Modeling User Activity Preference by Leveraging User Spatial Temporal Characteristics in LBSNs

Dingqi Yang, Daqing Zhang, *Member, IEEE,* Vincent W. Zheng, and Zhiyong Yu *Member, IEEE,*

***Abstract*—With the recent surge of Location Based Social Networks (LBSNs), activity data of millions of users has become attainable. This data contains not only spatial and temporal stamps of user activity, but also its semantic information. LBSNs can help to understand mobile users' spatial temporal activity preference, which can enable a wide range of ubiquitous applications, such as personalized context-aware location recommendation and group-oriented advertisement. However, modeling such user-specific spatial temporal activity preference needs to tackle high-dimensional data, i.e., user-location-time-activity quadruples, which is complicated and usually suffers from a data sparsity problem. In order to address this problem, we propose a spatial temporal activity preference (STAP) model. It first models the spatial and temporal activity preference separately, and then uses a principle way to combine them for preference inference. In order to characterize the impact of spatial features on user activity preference, we propose the notion of Personal Functional Region (PFR) and related parameters to model and infer user spatial activity preference. In order to model the user temporal activity preference with sparse user activity data in LBSNs, we propose to exploit the temporal activity similarity among different users and apply non-negative tensor factorization to collaboratively infer temporal activity preference. Finally, we put forward a context-aware fusion framework to combine the spatial and temporal activity preference models for preference inference. We evaluate our proposed approach on three real-world datasets collected from New York and Tokyo, and show that STAP consistently outperforms the baseline approaches in various settings.**

Keywords*—Spatial, Temporal, User Activity Preference, Tensor Factorization, Location Based Social Networks

I. Introduction

WITH the ubiquity of GPS-equipped smartphones, Location Based Social Networks have gained increasing popularity in recent years. In LBSNs, users interact not only with their friends by sending messages, sharing photos, but also with physical Points Of Interest (POIs) showing their presence in real-time, leaving their comments, etc.. These large-scale user generated digital footprints bring an unprecedented opportunity to understand the spatial and temporal features of user activity. In LBSNs, user activity is mainly represented by "check-in" which indicates that a user visited a POI at a certain time. Along with POI categories that are often associated with user activities, we can semantically characterize the activities of a user in a place. For example, a user, Jane, is having French food (i.e., being at a French restaurant) at [40.7586, -73.9791] at 21:10 on Friday. By mining these activity records, we are able to understand user spatial temporal activity preference which can then enable various location based applications. The most straightforward application is POI recommendation. For example, knowing Jane is currently interested in going to a Chinese restaurant, a recommendation of Chinese food in a nearby POI would be persuasive. Moreover, knowing a group of users' activity preference, realtime group-oriented advertisement can be better enabled. For example, a clothing store is offering a group discount, if we know five users in the area are interested in the clothing store, an invitation to them would be welcome by both business owners and customers.

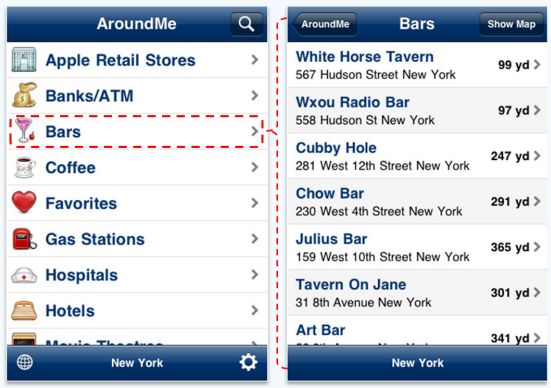


Fig. 1. Graphical user interface of AroundMe

In this paper, we try to answer the following question: "which activity is a mobile user interested in given her current context, including time and location?" Concretely, we aim at modeling user spatial temporal activity preference by leveraging user generated digital footprints in LBSNs. In LBSNs, user generated digital footprints usually contain rich semantic information on their activities. For example, a user's check-in at a French restaurant probably means that the user is having French food there. In current literature [1], [2], [3], [4], POI categories are often considered as the representation of user activities. In addition, as shown in Figure 1, a mobile application called "AroundMe" [1] implements a two-step user interface for users to explore the nearby places, which first lets them select activity category and then shows the specific POIs. Modeling user spatial temporal activity preference is able to improve the user experience of location based services. Taking

Manuscript received xxx; revised xxx.

D. Yang and D. Zhang are with the Department of Telecommunication Network and Services, Institut Mines-TELECOM/TELECOM SudParis, Evry 91011, France (e-mail: dingqi.yang@it-sudparis.eu; daqing.zhang@it-sudparis.eu). V. Zheng is with the Advanced Digital Sciences Center (ADSC), Singapore (e-mail: vincent.zheng@adsc.com.sg). Z. Yu is with the Fuzhou University, Fuzhou, Fujian, China (e-mail: yuzhiyong@fzu.edu.cn).

This work is partially supported by the EU FP7 Project SOCIETIES (No. 257493). (Corresponding Author: D. Zhang.)

