

# A Parallel Artificial Neural Network Learning Scheme based on Radio Wave Fingerprint for Indoor Localization

Chan Uk Park, Hong-Gi Shin, and Yong-Hoon Choi

School of Robotics  
Kwangwoon University,  
Seoul, Republic of Korea  
sau403@gmail.com, ghdl95@naver.com, yhchoi@kw.ac.kr

**Abstract**— Radio wave fingerprinting is known to be the best method for indoor positioning, and its performance depends greatly on the data comparison algorithm that is used. This paper implements a radio wave fingerprint positioning method with artificial neural network learning to improve the performance of a conventional radio fingerprint positioning algorithm based on the Euclidean distance. We propose a parallel learning method to reduce the error in the indoor height and an indoor positioning data augmentation method for data generalization. This method exhibits a higher performance than an existing Euclidean distance based positioning method. In particular, the data augmentation technique can be applied without depending on the specific positioning algorithm.

**Keywords**— *Indoor Positioning; Artificial Neural Network (ANN); Deep Learning; Data Augmentation*

## I. INTRODUCTION

Recently, the demand for more accurate positioning is increasing due to the increase in location-based services. Radio wave fingerprint positioning, which is known to be the best for indoor positioning, converts indoor wireless signal data into a database and compares it to the actual position. To accurately estimate the position, it is important to predict the actual position through the data in the database. Conventional methods estimate the position by comparing the data in the DB to that of the actual position using a mathematical method based on the Euclidean distance. In this paper, we propose an artificial neural network (ANN) learning method that can provide more accurate position estimation. When data is not sufficient and learning is not performed properly, data augmentation approach can improve positioning performance. In addition, we propose a learning structure suitable for indoor positioning. It is possible to confirm the effect of reducing the position error due to the regression problem of the artificial neural network through the learning structure suitable for indoor positioning. The feature of the proposed method can be easily applied in an existing positioning technique, and the position estimation performance can be confirmed to have improved for most positioning algorithms.

The organization of this paper is as follows. Section 2 describes the underlying technologies to help understand the paper. Section 3 describes the techniques proposed in this paper. Section 4 describes the performance evaluation experiments of

the proposed algorithm and the existing positioning algorithms. Finally, we conclude this paper in Section 5.

## II. RELATED WORKS

Section 2 describes the relevant techniques in this paper. First, the type of positioning algorithm and position calculation method are described, and the positioning algorithm that uses a comparison of the Euclidean distance, which shows the best performance in radio fingerprinting using existing machine learning, will be described. In addition, we describe the positioning method of radio wave fingerprint using an artificial neural network and explain the points to supplement previous studies.

### A. Triangulation method and radio wave fingerprint positioning technique

Existing positioning algorithms can be roughly divided into two methods: prediction through mathematical calculation and radio wave fingerprinting that compares existing positioning data. An example of a prediction method through a mathematical computation is a polygonal metric algorithm that is popularly known as triangulation [1]. However, due to factors such as attenuation and interference of the radio signal in indoor spaces, an accurate distance cannot be known through the strength of the radio signal. The radio wave fingerprinting method is a positioning algorithm in which signals of an access point (AP) are stored in a database and are compared with signals of a corresponding space to be positioned. This is suitable for an indoor space where attenuation and interference of radio signals occur [2].

### B. Fingerprint positioning based on Euclidean distance comparison

The algorithm that compares the positioning data with a database is one that has the greatest influence on the accuracy in estimating the position through data in radio wave fingerprint positioning. One of the methods for existing radio fingerprint positioning methods is to compare data with a database by taking the difference of the Euclidean distance [3]. However, as the number of APs in the environment increases, the dimension of the data to be compared increases, and thus the difference in

the Euclidean distance becomes ambiguous [3]. To solve this problem, existing studies have used the logarithmic Gaussian distance to achieve a high performance in high-order AP data [4]. In addition, research has been conducted on a method to classify data of a building to be located using Affinity Propagation Clustering and comparing the meaning distance with the related data [5]. When clustering is performed, the related data can be compared with the data in the cluster, and the size of the dimension comparison can be reduced to decrease the ambiguity in the meaning distance. The above study shows the highest performance in conventional positioning based on Euclidean distance comparison [6]. Therefore, it is used as a benchmark to measure the performance of this paper.

