Understanding Customer Behaviour in Urban Shopping Mall from WiFi Logs

Yuanyi Chen[1*], Jinyu Zhang[1*], Minyi Guo[1†] and Jiannong Cao[2]
[1]Department of Computer Science and Engineering, Shanghai Jiao Tong University
Email: *{cyyxz,zhangzhang}@sjtu.edu.cn, †guo-my@cs.sjtu.edu.cn
[2]Department of Computing, The Hong Kong Polytechnic University, Email: csjcao@comp.polyu.edu.hk

***Abstract*—Traditional ways of understanding customer behaviour are mainly based on predominantly field surveys, which are not effective as they require labor-intensive survey. As mobile devices and ubiquitous sensing technologies are becoming more and more pervasive, user-generated data from these platforms are providing rich information to uncover customer preference. In this study, we propose a shop recommendation model for urban shopping mall by exploiting user-generated WiFi logs to learn customer preference. Specifically, the proposed model consists of two phases: 1) offline learning customer's preference from their check-in activities; 2) online recommendation by fusing the learnt preference and temporal influence. We have performed a comprehensive experiment evaluation on a real dataset collected by over 39,000 customers during 7 months, and the experiment results show the proposed recommendation model outperforms state-of-the-art methods.**

I. Introduction

As a driver of local economies, urban shopping malls play a significant role in maintaining economic growth, offering employment and providing a better quality of life. Given increasing number of homogeneous urban shopping malls, the retailers who in-depth understand customer behaviour will gain advantages to support personal recommendation service, since most shopping decisions occur in the shop and only 1/3 of shopping decisions is planned beforehand [5].

Even though a number of shop recommendation approaches for urban shopping mall have been recently proposed [3], [1], these approaches suffer from a number of limitations. These limitations include the assumption that the level of customer's preference can be reflected by the check-in frequency, fail to model the relationship between customer's preference and their check-in activities. Understanding customer behaviour in urban shopping mall has also attracted enormous research from traditional marketing research [6], [4], which are mainly based on predominantly field surveys from small populations, thus limited to scalability and are powerless to capture information from survey avoiders. To tackle these challenges, we propose *MallRec*, a shop recommendation model for urban shopping mall based on customer's check-in activities. The idea behind our approach is customer's check-in activities can be viewed as a contexture of behaviour that is motivated by their intention and preference, then we can infer customer's preference from their history check-in activities.

The remainder of the paper is organized as follows: Section II describes the proposed recommendation model in detail. Section III reports and discusses the experimental results. Finally, we present our conclusion and future work in section IV.

II. Time-aware Recommendation Model for Urban Shopping Mall

The proposed model produces recommended shops by two phases: 1) offline modeling customer's preference from their check-in activities; 2) online recommendation 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 $S = \{s_1, ..., s_i, ...\}$, 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.

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

Definition 3. (**Check-in Activity**) A check-in activity is a 4-tuple $< u, p, ts, cst >$ that means customer u visits shop 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 shop p.

For generating customer's check-in activity, we need to map each WiFi logs to the corresponding shopping trajectories. Given a wifi log $S = \{s_1, ..., s_i, ...\}$, we first utilize fingerprint-based localization [9] to map each RSS sample $s_i \in S$ to the corresponding shop, then construct the shopping trajectories $L(S)$ based on chronological order and further extract customer's check-in activity according to *Definition 3*.

B. Offline Modelling Customer Preference

Since most customers have a finite amount of resources (e.g., money and time) for shopping, they tend to visit a shop by matching their preference. In this way, we model

978-1-5090-4338-5/17/$31.00 ©2017 IEEE

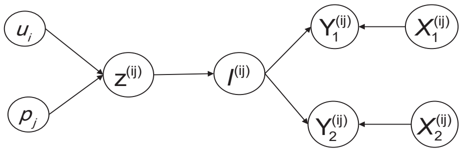


Fig. 1: Graphical model of learning customer's preference

customer's preference as the hidden factor of his/her check-in activities. Formally, let $Y_1^{(ij)}$ and $Y_2^{(ij)}$ denote the check-in frequency and average residence time of customer u_i to shop 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 1. The detailed description of variables in this figure is explained as follows:

