Pedestrian Flow Estimation Through Passive WiFi Sensing

Baoqi Huang, *Member, IEEE,* Guoqiang Mao, *Fellow, IEEE,* Yong Qin and Yun Wei

***Abstract*—In public places, even if pedestrians do not have their mobile devices connected with any WiFi access point (AP), WiFi probe requests will be broadcast, so that WiFi sniffers can be employed to crowdsource these WiFi probe packets for use. This paper tackles the problem of exploiting the passive WiFi sensing approach for pedestrian flow analysis. To be specific, a passive WiFi sensing model is first established based on a probabilistic analysis of interactions between WiFi sniffers and the moving pedestrian flow, capturing the main factors affecting pedestrian flow characteristics. On that basis, a sequential filtering algorithm is proposed based on the Rao-Blackwellized particle filter (RBPF) to produce simultaneous and efficient estimates of the pedestrian flow speed and pedestrian number utilizing the real-time sniffing results. In order to validate this study, an experimental pedestrian surveillance system using WiFi sniffers is deployed at the transfer channel of a metro station in Guangzhou, China. Extensive experiments are conducted to verify the passive sensing model, and confirm the effectiveness and advantages of the proposed algorithm. The pedestrian flow estimation not only helps to improve the safety and facility management and customer services, but also paves the way for introducing other novel applications.**

***Index Terms*—Pedestrian analysis, passive WiFi sensing, crowdsourcing, WiFi fingerprinting**

I. Introduction

NOWADAYS, WiFi network infrastructures (e.g. access points, APs) and WiFi enabled mobile devices (e.g. smartphones, tablets, etc.) have become pervasive in our daily lives. In light of IEEE 802.11 Standard, a mobile device actively and periodically broadcasts probe (request) frames across different channels for the purposes of associating with an AP and switching between different APs. Accordingly, if any nearby AP may receive a probe frame irrespective of whether the mobile device is associated with it, a probe (response) frame will be returned. Since probe frames involve some spatial-temporal information about the user carrying this mobile device, a special kind of WiFi APs, termed WiFi sniffers, can thus be leveraged to passively sense the user' behaviors and activities in public places [1], such as campus, shopping malls, metro stations, etc., by crowdsourcing the probe frames, thus enabling new possibilities for developing various novel applications in an automatic and non-intrusive manner, such as pedestrian flow or crowd analysis [2]–[5], tracking trajectories [6], unveiling social relationship [7]–[10], measuring queueing time [11], localization [12]–[16], understanding urban scenes [17], [18], etc. In fact, besides probe frames, some WiFi sniffers are able to sniff normal WiFi packets, so as to enhance their sensing capability.

In public places with massive pedestrian traffic, such as railway stations, underground pedestrian passages, and so on, it is of great importance to monitor pedestrian situations in real time for pedestrian safety and flow management. However, the popularly used video based approaches normally suffer from high computational complexities and small coverage, and incur low performance especially in the conditions of poor illuminations and high pedestrian densities [19]. In contrast, passive WiFi sensing based pedestrian flow analysis is to some extent immune to these limitations, and thus becomes attractive in the literature.

According to the existing pedestrian flow studies [20], [21], number and speed are two fundamental attributes of pedestrian flows. Firstly, learning the number of pedestrians located in or going through a specific region plays a vital role in understanding the criticality of a situation [20]. Existing passive WiFi sensing based studies [2], [3], [5] attempted to train a simple linear model mapping the number of detected mobile devices to the number of pedestrians in an area of interest, and reported the error rate of more than around 20%. Therefore, great efforts have to be devoted to improving the estimation accuracy. Secondly, since simply using pedestrian numbers does not allow for a complete assessment of the criticality of a situation, pedestrian flow speed is usually adopted as another key factor for the dynamic management of public infrastructures as well as providing convenient customer services [21]. In the active sensing approach, pedestrians share their real-time location information determined by GPS and other localization measurements through an APP installed on their smartphones, such that pedestrian speeds can be directly calculated [22], whereas in the passive sensing approach, due to the sparse and random nature of detecting mobile devices, it is challenging to estimate the speed of a pedestrian flow.

This paper tackles the problem of inferring two key characteristics (i.e. the speed and number) of a pedestrian flow using WiFi sniffers. First, a passive WiFi sensing model is established based on a probabilistic analysis on a moving pedestrian flow, and characterizes the interactions between WiFi sniffers and the pedestrian flow, enabling us to theoretically investigate the pedestrian flow characteristics. Second, based on the Rao-Blackwellized particle filter (RBPF) [23], [24], a sequential filtering algorithm is proposed to estimate both the speed and number of the pedestrian flow given WiFi sniffing results.

B. Huang is with College of Computer Science, Inner Mongolia University, Hohhot 010021, China. (e-mail:cshbq@imu.edu.cn)

G. Mao is with School of Computing and Communication, University of Technology Sydney and Data61, CSIRO, Sydney, Australia. (e-mail:guoqiang.mao@uts.edu.au)

Y. Qin is with State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China. (e-mail:yqin@bjtu.edu.cn)

Y. Wei is with Beijing Urban Construction Design & Development Group Co., Limited, Beijing 100037, China.(e-mail:luckyboy0309@163.com)

Therein, the WiFi fingerprint-based localization method [13], [25] is adopted to assist the proposed algorithm in obtaining location information of detected mobile devices.

For the purpose of performance evaluation, an experimental pedestrian surveillance system consisting of 5 WiFi sniffers was deployed at the transfer channel of the Yangji metro station in Guangzhou, China, and a sniffing dataset was collected at four typical scenarios during two work days. It is shown that, the accuracy of pedestrian number estimates based on the passive WiFi sensing model is significantly better than that using the existing linear model [2], confirming the superiority of the passive WiFi sensing model; moreover, the error rate of pedestrian flow speed estimates in the peak scenarios is below 10% with a probability of more than 0.8, revealing the advantages of the proposed algorithm in terms of the pedestrian flow speed estimation; in addition, a qualitative analysis indicates the consistency between the estimates of the proposed algorithm and the reality obtained by a video camera, thus verifying the effectiveness of the proposed algorithm.

To sum up, the major contributions of this paper include

1) A passive WiFi sensing model is formulated to characterize the interactions among various factors in relation to pedestrian flows and passive WiFi sniffers, including the pedestrian number, pedestrian flow speed, detected mobile device number, probabilities of a pedestrian carrying different number of mobile devices, layout of the surveillance area, and so on.
2) The RBPF based algorithm is developed to provide a unified and efficient solution by combining the pedestrian flow speed estimation based on particle filter and the pedestrian number estimation based on linear Kalman filter.
3) In order to provide more visual and informative results, the surveillance area can be partitioned into a set of small grids, so that a fine-grained estimate of the pedestrian number with respect to each grid is derived.
4) An experimental pedestrian surveillance system using WiFi sniffers is implemented for evaluating the passive sensing model and proposed algorithm in practice.

The rest of the paper is organized as follows. Section II surveys the literature in relation to the WiFi passive sensing. Section III establishes a passive WiFi sensing model for further pedestrian flow analysis. Section IV presents the RBPF based algorithm for estimating the pedestrian number and pedestrian flow speed. Section V report the field experimental results. We conclude this paper in Section VI.

II. Related Works

In this section, we shall briefly introduce the literature of pedestrian analysis based on the traditional video based approaches and the emerging wireless based approaches.

A. Video based Pedestrian Analysis

A recent paper [19] surveyed the existing studies on video based crowd scene surveillance, and reported the challenges in various video based techniques. Firstly, pixel-level approaches begin with edge detection and use edge features to train a model, and texture-level approaches are coarser grained than pixel-level approaches and aim to analyze image patches. These two low-level approaches aim to estimate the pedestrian number in a scene, rather than identify individuals, and can only achieve coarse-grained results. Secondly, object-level approaches are able to obtain more accurate results by identifying individual subjects in a scene, but are only suitable for sparse scenes. Thirdly, line counting approaches count the number of pedestrians crossing a line of interest rather than an entire scene or a region of interest. Therefore, the criticality of a situation in a scene, particularly pedestrian crowdness, cannot be thoroughly understood. Fourthly, density mapping approaches estimate the density of a scene rather than identifying the number of pedestrians in the scene, and require to train a model by using a number of labelled samples.

Besides the above individual limitations, a common limitation of all the video based techniques is that only the small area covered by single camera is considered. That is to say, when dealing with pedestrian analysis in a large area covered by multiple cameras, it is still an open problem, not to mention the possible coverage holes.

Based on the above discussion, it can be concluded that, besides the traditional limitations (e.g., illumination conditions, computational complexities), existing video based approaches are also restricted by large area surveillance, high pedestrian densities and cross-camera analysis.

B. Wireless based Pedestrian Analysis

First, with the advantages of low cost, large coverage, scalability, non-intrusive detection and convenience for target recognition, passive WiFi sensing has enabled various pedestrian analysis, including pedestrian number estimation, density estimation, target tracking, and so on.

In [1], the feasibility of the passive WiFi sensing based pedestrian flow analysis method was validated through field experiments on a university campus. Subsequently, the authors deployed motion sensors at the entrance and exit of a mall for pedestrian counting, and then estimated the mapping between the pedestrian number and the detected mobile device number, in turn the estimation of the number of pedestrians in the mall; the field experiments showed that, this method suffers higher error rates than 30% when the detected mobile device number is around 2000, and even worse, does not work when the number of detected mobile devices is less than 2000 (see [2] for details). Similarly, the authors in [3] adopted WiFi sniffers with directional antennas and video-based pedestrian number estimation method, so as to reduce the error rate close to 20%; in [5], a stereoscopic camera was installed at a specified calibration choke point and helped to reduce the error rate by 43.68% compared with the simple WiFi sniffing method.

