

RESEARCH

Open Access



A 3D mobile positioning method based on deep learning for hospital applications

Qingqing Zhang*  and Yuan Wang

* Correspondence: suiyue2959@163.com

School of Management, Xi'an
Polytechnic University, Xi'an 710048,
People's Republic of China

Abstract

In this study, a 3D positioning method is proposed for hospital applications, such as navigation within a hospital building. It employs deep learning algorithms to analyze the received signal strength from cellular networks and Wi-Fi access points in order to estimate the positions of mobile stations. A two-stage deep learning procedure (level classification and location determination) is constructed to obtain the exact position information (building level, longitude, and latitude) in multiple-level buildings. To evaluate the performance of the proposed method, an experiment was conducted in the hospital of Xi'an Polytechnic University. In total, 36,985 records, 42 sampling location points, 28 different cellular networks, and 289 different Wi-Fi access points were considered. A deep learning neural network was trained for the first stage of level classification. Three deep learning neural networks were trained to obtain the distinct location coordinates (longitude and latitude) for three different building levels. To compare the efficacy of heterogeneous networks, three kinds of neural networks with different inputs (only cellular, only Wi-Fi APs, and a conjunction of cellular and Wi-Fi APs) were implemented. The accuracy of level classification was shown to be 100% for only Wi-Fi APs as an input. The average distance error of the location determination for different floors was 0.28 m for only Wi-Fi APs and for the conjunction of Wi-Fi APs and cellular networks in the second stage.

Keywords: Indoor positioning, Deep learning, Mobile positioning method, Received signal strength

1 Introduction

Global Positioning Systems (GPS) are the most well-known tool in navigation and positioning frameworks. However, they do not usually work in the interior of buildings. In the urban environment, the propagation of GPS satellite signals is hindered by buildings. The “Urban Canyon” effect prevents GPS from accurately predicting indoor positioning.

Due to the complex indoor environment, indoor propagation of signals is more complicated than outdoor propagation. The positioning accuracy is required to be controllable within a few meters to provide users with the maximum utility. In view of the difficulties involved in indoor positioning and the excessive requirements for positioning accuracy, researchers have done a lot of work. These studies involve many

intersecting fields, such as wireless networks, sensor technology, and random signal processing. Ultra-wideband (UWB), radio-frequency identification (RFID), ZigBee, Wi-Fi, etc., were used in indoor positioning systems [1]. For hospital applications, it has been proposed to use UWB sensors in combination with GPS to achieve indoor and outdoor location tracking [2]. In order to track patients in a hospital, RFID technology has been used to solve the problems caused by patient mobility [2]. ZigBee technology has been used to locate certain patients suffering from mental disorders within the confines of the hospital [3]. However, these systems are not without some defects. It is an effective method to integrate UWB and GPS into a positioning system, but the stability of this system is poor because the signal transmission from UWB to GPS depends on the Wi-Fi network [2]. RFID is another positioning system, which needs to install the reader and the label of the item to be identified first. Bluetooth positioning technology requires the installation of multiple local area network access points, and the signal is easy to be affected [3, 4].

Wi-Fi is a popular component in wireless positioning technology in recent years due to various reasons. First, it is ubiquitous modern technology. It has extensive coverage, including shopping malls, schools, and hospitals, among others. Moreover, it is easy to set up a Wi-Fi access point (AP) in rural districts. Second, a MAC address is a unique identifier used to mark a specific Wi-Fi AP, and the signal strength indicator (RSSI) from the specific AP can be received by mobile devices. This implies that the RSSI of Wi-Fi APs could be featured for identification of the signal strength at a specific location. Nowadays, both android and iOS mobile phones have a function to detect Wi-Fi APs. The cellular network is another signal that can be automatically detected by mobile phones. Mobile positioning technology is an excellent method to integrate cellular networks and Wi-Fi APs.

With the development of the social economy, an increasing number of comprehensive buildings have been established in the world, such as large-scale shopping malls, general hospitals, and so on. Particularly in the case of general hospitals, it is necessary to develop an efficient indoor positioning system to help doctors and patients rapidly determine the exact location from the numerous departments, operating rooms, and treatment rooms. The existing indoor positioning system technologies are used in large parking lots, shopping malls, and hospitals to solve the location problem in 2D space. However, there is no pertinent discussion about the positioning technology of multi-story buildings. In this study, an indoor positioning system in 3D space is proposed to tackle the problem of navigation in multi-story buildings. This system can obtain the horizontal information of a plane space, as well as retrieve the vertical information of different floors.

The fingerprint-matching algorithm is commonly used for indoor positioning. The most basic algorithms are the nearest neighbor and naive Bayesian methods. In recent years, deep learning has been extensively used in various fields and exhibited excellent results [5]. In this study, the deep learning algorithm is used in indoor positioning technology to explore its effect on improving the positioning accuracy.

The innovations of this work are highlighted as follows:

- (1) A 3D mobile phone positioning system is proposed to locate multi-story buildings from three dimensions: plane and vertical.

- (2) A two-stage deep learning method is proposed to implement a 3D mobile phone positioning system. This can provide accurate information on the floor, longitude, and latitude of a location.

This study proposes a 3D positioning system for hospital applications, which is based on the integrated signal from cellular signals and Wi-Fi APs. The outline of the paper is as follows. [Section 2](#) illustrates the related work of indoor positioning systems. [Section 3](#) presents further details of the 3D mobile positioning system and deep neural networks used in the system. [Section 4](#) describes the experiment conducted in the campus hospital and illustrates the results. This is followed by the conclusions and exploration of potential future work in [Section 5](#).

