Towards People Counting Using Wi-Fi CSI of Mobile Devices

Masahide Mizutani
Grad. Sch. of Info. Sci.& Tech.
Osaka University
Osaka, Japan
m-mizutani@ist.osaka-u.ac.jp

Akira Uchiyama
Grad. Sch. of Info. Sci.& Tech.
Osaka University
JST, PRESTO
Osaka, Japan
uchiyama@ist.osaka-u.ac.jp

Tomoki Murakami
Access Network Service Systems Laboratories
Nippon Telegraph and Telephone Corporation
Kanagawa, Japan
tomoki.murakami.nm@hco.ntt.co.jp

Hirantha Abeysekera
Access Network Service Systems Laboratories
Nippon Telegraph and Telephone Corporation
Kanagawa, Japan
hirantha.abeysekera.eu@hco.ntt.co.jp

Teruo Higashino
Grad. Sch. of Info. Sci.& Tech.
Osaka University
Osaka, Japan
higashino@ist.osaka-u.ac.jp

***Abstract*—Monitoring crowdedness in public spaces such as shopping malls is useful for various services like marketing, safety, evacuation planning. For infrastructure-based people counting, Wi-Fi Channel State Information (CSI) has attracted attention as it does not require any additional deployment. However, the existing approaches assume multiple fixed infrastructures such as Wi-Fi access points (APs), which is different from standard AP deployment since they are installed for wireless network services. To solve this problem, our goal is to estimate the number of people (i.e. people counting) by using mobile devices such as smartphones and a small number of fixed APs. Since CSI represents the difference between the amplitude and the phase of the transmitted and received radio waves, we can estimate the change of the propagation environment from CSI. However, CSI is very noisy due to various factors that are not related to the number of people. In this paper, we focus on Sampling Frequency Offset (SFO), Carrier Frequency Offset (CFO), A/D Converter (ADC) delay, and quantization error, and design methods to mitigate their effects. Then, we propose people counting by using CSI variance as a location-independent feature for machine learning. From the evaluation results, we have confirmed that our method achieves the Root-Mean-Squared-Error (RMSE) of 0.49, which is much better than RMSE of 3.7 without denoising.**

***Index Terms*—Channel State Information, CSI, Wi-Fi, People Counting, Mobile Sensing**

I. Introduction

In public spaces such as shopping malls, stations, airports, information about crowdedness at each area is useful for various purposes like marketing, safety, and evacuation planning. Various approaches have been investigated to estimate the number of people (i.e. people counting). A camera-based approach [1] is one of the emerging methods because of the significant progress of deep learning. However, camera blind spots are a major drawback which may require additional camera installation. Privacy concern is another major drawback of the camera-based approach. Another approach is Wi-Fi probing which counts the number of Wi-Fi MAC addresses [2]. Although Wi-Fi probing does not require additional deployment, the estimation accuracy decreases depending on environments because each person does not always have one mobile device. Some people may have multiple mobile devices while some others, such as children do not have any devices.

On the other hand, Wi-Fi Channel State Information (CSI) has attracted attention as it can passively monitor crowdedness by monitoring the change of radio propagation at Wi-Fi Access Points (APs) [3]–[6]. CSI is information acquired between a transmitter and a receiver to improve wireless communication performance using Multiple-Input and Multiple-Output (MIMO) specified in 802.11n and later. For each subcarrier of Orthogonal Frequency Division Multiplexing (OFDM), the change in phase and amplitude between Tx and Rx antennas is reported by complex numbers as CSI. Compared to RSSI(Received Signal Strength Indicator)-based people counting [7], CSI provides more abundant information (i.e. phase and amplitude) on the radio propagation paths. Furthermore, we obtained CSI for each subcarrier, which further increases the amount of information.

The existing CSI-based people counting assumes multiple transmitters and receivers (e.g. APs) deployed in the target environment and builds a model to estimate crowdedness (the number of people or the crowdedness level) by machine learning. However, most of them require dense AP deployment (e.g. four APs in one lecture room), which is not realistic since APs are originally installed for wireless communication.