Moreover, the data collected by 29 public WiFi sniffers in Lower Manhattan, New York, were employed to infer the pedestrian number as well as the pedestrian flows cross different streets [26]. In addition, a vehicle trajectory tracking system was implemented in [6] by sniffing the mobile devices of vehicle occupants, and reported positioning errors as high as 67 meters due to the long measurement range, limited number of sniffers and sparse sniffing data.

In short, the existing studies based on passive WiFi sensing approach, though appear to be feasible in practice, are mainly restricted by the relatively low accuracy.

Second, efforts have been devoted to mining a variety of social information from WiFi sniffing data.

In [8], the preferred network list information contained in probe frames was firstly utilized to infer social connections between users. Similarly, this information was used to infer where participants come from in large events (such as elections) [9], and the analysis showed that the results were highly credible. However, the inclusion of the preferred network list information in probe frames has been controlled to protect privacy [27]. In addition, based on WiFi sniffing, researchers reported a queuing time estimation algorithm in [11], a device classification algorithm in [28], a pedestrian meeting event detection algorithm in [17], and an algorithm for mining potential social relationships and interaction patterns by exploiting empty data frames in [10].

In summary, neither the traditional video based approaches nor the emerging wireless based approaches are able to perfectly resolve the issues confronted by pedestrian analysis. However, the wireless based approaches demonstrate the potentials for addressing the issues of crowd estimation in large areas and/or with high pedestrian densities, which are the major shortages of video based approaches. As such, it is imperative to understand the rules governing the relationship between a pedestrian flow and WiFi sniffing results as well as to improve the accuracy of WiFi based pedestrian analysis.

III. A Passive WiFi Sensing Model for Pedestrian Flow Analysis

On the one hand, the invisible and indeterministic nature associated with WiFi signal propagations leads to uncertainties in matching the mobile devices detected by sniffers and pedestrians to be monitored. On the other hand, the event that a sniffer detects a mobile device is random and relies on the spatial-temporal evolution between the sniffer and the mobile device. Therefore, it is necessary to bridge the gap between the data collected by sniffers and the pedestrian flow of interest. To this end, we first present a probabilistic analysis of WiFi sniffing with a pedestrian flow, then establish a passive WiFi sensing model connecting WiFi sniffing data to the pedestrian flow, and finally discuss how to infer the characteristics of the pedestrian flow based on this model.

A. The Probabilistic Analysis

A mobile device is detectable if and only if its WiFi is switched on. In this paper, we only consider such detectable mobile devices, so that *any mobile device mentioned will be detectable unless otherwise specified.*

The following notations are used in the later analysis:

- L and W are the length and width of a surveillance area as shown in Fig. 1, respectively;
- v is the moving speed (meter per second) of a continuous pedestrian flow going through the surveillance area;
- ρ is the intensity of a spatial Poisson distribution that the pedestrian distribution is assumed to follow;

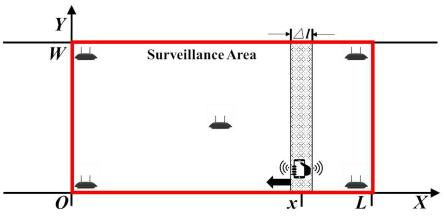


Fig. 1. The illustration of the problem

- b is the expected number of mobile devices carried by one pedestrian, namely

$$b = \sum_{i=1}^{\infty} i p_i, \tag{1}$$

  where p_i is the probability that a pedestrian carries i mobile devices;
- $q(t)$ is the probability that a mobile device can be detected by a sniffer during a time interval t.

Given a specific surveillance area, L and W are usually known constants, b is unknown constant and will be approximated through regression, the function $q(t)$ will be used for deducing the sensing model, and v and ρ are the two key variables to be estimated.

Since v and ρ reflect collective characteristics of a pedestrian flow, we make the following assumption

Assumption 1: Given a time interval which is sufficient for a pedestrian to go through the surveillance area with the length of L, the corresponding pedestrian flow speed v and intensity ρ are assumed to be constant.

In addition, for ease of analysis, we assume

Assumption 2: The detection event is independent during distinct differential intervals and among different mobile devices.

Consider a continuous pedestrian flow passing through a surveillance area, and without loss of generality, suppose that the pedestrian flow moves along the opposite direction of the x-axis, as shown in Fig. 1. We can obtain the following lemma describing the relationship between the probability of detecting a mobile device and the initial position of its holder.

Lemma 1: If a pedestrian carrying a mobile device goes through the surveillance area from x, the probability that the mobile device can be detected is

$$1 - e^{-c\frac{x}{v}}, \tag{2a}$$

where

$$c = \lim_{\Delta t \to 0} \frac{q(\Delta t)}{\Delta t}. \tag{2b}$$

Proof: According to Assumption 1, it takes an average of $\frac{x}{v}$ seconds for the pedestrian to pass through the surveillance area. The probability, i.e. $q(\frac{x}{v})$, can be calculated as follows

$$\begin{aligned} q\left(\frac{x}{v}\right) &= 1 - (1 - q(\Delta t))^{\frac{x}{v\Delta t}} \\ &\cong 1 - e^{-\frac{x}{v\Delta t} q(\Delta t)} \end{aligned}$$

with $\Delta t \to 0$. The lemma is proved. ■

It is noticeable that c actually denotes the first derivative of $q(t)$ at $t = 0$ and will function as an intermediate variable in the following formulations.

Moreover, we can obtain the expected number of detected mobile devices in the following lemma.

Lemma 2: If a pedestrian, carrying i mobile devices with probability of p_i $i = 0, 1, ...,$ goes through the surveillance area from x, the expected number of detected mobile devices associated with this pedestrian is

$$\sum_{i=1}^{\infty} \left(1 - e^{-c\frac{x}{v}}\right) ip_i. \tag{3}$$

Proof: Due to the independence assumption in Assumption 2, the expected number of detected mobile devices is $\left(1 - e^{-c\frac{x}{v}}\right) i$ when the pedestrian carries i mobile devices. Then, by applying the law of total probability, we can obtain (3) and prove the lemma. ∎

In light of Lemma 1 and 2, we can derive the following theorem describing the relationship between the expected number of detected mobile devices and various factors.

Theorem 1: If all the pedestrians currently in the surveillance area go through the surveillance area, the expected number of detected mobile devices associated with these pedestrians is

$$b\rho WL\left(1 - \frac{v}{cL}\left(1 - e^{-\frac{cL}{v}}\right)\right) \tag{4}$$

Proof: Let Δl be a differential length. In a differential stripe of $W \times \Delta l$ as shown in Fig. 1, it is assumed that the probability that there exists one pedestrian is given by $\rho W \Delta l$. The probability that there are more than one pedestrian can be ignored. Thus, according to Lemma 2, the expected number of mobile devices that are detected and located in $W \times \Delta l$ is $\sum_{i=1}^{\infty}(1 - e^{-c\frac{x}{v}})ip_i\rho W\Delta l$.

When these pedestrians go through the surveillance area, the expected number of detected mobile devices equals to the sum of Δl along the horizontal x-axis, namely $\sum_{i=1}^{\infty}\sum_{L}(1 - e^{-c\frac{x}{v}})i\rho W\Delta l p_i$.

When $\Delta l \to 0$, the summation along L can be substituted by integral, namely $\sum_{i=1}^{\infty} i\rho W p_i \left(\int_0^L \left(1 - e^{-c\frac{x}{v}}\right) \mathrm{d}\, x\right)$, which leads to (4). The theorem is proved. ∎

Evidently, Theorem 1 reveals how the expected number of detected mobile devices in the surveillance area is affected by various factors, including the intermediate variable c, the length L and width W of the surveillance area, the number b, the pedestrian flow speed v and the pedestrian density ρ. On this basis, one can investigate pedestrian flows from a theoretical perspective.

B. The Passive Sensing Model

Prior to presenting the sensing model, let us consider a pedestrian flow with the speed v and intensity ρ during an interval, say $[t_A, t_C]$, as shown in Fig. 2 which involves three snapshots at three distinct time instances, namely t_A, t_B and t_C. Accordingly, the pedestrian flow is divided into three parts with labels A, B and C. Define z_A, z_B and z_C to be the expected numbers of detected mobile devices in relation to the three parts.

It is clear that the pedestrians in the three parts enter and exit the surveillance area at different times. Consequently, pedestrians in the parts A or C incur different durations (when they stay inside of the surveillance area), whereas pedestrians in the part B have the same and maximum duration, i.e. $t_B - t_A$ or $\frac{L}{v}$; see the lengths of yellow bars associated with different pedestrians in Fig. 2.

Therefore, (4) in Theorem 1 is only suitable for determining z_A, namely

$$z_A = \rho WL\left(1 - \frac{v}{cL}\left(1 - e^{-\frac{cL}{v}}\right)\right) b. \tag{5a}$$

Since all the pedestrians in the part B have the maximum duration $\frac{L}{v}$, with the result that z_B can be evaluated by simply replacing L with L_B and x with L in (4)

$$z_B = \rho W L_B (1 - e^{-\frac{cL}{v}})b. \tag{5b}$$

Similarly, z_C can be derived by replacing x with $L - x$ in (4)

$$z_C = \rho WL\left(\frac{v}{cL} - \left(1 + \frac{v}{cL}\right) e^{-\frac{cL}{v}}\right) b. \tag{5c}$$

Given the sniffing data (i.e. the MAC address, time stamp, received signal strength (RSS), Channel ID and sniffer ID when a sniffer detects a mobile device) during $[t_A, t_C]$, it is intrinsically cumbersome to pick out the detected mobile devices to separately count z_A, z_B and z_C. Fortunately, by letting $z_{ABC} = z_A + z_B + z_C$, we have

$$z_{ABC} = \gamma(v) x_{AB}, \tag{6a}$$

where

$$x_{AB} = \rho W(L + L_B), \tag{6b}$$

$$\gamma(v) = (1 - e^{-\frac{cL}{v}})b. \tag{6c}$$

It follows from (6) that, the expected number of pedestrians in the parts A (or C) and B, i.e. x_{AB}, is related to the expected number of detected mobile devices during $[t_A, t_C]$, i.e. z_{ABC}, through the unknown function $\gamma(v)$; that is to say, x_{AB} (or equivalently ρ) can be estimated given z_{ABC} irrespective of whether the detected mobile devices belong to the parts A, B or C.

