Wi-Count : Passing People Counting with COTS WiFi Devices

Conference Paper · August 2018

DOI: 10.1109/ICCCN.2018.8487420

Yanni Yang
The Hong Kong Polytechnic University

Xiulong Liu
The Hong Kong Polytechnic University

Wi-Count: Passing People Counting with COTS WiFi Devices

Yanni Yang*, Jiannong Cao*, Xuefeng Liu†, Xiulong Liu*
*The Hong Kong Polytechnic University, Hong Kong, China.
Email: csynyang@comp.polyu.edu.hk, csjcao@comp.polyu.edu.hk, csxliu@comp.polyu.edu.hk
†Huazhong University of Science and Technology, Wuhan, China. Email: csxfliu@gmail.com

***Abstract*—People counting provides valuable information on population mobility and human dynamics, which plays a critical role for intelligent crowd control and retail management. Recently, people counting has been achieved via radio-frequency signals as human presence can influence the propagation of wireless signals, from which the information of the moving crowd can be extracted. However, most of the existing studies using wireless signals only apply to the scenario when people keep moving all the time. Besides, they require labour-intensive training phase for building the counting model.**

In the Wi-Count system, we take another approach, which is to count the people passing by the doorway with COTS WiFi devices. It can not only detect the passing direction, but also identify the number of people even when multiple persons pass by concurrently without regulating passing behavior and pre-trained counting model. The passing direction is recognized by modeling the effects of the bi-directional passing behavior on the phase difference of WiFi signals. In addition, the number of passing people is obtained through an enhanced signal separation algorithm for providing precise counting result. Extensive experiments show the average accuracy on passing direction detection and passing people counting are about 95% and 92% respectively.

I. Introduction

Driven by the developing human sensing technologies, the environment becomes more intelligent for sensing and observing human presence, which helps to ensure sufficient resource allocation and customized services. People counting provides essential information for various applications and services, *e.g.*, crowd control for big events and marketing investigation for retail sales. In particular, indoor activities require precise people counting for analyzing the crowd behavior, such as the visiting rush hours, so that they can discover more potentials in the marketing and drive the business to grow.

Current practices towards people counting are mainly realized in two ways, which are (1) counting people in the certain area, and (2) counting people passing by the doorway. On top of the above two means, existing techniques make use of infrared sensors, cameras, devices and radio-frequency signals for people counting. Infrared sensors [1], [2] deployed at the doorway count the entering or exiting people based on the assumption that people are moving through one by one under certain time interval. So, it cannot count multiple people passing by freely. Vision-based approach [3], [4] counts people by image object detection. While its performance can be weakened by object overlapping, poor lighting conditions and dead zones. It also raises the privacy concern among the massive customers for capturing their figures. For the device-based approach [5]–[7], in which devices (*e.g.*, mobile phones or RFID tags) shall be distributed or carried, requires active participation from the crowd. This could result in the reluctance among people for being counted. Recent advances in wireless human sensing witness the potential for using wireless signals, such as WiFi [8]–[10], RFID [11] and UWB radar [12] for human sensing and people counting. However, these works are based on the premise that all the human objects would keep moving all the time, which is unable to be met in practice. Thus, the counting result can only provide a rough estimate on the number of people, and it cannot be applied to scenarios where most of the people do not move frequently. In addition, they need extensive training phase for building the counting model which can be difficult for being adaptive to different environments.

In this paper, we propose the Wi-Count, to realize people counting in a non-intrusive, low-cost and accurate way by counting the people passing by the doorway: (1) **Non-intrusive**: Wi-Count leverages the effects of the passing behavior on the propagation of wireless signals to detect human presence, so it does not require people to carry any devices. It allows the human object to pass by freely without active cooperation for the counting process. (2) **Low-cost**: Wi-Count saves the cost on devices and labour resources. On the one hand, it only relies on the existing indoor WiFi infrastructure. On the other hand, it does not require labour-intensive training phase for building the people counting model. (3) **Accurate**: Wi-Count is not only capable of precisely detecting the bi-directional passing behavior for the single person, but also counting the number of people when multiple persons pass by the doorway at the same time.

Counting people passing by the doorway can be achieved by analyzing the phase information in the received WiFi signals. When people passing by, the human body, as a reflector, could influence the propagation of wireless signals from the transmitter to the receiver. To detect the binary moving direction, *i.e.*, entering and exiting, we model the passing behavior with respect to the change of phase information. The phase information can be extracted from the Channel State Information (CSI), which is available on many commercial WiFi devices. By performing theoretical analysis on the change of phase information, we find that there are distinctive patterns on the phase difference series for the entering and exiting behavior.

To identify the number of passing people, we treat each person as independent reflectors, as they are different in shape and walking habit. Then, the Independent Component Analysis (ICA) algorithm is applied on the phase difference series of all the subcarriers to reveal how many walking components are involved, which indicate the number of people.

However, to achieve precise counting results, several challenges are remained to be solved. First of all, the phase information extracted from the COTS WiFi devices suffers from different sources of noises, including phase difference ambiguity and random noises. This makes the observed passing pattern unclear and inconsistent for further analysis. Second, the human body is not a fixed-shape reflector while walking. It can result in fluctuations in the received phase information, which causing troubles for extracting the pattern for passing direction detection. Furthermore, the phase information on some subcarriers can be interfered by adjacent subcarries, leading to an ineffective result for using the signal separation algorithm to identify the number of multiple passing people.

To address the above challenges, calibration on the phase difference series is first performed to obtain clear phase information. To eliminate the phase difference ambiguity, we apply clustering algorithm on the phase difference time series to recover the original phase difference information. To remove the adverse effects of random noises and signal fluctuations for passing direction detection, we apply Savitzky-Golay filter [13] on the phase difference series for noise reduction. In order to enhance the performance of the signal separation algorithm for identifying the number of passing people, we transform the phase difference series of all the subcarriers into an input matrix. Afterwards, the input matrix is preprocessed through Principal Component Analysis (PCA) to remove the interference and correlation among adjacent subcarriers for achieving a more precise counting result.

