Subcarrier selection for efficient CSI-based indoor localization

Yu Taso¹, Shih-Chun Yeh², Yu-You Liang²*, Chu-Hsuan Wang² and Shih-Hau Fang²

¹Academia Sinica, Taiwan
²Yuan-Ze University, Taiwan

*Email: y.y.liang.hchs@gmail.com

Abstract. Indoor positioning systems have received increasing attention for supporting location-based services. In recent Wi-Fi networks, the rich information in the physical layer, known as channel state information (CSI), has been recognized an effective positioning characteristic rather than traditional received signal strength. However, the positioning performance depends on a very high-dimensional CSI due to all pairs of transceiver antenna, which may incur over-fitting problems. This paper proposes a subcarrier-selection approach based on information theoretic learning to compensate for over-fitting problems in CSI-based localization systems. After equalizing the histogram of CSIs, the proposed algorithm computes the information gain of each subcarrier and forms a new low-dimensional subset of CSIs to reduce the complexity and to decrease possible over-fitting caused by redundant CSIs. We demonstrate the effectiveness of the proposed algorithm through experiments. On-site experimental results demonstrate that the proposed approach outperforms traditional feature selection schemes.

1. Introduction
Indoor localization has become a crucial concern for various mobile applications in pervasive computing environments [1]. The massive deployment of Wi-Fi access points (APs) has made Wi-Fi a suitable technology for developing such indoor location systems. Numerous studies have addressed indoor location estimation by using existing Wi-Fi infrastructure. Among the various indoor Wi-Fi positioning systems, the fingerprinting-based approach, in which the user’s location is estimated by matching receive signal strength (RSS) with the radio map, is one of the most feasible solutions [2].

In recent Wi-Fi systems with the IEEE 802.11a/g/n standard, the Orthogonal Frequency Division Multiplexing (OFDM) and Multiple-input and multiple-output (MIMO) technology have been used, where data are modulated on multiple subcarriers in different frequencies and transmitted among different pairwise antenna to increase throughput [3-4]. Thanks to the development of firmware and open sources, the wireless MIMO channel response can be extracted from the receivers in the format of channel state information (CSI) [5-7].

CSI reveals a set of channel measurements depicting the amplitudes and phases of every subcarrier. Recently, CSI-based fingerprinting approaches, which uses PHY layer fine grained information, have been proposed for indoor localization. Instead of receive signal strength (RSS), these approaches exploit CSI as new positioning features, with the location granularity to improve the accuracy [8-14].
However, the positioning performance depends on a very high-dimensional CSI due to all pairs of transceiver antenna, which may incur over-fitting problems.

The studies in [15] proposed FIFS, which leverage both the frequency and spatial diversity and uses a correlation filter with the probability positioning algorithm. That is, FIFS represents the location by aggregating the power of all subcarriers. However, although the summation over all subcarriers can significantly reduce the dimensions, it does not consider the detail information within subcarriers. A similar approach, namely, MIMO, was presented in [16-17], in which the amplitude and phase value were subtracted for subsequent subcarrier to generate the location fingerprint. Although this approach effectively extracts the relatively robust fingerprints owing to the subtraction process, the high-dimension still maintains.

This paper presents a robust CSI-based indoor localization using histogram equalization (HEQ) [18] and information theoretic learning (ITL) [19] to reduce the required dimensions. Specifically, we propose a subcarrier-selection approach based on information theoretic learning to compensate for over-fitting problems. There are two steps of the proposed algorithm. First, we normalize CSIs using the HEQ technique by converting the coefficients into a reference to enhance the system robustness. Second, after equalizing the histogram of CSIs, the proposed algorithm computes the information gain of each subcarrier based on ITL and forms a new low-dimensional subset of CSIs to reduce the complexity and to decrease possible over-fitting caused by redundant CSIs. We demonstrate the effectiveness of the proposed algorithm through experiments. On-site experimental results demonstrate that the proposed approach provide compare or even better performance with a relatively lower-dimension subcarrier subset.

The rest of the paper is structured as follows. Section II illustrates the proposed localization algorithm. Section III presents the experimental setup and results. Finally, the conclusion is given in Section IV.

2. Proposed algorithm
To enhanced data throughput, Wi-Fi networks used OFDM (orthogonal frequency division multiplexing) and MIMO (multiple input multiple output) techniques in the 802.11a/g/n standard, where data are modulated on multiple channels in different frequencies and are transmitted simultaneously among multiple antenna pairs. In these systems, channel response can be extracted from the receivers in the format of channel state information (CSI), as indicated in figure 1.