In summary, as the passive sensing model, (6) will help us infer the characteristics of pedestrian flows.

C. Discussions

In order to apply the sensing model in practice, we have to address the following issues.

1) Duration of $t_C - t_A$: The duration of $t_C - t_A$ is minimized to be $\frac{L}{v}$ with $t_B = t_A$, and plays a vital role in applying the passive sensing model. Its determination should depend on the following issues. First, Assumption 1 is only acceptable given a relatively short duration. Second, a short duration results in a small amount of sniffing data, thus restricting the performance of related algorithms. Third, a large duration leads to a long time delay. Therefore, the duration should be traded off to balance the estimation accuracy (or processing overheads) and time delay.

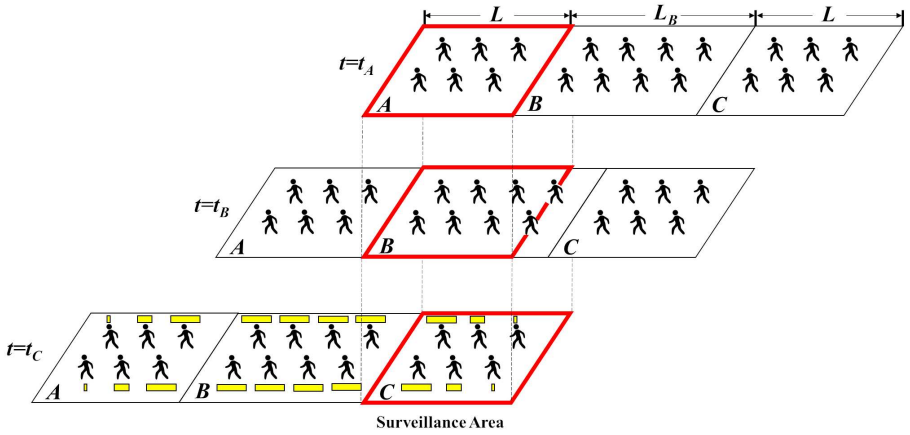


Fig. 2. The snapshots of a pedestrian flow at three time instances.

2) Localization: In the establishment of the passive WiFi sensing model (6), it is a prerequisite to determine whether a mobile device is inside the surveillance area or not at a specific time. Besides, it is also necessary to obtain the locations of detected mobile devices for estimating pedestrian flow speeds. Therefore, the popular WiFi fingerprint-based localization method [13], [25] will be adopted to assist the design and implementation of the algorithm proposed in the subsequent section.

3) Unknown function $\gamma(v)$: The unknown parameters b and c in $\gamma(v)$ can be regarded as constants at a certain circumstance, and thus can be regressed based on (6) by employing sniffing data with different values of v in real environments. Note that b and c are essentially affected by socio-economic factors, such as incomes, habitats, age distributions and so on, thus are environment specific and not universally applicable, and should be updated periodically to accommodate new circumstances.

4) Estimating the Pedestrian Number: With the knowledge of b and c, given the speed estimate, denoted $\hat{v}$, and the observation of z_{ABC} (by counting the unique MAC addresses within $[t_A, t_C]$), denoted by $\hat{z}_{ABC}$, the estimate of the expected pedestrian number x_{AB}, denoted $\hat{x}_{AB}$, can be directly calculated using (6). Furthermore, in view of the assumption of the spatial Poisson distribution, the average number of pedestrians located in the surveillance area can be estimated as

$$\hat{x} = \frac{L}{L+L_B}\hat{x}_{AB} = \frac{L}{(t_C - t_A)\hat{v}\gamma(\hat{v})}\hat{z}_{ABC}, \tag{7}$$

which will be employed to design the pedestrian flow estimation algorithm in the subsequent section.

IV. Proposed Algorithm Based on RBPF

This section presents the RBPF based algorithm for estimating both the pedestrian flow speed and pedestrian number.

A. Preliminaries

In order to smoothly output the estimates, the sliding time window mechanism is adopted with the window size equal to $t_C - t_A$ and a relatively short step size controlling the output frequency. The surveillance area is often as large as hundreds of square meters, and is thus partitioned into small grids (as shown in Fig. 3), so that each grid can be associated with a separate pedestrian number to provide more detailed spatial estimation and more visual information.

In addition, the area neighboring to the surveillance area is also included and divided into grids due to the following reasons. First, it is necessary to determine whether a mobile device is inside the surveillance area or not by using the WiFi fingerprint-based method, which requests the fingerprints in the non-surveillance area. Second, the proposed algorithm can be improved by utilizing the information in the area neighboring to the surveillance area.

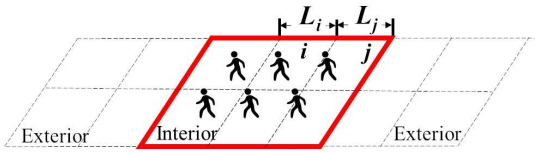


Fig. 3. The interior and exterior grids.

Considering the nonlinear nature of the problem, a two-stage algorithm based on RBPF is designed to estimate the pedestrian flow speed using particle filter in the first stage and then the pedestrian number using linear Kalman filter in the second stage. To this end, define the following notations:

- δ is the step size of the sliding time window;
- t_k is the end time of the k-th sliding time window and satisfies $t_{k+1} - t_k = \delta$;
- v_k is the pedestrian flow speed at t_k;
- n and m denote the numbers of interior and exterior grids, respectively, with the lengths $L_1, \cdots, L_{n+m}$ and widths $W_1, \cdots, W_{n+m}$, respectively;
- $\mathbf{l}_j$ denotes the coordinates of the j-th grid center with $j = 1, \cdots, n+m$;
- $p_{\mathbf{l}_j}(\cdot)$ is the probability density function of the RSS measurement vector at the j-th grid with $j = 1, \cdots, n+m$;
- $\mathbf{x}_k$ is an n-dimensional vector containing the expected pedestrian numbers in the n interior grids at t_k;

- $\mathbf{u}_k$ is an m-dimensional vector containing the expected numbers of detected mobile devices in the m exterior grids at t_k;
- $\mathbf{y}_k = \{\mathbf{y}_{k,i}\}_{i=1}^{s_k}$ contains s_k RSS measurement vectors in the sliding time window at t_k, with $\mathbf{y}_{k,i}$ being a RSS measurement vector averaged over one second and associated with a common MAC address;
- $\mathbb{I}_j^i$ is the index of the j-th RSS measurement vector from the i-th unique MAC address in $\mathbf{y}_k$ with $i = 1, \cdots, m_k$ and $j = 1, \cdots, r_i$, namely $\mathbf{y}_k = \{\mathbf{y}_{k,\mathbb{I}_j^i}\}_{i=j=1}^{i=m_k, j=r_i}$;
- N is the particle number.

The centers of the n interior grids and m exterior grids are employed as reference points to generate a fingerprint database in the offline site survey of the WiFi fingerprint-based method. Given a RSS measurement vector, a probabilistic approach is leveraged to calculate the weight of each grid; taking the $\mathbf{y}_{k,i}$ and the j-th grid for example, its weight is

$$w_{i,j} = p_{\mathbf{l}_j}(\mathbf{y}_{k,i}) \times L_j \times W_j. \tag{8}$$

where $L_j \times W_j$ reflects the influence grids with different sizes. Using the normalization of $w_{i,j}$, the weighted average of $\mathbf{l}_j$ with $j = 1, 2, \cdots, m+n$ is calculated as the location estimate.

B. Particle Filter Part

The speed estimation is equivalent to estimating the posterior probability $p(v_{0:k}|\mathbf{y}_{0:k})$, which can be factorized into the following recursive form according to Bayes's theorem

$$\begin{aligned} p(v_{0:k}|\mathbf{y}_{0:k}) \quad \propto \quad & p(\mathbf{y}_k|v_{0:k}, \mathbf{y}_{0:k-1}) p(v_k|v_{0:k-1}, \mathbf{y}_{0:k-1}) \\ & \times p(v_{0:k-1}|\mathbf{y}_{0:k-1}). \end{aligned} \tag{9}$$

As long as the distributions $p(\mathbf{y}_k|v_{0:k}, \mathbf{y}_{0:k-1})$ and $p(v_k|v_{0:k-1}, \mathbf{y}_{0:k-1})$ are available, (9) can be used to recursively estimating v_k based on particle filter.

Since the distribution $p(v_k|v_{0:k-1}, \mathbf{y}_{0:k-1})$ characterizes the time update or prediction of the pedestrian flow speed, by assuming the Markovian property in the pedestrian flow speed v_k, we can adopt an artificial evolution using kernel smoothing which guarantees the estimation convergence [29], namely

$$v_k = f v_{k-1} + (1-f)\bar{v}_{k-1} + e_{k-1}, \tag{10a}$$

where f is a weight between 0 and 1, $\bar{v}_{k-1}$ is the Monte Carlo mean of v_{k-1}, and $e_{k-1} \sim \mathcal{N}(0, (1-f^2)\sigma_{k-1}^2)$ with σ_{k-1}^2 being the Monte Carlo variance matrix of v_{k-1}.

