A Common Topic Transfer Learning Model for Crossing City POI Recommendations

Dichao Li, Zhiguo Gong, *Senior Member, IEEE*, and Defu Zhang

***Abstract*—With the popularity of location-aware devices (e.g., smart phones), large amounts of location-based social media data (i.e., user check-in data) are generated, which stimulate plenty of works on personalized point of interest (POI) recommendations using machine learning techniques. However, most of the existing works could not recommend POIs in a new city to a user where the user and his/her friends have never visited before. In this paper, we propose a common topic transfer learning graphical model–the common-topic transfer learning model (CTLM)–for crossing-city POI recommendations. The proposed model separates the city-specific topics (or features) of each city from the common topics (or features) shared by all cities, to enable the users' real interests in the source city to be transferred to the target city. By doing so, the ill-matching problem between users and POIs from different cities can be well addressed by preventing the real interests of users from being influenced by the city-specific features. Furthermore, we incorporate the spatial influence into our proposed model by introducing the regions' accessibility. As a result, the co-occurrence patterns of users and POIs are modeled as the aggregated result from these factors. To evaluate the performance of the CTLM, we conduct extensive experiments on Foursquare and Twitter datasets, and the experimental results show the advantages of CTLM over the state-of-the-art methods for the crossing-city POI recommendations.**

***Index Terms*—Graphical models, machine learning, recommender systems, transfer learning.**

I. Introduction

THE POPULARITY of location-aware devices (e.g., smart phones) makes it possible for users to freely capture and share their social activities through various location-based social networks (LBSNs), such as Foursquare[1] and Facebook Places.[2] The large amounts of user-contributed data in those LBSNs enable various mining tasks, among which point of interests (POIs) recommendations for tourists using machine learning techniques is one of the most active research problems in recent years. However, crossing-city POI recommendation techniques have not been well addressed because of the severe data sparsity for the crossing-city users.

Manuscript received October 26, 2017; revised February 11, 2018 and May 9, 2018; accepted July 27, 2018. This work was supported in part by the Fund of Science and Technology Development of Macau Government under Grant FDCT/007/2016/AFJ, and in part by the University of Macau under Grant MYRG2015-00070-FST, Grant MYRG2017-00212-FST, and Grant MYRG2018-00129-FST. This paper was recommended by Associate Editor Y. S. Y. S. Ong. *(Corresponding author: Zhiguo Gong.)*

D. Li and Z. Gong are with the Department of Computer Information Science, University of Macau, Macau, China (e-mail: yb57417@umac.mo; fstzgg@umac.mo).

D. Zhang is with the Department of Computer Science, Xiamen University, Xiamen 361005, China (e-mail: dfzhang@xmu.edu.cn).

Digital Object Identifier 10.1109/TCYB.2018.2861897

[1] https://foursquare.com/

[2] https://www.facebook.com/places/

Compared with the recommendation techniques in other domains (e.g., movie or product recommendations), the technique of POI recommendations faces a more serious data sparsity problem. As indicated in [1], users from New York only left 0.47% of their total check-ins in Los Angeles. Similarly, [2] observed that the sparsity of its user-POI matrix, which was extracted from Foursquare, reached as high as 99.87%, where each user visited only 55.94 POIs on average out of the total 46 194 POIs. In this paper, we aim at tackling the problem and providing some effective solutions to the crossing-city POI recommendations using transfer learning techniques, where the target city (containing the recommended POIs) and the source cities (visited by the users before) have very few or even no crossing-city users.

A. Motivation

To improve the effectiveness of POI recommendations, researchers have exploited many machine learning techniques which can be summarized into two broad categories: 1) pure collaborative filtering (CF) approaches [3] and 2) content-assisted CF techniques.

With respect to pure CF approaches, the intuition is that users with similar preferences tend to have similar behaviors. References [2] and [4]–[7] exploited the traditional matrix factorization techniques to predict new POIs for users, while Ference *et al.* [8] proposed extending a user's profile by combining it with profiles of similar users (i.e., his/her social friends) in order to deal with the data sparsity as they assumed that users preferred to visit the places that their friends visited in a city. However, it is impossible to adopt those techniques for the crossing-city POI recommendations due to a lack of crossing-city users who are essential to support the CF algorithms.

To address this problem, researchers have to seek assistance from the contents (textual descriptions) of POIs. In this respect, Bao *et al.* [1] extended the CF-based approach by mapping the users in the home city to a group of users (called "experts" in this paper) in the target city by the content similarities of their profiles (visited POIs in the past). In a similar manner, Zhao *et al.* [9] grouped both users in the home city and those in the target city into small communities in terms of their interests, and sought an optimal match between communities of the home city and the communities of the target city in order to fulfill the crossing-city POI recommendations. To bridge the crossing-city "gap," Gao *et al.* [10] extended the traditional factorization of the user-POI matrix by introducing a new matrix factor of content-latent, in addition to the two factors: 1) user-latent matrix and 2) POI-latent matrix. Jiang *et al.* [11] proposed an author-topic-based collaborative filtering method to facilitate the crossing city POI recommendations. In order to incorporate all of the features (i.e., users' profiles, POI contents, user/POI relationships) into a unified model, Hu and Ester [12], Yin *et al.* [13], [14], and Wang *et al.* [15] utilized a probabilistic generative topic model to derive the interests of users in the target city.

Although the aforementioned methods can alleviate the data sparsity of crossing-city POI recommendations to some extent, one common limitation of those existing solutions is ignoring the influence of the specific features in each city, which can cause an ill matching problem between the users in the source cities and the POIs in the target city. Directly matching a user's profile with POIs in a new city may lead to some ill recommendation. For example, an algorithm may recommend casino-related POIs in Singapore to a user from Las Vegas just because his/her profile is dominated by casino-related terms, even though his/her real interest may not be casino at all (the casino-related terms in his profile are solely caused by the special features of Las Vegas). Besides that, the spatial–temporal bias of each city may also play an important role in modeling users' check-in behaviors. On one hand, a user is more likely to go to bars or a cinema at night in a nightlife-biased city while he/she may prefer to visit some POIs in the daylight in cities where people tend to have activities in the daytime. Therefore, a city's temporal features may also influence users' check-in behavior for the time-aware POI recommendations. Further, each city has several regions and for the sake of convenience, users tend to visit the more accessible regions especially when they are in a new city. In this paper, we are going to make use of all those factors in our modeling the problem.

B. Contributions

By embedding the transfer learning techniques [16]–[19] into the topic modeling [20], [21], we propose a graphical model for the crossing-city POI recommendations. In the proposed approach, we assume a user's visit to POIs is not only driven by his/her real interests, but is also affected by the specificities of the city, and only his/her real interests can be transferred to the target city. To this end, we separate the specific topics of each city which stand for the unique features of this city, from the common topics which are shared by all cities in the collection, such that POIs under the same common topic reflect the same real interests of users. Under such an assumption, the users' real interests on POIs in the source city can be transferred to match with those POIs in the target city if they share the same common topic. On the other hand, POIs with strong specific features in the target city may also attract users' visit.

Moreover, to incorporate the spatial–temporal factors into our model, we introduce the regions' accessibility and temporal influence to learn the temporal topics in the city. Specifically, we split each day from Monday to Sunday into daytime and night; thus, we have 14 different time intervals. Given a target city and a specific time, the check-in records are used to mine the temporal common topics shared by the source cities and temporal city-specific topics. At the same time, we divide the city into several areas according to the geographical coordinates of POIs to discover the regions' accessibility as we assume that when a user travels to a new city he/she tends to visit the attractive areas for the sake of location convenience. To our best knowledge, this is the first work for the crossing-city POI recommendations by adopting the transfer learning techniques and jointly taking into account the common topics, the city-specific topics, and the spatial–temporal features into a unified model.

The main contributions of this paper can be summarized as follows.

1) We propose a novel transfer-learning-embedded graphical model for the crossing-city POI recommendations, which separates the city-specific topics of each city from the common topics shared by all cities and transfers users' real interests from the source cities to the target city by the medium of the common topics. With such handling, the ill-matching problem between users and POIs from different cities can be well addressed.
2) To incorporate the spatial–temporal features, we introduce the regions' accessibility and split time into slices to learn the temporal topics. In this way, the recommendation list can be refined by jointly considering the common topics, city-specific topics, and the spatial–temporal features.
3) We conduct extensive experiments on two real datasets to evaluate model performance. The results show that our proposed model is superior to the state-of-the-art methods in terms of both Precision and Recall. Besides, we study the impact of different factors by making comparison with three variants of the common-topic transfer learning model (CTLM).

C. Organization of This Paper

The rest of this paper is organized as follows. In Section II, we give some basic concepts and notations, Section III presents the CTLM model for the crossing-city POI recommendations in detail, and we demonstrate the experimental results in Section IV, Section V reviews the existing works related to POI recommendations, and finally we make the conclusion in Section VI.

II. Preliminary

Here are the notations used in this paper and the problem definition for the crossing-city POI recommendations.

Definition 1 (Check-In Record): A check-in record is a tuple (u, v, t, l_v, W_v, c), where u, v, t, l_v, W_v, and c denote the user, the POI, the time, the location of the POI, the contents of the POI, and the city of the POI, respectively. l_v is

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
CHECK-IN EXAMPLE

UserID: 10007	
POI Name: Boston Public Garden	
Time: Tue Jul 19 142511 +0000 2011	
Locations: Longitude: 42.3532527, Latitude: -71.0702542	
Categories: Garden and Park	
Tags: swan boats, gardens, flowers, scenic views, parks	

TABLE II
MATHEMATICAL NOTATIONS

Symbol	Description
U, V, W, C, T	the set of users, POIs, words, cities, and time slices
ω	indicator of source city or target city
R, K^{com}, K^{ω}	the number of the regions, the common topics, the city-specific topics
u, v, l, w, t	user, POI, location, word, and time of check-ins
z^{com}, z^{ω}	assignment of the common topics, the city-specific topics
r	region assignment in the target city
x^{ω}	a binary switch indicator if a check-in is generated from common topic it equals to 0, and 1 otherwise.
ϑ	a Multinomial distribution over regions in the target city
θ^{com}, θ^{ω}	Multinomial distributions over common topics, the city-specific topics
μ_r^{tar}, Σ_r^{tar}	the mean and covariance of region r in the target city
ϕ_z^{com}, ψ_z^{com}, φ_z^{com}	Multinomial distributions over users, POIs, and words specific to the common topics
ϕ_z^{ω}, ψ_z^{ω}, φ_z^{ω}	Multinomial distributions over users, POIs, and words specific to the city-specific topics z
λ^{ω}	Bernoulli distributions of switch indicators
α^{com}, α^{ω}	Dirichlet priors to θ^{com} and θ^{ω}
ζ	Dirichlet priors to ϑ
η^{com}, δ^{com}, β^{com}	Dirichlet priors to ϕ^{com}, ψ^{com}, and φ^{com}
η^{ω}, δ^{ω}, β^{ω}	Dirichlet priors to ϕ^{ω}, ψ^{ω}, and φ^{ω}
γ^{ω}	Beta priors to λ^{ω}

usually in the form of latitude and longitude, and W_v represents the bag of words describing the POI. Table I shows an example.

