Towards Accurate Indoor Localization using Channel State Information

Wei Kui[†], Shiling Mao[†], Xiaojun Hei✉[†], Fan Li[‡]
Huazhong University of Science and Technology, Wuhan, China, 430074
Beijing Institute of Technology, Beijing, China, 100081
Email: {kuiwei, mshil, heixj}@hust.edu.cn, fli@bit.edu.cn

***Abstract*—Indoor location-based mobile applications have been gaining momentum in reshaping the daily activities of Internet users. A large number of indoor localization techniques achieve the localization goal by analyzing the received signal strength indication (RSSI) of pervasive WiFi signals. Compared with RSSI, the channel state information (CSI) provides more comprehensive time and space information with more complex hardware and software cost. In this paper, we proposed two CSI-based indoor localization algorithms: 1) a localization algorithm based on the weighted linear discriminant analysis; 2) a localization algorithm based on two-dimensional principal component analysis. The experimental results show that the proposed algorithms outperform the basic Bayesian algorithm based on the principal component analysis on improving the localization accuracy and reducing the computational complexity.**

***Index Terms*—Indoor Localization; Channel State Information; Principal Component Analysis; Bayesian Inference**

I. Introduction

Traditional indoor positioning techniques are commonly device-dependent and may be impacted easily by the complex wireless environments based on analyzing the received signal strength indication (RSSI) of pervasive WiFi signals. The RSSI captures the aggregate effect of multipath transmissions and can not extract the information of each single path [1]. In recent years, the rich information of the underlying physical layer has been examined for indoor positioning, such as the channel state information (CSI). The open source CSI measurement tool running on off-the-shelve commercial WiFi devices [2] has stimulated increasing research efforts on many CSI-based localization algorithms. Nevertheless, the quality of the CSI data requires careful examinations [3]. In this paper, we propose two algorithms: 1) a localization algorithm based on the weighted linear discriminant analysis; 2) a localization algorithm based on a two-direction two-dimensional principal component analysis.

II. Algorithms

The positioning algorithm mainly consists of two phases (offline phase and online phase). In terms of functions, it includes three parts: data collection, feature extraction and localization. Suppose the training sample set is $X = [x_1, \ldots, x_N]^T, x_i = [x_{i1}, \ldots, x_{iM}]$, where N is the total number of training samples in the training set and M is the number of features in each sample.

A. Weighted PCA-based Localization

Data Pre-Processing. The data of different carriers have the distribution correlation, and the subcarrier distribution data center is offset, the distribution trends are relatively similar, and the data is delineated to highlight the difference between the data.

Data Weighting. The weighted principal component analysis (WPCA) [4] weight data samples to reflect the decisive role played by the sample in the population, so the choice of the weighting functions is important for the algorithm performance, and the weight should reflect the overall distribution trend of the data. We study two weighting functions including summative weighting and multiplicative weighting.

Weighted Principal Component Analysis. The basic idea of the WPCA algorithm is based on the role of the sample data in the identification process to weight the sample components and the sample itself, which strengthen the stability of the samples and the conventional samples and weaken the unconventional samples (outliers et al.). This method can obtain effective feature information to improve the accuracy of location identification.

Bayes Localization. The basic idea of the Bayes algorithm is as follows: calculate the probability that the sample data to be located belongs to all categories, and then select the known category with the highest probability as the category of data to be located.

B. Weighted LDA-based Localization

The basic idea of the Linear Discriminant Analysis (LDA) algorithm based on the Fisher discriminant criterion is to select the vector that maximizes the Fisher criterion function as the best projection direction, and the projected data can reach the maximum between-class dispersion and the minimum within-class dispersion. Both PCA and LDA are the subspace methods. The basic idea of PCA is how to express the original data with the least error eigenvector, keeping the original data as much as possible. The basic idea of LDA lies in how to optimize the existing data from the existing Data extracted from the difference between classes, LDA has a better recognition rate theoretically. The LDA algorithm is used to reduce the dimensionality of the original data and use the data as the input of the Bayesian algorithm. The WLDA-Bayesian algorithm based on the weighted linear discriminant is implemented by combining the weighting algorithms.

