IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 6, 2017, accepted August 14, 2017, date of publication August 29, 2017, date of current version September 19, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2744263

Mining Customer Preference in Physical Stores From Interaction Behavior

YUANYI CHEN[1,2], ZENGWEI ZHENG[1], SINONG CHEN[1], LIN SUN[1], AND DAN CHEN[1]
[1]Department of Computer Science and Technology, Zhejiang University City College, Hangzhou 310058, China
[2]Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Zengwei Zheng (zhengzw@zucc.edu.cn)

This work has been supported by Hangzhou Key Laboratory for IoT Technology & Application.

ABSTRACT An improved understanding of customer preference is crucial for successful business in physical stores. Online stores are capable learning customer preference from the click logs and transaction records, while retailers with physical store still lack effective methods to in-depth understand customer preference. Fortunately, user-generated data from mobile devices and social media are providing rich information to uncover customer preference. In this paper, we present a novel approach to mine customer preference in physical stores from their interaction behaviors. To demonstrate the utility of the proposed model, we conduct a store-type recommendation model for physical stores by jointly considering the learned customer preference and temporal influence. We have performed a comprehensive experiment evaluation on two real-world data sets, which are collected by more than 120 000 customers during 12 months from two urban shopping malls. Experimental results show the superiority of the proposed model not only in recommending interesting stores for customer, but also help retailers better understand customer preference.

INDEX TERMS Customer preference, interaction behaviour, store-type recommendation, physical stores.

I. INTRODUCTION

As the main retail business form in many cities, urban shopping centers play a significant role in maintaining vibrant economies, offering employment and providing a better quality of life. Recent years have witnessed a rapid development of urban shopping centers, for instance, more than 1,800 new shopping centers containing approximately 79.5 million m^2 will be added to the global inventory from 2014 to 2017 according to Cushman & Wakefield.[1] Nowadays, successful urban shopping centers not only need to offer retail shops but also provide various leisure and food facilities (e.g., cafes, game centers and theaters) for improving customer's comprehensive experience.

Given increasing number of homogeneous shopping centers, shopping center developers and retailers who in-depth understand customer's behavior will gain advantages to build excitement with customers and benefit a few services. More exactly, the benefit includes three sides: 1) For customer side, some context-aware personal services (e.g., personalized recommendation, optimal shopping route and targeted advertising) can be provided based on their shopping preferences mining from physical analytics; 2) For shop owner side, physical analytics are beneficial to targeted mobile advertising since the potential consumers can be found based on their shopping preferences; 3) For shopping mall manager side, physical analytics can real-time monitor people flow and discovery some correlations between shops and consumers, these information is useful for optimizing shopping mall layout. However, customer's behavior in urban shopping mall is little understood due to the following challenges:

- customer's behavior are personalized and diversified in such a complicated environment. For example, some customers visit to shopping centers for buying products that they need, another customers focus on the atmosphere and to enjoy the environment of shopping centers. As a result, it is quite challenging to discover customer's preference among among personalized customer behavior.
- customer's behavior in urban shopping mall a result of both personality and situational inuences [24]. Traditional marketing research has reported various factors have impact on customer behavior, such as demographic factors (e.g., gender [8], age [1] and ethnicity [10]) and the shopping center environment (e.g., background music [30], light and employee [15]). Therefore, the various factors that affect customer's behavior cannot be easily represented in a uniform feature space.

[1]http://www.cushmanwakefield.com/en/research-and-insight/2014/global-shopping-center-development-report-spring-2014/

Recently, a few studies (for a review see Section II) have been proposed to uncover customer preference in physical stores. The studies [1], [8], [15], [24], [30] from marketing research aimed to discover customer preference based on intercept surveys from small populations, thus are limited to scalability and labor-intensive. The literatures [18]–[20], [23], [25] uncovered customer preference from their interaction behaviours with physical stores using ubiquitous computing technologies (e.g., camera [18], Radio Frequency Identification [23] and smart glasses [20]). A major drawback of this kind of method is the huge cost to deploy store infrastructure for generating customer's interaction behaviours. With the development of Location Based Social Network (LBSN), a lot of techniques [6], [13], [37], [39] are proposed to uncover customer's preference by utilizing their check-in records in LBSN, but they cannot make recommendations for people who are not members of LBSN and will suffer data sparsity due to few check-ins.

Motivated by recent studies in LBSN [6], [13], we aim to understand customer's preference from their interaction behaviours (such as the residence time and check-in counts of visiting shop). In this study, we focus on mining customer preference in physical stores from their interaction behaviours. The idea behind our approach is customer's interaction behaviours imply their preference, as most people have a finite amount of resources (e.g., time and money), they tend to visit a store by matching their personal preference. To demonstrate the utility of the proposed model, we conduct a store-type recommendation model for physical stores by two phases: 1) offline modeling customer's preference. The offline modeling phase is designed to learn the preference of a customer towards a store based on his/her interaction behaviours. The principle is customer's interaction behaviours is motivated by customer's intention and preference, thus we model the preference of a customer towards a store as the hidden factor for his/her interaction behaviour with a latent variable model. 2) online recommendation. The online recommendation phase automatically produces top-k recommended stores by jointly considering the learnt preference and the temporal influence.

The remainder of the paper is organized as follows: Section II surveys related work on mining customer preference in physical stores. Section III describes the proposed store-type recommendation model in detail. Section IV reports and discusses the experimental results. Section V discusses the other points related to the proposed model. Finally, we present our conclusion and future work in Section VI.

II. RELATED WORK

In this section, we survey related works on mining customer preference in physical stores, including existing studies of mining customer preference using intercept surveys from traditional market research and using customer's interaction behaviours.

A. MINING CUSTOMER PREFERENCE USING INTERCEPT SURVEYS

In the marketing domain, it is of great interest to build a satisfactory relation with the customer, by assessing her/his preference and intention. For example, [3] divided consumers into four categories according to their shopping behavior: enthusiasts, traditionalists, grazers, and minimalists. Several studies [15], [30] analyzed the relationship between customer behavior and demographic factors, such as gender, age and ethnicity. The influence of situational factors (such as background music and shop brand) on customer behaviour are investigated in [1], [8], and [24]. The literatures in [8] and [30] aimed to analyze the influence of situational factors (e.g., background music, shop brand, and billboard image) on customer behaviour in physical stores. However, almost all studies from traditional marketing research utilized intercept surveys to collect customer's profiles and behavior information for inferring customer's preference, which is time-consuming and labor intensive. Moreover, intercept surveys are powerless to capture information from survey avoiders for inferring their preference.

TABLE 1. Compare of experimental datasets.

Methods	Participants	Collecting Time
[9]	21817	7 days
[25]	10	7 days
[19]	25	unknown
[18]	20	5 hours
[23]	10+	unknown
[11]	195	30 days
Our datasets	123,406	12 months

B. MINING CUSTOMERS PREFERENCE USING INTERACTION BEHAVIOUR

Recently, a few studies [9], [11], [18]–[20], [23], [25] have been proposed to mine customers preference from using their interaction behaviours. Typically, customer's interaction behaviours are extracted from shopping trajectories using various ubiquitous computing technologies, such as camera [18], RFID [23] and smart glasses [20]. However, these hardware-based technologies are lack of scalability (e.g., the comparison of experimental datasets in these literatures as shown in Table 1) due to the high cost of deployment and maintenance. On the contrary, WiFi-based technology is a promising method to collect customer's interaction behaviour since WLAN is available in most urban shopping mall. For instance, the work [14] considered customer's stay time in a store can reflect his/her preference towards this store to some extent. Fang *et al.* [5] estimated customer preference by linearly fusing three factors: residence time in a shop, check-in frequency, and matching between promotional activities and customer preference towards promotional activities. The work in [29] aimed to analyze customer behavior using an opt-in WiFi service, which first obtains customers behavior using location tracking and analytics technology based on WiFi access points (APs) that İlistenİ to transmissions from WiFi-enabled devices. Then, they map concepts of customer behavior (e.g., the residence time in a shop and check-in frequency, etc.) to concepts and key performance indicators commonly used in online store analytics, and finally using some marketing management technologies to learn customer preference.

Most of existing studies in POI recommendation [6], [13], [37], [39] in LBSN utilized user's check-in records to learn users preference in recent years. For instance, Zhang and Wang [35] recommended POIs by predicting POI rating using both POI content and POIs location. The work [33] proposed a time-aware location recommendation by considering the temporal pattern during POIs visiting. In [32], the author mined user preference by jointly considering user intrinsic interest and temporal context. Unfortunately, learning customer preference in shopping mall using existing POI recommendation methods will suffer the following three challenges: 1) it is insufficient to learn customer preference by only using the check-in records, other context information can help better reflect the level of customers interests; 2) cold-start problem in POI recommendation due to few check-ins; 3) existing POI recommendation methods can not learn preference for customers who are not members of the LBSN.

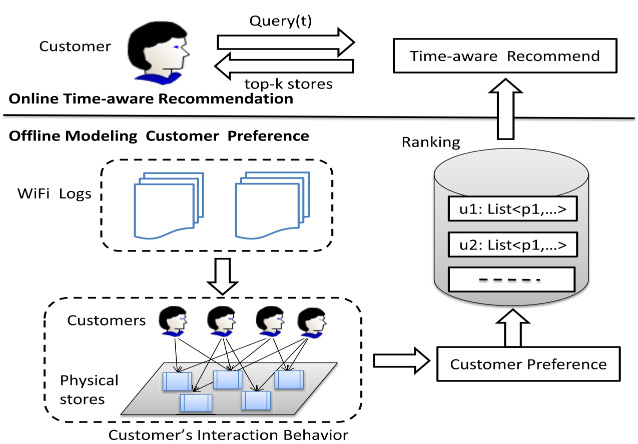


FIGURE 1. The architecture framework of the proposed recommendation model.

