an efficient feature set is produced. In the clustering step, the reduced-dimension training data is divided into several clusters using an efficient clustering method, which is based on K-means and AP. Later a cluster identification is employed for cluster matching and coarse localization. Ultimately, a DNN regression is trained for each cluster based on their similar data distributions. The accuracy of the model is evaluated using the validation dataset. If the accuracy of the proposed model is not sufficient, the clustering and regression parameters are modified in each iteration until convergence is achieved.

During the online mode, we provide the learned model to estimate the unknown locations of the test data. We first input the test data to the preprocessing level to transform the test data and reduce its dimensions. Then, its cluster or regions is identified using the cluster identification algorithm. Finally, the DNN regression of the related cluster is used to estimate the location.

4.1 Offline Mode

Let's assume there are L training data points which correspond to different RPs. Therefore, the radio map is formed as

$$\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_L]^T = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,M} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{L,1} & p_{L,2} & \cdots & p_{L,M} \end{bmatrix} \tag{6}$$

where each row $\mathbf{p}_l$ of radio map $\mathbf{P}$ is an M-dimensional fingerprint that corresponds to the training x-coordinates x_l and the training y-coordinates y_l, $l = 1, \ldots, L$.

$$\begin{aligned} \mathbf{x} &= [x_1, x_2, \ldots, x_L]^T, \\ \mathbf{y} &= [y_1, y_2, \ldots, y_L]^T \end{aligned} \tag{7}$$

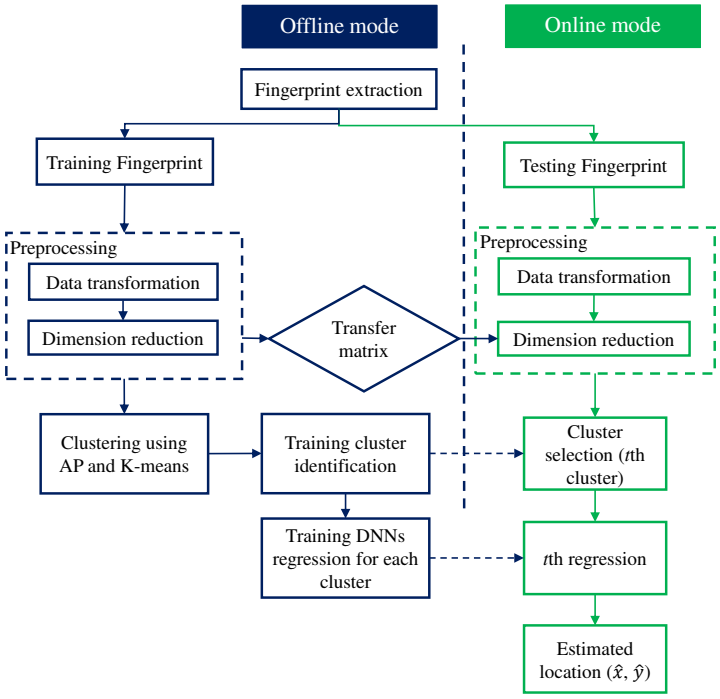


Fig. 2: Overview of the proposed position estimation scheme.

4.1.1 Pre-processing

The core step in data mining and machine learning is data pre-processing, which consists of data transformation and noise reduction, and dimensionality reduction with PCA in our method [39]. Each step is explained in detail as follows.

– *Data Transformation*
In data transformation, each feature value $p_{l,m}, m = 1, \ldots, M$ of each training fingerprint $\mathbf{p}_l$ is standardized [39] using

$$s_{l,m} = \frac{p_{l,m} - \mu_m}{\sigma_m} \tag{8}$$

where μ_m and σ_m are the mean and the standard deviation of the mth feature. With standardization, the training vector $\mathbf{p}_l$ converts to $\mathbf{s}_l$ where $l = 1, \ldots, L$, and the radio map $\mathbf{P}$ is changed into $\mathbf{S} \in \mathbb{R}^{L \times M}$.

– *Dimensionality reduction*
PCA is employed for denoising and dimension reduction in order to map the standardized radio map $\mathbf{S}$, which is in an M-dimensional space, to an D-dimensional space such that $D < M$, while the most relevant information is maintained [40]. In PCA, in order to find the principal components (PCs) in the new feature space, we need to compute the eigenvalues λ_i and eigenvectors $\mathbf{e}_i$ of the covariance matrix of $\mathbf{S}$, $\mathbf{\Sigma} = \mathbf{S}^T\mathbf{S}$, where $\mathbf{\Sigma} \in \mathbb{R}^{M \times M}$ and satisfying $\mathbf{\Sigma}\mathbf{e}_i = \lambda_i \mathbf{e}_i$. The details for computing the eigenvalues λ_i are presented in [40].
Let us assume there are I eigenvalues which are placed in descending order form the diagonal matrix $\mathbf{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_I]$. Choosing the D largest eigenvalues of $\mathbf{\Sigma}$, we form eigenvector matrix $\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_D]$ where $\mathbf{E} \in \mathbb{R}^{M \times D}$. Each column of $\mathbf{E}$ outlines the PCs, which are orthogonal to each other and in decreasing order. More precisely, the first eigenvector $\mathbf{e}_1$ is the direction that captures the maximum variance of data. The second eigenvector $\mathbf{e}_2$ is the direction that has the greatest variance among those that are orthogonal to the first eigenvector, and so on. Therefore, the low dimensionality radio map is $\mathbf{U} = (\mathbf{S})\mathbf{E}$, where $\mathbf{U} \in \mathbb{R}^{L \times D}$ and is defined as

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_L]^T \tag{9}$$

where $l = 1, 2, \ldots, L$.

4.1.2 Clustering

A clustering algorithm is required to split the whole area into several regions based on the collected RSS data. K-means is a very popular clustering algorithm that is extensively used due to its fast convergence. However, it is sensitive to the initial condition wherein the number of clusters is predefined, and a random set of initial exemplars is selected in advance. Therefore, in

K-means, many runs are needed to get a good clustering result. However, it does not guarantee i) that an appropriate initialization will occur during the repetitious running, and ii) unique clustering because we get different results with randomly chosen initial clusters.

In contrast, the affinity propagation (AP) clustering algorithm [41], has the initialization-independent property wherein all RSS samples have an equal chance to be a cluster head (CH). The optimal number of clusters is then obtained by passing iteratively two kinds of messages, named validity and responsibility, to maximize a fitness function until a good set of CHs emerges [41]. Therefore, APC can provide a good set of CHs with high speed. However, it can sometimes fail to converge, particularly for large similarity matrices. Considering the convergence property of K-means and the good performance of affinity propagation, a new clustering method is employed. The APC algorithm is first used to determine the optimal number of clusters and the initial CHs. Then, K-means is employed to create the final clustering results by iteration based on the initial CHs.