[1] http://www.aroundmeapp.com/

the 'AroundMe" application as an example, given a user's current GPS coordinates and current time, if we infer that the user is interested in going to a bar, the bar category should appear at the very top of the category list in the application. However, modeling user spatial temporal activity preference from user check-ins in LBSNs is not trivial.

- First, since the check-in data is usually sparse and is represented as user-location-time-activity quadruples that contains four data dimensions, it is difficult and complicated to directly discover the regularity from such sparse high-dimensional data.
- Second, to consider spatial dimension, the existing works usually segment a city into disjoint grid cells and discretely infer user preference in individual cells such as in [2]. This method may cause inaccuracy due to the discretization process. For example, when a user is located at the border of two adjacent cells, a movement with a very short distance may incur the change of cells and cause different preference inference results. However, due to the continuity of location dimension, it is not easy to model user spatial activity preference in a continuous manner.
- Third, different from the continuously sampled user activity data, check-ins are user voluntarily reported activities. Most of users do not regularly perform check-ins, due to the reasons such as lack of time and privacy concern, etc. Therefore, check-ins in LBSNs usually suffer from a data sparsity problem, which causes difficulties in modeling user activity preference.

Aiming at resolving the above research issues, we develop a novel user Spatial Temporal Activity Preference (STAP) model. First, in order to reduce the problem complexity, we separately consider the spatial and temporal characteristics of user activity preference in LBSNs. Second, to capture the spatial features, instead of segmenting a city into grid cells, we build Personal Functional Regions for each user using her check-ins, which can then be used to infer ones' spatial activity preference. Third, to resolve the data sparsity problem in capturing temporal features, we exploit other similar users' activities and collaboratively build one's temporal activity preference model. Finally, a context-aware fusion framework is proposed to combine them together.

In the following sections, we first describe two unique spatial and temporal features of user activities that we use to build individual spatial and temporal models, and then present our contribution in modeling user spatial temporal activity preference.

A. Observations From A Study of User Activities

In order to build a hybrid spatial temporal user preference model, we would like to consider the spatial and temporal features of user activity separately. To this end, we collect and investigate a check-in dataset from a well-known LBSNs Foursquare[2], and obtain the following observations:

[2]https://foursquare.com/

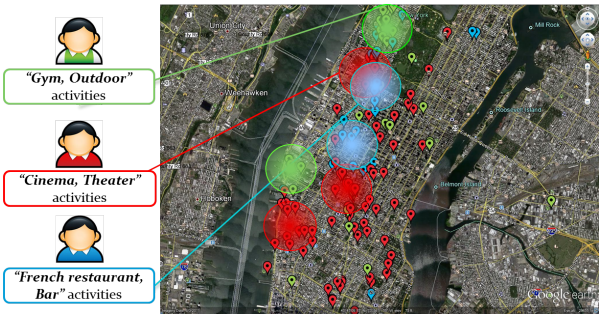


Fig. 2. Spatial distribution of three users' activities in Manhattan (Check-ins of three different users are plotted in red, green, blue colors, respectively)

Spatial specificity. Users' activities in LBSNs often show strong preference bias in their frequented regions. In other words, users only conduct a few types of activities (i.e., visit POIs of a few categories) in each of their frequented regions. Figure 2 shows check-ins of three New York users (represented by red, green, blue colors) in Manhattan in our dataset. First, we observe clearly that most of a user's check-ins happen in certain geographic areas, as plotted in the Figure. Such observation indicates that check-in behaviors have strong geographic preference and different users usually have their own frequented regions. Second, by investigating users' activities in their frequented regions, we discover that their activities are often limited to a few categories for majority of their frequented regions. We show the dominant activities in one frequented region of each user in Figure 2.

Temporal correlation. While users' activities in LBSNs can reflect their temporal activity preference, due to the sparsity of user check-ins, individual's digital footprints cannot well characterize a user's temporal activity preference. For example, if we consider weekly activity preference with hour granularity (i.e., 168 hours in a week), there are 103 hours on average for each user that we did not observe any activity in our dataset. However, we observe that some users' temporal activity preference may be very similar, which naturally fits the underlying assumption of collaborative filtering techniques, i.e., users who have similar temporal preference on some activities are likely to have similar temporal preference on others. For example, Figure 3 illustrates the activity category tag clouds of five similar users in different time slots in New York who share similar activity patterns. The selection of these users as a group is based on the community detection method proposed in [5]. We observe that they usually go to a coffee shop or a burger joint between 13:00 and 14:00 of weekday, stay at a bar between 21:00 and 22:00 on Friday, go to gym or outdoor places between 16:00 and 17:00 on weekend.

B. Our Contribution: STAP Model

Based on the previous observations, we introduce STAP model. Concretely, it first separately considers spatial and temporal features of user activity preference and then combines them together using a context-aware fusion framework.

1) ***Capturing Spatial Feature:*** Due to the continuity of location and sparsity of check-ins, it is impossible to observe

Fig. 3. Activity category tag clouds of five New York users who share similar activity patterns. (Larger font size implies a higher frequency, and vice versa.)

users' activities at all locations. However, spatial specificity suggests that users usually have activity preference bias in their frequented regions. Therefore, we may first try to estimate users' activity preference in their frequented regions, and then infer users' activity preference on their unvisited locations using interpolation methods.