Moreover, according to Assumption 1 and assuming the independence among different RSS measurements, we can have

$$\begin{aligned} & p(\mathbf{y}_k|v_{0:k}, \mathbf{y}_{0:k-1}) \\ = & \prod_{i=1}^{m_k} p(\mathbf{y}_{k,\mathbb{I}_1^i}, \cdots, \mathbf{y}_{k,\mathbb{I}_{r_i}^i}|v_k) \\ = & \prod_{i=1}^{m_k} \int p(\mathbf{y}_{k,\mathbb{I}_1^i}, \cdots, \mathbf{y}_{k,\mathbb{I}_{r_i}^i}|\Omega_i, v_k) p(\Omega_i|v_k)\, \mathrm{d}\,\Omega_i \\ = & \prod_{i=1}^{m_k} \int p(\mathbf{y}_{k,\mathbb{I}_1^i}, \cdots, \mathbf{y}_{k,\mathbb{I}_{r_i}^i}|\Omega_i) p(\Omega_i|v_k)\, \mathrm{d}\,\Omega_i \\ \propto & \prod_{i=1}^{m_k} \int p(\Omega_i|\mathbf{y}_{k,\mathbb{I}_1^i}, \cdots, \mathbf{y}_{k,\mathbb{I}_{r_i}^i}) p(v_k|\Omega_i)\, \mathrm{d}\,\Omega_i \\ \approx & \prod_{i=1}^{m_k} \sum_{\Theta_i} p(\Theta_i|\mathbf{y}_{k,\mathbb{I}_1^i}, \cdots, \mathbf{y}_{k,\mathbb{I}_{r_i}^i}) p(v_k|\Theta_i) \\ = & \prod_{i=1}^{m_k} \sum_{\Theta_i} \left(\prod_{j=1}^{r_i} p(\mathbf{l}_{\theta_{i,j}}|\mathbf{y}_{k,\mathbb{I}_j^i}) \right) p(v_k|\Theta_i) \\ \propto & \prod_{i=1}^{m_k} \sum_{\Theta_i} \left(\prod_{j=1}^{r_i} w_{\mathbb{I}_j^i, \theta_{i,j}} \right) p(v_k|\Theta_i), \end{aligned} \tag{10b}$$

where Ω_i denotes the set of r_i candidate locations corresponding to the r_i RSS measurement vectors associated with the i-th unique MAC address. Similarly, $\Theta_i = \{\mathbf{l}_{\theta_{i,j}}\}_{j=1}^{r_i}$ with $\theta_{i,j}$ taking values from $1, \cdots, n+m$, reflecting that the r_i candidate locations are restricted within the $n+m$ grid centers.

Evidently, the key step of calculating $p(\mathbf{y}_k|v_{0:k}, \mathbf{y}_{0:k-1})$ lies in the evaluation of $p(v_k|\Theta_i)$. The state variable v_k of each particle provides a distance estimate between any pair of location candidates in Θ_i, say $\mathbf{l}_{\theta_{i,j_1}}$ and $\mathbf{l}_{\theta_{i,j_2}}$; thus, we assume that the distance estimate follows a Gaussian distribution, $\sim \mathcal{N}(\|\mathbf{l}_{\theta_{i,j_1}} - \mathbf{l}_{\theta_{i,j_2}}\|, \sigma_r^2)$ (where σ_r^2 is empirically set as 4), which enables us to evaluate $p(v_k|\Theta_i)$.

To reduce the computational overheads and avoid degeneration of the calculations, the following strategies can be adopted. First, instead of the product of a sequence of weights, their geometrical mean is evaluated to avoid degeneration. Second, the dimension of Θ_i is reduced by only considering the 3 most probable candidate locations since more than 3 candidate locations contribute little to the accuracy. Third, the weights with respect to all the r_i RSS measurement vectors and all the grid centers can be stored for use in evaluating (10b) after being calculated during the online localization phase.

C. Kalman Filter Part

It follows from Assumption 1 that the pedestrian numbers in the n interior grids at t_k and t_{k+1} can be formulated as

$$\mathbf{x}_{k+1} = \mathbf{A}(v_k)\mathbf{x}_k + \mathbf{B}(v_k)\mathbf{u}_k + \boldsymbol{\epsilon}_k, \tag{11a}$$

where $\boldsymbol{\epsilon}_k$ is additive white Gaussian noise with covariance $\mathbf{Q}_k$; $\mathbf{A}(v_k)$ (or $\mathbf{B}(v_k)$) is the $n \times n$ (or $n \times m$) matrix describing how pedestrians move between neighboring interior grids (or between interior grids and exterior grids) during the interval δ given the speed v_k.

If pedestrians walk from j-th interior grid to i-th interior grid as shown in Fig. 3, then the associated two elements on the i-th, j-th rows and j-th column of $\mathbf{A}(v_k)$ are respectively

$$\frac{v_k\delta}{L_j}, 1-\frac{v_k\delta}{L_j}, \tag{11b}$$

where $\frac{v_k\delta}{L_j}$ represents the proportion of pedestrians leaving j-th grid for i-th grid; supposing $j>n$ (i.e. j-th grid is exterior), the element on the i-th row and $(j-n)$-th column of $\mathbf{B}(v_k)$ is

$$\frac{v_k\delta}{L_j}\times\frac{L}{(t_C-t_A)v_k\gamma(v_k)}=\frac{\delta L}{(t_C-t_A)L_j\gamma(v_k)}, \tag{11c}$$

where, similarly to (7), $\frac{L}{v_k(t_C-t_A)\gamma(v_k)}$ is used to calculate the expected number of pedestrians from that of detected mobile devices.

Moreover, establish the following measurement equation

$$g(\mathbf{y}_k)=\mathbf{C}(v_k)\mathbf{x}_k+\boldsymbol{\xi}_k, \tag{12a}$$

where

- $g(\mathbf{y}_k)$ returns the n-dimensional vector containing the expected numbers of detected mobile devices in the n interior grids given the measurements $\mathbf{y}_k$, and its j-th element is obtained by normalizing and summing the weights in (8) with respect to this grid;
- $\mathbf{C}(v_k)$ is a $n\times n$ diagonal matrix mapping the expected numbers of detected mobile devices from the pedestrian numbers in the n interior grids, namely

$$\mathbf{C}(v_k)=\frac{(t_C-t_A)v_k\gamma(v_k)}{L}\mathbf{I}, \tag{12b}$$

  where $\mathbf{I}$ is an identity matrix of a proper order;
- $\boldsymbol{\xi}_k$ is additive white Gaussian noise with covariance $\mathbf{R}_k$.

In summary, due to the linear relationship among $\mathbf{x}_k$ and $g(\mathbf{y}_k)$ in (11a) and (12a), subject to additive Gaussian noise, $\mathbf{x}_k$ can be solved by leveraging Kalman filter given v_k.

D. Summarizing the Algorithm

According to Model 1 in [24], we can obtain the RBPF based algorithm as described in Algorithm 1, where $\mathbf{A}_k, \mathbf{B}_k, \mathbf{C}_k$ are abbreviations for $\mathbf{A}(v_k), \mathbf{B}(v_k), \mathbf{C}(v_k)$.

The initial speed v_0 is defined to be randomly and uniformly distributed within the speed range $[v_{min}, v_{max}]$, such that

$$p_{v_0}(v_0)=\frac{1}{v_{max}-v_{min}}, \tag{15}$$

and the initial pedestrian number vector in the grids, i.e. $\mathbf{x}_0$, is assumed to be Gaussian

$$\mathbf{x}_0\sim\mathcal{N}(\bar{\mathbf{x}}_0,\bar{\mathbf{P}}_0), \tag{16}$$

where $\bar{\mathbf{x}}_0$ and $\bar{\mathbf{P}}_0$ can be empirically determined according to normal pedestrian behaviors.

Algorithm 1 The RBPF based pedestrian flow algorithm

1: Initialization: For $i=1,\cdots,N$, initialize the particles, $v_{0|-1}^{(i)}\sim p_{v_0}(v_0)$ and set $\{\mathbf{x}_{0|-1}^{(i)}, P_{0|-1}^{(i)}\}=\{\bar{\mathbf{x}}_0,\bar{\mathbf{P}}_0\}$.
2: k=0.
3: **while** True **do**
4: For $i=1,\cdots,N$, evaluate the importance weights $q_k^{(i)}=p(\mathbf{y}_k|v_{0:k}^{(i)},\mathbf{y}_{0:k-1})$ in (10b) and normalize

$$\tilde{q}_k^{(i)}=\frac{q_k^{(i)}}{\sum_{j=1}^N q_k^{(j)}}.$$

5: Particle filter measurement update (resampling): Resample N particles with replacement

$$\Pr(v_{k|k}^{(i)}=v_{k|k-1}^{(j)})=\tilde{q}_k^{(j)}.$$

6: Particle filter time update and Kalman filter:

1) Kalman filter measurement update: For $i=1,\cdots,N$,

$$\hat{\mathbf{x}}_{k|k}^{(i)} = \hat{\mathbf{x}}_{k|k-1}^{(i)}+\mathbf{K}_k^{(i)}(g(\mathbf{y}_k)-\mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}^{(i)}), \tag{13a}$$
$$\mathbf{P}_{k|k}^{(i)} = \mathbf{P}_{k|k-1}^{(i)}-\mathbf{K}_k^{(i)}\mathbf{C}_k\mathbf{P}_{k|k-1}^{(i)}, \tag{13b}$$
$$\mathbf{K}_k^{(i)} = \mathbf{P}_{k|k-1}\mathbf{C}_k^T(\mathbf{S}_k^{(i)})^{-1}, \tag{13c}$$
$$\mathbf{S}_k^{(i)} = \mathbf{C}_k\mathbf{P}_{k|k-1}^{(i)}\mathbf{C}_k^T+\mathbf{R}_k. \tag{13d}$$