To solve the problem, our goal is to estimate the number of people in the target area by using mobile devices such as smartphones. We assume a small number of fixed APs are deployed, and some cooperative users hold mobile devices. We can capture Wi-Fi CSI between each pair of an AP and a mobile device within the AP radio range. Then, we estimate the number of people in the AP radio range based on the reported CSI within the radio range (see Fig. 1).

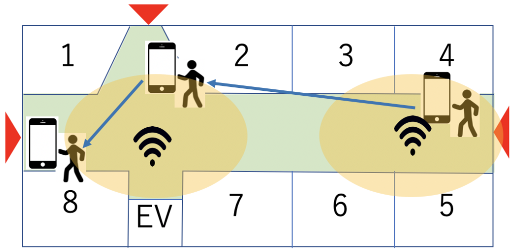


Fig. 1. Target Environment

The first challenge in our approach is the mobility of the devices because CSI is sensitive to changes in the radio propagation environment. The other challenge is CSI distortion due to CSI measurement noise. To tackle the first challenge, we propose variances of the amplitude and the phase as location-independent features. For the second challenge, we design a denoising scheme by modelling the CSI noise. Specifically, we focus on four types of noise: Sampling Frequency Offset (SFO), Carrier Frequency Offset (CFO), conversion delay in A/D converter, and quantization error. On the receiver side, SFO and CFO are introduced due to the offset of the oscillators between a transmitter and a receiver. To mitigate SFO and CFO, we employ the CSI ratio [8]. In the CSI ratio, SFO and CFO are removed by taking the ratio of CSI for each antenna pair by using the fact that CSI is simultaneously acquired for all antenna pairs, which means CSIs obtained by different antennas share the same SFO and CFO. We also mitigate A/D conversion delay by focusing on its linearity since it is due to the processing delay of A/D converter on the reception of the signal of each subcarrier. Moreover, the quantization error is caused by the loss of information during conversion from an analogue signal to CSI by the A/D converter. Since the quantization error becomes more significant as the signal weaker, we selectively use the subcarriers with large amplitudes. Finally, we build a model for people counting by using CSI variance as a feature by machine learning. From the real experiment up to 9 people, we have confirmed the RMSE is 0.49 while it is 3.7 without our denoising scheme. The result also shows that the performance of our variance-based people counting is independent of device positions.

Our contributions are summarized as below.

- To the best of our knowledge, this is the first study to explore the possibility of CSI-based people counting by using mobile devices.
- We propose a denoising scheme to mitigate the CSI measurement noise based on the CSI noise model.
- We show the effectiveness of our method through the preliminary experiment in the real environment.

II. Related Work

A. People Counting by CSI

Several methods for estimating the number of people using CSI obtained between indoor transceivers have been proposed [3]–[6]. In the document [3], Han Zou et al. estimate the number of people in the room using Transfer Kernel Learning (TKL), and they achieved the accuracy of 96% in the classification of 0 to 7 people. Wei Xi et al. defined a feature called Percentage of nonzero Elements (PEM) [4], quantified the fluctuation in the radio wave propagation, and estimated the number of people based on the Gray Model. In these existing studies, they fix both the transmitter and receiver to obtain CSI and estimate the number of people with supervised learning. For this reason, it is necessary to collect teacher data anew if any position of the transceiver changes. In literature [5], it reports that they use a learning model created by CSI acquired in one room for estimating the number of people in another room, and no significant performance degradation is observed. However, it is not clear how the congestion estimation performance changes according to the positional relationship between the transceivers.

B. People Counting with Mobile Phone

Li et al. Are working on a congestion estimation method using mobile terminals and fixed base stations [7]. Using Received Signal Strength Indicator (RSSI) obtained between the smartphone and the base station, they estimate the number of people in the room in increments of ten people from zero to fifty people. However, they installed as many as thirty-one base stations in a room of $8m \times 12m$, and there is a concern that the cost will be high. On the other hand, this research aims to count people for a wide range, even with a small number of base stations using CSI acquired between mobile terminals and fixed base stations.

