Performance of Indoor Localization based on Machine Learning Techniques Using CSI and RSSI Metrics

SONG Yuxuan

Department of Telecommunication service and usage, INSA Lyon

Abstract

As the indoor localization technology gains more and more attention over the world, considerable efforts are made on indoor localization systems based on new hardwares. So in this paper, we firstly show the interest in indoor localization and some classical techniques of it, like fingerprinting-based localization, ranging-based localization and angle-of-arrival(OAO)-based localization. Moreover, we deeply focus on the infrastructure-less Wi-Fi being and point out the existed problems and the involvement of machine learning techniques to solve them, where the Received Signal Strength Indicator (RSSI) and Channel State Information (CSI) are available in mainstream deep learning systems. Then for the next section, we present in detail the indoor localization techniques using different machine learnings such as Decision Tree (DT), K-nearest neighbor (KNN), DeepFi and so on. We then compare and discuss their algorithms in terms of accuracy, robust, time cost and calculation complexity. Meanwhile we give the reasons logically for the better performance. Finally we conclude and criticize briefly the indoor localization techniques.

Keywords: indoor localization, Channel state information, RSSI, machine learning

I. Introduction

Nowadays, the development of indoor localization has contributed to the large requirements of indoor positioning applications in daily life, which can be used for target tracking, foot traffic detection, surveillance, also for improving automation capabilities in warehouses and guidance in airport, hospitals, museum and so on. As a result, the popularity of the indoor localization has leaded to the pursuit of higher accuracy, robust and stability of this technique. However, compared with outdoor positioning, indoor localization is more challenging since it also has several problems such as multipath effect, shadowing, fading and so on [1]. To better solve these problems, many efforts are made on indoor localization systems based on new hardware with low cost and high accuracy. These recent works could be classified in 3 categories: Fingerprintingbased, Ranging-based and AOA-based. AOA refers to the angle of incidence at which radio signals from a transmitter arrive at the receiver, which is used in a geometric relationships to obtain the user location. It achieves high accuracy, but accurate AOA can only be obtained by directional antennas or an antenna array, which makes the receiver complex and expensive [2]. Ranging-based localization uses location metrics such as time-of-arrival (TOA) to estimate the distance between two nodes and leverages geometrical models for location estimation. However, it needs at least four reference positions to get a 3D position so it is not easy to apply in a real occasion. One more disadvantage is that the accuracy depends on the number of reference positions and the propagation model, which depends on the multipath effects. While Fingerprinting-based localization covers these shortages since it employs a constructed input database to be compared to the measurements associated with the sensor to localize [3]. Indeed, indoor localization based on fingerprinting is the most widely used, where the Wi-Fi signals are mostly involved, without the need to extra hardwares.

As for the part of wireless signal measurements, the Received Signal Strength Indicator (RSSI) characterizes the attenuation of the signals and represents the received power level at the receiver. Even though it has been adopted in most of the indoor localization systems since it can achieve meter-level localization accuracy, it has worse performance in complex environments .Indeed, the RSSI accuracy can be affected by several phenomenons such as the visibility conditions (i.e. Line of Sight (LOS) or Non LOS), as well as the multipath effects [4]. In the terrible case, a closer receiver may have an even lower RSSI than a distant one due to the existence of ceiling, floor, walls etc [5]. So currently, most of the RSSI-based positioning systems are utilized in simple environments. Unlike RSSI which is the value from MAC layer superimposition of multipath signals with fast changing phases, CSI adopts the PHY layer power feature which is able to discriminate multipath characteristics. It describes the amplitude and phase on each subcarrier in the frequency domain, but it also requires more calculation time which is its drawback, more details can be found in [6].

II. Indoor localization based on machine learning

II.1 Machine learning

In our localization context, we will be focusing on machine learning techniques based on finger-printing, which is the most dominant due to its low cost and high accuracy at the same time. Figure1 is a general method of Wi-Fi-based posi-tioning[7]. It requires a training phase (also called off-line phase) and a test phase (also called on-line phase). The main idea of training phase is to collect signal features (e.g. RSSI, CSI, FM…) from all possible locations in the planning area to build a database, while the test phase is the rest of process of localization, which is matching the measured fingerprints at a location with the database and return the result corresponding to the best-fitted fingerprint [8], [9]. The accuracy depends on the constructed and manipulated databases which will be treated by RSSI or CSI. However, this increases the compl-exity and the running time of the system, so the existed methods pursue to reduce the online complexity and remain the robust and accuracy at the same time.

II.2 Adaptation of machine learning in indoor localization

In this section, we present different indoor localization techniques using machine learning algorithms, where we will talk about the most used algorithms such as deep CNN, KNN, RF as well as DeeFi.