Extensive experiments show the average accuracy on passing direction detection and passing people counting are about 95% and 92% respectively.

The main contributions of our work are as follows:

- We propose Wi-Count, a non-intrusive, low-cost and accurate approach for people counting with the COTS WiFi devices. It only relies on the existing WiFi infrastructure and counts the human object without constraints on the moving behavior and labour-intensive model training.
- We present a model for detecting the bi-directional moving behavior based on the phase difference of WiFi signal. The moving direction can be recognized from the pattern of phase difference series.
- We apply source separation techniques for counting the multiple persons passing by concurrently, so that we can provide more accurate result for people counting and detailed information for observing the crowd behavior.

II. Related Work

There are various practices and studies into the problem of indoor people counting. Current works can be classified into four categories: (1) infrared-based approach; (2) vision-based approach; (3) device-based approach and (4) RF-based approach. In this section, we discuss the existing work on people counting with respect to the above four categories.

A. Infrared-based Approach

Infrared-based approach counts the people by detecting whether the light beam is blocked by people. Generally, multiple sets of infrared sensors are deployed for detecting the entering or existing directions passing by the door [1], [2]. However, they are only applicable to the single-person passing scenario, which means that they cannot tell the exact number of people when two or more people come across the beam at the same time. Thus, barriers are usually set around the door to allow only one person passing at a time, leading to extra deployment cost and inconvenience for the people.

B. Vision-based Approach

Vision-based approach uses pattern recognition techniques, *e.g.*, face [3] and head-shoulder detection [4], [14], to count human objects. Much work has been done to improve computer vision techniques for human detection [15]. Since people can dress up in different styles, many studies try to improve the robustness of the human detection algorithms with the help of machine learning. However, the dead zones and object overlapping still lead to the ineffectiveness for vision-based approach to get accurate counting result. In addition, people are reluctant to be captured by cameras everywhere, which could intrude their privacy without being notified.

C. Device-based Approach

Device-based approach realizes people counting by the means of spreading devices or sensors in the crowd, *e.g.*, RFID tags [16] or mobile phones [5]–[7]. Audio signals, Bluetooth and WiFi connection information are all used for tracking and counting the crowd in indoor and outdoor environment. On the one hand, the device-based approach raises the cost on people counting by allocating large number of devices. On the other hand, it requires people's active participation by operating the devices so that they can be detected, such as open the Bluetooth link or run a specific app.

D. RF-based Approach

Recently, RF-based human sensing has seen great potentials and possibilities on various applications, including localization, activity and gesture recognition and vital sign monitoring. Many kinds of wireless signals, such as WiFi [8]–[10], RFID [11] and UWB radar [12], have been leveraged for people counting. Since the presence of human object can affect the propagation of the wireless signals in the air, human sensing can be realized without attaching any sensor on the body. Due to the multipath effects of wireless signals, the moving object can still influence the propagation of the wireless signals even if they are in the none line-of-sight areas.

Existing work using RF signals for people counting leverages the RSS or CSI of the wireless signals to learn the relationship between the number of moving people and the variation in the wireless signals. Researchers in [17] deploy wireless sensor networks and estimate the rough crowd density based on the RSS with clustering algorithm. In [8], they count the people walking between the WiFi transmitter and receiver based on the RSS measurements. [9] finds out the monotonic relationship between the CSI measurements and the number of moving people in the certain area and count the crowd with Grey Verhulst Model. [18] proposes to derive the number of people through the statistical distribution of CSI measurements and applies semi-supervised regression algorithm to obtain the counting result. In [11], dozens of RFID tags are attached to the wall for counting the moving people.

The common limitation of the above methods is that they can only work when the human objects in the certain area keep moving, so the estimated results could be inaccurate when people have less mobility. Furthermore, they require extensive training phase for building the counting model which requires calibration for being adaptive to different indoor environments. Our work also utilizes the wireless signals, *i.e.*, WiFi, but from a different angle, which is to count the people passing by the doorway and convert this information as the number of people inside the certain areas. We can detect the bi-directional passing behavior and identify the exact number of passing human objects, so that to achieve precise counting result and obtain more information on human dynamics.

III. Model for Passing People Counting Based on WiFi Phase Difference

This section first introduces preliminary knowledge for the Channel State Information of WiFi signals, and then performs theoretical analysis on the effects of the passing behavior to the propagation of wireless signals. Afterwards, a model is built for detecting the bi-directional passing behavior based on the phase difference time series.

A. Preliminaries on CSI

In modern wireless network, the whole network spectrum is divided into orthogonal subcarriers using Orthogonal Frequency Division Multiplexing (OFDM). The PHY layer information, Channel State Information, underlying in each subcarrier reflects the linear combined effects, *e.g.*, reflection and scattering of the wireless signals along different propagation paths. Thus, the CSI can be represented as follows [19]:

$$H(f,t) = \sum_{i=1}^{n} a_i(f,t) \cdot e^{-j\psi(f,t)}, \tag{1}$$

where f denotes the central frequency of each subcarrier, n is the number of propagation paths. $|a_i(f,t)|$ and $\psi(f,t)$ represent the amplitude and phase values respectively. For m subcarriers, the CSI matrix for a given period is

$$\mathbf{H} = [H(f_1,t), H(f_2,t), ..., H(f_m,t)]. \tag{2}$$

In our work, we use the phase information for passing people counting. The phase information $\psi(f,t)$ captured by the COTS network interface card, *e.g.*, Intel 5300, contains the timing and phase offset [20], [21]. As a result, the measured $\widehat{\psi_j}$ for the subcarrier j can be expressed as

$$\widehat{\psi_j} = \psi_j + 2\pi \cdot f_j \cdot \alpha_j + \beta_j + Z, \tag{3}$$

in which ψ_j is the real phase, α_j and β_j are the timing and phase offset caused by Carrier Frequency Offset (CFO), Sampling Frequency Offset (SFO) and Packet Detection Delay (PDD). β_j is a constant value for the same NIC and Z denotes the minor random noises in the phase values. The phase errors in $\widehat{\psi_j}$ make it difficult to observe the real effects of the passing behavior from the phase information, so we need to remove those phase deviations for further use.