III. Channel State Information

In wireless LAN standards IEEE 802.11n and later, MIMO (Multiple-Input and Multiple-Output) is adopted to improve the quality of communication. OFDM (Orthogonal Frequency-Division Multiplexing) is used as a modulation scheme that uses multiple orthogonal subcarriers. MIMO leverages CSI (Channel State Information) for transmission signal control to improve the quality of the signal at the receiver. In CSI, the amplitude and phase difference for each subcarrier in this OFDM modulation is obtained. Let $\boldsymbol{X}(f,t)$ and $\boldsymbol{Y}(f,t)$ be the frequency domain representations of transmitted and received signals, respectively, with carrier frequency f. The relationship between $\boldsymbol{X}(f,t)$ and $\boldsymbol{Y}(f,t)$ are written as:

$$\boldsymbol{Y}(f,t) = \boldsymbol{H}(f,t) * \boldsymbol{X}(f,t) + \boldsymbol{N}(f,t), \tag{1}$$

where $\boldsymbol{H}(f,t)$ is the CSI and $\boldsymbol{N}(f,t)$ is the noise. Therefore, assuming that the number of subcarriers is S for N_{T_x} antennas of the transmitter and N_{R_x} antennas of the receiver, we can obtain $N_{T_x} * N_{R_x} * S$ pairs of the amplitude and the phase. Compared to RSSI, CSI contains more abundant information on conditions of the radio propagation paths. Since the channel states change due to dynamic components including human movement, CSI has been used for not only the improvement of communication quality but also applications such as activity recognition.

CSI is estimated by sending pilot signals from a transmitter to a receiver. In practice, the estimated CSI is distorted by various noise factors due to hardware imperfection. According to the existing work [8]–[11], we focus on the following four factors.

- **CFO and SFO:** Carrier Frequency Offset (CFO) and Sampling Frequency Offset (SFO) are introduced because the oscillators of the transmitter and the receiver are not exactly synchronized.
- **ADC Delay:** Since pilot signals are processed by the A/D converter of the receiver one by one for each subcarrier, ADC Delays are accumulated linearly with the processing order of the subcarriers.
- **Quantization Error:** A/D conversion of the pilot signals on the receiver introduces the quantization error, which amplifies the noise originally included in the CSI. The quantization error becomes larger as the signal amplitude smaller.

In the following section IV, we describe data processing to mitigate the above noise factors.

IV. Data Processing

A. CSI Ratio

Each element H of the CSI $\boldsymbol{H}$ is ideally represented as:

$$H(f,t) = \sum_{i=1}^{L} A_i e^{-2j\pi \frac{d_i(t)}{\lambda}}, \tag{2}$$

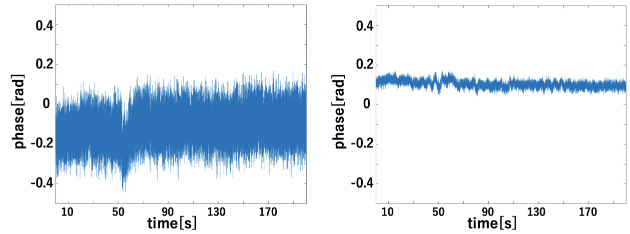
where L is the number of paths, A_i is the complex attenuation, $d_i(t)$ is the propagation length of the i-th path, and λ is the wave length of the subcarrier. However, in a real environment, a random offset θ_{offset} due to SFO and CFO is added, which is written as the following equation.

$$\begin{aligned} H'(f,t) &= e^{-j\theta_{offset}} \sum_{i=1}^{L} A_i e^{-2j\pi \frac{d_i(t)}{\lambda}} && (3) \\ &= e^{-j\theta_{offset}} H(f,t) && (4) \end{aligned}$$

To remove the random offset due to SFO and CFO, the ratio of CSI between a pair of antennas on the transceiver is used, which cancels out the phase offsets [8], [11]. This is because all the antennas of a wireless device (i.e. the transmitter or the receiver) are connected to the same oscillator, which results in the same SFO and CFO. The CSI ratio $H_{n,m}$ between anntenas m, n is defined as below.