### C. Fingerprint positioning technique using an ANN

ANN has been applied to positioning in [7] and [8]. Existing studies have calibrated the channel state information for the deep learning by modifying the driver of the wireless network interface card. Positioning using channel state information (CSI) data provides high location accuracy, but data collection is difficult because the AP hardware needs to be modified. In this paper, positioning is performed based on Received Signal Strength (RSS) data of Wi-Fi AP which is most used in radio wave fingerprint recognition research. Therefore, it has an advantage that it can be directly applied to existing research. Another study that applied ANNs to positioning is a study using a learning network that is a mixture of deep neural network (DNN) and hidden Markov model (HMM) [8]. The application of the learning network can improve positional accuracy within a given data, but the disadvantage is that location accuracy depends on the quality of the learning data. In this paper, we solve the weak point of the previous study by improving the quality of learning data.

## III. PROPOSED TECHNIQUE

Section 3 describes the positioning technique proposed in this paper. In this paper, we propose an ANN learning method that can provide more accurate position estimation.

### A. Data Augmentation Technique for Location Data

One weakness of ANN is that as learning progresses, the performance of the test decreases due over-fitting of the learning data. To prevent this phenomenon data generalization is performed [9]. In this paper, data generalization is performed using a data augmentation method suitable for wireless signal data.

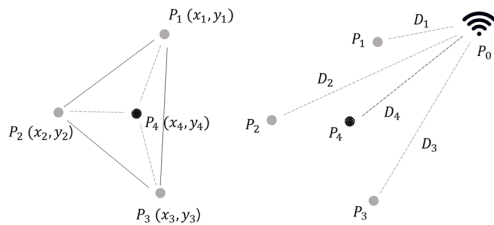


Figure 1. Data Augmentation Scheme for Indoor Positioning Location Data

Figure 1 shows an illustration of the data augmentation technique implemented in this paper. The new point  $P_4$  is the

center of gravity of the triangle  $\triangle(P_1, P_2, P_3)$  and thus, the equation (1) is satisfied.

$$P_4(x_4, y_4) = P_4\left(\frac{x_1+x_2+x_3}{3}, \frac{y_1+y_2+y_3}{3}\right) \quad (1)$$

When the three data (e.g., RSS values) for given points  $P_1, P_2$ , and  $P_3$ , are  $D_1, D_2$ , and  $D_3$ , the newly added data,  $D_4$  for new point,  $P_4$  is calculated by equation (2).

$$D_4 = \frac{D_1+D_2+D_3}{3} \quad (2)$$

Since  $P_4$  is located at the same distance from  $P_1, P_2, P_3$ , the radio signal strength information of  $P_4$  is calculated as the average value of radio signal strength of  $P_1, P_2, P_3$ . This data augmentation technique can produce the data the most similar to the actual data in the following situations.

- When there is no AP signal in  $\triangle(P_1, P_2, P_3)$
- When there is no environment that causes attenuation and reflection of radio signal inside  $\triangle(P_1, P_2, P_3)$
- When  $P_1, P_2, P_3$  are on the same floor of a building

However, it is not easy to recognize the location of each AP and the structure of the building with only the data set. Therefore, in this paper, the data augmentation method is performed for the data that exists in the same floor. Algorithm 1 and 2 describes the algorithm for data augmentation. Data Augmentation is performed in order from the first data. Using the data selected in order as a reference point, two data measured in the same floor in the closest order are further selected. After that, triangles are created from the selected 3 points and data expansion is performed. In this case, the information of the triangle used to prevent duplication of data is stored and duplicate inspection is performed in the future. The number of data generated by the above procedure is equal to the number of original data. Therefore, it is possible to obtain twice the learning data including the original data.

In order to test whether the extended data affects the positioning accuracy, several methods were used to perform positioning by using extended data and the original data.

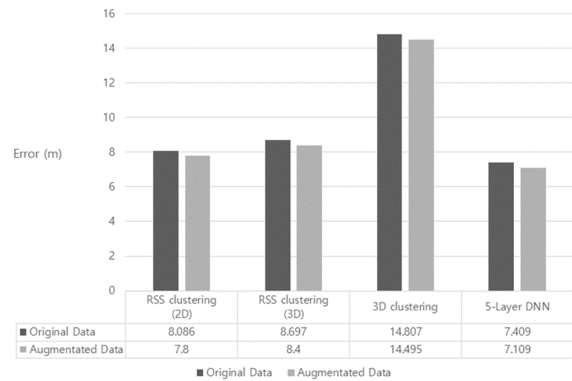


Figure 2. Position Accuracy Based on Data Augmentation Technique

Experimental results show that the data set that performed the data augmentation technique showed high accuracy overall.

This proves that effective generalization of the data set has been performed due to the data augmentation technique of this paper.