2 Related work

Usually, the algorithms for the indoor positioning technology can be divided into two categories: triangulation method and fingerprint method. The triangulation method uses a signal attenuation model to estimate the distance between the mobile device and all the detected APs, and this proportion is used to draw a circle. The intersection of all the circles is the specific location of the device we need to locate. The premise of this method is that the location of the AP must be known in advance. Once the environment changes, the triangulation method does not work [4].

The fingerprint method consists of two phases, the offline phase followed by the online phase. During the offline phase, sampling points (reference samples), which contain the RSSI values of all the detected APs and the coordinates of the known locations, are collected and stored. The collection of sampling points forms the fingerprint database of the surveyed area. During the online phase, the estimated location will be provided by the matching algorithms based on the comparison of the detected RSSI values and the corresponding APs in the database. Many such matching algorithms have been used in fingerprint technology.

The Euclidean distance is commonly used to measure the distance between the observed RSS vectors and sampling points. The matching algorithm estimates the location as the sampling points, which have the smallest distance to the observed signals [6]. Some researchers considered the location estimation to be a machine learning problem. The weighted k-nearest neighbor algorithms were proposed to estimate the position of the target node, based on Bluetooth technology. The estimated position error is approximately 1.8 m, which is too high [7]. To compare the efficacy of different machine learning algorithms, six different machine learning algorithms, including J48, Bayes Net, KNN, SMO, and Adaboost were used. J48 and Bagging with J48 which are included in Weka were used for the UJIIndoorLoc database [8]. A novel ensemble learning method was proposed to provide the building level and indoor localization in buildings. Extensive experiments were conducted in real-world office-like environments, as well as on Android smartphones. It achieved the best indoor landmark localization accuracy of almost 97% in office-like environments. This method can provide a basis for accurate indoor positioning [9]. An ensemble model consisting of fuzzy classifier and multi-layer perceptron was proposed for indoor parking localization [10]. This study employs deep learning algorithms to train the positioning system. Deep learning algorithms have been successfully used in many fields, such as image, transportation, and statistics [11–13].

They have also been used in wireless sensor networks, in an effort to implement positioning systems [14–18]. This study focuses on a 3D mobile positioning system, based on deep neural networks.

3 3D mobile positioning system and deep neural networks

The architecture and concepts of the proposed 3D mobile positioning system are illustrated in Section 3.1 and Section 3.2, respectively.

3.1 3D mobile positioning system

The proposed 3D positioning system includes a (1) Signal receiver, (2) Processor, (3) Performer, (4) Location server, and (5) Model server. The whole system is depicted in Fig. 1. Each component in the proposed system is presented in the following subsections.

As illustrated in Fig. 1, the signal from Wi-Fi and cellular networks are received first. Subsequently, the data pass through the receiver, database, processor, database, and performer.

3.1.1 Signal receiver

The receiver detects and receives the signals which are from the cellular networks and Wi-Fi APs. A mobile phone is a convenient device which can detect the signals both from cellular network and Wi-Fi AP. A mobile application (App) is required to collect the RSSIs and write to a list. All the RSSIs of cellular networks and Wi-Fi APs at one specific location are recorded with the corresponding beacons and MAC addresses. The matrices, constituted by the RSSIs, are the input sources for the neural network model in the process phase.

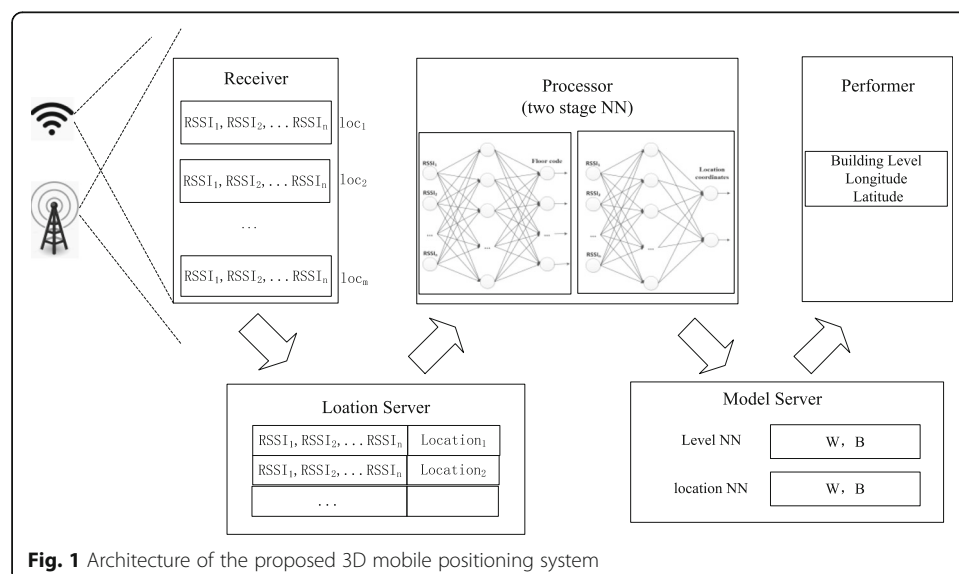


Fig. 1 Architecture of the proposed 3D mobile positioning system

3.1.2 Signal processor

The goal of the signal processor is to construct positioning models. To improve the positioning accuracy, the deep learning neural networks are used to be training algorithms. The received RSSIs (including cellular base stations and Wi-Fi APs) from mobile phones should be normalized before the data are used as an input for the training model. In order to resolve the 3D positioning problem for multilevel building, a two-stage deep learning neural network model is proposed. The first stage is level classification. In this stage, the network is trained for predicting building level (vertical indicator). The normalized RSSI is the input, and the building level is the output. The trained model will be called in the followed performance phase. Then, the building level will be predicted as the first step for the required mobile devices. The second stage is location determination which is trained for predicting the longitude and latitude (horizontal coordinates) for location in every building level. In the training of the second stage, the corresponding normalized RSSI is the input. The GPS coordinates are used as the output of the deep learning models. The building level information and GPS coordinates of the sampling locations are initially stored in the location server. When the deep learning neural networks are trained, these models are sent to the model server to be saved. Therefore, the signal processor component has two functions: normalization of the received signal and training of the deep learning model.