- $z^{(ij)}$ denote the intrinsic preference of u_i to p_j, which is a result of both personality and situational factors. Motivated by [2], we capture customer's intrinsic preference based on the widely used location co-occurrence. let $c_{ij} = 1$ if u_i has visited shop p_j and 0 otherwise. Then, we construct the check-in vector of u_i as $c(u_i) = \{c_{i1}, ..., c_{iN}\}$. Then the intrinsic preference $z^{(ij)}$ between u_i and p_j can be calculated by:

$$z^{(ij)} = \frac{\sum_{v \in U} sim(u_i, v) * c < v, p_j >}{\sum_{v \in U} sim(u_i, v)} \tag{1}$$

  where $sim(u_i, v)$ is the similarity between u_i and v, and estimated by cosine similarity between $c(u_i)$ and $c(v)$.
- $x_1^{(ij)}$ and $x_2^{(ij)}$ are two auxiliary variables for check-in frequency and residence time, respectively, such as the total number of visited shops or the average residence time of a customer.

Our model represents the relationships among these variables by modelling the conditional dependencies as shown in Figure 1, so the joint distribution decomposes as follows:

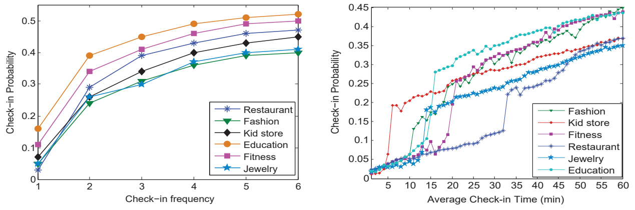
$$P(I^{(ij)}, Y_1^{(ij)}, Y_2^{(ij)} | u_i, p_j) = P(I^{(ij)} | u_i, p_j) \prod_{l=1}^{2} P(Y_l^{(ij)} | I^{(ij)}, X_l^{(ij)}) \tag{2}$$

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

$$P(I^{(ij)} | u_i, p_j) = (\eta z^{(ij)}, \sigma^2) \tag{3}$$

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

For modeling the dependency between $Y_l^{(ij)}$ and $I^{(ij)}, X_l^{(ij)} (l = 1, 2)$, we analyze the characteristics of customer's check-in activities (more details of customer's check-in activities are shown in Table I.). Figure 2 shows customer's check-in probability for different types of shops as a function of their check-in frequency and average check-in time. We observe that the distributions follow a similar power-law form, while the distribution parameters differ from different kinds of shops. In this way, we model the dependency between $Y_l^{(ij)}$ and $I^{(ij)}, X_l^{(ij)} (l = 1, 2)$ as follows:



(a) Check-in frequency

(b) Average check-in time

Fig. 2: Fraction of check-in probability as a function of check-in frequency(a) and average check-in time(b)

$$P(Y_l^{(ij)} | I^{(ij)}, X_l^{(ij)}) = (\alpha_l I^{(ij)} + \beta_l X_l^{(ij)})^{\theta_l} \tag{4}$$

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., α_1, β_1, θ_1, etc.) to avoid over-fitting, which can be regarded as Gaussian prior:

$$\begin{aligned} &P(\alpha_l, \beta_l) \propto e^{-(\lambda_l/2)(\alpha_l^2 + \beta_l^2)}, l = 1, 2 \\ &P(\theta_l) \propto e^{-(\lambda_{\theta_l}/2)(\theta_l)^2}, l = 1, 2 \\ &P(\eta) \propto e^{-(\lambda_\eta/2)\eta^2} \end{aligned} \tag{5}$$

