



Wi-Fi Attention Network for Indoor Fingerprint Positioning

Ting Zhang¹(✉), Yi Man¹, Xiaoshi Fang², Ligang Ren³, and Yue Ma¹

¹ Beijing University of Posts and Telecommunications, Beijing 100876, China
zhangtingzt@bupt.edu.cn

² Instrumentation Technology and Economy Institute, Beijing 100055, China

³ China United Network Communications Corporation,
Beijing Branch, Beijing 100038, China

Abstract. With the rapid development of human society, the scope and complexity of urban living space are increasing, and the demand for space position information is constantly improved by human activities. Especially in recent years, with the advent of the mobile Internet and the rapid popularization of smart phone terminals, the lifestyle and behavior habits of people are undergoing a huge change. Therefore, the positioning technology is also increasingly received by people. Among them, Wi-Fi fingerprint positioning is the most popular way. In this paper, we proposed Wi-Fi Attention Networks (WAN) for Wi-Fi fingerprint indoor positioning. In the WAN, a bidirectional LSTM is used to get a representation vector of Wi-Fi by summarizing the contextual information. Then, an attention mechanism is used to extract the Wi-Fi words which are important to the representation of Wi-Fi sequences and get a high-level representation vector. Finally, a fully connected network is used for classification. Experimental results demonstrate that WAN performs better than traditional machine learning methods on the publicly available dataset.

Keywords: Wi-Fi fingerprint positioning · Bidirectional LSTM · Attention networks

1 Introduction

In recent years, with the advent of the mobile Internet and the rapid popularization of smart phone terminals, the lifestyle and behavior habits of people are undergoing a huge change. Nowadays, people in daily life habitually through LBS (Location Based Service) to find restaurants, banks, takeout and so on. Therefore, positioning technology has become an increasingly significant problem.

At present, GPS (Global Positioning System) is the most widely used positioning system. However, in complex indoor environments, due to factors such as weak signal, environmental noise and multipath interference, GPS's positioning accuracy deteriorates rapidly and even cannot be located [1]. In order to solve the positioning problem in complex indoor environments, most of the solutions are based on local wireless sensor network [2]. In recent years, Wi-Fi wireless cards are installed on most

smartphones, tablets and laptops, providing a ready-made hardware platform for indoor positioning. Therefore, Wi-Fi fingerprint positioning has become the most widely used indoor fingerprint positioning [3].

1.1 Problem Definition

The Wi-Fi fingerprint positioning, as shown in Fig. 1, is roughly divided into two phases: the offline phase and the online phase. In the offline phase, by recording the RSS of Access Points (APs) at the Reference Points (RPs), a fingerprint database is built in the positioning area and the positioning model is trained. In the online phase, feature extraction is performed on the real-time Received Signal Strength (RSS) of Aps through the positioning model to estimate the device position of current moment. The positioning model generally adopt machine learning methods, which can divided into classification method which discrete result can be obtained and regression method which the continuous position coordinates can be obtained. In order to simplify the problem, in this paper, we adopt classification positioning model, and we propose a positioning model based on deep learning in offline phase, which has good performance on the publicly available dataset.

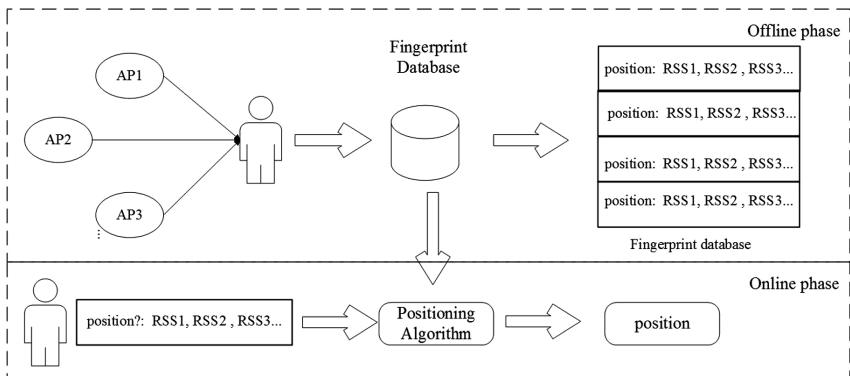


Fig. 1. Technology of Wi-Fi fingerprint positioning

2 Related Work

In recent year, some scholars have introduced machine learning into Wi-Fi finger-print positioning, such as KNN [4], WKNN [5] and SVM [6]. In this method, it mainly training a classification model based on machine learning to extract the mapping relationship between Wi-Fi RSS and device location. Wang et al. [7] used a four-layer FCN (fully connected network) for fingerprint positioning. The network structure is pre-trained by stacking and self-encoders, and the global fine-tuning is performed by backpropagation. The feature is automatically extracted from the wavy wireless signal and linearly transformed to calculate the position of target. FCN model can achieve