B. Model for Counting Passing People

In terms of the propagation of wireless signals, there are different multipaths traveling from the transmitter (*Tx*) to the receiver (*Rx*) except for the Line-of-Sight (*LoS*) path, due to the presence of multiple reflectors in the environment. As illustrated in Fig. 1(a), there are *LoS* path and multipaths reflected by the Reflector 1 and Reflector 2. For the i^{th} path, the phase shift can be represented as

$$\psi_i = \{l_i/\lambda\} \ mod \ 2\pi, \tag{4}$$

where l_i is the length of propagation path, λ is the wavelength of the wireless signal. Suppose that Reflector 2 is a moving object, then the received wireless signals consist of static propagation paths (P_s) and dynamic paths (P_d). Therefore, the overall phase change ψ in the received signals is the combination of static and dynamic shift of the phase values.

$$\psi = \{\frac{\sum_{i\in P_s} l_s + \sum_{i\in P_d} l_d}{\lambda}\} \ mod \ 2\pi \tag{5}$$

To detect the bi-directional passing behavior of the human object, we employ one transmitting antenna and two receiving antennas, which are available on commercial WiFi devices, and use the phase difference between the two receiving antennas to indicate the passing direction. As shown in Fig. 1(b), the transmitting and receiving antennas are horizontally displayed, and the distance between *Rx*$_1$ and *Rx*$_2$ is d. Then, the phase difference ψ_{21} between *Rx*$_1$ and *Rx*$_2$ can be formulated as

$$\begin{aligned}\psi_{21} = \psi_2 - \psi_1 &= \{\frac{L_{s_2} + L_{d_2}}{\lambda} - \frac{L_{s_1} + L_{d_1}}{\lambda}\} \ mod \ 2\pi \\ &= \{\frac{(L_{s_2} - L_{s_1}) + (L_{d_2} - L_{d_1})}{\lambda}\} \ mod \ 2\pi \\ &= \{\frac{L_0 + d \cdot sin\theta}{\lambda}\} \ mod \ 2\pi,\end{aligned} \tag{6}$$

where $L_s = \sum_{i\in P_s} l_s$, $L_d = \sum_{i\in P_d} l_d$. L_0 is the difference between L_{s_1} and L_{s_2}, representing the static phase shift caused by the static propagation paths, and the difference for the length of dynamic paths between the two receivers is mainly induced by the moving targets, which is $d \cdot sin\theta$, where θ is the angle of arrival of the wireless signals reflected by the human object to the two receiving antennas. However, the measured

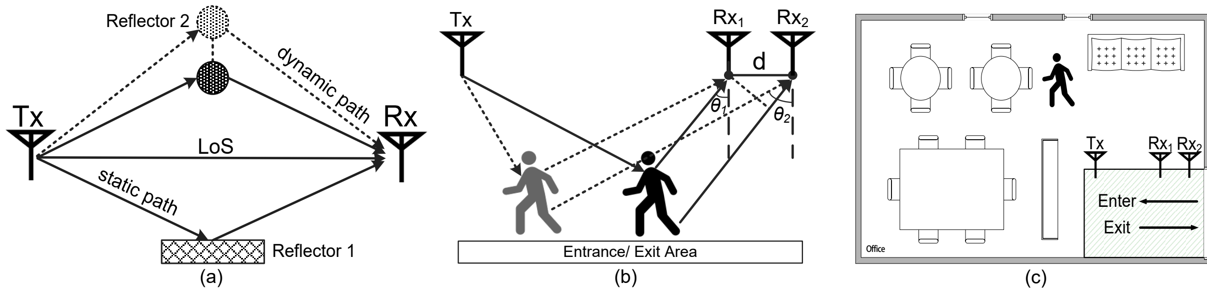

Fig. 1. (a) Propagation of wireless signals in indoor environment. (b) Model of applying phase difference for passing direction detection (c) Deployment of WiFi transmitter and receivers.

phase information includes several sources of errors, so the reported phase difference $\widehat{\psi_{21}}$ is formulated as

$$\begin{aligned}\widehat{\psi_{21}} &= \widehat{\psi_2} - \widehat{\psi_1} \\ &= \psi_{21} + 2\pi f \cdot \Delta\alpha + \Delta\beta + \Delta Z.\end{aligned} \tag{7}$$

For the above equation, $\Delta\alpha$ equals to zero, as the two receiving antennas on the same NIC card use the same clock and same down-converter frequency. So, $2\pi f \cdot \Delta\alpha$ can be removed, then $\widehat{\psi_{21}}$ can be shortened as

$$\begin{aligned}\widehat{\psi_{21}} &= \psi_{21} + \Delta\beta + \Delta Z \\ &= \{\frac{L_0 + d \cdot sin\theta}{\lambda} + \Delta\beta + \Delta Z\} \ mod \ 2\pi.\end{aligned} \tag{8}$$

In Eq. (8), $\Delta\beta$ is a constant value for the same NIC and ΔZ is the random noise with minor influence on the phase values. So, the change in the phase difference mainly resides in $d \cdot sin\theta$. When the human object moves towards left, *i.e.*, entering the area, θ will increase within the range $(0, \pi/2)$. If we let d approximate to the wavelength of the wireless signal, then the phase difference between Rx_1 and Rx_2 will monotonically increase as $d \cdot sin\theta$ goes up. Conversely, θ will decrease from $\pi/2$ to 0 when the human object moves to the right, *i.e.* exiting the area, making the phase difference decline. Therefore, we can identify the bi-directional passing behavior via the increasing and decreasing trend in the phase difference.

To verify the proposed model, we deploy WiFi devices in indoor environment as shown in Fig. 1(c). The transmitter and receivers are displaced horizontally in the entrance/exit area, with one-meter distance apart. The two receiving antennas are placed $12cm$ apart from each other since the wavelength of $2.4GHz$ WiFi signals is around $12.5cm$. The receiving antennas are equipped with directional antennas, orienting to the entrance/exit area, to avoid the disturbance of other moving objects on the received signals in the environment. We plot the phase difference when the person enters and exits the room in Fig. 2(a)-(b). The presence of the four separate time series is caused by the four-way phase ambiguity, which will be removed later. Despite of this, we can observe that there is an increasing trend in the phase difference series for the entering behavior, and a decreasing trend for the exiting behavior.