As mentioned in [9], the AP clustering algorithm requires two inputs to divide $\mathbf{U}$ into clusters: the similarities matrix $\mathbf{S}_{Sim}$ and the preference $pref$, which are defined as follows [41]:

$$\begin{aligned} s_{Sim}(\mathbf{u}_l, \mathbf{u}_{l'}) &= -||\mathbf{u}_l - \mathbf{u}_{l'}||^2 \\ pref &= \text{median}(s_{Sim}(\mathbf{u}_l, \mathbf{u}_{l'})) \end{aligned} \tag{10}$$

where $1 \leq l, l' \leq L$ and $l \neq l'$.

The cluster algorithm used in our research is summarized in Algorithm 1. The validity of clustering is measured by the silhouette which is a well-known measure of how similar a training RSS vector $\mathbf{u}_l,\ l = 1, \ldots, L$ is to its own cluster (cohesion) compared to other clusters (separation). It is averaged over all training RSS vectors and is defined [42] as

$$SI = \frac{1}{L} \sum_{l=1}^{L} \frac{d(\mathbf{u}_l) - f(\mathbf{u}_l)}{\max\{d(\mathbf{u}_l), f(\mathbf{u}_l)\}} \tag{11}$$

where $d(\mathbf{u}_l)$ is the average distance between $\mathbf{u}_l$ and all training RSS samples in other clusters and $f(\mathbf{u}_l)$ is the average distance between $\mathbf{u}_l$ and all training RSS samples in the same cluster. A clustering which has sufficient SI value is regarded as a valid clustering.

4.1.3 Cluster selection

To determine the cluster corresponding to a new data point, a cluster identification based on KD-tree is used, and its accuracy is evaluated based on the valid clustering. For this purpose, a subset of clustered training RSS samples is selected as a validation dataset, and it is supposed that their cluster ID are unknown. This subset is used to obtain the accuracy of the cluster identification algorithm. First, the KD-tree algorithm, which finds similar data

quickly, is employed for each cluster. Then the KD-tree uses each RSS sample of the validation dataset to find their K_{nn} nearest neighbors from each cluster; among them, the one that has the minimum distance is selected. Therefore, the predicted cluster ID of the validation RSS sample will be the cluster ID of its nearest neighbor.

To estimate the accuracy of clustering, we need to determine the error of cluster membership by comparing the predicted and the real cluster ID of the validation RSS samples. A threshold is considered for the accuracy of the cluster identification algorithm. If the accuracy of the valid clustering is less than the threshold, that clustering is ignored. Otherwise, its validity check and the number of clusters of the clustering which has the highest validity and sufficient accuracy requirements for cluster identification is selected as the best number of clusters.

Algorithm 1 Algorithm for clustering based on APC and K-means

Require:
Preference value $pref$
The training matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_L]^T$
1: Compute similarity matrix $\mathbf{S}_{Sim}$
2: Run APC algorithm to get CHs $C = \{C_1, C_2, ..., C_T\}$
3: Calculate the number of clusters T and the initial centers for the K-means clustering
4: Run K-means clustering

4.1.4 Regression

In order to have very precise localization, we use a DNN to solve the regression problem in each cluster.

Fig. 3 shows the full connected multi-layer neural network structure in our method, which consists of an input layer, hidden layers, and an output layer. The input layer consists of artificial input nodes and receives the initial data for further processing. After dimension reduction, each RSS vector is composed of D RSS values. Therefore, the number of input nodes is equal to the number of dimensions of each RSS vector. The output layer produces the required output. The position information of the MT is set as output. Therefore, the number of output nodes is two. The hidden layers are between the input and output layers, where the transitional computations are performed. Each hidden layer uses the output of the previous layer to perform a non-linear operation, which is defined as:

$$\mathbf{h}^k = \phi(\mathbf{W}^k \mathbf{h}^{k-1} + \mathbf{b}^k) \tag{12}$$

where $\mathbf{W}^k$ is a fully connected weight matrix that represents all the connections between each node of the $(k-1)$th layer and each node of the kth layer. $\mathbf{b}^k$ is the bias vector of the kth layer, $\mathbf{h}^{k-1}$ represents the output from the previous layer. The weights and biases in a neural network are initially set to random values but the model is trained using the back-propagation (BP)

method and the Adam optimizer [43] to minimize the loss function and the network parameters (i.e., weights and biases) are updated iteratively until convergence is achieved. $\phi(.)$ is the activation function and in our case, we use a Rectified Linear Unit (ReLU) [44] (i.e., $\phi(x) = \max(x; 0)$) in the hidden layers and a linear function (i.e., $\phi(x) = x$) in the output layer, since the localization is a regression problem.

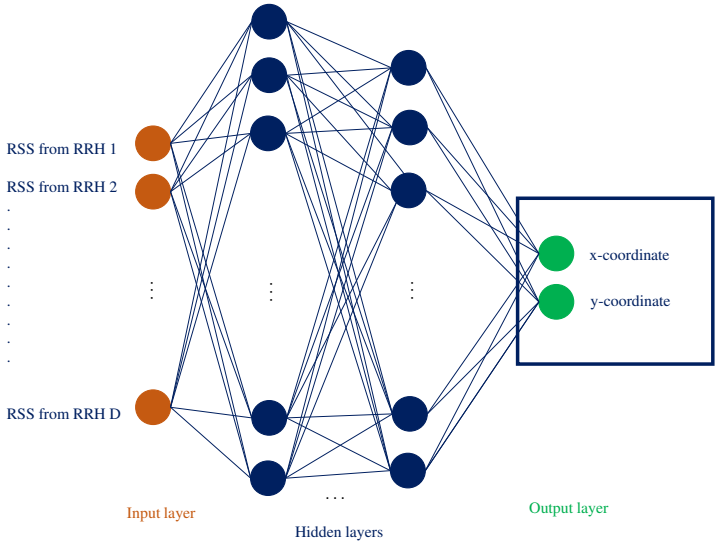


Fig. 3: The DNN structure.

4.2 Online Mode

In this phase, the position of a test user whose location is unknown is estimated. Let us suppose there are $\hat{L}$ test users. After using dimensionality reduction, the $\hat{L} \times D$ testing matrix $\hat{\mathbf{U}}$ is used to estimate the $\hat{L} \times 2$ location coordinates $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$. For each test user data $\hat{\mathbf{u}}$, the process of location estimation is described below.