Definition 2 (City Profile): Each city's profile is formally defined as a set of check-in records in that city, that is, $D_c = \{(u, v, t, l_v, W_v, c)\}$. The dataset D in our model consists of all city profiles, that is, $D = \{D_c : c \in C\}$, where C is the set of all cities in the collection.

Problem Definition 1: Given a set of cities associated with their profiles $D = \{D_c : c \in C\}$, a target user u, a target city c he/she has never visited before, and a specific time t, the crossing-city POI recommendations aim at offering the user a list of POIs in the target city that he/she would be interested in.

Table II summarizes all important notations used in the latter part of this paper.

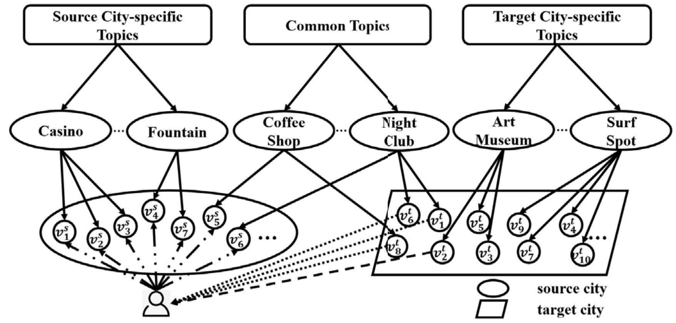


Fig. 1. Working principles of the transfer learning via common topics.

III. TRANSFER LEARNING EMBEDDED GRAPHICAL MODEL

In this section, we first give the principles of common-topic transfer learning, then introduce the generative process of the proposed model, and finally present the parameter inference techniques in detail.

A. Common Topics for Transfer Learning

The principles of transfer learning via common topics are presented in Fig. 1, where v_i^{sou} and v_j^{tar} denote POIs in the source city and target city, respectively. In the figure, each POI is generated by either a city-specific topic or a shared common topic, and POIs under the same topic share similar features. POIs in the city-specific topics attract visitors by the city-specific features while POIs in the common topics drive users to visit by their real interests. In this way, users' real interests are captured by those common topics. Clearly, there is no connection between the two cities but the POIs with similar contents may be clustered into the same common topic which enables transferring the users' real interests across cities. For example, although the user has no check-in records in the target city, the POI v_8^{tar} in the target city shares the same topic with v_5^{sou}, which is visited by the user in the source city. Therefore, v_8^{tar} could be recommended to the user based on his/her real interests. Besides that, each city has its own specific topics, in which the user may tend to visit the POIs. As a result, v_2^{tar} may be recommended to the user as well in the final result.

B. Model Description

Based on the basic idea, we propose a CTLM for the crossing-city POI recommendations, which incorporates the spatial–temporal features, and learns the common topics and city-specific topics in a unified model. The proposed model treats each city profile as a document, and users, POIs, geographical coordinates of POIs, and contents of POIs as vocabulary. We further assume each city has a mixture of the common topics and the city-specific topics which are multinomial distributions over users, POIs, and words, and each city has several regions over geographical coordinates, and a latent switch variable is introduced to indicate the generated route of the check-in records via either the common topics or the city-specific topics.

Fig. 2 presents the graphical model of the proposed approach for the crossing-city POI recommendations, where $\boldsymbol{r}$, $\boldsymbol{z}$, and $\boldsymbol{x}$ (the region, the topic, and the switch indicator) are the latent variables, $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{t}$, and $\boldsymbol{W_v}$ (user, POI, time, and contents) are the observed variables, R and K are the number of regions and the number of topics, letters within a dotted circle are the hyper-parameters, and the remains are the estimated parameters. Each city c has $|D_c|$ check-in records. Besides, the superscript sou, com, and tar denote source, common, and target. For convenience, $\boldsymbol{v}$ is represented a two-tuple $<v, l_v>$, where v denotes the id of the POI and i_v is its location. For each city c, the city indicator $\omega \in \{\text{sou}, \text{tar}\}$ indicates that the current city is either the source city or the target city. The generative process in time slot t is described as follows.

1) Generating Switch Indicators: To discover users' real interests, we introduce a switch indicator which can differentiate the check-in records generated from either the common topics or the city-specific topics during processing. The latent switch indicator x^ω is associated with each check-in record. In particular, $x^\omega = 0$ if the check-in record is generated via the common topic; otherwise, the check-in record is generated via the city-specific topic. Therefore, x^ω is generated from the Bernoulli distribution $x^\omega \sim \text{Bernoulli}(\lambda^\omega)$ given that λ^ω is generated by the conjugate prior $\lambda^\omega \sim \text{Beta}(\gamma^\omega, \gamma^\omega)$.

2) Generating Common Topics: If $x^\omega = 0$, the multinomial parameter θ_t^{com}, characterizing the common features shared by all cities in the collection at time t, is drawn from the Dirichlet distribution as $\theta_t^{\text{com}} \sim \text{Dir}(\alpha^{\text{com}})$, and the common topic z^{com} is drawn from the multinomial distribution as $z^{\text{com}} \sim \text{Multi}(\theta_t^{\text{com}})$ based on the whole collection of cities D. Conditioned on the topic z^{com}, a user u, a POI v, and words $w \in W_v$ are generated from the common topic-user multinomial distribution $u \sim \text{Multi}(\psi_{z^{\text{com}}}^{\text{com}})$, the common topic-POI multinomial distribution $v \sim \text{Multi}(\varphi_{z^{\text{com}}}^{\text{com}})$, and the common topic-word multinomial distribution $w \sim \text{Multi}(\phi_{z^{\text{com}}}^{\text{com}})$, respectively.

3) Generating City-Specific Topics: If $x^\omega = 1$, the city-specific multinomial parameter θ_{ct}^ω, modeling the city-specific features at time t in city c, is drawn from Dirichlet distribution as $\theta_{ct}^\omega \sim \text{Dir}(\alpha^\omega)$, and the city-specific topic z^ω is drawn as $z^\omega \sim \text{Multi}(\theta_{ct}^\omega)$ specific to city c. Conditioned on the topic z^ω, a user u, a POI v, and words $w \in W_v$ are generated from the city-specific topic-user multinomial distribution $u \sim \text{Multi}(\psi_{z^\omega}^\omega)$, the city-specific topic-POI multinomial distribution $v \sim \text{Multi}(\phi_{z^\omega}^\omega)$, and the city-specific topic-word multinomial distribution $w \sim \text{Multi}(\varphi_{z^\omega}^\omega)$, respectively.

4) Generating Regions: POIs can mutually influence each other in attracting users if they are geographically located together. In other words, travelers may visit a POI just for its easy access from other popular POIs nearby. We use region r to group the nearby POIs together and assume that POIs in the same region share the same accessibility (or popularity). The accessibility of regions in the target city c is modeled as

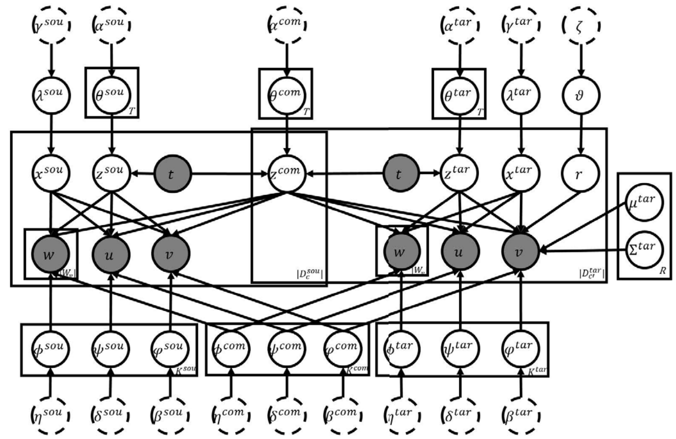


Fig. 2. Graphical representation of CTLM.

a multinomial distribution ϑ. Similar to [22], the geographical coordinates l_v is generated by a Gaussian distribution $l_v \sim \mathcal{N}(\mu_r^{\text{tar}}, \Sigma_r^{\text{tar}})$ for each region r in city c, which is formulated as

$$P\left(l_v|\mu_r^{\text{tar}}, \Sigma_r^{\text{tar}}\right) = \frac{1}{2\pi\sqrt{\left|\Sigma_r^{\text{tar}}\right|}} \times \exp\left(\frac{-\left(l_v - \mu_r^{\text{tar}}\right)^T\left(l_v - \mu_r^{\text{tar}}\right)}{2\Sigma_r^{\text{tar}}}\right) \quad (1)$$

where μ_{cr} and Σ_{cr} are, respectively, the mean vector and covariance matrix of region r.

Our CTLM jointly integrates the region accessibility, common topics, and city-specific topics in a unified way. Therefore, the common topics can transfer the knowledge learned from users' source cities to the target city and discover the co-occurrence patterns of users and POIs in the target city according to their "real" interests, while the city-specific topics can extract the most specific POIs of the target city for the users. Furthermore, the regions' accessibility enables the users to visit the POIs in the popular regions. The detailed generative process of CTLM is described in Algorithm 1.