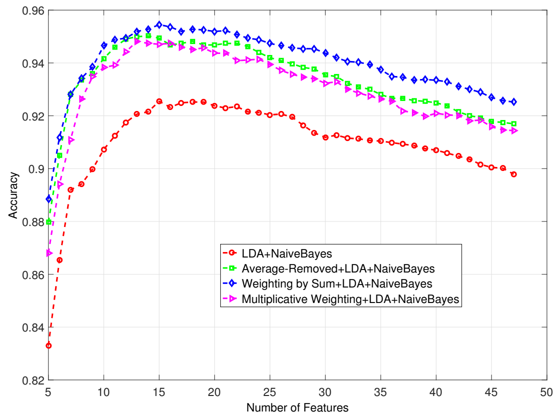

Fig. 2: Accuracy improvement with the increasing number of features for WLDA

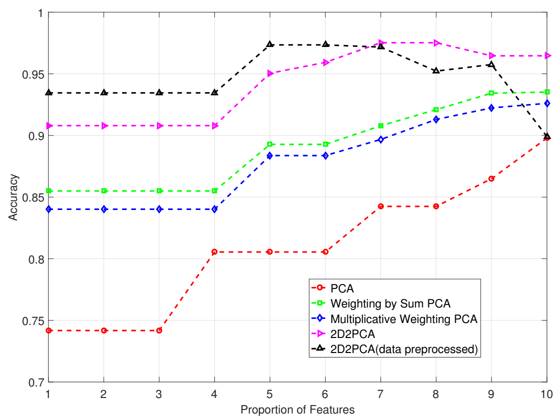

Fig. 3: Accuracy comparison: $(2D)^2PCA$ vs. WPCA

C. *Two-dimensional PCA-based Localization*

The two-dimensional principal component analysis ($2DPCA$) [5] calculates the covariance matrix directly from a two-dimensional matrix based on a two-dimensional matrix and obtains the optimal projection vector for feature extraction. Let $X \in R^{n\times d}$ be orthogonal unit column vector matrix, A is a "location image" of $m \times n$. In order to further reduce the dimension of feature image information extracted by $2DPCA$, we proposed a localization algorithm based on a two-direction two-dimensional principal component analysis ($(2D)^2PCA$) to calculate the image in the row and column direction respectively, then we can use optimal matrix X and projection image A reflecting line information to obtain the $m \times d$ matrix $Y = AX$, and the image A is projected using the optimal matrix Z reflecting the column information to generate a $q \times n$ matrix $B = Z_T A$, after obtaining the mapping matrix X, Z.

A single CSI data sample is susceptible to sample data fluctuations. The similarity of multiple samples and the stability of distribution can reflect the location characteristics better. Therefore, we use multiple pieces of sample data to construct two-dimensional positional image features, "location images". Suppose the number of selected samples is m, then the size of "location image" is $m\times n$, where n is the number of features in each CSI's record. "Location images" have both column and row dependencies. $(2D)^2PCA$ can remove row and column dependencies on the basis of preserving original information. So $(2D)^2PCA$ can be used to reduce the dimension of the "location images" and the location can be achieved by integrating the Bayes algorithm and the KNN algorithm.

III. Evaluation

We compare the two proposed algorithms with the traditional PCA-based localization algorithm. Fig. 1 shows that data weighting can effectively improve the location accuracy. Fig. 2 shows the accuracy improvement with the increasing number of features for the WLDA algorithm. Fig. 3 shows the accuracy comparison between the $(2D)^2PCA$ algorithm and the WPCA algorithm.

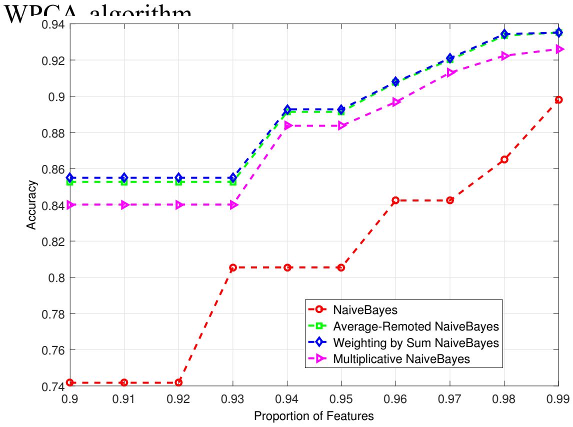

Fig. 1: Accuracy comparison: WPCA vs. PCA

IV. Conclusion

In this paper, we proposed the WLDA algorithm and the $(2D)^2PCA$ algorithm to combat against the fluctuation of indoor signals. The experimental results show that the location accuracy of the basic WPCA is up to 93.5%, the location accuracy of WLDA is up to 95%, the location accuracy of $(2D)^2PCA$ is up to 99.6%. The $(2D)^2PCA$ algorithm achieve the highest location accuracy than the WPCA algorithm.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (no. 61370231).

References

[1] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Computing Surveys*, vol. 46, no. 2, pp. 1–32, 2013.

[2] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: gathering 802.11n traces with channel state information," *ACM Sigcomm Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011.

[3] H. Cheng, X. Hei, and D. Wu, "An experimental study of harvesting channel state information of WiFi signals," in *IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*, May 2018.

[4] Z. Niu and X. Qiu, "Facial expression recognition based on weighted principal component analysis and support vector machines," in *International Conference on Advanced Computer Theory and Engineering*, 2010, pp. V3–174–V3–178.

[5] J. Yang, D. Zhang, A. F. Frangi, and J. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 26, no. 1, p. 131, 2004.