- *Step 1: Cluster selection*
  The cluster ID t of testing data point $\hat{\mathbf{u}}$ is determined using the cluster selection algorithm.
- *Step 2: Location estimation*
  The DNN regression model of cluster t is used to estimate the location $(\hat{x}, \hat{y})$ of fingerprint $\hat{\mathbf{u}}$.

5 Performance Evaluation

In this section, we compare using simulations, the location estimation using GPR [17], APC-GPR [9], and the proposed method in a DM-MIMO system with $M = 36$ single antenna RRHs, $L = 400$ training locations, and $\hat{L} = 16$ test users. Training users are distributed every 10 m in a grid configuration over the whole area of 200 m $\times$ 200 m. Test users are distributed in a random configuration, as shown in Fig. 4. For training, the RSS matrix $\mathbf{P}$ is generated using (4) with user transmit power $\rho = 21$ dBm, reference path loss $b_0 = -47.5$ dB and different shadowing noise variance $\sigma_z^2 = 1, 3, 5$ dB. Also, we set the path loss exponent to $\alpha = 0$ for $0 \leq d_{mk} < 10$ m, $\alpha = 2$ for 10 m $\leq d_{mk} < 50$ m, and $\alpha = 6.7$ for 50 m $\leq d_{mk}$, according to the 3GPP urban micro propagation model [45]. The PCA technique is applied on $\mathbf{P}$ to reduce the dimensions of RSS vectors and to generate the transformed RSS radio map $\mathbf{U}$. Then the clustering algorithm is applied to cluster all the training data using a similarity matrix and preference values which are generated by (10). When the optimal number of clusters is obtained, K-means algorithm is run for 100 times where the number of clusters is equal to 6. Also, the KD-tree is evaluated with different K_{nn}. Finally, the proposed DNN model is trained with different hidden layers and activation functions. The root-mean-squared error (RMSE) between the real coordinates (x_l, y_l) of the test users and their estimates $(\hat{x}_l, \hat{y}_l)$ is considered as a performance metric, which is defined as

$$RMSE = \sqrt{\frac{\sum_{l=1}^{\hat{L}} (x_l - \hat{x}_l)^2 + (y_l - \hat{y}_l)^2}{\hat{L}}} \tag{13}$$

The RMSE, is averaged over the Monte-Carlo realizations. Lower RMSE values indicate better location estimation performance.

5.1 Preprocessing

For an efficient and accurate clustering algorithm, we need to extract the essential RSS values received by RRHs by reducing the noise and the high dimensionality of the data. As mentioned, the PCA is employed in the offline mode to extract the more important feature set from the original RSS data set, while assuring the same level of positioning accuracy. Also, each RSS sample vector in the online mode is transformed into its low-dimension representation and is then compared with the corresponding low-dimension radio map. Using PCA, the optimal number of components that capture the greatest variance in the data are found. In this work, a 98% variance criterion is considered. Note that different variance thresholds may be chosen depending on the applications specific requirements. Fig. 5 shows how the variance is captured by principal components. We see that the first three components explain the majority of the variance in our data.

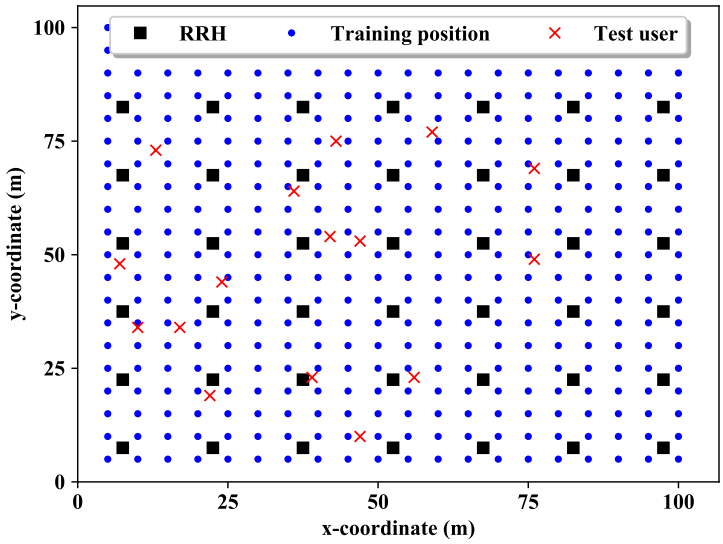


Fig. 4: Simulation setup with $M = 36$ single antenna RHHs, $L = 400$ training positions, and $\hat{L} = 16$ test users.

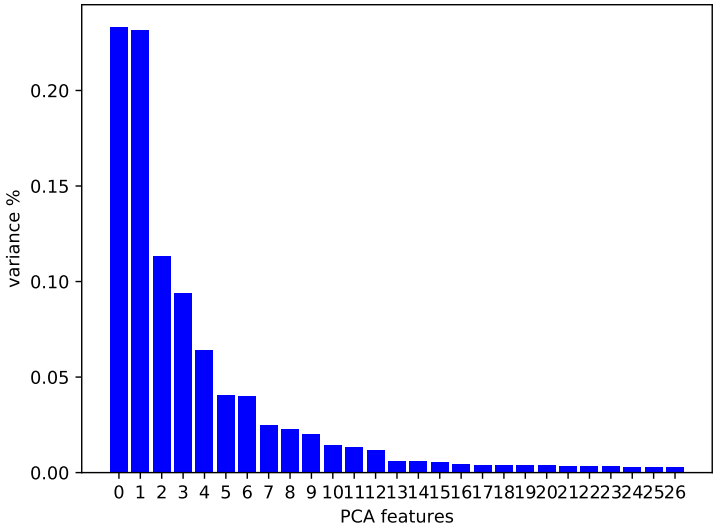


Fig. 5: Variance of the principal components.

From Fig. 6, we can see that with the first 27 components, 98% of the variance is contained. Therefore, the number of principal components is set at 27, and the dimension of the original data is reduced.