2) Particle filter time update (prediction): For $i=1,\cdots,N$, predict new particles according to (10a),

$$v_{k+1|k}^{(i)}\sim\mathcal{N}(fv_{k|k}^{(i)}+(1-f)\bar{v}_{k|k},(1-f^2)\sigma_{k|k}).$$

3) Kalman filter time update: For $i=1,\cdots,N$,

$$\hat{\mathbf{x}}_{k+1|k}^{(i)} = \mathbf{A}_k\hat{\mathbf{x}}_{k|k}^{(i)}+\mathbf{B}_k\mathbf{u}_k, \tag{14a}$$
$$\mathbf{P}_{k+1|k}^{(i)} = \mathbf{A}_k\mathbf{P}_{k|k}^{(i)}\mathbf{A}_k^T+\mathbf{Q}_k. \tag{14b}$$

7: k=k+1.
8: **end while**

The estimates, as expected means of the state variables and their covariances are given below.

$$\hat{v}_{k|k}=\sum_{i=1}^N\tilde{q}_k^{(i)}\hat{v}_{k|k}^{(i)}, \tag{17a}$$
$$\hat{\mathbf{P}}_{k|k}^v=\sum_{i=1}^N\tilde{q}_k^{(i)}\left(\hat{v}_{k|k}^{(i)}-\hat{v}_{k|k}\right)\left(\hat{v}_{k|k}^{(i)}-\hat{v}_{k|k}\right)^T, \tag{17b}$$
$$\hat{\mathbf{x}}_{k|k}=\sum_{i=1}^N\tilde{q}_k^{(i)}\hat{\mathbf{x}}_{k|k}^{(i)}, \tag{17c}$$
$$\hat{\mathbf{P}}_{k|k}=\sum_{i=1}^N\tilde{q}_k^{(i)}\left(\mathbf{P}_{k|k}^{(i)}+\left(\hat{\mathbf{x}}_{k|k}^{(i)}-\hat{\mathbf{x}}_{k|k}\right)\left(\hat{\mathbf{x}}_{k|k}^{(i)}-\hat{\mathbf{x}}_{k|k}\right)^T\right) \tag{17d}$$

In conclusion, the summation of $\hat{\mathbf{x}}_{k|k}$ returns the estimate of the pedestrian number at time t_k, revealing how many pedestrians are currently located in the surveillance area. Furthermore, dividing each element of $\hat{\mathbf{x}}_{k|k}$ by the corresponding grid area returns the pedestrian density associated with this grid, enabling us draw heat maps for the surveillance area.

TABLE I
THE DETAILS OF THE SNIFFING DATASET IN FOUR TYPICAL SCENARIOS.

Description	Date	Start Time	End Time	Duration	Number of Packets	Number of MACs
Evening Peak	Dec 26, 2018 (Wed)	17:47:51	18:45:46	57min 55sec	13714275	58940
Evening Off-peak	Dec 26, 2018 (Wed)	21:34:55	22:20:26	45min 31sec	6325523	16580
Morning Peak	Dec 27, 2018 (Thur)	07:48:23	08:52:12	63min 49sec	14105625	54101
Afternoon Off-peak	Dec 27, 2018 (Thur)	13:00:02	13:46:44	46min 42sec	5780439	21678

V. EXPERIMENTS

In this section, we conduct extensive experiments to validate the feasibility of the WiFi passive sensing approach as well as to evaluate the performance the proposed algorithm. Specifically, we shall first introduce the implementation of the experimental pedestrian surveillance system, and then present experimental results.

A. Overview of the Experimental System

The experimental pedestrian surveillance system was deployed at the transfer channel between Line 1 and Line 5 at the Yangji metro station in Guangzhou, China, which is divided into two lanes (termed *upper lane* and *lower lane*) by fixed fences, resulting in two bidirectional pedestrian flows.

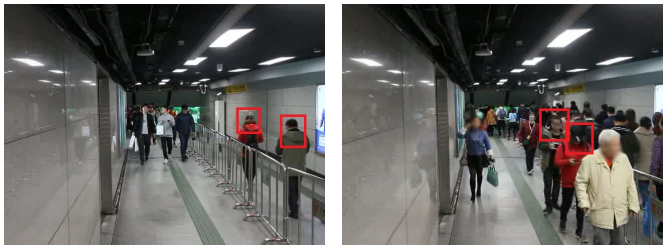
The hardware of the experimental system consists of five WiFi sniffers (*DS-AP-I* [30]), a computer and a 100 Mbps switch which connects the sniffers and computer to form a wired local area network. Specifically, four sniffers were installed at the corners of the surveillance area (i.e. a rectangle with the length of 18.1 meters and width of 5.22 meters), and the other one at the center.

The software of the experimental system involves a UDP server program installed on the computer to receive sniffing data (including Sniffer ID, MAC address, RSS, time stamps in millisecond and channel ID) uploaded by WiFi sniffers at a frequency of 10 Hz, and a Matlab program implementing the proposed RBPF based algorithm.

1) WiFi Sniffers: Each WiFi sniffer is customized to integrate nine dual-band WiFi modules, such that nine different channels can be simultaneously sniffed to crowdsource as many mobile devices and WiFi packets (including both probe requests and normal data packet) as possible.

In particular, five WiFi modules in a sniffer were scheduled to work in 2.4 GHz with each one polling three of the total 13 WiFi channels, and similarly, the other four modules in 5 GHz with each one polling two of the 12 WiFi channels. In addition, during channel polling, each module kept sniffing one channel for 300 milliseconds. Note that one channel may be polled by two different modules.

2) WiFi Fingerprint-based Localization: A fingerprint database was generated by taking RSS measurements from smartphones located at 76 interior reference points (i.e. the centers of the interior grids) in the transfer channel and 99 exterior reference points (i.e. the centers of the exterior grids) in its neighborhood. To reduce the overheads of the site survey, a gaussian process regression (GPR) based method reported in [31], [32] was adopted to generate the fingerprint database with RSS measurements from only 25% of the total 175 reference points. Two smartphones, i.e. an Android smartphone (Huawei Honor 6) and an iOS smartphone (Apple 7) were employed during the offline site survey. In addition, due to the fact WiFi signals in 5 GHz suffer from significantly larger attenuations than in 2.4 GHz, a fingerprint database using RSS measurements from both 2.4 GHz and 5 GHz will degrade localization performance. Since addressing such localization problem is beyond the scope of this paper, we only make use of the RSS measurements in 2.4 GHz for generating the fingerprint database and conducting localization.



(a) Experimenters in the upper lane (b) Experimenters in the lower lane

Fig. 4. The snapshots of the field experiment in the afternoon off-peak.

During the online localization, the RSS measurements from one mobile device are averaged over one second with respect to every sniffer, resulting a RSS measurement vector; if one sniffer does not receive any RSS measurements due to, e.g., collisions, the corresponding element in the vector is set to be -100 dBm. As such, any mobile device is localized at a maximum frequency of 1 Hz.

3) Sniffing Dataset: Sniffing dataset was collected in four typical scenarios, namely evening peak, evening off-peak, morning peak and afternoon off-peak, as listed in Tab. I. It can be found that the peak scenarios result in significantly more packets and MACs than the off-peak scenarios.

For the purposes of training the sensing model and evaluating the proposed algorithm, it is imperative to derive ground-truth pedestrian numbers and pedestrian flow speeds in the sniffing dataset, which is difficult due to the uncontrollability of both pedestrians and environments. To address this issue, a controlled experiment was conducted. As shown in Fig. 4, one experimenter carrying an iOS smartphone (Apple 6S) walked two laps along pedestrian flows in both lanes, and another experimenter followed closely behind the first experimenter and recorded by hand the times when they entered and exited the surveillance area, such that two samples of pedestrian flow speeds can be obtained with respect to each lane. This experiment was conducted twice in each of four typical scenarios. Meanwhile, a video camera was installed to record the pedestrian flows. In order to avoid MAC randomization and ensure more localization numbers, the first experimenter associated the smartphone with a WiFi AP and kept on

refreshing the WiFi list to trigger more probe packets.

As a result, 32 pedestrian flow speed samples were derived in total; moreover, a certain number of locations where RSS measurements from this smartphone were crowdsourced can be inferred as well, and can be further used as ground-truth for evaluating localization accuracy; additionally, four graduate students manually count pedestrians going through the surveillance area by playing the video in slow motion, such that 50 pedestrian number samples with speed labels were derived.

B. Analysis of Localization

The feasibility and performance of the pedestrian flow speed estimation depends on the localization performance in regards to detected mobile devices, including *localization number* (namely how many times a detected mobile device can be localized), *localization interval* (namely the time interval of successively localizing a detected mobile device) and *localization accuracy*. If a mobile device is distant from the surveillance area or wireless channel is busy, it may be detected by less than four sniffers in one second, and will be useless on account of poor localization performance; otherwise, its localization information will be used in the RBPF algorithm.

Due to MAC randomization, we count the localization numbers and localization intervals in terms of each unique MAC address, and plot their empirical cumulative density functions (CDFs) in Fig. 5(a) and 5(b). It can be observed that, for the total MAC addresses, more than 20% of them cannot be localized, and around 30% of them can only be localized once, and the remaining can be localized at least twice, enabling the pedestrian flow speed estimation. Intuitively, the larger is a localization interval, the better is the resulting speed estimation, but there is around 80% of the localization intervals with 1 second, which to some extent restricts the accuracy of the pedestrian flow speed estimation.

Moreover, the empirical CDFs of localization errors are plotted in Fig. 5(c). As can be seen, the median error is around 2.5 meters; this is a fair result considering the localization performance of WiFi fingerprint-based method and the device heterogeneity problem.