IV. Overview of Wi-Count

The proposed system, Wi-Count, consists of three modules, *i.e.*, phase collection and calibration, bi-directional passing behavior recognition and identification of the number of people passing by. The overview of the system is depicted in Fig. 3.

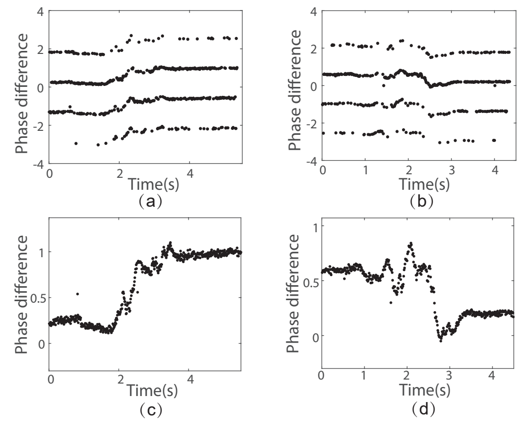

Fig. 2. Raw phase difference for (a) entering the door and (b) exiting the door. Phase difference after phase ambiguity removal for (c) entering the door and (d) exiting the door.

In the first module, the phase information is extracted from the CSI measurements along with the phase difference time series between two receiving antennas. Then, the phase ambiguity will be removed with cluster-based algorithm, *i.e.*, k-means, for the recovery of the original phase difference series, which will be articulated in the next section.

For the second module, the aim is to recognize the entering and existing direction. As the phase difference involves random noises and fluctuations caused by the periodic moving legs, it is first smoothed to get the general trend for the pattern extraction. For the settings in Fig. 1(c), if there is an increasing trend in the phase difference series, then the human object is entering the room. By contrast, the human object is exiting the room if the pattern presents a decreasing trend. We calculate the derivatives of the smoothed phase difference and detect the presence of local maximum or minimum to recognize the increasing and decreasing pattern.

To obtain an accurate counting result, we need to consider the multi-person scenario and count the number of people when multiple persons enter or exit the room concurrently. Multi-person scenario is quite common in shopping malls or exhibitions, where people would like to hang out with their friends. Capturing this information also provides opportunities

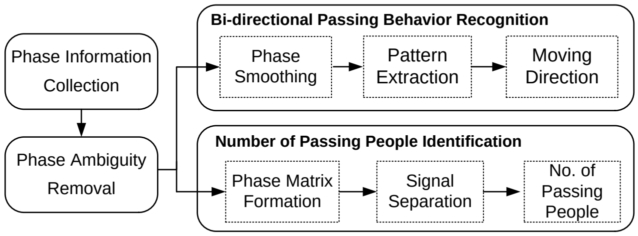

Fig. 3. System overview of Wi-Count

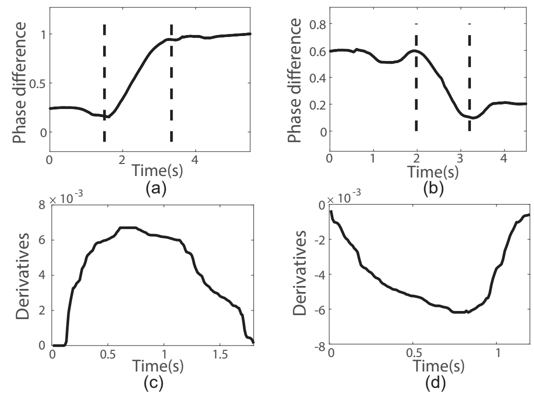

Fig. 4. Phase difference series after denoising for (a) entering behavior and (b) exiting behavior. Derivatives on the phase difference series of (c) entering segment and (d) exiting segment

for group detection and human dynamics analysis. To count the number of passing people, we treat each of the human objects passing by the doorway together as an individual component that affects the propagation of the wireless signals. Then, the overall received signals are the linear combination of the effects from different components. To provide enough input source signals, the phase difference series from all the subcarriers are regarded as different observations and leveraged for the formation of the input matrix. Then, the Independent Component Analysis (ICA) is applied on the input matrix to separate each component.

V. System Design

In this section, we first preprocess the raw phase difference time series and then detect the moving direction with the calibrated phase difference time series based on the proposed model. At last, we perform multi-person passing identification to count the concurrent passing people.

A. Phase Ambiguity Removal

For existing COTS wireless network interface card, the four-way phase ambiguity causes the real phase difference to be θ, $\theta + \pi/2$, $\theta - \pi/2$ or $\theta - \pi$ for $2.4GHz$ wireless signals [22] ($5GHz$ wireless signals have two-way phase ambiguity). In Fig. 2(a)-(b), the real phase difference is separated into four groups with $\pi/2$ spacing. One way to retrieve the real phase difference is to compare the difference between two consecutive phase values. If the difference is around $\pi/2$ or π, then it is expected to experience a phase shift and we can add the corresponding shift to the current phase to retrieve the original phase value. However, there are many outliers and noises in the phase difference, making the above approach ineffective in dealing with the phase ambiguity.

Here, we apply clustering algorithm to address the above problem. We gather certain amount of phase difference samples, for example, 100 samples and apply k-means clustering algorithm on them. For the four-way phase ambiguity, there would be four clusters with the spacing of each cluster's centroid to be $\pi/2$. Then, we sort the four centroids and choose the highest one as the baseline so that the samples in the other three clusters can add $\pi/2$, π and $3\pi/2$ respectively; and then they can be integrated into a single time series. The recovered phase difference series after phase ambiguity removal are shown in Fig. 2(c)-(d).