3.1.3 Performer

The function of the performer is based on the processor and model server. When the performer is activated, it receives the new RSSIs vectors and subsequently loads the trained models from the model server. Finally, the estimated location information (building level, longitude, and latitude) is provided.

3.1.4 Location server

The location server is a database, which is used to store RSSIs. The RSSIs are detected by the mobile receiver and wrote to the location server following the corresponding rule with sample points. The sample points are recorded as GPS coordinates.

3.1.5 Model server

The deep learning models trained in signal processor phase are stored in the model server. For the neural networks, the structure and parameters of the models (weights and biases) are stored as a database. The models will be called by performer module when it is needed to predict a location.

3.2 3D mobile positioning method

The 3D mobile positioning method includes (1) collection and normalization, (2) the two-stage neural network, and (3) de-normalization and estimation. Each step in the proposed method is detailed in the following subsections.

3.2.1 Collection and normalization

The data used in the 3D positioning system are collected by mobile receivers, which receive the RSSI of signals from cellular networks and Wi-Fi APs. Before training the

models, some sampling location points are collected. For every sampling location point, the mobile records received the RSSIs for a period of time. During this time period, one location point has multiple records of the RSSIs. The number of records at one location point depends on the writing interval of the mobile receiver, which can be adjusted as per the requirements.

The building level of the sampling location is recorded for a multiple-level building. The use of GPS to obtain the location coordinates (longitude and latitude) of an indoor sampling location does not work well. The transformation formula is used to assist the GPS in obtaining all the sampling location coordinates. First, the location coordinates of the specific location points (e.g., both ends of the building) should be obtained using GPS. The specific location points serve as reference points. The reference points should be at the end of the building. It is better to select the reference points and sample points on the same line. Take the basilica building as an example, selecting both ends of the building is the best choice for reference points. In the basilica building, L^* is noted as the left end point and R^* is noted as the right end point for every floor. The location coordinates of L^* and R^* are $\text{lon}(L^*)$, $\text{lat}(L^*)$ and $\text{lon}(R^*)$, $\text{lat}(R^*)$. The distance of the L^* between R^* is noted as $\text{long}(L^*, R^*)$. For a sampling location point S^* , which is on the same floor as the reference points, the distance between S^* and left L^* is noted as $d(L^*, S^*)$. The longitude and latitude of the points of a sampling location S^* are noted as $\text{lon}(S^*)$ and $\text{lat}(S^*)$, which are computed using Eqs. (1) and (2).

$$\text{lon}(S^*) = \text{lon}(L^*) - (\text{lon}(L^*) - \text{lon}(R^*)) \frac{d(L^*, S^*)}{\text{long}(L^*, R^*)} \quad (1)$$

$$\text{lat}(S^*) = \text{lat}(L^*) - (\text{lat}(L^*) - \text{lat}(R^*)) \frac{d(L^*, S^*)}{\text{long}(L^*, R^*)} \quad (2)$$

It is known that RSSI takes on a value between -150 and 0 . During computing, the input value should be normalized in order to eliminate the dimensional effect. The normalized value for RSSI is computed according to Eq. (3)

$$R_{\text{normalized}} = \frac{\text{RSSI}_{\text{origin}} - \text{RSSI}_{\text{min}}}{\text{RSSI}_{\text{max}} - \text{RSSI}_{\text{min}}} \quad (3)$$

where $R_{\text{normalized}}$ is the normalized value; $\text{RSSI}_{\text{origin}}$ is the received value, which is between -150 and 0 ; and RSSI_{min} and RSSI_{max} are the minimum and maximum values among the original collected data, respectively.

The location coordinates have not yet been normalized to $0-1$. The normalized value $\text{lon}_{\text{normalized}}$ and $\text{lat}_{\text{normalized}}$ are computed by Eqs. (4) and (5).

$$\text{lon}_{\text{normalized}} = \frac{\text{lon}_{\text{origin}} - \text{lon}_{\text{min}}}{\text{lon}_{\text{max}} - \text{lon}_{\text{min}}} \quad (4)$$

$$\text{lat}_{\text{normalized}} = \frac{\text{lat}_{\text{origin}} - \text{lat}_{\text{min}}}{\text{lat}_{\text{max}} - \text{lat}_{\text{min}}} \quad (5)$$

where $\text{lon}_{\text{normalized}}$ and $\text{lat}_{\text{normalized}}$ are the normalized values; and $\text{lat}_{\text{origin}}$ are the values received by solving (2) and (3); lon_{min} and lat_{min} are the minimum values among the

original collected longitudes and latitudes; and lon_{\max} and lat_{\max} are the maximum values.

3.2.2 The two-stage neural network

The processor component is the core of the proposed 3D positioning system. In this phase, the models are trained on the basis of the collected and normalized data. Here, deep learning algorithms are used in conjunction with neural networks to train the model to estimate the location of the building. The proposed 3D positioning system is a two-stage work, particularly for the multiple level buildings. The first stage is level classification, and the second stage is location determination. The GPS coordinates for indoor positioning are difficult to obtain, particularly in a vast building with multiple floors. Some locations share the same GPS coordinates, despite being on different floors in the building. The models in both the stages are trained by neural networks using the deep learning algorithm. The model and methods of the two stages are presented below.