According to the generative process, the probability that the target user u will visit the POI v in the target city c is computed by the following formula:

$$\begin{aligned} P(v, u|c, t) = &\left[\sum_r P(v|r)P(r|c)\right] \\ &\times \left[P(x = 0|c)\sum_{z^{\text{com}}} P\left(v, u|z^{\text{com}}\right)P\left(z^{\text{com}}|t\right)\right. \\ &\left. + P(x = 1|c)\sum_{z^\omega} P\left(v, u|z^\omega\right)P\left(z^\omega|c, t\right)\right] \end{aligned} \quad (2)$$

where $P(v|r)$ denotes the probability of v belonging to region r; $P(v, u|z^{\text{com}})$ and $P(v, u|z^\omega)$ are the probabilities that co-occurrence patterns of users and POIs are generated via the common topics and the city-specific topics, respectively;

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Algorithm 1 Probabilistic Generative Process in CTLM

for each common topic z^{com} **do**
  Draw a distribution over users $\psi^c \sim \text{Dirichlet}(\cdot|\delta^c)$;
  Draw a distribution over POIs $\varphi^c \sim \text{Dirichlet}(\cdot|\beta^c)$;
  Draw a distribution over words $\phi^c \sim \text{Dirichlet}(\cdot|\eta^c)$;
end for
for each city-specific topic $z^\omega, \omega \in \{sou, tar\}$ **do**
  Draw a distribution over users $\psi^\omega \sim \text{Dirichlet}(\cdot|\delta^\omega)$;
  Draw a distribution over POIs $\varphi^\omega \sim \text{Dirichlet}(\cdot|\beta^\omega)$;
  Draw a distribution over words $\phi^\omega \sim \text{Dirichlet}(\cdot|\eta^\omega)$;
end for
for each time slot t **do**
  Draw a distribution over common topics $\theta_t^{com} \sim \text{Dirichlet}(\cdot|\alpha^{com})$;
  for each city $c, \omega \in \{sou, tar\}$ **do**
    Draw a distribution over city-specific topics $\theta_{c,t}^\omega \sim \text{Dirichlet}(\cdot|\alpha^\omega)$;
    Draw a distribution over switch indicators $\lambda_c^\omega \sim \text{Beta}(\cdot|\gamma^\omega, \gamma^\omega)$;
    if $\omega = tar$ **then**
      Draw a distribution over regions $\vartheta_c^\omega \sim \text{Dirichlet}(\cdot|\zeta^\omega)$;
    end if
  end for
end for
for each city c, $\omega \in \{sou, tar\}$ **do**
  for each check-in record $(u, v, t, W_v) \in D_c$ **do**
    if $\omega = tar$ **then**
      Draw a region $r \sim \text{Multi}(\vartheta)$;
      Draw the location l_v of v $l_v \sim \mathcal{N}(\mu_r^{tar}, \Sigma_r^{tar})$;
    end if
    Draw a switch indicator $x_{ci}^\omega \sim \text{Bernoulli}(\lambda_c^\omega)$;
    if $x_{ci}^\omega = 0$ **then**
      Draw a common topic $z^{com} \sim \text{Multi}(\theta_t^{com})$;
      Draw a user $u \sim \text{Multi}(\psi_{z^{com}})$;
      Draw POI $v \sim \text{Multi}(\varphi_{z^{com}})$;
      for each word $w \in W_v$ **do**
        Draw word $w \sim \text{Multi}(\phi_{z^{com}})$;
      end for
    else
      Draw a city-specific topic $z^\omega \sim \text{Multi}(\theta_{c,t}^\omega)$;
      Draw user $u \sim \text{Multi}(\psi_{z^\omega}^\omega)$;
      Draw POI $v \sim \text{Multi}(\varphi_{z^\omega}^\omega)$;
      for each word $w \in W_v$ **do**
        Draw word $w \sim \text{Multi}(\phi_{z^\omega}^\omega)$
      end for
    end if
  end for
end for

$P(r|c)$ is the region distribution in city c; and $P(z^{\text{com}}|t)$ and $P(z^\omega|c, t)$ denote the common topic distribution for the whole collection and the city-specific topic distribution specific to city c in time slot t. $P(x = 0|c)$ and $P(x = 1|c)$ are the probability of the check-in record generated by the common topics and the city-specific topics, respectively.

C. Parameter Inference

To learn the estimated parameters in this model, it is difficult to maximize the marginal likelihood of the observed variables directly; therefore, collapsed Gibbs sampling [23] is utilized to jointly sample latent variables for each check-in record. According to the generative process, the joint distribution in the source city can be formulated as follows:

$$\begin{aligned} &P\big(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{W_v}, \boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^{\textbf{sou}}, \boldsymbol{x}^{\textbf{sou}}|t, \cdot\big) \\ &\quad= P\big(\boldsymbol{x}^{\textbf{sou}}|\cdot\big)P\big(\boldsymbol{z}^{\textbf{com}}|t, \cdot\big)P\big(\boldsymbol{z}^{\textbf{sou}}|t, \cdot\big) \\ &\quad\times \left[P\big(\boldsymbol{v}|\boldsymbol{z}^{\textbf{com}}, \cdot\big)P\big(\boldsymbol{u}|\boldsymbol{z}^{\textbf{com}}, \cdot\big)\prod_{w\in W_v} P\big(\boldsymbol{w}|\boldsymbol{z}^{\textbf{com}}, \cdot\big)\right]^{\mathbb{I}[x^{\text{sou}}=0]} \\ &\quad\times \left[P\big(\boldsymbol{v}|\boldsymbol{z}^{\textbf{sou}}, \cdot\big)P\big(\boldsymbol{u}|\boldsymbol{z}^{\textbf{sou}}, \cdot\big)\prod_{w\in W_v} P\big(\boldsymbol{w}|\boldsymbol{z}^{\textbf{sou}}, \cdot\big)\right]^{\mathbb{I}[x^{\text{sou}}=1]} \end{aligned} \tag{3}$$

where $\cdot$ denotes the hyper-parameters, and $\mathbb{I}[b]$ is the indicator function that is equal to 1 when b is true, and 0 otherwise.

Similarly, the joint distribution in the target city can be derived as well. Therefore, the joint distribution of the whole data is

$$\begin{aligned} &P\big(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{W_v}, \boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^{\textbf{sou}}, \boldsymbol{x}^{\textbf{sou}}, \boldsymbol{r}^{\textbf{tar}}, \boldsymbol{z}^{\textbf{tar}}, \boldsymbol{x}^{\textbf{tar}}|t, \cdot\big) \\ &\quad= P\big(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{W_v}, \boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^{\textbf{sou}}, \boldsymbol{x}^{\textbf{sou}}|t, \cdot\big) \\ &\quad\times P\big(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{W_v}, \boldsymbol{r}^{\textbf{tar}}, \boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^{\textbf{tar}}, \boldsymbol{x}^{\textbf{tar}}|t, \cdot\big). \end{aligned} \tag{4}$$

To calculate the parameters, the joint probability $P(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{W_v}, \boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^{\textbf{sou}}, \boldsymbol{x}^{\textbf{sou}}, \boldsymbol{r}^{\textbf{tar}}, \boldsymbol{z}^{\textbf{tar}}, \boldsymbol{x}^{\textbf{tar}}|t, \cdot)$ enables us to adopt the collapsed Gibbs sampling algorithm. Specifically, we obtain samples of the latent variable assignment $\boldsymbol{z}^{\textbf{com}}$, $\boldsymbol{z}^{\textbf{sou}}$, $\boldsymbol{x}^{\textbf{sou}}$, $\boldsymbol{r}^{\textbf{tar}}$, $\boldsymbol{z}^{\textbf{tar}}$, $\boldsymbol{x}^{\textbf{tar}}$ by integrating $\{\theta, \vartheta, \lambda, \psi, \phi, \varphi\}$ out.

By integrating out the parameter φ^{com}, φ^{sou}, and φ^{tar}, we obtain

$$\begin{aligned} &P\big(\boldsymbol{v}|\boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^\omega, \boldsymbol{x}^\omega, \beta^{\text{com}}, \beta^\omega\big) \\ &\quad= \left(\frac{\Gamma\big(\sum_v \beta_v^{\text{com}}\big)}{\prod_v \Gamma\big(\beta_v^{\text{com}}\big)}\right)^{K^{\text{com}}} \prod_{z^{\text{com}}} \frac{\prod_v \Gamma\big(n_{z^{\text{com}}v} + \beta_v^{\text{com}}\big)}{\Gamma\big(\sum_v \big(n_{z^{\text{com}}v} + \beta_v^{\text{com}}\big)\big)} \\ &\quad\times \left(\frac{\Gamma\big(\sum_v \beta_v^\omega\big)}{\prod_v \Gamma\big(\beta_v^\omega\big)}\right)^{K^\omega} \prod_{z^\omega} \frac{\prod_v \Gamma\big(n_{z^\omega v} + \beta_v^\omega\big)}{\Gamma\big(\sum_v \big(n_{z^\omega v} + \beta_v^\omega\big)\big)} \end{aligned} \tag{5}$$

where $n_{z^{\text{com}}v}$ and $n_{z^\omega v}$ are the number of times that POI v is assigned to the common topic z^{com} and city-specific topic z^ω, respectively. $\Gamma(\cdot)$ is the Gamma function. Similarly, $P(\boldsymbol{u}|\boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^\omega, \boldsymbol{x}^\omega, \delta^{\text{com}}, \delta^\omega)$ and $P(\boldsymbol{w}|\boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^\omega, \boldsymbol{x}^\omega, \eta^{\text{com}}, \eta^\omega)$ can be calculated by integrating their corresponding parameters as follows:

$$\begin{aligned} &P\big(\boldsymbol{u}|\boldsymbol{z}^{\textbf{com}}, \boldsymbol{z}^\omega, \boldsymbol{x}^\omega, \delta^{\text{com}}, \delta^\omega\big) \\ &\quad= \left(\frac{\Gamma\big(\sum_u \delta_u^{\text{com}}\big)}{\prod_u \Gamma\big(\delta_u^{\text{com}}\big)}\right)^{K^{\text{com}}} \prod_{z^{\text{com}}} \frac{\prod_u \Gamma\big(n_{z^{\text{com}}u} + \delta_u^{\text{com}}\big)}{\Gamma\big(\sum_u \big(n_{z^{\text{com}}u} + \delta_u^{\text{com}}\big)\big)} \\ &\quad\times \left(\frac{\Gamma\big(\sum_u \delta_u^\omega\big)}{\prod_u \Gamma\big(\delta_u^\omega\big)}\right)^{K^\omega} \prod_{z^\omega} \frac{\prod_u \Gamma\big(n_{z^\omega u} + \delta_u^\omega\big)}{\Gamma\big(\sum_u \big(n_{z^\omega u} + \delta_u^\omega\big)\big)} \end{aligned} \tag{6}$$

where $n_{z^{\text{com}}u}$ and $n_{z^{\omega}u}$ are the number of times that POI u is assigned to the common topic z^{com} and city-specific topic z^{ω}, respectively

$$P\left(\boldsymbol{w}|\boldsymbol{z}^{\mathbf{com}}, \boldsymbol{z}^{\omega}, \boldsymbol{x}^{\omega}, \eta^{\text{com}}, \eta^{\omega}\right) = \left(\frac{\Gamma\left(\sum_w \eta_w^{\text{com}}\right)}{\prod_v \Gamma\left(\eta_w^{\text{com}}\right)}\right)^{K^{\text{com}}} \prod_{z^{\text{com}}} \frac{\prod_w \Gamma\left(n_{z^{\text{com}}w} + \eta_w^{\text{com}}\right)}{\Gamma\left(\sum_w \left(n_{z^{\text{com}}w} + \eta_w^{\text{com}}\right)\right)} \times \left(\frac{\Gamma\left(\sum_w \eta_w^{\omega}\right)}{\prod_w \Gamma\left(\eta_w^{\omega}\right)}\right)^{K^{\omega}} \prod_{z^{\omega}} \frac{\prod_v \Gamma\left(n_{z^{\omega}w} + \eta_w^{\omega}\right)}{\Gamma\left(\sum_w \left(n_{z^{\omega}w} + \eta_w^{\omega}\right)\right)} \tag{7}$$

where $n_{z^{\text{com}}w}$ and $n_{z^{\omega}w}$ are the number of times that POI w is assigned to the common topic z^{com} and city-specific topic z^{ω}, respectively.