B. Bi-directional Passing Behavior Recognition

In Fig. 2(c)-(d), the raw series of the phase difference exhibit specific patterns for entering and exiting behavior. However, the human body is not a flat reflector with fixed shape while walking; leg movement also affects the wireless signals periodically, leading to fluctuations in the received signals. In order to extract the increasing and decreasing trend from the phase difference, we first apply Savitzky-Golay filter [13] on the phase difference. Savitzky-Golay filter is based on the least-squares polynomial approximation, which can smooth noises and maintain the contour of the time series. Here, we set the polynomial order as 3 and the length of frame as 50. The filtered phase difference series are illustrated in Fig. 4(a)-(b).

Then, a threshold is set to segment out the phase difference series affected by the moving behavior. As in Fig. 4(a)-(b), the series between the two vertical dashed lines are the affected series. To detect the increasing and decreasing trend from the phase difference series, we first calculate the first derivative of each sample points and smooth the derivatives with median filter. Then, the presence of peak or valley in the first derivative will be detected. Fig. 4(c)-(d) shows the smoothed derivatives of the affected series. The peak in the Fig. 4(c) corresponds to the increasing trend, while the valley in Fig. 4(d) reveals the decreasing trend. After finding out the specific pattern in the phase difference, we can detect whether the human object is entering or exiting the room.

C. Multi-Person Passing Identification

When multiple persons enter the room together, there is a similar pattern, an increasing trend as the single-person scenario. However, to get accurate counting result, we need to identify how many people are passing by the door at the same time. In fact, multi-person scenario is quite common in shopping malls and other indoor places where people can hang out with their friends. By identifying the multi-person passing behavior, it can also help to observe group behavior and human dynamics. Thus, after recognizing the passing direction, we need to figure out the exact number of passing people.

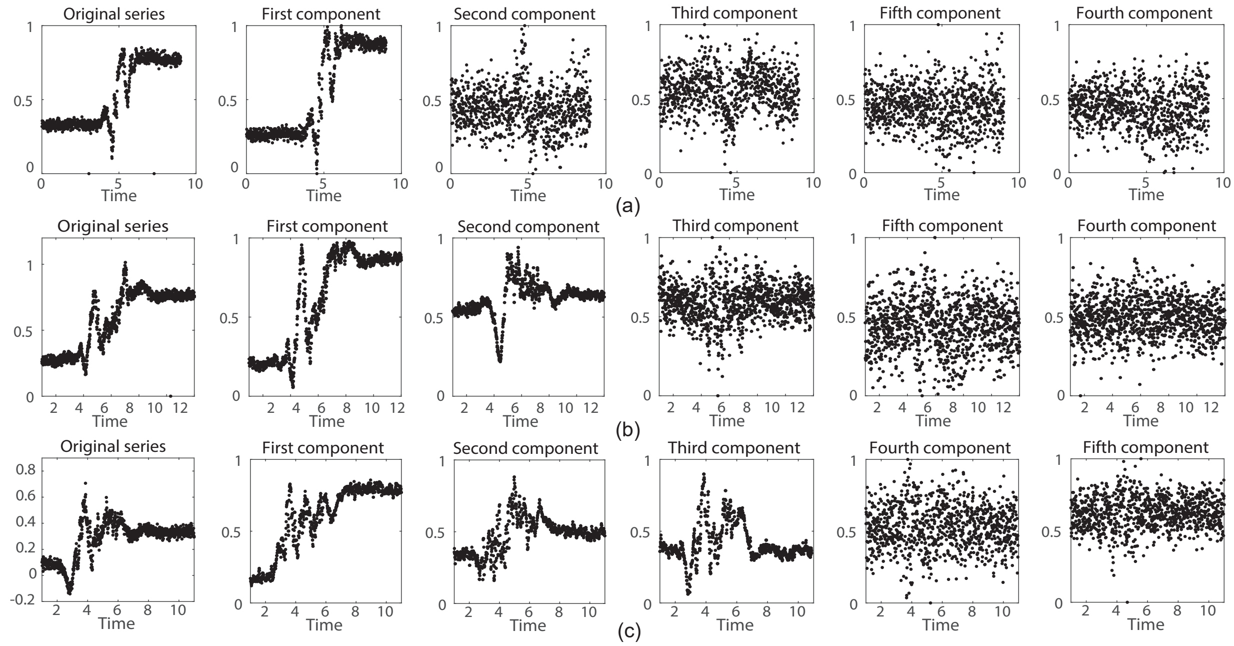

Fig. 5. Separated components when (a) one person enters the door; (b) two persons enter the door and (c) three persons enter the door.

Theoretically, we can treat each human object as independent reflectors to the wireless signals. The overall received wireless signals are the linear combination of different sources of reflected signals. Although some people would walk closely with each other, their moving behavior can still result in distinctive effects on the reflected signals as human objects are different in shape and walking habits. Previous work has leveraged the gait difference for human identification using wireless signals [23], [24], indicating that wireless signals can reveal different walking patterns. Therefore, we apply source separation techniques on the phase difference series to see how many separated sources are affecting the wireless signals. Here, we decompose the phase difference series with Independent Component Analysis (ICA) [25].

For the source separation problem, its target is to separate the mixed signals into individual sources. To separate n sources from the mixed signals, there must be at least n observations as the input. Denote $\mathbf{x} = [x_1, x_2, ..., x_n]$ as the input and $\mathbf{s} = [s_1, s_2, ..., s_m]$ as the source signals, then the mixed input signals can be represented as the linear combination of the source signals: $\mathbf{x} = \mathbf{A}\mathbf{s}$. A new matrix $\mathbf{W}$ is created to represent the estimated source signals as follows: $\widehat{\mathbf{s}} = \mathbf{W}\mathbf{x}, \mathbf{W} = \mathbf{A}^{-1}$. Then, by estimating $\mathbf{W}$, we can obtain the separated source signals $\widehat{\mathbf{s}}$. However, we only have one set of transceivers to get the phase difference series, while there would be two or more people passing by the door. To get enough source signals, we use the phase difference of multiple subcarriers as different source signals. For the Intel 5300 NIC, it can report 30 subcarriers. We assume that no more than five people could pass by concurrently due to the space limitation, so we only need to retrieve five components as the outputs of ICA. We then transform the phase difference series of all the 30 subcarriers into an input matrix with size $30 * k$ (k is the number of sample points). The input matrix $\mathbf{x}$ is:

$$\mathbf{x} = \begin{bmatrix} \Delta\psi_{1,1} & \Delta\psi_{1,2} & ... & \Delta\psi_{1,k} \\ \Delta\psi_{2,1} & \Delta\psi_{2,2} & ... & \Delta\psi_{2,k} \\ ... & ... & ... & ... \\ \Delta\psi_{30,1} & \Delta\psi_{30,2} & ... & \Delta\psi_{30,k} \end{bmatrix} \tag{9}$$