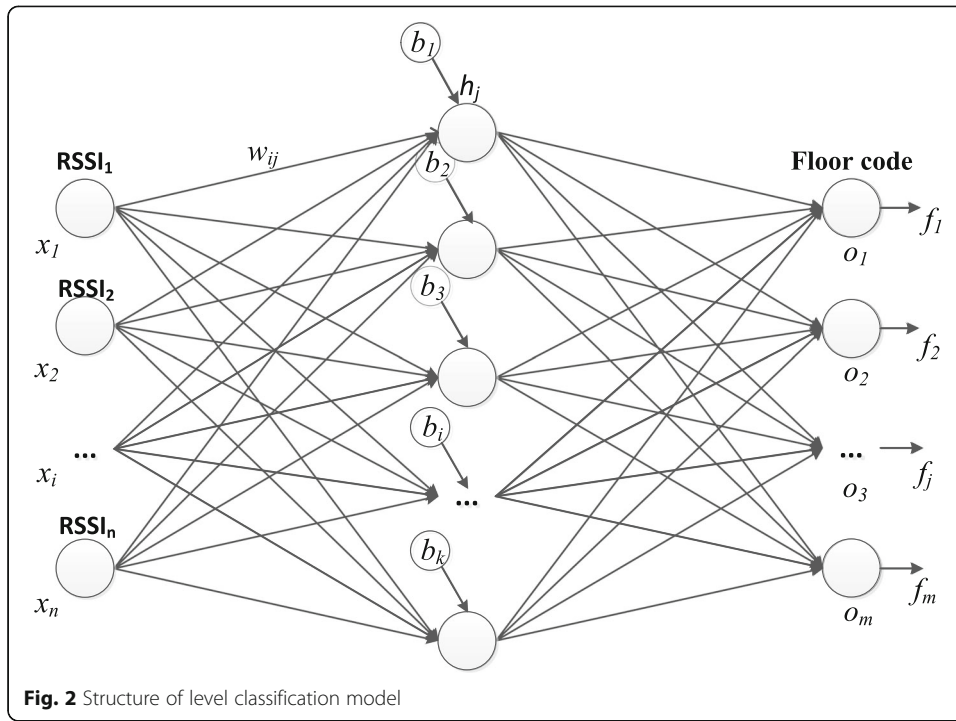
3.2.2.1 Level classification Level classification, which is the basis for location determination, is the first stage in the processor component of the proposed 3D positioning system. Some sampling location points in a building, despite being on different building levels, share the same GPS coordinates (longitude and latitude). Therefore, the first stage plays the role of separating locations in different building levels.

In this model, a three-layer forward neural network (one input layer, one hidden layer, and one output layer) is used. The inputs are the RSSIs collected at every sampling point, and the outputs are the corresponding building level information, which are encoded in 0–1 code. The number of inputs and outputs are the total number of RSSIs and total floors of the building, respectively. The number of hidden neurons is not definite; it can be retrieved by experience. All the neurons between neighboring layers are fully connected (see Fig. 2). The strength of connections is abstracted as weights, and every neuron in the hidden layer and the output layer has a bias, which is used to stimulate the stimulus pulse of the brain.

The input layer includes the normalized RSSIs of n_1 base stations and n_2 Wi-Fi APs. We concatenate them into a vector (x_1, x_2, \dots, x_n) , which is normalized with the original RSSIs. The coding method uses 0–1 coding. For example, in a building of 5 floors, if the position is on the second floor, then the output vector is (0, 1, 0, 0, 0). It is fully connected for all the nodes in the network. The weights between hidden layer and input layer are represented as w_{ij} (weight links input neuron h_i and hidden neuron x_j). The weights between hidden layer and output layer are represented as v_{ij} (weight links hidden neuron o_i and output neuron h_j). The bias of the neurons in the hidden layer and output layer are represented as b_i and b_i' , respectively.

The values of the hidden neuron h_i and the output neuron o_i are computed by Eqs. (6) and (7), respectively. The hidden layer is used to extract the intermediate information contained in the neural network model. The information retrieved by the hidden layer is then used as the input of the output layer (the subsequent layer).

$$h_i = \sum_{j=1}^n w_{ij}x_j + b_j \quad (6)$$



$$o_i = \sum_{j=1}^m v_{ij} h_j + b'_i = \sum_{j=1}^m v_{ij} \left(\sum_{j=1}^n w_{ij} x_j + b_i \right) + b'_i \quad (7)$$

The linear function is selected as the hidden layer activation function of each neuron (Eq. (8)), and the softmax function is selected as the output layer activation function through (Eq. (9)).

$$f(o_i) = o_i \quad (8)$$

$$f(o_i) = \frac{\exp(o_i)}{\sum_{j=1}^m \exp(o_j)} \quad (9)$$

Furthermore, the loss function is defined in Eq. (10). For the optimization of the level classification, the learning rate η and gradient descent method are used to update each weight and bias. The updates of w_{ij} , b_i , v_{ij} , b'_i are calculated by Eqs. (11), (12), (13), and (14), respectively.

$$L(w, v, b, b') = \frac{1}{2} (\hat{y} - y)^2 = \frac{1}{2} \left[\sum_{j=1}^m v_{ij} \left(\sum_{j=1}^n w_{ij} x_j + b_i \right) + b'_i - y \right]^2 \quad (10)$$

$$w_{ij} \leftarrow w_{ij} - \eta \frac{\partial L}{\partial w_{ij}} = w_{ij} - \eta (\hat{y} - y) x_j \sum_{k=1}^m v_{ki} \quad (11)$$

$$b_i \leftarrow b_i - \eta \frac{\partial L}{\partial b_i} = b_i - \eta (\hat{y} - y) \sum_{k=1}^m v_{ki} \quad (12)$$

$$v_{ij} \leftarrow v_{ij} - \eta \frac{\partial L}{\partial v_{ij}} = v_{ij} - \eta(\hat{y} - y)h_i = v_{ij} - \eta(\hat{y} - y) \left(\sum_{k=1}^n w_{ik}x_k + b_i \right) \quad (13)$$

$$b_i' \leftarrow b_i' - \eta \frac{\partial L}{\partial b_i'} = b_i' - \eta(\hat{y} - y) \cdot 1 \quad (14)$$