$$\begin{aligned} H_{n,m}(f,t) = \frac{H'_n(f,t)}{H'_m(f,t)} &= \frac{e^{-j\theta_{offset}} H_n(f,t)}{e^{-j\theta_{offset}} H_m(f,t)} && (5) \\ &= \frac{H_n(f,t)}{H_m(f,t)} && (6) \end{aligned}$$

We note that the CSI ratio can be defined between the antenna pairs of the receiver and the transmitter. Therefore, the CSI ratio is $({}_{N_{T_x} * N_{R_x}}C_2 * S)$-dimensional data with respect to the number of transmitter's antennas N_{T_x}, the number of receiver's antennas N_{R_x}, and the number of subcarriers S. Note that ${}_{N_{T_x} * N_{R_x}}C_2$ is a combination. The CSI ratio still retains the other features, canceling out SFO and CFO.



(a) Large quantization error (b) Small quantization error

Fig. 2. Sanitized CSI phases of different subcarriers

B. Sanitization

As described in Sec.III, ADC delays are accumulated linearly with the processing order of the subcarriers. Therefore, a phase change corresponding to the delay appears. If the phase difference per subcarrier due to the conversion delay in the A/D converter between subcarriers is δ, the CSI ratio of the k-th subcarrier is represented as:

$$H_{n,m,k}(f_k,t) = \frac{A_{n,k}}{A_{m,k}} e^{-2j\pi f_k (t+\delta(k-1))}, \tag{7}$$

where $A_{n,k}$ is the complex attenuation of the k-th subcarrier and f_k is the frequency of the k-th subcarrier. Since the ADC delay linearly increases, we can remove it by subtracting the phase difference due to $\delta(k-1)$ for the k-th subcarrier. For simplicity, we put $\theta = f_k(t+\delta(k-1))$. Then, we apply the least square method to estimate the phase delay y_k of k-th subcarrier due to the ADC Delay as below.

$$\begin{aligned} a &= \frac{\sum_{k=1}^{S}(k-\frac{S}{2})(\theta_k - \bar{\theta})}{\sum_{k=1}^{S}(k-\frac{S}{2})^2}, && (8) \\ b &= \bar{\theta} - a * \frac{S}{2}, && (9) \\ y_k &= a * k + b, && (10) \end{aligned}$$

where $\bar{\theta}$ is the average of θ of the all subcarriers. Finally, we subtract y_k from θ_k to obtain the sanitized phase.

C. Quantization Error Mitigation

The quantization error becomes larger as the pilot signal amplitude smaller. Figure 2 shows the sanitized phases of the CSI ratio of different subcarriers recorded at the same time by an antenna pair. The phase of the subcarrier shown in Fig. 2(a) largely fluctuates while the phase of the other subcarrier shown in Fig. 2(b) is much more stable. Therefore, we select s subcarriers with little quantization error to reduce the influence of the quantization error. Although averaging subcarriers may work to mitigate the quantization error, our method can obtain more information by selecting relatively reliable subcarriers.

V. People Counting by Regression

A. Overview

The whole process of the proposed method is shown in Fig. 3. We first apply denoising processes presented in Sec. IV, extract features, and input to an estimation model trained by machine learning.