The data are represented as $\Phi = U \times P$ samples of *customer-shop* pairs, denoted as $D = \{(i_1, j_1), ..., (i_N, j_M)\}$. 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 2, given all the observed variables, the joint probability is shown as:

$$\begin{aligned} &\prod_{l=1}^{2} P(\Phi | \eta, \alpha_l, \beta_l, \theta_l) P(\eta, \alpha_l, \beta_l, \theta_l) \\ &= \prod_{(i,j) \in D} P(I^{(ij)} | z^{(ij)}, \eta) P(\eta) \prod_{l=1}^{2} P(D | I^{(ij)}, X_l^{(ij)}, \alpha_l, \beta_l, \theta_l) \\ &P(\alpha_l, \beta_l, \theta_l) \\ &\propto \prod_{(i,j) \in D} \left(e^{-(1/2\delta^2)(\eta z^{(ij)} - I^{(ij)})^2} \prod_{l=1}^{2} (\alpha_l I^{(ij)} + \beta_l X_l^{(ij)})^{\theta_l} \right) \\ &e^{-(\lambda_\eta/2)\eta^2} \prod_{l=1}^{2} e^{-(\lambda_{\theta_l}/2)(\theta_l)^2} e^{-(\lambda_l/2)(\alpha_l^2 + \beta_l^2)} \end{aligned} \tag{6}$$

We maximize the likelihood function as shown in Equation 6 to estimate the unknown model parameters $\Sigma = \{\eta, \alpha_l, \beta_l, \theta_l | l = 1, 2\}$. Applying a logarithmic transformation to both sides of Equation 6, we obtain the following expression:

$$\begin{aligned} L((i,j) \in D, \eta, \alpha_l, \beta_l, \theta_l) = &\sum_{(i,j) \in D} -\frac{1}{2\sigma^2} (\eta z^{(ij)} - I^{(ij)})^2 \\ &+ \sum_{(i,j) \in D} \sum_{l=1}^{2} \theta_l log(\alpha_l I^{(ij)} + \beta_l X_l^{(ij)}) \\ &- \frac{\lambda_\eta}{2} \eta^2 - \sum_{l=1}^{2} \frac{\lambda_{\theta_l}}{2} \theta_l^2 - \sum_{l=1}^{2} \frac{\lambda_l}{2} (\alpha_l^2 + \beta_l^2) \end{aligned} \tag{7}$$

By adopting a coordinate ascent optimization algorithm, we have the following efficient updating rules to learn latent variables $\eta, \alpha_l, \beta_l, \theta_l$:

$$
\begin{aligned}
&I^{(ij)new} \leftarrow I^{(ij)old} - \frac{\partial L}{\partial I^{(ij)}} / \frac{\partial^2 L}{\partial (I^{(ij)})^2} \\
&\eta^{new} \leftarrow \eta^{old} - \frac{\partial L}{\partial \eta} / \frac{\partial^2 L}{\partial (\eta)^2}, \quad \alpha_l^{new} \leftarrow \alpha_l^{old} - \frac{\partial L}{\partial \alpha_l} / \frac{\partial^2 L}{\partial (\alpha_l)^2} \\
&\beta_l^{new} \leftarrow \beta_l^{old} - \frac{\partial L}{\partial \beta_l} / \frac{\partial^2 L}{\partial (\beta_l)^2}, \quad \theta_l^{new} = \theta_l^{old} - \frac{\partial L}{\partial \theta_l} / \frac{\partial^2 L}{\partial (\theta_l)^2}
\end{aligned}
\tag{8}
$$

C. *Online Time-aware Recommendation*

Human daily activities usually follow a regular temporal pattern, i.e., people usually eat dinner at 17:00-19:00, which means a customer may tend to check-in a restaurant rather than other kinds of shops during the time slot. Therefore, temporal influence plays an important role in mining customer's preference, which should be considered in making recommendation.

To extract the temporal pattern of customer's check-in activity, 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). We partition the total check-in activities into a few subsets according to the shop category and time slot of check-ins, and each subset represents customer's check-ins for a certain shop category at a specific time slot. Then, we calculate the check-in probability of customer u to shops that belong to c_p at time slot ts by:

$$
Pr(u, c_p|ts) = \frac{\psi(u, c_p|ts)}{\psi(u, ts)} \tag{9}
$$

where c_p is the category of shop p, $\psi(u, c_p|ts)$ is the check-in number for shops belong to c_p at time slot ts. Accordingly, $\psi(u, ts)$ is the total check-in number for all shops at time slot ts.