The training process of the neural network model is also the optimization process. The goal of optimization is to obtain the optimal weights and bias with which the error of the predicted value is the minimum. Therefore, the loss function is defined in order to measure the training error. The optimization process is described in (15) as

$$\left(w^*, v^*, b^*, b'^* \right) = \underset{w, v, b, b'}{\operatorname{argmin}} L(w, v, b, b') \quad (15)$$

To solve the optimization problem in (15), the gradient descent algorithm is used to obtain the optimal parameters (weights and biases). The algorithm is described as follows: The model parameters w_{ij} , b_i , v_{ij} , and b_i' are first initialized with a random number generator. The values of h_i , o_i , and \hat{y} are computed using Eqs. (6), (7), and (9), respectively. Subsequently, Eqs. (11)–(14) are used to adjust the model parameters w_{ij} , b_i , v_{ij} , and b_i' . This step is computed iteratively when one of the following conditions are met: the maximum iterations or the requisite error is attained.

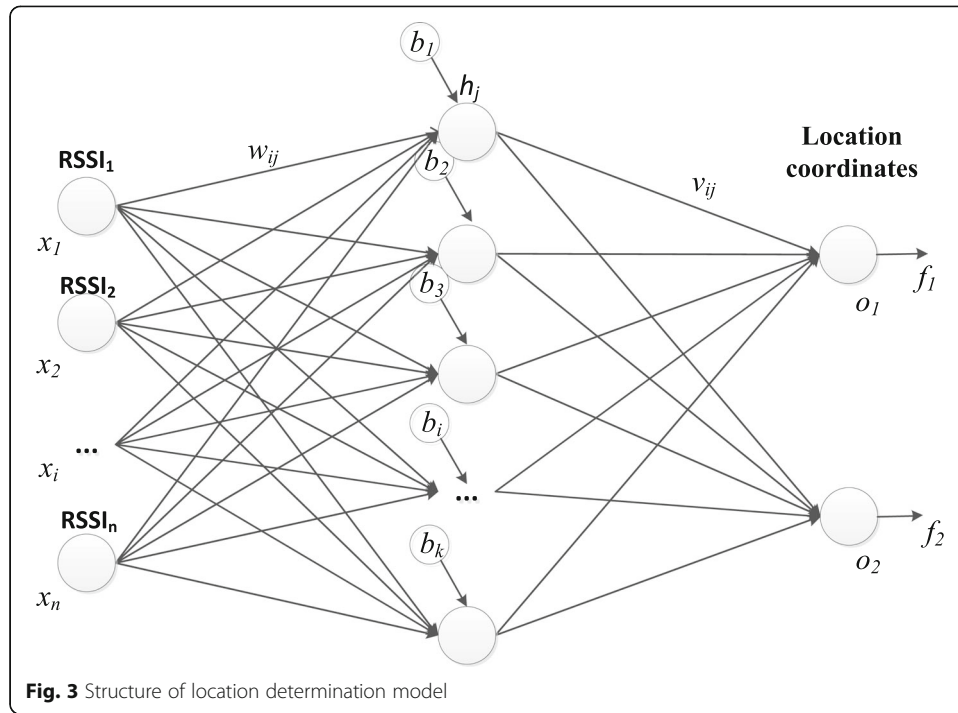
3.2.2.2 Location determination Location determination is the second stage in the data processing by the proposed 3D positioning system. On the basis of the first stage involving location determination, deep neural networks are trained separately for different floors. Therefore, the number of location determination models depends on the number of floors in the building. The structures and the optimization methods for these neural network models are identical. However, they have different inputs and outputs. The structure of the location determination model is presented in Fig. 3.

As depicted in Fig. 3, the structure of the neural network is the same as that of the level classification. It comprises three layers (one input layer, one hidden layer, and one output layer) of a deep neural network. The inputs are the normalized vectors of original RSSIs. Therefore, the same representation is used. The input vectors are represented as (x_1, x_2, \dots, x_n) , which are the same as those in the level classification neural network model. The hidden layer is (h_1, h_2, \dots, h_l) . The input layer and hidden layer are fully connected by weight w_{ij} and bias b_i . The value of h_i is derived from Eq. (6). The input layer and hidden layer are fully connected by weight v_{ij} and bias b_i' .

However, the output of the location determination model is different from that of the level classification model. For location determination, the output is the location coordinates (longitude and latitude). Therefore, the number of output neurons is two. Furthermore, the activation function for the output is linear (Eq. (7)). The output value, which is obtained from Eq. (8), is represented as (o_1, o_2) . Gradient descent (Eqs. (11)–(14)) is used as the optimization algorithm.

3.2.3 De-normalization and estimation

The de-normalized value is obtained from Eqs. (16) and (17) for the estimation of the location coordinates (longitude and latitude).



$$\text{lat}_{\text{denormalized}} = \text{lat} * (\text{lat}_{\text{max}} - \text{lat}_{\text{min}}) + \text{lat}_{\text{min}} \quad (16)$$

$$\text{lon}_{\text{denormalized}} = \text{lon} * (\text{lon}_{\text{max}} - \text{lon}_{\text{min}}) + \text{lon}_{\text{min}} \quad (17)$$

where lat is the value to be denormalized; lon_{min} and lat_{min} are the minimum values among the original collected longitudes and latitudes; and lon_{max} and lat_{max} are the maximum values. Here, lon_{min} , lat_{min} , lon_{max} , and lat_{max} are the same as in Eq. (4).