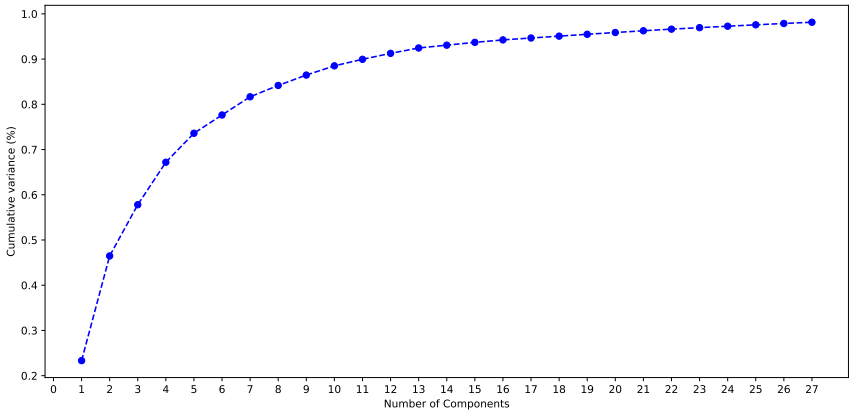


Fig. 6: The cumulative sum of PCA components' variance. The first component already contains more than 20% of the total variance, 27 components take into account 98% of the RSS.

5.2 Clustering

By running AP, the optimal number of clusters is equal to 6, which is considered in the k-means clustering as the input. In case there are 6 clusters, we have a maximum of silhouette value. Also, the KD-tree algorithm with $K_{nn} = 3$ is used for cluster identification as considered in [9].

5.3 Comparison of Different Localization Methods

Fig. 7 shows the average RMSE of the test user's location estimation as a function of different shadowing noise variance ranging from 1 dB to 5 dB for different methods. We can see that the average RMSE is increased by increasing the shadowing noise variance in all methods. When we apply PCA in GPR and AP-GPR methods in [17] and [9] respectively, although a similar increase in average RMSE is observed by increasing the shadowing noise variance, the methods where PCA is used have lower average RMSE than those with no PCA and also the proposed method in the current study has a significantly lower average RMSE compared to the others. Also, we can see that the proposed method has a superior performance compared to the previous methods. Using PCA reduces noise and the number of dimensions of the data, that leading to increase stability, and reduces the computational complexity.

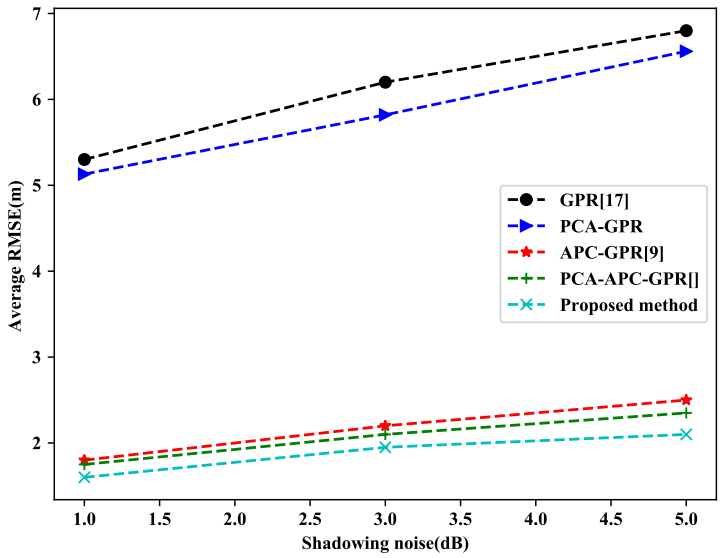


Fig. 7: Average RMSE of using GPR [17], AP-GPR [9] and the proposed methods with $M = 36$, when the shadowing noise variance is 1, 3, and 5 dB and $L = 400$.

6 Conclusion

We proposed an efficient and low dimension FP-based method using PCA, APC and k-means, and DNN to estimate the user's location based on RSS values in a DM-MIMO system. In the proposed method, after preprocessing the data such as denoising and dimension reduction, the whole testbed was first divided into clusters using the AP and k-means algorithms, which reduces the computational cost of online positioning. AP was chosen for clustering due to its initialization-independent property and a better selection of CHs and k-means was combined with AP due to its great convergence. Then, KD-tree was used for cluster identification to allow for a quick finding of the related cluster. Also, DNN was applied for further location estimation within each cluster. The proposed method was compared to the previous works in terms of localization accuracy. Numerical results have justified our proposed localization system over previous schemes. Also, through simulations, we showed that increasing the shadowing noise variance decreases localization performance.

Acknowledgment

The authors would like to thank NSERC for supporting this research.

References

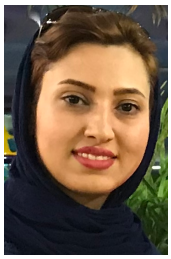
1. Liu, Y., Shi, X., He, S., & Shi, Z. (2017) Prospective Positioning Architecture and Technologies in 5G Networks. *IEEE Network* 31(6):115–121
2. F. Zafari, A. Gkelias, and K. K. Leung (2019) A survey of indoor localization systems and technologies. *IEEE Communications Surveys and Tutorials* 21(3):2568–2599
3. Alkhayrat, M., Aljnidi, M., & Aljoumaa, K. (2020) A comparative dimensionality reduction study in telecom customer segmentation using deep learning and PCA. *Journal of Big Data* 7(1):9
4. Nerem, R. S., & Larson, K. M. (2012) *Global positioning system: theory and practice*. Springer Science & Business Media.
5. Cui, Y., & Ge, S. S. (2003) Autonomous vehicle positioning with gps in urban canyon environments. *IEEE transactions on robotics and automation* 19(1):15–25
6. Chen, C. M., Blandino, S., Gaber, A., Desset, C., Bourdoux, A., & et al. (2017) Distributed massive MIMO: A diversity combining method for tdd reciprocity calibration. In: IEEE Global Communications Conference, pp. 1–7
7. Kamga, G. N., Xia, M., & Aïssa, S. (2016) Spectral-efficiency analysis of massive MIMO systems in centralized and distributed schemes. *IEEE Transactions on Communications* 64(5):1930–1941
8. Joung, J., Chia, Y. K., & Sun, S. (2014) Energy-efficient, large-scale distributed-antenna system (L-DAS) for multiple users. *IEEE Journal of Selected Topics in Signal Processing* 8(5):954–965
9. Moosavi, S. S., & Fortier, P. (2020) Fingerprinting localization method based on clustering and Gaussian process regression in distributed massive MIMO systems. In: International Symposium on Personal, Indoor and Mobile Radio Communications, IEEE
10. Hu, A., Lv, T., Gao, H., Zhang, Z., & Yang, S. (2014) An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems. *IEEE Journal of Selected Topics in Signal Processing* 8(5):996–1011
11. Shaikh, S. A., & Tonello, A. M. (2016) Localization based on angle of arrival in EM lens-focusing massive MIMO. In: IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin), IEEE, pp. 124–128
12. Lv, T., Tan, F., Gao, H., & Yang, S. (2016) A beamspace approach for 2-D localization of incoherently distributed sources in massive MIMO systems. *Signal Processing* 121:30–45
13. Shahmansoori, A., Garcia, G. E., Destino, G., Seco-Granados, G., & Wymeersch, H. (2015) 5G position and orientation estimation through millimeter wave MIMO. In: IEEE Globecom Workshops, IEEE, pp. 1–6
14. Guerra, A., Guidi, F., & Dardari, D. (2015) Position and orientation error bound for wideband massive antenna arrays. In: IEEE International Conference on Communication Workshop (ICCW), IEEE, pp. 853–858
15. Garcia, N., Wymeersch, H., Larsson, E. G., Haimovich, A. M., & Coulon,