Algorithm 1. Positioning Data Augmentation

```

1: function DATAUGMENTATION(PositioningDataArray)
2:   len = PositioningDataArray.length
3:   Let TriangleList be new linked list
4:   ▷ Prevent selection of the same triangle
5:   Let ExceptDataList be new linked list
6:   Let AugmentedDataList be new linked list      ▷ Newly added data
7:
8:   for i = 0 to len do
9:     D1 = PositioningDataArray[i]
10:    ExceptDataList.put(D1)
11:    D2 = SearchNearestPoint(D1,
12:      PositioningDataArray, ExceptDataList)
13:    ExceptDataList.put(D2)
14:
15:    for j = 0 to TriangleList.length do
16:      if D1 in TriangleList & D2 in TriangleList then
17:        ExceptDataList.putAll(TriangleList[j])
18:      end if
19:    end for
20:
21:    Let CenterDot be new PositioningData
22:    CenterDot.x = (D1.x + D2.x)
23:    CenterDot.y = (D1.y + D2.y) / 2
24:    D3 = SearchNearestPoint(CenterDot, PositioningDataArray, Except-
25:      DataList)
26:    ExceptDataList.put(D3)
27:
28:    Let T = new Triangle(D1, D2, D3)
29:    TriangleList.put(T)
30:
31:    Let D4 be new PositioningData
32:    D4.x = (D1.x + D2.x + D3.x) / 3
33:    D4.y = (D1.y + D2.y + D3.y) / 3
34:    AugmentedDataList.put(D4)
35:  end for
36:  return AugmentedDataList
37: end function

```

Algorithm 2. Search Nearest Positioning Data

```

1: function SEARCHNEARESTPOINT(RefPoint, PositioningDataArray, Ex-
2:   ceptDataList)
3:   ▷ params : Base points, datasets, and selective prevention datasets
4:   NearstDistance = Integer.MAX
5:   NearstPoint be new PositioningData
6:
7:   for locData in PositioningDataArray do
8:     if locData.z != RefPoint.z then
9:       continue
10:    end if
11:
12:    Distance =  $\sqrt{(RefPoint.x - locData.x)^2 + (RefPoint.y - locData.y)^2}$ 
13:
14:    if NearstDistance > Distance & locData not in ExceptDataList
15:      & Distance ≠ 0 then
16:        NearstDistance = Distance
17:        NearstPoint = locData
18:      end if
19:    end for
20:  return NearstPoint
21: end function

```

### B. Parallel learning for interlayer identification

The machine learning algorithm can be divided into a classification or a regression problem according to the result value that is to be searched. Figure 3 shows an example of the classi-

fication and regression problems in radio wave fingerprint positioning.

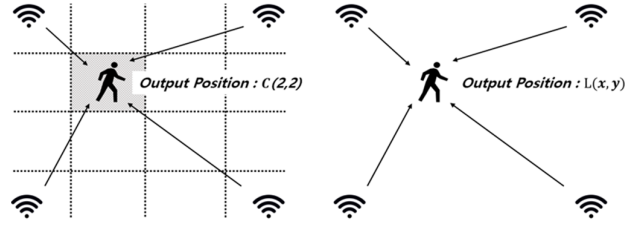


Figure 3. Classification and regression problem of radio wave fingerprint positioning

The left-hand side of Figure 3 shows an example of radio wave fingerprint positioning of the classification problem. This algorithm is a method for estimating the position of the test data by dividing the place where the positioning is performed into an arbitrary lattice based on the learning data. The right side of Figure 3 shows an example of a regression problem. This algorithm is an algorithm that estimates the location of test data by converting the position of each measured data into x, y coordinates. In this way, positioning the radio wave fingerprint can be approached in two ways of classification and regression. In the case of the classification problems, the performance evaluation is complicated due to the difference in the accuracy depending on the interval of the lattice.

In this paper, we use a classification technique to distinguish the floors of a building, and use regression techniques to estimate locations within the same floor. We perform parallel learning on the x, y and z axes with the following structure as shown in Figure 4.

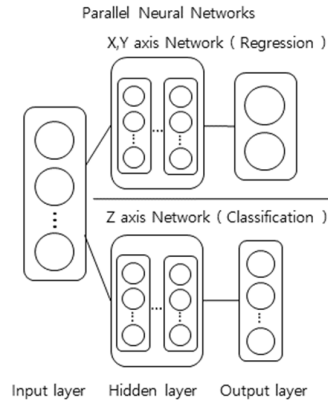


Figure 4. Multi-floor parallel learning structure diagram

TABLE I. ACCURACY BASED ON PARALLEL LEARNING

	5-layer DNN (2D)	5-layer DNN (3D)	5-layer P-DNN
Mean error (m)	7.40	7.84	7.28
Remark	-	Floor detection error: 0.4m	Floor detection probability: 94.4%

Table 1 shows the learning accuracy of parallel learning. Experiments were conducted to investigate the difference between 2D positioning performance and 3D positioning performance and the parallel learning positioning performance of this paper when learning the same data set [10] in the same neural network model with the same network structure. Experimental results show that the 3D positioning performance is less accurate than the 2D positioning due to the z-axis error, and the z-axis error can be eliminated when the learning is improved through parallel learning. The parallel learning architecture proposed in this paper can improve the positioning performance by eliminating the errors that occur in the inter-floor division

#### IV. EXPERIMENT

Section 4 evaluates the performance of the proposed method. The proposed method is compared with the existing indoor positioning methods.