4 Practical experimental results and discussion

In this section, the practical experimental results are presented and discussed. The practical experimental environments are illustrated in Section 4.1 and the practical experimental results are detailed in Section 4.2. The results for different neural networks are discussed in Section 4.3.

4.1 Practical experiment environment

To validate the proposed 3D positioning method, we conducted an experiment in the school hospital of Xi'an Polytechnic University. The school hospital is a three-story building, containing 37 rooms, including emergency, internal medicine, otolaryngology, X-ray, injection, treatment, pharmacy, and inpatient department. All the rooms are located on these three floors. After considering the significance of each room, every door was used as a sampling location point. Furthermore, the length of the building was measured to be 47 m. There are 13, 14, and 13 rooms on the 3rd, 2nd, and 1st floors, respectively.

In this experiment, an Android application was implemented and installed on mobile stations (e.g., Huawei honor running Android platform 8.0.0). It was tasked with collecting the RSSIs from cellular networks and Wi-Fi networks every second.

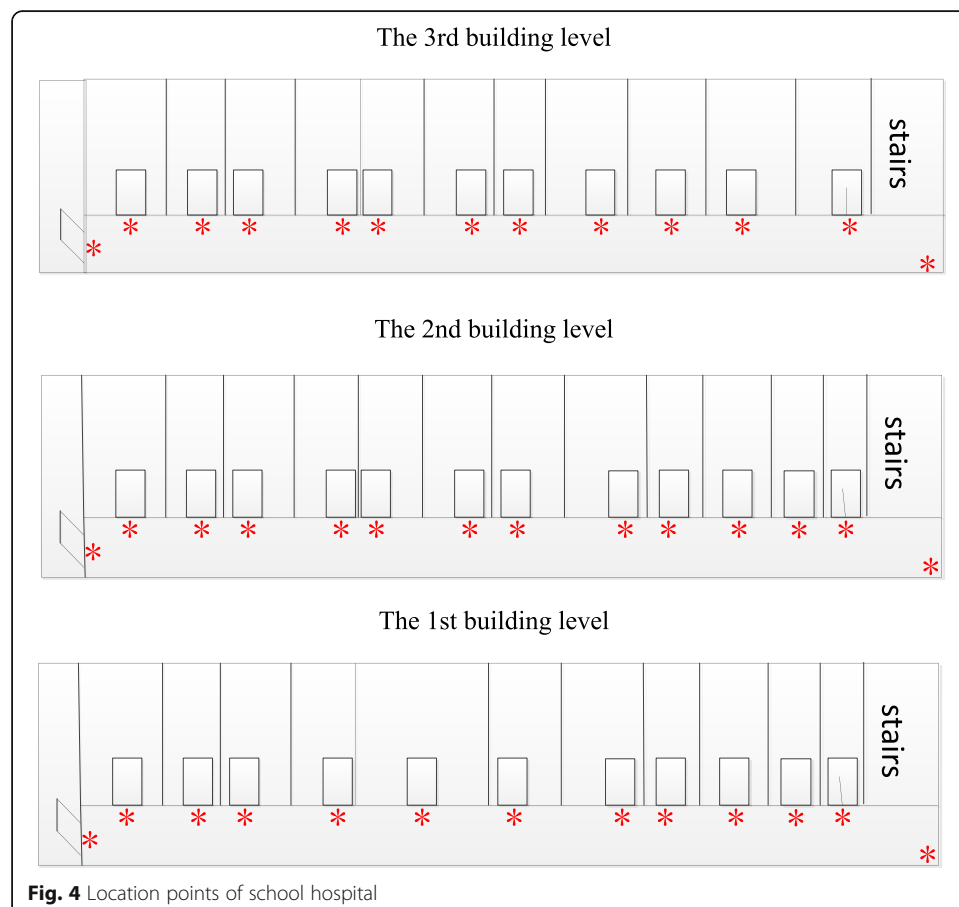
The mobile receiver was situated on the building. Every room was labeled as a sampling point from which data was collected. In addition, 3 sampling points were allocated to the corner of stairs and 2 points to the stairway. A total of 42 sampling points were labeled (see Fig. 4).

We allotted a time of approximately 30 s for the sampling of each location point. It is guaranteed that there are at least 30 records for every location point sampled. Finally, a total of 1527 records were collected. In order to maintain the reliability of the records, the first and the last record was deleted in case of an observation having a null value.

4.2 Experimental results

Two-stage neural networks were used in this experiment. To compare the classification accuracy of different inputs, three neural networks were used in every stage. The inputs are the RSSI of only the cellular network, only the Wi-Fi AP, and the combination of cellular network and Wi-Fi AP. A total of 27 cellular networks and 287 Wi-Fi APs were received. Therefore, the inputs of the neural networks were 27, 287, and 314, respectively.

For the first stage of level classification, the number of neurons in the hidden layer was set to 20. The number of output neurons was 5, which included those on the 1, 2, 3, 2.5 (location between 2nd floor and 3rd floor), and 1.5 floors (location between 1st floor and 2nd floor). The training and testing data were separated by half and half. A



two-fold cross validation was applied on both the level classification and location determination. The accuracy was used to determine the reliability of the model. The accuracies of the level classification by two-fold cross validation are presented in Tables 1 and 2.

From Tables 1 and 2, it can be seen that the accuracies of both the cross validations are the same. The highest accuracy is 100%, which is obtained from the union of cellular networks and Wi-Fi AP as an input and only Wi-Fi AP as an input. The accuracy for only a cellular signal as an input is 92%.

In the second stage, three neural networks with different input (only cellular network, only Wi-Fi AP, and the combination of a cellular network and Wi-Fi AP) were used. The hidden layer neurons were set to 20. The training data and testing data were separated by half and half. A two-fold cross validation was applied on the location determination.