M. (2017) Direct localization for massive MIMO. *IEEE Transactions on Signal Processing* 65(10):2475–2487
16. Wei, T., Zhou, A., & Zhang, X. (2017) Facilitating robust 60 GHz network deployment by sensing ambient reflectors. In: 14th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 17), pp. 213–226
17. Savic, V., & Larsson, E. G. (2015) Fingerprinting-based positioning in distributed massive MIMO systems. In: Vehicular Technology Conference (VTC Fall), pp. 1–5
18. Prasad, K. S., Hossain, E., Bhargava, V. K., & Mallick, S. (2018) Analytical approximation-based machine learning methods for user positioning in distributed massive MIMO. *IEEE Access* 6:18431–18452
19. Larsson, J. (2015) Distance estimation and positioning based on Bluetooth low energy technology
20. Li, M. & Lu, Y. (2008) Angle-of-arrival estimation for localization and communication in wireless networks. In: 16th European Signal Processing Conference, IEEE, pp. 1–5
21. C. Gao, G. Zhao and H. Fourati (2019) *Cooperative Localization and Navigation: Theory, Research and Practice*, 1st edn. CRC Press
22. Chan, Y., & Ho, K. (1994) A simple and efficient estimator for hyperbolic location. *IEEE Transactions on signal processing* 42(8):1905–1915
23. Vo, D.Q., & De, P. (2015) A survey of fingerprint-based outdoor localization. *IEEE Communications Surveys & Tutorials* 18(1):491–506
24. Ergen, S. C., Tetikol, H. S., Kontik, M., Sevlian, R., Rajagopal, R., & Varaiya, P. (2013) RSSI-fingerprinting-based mobile phone localization with route constraints. *IEEE Transactions on Vehicular Technology* 63(1):423–428
25. Zhang, A., Yuan, Y., Wu, Q., Zhu, S., & Deng, J. (2017) Wireless RSSI fingerprinting localization. *Signal Processing* 131:235–244
26. Zhang, A., Yuan, Y., Wu, Q., Zhu, S., & Deng, J. (2015) Wireless localization based on RSSI fingerprint feature vector. *International Journal of Distributed Sensor Networks* 11(11):528747
27. Wang, X., Gao, L., Mao, S., & Pandey, S. (2016) CSI-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology* 66(1):763–776
28. Chen, H., Zhang, Y., Li, W., Tao, X., & Zhang, P. (2017) ConFi: Convolutional neural networks based indoor wi-fi localization using channel state information. *IEEE Access* 5:18066–18074
29. Mao, G., & Fidan, B. (2009) *Localization Algorithms and Strategies for Wireless Sensor Networks: Monitoring and Surveillance Techniques for Target Tracking*. NY, USA: Hershey
30. Vieira, J., Leitinger, E., Sarajlic, M., Li, X., & Tufvesson, F. (2017) Deep convolutional neural networks for massive mimo fingerprint-based positioning. In: IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), IEEE, pp. 1–6
31. Kumar, S., Hegde, R. M., & Trigoni, N. (2016) Gaussian process regression

for fingerprinting based localization. *Ad Hoc Networks* 51:1–10
32. Yiu, S., & Yang, K. (2015) Gaussian process assisted fingerprinting localization. *IEEE Internet of Things Journal* 3(5):683–690
33. Yang, A., Jing, Y., Xing, C., Fei, Z., & Kuang, J. (2015) Performance analysis and location optimization for massive MIMO systems with circularly distributed antennas. *IEEE Transactions on Wireless Communications* 14(10):5659–5671
34. Prasad, K. S. V., Hossain, E., & Bhargava, V. K. (2017) A numerical approximation method for RSS-based user positioning in distributed massive MIMO. In: IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), IEEE, pp. 1–6
35. Zheng, K., Ou, S., & Yin, X. (2014) Massive MIMO channel models: A survey. *International Journal of Antennas and Propagation* 2014
36. Ngo, H. Q., Ashikhmin, A., Yang, H., Larsson, E. G., & Marzetta, T. L. (2017) Cell-free massive MIMO versus small cells. *IEEE Transactions on Wireless Communications* 16(3):1834–1850
37. Gunnarsson, S., Flordelis, J., Van der Perre, L., & Tufvesson, F. (2018) Channel hardening in massive mimo-a measurement based analysis. In: IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), IEEE, pp. 1–5
38. Zanella, A. (2016) Best practice in rss measurements and ranging. *IEEE Communications Surveys Tuts* 18(4):2662–2686
39. Obaid, H. S., Dheyab, S. A., & Sabry, S. S. (2019) The impact of data pre-processing techniques and dimensionality reduction on the accuracy of machine learning. In: Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON), IEEE, Jaipur, pp. 279–283
40. Jolliffe, I.T. (2002) *Principal Component Analysis, Second Edition*. 2nd ed. Springer series in statistics
41. Frey, B. J., & Dueck, D. (2007) Clustering by passing messages between data points. *Science* 315(5814):972–976
42. Rousseeuw, P. J. (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* 20:53–65
43. Zhang, Z. (2018) Improved Adam optimizer for deep neural networks. In: 26th International Symposium on Quality of Service (IWQoS), IEEE, pp. 1–2
44. Soltanolkotabi, M. (2017) Learning relus via gradient descent. In: Advances in neural information processing systems, pp. 2007–2017
45. Access, E. U. T. R. (2010) *Further advancements for E-UTRA physical layer aspects (release 9)*. Document TR 36.814, 3GPP

Declarations

– Funding: It is provided by Natural Sciences and Engineering Research Council of Canada (NSERC).
– Conflicts of interest: There is no Conflicts of interest.
– Availability of data and material: Not applicable
– Code availability: Not applicable



Seyedeh Samira Moosavi received her M.Sc. degree in Information Technology Engineering, Network Engineering from the Shiraz University of Technology (SUTech), Shiraz, Iran, in 2015. She is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering at Laval University, Quebec, Canada. Her current research focus is on machine learning methods and their applications in 5G technology.