Deep Convolutional Neural Networks (CN-N) for indoor localization, which changes the problem of positioning to an image classification problem. It shifts the complexity of the online

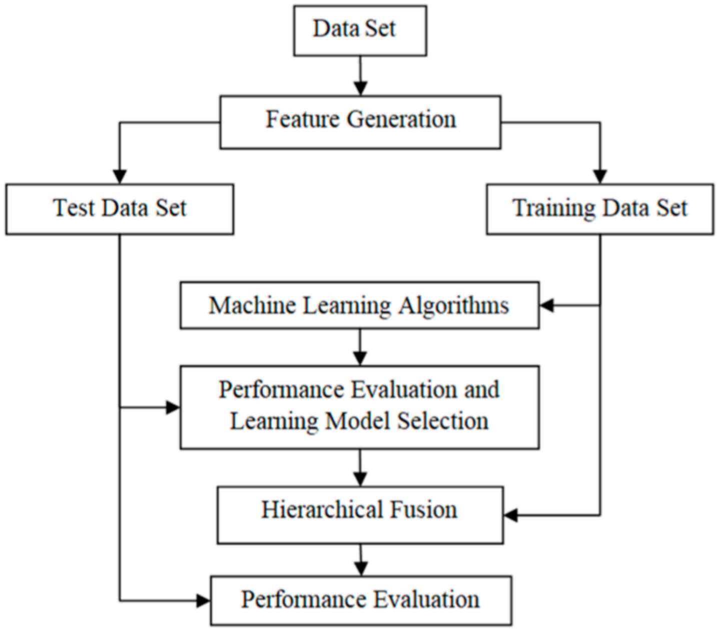

Figure1 : Wi-Fi based positioning system[7]

prediction to an offline preprocessing step. Since CNN has a great performance in image classi-fication, an accurate estimated localization is obtained by recognizing a 3D radio image-based region which is constructed based on RSSI fingerprints. Hybrid Wireless (HW) fingerprint using CNN in [10], which the amelioration is that it covers the unstability of the RSSI by using HW fingerprint which is constructed by a ratio fingerprint obtained by calculating the ratio of RSSI between different Access Points (APs) and the RSSI fingerprint. Then the CNN learns the useful features from HW-fingerprint by building a pixel grayscale image, where we consider the brightness is 0-255 and is distributed by each feature of the HW-fingerprint. The samples from the same reference points have the similar brightness and pixel distribution which leads us to visualize the treatment of indoor localization by image classification. Finally we can get the probability that the input data belongs to a location by the calculation in convolution layer, pooling layer, fully connected layer and softmax layer [10]. Besides, there is also an optimized CNN technique in [11].

K nearest neighbors (KNN) is first applied in [12] which has only one hyper-parameter K value in its algorithm. We need to calculate the distance of the input data to others and the mean value during the test phase. The result of the positioning is related to the value of the k nearest neighbors based on RSSI.

Random Forest (RF) is a machine learning consisting of many decision trees (DTs). It has similar effect to KNN since the different DTs can be seen as the nearest neighbors in KNN.

However, the difference is that the DT builds classification models in a tree structure by putting the most important feature as a root, then splitting the data into smaller subsets with the same-important-level-features and repeating the two steps until we finally have a tree with decision nodes. In [13] it has a new technique named Hybloc applies a hybrid system based on RF ensembles utilizing Gaussian Mixture Model (GMM) soft clustering to partition the input data. Since GMM distribution is similar to the Wi-Fi propagation characteristics in nature, it has a good performance for RSSI samples clustering. Thus it helps the construction of RF ensembles by the inclusion of relevant samples, which decreases the computational time and enhances the accuracy at the same time.

DeepFi is a novel deep learning based indoor fingerprinting theme that uses CSI information in[14]. After we have the normalized CSI values, it adopts a new approach to represent fingerprints using the weights which are analyzed in a deep network with four hidden layers instead of directly stocking the CSI values in the traditional way[15]. For the test phase, a probabilistic data fusion method based on radial basis function is developed for online location estimation using multiply packets. The result of the positioning corresponding to the probability model based on Bayes' law which is given by

$$Pr(L_i|v) = \frac{Pr(L_i)Pr(v|L_i)}{\sum_{i=1}^{N} Pr(L_i)Pr(v|L_i)}$$

Where L_i is reference location i, $Pr(L_i)$ is the prior probability that the mobile device is determined to be at reference location i.

II.3 Discussion of the performance

CNN achieves high accuracy in image recognition and positioning, but this also requires a high computational cost which means it is slower to obtain a result.

KNN has the simplest manipulation but its drawback is when the datasets increase, the calculation process will be overloaded since it needs the distance values from one another.

RF and DT are frequently used for data mining and statistical learning. The limitation is that trees can be very non-robust which means a small change in the training data could cause a huge change in the tree. Moreover the over-complex trees lead to the problem of overfitting.