The distance error was used to measure the capability of the model. When the location coordinates were estimated by the neural network, they were denormalized first. Subsequently, the distance was transformed using Eqs. (18) and (19). The two points A and B were hypothesized. The location coordinates (longitude and latitude) of A and B were recorded as (latA, lonA) and (latB, lonB). The distance between A and B was denoted by *distance*.

$$C = \sin(\text{latA}) * \sin(\text{latB}) + \cos(\text{latA}) * \cos(\text{latB}) * \cos(\text{lonA} - \text{lonB}) \quad (18)$$

$$\text{distance} = R * \arccos(C) * \text{Pi} / 180, \quad (19)$$

where R is the radius of the earth.

The mean error of the distance is obtained by two-fold cross validation and is presented in Tables 3 and 4.

The results presented in Tables 3 and 4 are consistent. The mean of the distance error obtained in the model with only the cellular signal as input is approximately 3.7–4.3 m. The distance error is so large that it would guide a user to the wrong room in this kind of building. Among the models with only Wi-Fi AP as input and with combination of cellular network and Wi-Fi AP as input, the mean distance error of the model is approximately 0.1–0.4 m. This error is acceptable in an actual scenario. Both trained models (only Wi-Fi as input and combination of cellular and Wi-Fi as input) are efficient in meeting their positioning requirements.

4.3 Discussions

In practical experimental environments, there are 36,985 records, 42 sampling location points, 28 different cellular networks, and 289 different Wi-Fi access points. All these are collected in the multilevel hospital building of the Xi'an Polytechnic University.

Table 1 The accuracy of level classification in the first-fold cross validation

	Cell + Wi-Fi	Cell	Wi-Fi
Accuracy	100%	92%	100%

Table 2 The accuracy of level classification in the second-fold cross validation

	Cell + Wi-Fi	Cell	Wi-Fi
Accuracy	100%	92%	100%

4.3.1 The two-stage neural network analysis

In the experiment, one deep learning neural network was trained for the first stage of level classification. The location coordinates (longitude and latitude) for three different levels were individually obtained by three deep learning neural networks in the second stage. The optimal accuracy of level classification was found to be 100%, as listed in Tables 1 and 2. This lays a good foundation for the follow-up work in the two-stage method for multilevel buildings. The deep learning neural network plays a pivotal role in this step.

In the second stage, the mean distance error corresponding to only Wi-Fi APs and the conjunction of Wi-Fi APs and cellular networks was 0.28 m for different floors. In the experiment, the door of every room in the building was located. In China, a typical single leaf door has a width of 0.8 m. Therefore, an error of 0.28 m will not cause the system to guide the user inaccurately in the navigation application. The second stage uses multiple deep learning neural networks, which are reliable options. The two-stage neural network is effective for multilevel buildings in the location positioning systems.

4.3.2 The comparison of heterogeneous networks

To compare the effectiveness of heterogeneous networks, three experiments with three different input networks (only cellular, only Wi-Fi APs, and a conjunction of cellular and Wi-Fi APs) were conducted. The accuracy of the level classification was 100% when only Wi-Fi APs and a combination of cellular network and Wi-Fi APs were used as the inputs. The distance error was used to determine the location. The average distance error in different floors was 0.28 m for only Wi-Fi APs and a combination of Wi-Fi APs and cellular networks. However, for only cellular networks, the results at both stages were not satisfied. All the distance errors were greater than 1 m. This could lead to the user being guided to the wrong room.

Several cellular network-based wide area location systems have been proposed in recent years. The technological methods of location determination involve measuring the signal strength, the angle of signal arrival, and/or the time difference of signal arrival. However, the accuracy of wide area location systems is highly limited by the cell size. Moreover, the effectiveness of systems in an indoor environment is also limited by the multiple reflections experienced by the radio frequency signal. Using cellular networks as the minor feature in deep learning neural networks will not change these factors.

Table 3 The mean distance error of location determination in the first-fold cross validation

Building floor	Cell	Cell + Wi-Fi	Wi-Fi
1	3.81318	0.24169	0.19428
2	4.23094	0.41723	0.44292
3	3.87375	0.17962	0.19766
Total	3.98238	0.28354	0.28406

Table 4 The mean distance error of location determination in the second-fold cross validation

Building floor	Cell	Cell + Wi-Fi	Wi-Fi
1	3.75281	0.30836	0.30808
2	4.04372	0.188962	0.22982
3	3.93105	0.35378	0.400371
Total	3.91596	0.28118	0.311244

4.3.3 The comparison of different building levels

In the second stage, three location determination neural networks were trained. However, the distance error on the second floor is larger than that on the other two building levels. Identical results are produced by three neural networks using different inputs. This result could be caused due to measuring errors.

5 Conclusion and future work

The proposed 3D mobile system is based on the RSSIs from cellular networks and Wi-Fi APs. Deep learning is used to train the model. For multiple-story buildings, the two-stage location model is theoretically reasonable and practical, which is verified experimentally. It demonstrates the validity of this model for dealing with practical problems.

This 3D positioning system is designed particularly for multiple-story buildings. It aims to obtain the building level, longitude, and latitude for a specific location. This system can recognize the horizontal information of the plane space, as well as the vertical information of different floors.

There are still some defects in the systems. Although this experiment is conducted on a simple building, the implementation of the two-stage 3D indoor positioning method in multiple level buildings is based on the same logic. For irregular buildings, such as cylindrical buildings, the calculation method of the latitude and longitude for the reference point can be designed more sensitively to ensure the accuracy of the latitude and longitude. Furthermore, an additional condition should be considered. The collected RSSIs of Wi-Fi APs for training models are affected by many factors, such as temperature and air humidity. Therefore, the positioning error maybe slightly different at different times. Further collection of data may optimize the system in the future.