The research challenge here is to discover all those regions for each user and quantitatively measure such preference bias, and then continuously infer user spatial activity preference. Intuitively, for a specific user, a good region should be an area where the user frequently visits and has strong preference bias (i.e., among various categories of activity available there, the user often conducts a few categories of activities). Because equally conducting all categories of activities in that area means the user has no obvious activity preference there. Therefore, we propose Personal Functional Region (PFR). Concretely, we first define a frequented region of a user as an area with a center and a radius where the user performs more than a certain percentage of all her check-ins. We then measure her activity preference bias in such a region using an entropy-based measure called ratio of preference bias that describes how deterministic the user's activities are in that area. Finally, we define Personal Functional Region as a user frequented region where the user has strong preference bias on certain activities (i.e., the ratio of preference bias is higher than a threshold). Knowing the PFRs of a user, we can then continuously infer her spatial activity preference on her unvisited locations using interpolation methods.

To find out one's PFRs, due to the continuity of location, it is impossible to exhaustively enumerate all regions and identify PFRs. Therefore, we propose a greedy clustering based approach to discover PFRs using user's historical check-ins. The basic idea is to start from one's most frequently visited POI and its neighboring area, because such a region are most likely to be a frequented region, which is a necessary condition for a PFR. Specifically, it first scans from the most checked POI and considers the POI's GPS coordinates as a region center. By evaluating the nearby (i.e., within a certain radius) user activity frequency and ratio of activity preference bias, it then determines whether such a region is a PFR using a threshold based criterion.

*2) **Capturing Temporal Feature**:* Temporal correlation shows that some users may have similar temporal activity patterns. It suggests that user temporal activity preference can be collaboratively inferred. Collaborative filtering techniques are widely adopted when tackling sparse data, especially in recommendation systems to predict user preference using limited and sparse user historical data. To collaboratively build user temporal activity preference model using sparse check-ins, we need to first find the latent correlation, i.e., users with the similar temporal activity preference as shown in Figure 3, and then infer one's temporal activity preference with the help of the preference of similar users.

The research challenge here is how to discover and leverage the latent correlation. Instead of manually identifying and explicitly describing such correlations, low-rank approximation techniques, such as matrix/tensor factorization, are usually adopted to discover the latent correlation on such multi-facet data. In this work, we use a three-way tensor to model user temporal activities (i.e., user-time-activity), where we consider the weekly user activity pattern with hour granularity because users often exhibit different daily patterns in a week and it is meaningless to measure user activity duration in seconds or minutes in our case. Using non-negative tensor factorization techniques, we are able to discover the latent correlation between user, time and activity factors. By recovering a tensor from these factors, we obtain the approximated non-negative preference measure for each user-time-activity triplet.

Moreover, compared to the continuously sampled user activity data, check-ins are user voluntarily reported activities. Such a property implies that the sequential patterns of check-ins are not reliable as shown in [6]. Hence, we ignore the sequential patterns of user activities in STAP model and we later consider sequential pattern mining approaches as baselines in evaluation.

*3) **Fusion of Spatial and Temporal Feature**:* Due to the complexity of handling the user-location-time-activity data directly, we separately consider spatial and temporal features of user activities and then propose a context-aware fusion framework integrated in STAP to infer user activity preference. The existing works mainly leverage weighted average methods. However, since the ability of spatial and temporal models varies depending on users' contexts (i.e., time and locations), it is difficult to identify the optimal weights for fusion across different contexts. Therefore, we propose to simply use 1/0 weight to combine spatial and temporal models together dynamically according to users' current contexts. Concretely, we first calculate the activity preference inference success rate of both spatial and temporal models on a validation dataset for different contexts. When inferring user activity preference, we then choose the better model for the given context by comparing the success rate of the two models.

We experimentally evaluate STAP using three check-in datasets collected from two LBSN services, i.e., Foursquare and Gowalla[3]. The experiment results show that the STAP model achieves consistently good performance with all three datasets and outperforms various baseline approaches, which verifies the generality and advantages of our solution in modeling spatial-temporal activity preference with sparse check-in data.

The rest of the paper is organized as follows. We first present the related work in Section II, we then formally define the spatial temporal activity preference inference problem in Sec-

[3]http://blog.gowalla.com/

tion III. The proposed spatial and temporal activity preference models are presented in Section IV and V, followed by the context-aware fusion framework in Section VI. Experimental evaluation is shown in Section VII. We conclude our work in Section VIII.

II. Related Work

Due to the increasing popularity of LBSNs, users generate tremendous amount of digital footprints [7] in their daily life. Research on understanding user activity by mining these digital footprints has attracted extensive attention in recent years. Since the objective of this paper is to infer user spatial temporal activity preference in LBSNs, we first briefly survey these research works on user activities from two perspectives: 1) user mobility perspective which focuses on modeling user mobility patterns by leveraging spatial temporal regularities; 2) user preference perspective which usually focuses on inferring user preference on the unvisited POIs. We then present the research works considering POI categories as user activities, as well as the related works for our spatial and temporal models.