C. Analysis of MAC Randomization

Due to privacy issue, more and more mobile devices are adopting MAC randomization, but since the implementation is different across various brands and models of mobile devices, thoroughly and precisely analyzing MAC randomization of every device is tedious, and even impossible. However, as far as the pedestrian flow estimation based on passive WiFi sensing is considered, it is not the specific random MAC address of a mobile device but the update frequency of the random MAC address that affects the proposed pedestrian estimation algorithms. Therefore, we shall present an indirect analysis of MAC randomization based on the sniffing dataset.

Clearly, the worst case happens when every mobile device generates a new random MAC address in each probe request (i.e. each active scan). Fortunately, Fig. 5(a) illustrates that around 50% of the detected MAC addresses can be localized at least twice, and this ratio will be increased if removing those unlocalized ones; that is to say, around 50% of the detected MAC addresses are used in two or more active scans, implying that the real situation is far away from the worst case. Thus, it is feasible to infer the pedestrian number estimation from MAC addresses.

D. Performance Evaluation

We first present the result on regressing the passive sensing model, and then respectively evaluate the performance from quantitative and qualitative aspects.

In the experiments, the particle number $N = 100$, the sliding window size $t_C - t_A = 35$ seconds, the moving step size $\delta = 1$ second, the coefficient $f = 0.95$, $v_{max} = 1.4$ m/s and $v_{min} = 0.4$ m/s according to the field experiment in Hong Kong metro stations [33].

1) Regression of the Sensing Model: Using the nonlinear least squares regression method (i.e. the Matlab routine *lsqnonlin*), we obtained the estimates of the unknown parameters in (6c), namely $c = 0.072$ and $b = 0.7562$, revealing that the expected number of mobile devices carried by one pedestrian is 0.7562 in the considered scenarios.

Note that the value of b is deviated from the average number of mobile devices per person in reality due to the following facts. First, as aforementioned, we only consider detectable mobile devices with WiFi switched on, namely that mobile devices with WiFi disabled are not counted here, thus making b be decreased. Second, b tends to be increased due to MAC randomization. However, b is only a parameter used for describing the passive sensing model, and is not necessarily equal to the average number of mobile devices per person.

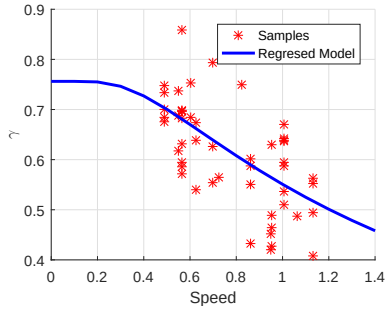
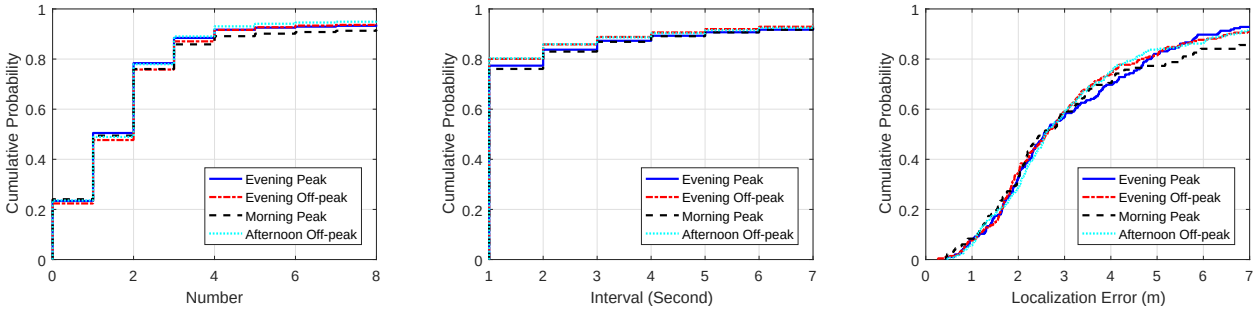


Fig. 6. The illustration of $\gamma(v)$ in the passive sensing model.

The value of γ is plotted with respect to different v in Fig. 6. As can be seen, γ decreases from around 0.75 to 0.45, with v rising from 0 to 1.4 m/s, due to the fact that pedestrians with a higher speed will go through the transfer channel in a shorter time, resulting a less number of detected mobile devices, consequently a lower γ. More importantly, it can be observed that γ appears to be constant when the speed is below 0.4 m/s, which is attributable to the fact that pedestrians with an extremely low speed will stay in the transfer channel for a sufficiently long time, such that no more mobile devices can be detected, thus keeping γ constant.



(a) The CDF of localization numbers (b) The CDF of localization intervals (c) The CDF of localization results

Fig. 5. Statistics of sniffing dataset in terms of localization.

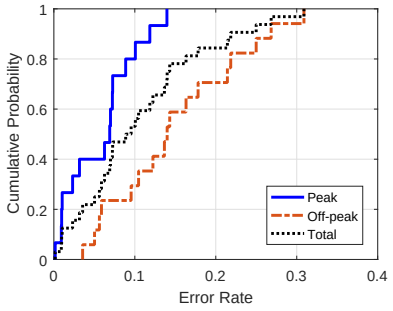


Fig. 7. The CDF of pedestrian flow speed estimation results.

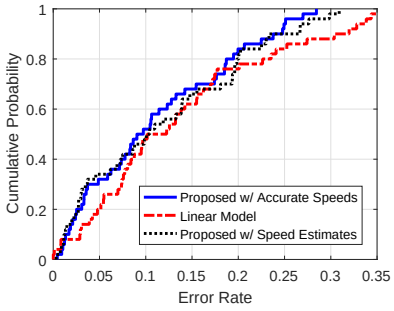


Fig. 8. The CDF of pedestrian number estimation results.

2) Quantitative Evaluation: In the first place, in order to evaluate the performance of the pedestrian flow speed estimation, define $\frac{|v-\hat{v}|}{v}$ to be the error rate. The empirical CDF of the overall error rates as well as those respectively in the peak and off-peak scenarios are plotted in Fig. 7.

As can be seen, the error rate in the peak scenarios is apparently better than that in the off-peak scenarios, and specifically, is below 0.1 in around 80% cases. This is attributable to the following reasons: (1) more measurements available in the peak scenarios result in more accurate estimation; (2) pedestrians in the peak scenarios tend to be synchronized to walk at nearly a same slow speed, whereas pedestrians in the off-peak scenarios often walk freely with varying speeds.

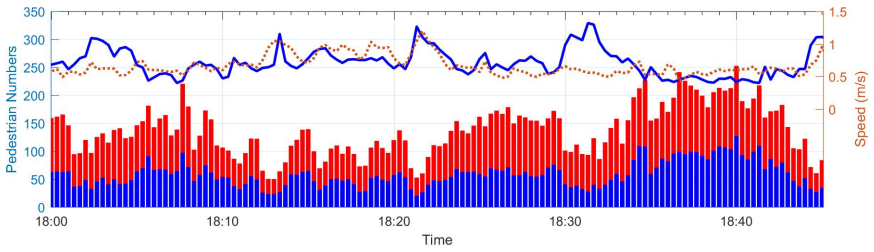
In the second place, since the proposed algorithm estimates the pedestrian number in the transfer channel at any time instant, the ground-truth of which cannot be derived, especially in crowded situations, indicating that it is hard to conduct a quantitative performance evaluation. Alternatively, we can use the 50 samples involving pedestrian numbers to quantitatively evaluate the accuracy of the pedestrian number estimation method (7), which is actually the basis of the proposed algorithm. Besides, (7) with the accurate speed v and the linear model based method in [2] are also implemented for comparison.

Due to the limited number of total samples, the five-fold cross validation method is adopted to generate as many results as possible for plotting. The empirical CDFs of the error rates in pedestrian number estimation are plotted in Fig. 8. As can be seen, the usage of speed estimates only slightly deteriorates the pedestrian number estimation, confirming the effectiveness of the sensing model as well as the pedestrian number estimation method; more importantly, no matter whether an accurate speed is known or not, the method based on the sensing model outperforms the linear model based method reported in [2] in most cases.

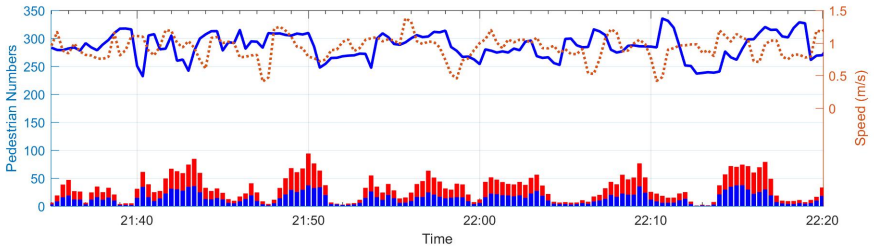
3) Qualitative Evaluation: To further validate the proposed algorithm, we conduct qualitative performance evaluations through the numbers of pedestrians located in the transfer channel and the pedestrian flow speeds, which are both estimated by the proposed algorithm. For ease of presentation, we pick out 45 minutes from the total duration of each typical scenario, and plot the results every other 20 seconds in Fig. 9.

First of all, it is shown that the pedestrian numbers in the peak scenarios are significantly greater than those in the off-peak scenarios, which is certainly consistent with the real situations.