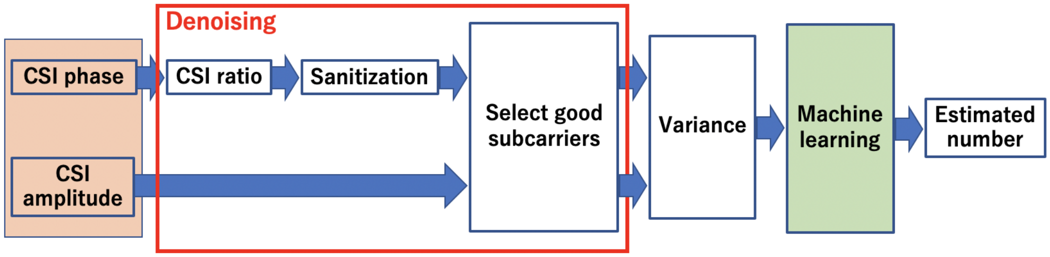


Fig. 3. Overview of system flow

B. Feature Extraction

Different from the existing people counting using Wi-Fi CSI, our goal is to employ mobile devices for Wi-Fi people counting. CSI itself changes depending on the position of the mobile device due to the change of the propagation path. The amplitude and the phase obtained by CSI include effects due to dynamic path changes in the propagation environment and static effects such as reflection from walls and interference in the propagation environment. Since the amplitude and the phase change when the propagation path changes, they are almost constant as long as the propagation environment is stable regardless of the positions of the transmitter and the receiver. This means the variances of the amplitude and the phase increase according to the change of the propagation environment. Obviously, people movement affects the propagation paths, leading to the CSI change. As the number of people increases, the frequency of the path change also increases. Therefore, the variances of the amplitude and the phase become larger with the increase in the number of people. From the above observation, we use the variances of the amplitude and the phase as the location-independent feature. For the calculation of the variances, we empirically set the width of the time window as 10 seconds without any overlap.

C. Regression model

For people counting, we build a regression model. The input to the model is the variances of the CSI phase and amplitude denoised by the processes presented in Sec. IV. We employ Gaussian Process Regression (GPR) after comparison among linear, tree, Support Vector Regression, and GPR.

VI. Evaluation

A. Experiment Settings

In the experiment, we use a Wi-Fi AP (ELECOM WRC-1167GHBK-S) as a transmitter and seven laptops with Intel 5300 Wi-Fi interfaces as receivers. The AP has two antennas, while each receiver has three antennas. We used 802.11n CSI Tool[1] [12] to obtain CSI. Since it provides CSI of 30 subcarriers, the obtained CSI has 180 elements in one packet. CSI was recorded every 10 [ms] by using the ping command from the laptops using 5 [GHz] band. To see the effect of different positions, we collected CSI with multiple laptops

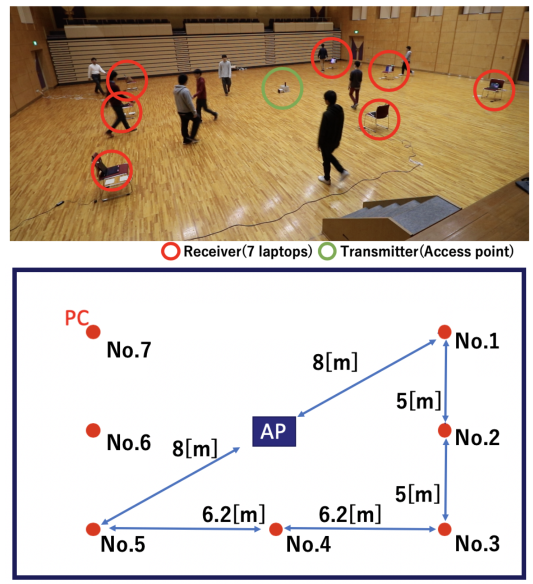


Fig. 4. Experiment Environment

simultaneously, as shown in Fig. 4. We note that we could not use the data of No.3 and No.4 due to hardware failure. We obtained CSI with 0, 3, 6, and 9 people randomly walking in the room. We also note that people followed the random walking pattern where s/he chooses the random direction and walks straight until s/he reaches the wall. When s/he reaches the wall, s/he chooses the next direction randomly and continues walking. To evaluate small changes in positions, we obtained CSI when each laptop was shifted to 0.5 [m] to the left. For each experiment setting, we recorded CSI for 200 [sec]. For building the regression model, we additionally recorded CSI for 200 [sec] at No.1. Since we use the 10-second time window for variances without overlap, as we mentioned in Sec. V-C, we use 20 test cases for each receiver and each scenario.