higher positioning accuracy and enhances system robustness, but the multi-layer FCN has too many parameters, which need a large number of data to train the parameters. More importantly, FCN does not take advantage of positional information between RSS.

If we sort each AP and RSS pairs according to the RSS value, the connection between each AP and RSS pairs and its surrounding pairs is relatively close, just like the adjacent word in a sentence. Although in other fields, deep learning model such as CNN, LSTM and Attention model perform well in sequence modeling, it is still not well applied in Wi-Fi fingerprint positioning. Therefore, in the paper, we proposed a fingerprint positioning method based on Bi-LSTM and Attention model.

3 Wi-Fi Attention Network

The overall architecture of the Wi-Fi Attention Network is show in Fig. 2 which refers to Yang's Hierarchical Attention architecture [8]. It consists of two part: a Wi-Fi sequence encoder and a Wi-Fi attention layer. We introduce the details in the following sections.

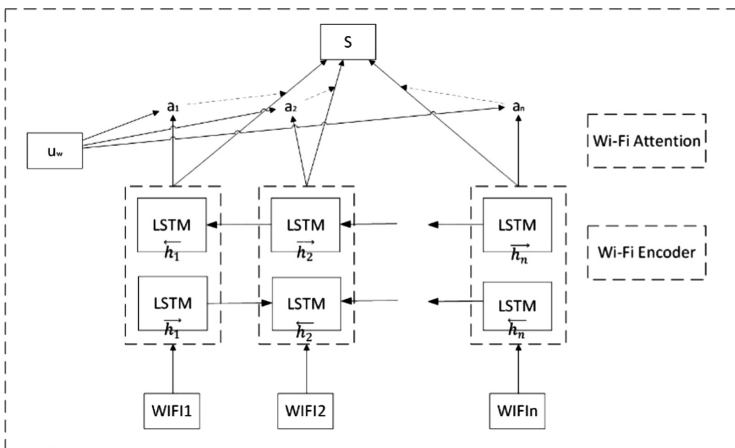


Fig. 2. The architecture of Wi-Fi Attention Network

3.1 Wi-Fi Encoder

Each Wi-Fi sequence contains several APs and its RSS, drawing on the idea of text classification, we treat an AP and its RSS pairs as a Wi-Fi word, and treat each Wi-Fi sequence as a Wi-Fi sentence. Assume that a Wi-Fi sentence contains n Wi-Fi words w_i , with $i \in [0, T]$ represents the i_{th} Wi-Fi word in the Wi-Fi sentence. We first embed the Wi-Fi words to vectors through an embedding matrix W_e , $x_i = W_e * w_i$. Then, we use a bidirectional LSTM [9] to get representation vector of Wi-Fi words by summarizing information from both directions, in this way, we can incorporate the

contextual information of Wi-Fi sentence. The bidirectional LSTM contains the forward LSTM which get the Wi-Fi sentence s from x_1 to x_n and a backward LSTM which reads from x_n to x_1 .

$$x_i = W_e * w_i \quad (1)$$

$$\vec{h}_i = \overrightarrow{LSTM}(x_i) \quad (2)$$

$$\overleftarrow{h}_i = \overleftarrow{LSTM}(x_i) \quad (3)$$

We obtain a complete representation vector for as given word w_i by concatenating the forward hidden state \vec{h}_i and backward hidden state \overleftarrow{h}_i . In this way, h_i can summarize the information of the whole Wi-Fi sentence centered around w_i .

$$h_i = \left[\vec{h}_i, \overleftarrow{h}_i \right] \quad (4)$$

3.2 Wi-Fi Attention

Not all Wi-Fi words have equivalent effect on the representation of the Wi-Fi sentence. Therefore, we introduce attention mechanisms to extract Wi-Fi words which are important to the representation of the Wi-Fi sentence, and summarize the representation of those Wi-Fi words according to its significance to form a Wi-Fi sentence vector.