From user mobility perspective, various studies have been conducted for location prediction. In LBSNs, location prediction in terms of POIs aims at predicting the specific POI that users will visit next. For example, Chang et al. [8] incorporated various features in LDA model for next POI prediction. Gao et al. [6] proposed a social-historical model based on Hierarchical Pitman-Yor process for predicting the next check-in of a user. Noulas et al. [9] extracted user specific features and global mobility features and built a prediction model for next POI. Different from these works that try to predict users' future locations, our objective is to infer the users' current activity preference based on their current context, i.e., time and location.

From user preference perspective, most research efforts focused on location recommendation. Location recommendation tries to suggest POIs to users by estimating user preference on individual POI. For example, Bao et al. [10] developed a collaborative filtering based location recommender system by estimating user preference on POIs in an unfamiliar city. Noulas et al. [11] tackled new venue recommendation problem and proposed a new model based on personalized random walk approaches. Chen et al. [12] proposed a multi-center Gaussian model to capture the geographical influence and combined matrix factorization methods to perform the location recommendation. Yang et al. [13] estimated user preference on location using the hybrid preference extracted from user check-ins and text-based tips using statistic and sentiment analysis techniques. Biancalana et al. [14] described a context-aware social recommender system by identifying user preference and information needs and then making recommendations related to POIs. Different from location recommendation problems that focus on estimating user preference on specific POIs, the problem formulated in this paper is to model user spatial temporal preference on activities.

Besides the above two categories of research, papers in [1], [2], [3], [4] consider POI categories as the semantic interpretation of user activities (e.g., food, shopping, entertainment). For instance, Lian et al. [1] clustered users based on their temporal activities and activity transition in order to collaboratively predict user activities. Pianese et al. [3] clustered user activities for user routine detection and predicted user future activities as well as locations. Ye et al. [2] used the mixed hidden Markov model to predict user's next activity. Huang et al. [15] proposed a method to automatically identify user activity using the spatial temporal attractiveness of POIs. These works focus on predicting users' next activities which is different from our objective that is to infer the users' current activity preference according to their current contexts.

To capture spatial specificity of user activity preference, we define Personal Functional Region (PFR) in a city. The concept of urban functional region [16] has been studied for years. The availability of user digital footprints brings a novel opportunity to discover city functional regions. For example, Yuan et al. [17] proposed a framework to discover and semantically annotate urban functional regions using human mobility and POIs in a city. Kurashima et al. [18] proposed a method called Geo Topic Model to discover different activity areas in a city and user's interest for the purpose of location recommendation. Combining LBSN data and cellular phone call data, Noulas et al. [4] investigated the crowd activity in urban environments and formulated urban activity inference problem in a supervised learning framework. Although these works managed to find out the common functional regions in a city, the empirical study shows that different users often have different activity preference in the same area. Based on this observation, we propose PFR which is able to capture individual's spatial activity preference.

To capture temporal correlation of user activity preference, we adopt tensor factorization techniques which are widely used in context-aware recommendation. Compared to classical matrix factorization approaches, tensor factorization considers context as an additional dimension. For example, Karatzoglou et al. [19] proposed a multiverse recommendation approach by incorporating context into factorization model using tensor. Yang et al. [20] developed a multi-tuple tensor factorization algorithm to improve the location search quality. Balázs et al. [21] developed a novel tensor factorization technique for context-aware recommendation with better runtime performance. Symeonidis et al. [22] proposed a Geo-social recommender system by leveraging the Higher Order Singular Value Decomposition (HOSVD) technique in tensor factorization. Zheng et al. [23], [24] leveraged a ranking-based collective tensor factorization techniques to recommend both POIs and activities. We advocate for using tensor factorization techniques to collaboratively model user temporal activity preference. Since non-negative tensor decomposition can help to convey the interpretable results as probabilities, we adopt it in this work to solve our problem.

III. Problem Definition

The objective of this work is to model and infer user spatial temporal activity preference. In LBSNs, users visit diverse categories of POIs. Since POI categories usually imply the activities that users conduct there, we consider POI categories

TABLE I. NOTATION

Symbol	Description
U	set of users
u	a user, $u \in U$
V	set of venues
$\mathcal{A}_u$	check-in activities of user u
$\mathcal{A}_{u,r}$	check-in activities of user u in region r
C	set of POI categories
C_l^d	existing activity categories within center l and radius d
$C_{u,r}$	user u conducted activity categories in region r_u
c	an activity category, $c \in C$
l	GPS coordinates
T	set of time slots
t	a time slot, $t \in T$
r	a region with center l and radius d
r_u	a personal functional region of user u
$\mathcal{R}_u$	personal functional regions of user u
$\psi_{u,r}$	user u's activity distribution in region r
$\Psi_{u,l}$	spatial activity preference of user u at location l
$\Psi_{u,t}$	temporal activity preference of user u at time t
$\Psi_{l,t}$	spatial temporal activity preference of user u at time t and location l

as user activities. Therefore, we are interested in the following problem: given a set of users' historical behaviors, i.e., check-ins, the objective is to infer their interest in activities (visiting certain categories of venues[4]) for a given time, around the current geo-location.