Intuitively, the more sources of observation are captured, the better separation result can be attained. However, due to the hardware imperfection in the COTS WiFi devices, some of the subcarriers contain noises and interferences from adjacent carriers. A possible way is to apply smoothing algorithms to denoise the input matrix, but the side effect is that the minor changes in the signals caused by multiple persons' moving behavior will be removed as well. Besides, since the carrier frequency difference between two consecutive subcarriers is quite small, we need to consider the effect of the correlation among different subcarriers on the result of ICA.

To enhance the performance of ICA, we leverage Principal Component Analysis to discard the dimensions with less dominance by analyzing the eigenspace [26]. Furthermore, we can also remove the pair-wise dependency among subcarriers. Denote the eigenvectors of the covariance matrix of the input matrix as $\mathbf{R}$, so that $\mathbf{R}^T(\mathbf{x} * \mathbf{x}^T)\mathbf{R} = \Lambda$, where Λ is the diagonal matrix of eigenvalues. The smallest q eigenvalues indicate the noise space $\mathbf{E_n}$, and we discard the eigenvectors in $\mathbf{R_n}$ and only employ the eigenvectors $\mathbf{R_s}$ in the signal space $\mathbf{E_s}$ as the input for ICA. The eigenvectors whose eigenvalues account for less than 5% of the sum of all eigenvalues are discarded. Then, the estimated source signals can be derived as: $\widehat{\mathbf{s}} = \mathbf{W} * \mathbf{R_s}^T$.

Here, we apply the FastICA algorithm which can perform ICA efficiently on the preprocessed input matrix [27]. For the estimation of $\mathbf{W}$, FastICA employs the approximation of negentropy to maximize the nongaussianity, which is a

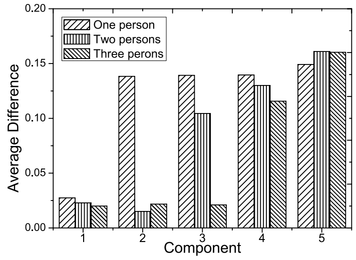

Fig. 6. Average difference among all the difference between consecutive values for difference number of passing people

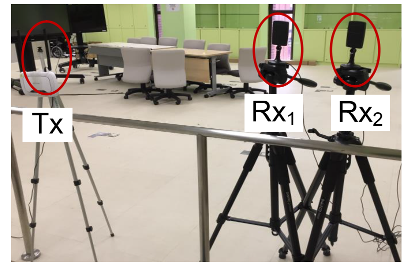

Fig. 7. Wi-Count deployment in real environment

measure of independency among different sources. For the input parameter of the FastICA algorithm, we set the number of separated components to be five owing to the upper limit of the number of people. As an example, we show three sets of separated sources in Fig. 5 with different numbers of people entering the room. For the raw phase difference series, there are more variations in the signals for more passing people. For one-person case (Fig. 5(a)), the first component reflects the entering behavior, while others are all noises. For the two-person case (Fig. 5(b)), the first component is the general trend of entering behavior, the second one reveals the walking behavior of the second human object. Similarly, when there are three human objects, the separated sources will have three effective components representing the entering and walking behavior, as shown in Fig. 5(c). Therefore, we can count the number of passing people by detecting effective components from the separated components.

The way to determine how many effective components exist comes from the observation that the points in the effective component are consistent with each other, while the random noises are distributed out-of-order. So, the difference between two consecutive points of effective components is smaller than that of the noises. We add all the difference between two consecutive points and obtain the average value. The average differences for effective components and random noises are shown in Fig. 6. It shows that the average difference of effective components is much smaller than that of random noises. As there is always one effective component, we use it as the benchmark (D_{base}), if the difference of other components (i) meets the requirement $D_i > \alpha \cdot D_{base}$, then the components would be regarded as random noises. Here, the parameter α is set to be 3 empirically. Then, the number of people is the number of the effective components.

VI. Evaluation

In this section, we first introduce the deployment of Wi-Count and the data collection phase. Then the evaluation metrics are given for passing direction detection and number of passing people identification, and the counting performance is evaluated under different settings and scenarios.

A. Deployment and Data Collection

To evaluate the performance of our approach, we implement our system with commercial off-the-shelf devices, *i.e.*, a TP-Link wireless router, a laptop equipped with Intel 5300 NIC. There are three antennas on the Intel 5300 NIC, we only use the first two as the receivers and the antenna on the TP-Link wireless router is regarded as the transmitter. They are horizontally placed around the doorway with 1-meter distance. The CSI information is extracted through the CSI tool [28] which modifies the firmware under Linux system for exposing the CSI information.

We collect phase difference series under the above deployment in real environments. Figure 7 shows examples of the deployment environment of our system. We recruit 8 volunteers, including 5 males and 3 females, and make them enter and exit the rooms back and forth. They are required to move with different walking speeds, *i.e.*, slow, normal and fast walking. For the multi-person scenario, they enter and exit the room in two group patterns, *i.e.*, one after another or side by side. We spent around 20 days distributed in three months for the data collection phase. While collecting the phase information, the sampling rate is set to be $100p/s$, $200p/s$ and $400p/s$ respectively.

B. Evaluation Metrics

To evaluate the performance on passing direction detection and number of passing people identification, the following metrics are defined accordingly.

Evaluation on Passing Direction Detection: There are two possible results for detecting the passing direction, which are entering and exiting. We define the *accuracy of passing direction detection* as the ratio between all the correct instances and the total number of instances for each case.