$$u_i = \tanh(W_w h_i + b_w) \quad (5)$$

$$\alpha_i = \frac{\exp(u_i^T u_w)}{\sum_i \exp(u_i^T u_w)} \quad (6)$$

$$s = \sum_i \alpha_i h_i \quad (7)$$

At first, we use a single-layer MLP to get the u_i as a hidden representation of h_i . Then, we calculate the importance of this Wi-Fi words as the similarity between u_i and u_w which is a context vector and get a normalized weight vector α_i through a softmax function. Finally, we calculate the Wi-Fi sentence representation vector s as a weighted sum of Wi-Fi words. The context vector u_w is the high-level representation of the Wi-Fi sentence, which is randomly initialized and learned together during training process.

3.3 Wi-Fi Classification

The Wi-Fi sentence representation vector s is a high-level representation of this Wi-Fi sequence and can be used as features for the Wi-Fi classification:

$$p = \text{softmax}(W_s s + b_s) \quad (8)$$

Then, we choose then $\max p$ as the classification result of this Wi-Fi sequence.

4 Experiments

4.1 Dataset

We evaluate the effectiveness of our Wi-Fi Attention Network model on the public dataset of BDCI 2017 Task 1, to accurately locate the user’s store in the mall [10]. This dataset has 1.1 million training sample and 48.4 thousand test sample, the target of this competition is to locate the user’s current store accurately. The dataset including three kind of information. The first information is the user information, including user id and its historical consumption record. The second information is the shop information, such as average consumption price of this shop. The last information is the consumption information, including Wi-Fi sequence information, longitude and latitude information. In order to evaluate the effectiveness of our Wi-Fi Attention Network model, we only use the Wi-Fi sequence information. This Wi-Fi sequence information is Wi-Fi list consist of several AP information, for each AP, it has three items: id, signal and flag. Among them, id is the unique identifier of APs, it can distinguish between different APs. Signal is the RSS of APs, which vary from -104 to 0 , 0 represent the max value of AP’s RSS. Flag indicates whether this AP is connected by the device. We use 80% of the dataset for training, 10% for validation and the remaining 10% for test.

4.2 Preprocessing

We split Wi-Fi sequence into Wi-Fi words as shown in Fig. 3. The original Wi-Fi sequence has ten AP information, each AP information has three parts: id, signal and flag. In order to achieve good effect, we sort each AP information in descending order according to the RSS value. Then we can get the preprocessed Wi-Fi sequence.

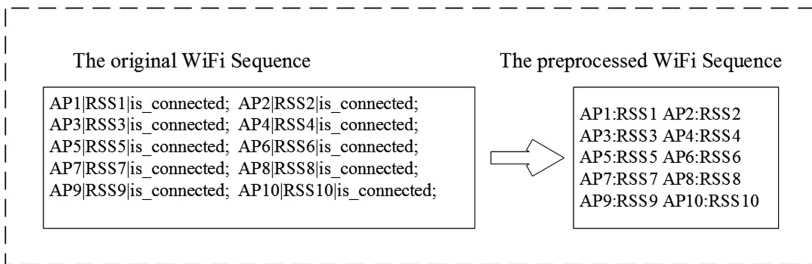


Fig. 3. The preprocessing of Wi-Fi sequence

4.3 Evaluation

We train our Wi-Fi Attention Network on the preprocessed dataset and compare it with other models. The hyper parameters as shown in Table 1, which are tuned on the validation dataset. In our model, we set the LSTM dimension to be 100, in this situation, the bidirectional LSTM produce a 200 dimensions vector. And we set the Wi-Fi word embedding dimension to be 100. For training, we use RMSProp optimizer, batch size of 128, training epochs of 20 and learning rate of 0.001.