Formally, given a set of users U and a set of venues V related to a set of categories C, each venue belongs to a category c, where $c \in C$. Let C_l^d denote the existing location categories within d km from the Geo-location l (represented by GPS coordinates). Each check-in can then be represented by a quadruple (u, l, c, t), representing the user u conduct activity c, at time t when user's position is l. Let $\mathcal{A}_u$ denote the check-in activities of the user u. The problem of modeling user spatial temporal activity preference can then be formulated as: Knowing the historical activities of users U, i.e.,$\{\mathcal{A}_u | u \in U\}$, given a user u whose current position is l at current time t, our aim is to infer u's preference in visiting the nearby venue categories C_l^d. The notations used in this paper are summarized in Table I.

IV. User Spatial Activity Preference Modeling

In order to model user spatial activity preference in a continuous manner, we propose Personal Functional Region by considering the spatial specificity feature of user activity. In this section, we first give the definition of Personal Functional Region and then propose a PFR discovery algorithm by mining users' historical activities. Finally, we show how to infer user activity preference using PFRs.

A. Personal Functional Regions

The definition of Personal Functional Region is based on the spatial specificity feature of user activity preference, which shows that users usually perform certain specific activities in their frequented regions. Therefore, in the following, we first define user frequented regions and then define the ratio of preference bias in a frequented region to quantitatively characterize user activity preference bias. Afterwards, we give the definition of Personal Functional Regions.

Definition 1 (Frequented Region): A region is a geographical area with a center l and a radius d. Region r is a Frequented Region of user u if and only if user u has performed more than s_{freq} of her total check-ins, i.e., the fraction of u's check-ins activities in r is greater than or equal to the threshold s_{freq}.

$$freq = \frac{\mathcal{A}_{u,r}}{\mathcal{A}_u} \geq s_{freq} \tag{1}$$

In this definition, l and d determine the location and the size of the region. The threshold s_{freq} determines the lower bound of the frequency u visits r. Note that we use circular region for frequented region representation in this paper due to its simplicity. However, in urban planning community, the most popular methods of describing functional regions in a city are based on its road segmentation and are usually non-overlapped as in [17]. However, based on the geographical distribution of individual's activities, PRFs may have more complex geographical representation, such as polygonal areas. We will investigate different geographical representations of PFRs in our future work.

Functional regions in a city are usually characterized by the distribution of venue categories. For example, a region with lots of stores and restaurants is likely a commercial center; a region with many monuments and historical sites is probably a tourism spot. To describe individual's functional regions, we need to quantitatively measure users' activity diversity in their frequented regions. Since users' activities in their frequented regions usually fall into a few categories rather than all the existing categories C_l^d, we are inspired by the definition of relative redundancy in information theory and propose ratio of preference bias to characterize one's activity preference. Specifically, we measure how deterministic one's activity is in a frequented region r by calculating the difference between the entropy of the users' actual activity distribution and the maximum entropy of the activity distribution in r. The maximum entropy of the activity distribution $H_{max}(|C_l^d|)$ is calculated as follows:

$$H_{max}(|C_l^d|) = \log_2 |C_l^d| \tag{2}$$

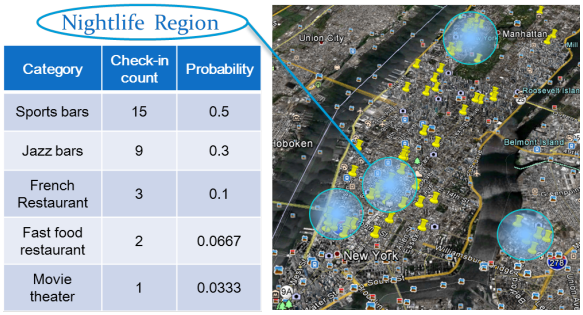
It corresponds to the situation that a user u visits all the location categories C_l^d in the region r equally, which means that u does not have obvious preference bias on activities in r. With this reference, we define the ratio of preference bias as follows:

Definition 2 (Ratio of Preference Bias): For a user u and one of her frequented regions r, the ratio of preference bias measures the fractional difference between the entropy of u's activity distribution $\psi_{u,r}$ in r and the maximum entropy of the activity distribution in r, which is calculated as follows:

$$ratio_{PB} = 1 - \frac{H(\psi_{u,r})}{H_{max}(|C_l^d|)} \tag{3}$$

Higher value of $ratio_{PB}$ implies the stronger activity preference bias of u in r, and vice versa. Since the PFR of a user should be an area where the user has strong activity preference bias, we thus define Personal Functional Region as :

[4]Since "venue" is used to represent a POI in Foursquare, we do not differentiate these terms throughout this paper.



Category	Check-in count	Probability
Sports bars	15	0.5
Jazz bars	9	0.3
French Restaurant	3	0.1
Fast food restaurant	2	0.0667
Movie theater	1	0.0333

Fig. 4. Example of a "nightlife" PFR of a user

Definition 3 (Personal Functional Region): A Personal Functional Region (PFR) r_u for user u is a user frequented region where u's activities have higher ratio of preference bias, i.e., $ratio_{PB}$ is greater than or equal to a threshold $s_{ratio_{PB}}$.