Evaluation on Multi-Person Passing Identification: To evaluate the performance of the source separation method on counting the number of people passing by, we define the *error of multi-person counting* as the difference between the estimated number of people and the real number of people, and the *accuracy of multi-person counting* as the percentage of the correctly identified instances over the total number of instances for different number of passing people.

C. Passing Direction Detection

The performance of passing direction detection is evaluated with different scenarios, *i.e.*, walking speed, sampling rate and number of people under two environments. In the following, the accuracy on passing direction detection will be discussed in detail in terms of the impact of walking speed, sampling rate and the number of people.

1) Impact of walking speed: Volunteers are asked to enter or exit the doorway with different walking speeds, including slow (about $0.7m/s$), normal (about $1m/s$) and fast (about $1.3m/s$) walking. The accuracy of passing direction detection with different walking speeds is shown in Fig. 8(a). The average accuracy among the three walking speeds is around 94%. The detection accuracy of slow and normal walking speeds is quite similar with each other, while walking with faster speed can lead to more detection failure. This is because that the effective time series will be less affected by the moving behavior when the object is moving too fast and the corresponding time span will be shorter which makes the extraction of the signal trend suffer from more disturbance.

2) Impact of sampling rate: In addition to the walking speed, we also investigate the impact of different sampling rates on the accuracy of passing direction detection. In fact, the sampling rate is relevant to the walking speed. If the sampling rate is too low, the effective time series of the passing behavior will be too short to be detected. By contrast, if the sampling rate is too high, then it can lead to more noises in the wireless signals. Figure 8(b) depicts the accuracy on passing direction detection under different sampling rates, which are 100 packets per second (p/s), $200p/s$ and $400p/s$. It shows that the sampling rate of $200p/s$ can achieve the best performance with around 95.5% detection accuracy. In our case, the distance between the transmitter and two receivers is about 1 meter. The sampling rate can be adaptively adjusted with the change on the distance, like applying higher sampling rate on longer distance scenario.

3) Impact of the number of people: The above experiments are done with a single human object. Other than the single-person scenario, we also explore the effect of the number of people on the performance of passing direction detection. The result is shown in Fig. 8(c) for 2, 3 and 4 human objects moving in or out the door together. Although the accuracy on passing direction detection drops with the rising of the number of people, the result is still above 90%. More people passing by the door induces more variations in the wireless signals owing to the superposition of multiple human objects as the reflector and various walking patterns, leading to more misinterpretation on the pattern of the phase difference.

D. Multi-People Passing Identification

To evaluate the performance on people counting, both the estimation error on the number of people and the overall accuracy are calculated. Besides, we investigate the impact of the PCA for the preprocessing of the input matrix to ICA, the impact of the threshold for identifying effective components and the group passing patterns on the counting result.

1) Impact of α: In the previous section, we set a threshold α for identifying the effective components. Here, we try to find out the optimal value of α. The value of α is iterated from 1 to 5 with the interval of 0.5, and the accuracy on counting passing people is shown in Fig. 9(a). The accuracy goes up when α changes from 1 to 2.5, and then experiences slight fluctuation and reaches the highest peak when α is 3. Then,

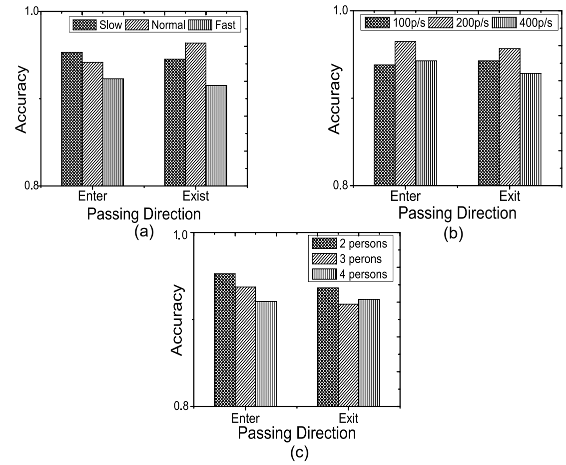

Fig. 8. Accuracy on passing direction detection with different (a) walking speeds; (b) sampling rates and (c) number of people

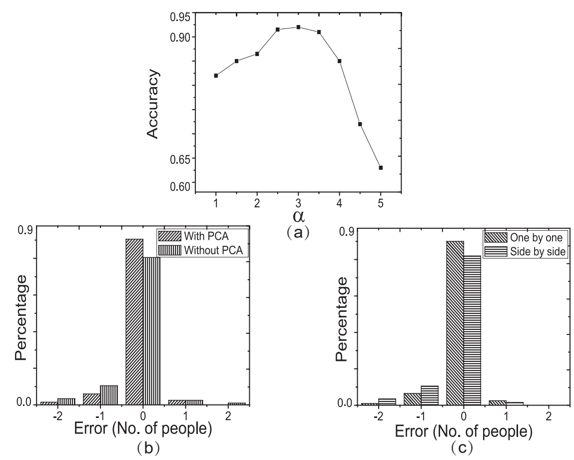

Fig. 9. Performance on multi-person passing identification (a) with difference α; (b) with and without PCA and (c) with different group patterns

the accuracy decreases sharply after 4. From Fig. 6, we can also draw the similar conclusions, the possible range for α is between 2 to 4. Therefore, α is set to be 3 from empirical observation and experiments.

2) Impact of PCA: To enhance the performance of ICA for estimating the number of passing people, we employ principal component analysis for preprocessing the input matrix. We compare the counting performance with and without PCA, and the result is shown in Fig. 9(b). With PCA, around 90% of the counting result is correct, while the percentage goes down to around 80% when the PCA is not applied. This implies that simply using all the 30 subcarriers as the input matrix has some adverse effects on the result since some subcarriers are more vulnerable to noises rather than the moving behavior.

3) Impact of group passing pattern: Next, we investigate how our counting method goes with different group passing patterns. Volunteers are required to enter or exit the door with two patterns, which are one by one and side by side. For the one-by-one pattern, they are close to each other with around $0.5m$ spacing. While for the side-by-side pattern, there is less vertical space among them. The error for the counting result is illustrated in Fig. 9(c), which tells that there are more errors on the estimation of the number of people when they are passing side-by-side. This is due to the reason that, they may have less distinctive vibrations on the wireless signals. However, since people cannot completely copy other's moving behavior, our method can still figure out the exact number of people passing by for most of the cases (85%).