![Figure 1. Main idea of CSI in a MIMO system.](image)

Because CSI reports each channel state for each antenna pair, we assume that a complete measured CSI signal in the MIMO wireless network is given as equation (1):

\[ h = [h_{11}, \cdots, h_{1M}, \cdots, h_{NM}, \cdots, h_{NM}]^T \] (1)

To localize a user, a typical fingerprinting location algorithm estimates the likelihood of real-time CSI observations for all candidate locations and selects the one having maximum likelihood as the
result. Let $R$ be the entire set of reference locations; the CSI-fingerprinting localization problem can be formulated as equation (2)

$$r^* = \arg \max_{r \in R} P(h|\lambda_r)$$

(2)

where $\lambda_r$ is the probabilistic model of the $r$-th reference location, which can be obtained on the basis of the training data during the offline stage [20]. A challenging issue in CSI-based localization is the very high-dimensional CSI. In this study a mechanism is designed to select the important CSIs as positioning feature.

First, we normalize the CSIs using the histogram equalization (HEQ) technique (i.e., converting the coefficients into a reference). HEQ provides a transformation that converts the probability density function of an original variable into a reference probability density function. This study uses a typical Gaussian as the reference. Then, the proposed approach ranks CSIs in descending order of their InfoGain values which are calculated as equation (3):

$$\text{InfoGain}(h_i) = I(D) - I(h_i), \ i = 1, \cdots, NM$$

(3)

where $I(D)$ is the information incurred in the entire training database $D$ (entropy of the reference locations without $h_i$), and $I(h_i)$ computes the conditional entropy of the location given $h_i$. This approach calculates the discriminative ability for each subcarrier and the top are considered the best. As the above equation indicates, the feature with larger information gain is assigned a higher rank, representing more contribution in indoor localization.

![Figure 2](image2.png)

**Figure 2.** Part of the fifth floor plane of the test-bed building, where we had performed the experiments.

### 3. Performance evaluation

#### 3.1. Experimental setup

This study involved experiments conducted on the fifth floor of the telecommunications building at Yuan-Ze University. Figure 2 shows where the experiments were performed. The size of the test-bed was 14 (m) × 10 (m). One access point (ASUS RT-N12E) with two transmitted antennae (M=2) was deployed at the corner. After installing the Linux 802.11n driver, which is built on the Intel Wi-Fi Wireless Link 5300 MIMO radios, the gain and phase of the signal path between a single transmit–receive antenna pair is available in the measured device (Lenovo x200). In the experiments, the laptop measured CSIs using three receiver antenna (N=3), and thus there were 6 links available as positioning features. Each reported channel state contains 30 subcarrier channel groups. Each channel state matrix entry is a complex number and only the magnitudes were used in this study.

Figure 3 shows the realistic RSS and CSI measurements over 100s at two different locations. In figures 3(b) and 3(d), CSIs are measured over 100s at a fixed location, indicated by a red and blue circles in figure 2, respectively. The wider curve represents the higher time variation of an antenna pair. Similarly, figures 3(a) and 3(c) show the measured RSSs at the same locations. Figure 3 clearly
reveals the changing CSI patterns measured from two different locations. It indicates the possibility of determining the location on the basis of CSI. Moreover, the figures verify that CSI provides fine-grained information, for each subcarrier and each Tx-Rx pair, whereas RSS provides coarse-grained information only.

![Figure 3. The measured RSSs and CSIs over 100s at two different locations.](image)

**Figure 3.** The measured RSSs and CSIs over 100s at two different locations.

### 3.2. Performance evaluation

Figure 4 displays the mean positioning error versus the number of subcarriers based on six algorithms, including RSS, CSI, HEQ, MIMO, FIFS, and the proposed ITL algorithm. The HEQ approach indicates the direct normalization of CSI, and the RSS and FIFS methods are constant because the performance is independent of the number of subcarriers. This figure first shows that the error generally decreases as the number of channels increases. Moreover, this figure indicates the effectiveness of ITL, which considerably reduce the error while using the lowest number of subcarriers. This shows that ITL can select the useful location-related information among CSIs more effectively.

Next, figure 5 compare the performance of different subcarrier selection methods on mean positioning error. In this figure, the Maxmean method selects the most stronger CSIs (in terms of the maximal mean magnitude over time) while the Minvar approach selects the most stable ones (in terms of the minimal variance over time). This figure shows that the proposed ITL algorithm outperforms traditional selection approaches. Using 40 subcarriers, ITL reduces the mean localization errors by 40% and 22%, respectively, compared to Maxmean and Minvar. The trend is similar using 50 subcarriers. The results again demonstrate that ITL constructs a more suitable subset of subcarriers according to information gain, thus achieving more accurate location estimations with fewer numbers of CSIs. The computed information gain of CSIs is shown in figure 4.
4. Conclusions

This paper proposes a subcarrier-selection approach based on ITL to compensate for over-fitting problems in CSI-based localization systems. After equalizing the histogram of CSIs, the proposed algorithm computes the information gain of each subcarrier and forms a new low-dimensional subset of CSIs to reduce the complexity and to decrease possible over-fitting caused by redundant CSIs. We demonstrate the effectiveness of the proposed algorithm through experiments. The experimental results demonstrate that ITL constructs a more suitable subset of subcarriers according to information gain, thus achieving more accurate location estimations with fewer numbers of CSIs.