A PFR r_u is then represented by the center l, radius d, u's activity distribution $\psi_{u,r}$ and ratio of preference bias $ratio_{PB}$. As the threshold $s_{ratio_{PB}}$ denotes the lower bound of user activity preference bias in PFRs, we need to identify the boundary of $ratio_{PB}$ in order to properly choose the threshold $s_{ratio_{PB}}$. We show in Proposition 1 that $ratio_{PB}$ is actually bounded to $[0, 1]$.

Proposition 1: Given any u and any of her frequented region r, the ratio of preference bias $ratio_{PB} \in [0, 1]$.

The proof of Proposition 1 can be found in Appendix A. We now use an example to show how to discover a user's PFR by computing the $ratio_{PB}$ as follows:

Example 1: A "nightlife region" for a user is shown in figure 4. This user visits mainly "Sports bars" and "Jazz bars". There are totally 20 categories ($|C_l^d| = 20$) in this region. The ratio of preference bias is then calculated as follows:

$$H(\psi_{u,r}) = -\sum_{c_i \in \psi_{r_u}} p(c_i) \log_2 P(c_i) = 1.78 \tag{4}$$

$$H_{max}(|C_l^d|) = \log_2 |C_l^d| = 4.32 \tag{5}$$

$$ratio_{PB} = 1 - \frac{H(\psi_{u,r})}{H_{max}(|C_l^d|)} = 0.59 \tag{6}$$

B. PFR Discovery Algorithm

According to above definitions, we need to determine four parameters to specify a PFR based on users' historical activities. Among these four parameters, l refers to the center location of the region; d decides the size of PFR; s_{freq} shows how active a user is in her PFRs; $s_{ratio_{PB}}$ represents the user's activity preference bias degree in her PFRs.

The basic idea of efficiently discovering PFRs is to start from one's most frequently visited POI and its surrounding area, because such a region is most likely to be a frequented region, which is a necessary condition for being a PFR. Therefore, we propose a greedy clustering algorithm to discover PFRs from one's check-in activities as shown in Algorithm 1. Concretely, given a user's historical check-ins, we first scan

Algorithm 1 PFR Discovery Algorithm

Require: User u's check-ins $\mathcal{A}_u$ and parameters $d, s_{freq}, s_{ratio_{PB}}$

1: Sort $\mathcal{A}_u$ in descend order according to visiting frequency
2: Initialize remainder check-in set, $\mathcal{A}_{rest} = \mathcal{A}_u$
3: Initialize user u's PRF set, $\mathcal{R}_u = \emptyset$
4: **for** $v \in \mathcal{A}_u$ **do**
5: **if** $v \in \mathcal{A}_{rest}$ **then**
6: Select r with center $v.l$ and radius d
7: Find check-ins $\mathcal{A}_{rest}$ in r denoted as $\mathcal{A}_{u,r}$
8: Calculate $freq$ in the region r
9: **if** $freq \geq s_{freq}$ **then**
10: Calculate $ratio_{PB}$ in r based on $\psi_{u,r}$
11: **if** $ratio_{PB} \geq s_{ratio_{PB}}$ **then**
12: Add r in $\mathcal{R}_u$ with $l, d, \psi_{u,r}, ratio_{PB}$
13: Remove $\mathcal{A}_{u,r}$ from $\mathcal{A}_{rest}$
14: **end if**
15: **end if**
16: **end if**
17: **end for**
18: **return** $\mathcal{R}_u$

from the most checked venue and consider all the visited venues whose distance is less than d kilometers from the selected venue as a region r (Line 1-6). When calculating the visiting frequency, an activity can only be counted once. We use $\mathcal{A}_{rest}$ to denote the un-counted check-ins. If this region is a user's frequented region (i.e., visiting frequency $freq$ is equal or higher than the threshold s_{freq}), we calculate ratio of preference bias $ratio_{PB}$ (Line 7-10). If $ratio_{PB}$ is equal or higher than the threshold $s_{ratio_{PB}}$, we choose r as a PFR for the user (Line 11-12) and remove the counted check-ins $\mathcal{A}_{u,r}$ from $\mathcal{A}_{rest}$ (Line 13). After examining each check-in venue in all the regions, we get a set of PFRs $\mathcal{R}_u$ for the user u.