III. STORE-TYPE RECOMMENDATION MODEL FOR PHYSICAL STORES

As is shown in Figure 1, the proposed recommendation model produces top-K recommended stores for a target customer by two phases: 1) offline modeling customer preference from their interaction behaviours, which is extracted from WiFi logs that collected by passive crowdsourcing; 2) online recommendation. We make top-K recommended stores by jointly considering the learnt customer preference and temporal influence.

A. PRELIMINARY

For ease of the following presentation, we define the key data structures and notations used in the proposed model.

Definition 1 (WiFi Log): A WiFi log is a set of scanned WiFi records and denote by $S = \{s_1, \ldots, s_i, \ldots\}$, s_i is a triple $< u, t_i, R_i >$ which means the RSS sample R_i is collected by customer u at time t_i, where $R_i = (r_i^1, r_i^2, \ldots r_i^K)$ is a K-dimension vector which means the RSS values collected from surrounding WiFi APs, K is the number of WiFi APs.

Definition 2 (Shopping Trajectory): A shopping trajectory is a sequence of stores that are consecutively visited by a customer, denote by $L =< l^1 \longrightarrow l^2 \longrightarrow \ldots >$, $l^i =< u, p, t_s, t_e >$ is a 4-tuple, t_s and t_e is the start time and end time for visiting store p.

Definition 3 (Interaction Behaviour): Interaction behaviour demotes by a 4-tuple $< u, p, ts, cst >$, which means customer u visits store p at time slot ts, and cst is the residence time of this visit.

Definition 4 (Customer Preference): Customer preference $I^{(up)}$ indicates the interest of customer u towards store p.

B. EXTRACTING INTERACTION BEHAVIOUR FROM WiFi LOGS

As mentioned above, our approach utilizes existing WLAN infrastructure to collect customer's interaction behaviour from WiFi logs, which is infrastructure-free and no customer involvement. More exactly, we utilize WiFi probe requests collect customer's WiFi logs with a non-intrusive way, which has been utilized in many successful applications, such as passive localization [16] and building facility planning [22]. Mobile phones will broadcast WiFi probe requests every few seconds as reported in [16]. Therefore, using WiFi probe requests to collect customer's WiFi logs can track all mobile devices that connect to WLAN. In our experiment, every device that connects to WLAN at each store has agreed to this data collection as part of the sign-on agreement. For privacy issue, we collect customer's interaction behaviour as hashed entities with no additional knowledge about them, and finish collecting data when customer leave the shopping mall. We believe that this is a privacy-safe application.

For generating customer's interaction behaviour, we need to map the WiFi logs to the corresponding shopping trajectories (as shown in Figure 2). The problem can be described as: given a WiFi log $S = \{s_1, \ldots, s_i, \ldots\}$, generate the corresponding shopping trajectories $L(S) =< l^1 \longrightarrow \ldots \longrightarrow l^i \longrightarrow \ldots >$. We utilize fingerprint-based localization [38] to map WiFi logs to shopping trajectories. For constructing the location fingerprint map, we develop a mobile application to collect 100 WiFi RSS samples with sample rate 1 Hz in each store. After mapping all elements $s_i \in S$ to the corresponding store, we construct the shopping trajectories $L(S)$ based on chronological order and further extract customer's check-in activity according to *Definition 3*.

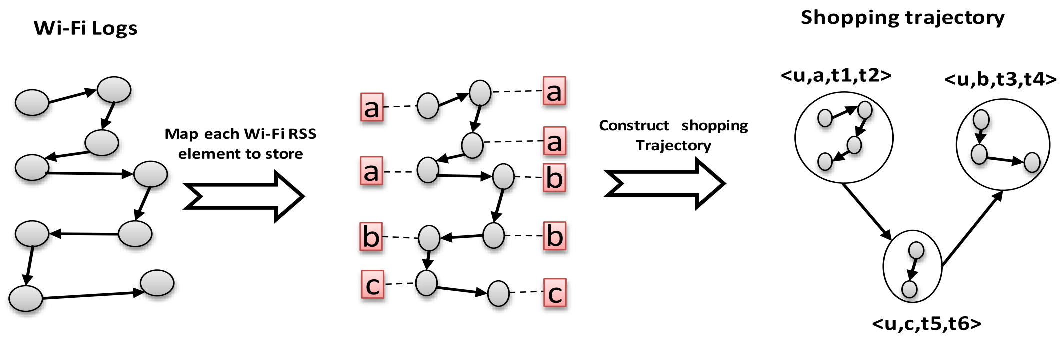


FIGURE 2. Mapping WiFi logs to shopping trajectories.

C. OFFLINE MODELING CUSTOMER PREFERENCE

The principle underlying our model is customer's preference to a store directly impacts the check-in frequency and residence time, which has theoretical foundations and empirical evidence from traditional marketing research [4]. The reason is people have a finite amount of resources (e.g., money and time) for shopping, they tend to visit store by matching their personal preference. In this way, we model the preference of a customer towards a store as the hidden factor of his/her interaction behaviours.

Formally, let $Y_1^{(ij)}$ and $Y_2^{(ij)}$ denote the check-in frequency and average residence time of customer u_i to store p_j, $I^{(ij)}$ denote the preference of u_i to p_j. Then, we utilize a graphical model to combine the influence of u_i and p_j to $I^{(ij)}$, as well as the influence of $I^{(ij)}$ to $Y_1^{(ij)}$ and $Y_2^{(ij)}$, as shown in Figure 3. The detailed description of variables in this figure is explained as follows:

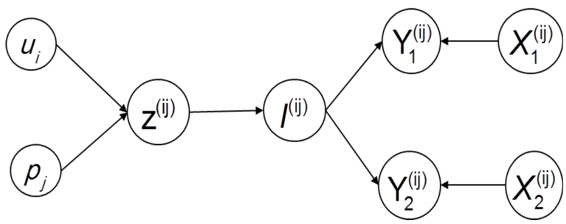


FIGURE 3. Graphical model of learning customer preference.

- $z^{(ij)}$ denote the intrinsic preference of u_i to p_j, which is a result of both personality and situational factors. Since the intrinsic preference is implicit and influenced by various factors, it is difficult to directly infer $z^{(ij)}$. In this way, we capture customer's intrinsic preference based on the widely used location co-occurrence in LBSN [6], [13]. Specifically, let $c < u_i, p_j >= 1$ if u_i has visited store p_j, and $c < u_i, p_j >= 0$ otherwise. Then, we construct customer's check-in vector as:

$$c(u_i) = \{c < u_i, p_1 >, \ldots, c < u_i, p_N >\} \tag{1}$$

Then the intrinsic preference $z^{(ij)}$ between u_i and p_j can be calculated by:

$$z^{(ij)} = \frac{\sum_{v \in U} sim(u_i, v) * c < v, p_j >}{\sum_{v \in U} sim(u_i, v)} \tag{2}$$

where $sim(u_i, v)$ is the similarity between customer u_i and customer v, and estimated by cosine similarity between $c(u_i)$ and $c(v)$.

- $I^{(ij)}$ is the preference between customer u and store p, which is a hidden factor for customer's check-in activities and influenced by customer's intrinsic preference.
- $x_1^{(ij)}$ and $x_2^{(ij)}$ are two auxiliary variables for check-in frequency and residence time, respectively. For example, the total number of stores that a customer has visited, or the average residence time of a customer to all shops. The auxiliary variables capture the tendency of a customer to shopping, which can moderate the effect of customer's preference on interaction behaviours.

Our model represents the relationships among these variables by modeling the conditional dependencies as shown in Figure 3, so the joint distribution decomposes as follows:

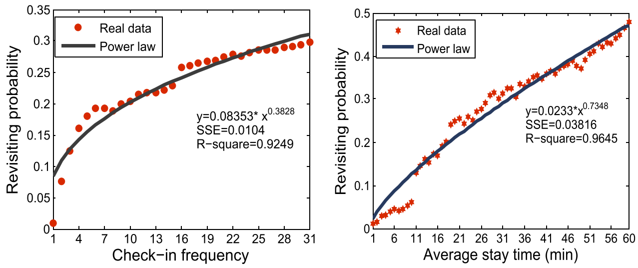
$$P(I^{(ij)}, Y_1^{(ij)}, Y_2^{(ij)} | u_i, p_j) = P(I^{(ij)} | u_i, p_j) \prod_{l=1}^{2} P(Y_l^{(ij)} | I^{(ij)}, X_l^{(ij)}) \tag{3}$$

Given the intrinsic preference between customer u_i and store p_j, we model the conditional probabilities $P(I^{(ij)}|u_i, p_j)$ using the widely-used Gaussian distribution:

$$P(I^{(ij)} | u_i, p_j) = (\eta z^{(ij)}, \sigma^2) \tag{4}$$

where η is a coefficient and σ^2 is the variance of Gaussian model, which is set to 0.5 in experiments.

For model the dependency between $Y_l^{(ij)}$ and $I^{(ij)}$, $X_l^{(ij)} (l = 1, 2)$, we study two anonymized datasets that consists of more than 3 million interaction behaviours from 123,406 customers, more details of the datasets are shown in Table 3. Figure 4a shows customer's revisiting probability to a store as a function of the check-in frequency. From this figure, we observe that: 1) over 16% of customers will revisit a store if they have visited the store more than 4 times;



(a) (b)

FIGURE 4. Fraction of revisiting probability as a function of check-in frequency(a) and average stay time(b). (a) Check-in frequency. (b) Average stay time.