Paul Fortier (S'78-M'89-SM'00) received his B.Sc. degree and his M.Sc. degree in Electrical Engineering from Laval University in 1982 and 1984, respectively and his M.S. degree in Statistics and his Ph.D. degree in Electrical Engineering from Stanford University in 1987 and 1989, respectively. Since 1989, he has been with the Department of Electrical and Computer Engineering at Université Laval where he is currently a full professor. From 1991 to 1996, he was program director for the B.Sc. degree in Computer Engineering and from 1997 to 2003 he was Chairman of the Department of Electrical and Computer Engineering. From 2003 to 2007, he was Associate Dean for Development and Research at the Faculty of Science and Engineering. From 2007 to 2009, he was Vice-president Scientific Affairs and Partnerships at FQRNT (Quebec's granting agency). From 2010 to 2012, he was Vice-President for Research and Innovation at Laval University. He is currently Chair of the Department of Electrical and Computer Engineering at Laval University. His research interests include digital signal processing for communications and the study of complexity and performance tradeoffs in hardware implementations, with applications in wireless communications. He has been involved in the organization of national and international conferences and workshops in these fields. He has done consulting work for several companies and government agencies in Canada. Dr. Fortier is a Fellow of the Engineering Institute of Canada, a Fellow of the Canadian Academy of Engineering and a Senior Member of the Institute of Electrical and Electronics Engineers (IEEE).

Figures

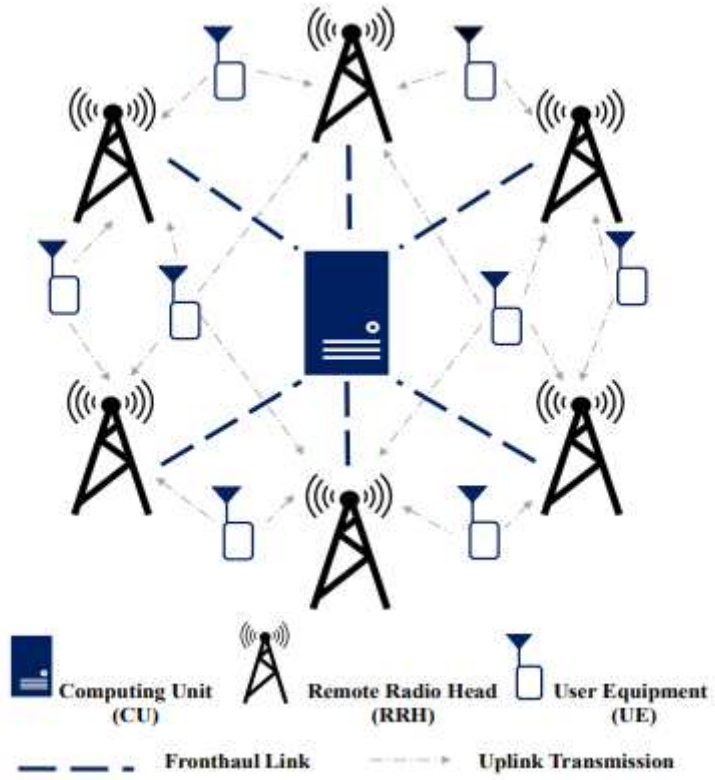


Figure 1

Multi-user DM-MIMO system model for location estimation.

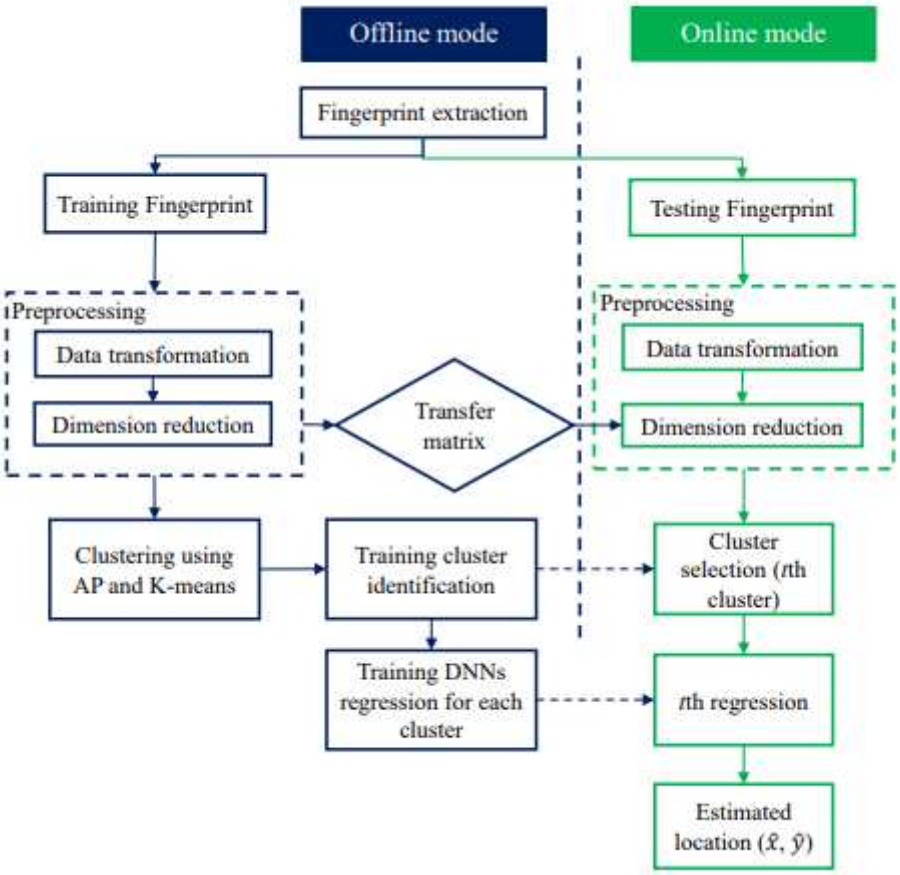


Figure 2

Overview of the proposed position estimation scheme.