Given customer u and a unvisited shop $\widehat{p}$ at time slot ts, we calculate the recommendation score using collaborative filtering:

$$
score(u, \widehat{p}, ts) = \overline{I_u} + \frac{\sum_{v \in U} sim(u,v)(I^{v\widehat{p}} - \overline{I_v}) Pr(u, c_{\widehat{p}}|ts)}{\sum_{v \in U} sim(u,v)} \tag{10}
$$

where $\overline{I_u}$ and $\overline{I_v}$ are the average score of customer u, v to all shops, respectively. $I^{v\widehat{p}}$ is the preference of customer v towards $\widehat{p}$, U is the nearest neighbour set of u. $sim(u,v)$ is the Pearson correlation between customer u, v, calculated by:

$$
sim(u,v) = \frac{\sum_{i \in P_{uv}} (I^{ui} - \overline{I_u})(I^{vi} - \overline{I_v})}{\sqrt{\sum_{i \in P_{uv}} (I^{ui} - \overline{I_u})^2 \sum_{i \in P_{uv}} (I^{ui} - \overline{I_u})^2}} \tag{11}
$$

where P_{uv} is the shops that have been visited by both customer u and v.

TABLE I: Statistics of dataset

Number of Shops	211
Number of Customers	39,038
Number of Check-ins	89,794
Average No. of Check-in shops per customer	41
Average No. of Check-in customers per shop	426
Data density	1.09%

TABLE II: Methods for comparison

Method	**Description**
RBCA [1]	Rule-based recommendation algorithm
TSO [3]	Time-based Slope One
LVM	Our method without fusing temporal influence
LCCF [7]	Collaborative filtering using location co-occurrence
TA-LCCF [8]	LCCF with fusing temporal influence

III. 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 shops to customers. We first 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) Dataset: We gather an anonymized dataset from registered customers using an opt-in WiFi network in an urban shopping mall during 7 months. For removing noise data, we filter out the mall workers and shop employees based on the check-in frequency. Specifically, we consider a customer as a mall worker or shop employee if her/his check-ins are more than 30 during seven months. After preprocessing, the dataset consists of 89,794 check-ins from 39,038 customers on 211 shops, more details of the dataset are shown in Table I.

2) Comparative Algorithms: We compare the proposed recommendation model (MallRec) with 5 start-of-art methods for shop recommendation in urban shopping mall, as shown in Table II.

3) Evaluation Metric: We adopt Recall@K as the measurement metric, where k is the number of the recommendation results. Let $hit@K$ denotes a single test case as either the value 1 if s_i appears in the top-k results, or else the value 0. The overall $Recall@K$ are defined by averaging all test cases:

$$
Recall@K = \frac{\#hit@K}{|D_{te}|} \tag{12}
$$

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

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.

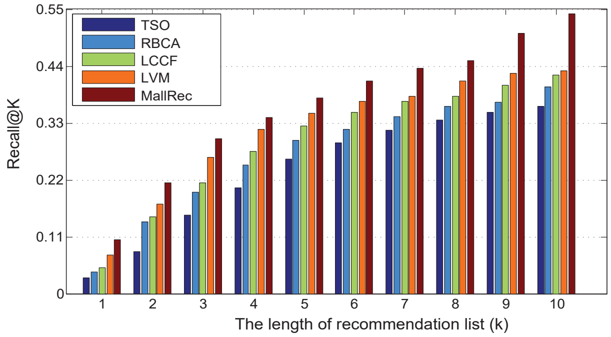
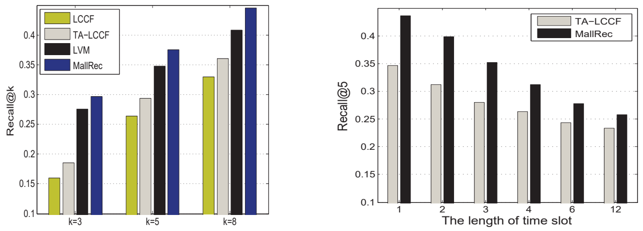