2) the distribution follows a roughly power-law form. Figure 4b shows customer's revisiting probability to a store as a function of the average stay time. We also observe the distribution follows a roughly power-law form, which is also reported in [7]. In this way, we model the dependency between $Y_l^{(ij)}$ and $I^{(ij)}$, $X_l^{(ij)}(l = 1, 2)$ as follows:

$$P(Y_l^{(ij)}|I^{(ij)}, X_l^{(ij)}) = (\alpha_l I^{(ij)} + \beta_l X_l^{(ij)})^{\theta_l} \tag{5}$$

where α_l and β_l are the coefficients, θ_l is the parameter of power law distribution, $l = 1, 2$.

We further add L_2 regularizes on these hyper parameters (e.g., $\alpha_1, \beta_1, \theta_1$) to avoid over-fitting, which can be regarded as Gaussian prior:

$$\begin{aligned} P(\alpha_l, \beta_l) &\propto e^{-(\lambda_l/2)(\alpha_l^2+\beta_l^2)}, \quad l = 1, 2 \\ P(\theta_l) &\propto e^{-(\lambda_{\theta_l}/2)(\theta_l)^2}, \quad l = 1, 2 \\ P(\eta) &\propto e^{-(\lambda_\eta/2)\eta^2} \end{aligned} \tag{6}$$

The data are represented as $\Phi = U \times P$ samples of *customer-store* pairs, denoted as $D = \{(i_1, j_1), \ldots, (i_N, j_M)\}$. During training phase, the variables $z^{(ij)}, Y_1^{(ij)}, Y_2^{(ij)}, X_1^{(ij)}$ and $X_2^{(ij)}$ are all visible, $(i, j) \subseteq \Phi$. According to Equation 3, given all the observed variables, the joint probability is shown as:

$$\begin{aligned} &\prod_{l=1}^{2} P(\Phi|\eta, \alpha_l, \beta_l, \theta_l)P(\eta, \alpha_l, \beta_l, \theta_l) \\ &\quad = \prod_{(i,j)\in D} P(I^{(ij)}|z^{(ij)}, \eta)P(\eta) \\ &\quad\quad \times \prod_{l=1}^{2} P(D|I^{(ij)}, X_l^{(ij)}, \alpha_l, \beta_l, \theta_l)P(\alpha_l, \beta_l, \theta_l) \\ &\quad \propto \prod_{(i,j)\in D} \left(e^{-(1/2\delta^2)(\eta z^{(ij)} - I^{(ij)})^2} \prod_{l=1}^{2} (\alpha_l I^{(ij)} + \beta_l X_l^{(ij)})^{\theta_l} \right) \\ &\quad\quad \times e^{-(\lambda_\eta/2)\eta^2} \prod_{l=1}^{2} e^{-(\lambda_{\theta_l}/2)(\theta_l)^2} e^{-(\lambda_l/2)(\alpha_l^2+\beta_l^2)} \end{aligned} \tag{7}$$

We maximize the likelihood function as shown in Equation 7 to estimate the unknown model parameters $\Sigma = \{\eta, \alpha_l, \beta_l, \theta_l | l = 1, 2\}$. We set $\sigma^2 = 0.5$ as suggested by [28]. As for the hyper parameters $\lambda_\eta, \lambda_{\theta_l}, \lambda_l$, we select the optimal values ($\lambda_1 = \lambda_2 = \lambda_\eta = \lambda_{\theta_1} = \lambda_{\theta_2} = 0.01$) by conducting grid search and 5-fold cross-validation.

Applying a logarithmic transformation to both sides of Equation 7, we obtain the following expression:

$$\begin{aligned} &L((i, j) \in D, \eta, \alpha_l, \beta_l, \theta_l) \\ &\quad = \sum_{(i,j)\in D} -\frac{1}{2\sigma^2}(\eta z^{(ij)} - I^{(ij)})^2 \\ &\quad\quad + \sum_{(i,j)\in D} \sum_{l=1}^{2} \theta_l log(\alpha_l I^{(ij)} + \beta_l X_l^{(ij)}) \\ &\quad\quad - \frac{\lambda_\eta}{2}\eta^2 - \sum_{l=1}^{2} \frac{\lambda_{\theta_l}}{2}\theta_l^2 - \sum_{l=1}^{2} \frac{\lambda_l}{2}(\alpha_l^2 + \beta_l^2) \end{aligned} \tag{8}$$

Note the function L (see in Equation 8) is concave, then we optimize the parameters $\eta, \alpha_l, \beta_l, \theta_l$ and variable $I^{(ij)}$ with a stochastic gradient descent algorithm. The coordinate-wise gradients are:

$$\begin{aligned} \frac{\partial L}{\partial I^{(ij)}} &= \frac{1}{\sigma^2}(\eta z^{(ij)} - I^{(ij)}) + \sum_{l=1}^{2} \frac{\theta_l \alpha_l}{\alpha_l I^{(ij)} + \beta_l X_l^{(ij)}} \\ \frac{\partial L}{\partial \eta} &= -\frac{1}{\sigma^2} \sum_{(i,j)\in D} z^{(ij)}(\eta z^{(ij)} - I^{(ij)}) - \lambda_\eta \eta \\ \frac{\partial L}{\partial \alpha_l} &= \sum_{(i,j)\in D} \frac{\theta_l I^{(ij)}}{\alpha_l I^{(ij)} + \beta_l X_l^{(ij)}} - \lambda_l \alpha_l \\ \frac{\partial L}{\partial \beta_l} &= \sum_{(i,j)\in D} \frac{\theta_l X_l^{(ij)}}{\alpha_l I^{(ij)} + \beta_l X_l^{(ij)}} - \lambda_l \beta_l \\ \frac{\partial L}{\partial \theta_l} &= \sum_{(i,j)\in D} log(\alpha_l I^{(ij)} + \beta_l X_l^{(ij)}) - \lambda_{\theta_l} \theta_l \end{aligned} \tag{9}$$

We use a coordinate ascent optimization scheme to update $\eta, \alpha_l, \beta_l, \theta_l$ and $I^{(ij)}$. More exactly, using a Netwton-Raphson scheme to update these parameters in each iteration:

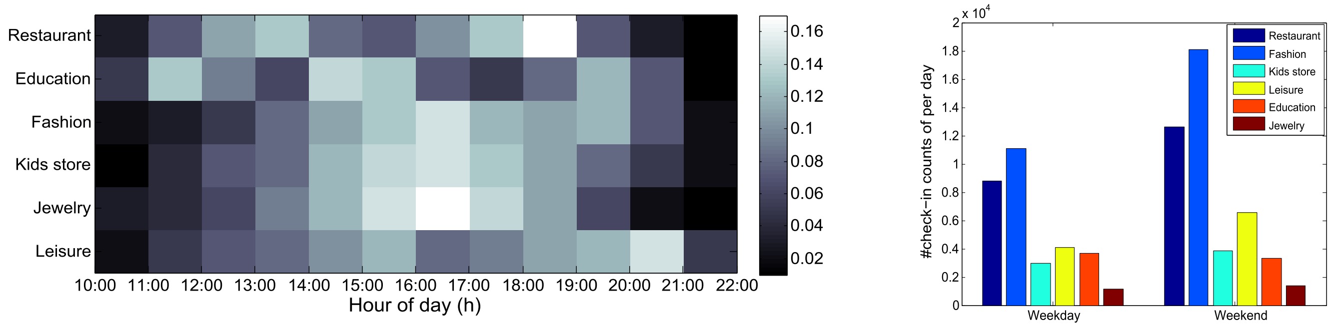
$$I^{(ij)new} = I^{(ij)old} - \frac{\partial L}{\partial I^{(ij)}} / \frac{\partial^2 L}{\partial (I^{(ij)})^2} \tag{10}$$

$$\eta^{new} = \eta^{old} - \frac{\partial L}{\partial \eta} / \frac{\partial^2 L}{\partial (\eta)^2} \tag{11}$$

$$\alpha_l^{new} = \alpha_l^{old} - \frac{\partial L}{\partial \alpha_l} / \frac{\partial^2 L}{\partial (\alpha_l)^2} \tag{12}$$

$$\beta_l^{new} = \beta_l^{old} - \frac{\partial L}{\partial \beta_l} / \frac{\partial^2 L}{\partial (\beta_l)^2} \tag{13}$$

$$\theta_l^{new} = \theta_l^{old} - \frac{\partial L}{\partial \theta_l} / \frac{\partial^2 L}{\partial (\theta_l)^2} \tag{14}$$



(a) (b)

FIGURE 5. The hourly (a) and daily (b) distribution of customer's check-in activities. (a) Hourly distribution. (b) Daily distribution.

Where the second order derivatives are given by:

$$\begin{aligned}
\frac{\partial^2 L}{\partial (I^{(ij)})^2} &= -\frac{1}{\sigma^2} - \sum_{l=1}^{2} \frac{\theta_l \alpha_l^2}{(\alpha_l I^{(ij)} + \beta_l X_l^{(ij)})^2} \\
\frac{\partial^2 L}{\partial (\alpha_l)^2} &= -\lambda_l - \sum_{(i,j)\in D} \frac{\theta_l (I^{(ij)})^2}{(\alpha_l I^{(ij)} + \beta_l X_l^{(ij)})^2} \\
\frac{\partial^2 L}{\partial (\beta_l)^2} &= -\lambda_l - \sum_{(i,j)\in D} \frac{\theta_l (x_l^{(ij)})^2}{(\alpha_l I^{(ij)} + \beta_l X_l^{(ij)})^2} \\
\frac{\partial^2 L}{\partial (\theta_l)^2} &= -\lambda_2
\end{aligned} \tag{15}$$

Algorithm 1 The Algorithm for Optimizing Parameters

Require: Data samples $D = \{(u_1, p_1), \ldots, \ldots, (u_N, p_M)\}$.
Ensure: Model parameters $\Sigma = \{\eta, \alpha_l, \beta_l, \theta_l | l = 1, 2\}$.
1: Generate customer check-in vectors according to Equation 1.
2: Calculate the intrinsic preference between customers and stores according to Equation 2.
3: **while** not converged **do**
4: **for** each Newton-Raphson step **do**
5:     **for** $(i, j) \in D$ **do**
6:         Update $I^{(ij)}$ according to Equation 10.
7:     **end for**
8:     **for** $l = 1, 2$ **do**
9:         Update $\alpha_l, \beta_l, \theta_l$ according to Equation 12 $\sim$ 14.
10:     **end for**
11: **end for**
12: Update η according to Equation 11.
13: **endwhile**
14: **return** $\Sigma = \{\eta, \alpha_l, \beta_l, \theta_l | l = 1, 2\}$.