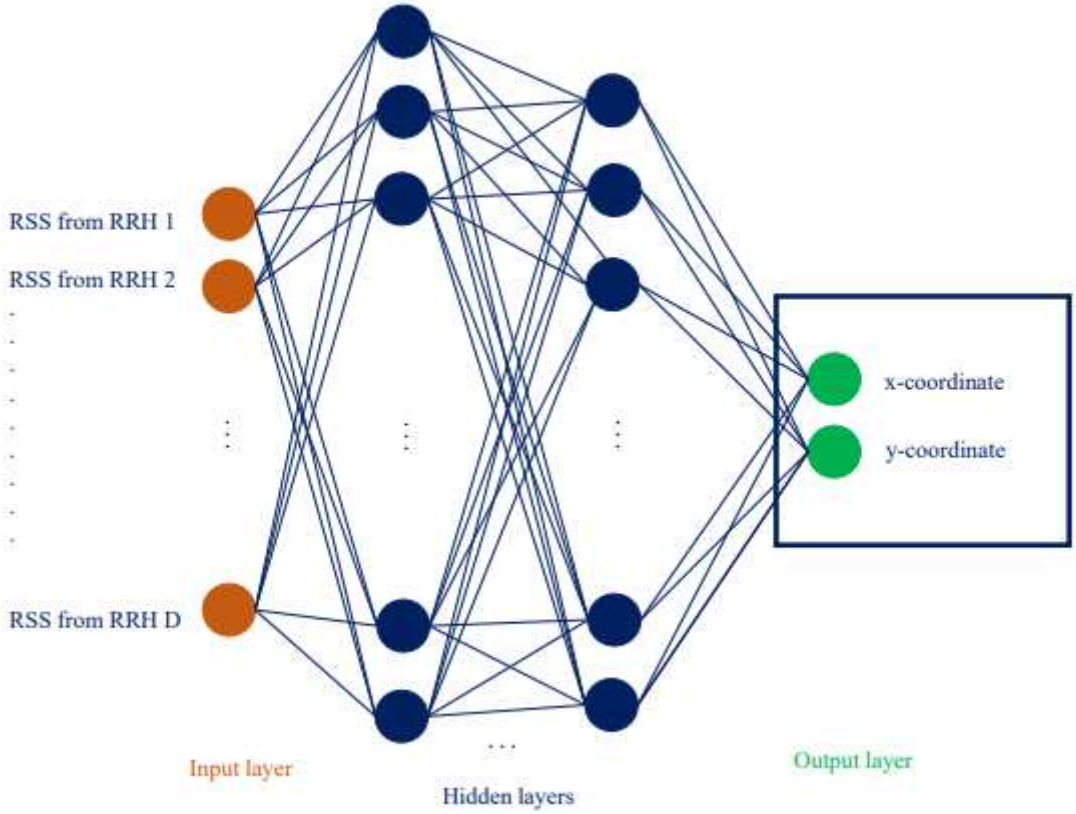


Figure 3

The DNN structure.

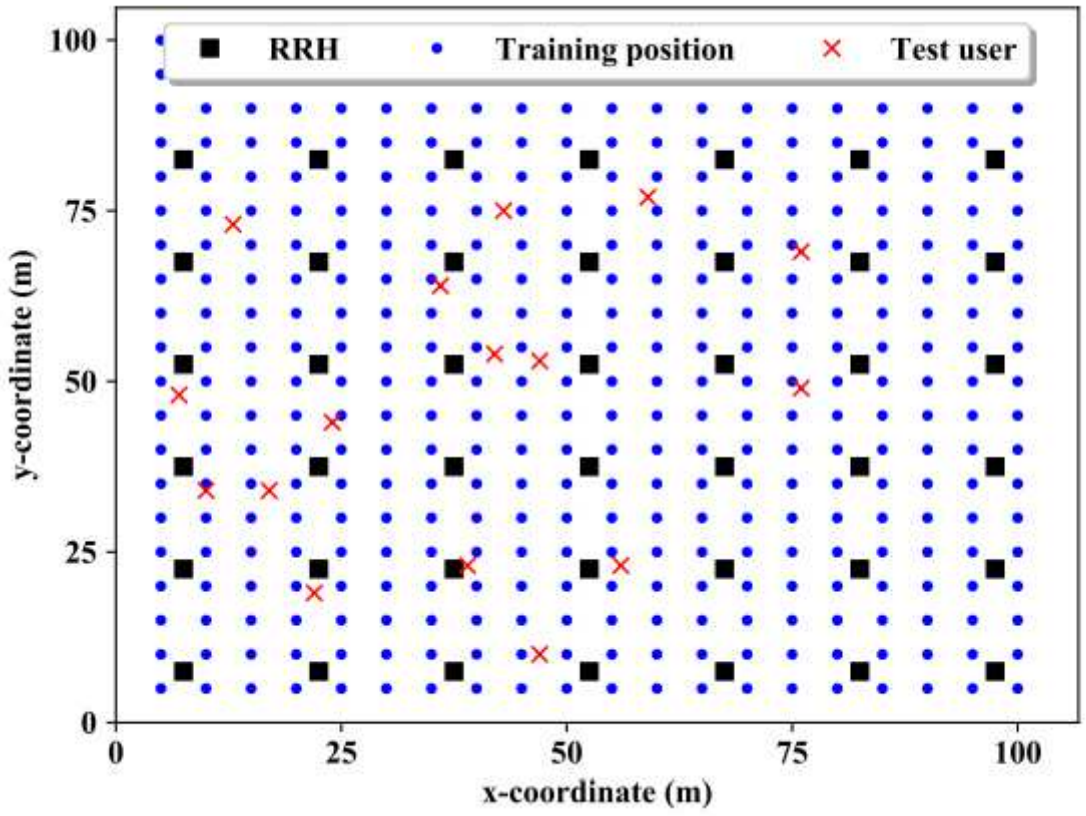


Figure 4

Simulation setup with M = 36 single antenna RHHs, L = 400 training positions, and L^ = 16 test users.