Since DeepFi uses CSI, it has an extra treatment for the input data compared with RSSI.

Many existing indoor localization systems use RSSI as fingerprints due to its simplicity and low hardware requirements. But the main drawback is that it has a high variability over time for a fixed location because of the multipath effects. Even though CSI costs more time, the accuracy is higher than RSSI.

III. Conclusion

In this paper, different indoor localization techniques are presented and criticized. Most indoor localization applications today are based on RSSI fingerprints. However RSSI measurements can be significantly affected by noise and environmental changes. In that case, RSSI performs better in simple environments while CSI achieves higher accuracy in complex occasions such as the environments with human mobility and temporal dynamics but also has more calculation time cost. As for the different indoor localization techniques, many of them are done based on CNN since the experiments have proved that it outperforms KNN, DT and RF with robust and higher accuracy. But it should be applied more in the complex environments because its performance also depends on the large amount of training data which can be a little waste in simple situations. So KNN is the most recommended technique in the simple and stable environments since it has the simplest algorithm. The only one hyper-parameter which needs to be defined is a suitable K value because big K value would cause underfitting while small K value causes overfitting.

Thanks to the more and more efforts on exploring indoor localization techniques, we now have many different new techniques based on the machine learning with great ameliorations. For example, HW-fingerprint is related to RSSI but with more stability by calculating the ratio of RSSI, DeepFi is a technique based on CSI but reduces the complexity of the input database treatment and also, the technique Hybloc which combines the GMM soft clustering with the RF to reduce the data partitioning time in training of RF and enhance the accuracy.

References

[1]H. Liu, H. Darabi, P. Banerjee, and L. Jing, "Survey of wireless indoor positioning techniques and systems," IEEE Trans. Syst., Man, Cybern. C, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.

[2]Y.Hou, X.Yang and Q.H.Abbasi, "Efficient AoA-based wireless indoor localization for hospital outpatients using mobile devices", Sensors 2018, 18, 3698.

[3]Gu, Y.; Chen, M.; Ren, F.; Li, J. HED "Handling environmental dynamics in indoor WiFi fingerprint localization." In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), Doha, Qatar, 3–6 April 2016; pp. 1–6.

[4]Zimu Zhou, Zheng Yang, Chenshu Wu, Wei Sun, Yunhao Liu," LiFi: Line-Of-Sight identification with WiFi," INFOCOM 2014: 2688-2696

[5]S.Shue, J.M.Conrad, "Reducing the Effect of Signal Multipath Fading in RSSI-Distance Estimation using Kalman Filters", in Proceedings of the 19th Communications & Networking Symposium, 2016, 1-6

[6]Yang, Z.; Zhou, Z.; Liu, Y. "From RSSI to CSI: Indoor localization via channel response." ACM Comput. Surv. 2013, 46, 1–32.

[7]AÇ.Seçkin, A.Coskun,"Hierarchical Fusion of Machine Learning Algorithms in Indoor Positioning and Localization".Appl. Sci. 2019, 9, 3665.

[8]M. M. Atia, A. Noureldin, IEEE, and M. J. Korenberg, "Dynamic online-calibrated radio maps for indoor positioning in wireless local area networks," IEEE Trans. Mobile Comput., vol. 12, no. 9, pp. 1774–1787, Sept. 2013.

[9]W. Au, C. Feng, S. Valaee, S. Reyes, S. Sorour, S. N. Markowitz, D. Gold, K. Gordon, and M. Eizenman, "Indoor tracking and navigation using received signal strength and compressive sensing on a mobile device," IEEE Trans. Mobile Comput., vol. 12, no. 10, pp. 2050–2062, Oct. 2013.

[10]W.Njima, I.Ahriz, R.Zayani, M.Terre and R.Bouallegue. "Deep CNN for Indoor Localization in IoT-Sensor Systems,"July, 2019.

[11]Z.Liu, B.Dai, X.W and X.Li, "Hybrid Wireless Fingerprint Indoor Localization Method Based on a Convolutional Neural Network", Sensors 2019, 19, 4597.

[12]Masand, B.; Linoff, G.; Waltz, D. Classifying news stories using memory based reasoning. In Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Copenhagen, Denmark, 21–24 June 1992; pp. 59–65.

[13]B.A. Akram, A.H.Akbar and O.Shafiq, "Hybloc: Hybrid Indoor Wi-Fi Localization Using Soft Clustering-Based Random Decision Forest Ensembles," IEEE Access, vol. 6, pp. 2169-3536, July, 2018.

[14]X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," in Proc. IEEE WCNC 2015, New Orlean, LA, Mar. 2015, pp. 1666–1671.

[15]G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, no. 5786, pp. 504–507, July 2006.