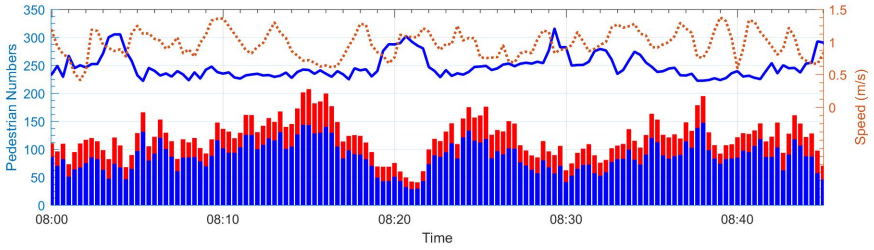
It is intuitive that with a train arriving at the metro station of either Line 1 or Line 5, the pedestrians tend to increase in the transfer channel, resulting in a peak in the pedestrian number. As can be seen from Fig. 9, in the afternoon off-peak scenario, peaks in the blue bars (i.e. the lower lane) appear at 13:01, 13:02, 13:07, 13:08, 13:10, 13:11, etc., because subway trains drive from two opposite directions; similarly, in the evening off-peak scenario, peaks in the blue bars (i.e. the lower lane) appear with larger intervals than those in the afternoon off-peak scenario, due to the fact the subway train frequency is often reduced in the evening off-peak times. However, the peaks of the pedestrian numbers in the peak scenarios are not as distinct as in the off-peak scenarios, on account of the high pedestrian density and subway train frequencies.



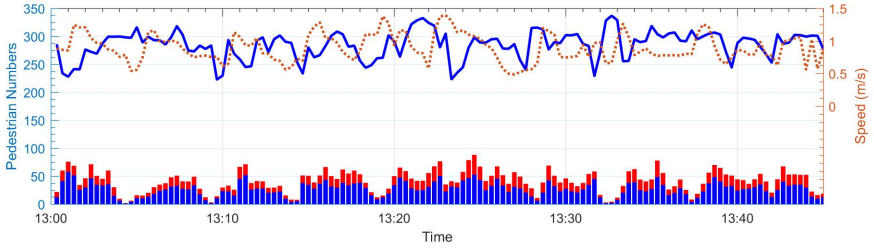
(a) Evening peak on 26 Dec 2018



(b) Evening off-peak on 26 Dec 2018



(c) Morning peak on 27 Dec 2018



(d) Afternoon off-peak on 27 Dec 2018

Fig. 9. Estimates of the pedestrian numbers and pedestrian flow speeds in the four typical scenarios (The bars denote the pedestrian number estimates and the curves denote the pedestrian flow speed estimates; the blue bars and curves correspond to the lower lane and the red ones to the upper lane).

Moreover, the pedestrian numbers in the evening scenarios appear to be symmetric in both lanes, whereas in the daytime scenarios the pedestrian numbers in the lower lane are usually greater than those in the upper lane; this interesting observation reflects the habits of citizens transferring between Lane 1 and Lane 5.

Additionally, it can be found that the pedestrian flow speeds are often very low in the peak scenarios, but are relatively high and incur fluctuations in the off-peak scenarios. However, all these observations comply with the common sense that: (1) there exists a restrictive relationship between the pedestrian number (or density) and the pedestrian flow speed in pedestrian flows; (2) with sparse pedestrians in the off-peak scenarios, pedestrians may walk freely at different speeds, and the estimation of pedestrian flow speeds is easily degraded due to insufficient RSS measurements.

In order to have a clear and intuitive view, we select 2 time instants from each typical scenario, and show corresponding snapshots from the videos in Fig. 10. To avoid biases caused by the experimenters, we do not select time instants when they were inside of the transfer channel. Besides, since the proposed algorithm suffers from a short delay due to the sliding window mechanism, we plot in Fig. 11 the heat maps of the pedestrian densities at 10 seconds behind the time of the snapshots.

It can be seen that, the three sources, i.e. pedestrian numbers, the heat maps and snapshots, demonstrate good consistence at all the eight time instants, which to some extent validates the effectiveness of the proposed algorithm. Furthermore, in the crowded situations (e.g. Fig. 10(b) and Fig. 10(e)), it is hard for either us or the video based pedestrian counting approaches

(a) Evening peak (18:02:30) (b) Evening peak (18:40:40) (c) Evening off-peak (21:50:00) (d) Evening off-peak (22:12:20)

(e) Morning peak (08:15:24) (f) Morning peak (08:21:00) (g) Afternoon off-peak (13:30:40) (h) Afternoon off-peak (13:32:30)

Fig. 10. Snapshots from the video taken at the transfer channel in the four typical scenarios.

to tell any differences, whereas the proposed algorithm shows that the pedestrian numbers have a difference of around 50, indicating the advantage of the proposed algorithm in comparison with video based approaches.

In summary, the extensive experiments not only validate the passive WiFi sensing model, but also confirm the effectiveness of the proposed algorithm in comparison with the existing studies in terms of the estimation of pedestrian numbers.

VI. Conclusion

This paper dealt with the problem of inferring pedestrian flow characteristics via a passive WiFi sensing approach. To be specific, the passive WiFi sensing model was initially established by probabilistically analyzing the interactions between a moving pedestrian flow and WiFi sniffers, and after that, the sequential filtering algorithm based on RBPF was proposed to simultaneously estimate the pedestrian number and the pedestrian flow speed given real-time WiFi sniffing data. The experimental pedestrian surveillance system, deployed in the transfer channel of a metro station in Guangzhou, China, confirmed the feasibility of monitoring pedestrian flows through WiFi sniffers as well as the effectiveness of the proposed algorithm.

However, during the design and implementation of the experimental system, we have identified several problems and will consider them in future works. First, it is worthwhile to extend the study and results in the transfer channel to a large-scale surveillance area. Second, cracking randomized MACs will be helpful in both the pedestrian number estimation and pedestrian flow speed estimation. Third, existing solutions for WiFi localization with heterogenous mobile devices do not perform well in our problem due to the small amount of the sniffers (i.e. 5), so that it is necessary to develop a new robust localization algorithm. Last but not least, it is attractive and meaningful to fuse, e.g., video based approaches with the wireless based approach, so as to leverage the output of video based approaches to automatically update the parameters in the sensing model.

VII. Acknowledgement

We thank the colleagues and students who contributed to setting up the experimental system, collecting field data and manually labelling samples from videos. This work is supported by the National Natural Science Foundation of China (Grants No. 41871363 and 61461037), the Natural Science Foundation of Inner Mongolia Autonomous Region of China (Grant No. 2017JQ09), the ”Grassland Elite” Project of the Inner Mongolia Autonomous Region (Grant No. CYYC5016) and Chinese Scholarship Council (CSC).

References

[1] Y. Fukuzaki, M. Mochizuki, K. Murao, and N. Nishio, “A pedestrian flow analysis system using wi-fi packet sensors to a real environment,” in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, ser. UbiComp ’14 Adjunct. New York, NY, USA: ACM, 2014, pp. 721–730.

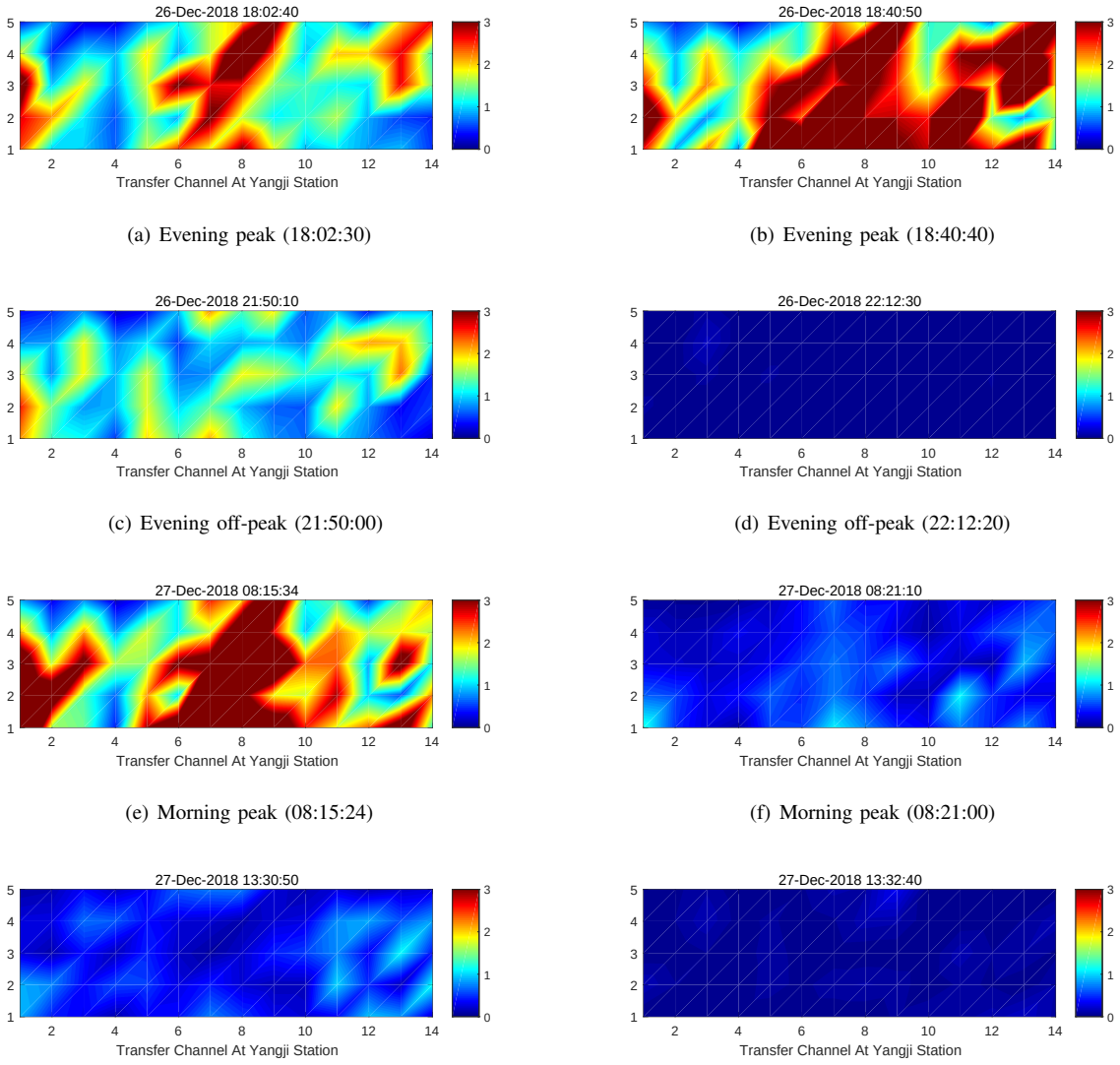
[2] ——, “Statistical analysis of actual number of pedestrians for wi-fi packet-based pedestrian flow sensing,” in *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, ser. UbiComp/ISWC’15 Adjunct. New York, NY, USA: ACM, 2015, pp. 1519–1526.