Next, we need to evaluate $P(z^{\text{com}}|t, \alpha^{\text{com}})$, $P(z^{\omega}|c, t, \alpha^{\omega})$, and $P(r|c, \zeta)$

$$P\left(\boldsymbol{z}^{\mathbf{com}}|t, \alpha^{\text{com}}\right) = \left(\frac{\Gamma\left(\sum_z \alpha_z^{\text{com}}\right)}{\prod_z \Gamma\left(\alpha_z^{\text{com}}\right)}\right) \prod_z \frac{\prod_z \Gamma\left(n_{tz} + \alpha_z^{\text{com}}\right)}{\Gamma\left(\sum_z \left(n_{tz} + \alpha_z^{\text{com}}\right)\right)} \tag{8}$$

where n_{tz} is the number of times that common topic z has been sampled at time slice t

$$P\left(\boldsymbol{z}^{\boldsymbol{\omega}}|t, c, \alpha^{\omega}\right) = \left(\frac{\Gamma\left(\sum_z \alpha_z^{\omega}\right)}{\prod_z \Gamma\left(\alpha_z^{\omega}\right)}\right)^{|C*T|} \prod_z \frac{\prod_z \Gamma\left(n_{ctz} + \alpha_z^{\text{com}}\right)}{\Gamma\left(\sum_z \left(n_{ctz} + \alpha_z^{\text{com}}\right)\right)} \tag{9}$$

where n_{ctz} is the number of times that city-specific topic z has been sampled specific to city c at time slice t; $|C|$ and $|T|$ are the number of cities and time slices, respectively

$$P(\boldsymbol{r}|c, \zeta) = \left(\frac{\Gamma\left(\sum_r \zeta_r\right)}{\prod_r \Gamma(\zeta_r)}\right) \prod_r \frac{\prod_r \Gamma(n_{cr} + \zeta_r)}{\Gamma\left(\sum_r (n_{cr} + \zeta_r)\right)} \tag{10}$$

where n_{cr} is the number of times that region r has been sampled specific to city c.

Finally, we integrate out λ^{ω} to derive $P(\boldsymbol{x}|\gamma^{\omega})$

$$P\left(\boldsymbol{x}^{\omega}|\gamma^{\omega}\right) = \left(\frac{\Gamma(2\gamma^{\omega})}{(\Gamma(\gamma^{\omega}))^2}\right)^{|C|} \prod_c \frac{\Gamma\left(n_{cx_0^{\omega}} + \gamma^{\omega}\right)\Gamma\left(n_{cx_1^{\omega}} + \gamma^{\omega}\right)}{\Gamma\left(n_{cx_0^{\omega}} + n_{cx_1^{\omega}} + 2\gamma^{\omega}\right)} \tag{11}$$

where $n_{cx_0^{\omega}}$ and $n_{cx_1^{\omega}}$ are the number of times that $x^{\omega} = 0$ and $x^{\omega} = 1$ have been sampled specific to city c, respectively.

We iteratively generate latent variables for each check-in record according to the conditional probability, and finally calculate the estimated parameters based on the latent variables. Here, we sample the switch indicator and the topic assignment together, and then sample the region assignment for the location.

1) Sampling Switch Indicator and Common Topic Assignment Jointly: If $x_{ci}^{\omega} = 0$, the user, POI, and words in the ith check-in record in city c are generated by the common topics, and the jointly conditional distribution is computed as

$$P\left(x_{ci}^{\omega} = 0, z_{ci}^{\text{com}} = z^{\text{com}}|\boldsymbol{x}_{\neg ci}^{\boldsymbol{\omega}}, \boldsymbol{z}_{\neg ci}^{\mathbf{com}}, \text{rest}\right) \propto \frac{n_{tz^{\text{com}}}^{\neg ci} + \alpha^c}{\sum_{z^{c'}} (n_{tz^{c'}}^{\neg ci} + \alpha^{c'})} \frac{n_{z^{\text{com}}u_{ci}}^{\neg ci} + \delta_{ci}^c}{\sum_u \left(n_{z^{\text{com}}u}^{\neg ci} + \delta_u^c\right)} \frac{n_{z^{\text{com}}v_{ci}}^{\neg ci} + \beta_{ci}^c}{\sum_v \left(n_{z^{\text{com}}v}^{\neg ci} + \beta_v^c\right)} \times \prod_{w \in W_{ci}} \frac{n_{z^{\text{com}}w}^{\neg ci} + \eta_w^c}{\sum_{w'} \left(n_{z^{\text{com}}w'}^{\neg ci} + \eta_{w'}^c\right)} \frac{n_{cx_0^{\omega}}^{\neg ci} + \gamma^{\omega}}{n_{cx_0^{\omega}}^{\neg ci} + n_{cx_1^{\omega}}^{\neg ci} + 2\gamma^{\omega}} \tag{12}$$

where the notation "$\neg ci$" indicates that the ith check-in record specific to city c is excluded from the count.

2) Sampling Switch Indicator and City-Specific Topic Assignment Jointly: If $x_{ci}^{\omega} = 1$, the user, POI, and words in the ith check-in record in city c are generated by the city-specific topics, and the latent variables $\boldsymbol{x}^{\boldsymbol{\omega}}$ and $\boldsymbol{z}^{\boldsymbol{\omega}}$ can be sampled by the following probability:

$$P\left(x_{ci}^{\omega} = 1, z_{ci}^{\omega} = z^{\omega}|\boldsymbol{x}_{\neg ci}, \boldsymbol{z}_{\neg ci}^{\boldsymbol{\omega}}, \text{rest}\right) \propto \frac{n_{ctz^{\omega}}^{\neg ci} + \alpha^{\omega}}{\sum_{z^{\omega'}} \left(n_{ctz^{\omega'}}^{\neg ci} + \alpha^{\omega'}\right)} \frac{n_{z^{\omega}u_{ci}}^{\neg ci} + \delta_{ci}^{\omega}}{\sum_u \left(n_{z^{\omega}u}^{\neg ci} + \delta_u^{\omega}\right)} \frac{n_{z^{\omega}v_{ci}}^{\neg ci} + \beta_{ci}^{\omega}}{\sum_v \left(n_{z^{\omega}v}^{\neg ci} + \beta_v^{\omega}\right)} \times \prod_{w \in W_{ci}} \frac{n_{z^{\omega}w}^{\neg ci} + \eta_w^c}{\sum_{w'} \left(n_{z^{\omega}w'}^{\neg ci} + \eta_{w'}^c\right)} \frac{n_{cx_1^{\omega}}^{\neg ci} + \gamma^{\omega}}{n_{cx_0^{\omega}}^{\neg ci} + n_{cx_1^{\omega}}^{\neg ci} + 2\gamma^{\omega}}. \tag{13}$$

3) Sampling Region Assignment: For the location of the ith check-in record in target city c, a latent region is drawn from the following probability:

$$P(r_{ci} = r|\boldsymbol{r}_{\neg ci}, \text{rest}) \propto \frac{n_{cr}^{\neg ci} + \zeta}{\sum_{r'} \left(n_{cr'}^{\neg ci} + \zeta\right)} P\left(v|\mu_r^{\text{tar}}, \Sigma_r^{\text{tar}}\right) \tag{14}$$

where v indicates the geographical coordinates.

At each iteration, we sample the latent regions so that we can update μ_r^{tar}, Σ_r^{tar} as follows:

$$\mu_r^{\text{tar}} = \frac{1}{\left|S_r^{\text{tar}}\right|} \sum_{v \in S_r^{\text{tar}}} v$$
$$\Sigma_r^{\text{tar}} = \frac{1}{\left|S_r^{\text{tar}}\right| - 1} \sum_{v \in S_r^{\text{tar}}} \left(v - \mu_r^{\text{tar}}\right)\left(v - \mu_r^{\text{tar}}\right)^T \tag{15}$$

where S_r^{tar} denotes the set of POIs assigned to region r in the target city, and v indicates the geographical coordinates.

After a sufficient number of sampling iterations, the parameters can be learned by the approximated posterior and they are formulated by the following formulas:

$$\hat{\theta}_{tz}^{\text{com}} = \frac{n_{tz} + \alpha^{\text{com}}}{\sum_{z'} \left(n_{tz} + \alpha^{c'}\right)}, \quad \hat{\theta}_{ctz}^{\omega} = \frac{n_{cz} + \alpha^{\omega}}{\sum_{z'} \left(n_{ctz} + \alpha^{\omega'}\right)}$$
$$\hat{\psi}_{zu}^{\omega} = \frac{n_{zu} + \delta_u^{\omega}}{\sum_{u'} \left(n_{zu'} + \delta_{u'}^{\omega}\right)}, \quad \hat{\varphi}_{zv}^{\omega} = \frac{n_{zv} + \beta_v^{\omega}}{\sum_{v'} \left(n_{zv'} + \beta_{v'}^{\omega}\right)}$$
$$\hat{\phi}_{zw}^{\omega} = \frac{n_{zw} + \eta_w^{\omega}}{\sum_{w'} \left(n_{zw'} + \eta_{w'}^{\omega}\right)}, \quad \hat{\lambda}_c^{\omega} = \frac{n_{cx_0}^{\omega} + \gamma^{\omega}}{n_{cx_0}^{\omega} + n_{cx_1}^{\omega} + 2\gamma^{\omega}}$$
$$\hat{\vartheta}_{cr} = \frac{n_{cr} + \zeta_r}{\sum_{r'} (n_{cr'} + \zeta_{r'})}. \tag{16}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE III
STATISTICS OF DATASETS

	Foursquare	San Francisco	Los Angeles	Twitter	Boston	San Diego
Total users	4163	1946	1987	114,508	1940	1378
Total POIs	21,142	7858	7907	62,462	456	481
Vocabulary size	641	389	352	7490	804	820
Total cities	4,368	–	–	4955	–	–
Total check-ins	483,813	59949	51145	1,434,668	8744	7055

D. Recommendation Based on CTLM

The task of crossing-city POI recommendations is to offer a list of POIs in the target city based on the users' check-in records in the source cities. According to the conditional independence and (2), given the target city c and a specific time t, the probability that the user u will visit the POI v in the target city can be computed as

$$\begin{aligned} &P(v, u|c, t, \Theta)\\ &= \Bigg[P(x=0|\lambda_c)\sum_{z=1}^{K^{\text{com}}} P\big(v, u|z, \varphi_z^{\text{com}}, \psi_z^{\text{com}}\big)P\big(z|\theta^{\text{com}}, t\big)\\ &\quad + P(x=1|\lambda_c)\sum_{z=1}^{K^{\text{tar}}} P\big(v, u|z, \varphi_z^{\text{tar}}, \psi_z^{\text{tar}}\big)P\big(z|\theta_c^{\text{tar}}, t\big)\Bigg]\\ &\quad \times \Bigg[\sum_r P(v|r)P(r|\vartheta_{cr})\Bigg]. \end{aligned} \tag{17}$$

To estimate the probability $P(v|r)$ based on the geographical coordinates associated with v, we introduce multinomial distribution ψ_{cr}^{v}, describing the POI distributions specific to region r. The distribution can be estimated in the region sampling process, and to avoid overfitting we compute it as the following form:

$$\hat{\psi}_{cr}^{v} = \frac{n_{crv} + \zeta^{v}}{\sum_{v'}\big(n_{crv'} + \zeta^{v'}\big)} \tag{18}$$

where n_{crv} counts the number of times POI v which is assigned to region r in city c.