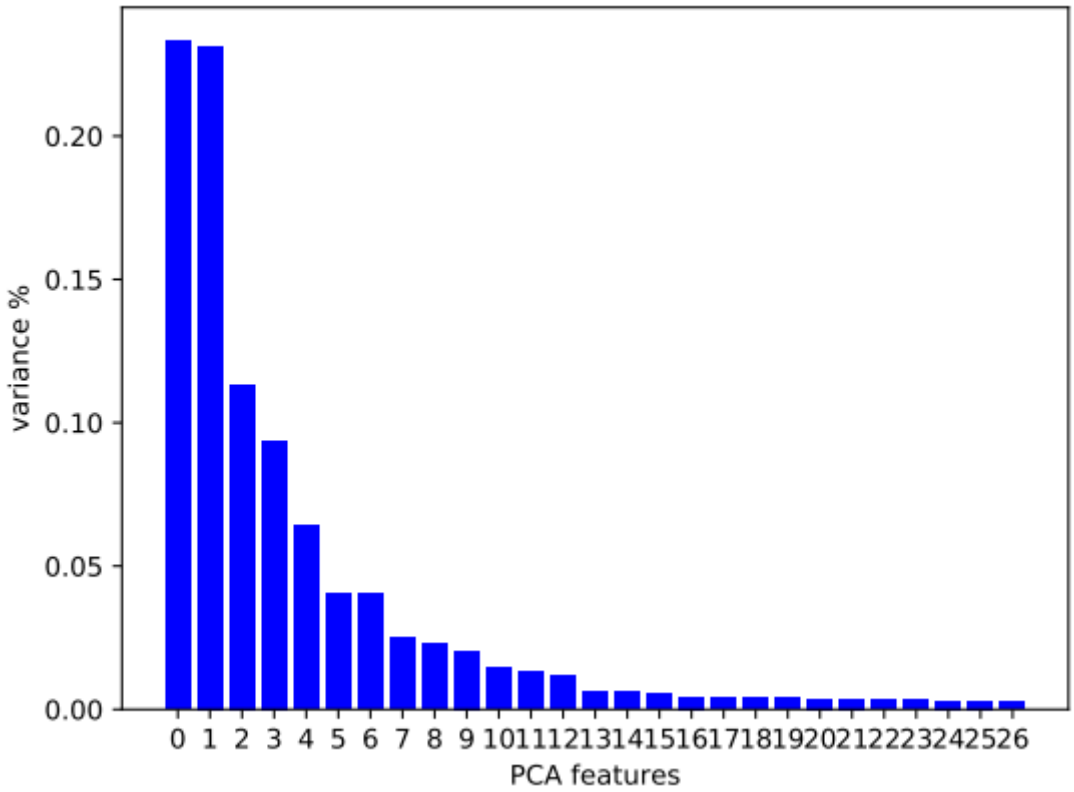


Figure 5

Variance of the principal components.

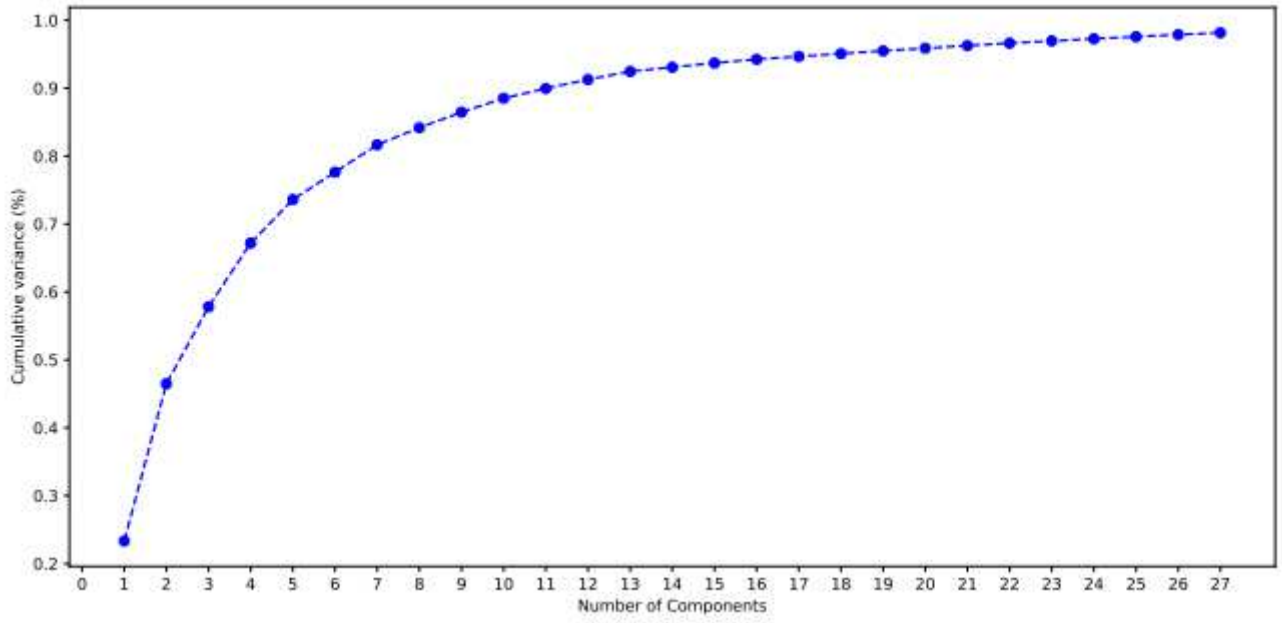


Figure 6

The cumulative sum of PCA components' variance. The first component already contains more than 20% of the total variance, 27 components take into account 98% of the RSS.

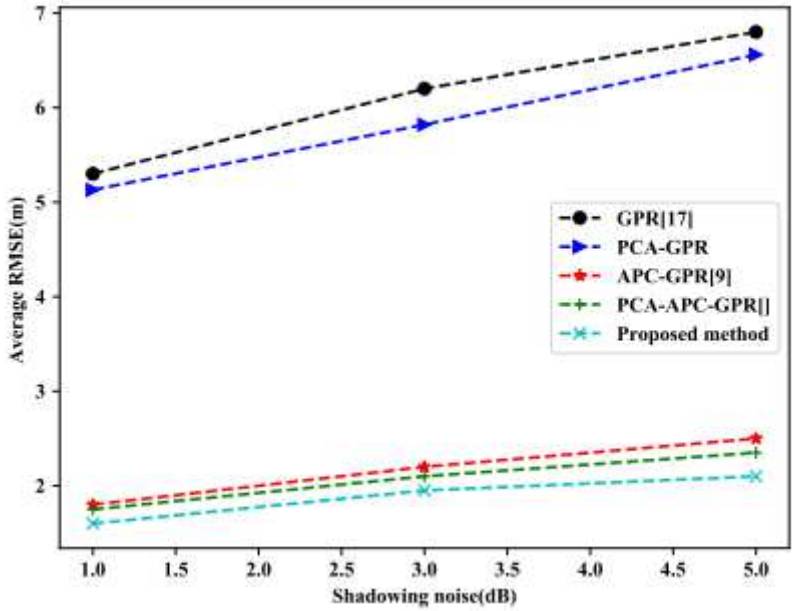


Figure 7

Average RMSE of using GPR [17], AP-GPR [9] and the proposed methods with M = 36, when the shadowing noise variance is 1, 3, and 5 dB and L = 400.