C. Spatial Preference Inference Using PFRs

After discovering users' PFRs, the next issue turns to inferring user activity preference using PFRs. Knowing user u's current location l, we first estimate the preference influence of individual PFRs and then combine activity preference distribution $\psi_{u,r}$ of all PRFs using weighted average methods. Some previous works [12], [25], [26] have studied the probability of location visiting w.r.t. the travel distance, and found that it is inversely proportional. We advocate for their finding and based on that, we propose the following weight function:

$$w_{l,r_u} = \begin{cases} d^{-1}, & if \quad d_{l,r_u} \leq d \\ d_{l,r_u}^{-1}, & if \quad d_{l,r_u} > d \end{cases} \tag{7}$$

Specifically, for the PFRs whose distance d_{l,r_u} from l is less than or equal to the radius d of r_u, the user is currently in these PFRs and we consider their influence equally. For other PFRs whose distance d_{l,r_u} from l is greater than the radius d of r_u, their influence is proportional to d_{l,r_u}^{-1}. Therefore, the spatial activity preference $\Psi_{u,l}$ of user u at location l can be

calculated as follows:

$$\Psi_{u,l} = \sum_{r_u \in \mathcal{R}_u} \psi_{u,r_u} \cdot w_{l,r_u} \tag{8}$$

V. User Temporal Activity Preference Modeling

Due to the sparsity of check-in data and the temporal correlation of user activity preference, we exploit other similar user's activities and collaboratively build a user's temporal activity preference. Concretely, we first model user temporal activities using a three-way tensor and then leverage tensor factorization techniques to decompose the tensor into three factors, i.e., user, time and activity factors. By recovering a tensor using these factors, we obtain the preference measure for each user-time-activity triplet. In order to avoid the negative value in the recovered tensor which is meaningless for preference measure, we add non-negative constraint into the factorization process. The non-negative constraint can help to make the results interpretable [27] as probability. In the following, we first present tensor factorization model, and then explain how to infer user temporal activity preference using non-negative tensor factorization techniques.

A. Tensor Factorization Model

In this work, we build a user-time-activity tensor based on users' historical check-ins. Since we consider venue categories as user activity categories, the tensor is denoted as *u-t-c*, (i.e., user-time-category). Tensor factorization techniques intend to decompose such a tensor into multiple factors. Let $\hat{U}$, $\hat{T}$ and $\hat{C}$ denote the user, time and activity category feature matrices, with size of $|U| * f$, $|T| * f$ and $|C| * f$, respectively. In a sense, these matrices comprise computerized groups of user, time and activity dimension according to users' activities modeled by tensor. For example, a feature dimension for user matrix measures how much a user likes a certain group of temporal activities on the corresponding time and activity feature dimension. Note that f is called latent space dimension (or factorization dimension) which is the most important parameter in tensor factorization. It controls the number of features involved in the factorization process. The decomposition is formulated as follows:

$$\hat{Y} = \hat{O} \times_U \hat{U} \times_T \hat{T} \times_C \hat{C} \tag{9}$$

where $\times_n$ is the mode-n tensor product with matrix. The core tensor $\hat{O}$ with dimension $f*f*f$ handles the correlation among different factors. The value of each element in $\hat{Y}$ is calculated as:

$$\hat{y}_{u,t,c} = \sum_{\hat{u}} \sum_{\hat{t}} \sum_{\hat{c}} \hat{o}_{\tilde{u},\tilde{t},\tilde{c}} \cdot \hat{u}_{u,\tilde{u}} \cdot \hat{t}_{t,\tilde{t}} \cdot \hat{c}_{c,\tilde{c}} \tag{10}$$

where $\tilde{u}, \tilde{t}, \tilde{c} \in \{1, ..., f\}$ are indices of latent features. This model is called Tucker decomposition model [28]. For simplicity, we assume that the core tensor $\hat{O}$ is a diagonal tensor:

$$\hat{c}_{\tilde{u},\tilde{t},\tilde{c}} = \begin{cases} 1, & if \quad \tilde{u} = \tilde{t} = \tilde{c} \\ 0, & else \end{cases} \tag{11}$$

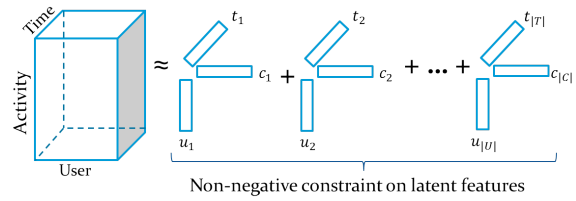


Fig. 5. Non-negative tensor factorization using Canonical decomposition model

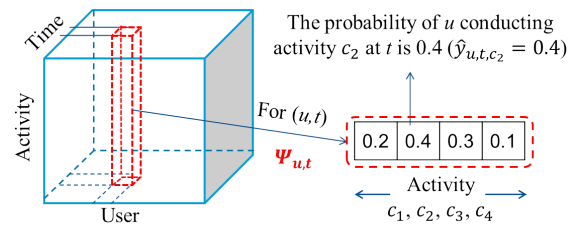


Fig. 6. Temporal activity preference inference

We then obtain Canonical decomposition model with each element calculated as:

$$\hat{y}_{u,t,c} = \sum_{\tilde{f}} \hat{u}_{u,\tilde{f}} \cdot \hat{t}_{t,\tilde{f}} \cdot \hat{c}_{c,\tilde{f}} \tag{12}$$

where $\tilde{f} \in \{1, ..., f\}$ is the index of latent space. In this work, we adopt non-negative tensor factorization using Canonical decomposition model which can be efficiently calculated within relatively short running time. Figure 5 illustrates the factorization model. We decompose a tensor into three factors (i.e., $\hat{U}$, $\hat{T}$ and $\hat{C}$) and try to optimize the loss function between the recovered tensor $\hat{Y}$ and the original *u-t-c* tensor.