Finally, the recommendation score in (17) can be computed as

$$\begin{aligned} &P(v, u|c, t, \Theta)\\ &= \Bigg[\sum_r \hat{\psi}_{cr}^{v}\hat{\vartheta}_{cr}\Bigg]\Bigg[\hat{\lambda}_c^{t}\sum_{z=1}^{K^{\text{com}}} \hat{\varphi}_{zv}^{\text{com}}\hat{\psi}_{zu}^{\text{com}}\hat{\theta}_{tz}^{\text{com}}\\ &\quad + \Big(1-\hat{\lambda}_c^{t}\Big)\sum_{z=1}^{K^{\text{tar}}} \hat{\varphi}_{zv}^{\text{tar}}\hat{\psi}_{zu}^{\text{tar}}\hat{\theta}_{ctz}^{\text{tar}}\Bigg]. \end{aligned} \tag{19}$$

After the probability is computed, the top-k POIs with the highest scores are recommended to the user.

E. Time Complexity Analysis

We analyze the time complexity of the inference process. Let I be the number of iterations conducted by collapsed Gibbs sampling until convergence. In each iteration, we need to go through all of the check-in records. Each check-in record requires assigning a switch indicator, a latent topic, and a latent region if the check-in record is in the target city. Therefore, the overall time complexity is $O(I(K\sum_c |D_c| + R|D_{\text{tar}}|))$, where K is the number of topics, R is the number of regions, D_c is the number of check-ins in city c, and $|D_{\text{tar}}|$ is the number of check-in records in the target city.

IV. EXPERIMENTS

In this section, we conduct extensive experiments to evaluate the effectiveness of our proposed model and make a comparison with the state-of-the-art POI recommendation techniques on two real-world datasets.

A. Experimental Settings

1) Datasets: In this paper, we use Foursquare and Twitter datasets which are publicly available[3] to conduct our experiments and each check-in record is formulated as user-ID, POI-ID, time, contents, location, and city. The contents in our experiments refer to the textual categories of POIs in the Foursquare dataset, and both the textual categories and the textual tags of POIs in the Twitter dataset, respectively. The statistics of the two datasets are demonstrated in Table III.

In the scenario of crossing-city POI recommendations, we select San Francisco and Los Angeles from the Foursquare dataset, and Boston and San Diego from the Twitter dataset as the target cities since they have the most ground truth users (i.e., crossing-city users). First, we choose the crossing-city users who have visited both the target city and the source cities, and regard their check-ins in the target city as the ground truth. Then, we divide the ground-truth into two disjoint groups: one for tuning parameters (number of regions and number of topics) and the other for algorithm evaluation. The group for algorithm evaluation is further split into five disjoint sets to conduct five-fold cross-validation. In this way, we can have five different testing results and all of the results together can also be used for the significance testing. Compared with the ground truth, the measure metrics of F1, MAP, and nDCG are reported, and then t-testing is conducted to determine whether the differences between CTLM and the state-of-the-art algorithms are statistically significant.

2) State-of-the-Art Algorithms: We compare our proposed method with the following five state-of-the-art approaches.

1) LCA-LDA [13] was proposed to provide out-of-town POI recommendations. This method employed the topic model to incorporate multiple factors, that is, personal interests and local preferences of each city by the medium of POIs and their contents.

[3] https://sites.google.com/site/dbhongzhi/

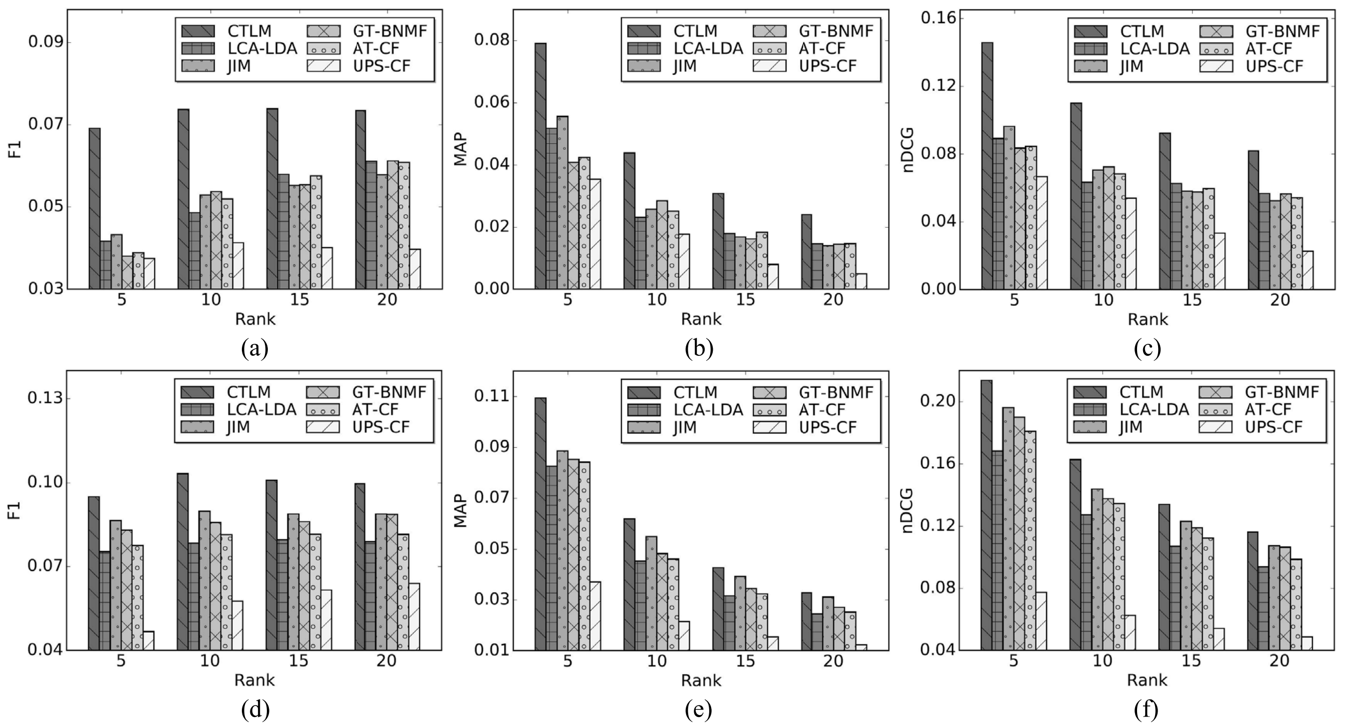


Fig. 3. Performance of all algorithms on the Foursquare. (a) F1 on San Francisco. (b) MAP on San Francisco. (c) nDCG on San Francisco. (d) F1 on Los Angeles. (e) MAP on Los Angeles. (f) nDCG on Los Angeles.

2) JIM [14] was developed to jointly model users' check-in behaviors for out-of-town POI recommendations, which incorporated temporal effect, geographical-social influence, and content effect together.
3) GT-BNMF [2] combined topic model and Bayesian non-negative matrix factorization to offer personalized POI recommendations. They jointly took user interests, POI popularity, geographical influence, and user mobility behaviors into consideration.
4) AT-CF [24] was an author-topic-based collaborative filtering method to facilitate the crossing city POI recommendations. In this model, the author-topic technique was exploited to extract users' preference, and then the similar users in the target city were detected to overcome the data sparsity problem.
5) UPS-CF [8] was a user-based CF method to make POI recommendations for the out-of-town users. To overcome the data sparsity problem, it linearly integrated the interests of the target users with those of their social friends.

To study the impact of the factors considered in our proposed model, we also construct the following three variants of our CTLM as baselines for the comparison.

1) CTLM-1 is the simplified version of CTLM by only taking into account the city-specific topics and the spatial influence of regions in each city, but ignoring the common topics shared by all cities.
2) CTLM-2 is the simplified version of CTLM by only taking into account the common topics and the spatial influence of regions in each city, but ignoring the city-specific topic of each city.
3) CTLM-3 ignores the spatial influence of POIs and only models each city as a mixture of common topics and city-specific topics to generate the POIs, users, and contents.

3) Parameter Settings: Based on the tuning results, the number of topics (i.e., the number of common topics K^{com}, the number of the source city-specific topics K^{sou}, and the number of the target city-specific topics K^{tar}) is set to 20 and the number of regions R is set to 30 on the Foursquare dataset, while they are all set to 10 on the Twitter dataset. In both datasets, hyper-parameters $\{\alpha^{\text{com}}, \alpha^{\omega}, \zeta, \beta, \eta, \delta, \gamma\}$ are empirically set as $\{50/K^{\text{com}}, 50/K^{\omega}, 50/R, 0.01, 0.01, 0.01, 0.5\}$.