Algorithm 1 shows the learning procedure for optimizing the parameters. First, as shown in Lines 1 $\sim$ 2, we calculate the intrinsic preference between customers and stores using location co-occurrence from their check-in activities. Then, as depicted in Line 3 $\sim$ 13, we optimize model parameters $\Sigma = \{\eta, \alpha_l, \beta_l, \theta_l | l = 1, 2\}$ using Newton-Raphson until converged.

D. ONLINE TIME-AWARE RECOMMENDATION

Mining customer preference in physical stores can facilitate a few personalized applications (e.g., store-type recommendation, detected potential customers and targeted advertising). Due to space constraints, we briefly introduce an important application: top-K store recommendation. More exactly, given historical interaction behaviours of a group of customers, the goal is to recommend top-K stores with the maximum visiting probability for customers.

Formally, let M be the number of customers and N the number of stores, the preference database is represented by a $N \times M$ preference matrix R, each row of R represents a customer, each column of R represents an store, each element r_{ij} of R represents the preference of customer i towards store j. Then, the recommendation problem can be formulated as inferring missing values of a partially observed Customer-Store preference matrix R. Traditionally, the Regularized Singular Value Decomposition [17] model is employed to predict missing values. The basic idea is using low-rank matrix factorization approach seeks to approximate the preference matrix R by a multiplication of f-rank factors $R = U^T V$, where $U \in R^{f \times M}$ and $V \in R^{f \times N}$. The objective function is equivalent to minimizing the sum of squared errors with quadratic regularization terms as follows:

$$L = \min_{U,V} \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} c_{ij}(r_{ij} - u_i^T v_j)^2 + \frac{\lambda_u}{2} ||U||_F^2 + \frac{\lambda_v}{2} ||V||_F^2 \tag{16}$$

where u_i and v_j are column vectors with f values, c_{ij} is the indicator function that is equal to 1 if customer i visited store j and equal to 0 otherwise, λ_u, λ_v represent the regularization parameters, and $||.||_F$ is the Frobenius norm of matrices.

However, human daily activities usually follow a regular temporal pattern [27], i.e., people usually eat dinner at 17:00-19:00, which means a customer may tend to visit a restaurant rather than other kinds of stores during the time slot. To show the temporal pattern of

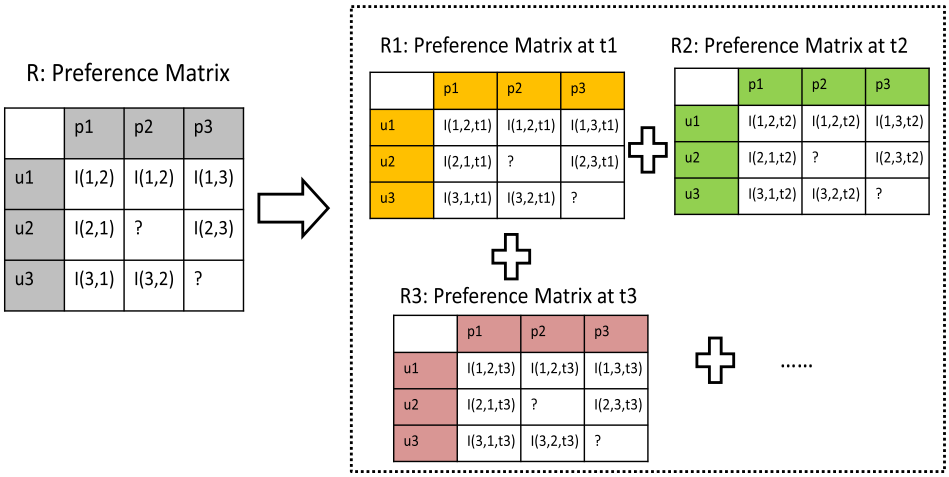


FIGURE 6. **Partition customer preference matrix *R* into a series of preference matrices at different timeslots.**

customer's check-in activity in physical stores, we analyze the hourly and daily distribution using millions of interaction behaviours, more details of this dataset are shown in Table 3. Figure 5a shows the hourly distribution of customer's interaction behaviours, we observe different kinds of stores have different visiting temporal pattern. For instance, customers prefer to check-in a restaurant at noon (12:00-14:00) and dinner time (17:00-19:00), the probability of check-in stores that belong to leisure (such as watch movie or play games) at night is greater than in the day. Figure 5b shows the daily check-in distribution, we observe the probability of check-in stores that belong to restaurant and fashion at the weekend is greater than at weekday. The reason is that people usually have more time to dinner together and go shopping on weekend. From the two figures, we can see that temporal influence plays an important role in mining customer's interaction behaviours, which should be considered in making recommendation.

To extract the temporal pattern of customer's interaction behaviours, we divide days into two categories: Weekday and Weekend, and further divide a day 12 hourly slots (since the operation hours of the shopping mall are 10:00 am-10:00 pm). To this end, we generate the total number of hashed time slots is 24, denote as $T = \{Weekday\_1, Weekday\_2, \ldots, Weekend\_12\}$. For instance, if a customer visited a restaurant at 1:12 pm, 3/15/2016, the time slot of this visit is $Weekday_3$. We partition the total interaction behaviours into a few subsets according to the store category and time slot of check-ins, and each subset represents customer's interactions for a certain store category at a specific time slot. Then, we calculate the check-in probability of customer i to stores that belongs to category ρ_j at time slot t by:

$$Pr(i, \rho_j|t) = \frac{\psi(i, \rho_j|t)}{\psi(i, t)} \tag{17}$$

where ρ_j is the category of store j, $\psi(i, \rho_j|t)$ is the check-ins for stores belong to ρ_j at time slot t. Accordingly, $\psi(i, t)$ is the total check-ins for all stores at time slot t.

To fuse the temporal influence of customer's interaction behaviours, we partition the raw Customer-Store preference matrix R into a series of preference matrix $\{R_1, R_2, \ldots, R_{24}\}$ at different timeslots (as shown in Figure 6), and each element r_{ij} of R_t represents the preference of customer i towards store j at timeslot t, as defined:

$$r_{ijt} = Pr(i, \rho_j|t) * I^{(ij)} \tag{18}$$

Then, we utilize a time-aware matrix factorization to predict missing values of R_t where the objective function is equivalent to minimizing the sum of squared errors with quadratic regularization terms as follows:

$$L = \min_{U,V} \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{t=1}^{24} c_{ijt}(r_{ijt} - u_{it}^T v_j)^2 + \frac{\lambda_u}{2}||U_t||_F^2 + \frac{\lambda_v}{2}||V||_F^2 \tag{19}$$

where c_{ijt} is the indicator function that is equal to 1 if customer i visited store j at timeslot t and equal to 0 otherwise, and λ_u, λ_v represent the regularization parameters.

By adopting a stochastic gradient descent algorithm, for each observed preference r_{ijt}, we have the following efficient updating rules to learn latent variables u_{it}, v_j:

$$u_{it} \leftarrow u_{it} + \gamma_1(e_{ijt} v_j - \lambda_u u_{it}) \tag{20}$$

$$v_j \leftarrow v_j + \gamma_2(e_{ijt} u_{it} - \lambda_v v_j) \tag{21}$$

where $e_{ijt} = r_{ijt} - u_{it}^T v_j$ and γ_1, γ_2 are learning rate.

Given customer i and a unvisited store $\widehat{j}$ at time slot t, we calculate the recommendation score using the learnt latent variables u_{it}, v_j:

$$score(i, \widehat{j}, t) = u_{it}^T v_{\widehat{j}} \tag{22}$$

IV. EXPERIMENT EVALUATION

In this section, we report on the results of a series of experiments conducted to evaluate the performance of the proposed model to recommend top-K stores to customers. We first

TABLE 2. Statistics of store categories.

	Restaurant	Fashion	Kids store	Leisure	Education	Jewellery
Mall 1	69	56	26	23	15	19
Mall 2	37	46	13	18	12	8

describe the settings of experiments including data sets, comparative algorithms and evaluation metric. Then, we report and discuss the experimental results.

A. EXPERIMENTAL SETTINGS

1) DATA SETS

Our experimental environment is two inner-city shopping malls: one is with 5 floors and covered over 300,000 m^2, which contains 208 stores and these stores belong to 6 categories given by the mall owner; another contains 3 floors with 134 stores, which also consist of 6 categories of stores. More details of the two shopping malls are shown in Table 2.

As mentioned in Section III-B , We gather two anonymized WiFi logs dataset from registered customers using an opt-in WiFi network in the urban shopping mall during 12 months. For removing noise data, we perform two preprocessing steps: (1) we filter out the mall workers and shop employees based on the check-in frequency. Empirically, we consider a customer as a mall worker or store employee if her/his check-ins are more than 100 during 12 months; (2) we remove the abnormal visiting with the residence time is less than 1 minute. After preprocessing, two datasets consists of 3,860,749 interaction behaviours from 123,406 customers, more details of the dataset are shown in Table 3.