In the following evaluation, we use the Root Mean Squared Error (RMSE) defined as:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(c_i - \hat{c}_i)^2}, \tag{11}$$

where c_i is the estimated number of people and $\hat{c}_i$ is the ground

[1]https://dhalperi.github.io/linux-80211n-csitool/installation.html

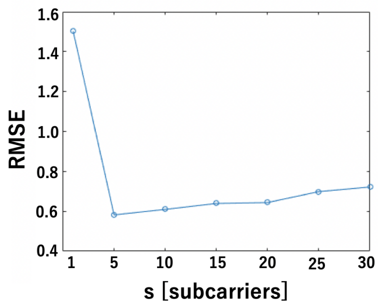


Fig. 5. # of subcarriers vs. RMSE

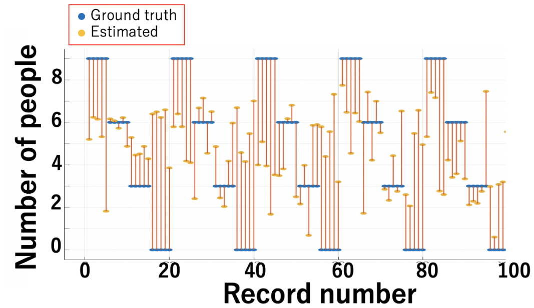


Fig. 6. Example of people counting without denoising

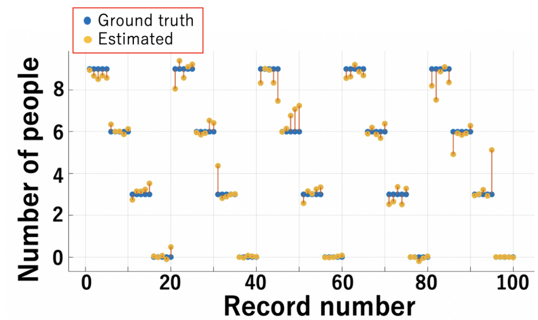


Fig. 7. Example of people counting with denoising

truth in the i-th test case.

B. Results

1) Effect of Subcarrier Selection: We change the parameter s to see the effect of the subcarrier selection. We note that the number of variances used as input greatly depends on s. For example, 20 variances are available for $s = 1$ while 600 variances are used for $s = 30$. We build different regression models for each s. Therefore, there is a trade-off between s and the performance.

Figure 5 shows RMSE for different s. We see that RMSE gradually increases with an increase of s. This result indicates that the effect of the quantization error gradually increases for $s > 5$ and $s = 5$ is the best trade-off between the quantization error and the number of variances. RMSE for $s = 1$ is the worst because the number of variances used as input to the regression model is tiny. From the result, we use $s = 5$ in the following evaluation.

2) Effect of Denoising: To see the effect of our denoising scheme, we compared the results after applying each denoising. First, we estimated the number of people using the CSI phase and amplitude variances without any denoising. First, we found the RMSE without denoising is 3.7. After taking the CSI ratio to remove SFO and CFO, RMSE significantly decreased to 0.99. Then, the sanitization slightly improved RMSE to 0.72. Finally, we selected top-5 subcarriers with small quantization errors, which results in RMSE of 0.49.

Figures 6 and 7 show examples of the estimation results for 100 test cases without denoising and with denoising, respectively. The result without denoising has a large error, which seems overfitted to output the average number of people. On the other hand, the estimation result is very close to the ground truth by applying our denoising scheme. From the result, we confirmed that our denoising scheme works well.

3) Effectiveness of CSI Variance: To see the effect of the difference in positions of receivers, we changed the receiver positions while using the same regression model. The bar colors in Figs. 8 and 9 indicate the distance from the original receiver positions. As shown in Fig. 8, even 0.5 [m] difference degrades the performance if we use the CSI raw data as a feature. We also see RMSE varies with the change of the positions. This is because the regression model learns static environment factors which are location dependent. On the other hand, as shown in Fig. 9, RMSE of regression by the CSI variance is much more stable against different positions although we see small errors less than 2.0.

4) Combination of Amplitude and Phase: Finally, to see the effect of the amplitude and the phase as features, we compared the different combinations of them. RMSE is 0.91 for the phase variance, while RMSE is 0.59 for the amplitude variance. These are worse than RMSE of 0.49 using both of the variances. From the result, we have confirmed the effectiveness of the combination of the phase and the amplitude obtained by CSI.