[3] J. Weppner, B. Bischke, and P. Lukowicz, “Monitoring crowd condition in public spaces by tracking mobile consumer devices with wifi interface,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, ser. UbiComp ’16. New York, NY, USA: ACM, 2016, pp. 1363–1371.

[4] K. Li, C. Yuen, S. S. Kanhere, K. Hu, W. Zhang, F. Jiang, and X. Liu, “Understanding crowd density with a smartphone sensing system,” in *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)*, Feb 2018, pp. 517–522.

[5] F.-J. Wu and G. Solmaz, “Crowdestimator: Approximating crowd sizes with multi-modal data for internet-of-things services,” in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys ’18. New York, NY, USA: ACM, 2018, pp. 337–349.

[6] A. B. M. Musa and J. Eriksson, “Tracking unmodified smartphones using wi-fi monitors,” in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys ’12. New York, NY, USA: ACM, 2012, pp. 281–294.



(a) Evening peak (18:02:30)

(b) Evening peak (18:40:40)

(c) Evening off-peak (21:50:00)

(d) Evening off-peak (22:12:20)

(e) Morning peak (08:15:24)

(f) Morning peak (08:21:00)

(g) Afternoon off-peak (13:30:40)

(h) Afternoon off-peak (13:32:30)

Fig. 11. Heat maps of the pedestrian densities in the four typical scenarios. The lower lane corresponds to the pedestrian flows coming out of the snapshots and the lower part of the heat maps, and vice versa.

[7] M. V. Barbera, A. Epasto, A. Mei, V. C. Perta, and J. Stefa, "Signals from the crowd: Uncovering social relationships through smartphone probes," in *Proceedings of the 2013 Conference on Internet Measurement Conference*, ser. IMC '13. New York, NY, USA: ACM, 2013, pp. 265–276.

[8] M. Cunche, M.-A. Kaafar, and R. Boreli, "Linking wireless devices using information contained in wi-fi probe requests," *Pervasive and Mobile Computing*, vol. 11, pp. 56 – 69, 2014.

[9] 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*, April 2016, pp. 1–9.

[10] H. Hong, C. Luo, and M. C. Chan, "Socialprobe: Understanding social interaction through passive wifi monitoring," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, ser. MOBIQUITOUS 2016. New York, NY, USA: ACM, 2016, pp. 94–103.

[11] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin, "Measuring human queues using wifi signals," in *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*, ser. MobiCom '13. New York, NY, USA: ACM, 2013, pp. 235–238.

[12] P. E. L. de Teruel, F. J. Garcia, and O. Canovas, "Validating passive localization methods for occupancy sensing systems in wireless environments: A case study," *Procedia Computer Science*, vol. 94, pp. 57 – 64, 2016, the 11th International Conference on Future Networks and Communications (FNC 2016) / The 13th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2016) / Affiliated Workshops.

[13] H. Zou, B. Huang, X. Lu, H. Jiang, and L. Xie, "A robust indoor positioning system based on the procrustes analysis and weighted extreme learning machine," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1252–1266, 2016.

[14] H. Zou, M. Jin, H. Jiang, L. Xie, and C. J. Spanos, "Winips: Wifi-based non-intrusive indoor positioning system with online radio map construction and adaptation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8118–8130, Dec 2017.

[15] Y. Tao and L. Zhao, "A novel system for wifi radio map automatic adaptation and indoor positioning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10 683–10 692, Nov 2018.

[16] M. Zhou, Y. Wang, Y. Liu, and Z. Tian, "An information-theoretic view of wlan localization error bound in gps-denied environment," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2019.

[17] G. Vanderhulst, A. Mashhadi, M. Dashti, and F. Kawsar, "Detecting human encounters from wifi radio signals," in *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*, ser.

MUM '15. New York, NY, USA: ACM, 2015, pp. 97–108.
[18] P. Xu, F. Davoine, J.-B. Bordes, H. Zhao, and T. Denaux, "Multimodal information fusion for urban scene understanding," *Mach. Vision Appl.*, vol. 27, no. 3, pp. 331–349, Apr. 2016.
[19] J. M. Grant and P. J. Flynn, "Crowd scene understanding from video: A survey," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 13, no. 2, pp. 19:1–19:23, Mar. 2017.
[20] C. Nicholson and B. Roebuck, "The investigation of the hillsborough disaster by the health and safety executive," *Safety Science*, vol. 18, no. 4, pp. 249 – 259, 1995, engineering for Crowd Safety.
[21] J. Fruin, "Crowd disasters - a systems evaluation of causes and countermeasures," *Inc. US National Bureau of Standards*, pp. 81–3261, 1981.
[22] M. Wirz, T. Franke, D. Roggen, E. Mitleton-Kelly, P. Lukowicz, and G. Tröster, "Probing crowd density through smartphones in city-scale mass gatherings," *EPJ Data Science*, vol. 2, no. 1, p. 5, Jun 2013.
[23] C. Andrieu and A. Doucet, "Particle filtering for partially observed gaussian state space models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 64, no. 4, pp. 827–836, 2002.
[24] T. Schon, F. Gustafsson, and P. . Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, July 2005.
[25] P. Bahl and V. N. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, vol. 2, March 2000, pp. 775–784 vol.2.
[26] M. Traunmueller, N. Johnson, A. Malik, and C. E. Kontokosta, "Digital traces: Modeling urban mobility using wifi probe data," in *6th International Workshop on Urban Computing, ACM KDD 2017*, 2017.
[27] J. Freudiger, "How talkative is your mobile device?: An experimental study of wi-fi probe requests," in *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, ser. WiSec '15. New York, NY, USA: ACM, 2015, pp. 8:1–8:6. [Online]. Available: http://doi.acm.org/10.1145/2766498.2766517
[28] A. E. C. Redondi, D. Sanvito, and M. Cesana, "Passive classification of wi-fi enabled devices," in *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, ser. MSWiM '16. New York, NY, USA: ACM, 2016, pp. 51–58. [Online]. Available: http://doi.acm.org/10.1145/2988287.2989161
[29] N. Kantas, A. Doucet, S. Singh, and J. Maciejowski, "An overview of sequential monte carlo methods for parameter estimation in general state-space models," *IFAC Proceedings Volumes*, vol. 42, no. 10, pp. 774 – 785, 2009, 15th IFAC Symposium on System Identification.
[30] "Shenzhen Daison Intelligence Technology Co. Ltd." http://www.daison-intelligence.com/.
[31] H. Zhao, B. Huang, and B. Jia, "Applying kriging interpolation for wifi fingerprinting based indoor positioning systems," in *WCNC '16*. IEEE, 2016, pp. 1–6.
[32] B. Huang, Z. Xu, B. Jia, and G. Mao, "An online radio map update scheme for wifi fingerprint-based localization," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6909–6918, 2019.
[33] W. H. Lam and C.-y. Cheung, "Pedestrian speed/flow relationships for walking facilities in hong kong," *Journal of transportation engineering*, vol. 126, no. 4, pp. 343–349, 2000.

Baoqi Huang (S'08-M'12) received the Ph.D. degree in information engineering from the Australian National University, Canberra, Australia, in 2012. From May 2013 to April 2014, he worked as Research Fellow in Nanyang Technological University, Singapore. He is with the College of Computer Science, IMU, where he is currently a professor. He was a recipient of the Chinese Government Award for Outstanding Chinese Students Abroad in 2011. His research interests include wireless networks, sensor networks, mobile computing, ubiquitous computing and etc.

Guoqiang Mao (S'98-M'02-SM'08-F'18) joined the University of Technology Sydney in February 2014 as Professor of Wireless Networking and Director of the Center for Real-time Information Networks. Before that, he was with the School of Electrical and Information Engineering, the University of Sydney. He has published over 200 papers in international conferences and journals, which have been cited more than 7000 times. He is an editor of the IEEE Transactions on Intelligent Transportation Systems (since 2018), IEEE Transactions on Wireless Communications (since 2014), IEEE Transactions on Vehicular Technology (since 2010) and received Top Editor award for outstanding contributions to the IEEE Transactions on Vehicular Technology in 2011, 2014 and 2015. He is a co-chair of IEEE Intelligent Transport Systems Society Technical Committee on Communication Networks. He has served as a chair, co-chair and TPC member in a number of international conferences. He is a Fellow of IET. His research interest includes intelligent transport systems, applied graph theory and its applications in telecommunications, Internet of Things, wireless sensor networks, wireless localization techniques and network modeling and performance analysis.

Yong Qin received the Ph.D. degree from the China Academy of Railway Sciences, Beijing, China, in 1999. He is currently with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. His current research interests include intelligent transportation systems, railway operation safety and reliability, rail network operation management, and traffic model.

Yun Wei received his B.S. from Nanjing University of Aeronautics and Astronautics in 2005, M.S. from Southeast University in 2008, and Ph.D. from Southeast University in 2013. He is the vice director of the Urban Railway Green and Safe Construction National Engineering Laboratory. His research field includes intelligent vision analysis and pattern recognition. He received the Beijing Science and Technology New Star Award and the Exceptional Talent Award of Beijing. He is a member of the Comprehensive Intelligent Transportation Systems Technical Committee of Chinese Association of Automation.