TABLE 3. Statistics of customer's interaction behaviours.

Urban shopping mall	Mall 1	Mall 2
Number of Stores	208	134
Number of Customers	75,541	47,865
Number of Interactions	2,568,394	1,292,355
Average No. of Interactions stores per customer	34	27
Average No. of Interactions customers per store	12348	9644

2) COMPARATIVE ALGORITHMS

We compare the proposed recommendation model with the following five methods, where the first four models are the well-known existing methods for physical store recommendation, and the last model corresponds to the proposed model without fusing temporal influence.

- **Collaborative Filtering based Location Co-occurrence (LCCF).** LCCF [31] calculates the recommendation score of a unvisited store by considering other customer's check-in records on the store. Let $c < u, p >= 1$ if u has visited store p at time slot ts, and $c < u, p >= 0$ otherwise; $c < u, P >= \{c < u, p_1 >, \ldots, c < u, p_N >\}$ is the store check-in vector of customer u. Then, the recommendation score between u and a unvisited store $\widehat{p}$ is calculated by
$$score(u, \widehat{p}) = \frac{\sum_{v \in U} sim(u, v) * c < v, p >}{\sum_{v \in U} sim(u, v)} \quad (23)$$
where $sim(u, v)$ is the similarity between customer u and customer v, and calculated using the cosine similarity between $c < u, P >$ and $c < v, P >$.
- **Time-aware Collaborative Filtering based on Location Co-occurrence (TA-LCCF).** Similar to LCCF, TA-LCCF [33] calculates the recommendation score based on location co-occurrence with fusing temporal influence. Let $c < u, p, ts >= 1$ if u has visited store p at time slot ts, and $c < u, p, ts >= 0$ otherwise; $c < u, P, ts >= \{c < u, p_1, ts >, \ldots, c < u, p_N, ts >\}$ is the store check-in vector of customer u at time slot ts. Then, the recommendation score between u and a unvisited store $\widehat{p}$ at time slot ts is calculated by
$$score(u, \widehat{p}) = \frac{\sum_{v \in U} sim(u, v, ts) * c < v, p, ts >}{\sum_{v \in U} sim(u, v, ts)} \quad (24)$$
where $sim(u, v, ts)$ is the similarity between customer u and customer v at time slot ts, and calculated using the cosine similarity between $c < u, P, ts >$ and $c < v, P, ts >$.
- **Matrix Factorization based on Check-in Frequency (MFCF).** MFCF utilizes the check-in frequency to reflect customer preference. Let M be the number of customers and N the number of physical stores, the preference database is represented by a $N \times M$ preference matrix R. Each element r_{ij} of R represents the check-in frequency of customer i towards store j. Then, the Regularized Singular Value Decomposition [17] model is employed to predict missing values of R for making recommendation.
- **Rule-based recommendation algorithm (RBCA).** RBCA [5] estimates customer preference by linearly fusing three factors: residence time in a store, check-in frequency, and matching between promotional activities and customer preference towards promotional activities. Recommendation rules are extracted according to two assumptions, one is the higher a customer preference towards a store, the more likely he/she is to check-in, another is a customer will enter the store with promotional offers when her preferences towards two stores are the same.
- **Time-based Slope One(TSO).** In [14], customer preference is extracted directly from residence time for a store. Specifically, customer preference is generated by using a logarithmic function to map residence time of a store to

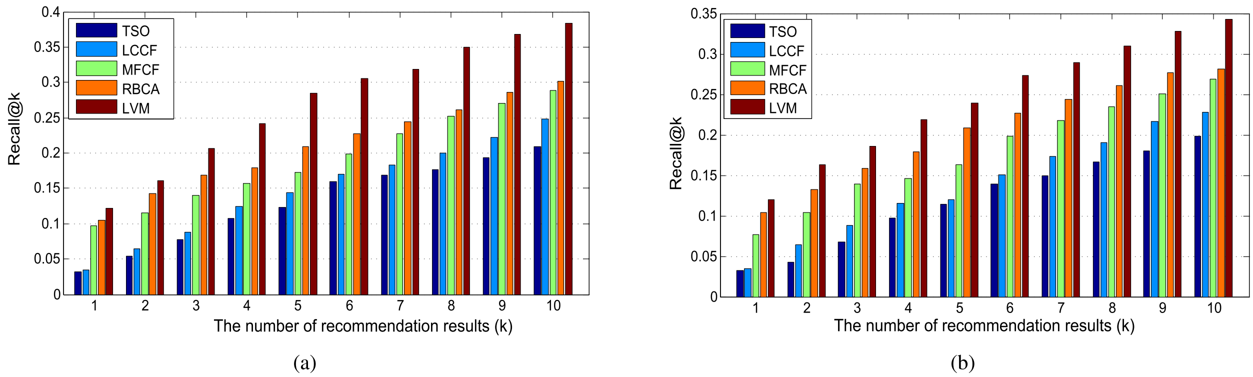


FIGURE 7. Effectiveness of top-k recommendations. (a) Top-k recommendation performance on Mall 1. (b) Top-k recommendation performance on Mall 2.

recommendation score (as shown in Equation 25), then Slope-One [12] is used to make recommendation.

$$I^{up} = min(6, ln(t/10 + 1)) \tag{25}$$

where t is the total time that customer u spent in store p with a unit of a second.

- **Matrix Factorization without fusing temporal influence (LVM)**. As a component of the proposed recommendation model, LVM means our model without fusing temporal influence in online recommendation. Given customer u and a unvisited store $\widehat{p}$ at time slot ts, LVM calculate the recommendation score as:

$$score(i, \widehat{j}) = u_i^T v_{\widehat{j}} \tag{26}$$

where $u + i, v_{\widehat{j}}$ are latent variables that obtained by optimizing the objective Equation 16.

3) EVALUATION METRIC

We provide the experiment results we obtained on the average after tenfold cross-validation. In each experiment, we randomly select 30% of interaction behaviours as the test set D_{te}, and use the rest 70% interaction behaviours as the training set D_{tr}. To evaluate the recommendation effectiveness of our proposed method, we adopt Recall@k as the measurement metric, where k is the number of the recommendation results. For each test case $(u, s_i) \in D_{te}$:

(1). We randomly select 10 stores that unvisited by customer u, and compute the recommendation score for s_i and the additional selected 10 stores;

(2). We form a ranked list by ordering all the 10 stores according to their recommendation scores. Let *ind* denote the rank of the test item s_i within this list;

(3). We form a top-k recommendation list by picking the k top ranked items from the list. If $ind < k$ we have a hit (i.e., the test item s_i is recommended to the customer). Otherwise we have a miss. Clearly, the probability of a hit increases with the increasing value of k. When $k = 101$ we always have a hit.

Let $\#hit@k$ denotes a single test case as either the value 1 if the test item s_i appears in the top-k results, or else the value 0. The overall $Recall@k$ are defined by averaging all test cases:

$$Recall@k = \frac{\#hit@k}{|D_{te}|} \tag{27}$$

where $\#hit@k$ denotes the number of hits in the test set, and $|D_{te}|$ is the number of all test cases.

We conduct three groups of experiments. The first group is to evaluate the performance of the proposed methods utilizing customer's interaction behaviours to learn their preference, and we compare methods LCCF, MFCF, RBCA, TSO and LVM. The second group is to evaluate the recommendation effectiveness by fusing temporal influence, and we compare methods LCCF, TA-LCCF, LVM and TA-LVM (the proposed method). The third group is a case study that was performed for two purposes: 1) to verify the accuracy of extracting customer's interaction behaviour from WiFi logs. Since our method learns customer preference from his/her interaction behaviours, it is useful to show the accuracy of extracting customer's interaction behaviour from WiFi logs; 2) to show the validity of the store recommendation in physical stores. We show it is useful to recommend customers some interesting stores based on their preference, since most customers have no clear idea about what they want to do in a shopping mall [21].

B. EXPERIMENTAL RESULTS

In this subsection, we first report the performance of the proposed model on the recommendation effectiveness and then discuss the temporal influence for different recommendation models.

1) EFFECTIVENESS OF RECOMMENDATIONS

Figure 7 reports the performance of the recommendation models on the dataset. We show only the performance where the length (k) of recommendation list is in the range $[1 \ldots 10]$, because there are 211 stores in total and a greater value of k is usually ignored for a typical top-K recommendation task. It is apparent that these models have significant performance disparity in terms of top-k recall.