4) Evaluation Metrics: To evaluate the performance of different approaches for the crossing-city POI recommendations, we adopt four standard measuring metrics, that is, F1, MAP, and nDCG. The definitions are formulated as follows:

$$\text{F1@}k = \frac{2 \times \text{Precision@}k \times \text{Recall@}k}{\text{Precision@}k + \text{Recall@}k}$$

$$\text{MAP@}k = \frac{\sum_{u=1}^{U} \text{AP@}k}{U}, \text{AP@}k = \frac{1}{k}\sum_{i=1}^{k} \text{Precision@}k$$

$$\text{nDCG@}k = \frac{\text{DCG@}k}{\text{IDCG@}k}, \text{DCG@}k = \sum_{i=1}^{k} \frac{2^{\text{rel}_i} - 1}{\log_2(i+1)}$$

where Recall@k and Precision@k are defined as $\text{Recall@}k = \sum_u (|S_{u,k} \cap S_{u,\text{visited}}|/|S_{u,\text{visited}}|)$ and $\text{Precision@}k = \sum_u (|S_{u,k} \cap S_{u,\text{visited}}|/k)$ ($S_{u,k}$ denotes the set of top-k POIs in the target city recommended to the user u, $S_{u,\text{visited}}$ indicates the ground truth of user u in the testing set), IDCG@k is the DCG@k value when the recommended POIs are ideally ranked, and rel_i refers to the graded relevance of the result ranked at position i.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

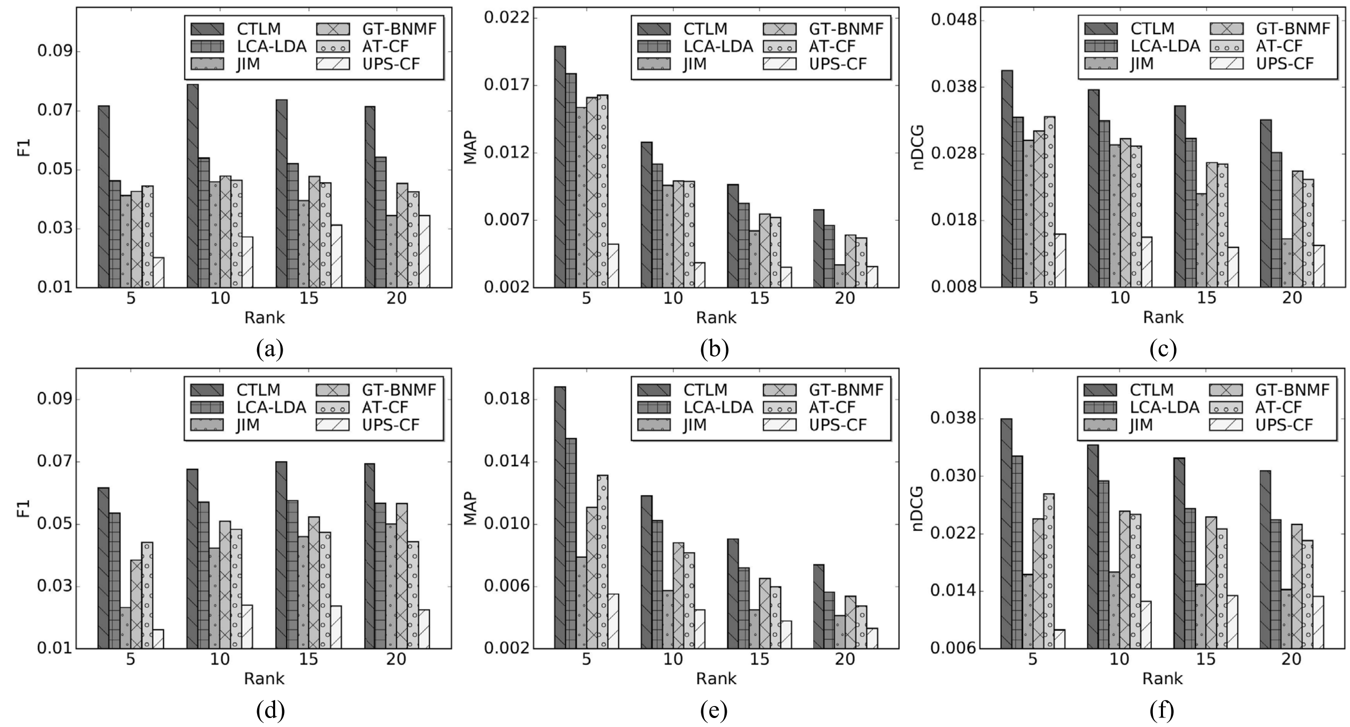


Fig. 4. Performance of all algorithms on Twitter. (a) F1 on Boston. (b) MAP on Boston. (c) nDCG on Boston. (d) F1 on San Diego. (e) MAP on San Diego. (f) nDCG on San Diego.

To test whether the difference between CTLM and the state-of-the-art algorithms is statistically significant, we adopt *t*-testing [25] and a *p*-value can be found from Student's *t*-distribution. If the calculated *p*-value is below the threshold chosen for statistical significance (usually 0.10, 0.05, or 0.01 level), then the differences are statistically significant.

B. Experimental Results

In this part, we demonstrate the experimental results of the aforementioned approaches conducted on the Foursquare and Twitter datasets.

1) Performance Comparisons: We make a comparison between CTLM and the five state-of-the-art POI recommendation methods. Fig. 3 shows the F1@k ($k = 5, 10, 15, 20$) of all the methods on the Foursquare and Twitter datasets, respectively.

On the Foursquare dataset, Fig. 3(a)–(c) shows the performance of all methods on San Francisco and Fig. 3(d)–(f) presents the results on Los Angeles. F1@5 of CTLM are about 0.07 and 0.09 on San Francisco and Los Angeles, respectively.

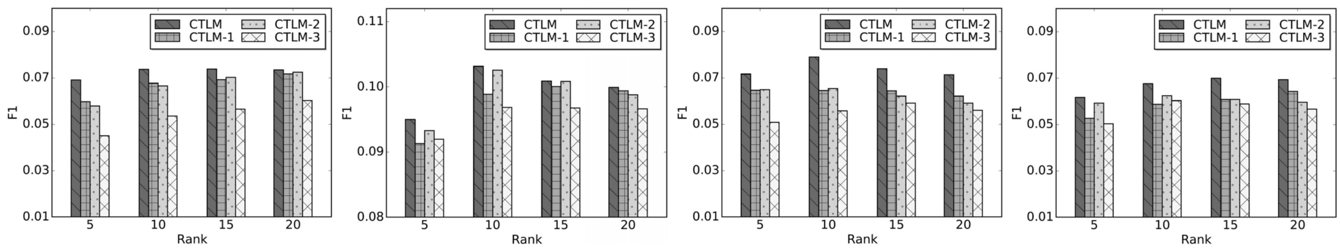
It is observed from the figures that CTLM significantly outperforms the other four methods in terms of both recall and Precision. For example, compared with LCA-LDA, JIM, GT-BNMF, AT-CF, and UPS-CF, CTLM improves F1@10 by 51.7%, 39.3%, 37.2%, 41.8%, and 78.3% on San Francisco, and 31.6%, 15.0%, 20.3%, 26.7%, and 78.8% on Los Angeles. The observations are summarized as follows.

1) CTLM outperforms LCA-LDA, which indicates the advantage of separating the specificities from the common features of all cities to overcome the ill-matching problem. Besides that, LCA-LDA ignores the factor of regions' accessibility compared with CTLM.
2) Both JIM and GT-BNMF drop behind CTLM although both of them consider the effect of topics and regions. On one hand, this is due to the CTLM separating the specificities from the common features of all cities, which can overcome the ill-matching problem between users from different cities in terms of their real interests. On the other hand, both JIM and GT-BNMF model users' mobility patterns by learning their regional preference. However, for the crossing-city POI recommendations, the users have no check-in records in the target city, as the result users' regional preference is not available.
3) AT-CF performs worse than CTLM and LCA-LDA since it depends on users' interests to provide crossing city recommendations and fails to consider the effect of city-specific features when users travel to a new city.
4) The performance of UPS-CF is the worst among all of the algorithms. The main reason is that it employs the CF method and only exploits users' visited POIs to make crossing-city POI recommendations without incorporating the contents, which suffers from serious data sparsity problem as there are few crossing-city users.
5) In terms of nDCG and MAP, our algorithm outperforms the state-of-the-art algorithms.

For the Twitter dataset, Fig. 4(a)–(c) shows the performances of all methods on Boston, and Fig. 4(d)–(f) presents the results on San Diego. Compared with LCA-LDA, JIM, GT-BNMF, AT-CF, and UPS-CF, CTLM improves F1@10 by 46.3%, 72.1%, 64.9%, 70.0%, and 189.1% on Boston, and

TABLE IV
SIGNIFICANCE TESTING OF ALGORITHMS

	p-value of F1 on San Francisco				p-value of F1 on Boston			
	5	10	15	20	5	10	15	20
LCA-LDA	**2.10e-05**	**1.08e-05**	**0.0002**	**0.0086**	**0.0014**	**0.0006**	**0.0002**	**0.0029**
JIM	**0.0003**	**0.0030**	**0.0025**	**0.0019**	**0.0113**	**0.0002**	**1.09e-05**	**5.27e-07**
GT-BNMF	**7.16e-06**	**0.0057**	**0.0001**	**0.0005**	**0.0003**	**0.0007**	**6.569e-05**	**3.73e-06**
AT-CF	**1.36e-06**	**3.20e-06**	**1.83e-05**	**0.0001**	**0.0003**	**2.39e-05**	**2.79e-06**	**8.25e-07**
UPS-CF	**8.55e-07**	**7.83e-08**	**2.21e-08**	**1.67e-08**	**1.48e-06**	**3.65e-07**	**5.85e-08**	**7.96e-08**



(a) (b) (c) (d)

Fig. 5. Impact of different factors. F1 on (a) San Francisco, (b) Los Angeles, (c) Boston, and (d) San Diego.

TABLE V
SIGNIFICANCE TESTING OF VARIANTS

	p-value of F1 on San Francisco				p-value of F1 on Boston			
	5	10	15	20	5	10	15	20
CTLM-1	**0.0048**	**0.0150**	**0.0255**	0.2891	0.3197	**0.0251**	0.1073	0.1103
CTLM-2	**0.0009**	**0.0036**	**0.0965**	0.5934	0.3802	**0.0235**	**0.0387**	**0.0170**
CTLM-3	**1.55e-05**	**2.90e-05**	**0.0003**	**0.0021**	**0.0014**	**0.0015**	**0.0221**	**0.0105**

18.5%, 60.0%, 32.7%, 40.0%, and 181.5% on San Diego. Fig. 4 shows similar trends as in Fig. 3. However, it is worth noting that compared with UPS-CF, the improvement of CTLM is much higher on the Twitter dataset than on the Foursquare dataset, showing the significance of social friends for the crossing-city POI recommendations as there is no social relationship in the Twitter dataset and UPS-CF degenerates to the CF-based method.

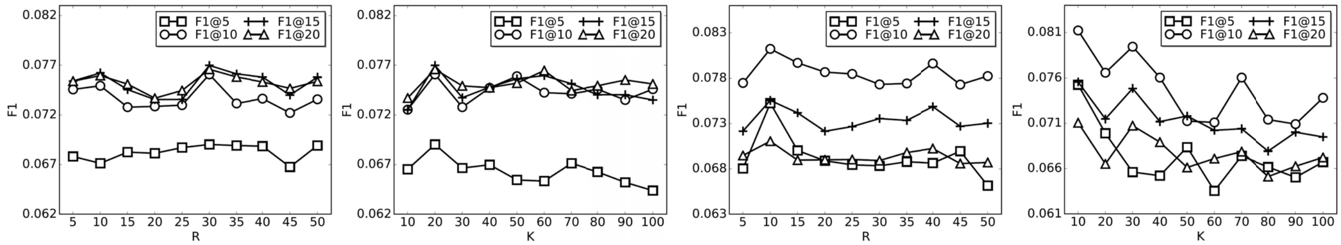
Table IV shows the t-testing results on San Francisco from the Foursquare dataset and Boston from the Twitter dataset, respectively, and we find that the p-value is very small even smaller than 0.01, indicating that the superiority of CTLM over the-state-of-the-art algorithms is statically significant. Due to space limitations, we only show the t-testing result on San Francisco and Boston, and the t-testing results for Los Angeles and San Diego are significant as well.