VII. Discussion

The presence of moving human objects will influence the propagation of wireless signals. Therefore, we discuss the effects of the surrounding human activities on the counting result. Directional antennas are employed, and the displacement of the transmitter and receivers are only targeted at the passing human object. Hence, the effects of the surrounding changes will be lowered to the least. We allow several volunteers to walk around in the room to see its effects on the wireless signals received by the two antennas. Experiments are done to evaluate the performance of passing direction detection and multi-person passing identification. The average accuracy is still around 90% for bi-directional passing detection and the percentage of identifying the correct number of people is about 86%, indicating a decent performance with the surrounding human activities.

VIII. Conclusion

In this paper, we propose a passing people counting system, Wi-Count using WiFi signals. The number of people is counted when human objects pass by the doorway in a low-cost, accuracy and non-intrusive manner. Wi-Count can be deployed with the existing indoor WiFi infrastructure, and people can be detected without active participation and carrying any devices. We present a physical model to represent the effects of the bi-directional passing behavior on the wireless signals with respect to the phase information. So, the passing direction can be recognized by the specific pattern in the phase information. Wi-Count not only detects the passing direction but also counts the number of people when multiple persons pass by concurrently through an enhanced counting algorithm, so that to provide more precise counting results. Extensive experiments verify the correctness of the physical model and the performance of our people counting approach. The average accuracy on passing direction detection and passing people counting are around 95% and 92% respectively, and the system is robust to the surrounding moving human objects.

Acknowledgment

The work presented in this paper was supported in part by the NSFC Key Grant with Project No. 61332004 and Hong Kong RGC General Research Fund under Grant PolyU 152244/15E.

References

[1] F. Wahl, M. Milenkovic, and O. Amft, "A distributed pir-based approach for estimating people count in office environments," in *CSE 2012*. IEEE.

[2] A. Sikdar, Y. F. Zheng, and D. Xuan, "An iterative clustering algorithm for classification of object motion direction using infrared sensor array," in *TePRA 2015*. IEEE.

[3] M. Li, Z. Zhang, K. Huang, and T. Tan, "Estimating the number of people in crowded scenes by mid based foreground segmentation and head-shoulder detection," in *ICPR 2008*. IEEE.

[4] J. D. Nichols, L. L. Bailey, N. W. Talancy, E. H. Campbell Grant, A. T. Gilbert, E. M. Annand, T. P. Husband, J. E. Hines *et al.*, "Multi-scale occupancy estimation and modelling using multiple detection methods," *Journal of Applied Ecology*, vol. 45, no. 5, pp. 1321–1329, 2008.

[5] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *PerCom) 2013*. IEEE.

[6] L. Schauer, M. Werner, and P. Marcus, "Estimating crowd densities and pedestrian flows using wi-fi and bluetooth," in *Mobiquitous 2014*. ICST.

[7] P. G. Kannan, S. P. Venkatagiri, M. C. Chan, A. L. Ananda, and L.-S. Peh, "Low cost crowd counting using audio tones," in *Sensys 2012*. ACM.

[8] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only wifi power measurements," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1381–1393, 2015.

[9] 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*. IEEE.

[10] C. Xu, B. Firner, R. S. Moore, Y. Zhang, W. Trappe, R. Howard, F. Zhang, and N. An, "Scpl: Indoor device-free multi-subject counting and localization using radio signal strength," in *IPSN 2013*. IEEE.

[11] H. Ding, J. Han, A. X. Liu, J. Zhao, P. Yang, W. Xi, and Z. Jiang, "Human object estimation via backscattered radio frequency signal," in *Infocom 2015*. IEEE.

[12] J. W. Choi, X. Quan, and S. H. Cho, "Bi-directional passing people counting system based on ir-uwb radar sensors," *IEEE Internet of Things Journal*, 2017.

[13] R. W. Schafer, "What is a savitzky-golay filter?[lecture notes]," *IEEE Signal processing magazine*, vol. 28, no. 4, pp. 111–117, 2011.

[14] X. Zhao, E. Delleandrea, and L. Chen, "A people counting system based on face detection and tracking in a video," in *AVSS 2009*. IEEE.

[15] S. A. M. Saleh, S. A. Suandi, and H. Ibrahim, "Recent survey on crowd density estimation and counting for visual surveillance," *Engineering Applications of Artificial Intelligence*, vol. 41, pp. 103–114, 2015.

[16] T. Germa, F. Lerasle, N. Ouadah, and V. Cadenat, "Vision and rfid data fusion for tracking people in crowds by a mobile robot," *Computer Vision and Image Understanding*, vol. 114, no. 6, pp. 641–651, 2010.

[17] Y. Yuan, C. Qiu, W. Xi, and J. Zhao, "Crowd density estimation using wireless sensor networks," in *MSN 2011*. IEEE.

[18] X. Guo, B. Liu, C. Shi, H. Liu, Y. Chen, and M. C. Chuah, "Wifi-enabled smart human dynamics monitoring."

[19] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Mobicom 2015*. ACM.

[20] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "Rt-fall: A real-time and contactless fall detection system with commodity wifi devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, 2017.

[21] X. Wang, C. Yang, and S. Mao, "Phasebeat: Exploiting csi phase data for vital sign monitoring with commodity wifi devices," in *ICDCS 2017*. IEEE.

[22] Y. Zhuo, H. Zhu, H. Xue, and S. Chang, "Perceiving accurate csi phases with commodity wifi devices," in *Infocom 2017*. IEEE.

[23] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *Ubicomp 2016*. ACM.

[24] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "Wfid: passive device-free human identification using wifi signal," in *Mobiquitous 2016*. ACM.

[25] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2004, vol. 46.

[26] B. A. Draper, K. Baek, M. S. Bartlett, and J. R. Beveridge, "Recognizing faces with pca and ica," *Computer vision and image understanding*, vol. 91, no. 1-2, pp. 115–137, 2003.

[27] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000.

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