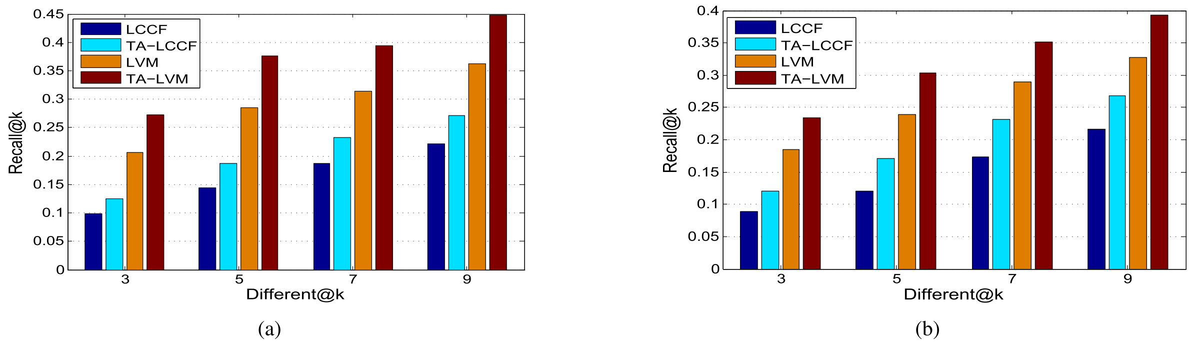


FIGURE 8. **Impact of temporal influence. (a) Different@*k* on Mall 1. (b) Different@*k* on Mall 2.**

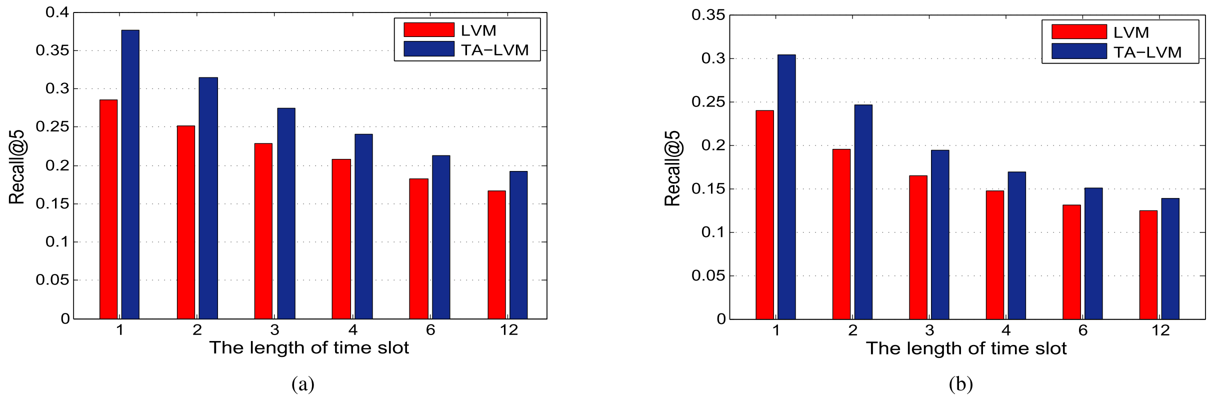


FIGURE 9. **Performance of varying length of time slot. (a) Recall@5 on Mall 1. (b) Recall@5 on Mall 2.**

From this figure, we also observe: (1) LVM always perform better than other competitor recommendation models (LCCF, RBCA, TSO), showing the advantage of using latent variable model to learn customer's preference. For example, the recall of our proposed method with fusing temporal influence (LVM) is about 0.347 for top-8 store recommendation in Mall 1, (i.e., the LVM model has a probability of 34.7% of placing a store within target customer's check-in list in the top-8), while 0.25 for RBCA, 0.247 for MFCF, 0.186 for LCCF and 0.161 for TSO. Similar results are also observed in top-*k* recommendation for Mall 2 (for example, the recall@10 of different methods are 0.342(LVM), 0.27(RBCA), 0.252(MFCF), 0.23(LCCF) and 0.2(TSO)), showing again our proposed model outperforms other competitor recommendation models significantly; (2) TSO performs worst among all recommendation models, which suggests that only utilizing the residence time in a store is insufficient to reflect the level of customer's interests. Similarly, the results of MFCF and LCCF suggest that only utilizing the check-in frequency is also insufficient to reflect the level of customer's interests.

2) IMPACT OF TEMPORAL INFLUENCE

We compare the recommendation effectiveness of two recommendation models (LCCF and LVM) by fusing temporal influence in Figure 8. Clearly, the proposed MallRec and LVM models outperform TA-LCCF and LCCF models significantly. For instance, the Recall@5 of Mall 1 for LCCF 14.3% that drops 14% compare to LVM, while the Recall@7 of Mall 2 for TA-LCCF 23.6% that drops 11.4% compare to TA-LVM. The results suggest that, the proposed models can better uncover customer preference by modeling as a hidden factor of customer's check-in activities with latent variable model. On the contrary, the well-known location co-occurrence in LBSN is not obvious for customer's check-in activities in physical stores. From Figure 8, we can also observe the two models (TA-LCCF and TA-LVM)) by fusing temporal influence perform better than the baseline methods (e.g., for Recall@9 of Mall 1, TA-LCCF and TA-LVM increase 4% and 8.2% performance compare to LCCF and LVM, respectively.), showing temporal influence plays a vital role in mining customer's preference and is vital for store recommendation.

Table 4 presents the Recall@5 of for recommendation models (LCCF,LVM,TA-LCCF and TA-LVM) on weekday and weekend. An interesting observation from this figure is that the advantages by fusing temporal influence on weekend is more significant than on weekday, showing temporal pattern among customer's check-in activities on weekend are more stable and obvious than weekday. For example, the performance improved of Mall 1 with LCCF is 2.4% on weekday with incorporating temporal influence while 5.8%

TABLE 4. Performance comparison: Weekday vs Weekend.

Method	LCCF		TA-LCCF		LVM		TA-LVM	
	weekday	weekend	weekday	weekend	weekday	weekend	weekday	weekend
Mall 1	21.2%	22.3%	24.6%	28.1%	30.2%	31.3%	34.4%	37.7%
Mall 2	17.4%	18.5%	20.7%	24.3%	27.1%	28.7%	30.5%	35.1%

TABLE 5. Performance of different store categories.

Category	Restaurant		Fashion		Kids store		Leisure		Education		Jewelery	
	LVM	TA-LVM	LVM	TA-LVM	LVM	TA-LVM	LVM	TA-LVM	LVM	TA-LVM	LVM	TA-LVM
Mall 1	32.4%	44.1%	31.2%	36.4%	26.1%	30.7%	29.1%	40.6%	27.3%	35.8%	25.4%	30.4%
Mall 2	27.4%	36.2%	25.1%	28.5%	21.4%	23.7%	24.3%	34.4%	23%	32.8%	22.1%	25.1%

on weekend, the performance improved of Mall 2 with LVM is 2.9% on weekday with incorporating temporal influence while 6.4% on weekend.

Figure 9 reports the effect on the length of time slot for the proposed recommendation model (TA-LVM), which controls the time granularity of time-aware recommendations. A larger length of time slot implies that the recommendation results will be less time-specific. From this figure, we can observe the Recall@5 for the shopping malls drop with the time slot length increases. The reason is that increasing the length of time slots will bring in more ground truth stores for a customer at each time slot. Since the length of recommendation list (k) unchanged, the recall will decrease when increasing the length of time slot. In summary, the performance improvement decreases as increasing the length of time slot, since increasing the length of time slot will weaken the advantage by fusing temporal influence.

Table 5 reports the top-5 recommendation performance of the proposed methods (LVM and TA-LVM) for different categories of stores. From this figure, we can observe: (1) the performance for different store categories is significantly disparity for the two models, and TA-LVM achieves better performance for all store categories. For instance, the Recall@5 of TA-LVM increases the performance of 11.7% for Restaurant compares to LVM in Mall 1; (2) the performance improvement by fusing temporal influence is diverse for different store categories. More exactly, the performance improvement for three store categories (Restaurant, Education and Leisure) can reach 8%, while has a small effect for other kinds of stores (Fashion, Kids store and Jewelery). The results suggest that, customer's check-in activities for stores that belong to the three categories (Restaurant, Education and Leisure) have a stronger temporal pattern than the other three kinds of stores (Fashion, Kids store and Jewelery). We further find the average check-in time of the first three kinds of stores are much higher than the second three kinds of stores, justifying that there is a positive correlation between the revisit probability and the average check-in time, which is reported in [4].

3) A CASE STUDY

As mentioned above, this case study was performed for two purposes: 1) to verify the accuracy of extracting customer's interaction behaviour from WiFi logs; 2) to show the validity of store recommendation in physical stores.

a: DATASET

To evaluate the performance, We develop a mobile application to collect WiFi logs with a sampling rate of 0.2 Hz, each RSS record is represented by a tuple: $< s, o >$. Specifically, s is the store and $o = (M, t, R)$, M is the MAC address of collection device and t is the collection time, $R =< r^1, r^2, \ldots, r^K >$ is the scanned RSS record from surround WiFi APs. The pre-defined information includes the check-in time and check-out time of each store, which can be regarded as ground-truth data to evaluate the performance of extracting customer's interaction behaviour. Totally, we collect 65 WiFi logs for experiment evaluation by 8 participants (including 5 males and 3 females) over 2 weeks in Mall 1, in which one WiFi log includes an average of 10 stores and 1742 RSS records.

b: EVALUATION METRIC

We utilize mapping accuracy and trajectory distance to evaluate the performance of extracting customer's interaction behaviour from WiFi logs, and utilize $NDCG@k$ to evaluate the performance of top-k recommendation.

- **Mapping Accuracy.** Let p_i denotes the ground truth store when collecting RSS record R, $\widehat{p_i}$ denote the mapping store from R obtains by fingerprint-based localization, the mapping accuracy is defined as:

$$MA = \frac{\sum_{i=1}^{Te} I(p_i, \widehat{p_i})}{Te} \tag{28}$$

Where $I(p_i, \widehat{p_i})$ is an indicator function that return 1 if $\widehat{p_i} = p_i$, Te is the RSS records for evaluation.