Fig. 3: Top-*k* performance on our data set



(a) Different@*k* (b) Varying length of timeslot

Fig. 4: Impact of temporal influence on shop recommendation

1) Effectiveness of Recommendations: Figure 3 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...10], because there are 211 shops in total and a greater value of k is usually ignored for a typical top-K recommendation task. From this figure, we also observe: 1) TSO performs worst among all recommendation models, showing only utilizing the residence time is insufficient to reflect the level of customer's interests. Similarly, the results of LCCF suggest that only utilizing the check-in frequency is also insufficient to learn customer's interests; 2) MallRec perform better than other competitor methods (LCCF, RBCA, TSO), showing the advantage of using latent variable model to learn customer's preference and fusing temporal influence to make recommendation. For example, the recall of MallRec is about 0.51 when k = 9 (i.e., the model has a probability of 51% of placing a shop within target customer's check-in list), while 0.4 for LCCF, 0.37 for RBCA and 0.35 for TSO.

2) Impact of Temporal Influence: We compare the recommendation effectiveness of two recommendation models (LCCF, LVM) by fusing temporal influence in Figure 4a. From this figure, we can see the two models (TA-LCCF and MallRec)) with fusing temporal influence perform better than the baseline methods (LCCF and LVM), showing temporal influence plays a vital role in analyzing customer's check-in activities and is vital for shop recommendation.

Figure 4b reports the effect on the length of time slot for two recommendation models (TA-LCCF, and MallRec), 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 two models drop with the time slot length increases. The reason is that increasing the length of time slots will bring in more ground truth shops 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.

IV. Conclusion and FUTURE WORK

In this paper, we propose a time-aware recommendation model that recommends a customer a set of shops at a specific time slot by learning customer's preference from their check-in activities. Customer's check-in activity are generated from user-generated WiFi logs. The proposed model firstly models customer's preference as a hidden factor of his/her check-in activity using a latent variable model, then produces top-K recommended shops by jointly considering the learnt preference and temporal influence. Experimental results show that the proposed model significantly outperforms state-of-art methods in recommendation effectiveness.

As future work, we plan to 1) analysis the aspects that a customer most concerned about when checking a shop by exploiting shop's online textual reviews; 2) facilitate more context-aware applications (e.g., detecting target customers and optimizing promotion strategy) in shopping malls by modeling customer's preference from multi-modal information: check-in activities and online reviews.

V. Acknowledge

This work is sponsored by the National Basic Research 973 Program of China (No. 2015CB352403), the Program for National Natural Science Foundation of China / Research Grants Council (NSFC/RGC)(612191030), the Scientific Innovation Act of STCSM (13511504200), and EU FP7 CLIMBER project (PIRSES-GA-2012-318939).

References

[1] B. Fang, S. Liao, K. Xu, H. Cheng, C. Zhu, and H. Chen. A novel mobile recommender system for indoor shopping. *Expert Systems with Applications*, 39(15):11992–12000, 2012.

[2] D. Lian, Y. Ge, F. Zhang, N. J. Yuan, X. Xie, T. Zhou, and Y. Rui. Content-aware collaborative filtering for location recommendation based on human mobility data. In *ICDM*, pages 261–270. IEEE, 2015.

[3] Z. Lin. *Indoor Location-based Recommender System*. PhD thesis, University of Toronto, 2013.

[4] G. Mohan, B. Sivakumaran, and P. Sharma. Impact of store environment on impulse buying behavior. *European Journal of Marketing*, 47(10):1711–1732, 2013.

[5] P. M. Reyes and G. V. Frazier. Goal programming model for grocery shelf space allocation. *European Journal of Operational Research*, 181(2):634–644, 2007.

[6] J. Sit, H. Y. Wong, and D. Birch. An exploratory study on service dimensions of regional shopping centres: A segmentation approach. In *ANZMAC*. Australian and New Zealand Marketing Academy (ANZMAC), 2002.

[7] M. Ye, P. Yin, and W.-C. Lee. Location recommendation for location-based social networks. In *SIGSPATIAL*, pages 458–461. ACM, 2010.

[8] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann. Time-aware point-of-interest recommendation. In *SIGIR*, pages 363–372. ACM, 2013.

[9] Z. Zheng, Y. Chen, T. He, F. Li, and D. Chen. Weight-rss: a calibration-free and robust method for wlan-based indoor positioning. *International Journal of Distributed Sensor Networks*, 2015:55, 2015.