Table 1. Table parameters of WAN.

| Parameters | Values | Explanation |
|---------------|---------|-----------------------------------|
| Embedding dim | 100 | Wi-Fi word embedding dimension |
| Size of LSTM | 100 | Bidirectional LSTM dimension |
| Epochs | 20 | Model training epochs |
| Learning rate | 0.001 | Adjustment speed of loss function |
| Batch size | 128 | Sample size in one training |
| Optimizer | RMSProp | Optimizer function |

In order to prove the effect of our model, we compare it with several methods, including traditional approaches such as KNN (Bahdanau et al. 2014) and SVM (Zhu et al. 2013) and other neural network model, LSTM, GRU (Bahdanau et al. 2014) without attention structure. The experimental results are shown in Table 2, we refer our models as WAN (Wi-Fi Attention Network), and the result show that WAN gives the best performance across this dataset on precision, recall and F1 score. From Table 2 we can see that WAN outperforms the traditional approaches KNN and SVM model by 5.0% and 4.2% respectively on precision. More importantly, WAN can significantly improve over LSTM 1.5% and GRU 1.4% on precision.

This result proves our hypothesis that better performance can be obtained by incorporating knowledge of Wi-Fi fingerprint sequence structure in the WAN architecture. The intuition underlying WAN is that not all parts of a Wi-Fi fingerprint sequence are equally relevant for indoor positioning and determining the significant parts involves modeling the interactions of the Wi-Fi parts, not just their presence in isolation. That is to say, if we sort each AP according to its RSS value, the connection between AP and its surrounding AP is relatively close. This relationship can help us improve the performance of Wi-Fi fingerprint positioning.

Table 2. The target of different architecture.

| | Methods | Precision | Recall | F1 |
|----------------------|---------|--------------|--------------|--------------|
| Bahl et al. 2012 | KNN | 86.34 | 83.95 | 85.13 |
| Zhu et al. 2013 | SVM | 87.12 | 84.78 | 85.93 |
| Bahdanau et al. 2014 | LSTM | 89.72 | 85.93 | 87.78 |
| Bahdanau et al. 2014 | GRU | 89.83 | 86.44 | 88.10 |
| This paper | WAN | 91.29 | 89.84 | 90.56 |

5 Conclusion

In this paper, we proposed Wi-Fi Attention Networks (WAN) for Wi-Fi fingerprint indoor positioning. In the WAN, a bidirectional LSTM is used to get a representation vector of Wi-Fi by summarizing the contextual information. Then, an attention mechanism is used to extract the Wi-Fi words which is important to the representation

of Wi-Fi sequences and get a high-level representation vector. Finally, a fully connected network is used for classification. Experimental results demonstrate that WAN performs than other models on the publicly available dataset.

Acknowledgement. This work is supported by the National Key R&D Program of China (No. 2018YFB1201500), the National Natural Science Foundation of China under (Grant No.61471055), and the Beijing Natural Science Foundation under Grant No. L171011, Beijing Major Science and Technology Special Projects under Grant No. Z181100003118012 and China Railway with project No. BX37.

References

1. Sithole, G., Zlatanova, S.: Position, location, place and area: an indoor perspective. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf.* **III-4**, 89–96 (2016)
2. Xia, S., Liu, Y., Yuan, G., et al.: Indoor fingerprint positioning based on Wi-Fi: an overview. *Int. J. Geo Inf.* **6**(5), 135 (2017)
3. Gentile, C., Alsindi, N., Raulefs, R., et al.: *Geolocation Techniques: Principles and Applications*. Springer, New York (2012). <https://doi.org/10.1007/978-1-4614-1836-8>
4. Bahl, P., Padmanabhan, V.N.: RADAR: an in-building RF-based user location and tracking system. *Proc. IEEE INFOCOM* **2000**(2), 775–784 (2000)
5. Ni, L.M., Liu, Y., Lau, Y.C., et al.: LANDMARC: indoor location sensing using active RFID. *Wirel. Netw.* **10**(6), 701–710 (2004)
6. Zhu, Y.J., Deng, Z.L., Liu, W.L., Xu, L.M., Fang, L.: Multi-classification algorithm for indoor positioning based on support vector machine. *Comput. Sci.* **39**(4), 32–35 (2012)
7. Wang, X., Gao, L., Mao, S., et al.: DeepFi: deep learning for indoor fingerprinting using channel state information. In: *Wireless Communications and NETWORKING Conference*, pp. 1666–1671. IEEE (2015)
8. Yang, Z., Yang, D., Dyer, C., et al.: Hierarchical attention networks for document classification. In: *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 1480–1489 (2017)
9. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. *Computer Science* (2014)
10. Information on <http://www.datafountain.cn/?spm=5176.100067.444.1.62dd1226aO75xl#/competitions/279/intro>