- **Trajectory Distance.** Let $C_{ij} = \{p_1, p_2, \ldots, p_k\}$ is the common store set of shopping trajectory $Traj_i$ and $Traj_j$, then the longest common sub-sequence of $Traj_i$ and $Traj_j$ is defined as Equation 29.

$$LCSS(C_{ij}) = \begin{cases} 0 & \text{k} = 0 \\ LCSS(Rest(C_{ij})) + 1 & \triangle < \theta \end{cases} \tag{29}$$

where $\triangle = |t_i(p_k) - t_j(p_k)|$, $\triangle < \theta$ means the stay time difference of p_k in $Traj_i$ and $Traj_j$ is less than a

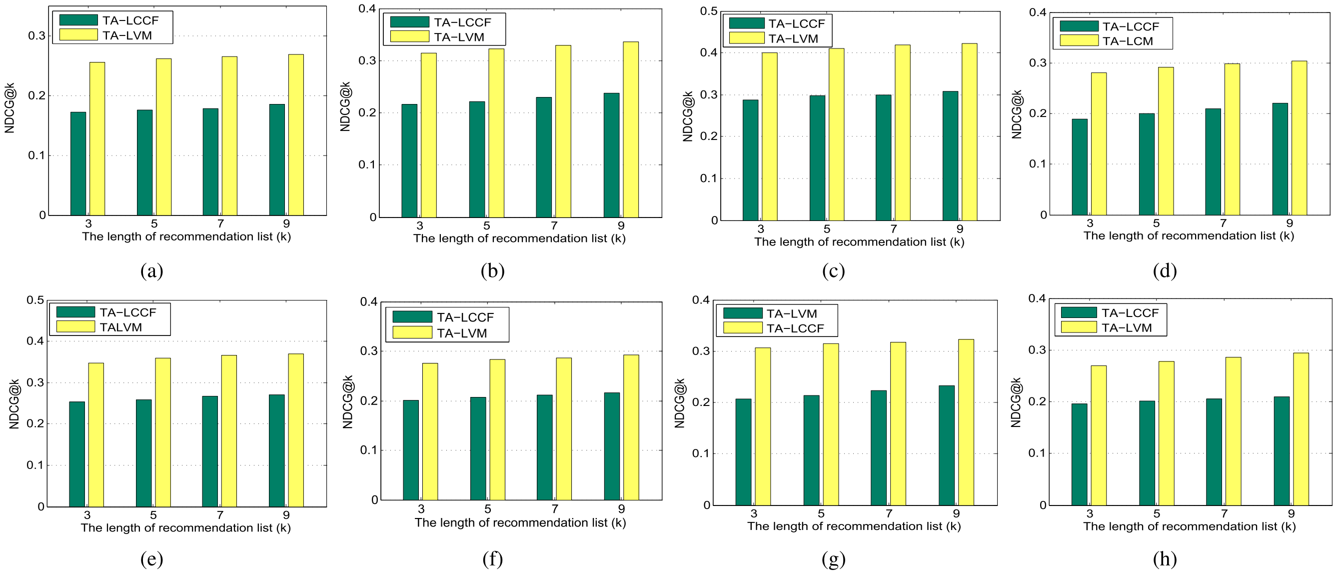


FIGURE 10. NDCG@*k* for each customer. (a) customer 1. (b) customer 2. (c) customer 3. (d) customer 4. (e) customer 5. (f) customer 6. (i) customer 7. (j) customer 8.

threshold, $Rest(C_{ij}) = \{p_1, p_2, \ldots, p_{k-1}\}$. Following the work by [26], we define the distance of $Traj_i$ and $Traj_j$ as Equation 30.

$$Dist(Traj_i, Traj_j) = 1 - \frac{LCSS(C_{ij})}{min\{l_i, l_j\}} \tag{30}$$

where l_i and l_j are the store number of $Traj_i$ and $Traj_j$, respectively.

- **NGCG@k [36].** Let $corr_i$ denotes a relevance value, $NGCG@k$ is calculated as:

$$NGCG@k = \frac{DCG(k)}{IDCG(k)}$$

$$DCG(k) = corr_1 + \sum_{i=2}^{k} \frac{corr_i}{log_2 i} \tag{31}$$

where $IDCG(k)$ is the $DCG(k)$ value of ideal ranking list.

TABLE 6. Results of mapping WiFi logs to customer's interaction behaviour.

Trajectory Distance	0	0.25	0.5	0.75	1
Cumulative Probability	68.3%	89.3%	95.7%	98.6%	100%

c: RESULTS

As shown in Table 6, the mapping accuracy is 94.7% and the percentage that trajectory distance is less than 0.5 is 95.7%. The results show extracting customer's interaction behaviour from WiFi logs is available and accurate.

We perform a user case study on the usefulness of store recommendation. A group including 8 customers (5 males and 3 females) participated in the case study. The users get store recommendation list through TA-LCCF and TA-LVM respectively, then they specify the ideal recommendation list. Based on the recommendation list from recommendation model and the ideal recommendation list given by customers, we calculate the $NDCG@k$ to evaluate the recommendation utility. For instance, if the recommendation list of a customer is $[p_1, p_2, p_4, p_3]$ for four stores, while the ideal recommendation list is $[p_2, p_1, p_3, p_4]$ given by the user, then we can calculate the $NDCG@4 = 0.907$ according to Equation 31.

Figure 10 shows the $NDCG@k$ for each customer with two recommendation models: TA-LCCF and TA-LVM. The $NDCG@k$ is calculated as follows: for a specific customer, the recommendation model firstly derives his/her shopping preference from check-in activities, then generates top-k recommendation list using TA-LVM and TA-LCCF, respectively. For the recommendation results, each user has an ideal rank list answer in his or her mind. Based on the recommendation list from recommendation model and customer's ideal rank list, the $NDCG@k$ can be calculated as Equation 31. We can observe that for all participants the $NDCG@k$ of our proposed model (TA-LVM) is better than the compared model (TA-LCCF), the performance improvement is about 10%. The results suggest that, merely using check-in frequency is insufficient to reflect customer's preference, the average check-in time plays an important role in mining customer's preference. In particular, some of the participants indicated that our recommendation method is conducive to discover new interesting stores when the recommendation results are associated with temporal context.

V. DISCUSSION

This paper proposes a store-type recommendation model (TA-LVM) for physical stores by exploiting interaction behaviours (e.g., the check-in frequency and average stay time) to learn customer's preference. Given increasing number of homogeneous physical stores, the proposed store-type recommendation model can help retailers gain advantages to build excitement with customers. Since most shopping decisions occur in the store and only 1/3 of shopping decisions is planned beforehand [21],personal recommendation service can assist the window shopping customers find new interesting stores based on their shopping preference and intention, which is also confirmed by a user case study.

Different to exist approaches, TA-LVM involves zero-effort for collecting customer's interaction behaviours from WiFi logs with a non-intrusive way. Zero-effort means the data collection is infrastructure-free and no customer involvement, thus can be built as a 3rd-party mobile application. We have collected more than 3,800,000 interaction behaviours from 123,406 customers during 12 months from two urban shopping malls, and found that if a customer can freely use WiFi service, he is willing to participate in data collection (only need to enable WiFi service) in the shopping journey. We extract customer's interaction behaviours based on indoor fingerprint-based localization. According to [34], one major drawback of fingerprint-based localization is that constructing RSS fingerprint map is time-consuming and labor-intensive. Fortunately, the cost for constructing RSS fingerprint map is acceptable in our scenario. For example, we collect 100 WiFi RSS samples with sample rate 1 Hz in each store for building RSS fingerprint map, thus 10 hours are enough for constructing the fingerprint map of two large urban shopping malls by one user. In addition, we collect customer's interaction behaviour as hashed entities with no additional knowledge about them, and finish collecting data when customers leave the shopping mall. We believe that this is a privacy-safe application.

A limitation of the proposed model is the learnt preference is store-level based on customer's interaction behaviours. However, a customer visits a store by matching personal preference with the service content of that store. A customer would have her/his own preference for the choice of stores, and the personal preference can be represented by his/her opinion to various aspects of stores (e.g., environment, price and service). For the aspect-level preference, we plan to extract the aspects that a customer most concerned about when checking a store by exploiting store's online textual reviews. For example, consider a typical review written by a customer regarding a restaurant: "both the taste and service are excellent, while the price is little expensive!." This comment shows the customer's opinions towards three aspects of the restaurant, like "taste:good," "service:good" and "price:bad," such information can provides meaningful semantics about customer's opinion towards various aspects of stores [2] thus is helpful to provide better personalized recommendation.

VI. CONCLUSION

In this paper, we present a novel approach to mine customer preference in physical stores from their interaction behaviours, which are generated from WiFi logs with a non-intrusive way. Using millions of interaction behaviours, we observe that customer's preference is generally influenced by intrinsic preference as well as temporal influence. Based on this observation, we firstly model customer's preference as a hidden factor of his/her interaction behaviours by a latent variable model. To demonstrate the utility of the proposed model, we conduct a store-type recommendation model for physical stores by jointly considering the learnt customer preference and temporal influence. Experimental results show that the proposed model significantly outperforms state-of-art methods in recommendation effectiveness, showing our model can effectively learn customer's preference.

As future work, we plan to 1) derive customer's aspect-level preference by extracting the aspects that a customer most concerned about when checking a store from online textual reviews; 2) facilitate more context-aware applications in shopping malls based on the proposed recommendation model. I.e., detecting target customers and optimizing promotion strategy.

REFERENCES

[1] Z. U. Ahmed, M. Ghingold, and Z. Dahari, "Malaysian shopping mall behavior: An exploratory study," *Asia–Pacific J. Marketing Logistics*, vol. 19, no. 4, pp. 331–348, 2007.

[2] A. Bagheri, M. Saraee, and F. De Jong, "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews," *Knowl.-Based Syst.*, vol. 52, pp. 201–213, Nov. 2013.

[3] P. H. Bloch, N. M. Ridgway, and S. A. Dawson, "The shopping mall as consumer habitat," *J. Retailing*, vol. 70, no. 1, pp. 23–42, 1994.