Abbreviations

GPS: Global Positioning System; RSSI: Received signal strength indicator; AP: Access point; NN: Neural network

Acknowledgements

The authors thank the anonymous reviewers and editors for their efforts in valuable comments and suggestions.

Authors' contributions

All the authors participated in writing the article and revising the manuscript. The authors read and approved the final manuscript.

Funding

This research program was supported by the Shaanxi's Scientific and Technological Commission (Project No. 2019KRM141), the Natural Science Foundation of the Education Department of Shaanxi Province, China (Project No. 19JK0373), and Scientific and Technological Department of Xi'an City (Project No. 201805072RK35F6).

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests.

Received: 2 January 2020 Accepted: 14 August 2020

Published online: 04 September 2020

References

1. G.Y. Chen, M. Gan, C.L.P. Chen, et al., A two-stage estimation algorithm based on variable projection method for GPS positioning. *IEEE Trans Instrument Measure* **67**(11), 1–8 (2018). <https://doi.org/10.1109/TIM.2018.2826798>
2. L Jiang, L N Hoe, L L Loon. Integrated UWB and GPS location sensing system in hospital environment. *Industrial Electronics and Applications (ICIEA)*, 2010 the 5th IEEE Conference on. IEEE, 2010. DOI: <https://doi.org/10.1109/ICIEA.2010.5516828>
3. J T Gong, C W Tan, and L J Liu. A mental patient positioning management system in hospital based on ZigBee. 2017 International Conference on Robots & Intelligent System (ICRIS) IEEE Computer Society, 2017. DOI: <https://doi.org/10.1109/ICRIS.2017.9>
4. J Chai. Patient positioning system in hospital based on Zigbee. *International Conference on Intelligent Computation & Bio-medical Instrumentation*. IEEE, 2012. DOI: <https://doi.org/10.1109/ICBIM.2011.72>
5. J Chen, C Dong, Chen DO, G HE, X ZHANG. A method for indoor Wi-Fi location based on improved back propagation neural. *Turkish Journal of Electrical Engineering & Computer Sciences*. 2019, 27(4). DOI: <https://doi.org/10.3906/elk-1812-13>
6. C.H. Chen, B.Y. Lin, C.H. Lin, Y.S. Liu, C.C. Lo, A green positioning algorithm for Campus Guidance System. *Int J Mobile Commun* **10**(2), 119–131 (2012). <https://doi.org/10.1504/ijmc.2012.045669>
7. Y C Pu, P C You. Indoor positioning system based on BLE location fingerprinting with classification approach. *Applied Mathematical Modelling*, 2018: S0307904X18302841. DOI: <https://doi.org/10.1016/j.apm.2018.06.031>
8. Q Pu, M Zhou, F Zhang, et al. Group power constraint based Wi-Fi access point optimization for indoor positioning. *KSII Transactions on Internet and Information Systems*, 2018, 12(5):1951-1972. DOI: <https://doi.org/10.3837/tiis.2018.05.003>
9. Z Zhao, J L Carrera V., T Braun, Z Pan, Conditional probability-based ensemble learning for indoor landmark localization. *Comput Commun*, Volume 145, 2019, Pages 319-325, ISSN 0140-3664, <https://doi.org/10.1016/j.comcom.2019.07.007>
10. N Hernández, J M Alonso, M Ocaña. Fuzzy classifier ensembles for hierarchical WiFi-based semantic indoor localization. *Expert Systems with Applications*, 2017, 90. DOI: <https://doi.org/10.1016/j.eswa.2017.08.007>
11. X Ke, J Zou, Y Niu. End-to-end automatic image annotation based on deep CNN and multi-label data augmentation. *IEEE Trans Multimedia*, 2019:1-1. DOI: <https://doi.org/10.1109/TMM.2019.2895511>
12. C H Chen, A cell probe-based method for vehicle speed estimation, *IEICE transactions on fundamentals of electronics. Communications and Computer Sciences*, vol. E103-A, no. 1, pp. 265-267, January 2020. DOI: <https://doi.org/10.1587/transfun.2019TSL0001>
13. C H Chen, F Song, F J Hwang, L Wu, A probability density function generator based on neural networks, *Physica A: Statistical Mechanics and its Applications*, vol. 541, Article ID 123344, March 2020. (SCI/El, 0378-4371) DOI: <https://doi.org/10.1016/j.physa.2019.123344>
14. H Heng, Z Xie, Y Shi, N Xiong. Multi-step data prediction in wireless sensor networks based on one-dimensional CNN and bidirectional LSTM. *IEEE Access*. 2019, 1-1. DOI: <https://doi.org/10.1109/ACCESS.2019.2937098>
15. H Luo, C Chen, L Fang, et al. High-resolution aerial images semantic segmentation using deep fully convolutional network with channel attention mechanism. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **12**(9) (2019): 3492-3507. DOI: <https://doi.org/10.1109/JSTARS.2019.2930724>
16. H Cheng, Z Xie, L Wu, et al. Data prediction model in wireless sensor networks based on bidirectional LSTM. *EURASIP Journal on Wireless Communications and Networking*, 2019, 2019(1):203. DOI: <https://doi.org/10.1186/s13638-019-1511-4>
17. X W Shi, H Q Zhang. Research on indoor location technology based on back propagation neural network and Taylor series. *Control & Decision Conference*. IEEE, 2012. DOI: <https://doi.org/10.1109/CCDC.2012.6244303>
18. L. Wu, C.H. Chen, Q. Zhang, A mobile positioning method based on deep learning techniques. *Electronics* **8**(1), 59 (2019). <https://doi.org/10.3390/electronics8010059>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)