2) Factor Analysis: To further explore the superiority of the techniques used in our model CTLM, we compare it with our designed baseline methods, that is, CTLM-1, CTLM-2, and CTLM-3, respectively. Fig. 5 reports F1 on Foursquare and Twitter datasets, respectively. Observed from the results, CTLM is consistently superior to the other three methods on both the Foursquare dataset and Twitter dataset. For example, the improvements in terms of F1@5, are 10.8%, 10.4%, and 40.9% compared with CTLM-1, CTLM-2, and CTLM-3 on Boston from the Twitter dataset. Besides, the results show that the factors considered are important to improve the performance of CTLM. We find that CTLM, CTLM-1, and CTLM-2 consistently outperform CTLM-3, indicating the benefits of introducing the spatial influence of regions in each city.

To validate the significance of factors considered in CTLM, we conduct t-testing by comparing CTLM with the other three variants on San Francisco from the Foursquare dataset and on Boston from the Twitter dataset, and the results are listed in Table V. Observed from the results, the difference between CTLM and CTLM-3 on both cities is statistically significant, indicating that spatial influence is the most important factor to improve performance of CTLM. In addition, the city-specific topic factor is more statistically significant compared with the common topic factor since p-values for CTLM-2 are smaller than those for CTLM-1.

3) Parameters Analysis: In our model, there are four parameters, for example, the number of common topics (K^{com}), the number of the source city-specific topics (K^{sou}), the number of the target city-specific topics (K^{tar}), and the number of regions (R) in the target city, and tuning them is critical to the performance of the CTLM model. As for the hyper-parameters α^{com}, α^{ω}, ζ, γ, β, δ, and η, we empirically set them as fixed values. To study the effect of these parameters in CTLM, we conduct experiments on San Francisco from the Foursquare and on Boston from the Twitter dataset by varying these parameters and we fix $K^{\text{com}} = K^{\text{sou}} = K^{\text{tar}}$ for simplicity. The experimental results are shown in Fig. 6.

We first study the impact of the number of regions R in the target city on the performance of recommendations. Fig. 6(a) and (b) shows the F1 for San Francisco on the



(a) (b) (c) (d)

Fig. 6. Impact of parameters. (a) Impact of the region number for F1 on San Francisco. (b) Impact of the topic number for F1 on San Francisco. (c) Impact of the region number for F1 on Boston. (d) Impact of the topic number for F1 on Boston.

Foursquare dataset by varying R from 5 to 50 and K from 10 to 100, respectively. Observed from the results, our algorithm performs best when the number of regions reaches 30 and the number of topics is 20. As a result, we set R as 30 and K as 20 on the Foursquare dataset.

Fig. 6(c) and (d) reports the performance of CTLM on the Twitter dataset. Obviously, F1 reaches a peak when both the number of regions and the number of topics are 10. Therefore, the number of regions and the number of topics are set as 10 and 10 on the Twitter dataset.

4) Latent Topics Analysis: In this section, we demonstrate the city-specific topics in Boston and common topics shared by all cities on the Twitter dataset, which are discovered by CTLM.

Table VI shows four city-specific topics in Boston and four common topics shared on the Twitter dataset. Each topic is represented by the top ten words with highest probabilities. Observed from Table VI, the city-specific topics in Boston learned by CTLM are semantically coherent to describe the city-specific POIs while the common topics indicate the ubiquitous functioned POIs across different cities. For example, the four city-specific topics in Boston are about Faneuil Hall Marketplace, Chinatown, Charles River, and Boston Common, respectively. In contrast, the four common topics shared by all of the cities represent nightlife activities, coffee shop, movies, and outdoor activities, respectively. The differences between the city-specific topics and the common topics are that it is necessary to separate the two kinds of topics for the crossing city POI recommendations. On one hand, the words representing the common topics are more frequent than those indicating the city-specific topics; as a result, they will dominate the discovered topics and overwhelm the city-specific features if mixed together. On the other hand, if the words representing city-specific topics are mixed with those about common topics, users' real interests will become difficult to identify.

5) Recommendation Cases Illustration: In this section, we present some crossing-city recommendation examples in Table VII, where both users come from San Jose and their target city is San Francisco. Observed from their check-in records in San Jose, they have different interests, that is, user 77 is interested in sports and movies while user 234 prefers to visit music-related POIs, but both have visited the city-specific POI 4231. The ground truth in San Francisco indicates that their check-in behavior is not only driven by their real interests but also by the city-specific features; thus, our proposed algorithm recommends a list of POIs balancing the two aspects. For example, CTML recommends to both of the users the bridge-related POI which is famous in San Francisco. Furthermore, based on their different tastes, sport-related POI is recommended to user 77 while music-related POI is selected for user 234.

V. Related Work

Recently, the availability of users' real-time locations fosters the development of LBSNs and many techniques have been exploited to improve POI recommendations. Based on the techniques employed in POI recommendations, the works can be divided into two broad categories, that is, pure CF approaches and content-assisted CF techniques.

With respect to pure CF approaches, Ye *et al.* [26], [27] proposed a unified framework by fusing factors of user preference, social influence, and geographical influence to perform collaborative POI recommendations as they assumed that in a new city, users intend to visit the places that their friends visited. In a similar way, the CF-based frameworks were proposed in [8] and [28] to provide POI recommendations based on the check-in records generated by both social friends and similar users in a unified way. Yuan *et al.* [29], [30] proposed a time-aware collaborative recommendation model to recommend POIs to a given user at a specified time in a day by incorporating temporal and geographical influences. Instead of using check-in records directly, matrix factorization [31] is utilized to provide POI recommendations. Liu *et al.* [2], [4], [5] proposed a Bayesian non-negative matrix factorization framework by considering user preferences, POI popularity, geographical influence, and user mobility behaviors to imitate the decision process of a user's visit to a POI while IRenMF [32] proposed utilizing a weighted matrix factorization framework to server CF with implicit feedback and Lian *et al.* [6] extended their work by incorporating spatial clustering phenomenon. Besides, Li *et al.* [7] performed a similar work called Rank-GeoFM by incorporating temporal influence. To improve POI recommendation accuracy for the cold-start problem, Li *et al.* [33] incorporated check-ins of users' friends, that is, social friends, location friends, and neighboring friends, into the matrix factorization model to deal with the challenge from data sparsity. Pham *et al.* [34] exploited the influence between POIs to make accurate out-of-town region recommendations and proposed sweeping line-based methods to improve the efficiency of searching for the best region. Liu *et al.* [35]

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE VI
LATENT TOPICS ANALYSIS ON THE TWITTER DATASET

City-specific Topics in Boston				Common Topics Shared in Twitter			
topic 1	topic 2	topic	topic 4	topic 1	topic 2	topic3	topic 4
trail	chinese	bridge	playground	bar	coffee	theater	park
market	dumplings	river	boston	music	caf	movie	outdoor
historic site	asian	charles	common	nightclub	tea	cineplex	garden
shopping	boston	cambridge	harvard	live	starbuck	multiplex	playground
national site	pork	dock	ski	lounge	bakery	photobooth	dog run
street	noodles	boating	university	beer	espresso	cinema	scenic
freedom	chinatown	marina	history	dancing	dessert	popcorn	kids
performers	brunch	harbor	museum	rock	sandwiches	photo	views
fanueil	cuisine	sailing	library	gay	breakfast	imax	water
quincy	hot tea	kayak	presidential	cocktail	free	food	bbq

TABLE VII
RECOMMENDATION EXAMPLES FOR USERS TO SAN FRANCISCO

User ID	User 77		User 234	
	The Common POIs	The City-specific POIs	The Common POIs	The City-specific POIs
Check-in Records	2474(baseball stadium, baseball field, general entertainment) 4321(movie theater, multiplex)	4656(historic site, plaza, monument, landmark) 4231(music venue, rock club, multiplex)	14630(concert hall, music venue) 14640(nightlife spot, music venue, arts)	4231(music venue, rock club, multiplex) 14704(shop, services, electronics store)
Ground Truth	**183(baseball stadium)** **884(movie theater multiplex)**	**185(bridge)** **2309(monument, shop, landmark, farmer market)**	11839(building, tech startup) **4659(music venue, rock club)**	**185(bridge)** 2309(monument, shop, landmark, farmer market)
POI List	**183(baseball stadium)** **884(movie theater, multiplex)**	**185(bridge)** **2309(monument, shop, landmark, farmer market)**	**4659(music venue, rock club)** 426(coffee shop, breakfast spot)	**185(bridge)** 2252(playground, scenic lookout, park)

provided an all-around evaluation of the state-of-the-art POI recommendation models for different scenarios. Recently, CF-based deep learning techniques are implemented in the POI recommendation systems. Specifically, Wang *et al.* [36] proposed a collaborative deep learning model to jointly perform deep representation learning for the content information and the CF for the ratings (feedback) matrix. Liu *et al.* [37] extended RNN and proposes a novel spatial temporal recurrent neural networks model, which can enable the spatial and temporal contextual information transition in each layer.

To improve the performance of POI recommendations, an alternative approach is to incorporate contents of POIs in the algorithm, which can overcome the data sparsity to some extent. Many recent studies [26], [38]–[40] fused contents of POIs into the collaborative recommendation framework. The model proposed in [1] extended the CF-based approach by mapping the users in the home city to a group of users in the target city by the content similarities of their profiles. Similarly, Zhao *et al.* [9] grouped both users in the home city and those in the target city into communities by inferring their interests from contents, and sought an optimal match between communities of the home city and those of the target city in order to fulfill the crossing-city POI recommendations. Gao *et al.* [10] integrated the user-POI matrix, user-word matrix, and POI-word matrix into a unified matrix factorization framework to support POI recommendations, while Xie *et al.* [41] embedded the POI-POI, POI-region, POI-time, and POI-word graphs into a unified model assisting with embedding learning techniques. Hu and Ester [12], Kurashima *et al.* [42], and Liu *et al.* [43] proposed a spatial topic model by capturing users' mobility patterns and their interests for location recommendations. Similarly, works [44], [45] considered spatial–temporal–semantic information of tweets for the task of location predictions. Jiang *et al.* [11], [24] designed an author topic model-based collaborative filtering method to facilitate the crossing city POI recommendations. In this model, the author topic model was applied to extract the users' preference, then the similar users in the target city were detected to overcome the data sparsity problem. In their recent work [46], they exploited the textual descriptions of POIs to learn users personalized interests, based on which CF-based method is applied to recommend a travel sequence. Yin *et al.* [13] incorporated contents of POIs into a topic model framework to learn users' interests and local preferences to provide out-of-town POI recommendations. Recently, Yin *et al.* [14], [47] and Wang *et al.* [15], [48] further developed their model by introducing the geographical influence into the topic model to learn users' mobility patterns for both hometown and out-of-town POI recommendations.