References

[1] Y. Gu, A. Lo, and J. Niemegeers, “A Survey Of Indoor Positioning Systems For Wireless Personal Networks,” IEEE Communications Surveys & Tutorials, vol. 11, no. 1, pp 13–32, 2009

[2] A. S. Paul and E. A. Wan, “RSSI-Based Indoor Localization And Tracking Using Sigma-Point Kalman Smoothers,” IEEE J. Select. Topics Signal Process, vol. 3, pp 860–73, Oct 2009

[3] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, “An Overview Of MIMO Communications - A Key To Gigabit Wireless,” Proc. IEEE, vol. 92, pp 198–218, Feb 2004

[4] P. Xia, S. Zhou, and G. B. Giannakis, “Adaptive MIMO-OFDM Based On Partial Channel State Information,” IEEE Trans. Signal Process., vol. 52, pp 202–13, Jan 2004

[5] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, “Predictable 802.11 Packet Delivery From Wireless Channel Measurements,” SIGCOMM Comput. Commun. Rev., vol. 41, no. 4, pp 159–70, 2011

[6] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, “Tool Release: Gathering 802.11n Traces With Channel State Information,” SIGCOMM Comput. Commun. Rev., Jan 2011

[7] S. Sen, R. R. Choudhury, B. Radunovic, and T. Minka, “Precise Indoor Localization Using PHY Layer Information,” in ACM HotNets, pp 18:1–18:6, 2011

[8] K. Wu, J. Xiao, Y. Yi, M. Gao, and L. Ni, “FILA: Fine-Grained Indoor Localization,” in INFOCOM, pp 2210–18, March 2012

[9] Z. Yang, Z. Zhou, and Y. Liu, “From RSSI To CSI: Indoor Localization Via Channel Response,” ACM Comput. Surv., vol. 46, no. 2, pp 25:1–25:32, November 2013

[10] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. Ni, “CSI-Based Indoor Localization,” IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 7, pp 1300–09, July 2013

[11] Z.-P. Jiang, W. Xi, X. Li, S. Tang, J.-Z. Zhao, J.-S. Han, K. Zhao, Z. Wang, and B. Xiao, “Communicating Is Crowdsourcing: Wi-Fi Indoor Localization With CSI-Based Speed
Estimation,” *JCST*, vol. 29, no. 4, pp 589–604, 2014

[12] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, “SpotFi: Decimeter Level Localization Using Wi-Fi” in *SIGCOMM*, pp 269–82, 2015

[13] F. Wen and C. Liang, “Fine-Grained Indoor Localization Using Single Access Point With Multiple Antennas,” *IEEE Sens. J.*, vol. 15, pp 1538–44, March 2015

[14] X. Wang, L. Gao, S. Mao, and S. Pandey, “CSI-Based Fingerprinting For Indoor Localization: A Deep Learning Approach,” *IEEE Trans Veh Technol*, vol. PP, no. 99, pp 1–1, 2016

[15] J. Xiao, K. Wu, Y. Yi, and L. Ni, “FIFS: Fine-Grained Indoor Fingerprinting System,” in *ICCCN*, pp 1–7, July 2012

[16] Y. Chapre, A. Ignjatovic, A. Seneviratne, and S. Jha, “CSI-MIMO: Indoor Wi-Fi Fingerprinting System,” in *Proc. IEEE LCN*, pp 202–09, Sept 2014

[17] Y. Chapre, A. Ignjatovic, A. Seneviratne, and S. Jha, “CSI-MIMO: An Efficient Wi-Fi Fingerprinting Using Channel State Information With MIMO,” *Pervasive and Mobile Computing*, vol. 23, pp 89–103, July 2015

[18] S. H. Fang, C. H. Wang, and Y. Tsao, “Compensating For Orientation Mismatch In Robust Wi-Fi Localization Using Histogram Equalization,” *IEEE Trans. Veh. Technol.*, vol. 64, pp 5210–20, Nov 2015

[19] J. C. Principe, *Information Theoretic Learning: Renyi’s Entropy and Kernel Perspectives*. Springer Verlag, 2010

[20] S. H. Fang, W.H. Chang, Y. Tsao, H. C. Shih, and C. W. Wang, “Channel State Reconstruction Using Multilevel Discrete Wavelet Transfrom for Improved Fingerprinting-Based Indoor Location,” *IEEE Sensors Journal*, 2016