VII. Conclusion

In this paper, we proposed a method of towards people counting using mobile devices. We design a denoising scheme based on the noise model for the phase offsets, ADC delay.

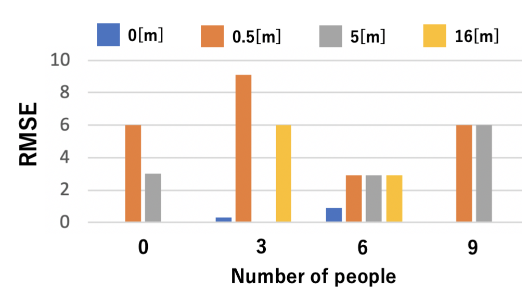


Fig. 8. Regression using CSI raw data

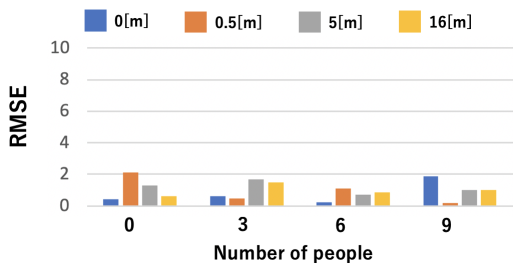


Fig. 9. Regression using CSI variance

Also, the quantization error is mitigated by the subcarrier selection. To solve the challenge of device mobility, we proposed the variance as a location-independent feature. The experiment results showed the effectiveness of our method, achieving RMSE of 0.49 for 0-9 people.

Our future work includes real experiments in public spaces, such as shopping malls. Also, we are planning to clarify the definition of the target area for open spaces by collecting data with accurate tracking devices such as LiDAR. Furthermore, we design a method to combine estimation results from multiple mobile devices.

Acknowledgments

This work was partly supported by JST, PRESTO Grant Number JPMJPR1932, Japan.

References

[1] A. C. Davies, J. H. Yin, and S. A. Velastin, "Crowd monitoring using image processing," *Electronics & Communication Engineering Journal*, vol. 7, no. 1, pp. 37–47, 1995.

[2] A. Di Luzio, A. Mei, and J. Stefa, "Mind your probes: De-anonymization of large crowds through smartphone wifi probe requests," in *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*. IEEE, 2016, pp. 1–9.

[3] H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. Spanos, "Freecount: Device-free crowd counting with commodity wifi," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[4] W. Xi, J. Zhao, X.-Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang, "Electronic frog eye: Counting crowd using wifi," in *Infocom, 2014 proceedings IEEE*. IEEE, 2014, pp. 361–369.

[5] S. Di Domenico, M. De Sanctis, E. Cianca, and G. Bianchi, "A trained-once crowd counting method using differential wifi channel state information," in *Proceedings of the 3rd International on Workshop on Physical Analytics*. ACM, 2016, pp. 37–42.

[6] Y.-K. Cheng and R. Y. Chang, "Device-free indoor people counting using wi-fi channel state information for internet of things," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[7] H. Li, E. C. Chan, X. Guo, J. Xiao, K. Wu, and L. M. Ni, "Wi-counter: smartphone-based people counter using crowdsourced wi-fi signal data," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 4, pp. 442–452, 2015.

[8] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, p. 121, 2019.

[9] J. Yuan and M. Torlak, "Joint cfo and sfo estimator for ofdm receiver using common reference frequency," *IEEE Transactions on Broadcasting*, vol. 62, no. 1, pp. 141–149, 2016.

[10] X. Tang and G. Fu, "A ml-based high-accuracy estimation of sampling and carrier frequency offsets for ofdm systems," in *Computer Engineering and Technology: 19th CCF Conference, NCCET 2015, Hefei, China, October 18-20, 2015, Revised Selected Papers*, vol. 592. Springer, 2016, p. 85.

[11] T.-C. Tai, K. C.-J. Lin, and Y.-C. Tseng, "Toward reliable localization by unequal aoa tracking," in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2019, pp. 444–456.

[12] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 1, pp. 53–53, Jan. 2011. [Online]. Available: http://doi.acm.org/10.1145/1925861.1925870