All of the existing works [1], [8], [9], [13]–[15], [47], and [48] for out-of-town recommendations are different from our proposed CTLM as our model distinguishes the specific topics of each city from the common topics shared by all cities in the collection and transfers users' real interests from the source cities to the target city by the medium of the common topics to address the ill-matching problem. Moreover, we introduce the spatial influence of regions in each city to meet the users' needs for accessibility in the target city.

VI. Conclusion

In this paper, we propose a CTLM for the crossing-city POI recommendations by adopting the transfer learning techniques. The proposed model separates the city-specific topics of each city from the common topics shared by all cities, and the users' real interests in the source city can be transferred to the target city by the medium of common topics; as a result, the ill-matching problem between users and POIs across cities can be well addressed. Furthermore, we incorporate the spatial–temporal influence with the topic model in a unified way; as a result, the co-occurrence patterns of users and POIs are modeled as the aggregated result from these factors. We conduct extensive experiments to evaluate the performance of CTLM on Foursquare and Twitter datasets, and the results validate the advantages of CTLM for the crossing-city POI recommendations. Also, to investigate the impact of different factors considered in our model, we conduct experiments on three variants of CTLM and find that all of those factors are essential to improve the crossing-city POI recommendations.

Acknowledgment

The authors would like to thank the anonymous reviewers and editor for their insightful comments on the revision of this paper.

References

[1] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *Proc. ACM GIS*, 2012, pp. 199–208.

[2] B. Liu, Y. Fu, Z. Yao, and H. Xiong, "Learning geographical preferences for point-of-interest recommendation," in *Proc. ACM SIGKDD*, 2013, pp. 1043–1051.

[3] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proc. ACM WWW*, 2001, pp. 285–295.

[4] B. Liu and H. Xiong, "Point-of-interest recommendation in location based social networks with topic and location awareness," in *Proc. SDM*, 2013, pp. 396–404.

[5] B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, "A general geographical probabilistic factor model for point of interest recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 5, pp. 1167–1179, May 2015.

[6] D. Lian *et al.*, "GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation," in *Proc. ACM SIGKDD*, 2014, pp. 831–840.

[7] X. Li, G. Cong, X.-L. Li, T.-A. N. Pham, and S. Krishnaswamy, "Rank-GeoFM: A ranking based geographical factorization method for point of interest recommendation," in *Proc. ACM SIGIR*, 2015, pp. 433–442.

[8] G. Ference, M. Ye, and W.-C. Lee, "Location recommendation for out-of-town users in location-based social networks," in *Proc. ACM CIKM*, 2013, pp. 721–726.

[9] Y.-L. Zhao, L. Nie, X. Wang, and T.-S. Chua, "Personalized recommendations of locally interesting venues to tourists via cross-region community matching," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, p. 50, 2014.

[10] H. Gao, J. Tang, X. Hu, and H. Liu, "Content-aware point of interest recommendation on location-based social networks," in *Proc. AAAI*, 2015, pp. 1721–1727.

[11] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model-based collaborative filtering for personalized POI recommendations," *IEEE Trans. Multimedia*, vol. 17, no. 6, pp. 907–918, Jun. 2015.

[12] B. Hu and M. Ester, "Spatial topic modeling in online social media for location recommendation," in *Proc. ACM RecSys*, 2013, pp. 25–32.

[13] H. Yin, Y. Sun, B. Cui, Z. Hu, and L. Chen, "LCARS: A location-content-aware recommender system," in *Proc. ACM SIGKDD*, 2013, pp. 221–229.

[14] H. Yin, X. Zhou, Y. Shao, H. Wang, and S. Sadiq, "Joint modeling of user check-in behaviors for point-of-interest recommendation," in *Proc. ACM CIKM*, 2015, pp. 1631–1640.

[15] W. Wang *et al.*, "Geo-SAGE: A geographical sparse additive generative model for spatial item recommendation," in *Proc. ACM SIGKDD*, 2015, pp. 1255–1264.

[16] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[17] O. Jin, N. N. Liu, K. Zhao, Y. Yu, and Q. Yang, "Transferring topical knowledge from auxiliary long texts for short text clustering," in *Proc. ACM CIKM*, 2011, pp. 775–784.

[18] Y. Wei, Y. Zheng, and Q. Yang, "Transfer knowledge between cities," in *Proc. SIGKDD*, 2016, pp. 1905–1914.

[19] P. Hao, G. Zhang, L. Martinez, and J. Lu, "Regularizing knowledge transfer in recommendation with tag-inferred correlation," *IEEE Trans. Cybern.*, to be published.

[20] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Jan. 2003.

[21] C. Chemudugunta, P. Smyth, and M. Steyvers, "Modeling general and specific aspects of documents with a probabilistic topic model," in *Proc. Adv. Neural Inf. Process. Syst.*, 2007, pp. 241–248.

[22] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. ACM SIGKDD*, 2011, pp. 1082–1090.

[23] T. L. Griffiths and M. Steyvers, "Finding scientific topics," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. S1, pp. 5228–5235, 2004.

[24] S. Jiang, X. Qian, J. Shen, and T. Mei, "Travel recommendation via author topic model based collaborative filtering," in *Proc. Int. Conf. Multimedia Model.*, 2015, pp. 392–402.

[25] R. V. Hogg and A. T. Craig, *Introduction to Mathematical Statistics*, 5th ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1995.

[26] M. Ye, P. Yin, and W.-C. Lee, "Location recommendation for location-based social networks," in *Proc. ACM SIGSPATIAL*, 2010, pp. 458–461.

[27] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proc. ACM SIGIR*, 2011, pp. 325–334.

[28] H. Li, R. Hong, S. Zhu, and Y. Ge, "Point-of-interest recommender systems: A separate-space perspective," in *Proc. IEEE ICDM*, 2015, pp. 231–240.

[29] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in *Proc. ACM SIGIR*, 2013, pp. 363–372.

[30] Q. Yuan, G. Cong, and A. Sun, "Graph-based point-of-interest recommendation with geographical and temporal influences," in *Proc. ACM CIKM*, 2014, pp. 659–668.

[31] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[32] Y. Liu, W. Wei, A. Sun, and C. Miao, "Exploiting geographical neighborhood characteristics for location recommendation," in *Proc. ACM CIKM*, 2014, pp. 739–748.

[33] H. Li, Y. Ge, and H. Zhu, "Point-of-interest recommendations: Learning potential check-ins from friends," in *Proc. ACM SIGKDD*, 2016, pp. 975–984.

[34] T.-A. N. Pham, X. Li, and G. Cong, "A general model for out-of-town region recommendation," in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 401–410.

[35] Y. Liu, T.-A. N. Pham, G. Cong, and Q. Yuan, "An experimental evaluation of point-of-interest recommendation in location-based social networks," *Proc. VLDB Endowment*, vol. 10, no. 10, pp. 1010–1021, 2017.

[36] H. Wang, N. Wang, and D.-Y. Yeung, "Collaborative deep learning for recommender systems," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2015, pp. 1235–1244.

[37] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: A recurrent model with spatial and temporal contexts," in *Proc. AAAI*, 2016, pp. 194–200.

[38] X. Liu, Y. Liu, K. Aberer, and C. Miao, "Personalized point-of-interest recommendation by mining users' preference transition," in *Proc. ACM CIKM*, 2013, pp. 733–738.

[39] J.-D. Zhang and C.-Y. Chow, "GeoSoCa: Exploiting geographical, social and categorical correlations for point-of-interest recommendations," in *Proc. ACM SIGIR*, 2015, pp. 443–452.

[40] J.-D. Zhang, C.-Y. Chow, and Y. Zheng, "ORec: An opinion-based point-of-interest recommendation framework," in *Proc. ACM CIKM*, 2015, pp. 1641–1650.

[41] M. Xie *et al.*, "Learning graph-based POI embedding for location-based recommendation," in *Proc. ACM CIKM*, 2016, pp. 15–24.

[42] T. Kurashima, T. Iwata, T. Hoshide, N. Takaya, and K. Fujimura, "Geo topic model: Joint modeling of user's activity area and interests for location recommendation," in *Proc. ACM WSDM*, 2013, pp. 375–384.

[43] Y. Liu, M. Ester, B. Hu, and D. W. Cheung, "Spatio-temporal topic models for check-in data," in *Proc. IEEE ICDM*, 2015, pp. 889–894.

[44] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Who, where, when and what: Discover spatio-temporal topics for twitter users," in *Proc. ACM SIGKDD*, 2013, pp. 605–613.

[45] Q. Yuan, G. Cong, K. Zhao, Z. Ma, and A. Sun, "Who, where, when, and what: A nonparametric Bayesian approach to context-aware recommendation and search for Twitter users," *Trans. Inf. Syst.*, vol. 33, no. 1, p. 2, 2015.

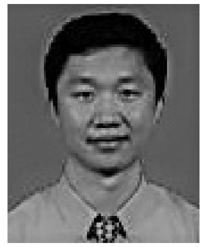
[46] S. Jiang, X. Qian, T. Mei, and Y. Fu, "Personalized travel sequence recommendation on multi-source big social media," *IEEE Trans. Big Data*, vol. 2, no. 1, pp. 43–56, Mar. 2016.

[47] H. Yin *et al.*, "Adapting to user interest drift for POI recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 10, pp. 2566–2581, Oct. 2016.

[48] H. Wang *et al.*, "A location-sentiment-aware recommender system for both home-town and out-of-town users," in *Proc. ACM SIGKDD*, 2017, pp. 1135–1143.

Dichao Li is currently pursuing the Ph.D. degree in computer science with the Department of Computer and Information Science, University of Macau, Macau, China, under the supervision of Prof. Z. Gong.

Her current research interest includes data mining.



Zhiguo Gong (M'10–SM'16) received the Ph.D. degree in computer science from the Institute of Mathematics, Chinese Academy of Science, Beijing, China.

He is currently a Professor with the Faculty of Science and Technology, University of Macau, Macau, China. His current research interests include machine learning, data mining, database, and information retrieval.

Defu Zhang received the bachelor's and master's degrees in computational mathematics from Xiangtan University, Xiangtan, China, and the Ph.D. degree in computer science and technology from the School of Computer Science, Huazhong University of Science and Technology, Wuhan, China,

He visited the City University of Hong Kong, Hong Kong, and the University of Macau, Macau, China, as a Research Fellow nine times. He finished 11 projects for Huawei Co., Shenzhen, China; Longtop Co., Xiamen, China; and Yilianzong Co., Beijing, China. He has published over 40 SCI papers with an H-index of 27. His current research interests include computational intelligence, data mining, big data, and cloud computing.