[4] Y. Cai and R. Shannon, "Personal values and mall shopping behaviour: The mediating role of intention among Chinese consumers," *Int. J. Retail Distrib. Manage.*, vol. 40, no. 4, pp. 290–318, 2012.

[5] B. Fang, S. Liao, K. Xu, H. Cheng, C. Zhu, and H. Chen, "A novel mobile recommender system for indoor shopping," *Expert Syst. Appl.*, vol. 39, no. 15, pp. 11992–12000, 2012.

[6] G. Ference, M. Ye, and W.-C. Lee, "Location recommendation for out-of-town users in location-based social networks," in *Proc. 22nd ACM Int. Conf. Inf. Knowl. Manage.*, 2013, pp. 721–726.

[7] A. Galati, K. Djemame, and C. Greenhalgh, "Analysis of human mobility patterns for opportunistic forwarding in shopping mall environments," *Soc. Netw. Anal. Mining*, vol. 5, no. 1, pp. 1–14, 2015.

[8] V. Jackson, L. Stoel, and A. Brantley, "Mall attributes and shopping value: Differences by gender and generational cohort," *J. Retailing Consum. Services*, vol. 18, no. 1, pp. 1–9, 2011.

[9] T. Kanda, D. F. Glas, M. Shiomi, H. Ishiguro, and N. Hagita, "Who will be the customer?: A social robot that anticipates people's behavior from their trajectories," in *Proc. 10th Int. Conf. Ubiquitous Comput.*, 2008, pp. 380–389.

[10] E. Y. Kim and Y.-K. Kim, "The effects of ethnicity and gender on Teens' mall shopping motivations," *Clothing Textiles Res. J.*, vol. 23, no. 2, pp. 65–77, 2005.

[11] S. Lee, C. Min, C. Yoo, and J. Song, "Understanding customer malling behavior in an urban shopping mall using smartphones," in *Proc. ACM Conf. Pervasive Ubiquitous Comput. Adjunct Publication*, 2013, pp. 901–910.

[12] D. Lemire and A. Maclachlan, "Slope one predictors for online rating-based collaborative filtering," in *Proc. SDM*, vol. 5. 2005, pp. 1–5.

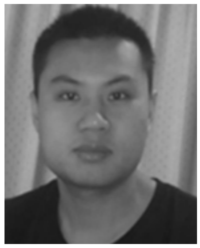
[13] D. Lian *et al.*, "Content-aware collaborative filtering for location recommendation based on human mobility data," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2015, pp. 261–270.

[14] Z. Lin, "Indoor location-based recommender system," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Toronto, Toronto, ON, Canada, 2013.

[15] G. Mohan, B. Sivakumaran, and P. Sharma, "Impact of store environment on impulse buying behavior," *Eur. J. Marketing*, vol. 47, no. 10, pp. 1711–1732, 2013.

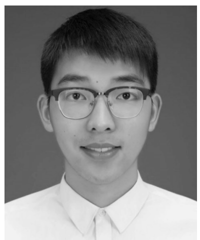
[16] A. B. M. Musa and J. Eriksson, "Tracking unmodified smartphones using Wi-Fi monitors," in *Proc. 10th ACM Conf. Embedded Netw. Sensor Syst.*, 2012, pp. 281–294.

[17] A. Paterek, "Improving regularized singular value decomposition for collaborative filtering," in *Proc. KDD Cup Workshop*, 2007, pp. 5–8.
[18] M. C. Popa, L. J. M. Rothkrantz, C. Shan, T. Gritti, and P. Wiggers, "Semantic assessment of shopping behavior using trajectories, shopping related actions, and context information," *Pattern Recognit. Lett.*, vol. 34, no. 7, pp. 809–819, 2013.
[19] M. Radhakrishnan, S. Eswaran, A. Misra, D. Chander, and K. Dasgupta, "IRIS: Tapping wearable sensing to capture in-store retail insights on shoppers," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2016, pp. 1–8.
[20] S. Rallapalli, A. Ganesan, K. Chintalapudi, V. N. Padmanabhan, and L. Qiu, "Enabling physical analytics in retail stores using smart glasses," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 115–126.
[21] P. M. Reyes and G. V. Frazier, "Goal programming model for grocery shelf space allocation," *Eur. J. Oper. Res.*, vol. 181, no. 2, pp. 634–644, 2007.
[22] A. J. Ruiz-Ruiz, H. Blunck, T. S. Prentow, A. Stisen, and M. B. Kjærgaard, "Analysis methods for extracting knowledge from large-scale WiFi monitoring to inform building facility planning," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2014, pp. 130–138.
[23] L. Shangguan, Z. Zhou, X. Zheng, L. Yang, Y. Liu, and J. Han, "ShopMiner: Mining customer shopping behavior in physical clothing stores with COTS RFID devices," in *Proc. 13th ACM Conf. Embedded Netw. Sensor Syst.*, 2015, pp. 113–125.
[24] S. Shim and M. A. Eastlick, "The hierarchical influence of personal values on mall shopping attitute and behavior," *J. Retailing*, vol. 74, no. 1, pp. 139–160, 1998.
[25] T. Tajima, Y. Iida, and T. Kato, "Analysis of customer preference through unforced natural passive observation," in *Proc. Int. Conf. Hum.-Comput. Interact.*, 2013, pp. 466–474.
[26] M. Vlachos, M. Hadjieleftheriou, D. Gunopulos, and E. Keogh, "Indexing multi-dimensional time-series with support for multiple distance measures," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2003, pp. 216–225.
[27] Y. Wang *et al.*, "Regularity and conformity: Location prediction using heterogeneous mobility data," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2015, pp. 1275–1284.
[28] R. Xiang, J. Neville, and M. Rogati, "Modeling relationship strength in online social networks," in *Proc. 19th Int. Conf. World Wide Web*, Raleigh, NC, USA, Apr. 2010, pp. 981–990.
[29] A. Yaeli *et al.*, "Understanding customer behavior using indoor location analysis and visualization," *IBM J. Res. Develop.*, vol. 58, nos. 5–6, pp. 3:1–3:12, Sep./Nov. 2014.
[30] R. Yalch and E. Spangenberg, "Effects of store music on shopping behavior," *J. Consum. Marketing*, vol. 7, no. 2, pp. 55–63, 1990.
[31] M. Ye, P. Yin, and W.-C. Lee, "Location recommendation for location-based social networks," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geogr. Inf. Syst.*, 2010, pp. 458–461.
[32] H. Yin, B. Cui, L. Chen, Z. Hu, and X. Zhou, "Dynamic user modeling in social media systems," *ACM Trans. Inf. Syst.*, vol. 33, no. 3, p. 10, 2015.
[33] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in *Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2013, pp. 363–372.
[34] Y. Chen, M. Guo, J. Shen, and J. Cao, "A graph-based method for indoor subarea localization with zero-configuration," in *Proc. 13th IEEE Int. Conf. Ubiquitous Intell. Comput.*, Jul. 2016, pp. 236–244.
[35] C. Zhang and K. Wang, "POI recommendation through cross-region collaborative filtering," *Knowl. Inf. Syst.*, vol. 46, no. 2, pp. 369–387, 2016.
[36] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with GPS history data," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 1029–1038.
[37] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Towards mobile intelligence: Learning from GPS history data for collaborative recommendation," *Artif. Intell.*, vols. 184–185, pp. 17–37, Jun. 2012.
[38] Z. Zheng, Y. Chen, T. He, F. Li, and D. Chen, "Weight-RSS: A calibration-free and robust method for WLAN-based indoor positioning," *Int. J. Distrib. Sensor Netw.*, vol. 11, p. 55, Jan. 2015.
[39] D. Zhou, B. Wang, S. M. Rahimi, and X. Wang, "A study of recommending locations on location-based social network by collaborative filtering," in *Advances in Artificial Intelligence*. Toronto, ON, Canada: Springer, 2012, pp. 255–266.



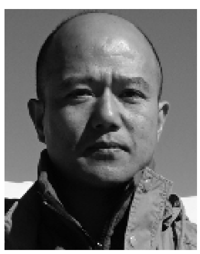
YUANYI CHEN received the B.Sc. degree from Sichuan University in 2010, the master's degree from Zhejiang University in 2013. From 2014 to 2015, he was a jointly supervised Ph.D. student with The Hong Kong Polytechnic University. He is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering, Shanghai Jiao Tong University. His research topic includes Internet of Things, ubiquitous computing, and mobile computing.

ZENGWEI ZHENG received the B.S. and M.Sc. degrees in computer science and western economics from Hangzhou University, China, in 1991 and 1998, respectively, the Ph.D. degree in computer science and technology from Zhejiang University, China, in 2005. From 2000 to 2003 and from 2004 to 2008, he was an Assistant Professor and an Associate Professor with the Department of Computer Science and Engineering, Zhejiang University City College, respectively, where he is currently a Professor and the Head of Department of Scientific Research. His research interests include wireless sensor network, location-based service, Internet of Things, and pervasive computing.



SINONG CHEN received the B.S. degree from Wuhan University in 2015. He is currently pursuing the master's degree with the College of Computer Science and Technology, Zhejiang University City College. His research interests include indoor localization and smart transportation.

LIN SUN received the Ph.D. degree in computer science and technology from Zhejiang University City College, China, in 2010. He has been a Lecturer of computer science with the Zhejiang University City College. His research interests include computer vision, biometric authentication, and pattern recognition.



DAN CHEN received the B.S. degree in electrical engineering and automation from Tianjin Polytechnic University in 1996 and the M.E. degree in computer Science from Zhejiang University in 2004. Since 2008, he has been an Associate Professor of computer science with the Zhejiang University City College. His research interests include computer network, mobile computing, IoT, and computer architecture.