##### A. Experiment environment

The data set used in the experiment is an open data set for indoor positioning studies and consists of 4,648 data collected from 21 devices in the Tampere University building in Finland. It also provides benchmark results for existing positioning algorithms to facilitate future research [10]. Experiments were carried out to measure the error in location estimation accuracy, floor classification performance, and learning time for location estimation for each positioning algorithm.

##### B. Experiment result

TABLE II. POSITION ACCURACY ACCORDING TO VARIOUS POSITIONING ALGORITHMS

<i>Algorithm</i>	<i>2D Error (m)</i>	<i>3D Error (m)</i>	<i>Floor Detection Probabilities (%)</i>
5-Layer DNN	7.409	7.84	N/A
<b>Data Augmented 5-Layer P-DNN (proposed)</b>	<b>6.943</b>	<b>7.109</b>	<b>94.40</b>
RSS clustering (affinity propagation) [6]	8.086	8.697	90.81
Data Augmented RSS clustering (affinity propagation) [6]	7.8	8.4	91.16
3D clustering (k-means) [6]	14.807	19.519	81.54
Data Augmented 3D clustering (k-means) [6]	14.495	19.484	82.27

The algorithms used in the experiment are the DNN, Parallel DNN (P-DNN), 3D clustering, and RSS clustering. Each algorithm has been tested before and after applying the data

augmentation method proposed in this paper. Experimental results show that the P-DNN that uses the data augmentation algorithm and the parallel learning method proposed in this paper shows the highest performance.

#### V. CONCLUSION

In this paper, an efficient data augmentation method for indoor positioning and a parallel learning method for floor-to-floor distinction have been proposed. The advantage of this paper is that it does not apply any filter to the RSS data, but it shows better accuracy than the benchmarked conventional positioning study when it is learned only by DNN. The weak point of this study is that the learning accuracy is not always constant and that it takes a lot of learning time to get high accuracy due to the nature of the deep artificial neural network.

#### ACKNOWLEDGEMENT

This research was supported by Basic Science Research Program through the National Research Foundation (NRF) of Korea funded by the Ministry of Education (NRF-2016R1D1A1B03934507).

#### REFERENCES

- [1] M. Armstrong, J. Wright, "Mobile call termination in the UK", UCL, Sep. 2007.
- [2] S. He, S.-H.G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons", IEEE Communications Surveys & Tutorials, vol. 18, pp. 466-490, 2016.
- [3] C. Feng, W. S. A. Au, S. Valaee, Z. Tan, "Received-signal-strength-based indoor positioning using compressive sensing", IEEE Trans. Mobile Comput., vol. 11, no. 12, pp. 1983-1993, Dec. 2012.
- [4] S. Shrestha, J. Talvitie, and E.S. Lohan "Deconvolution-based indoor localization with WLAN signals and unknown access point locations", in Proc. of IEEE ICL-GNSS, Jun 2013, Italy
- [5] Zengshan Tian, Xiaomou Tang, "Fingerprint indoor positioning algorithm based on affinity propagation clustering", EURASIP Journal on Wireless Communications and Networking 2013
- [6] A. Cramariuc, H. Huttunen, and E.-S. Lohan, "Clustering benefits in mobile-centric WiFi positioning in multi-floor buildings," in International Conference on Localization and GNSS (ICLGNSS), Jun 2016
- [7] Xuyu Wang, Lingjun Gao, Shiwen Mao, Santosh Pandey, "CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach", Vehicular Technology IEEE Transactions on, vol. 66, pp. 763-776, 2017, ISSN 0018-9545.
- [8] W. Zhang, K. Liu, W. Zhang, Y. Zhang, J. Gu, "Wi-Fi positioning based on deep learning, in: International Conference on Information and Automation", 2014, pp. 1176-1179.
- [9] C. Zhang, S. Bengio, M. Hardt, B. Recht, O. Vinyals, "Understanding deep learning requires rethinking generalization", Proceedings of the International Conference on Learning Representations (ICLR) (2017)
- [10] Elena Simona Lohan, Helena Leppäkoski, Philipp Richter, Zhe Peng, "Wi-Fi Crowdsourced Fingerprinting Dataset for Indoor Positioning", MDPI Data, 2017