B. Temporal Preference Inference

We adopt the non-negative tensor factorization implementation[5] in [29]. It adds non-negative constraint to Alternative Least Square based tensor factorization algorithms and uses Canonical decomposition model.

By recovering $\hat{Y}$ from $\hat{U}$, $\hat{T}$ and $\hat{C}$ using Equation 12, we obtain a non-negative tensor describing users' temporal activity preference. In order to infer user u's preference (probability of conducting an activity) at time t, we normalize $\hat{Y}$ as follows:

$$\sum_{c=1}^{|C|} \hat{y}_{u,t,c} = 1, \ \forall u \in U \ and \ \ \forall t \in T \tag{13}$$

For the given u and t, the sum of all activities' preference measure (i.e., probability) is normalized to one. This is for the later fusion with spatial activity preference which is represented by probability in the value range of $[0, 1]$. Figure 6 shows an example with four activity categories, where the normalized preference measure can be regarded as the probability that u conducts activity c at time t. Thus, we obtain temporal activity preference $\Psi_{u,t}$ as follows:

$$\Psi_{u,t} = \{\hat{y}_{u,t,c} | c \in C\} \tag{14}$$

[5]https://sites.google.com/site/jingukim/home#ntfcode

VI. Context-aware Fusion Framework

Given one's spatial and temporal activity preference, i.e., $\Psi_{u,l}$ and $\Psi_{u,t}$, the fusion framework tries to combine them together to obtain the user spatial temporal activity preference. The most straightforward approach is to merge them using two static weights. Since the performance of spatial and temporal models varies over time and locations, the simple weighted average cannot always get the best of the two models (later proved in evaluation). However, it is difficult to dynamically assign the two weights according to the user context. Therefore, by conducting the study on a validation dataset, we simply select the model with higher accuracy for activity preference inference according to a user's current context (i.e., location l and time t). In this paper, we propose a context-aware fusion framework to take advantage of both spatial and social models. Specifically, we first define the success rate of a preference model as the frequency of correct inference for the Top 1 activity. Then, for each user, we use two matrices to calculate the success rate of both spatial and temporal models on different contexts using a validation dataset. When inferring user activity preference, the model with higher success rate is used.

A. Success Rate Calculation of Preference Model

The objective of calculating success rate is to get the inference accuracy of both preference models under different contexts, i.e., time and involved PFRs. Let M_{tem} and M_{spa} denote the matrices of spatial and temporal success rate, respectively. Each row of the matrices corresponds to a time slot t and each column represents one functional region r_u of user u. Algorithm 2 shows the process of building M_{tem} and M_{spa}. We first initialize M_{tem} and M_{spa} by assigning each element to 0 (Line 1). For each check-in activity in the validation dataset, we infer u's spatial and temporal activity preference $\Psi_{u,l}$ and $\Psi_{u,t}$ and then get the most probable activity c_l and c_t (Line 2-4). We also get u's local PFRs $R_{u,local} = \{r_u \in R_u | d_{l,r_u} \leq d\}$ where the user is currently in (Line 5). If the user's $R_{u,local}$ is not empty, we get the user's current context, i.e., time $v.t$ and local PFRs $R_{u,nearby}$ (Line 6). Afterwards, if the spatial model infers the correct activity, we augment the success rate in M_{spa} for current context, i.e., time slot and local PFRs, by 1 (Line 7-9). Specifically, $M_{spa}(v.t, R_{u,nearby})$ represents the numbers in $v.t$ row and r_u ($r_u \in R_{u,nearby}$) column(s) of M_{spa}. If the temporal model infers the correct activity, we do the same for M_{tem} (Line 10-12).

B. Fusion Criterion

Knowing the success rate of each model under different contexts, we choose the model with the higher success rate. Algorithm 3 shows this process. Specifically, for a given user u and her context, i.e., time t and location l, we obtain matrices M_{tem} and M_{spa} of u generated by Algorithm 2. Then, we find u's local PFRs $R_{u,local}$ based on her current location l (Line 1). If $R_{u,local}$ is not empty, we calculate the overall success rate for both spatial and temporal models as follows (Line 2-4):

$$rate_{spa} = \sum_{r_u \in R_{u,local}} M_{spa}(t, r_u) \tag{15}$$

$$rate_{tem} = \sum_{r_u \in R_{u,local}} M_{tem}(t, r_u) \tag{16}$$

We then use the one with higher success rate as final preference distribution (Line 5-10). In case of equality, we randomly choose one from $\Psi_{u,t}$ and $\Psi_{u,l}$ (Line 11-13). If u's local PFR set $R_{u,local}$ is empty, we consider the preference distribution of temporal model as the final result (Line 15), because the spatial model is considered to be unconfident in this case.

Algorithm 2 Context-aware success rate calculation

Require: User u's spatial and temporal preference distribution $\Psi_{u,l}$ and $\Psi_{u,t}$, PRFs $\mathcal{R}_u$, activities in validation dataset $\mathcal{A}_{u,valid}$

1: Initialize M_{tem} and M_{spa} with 0
2: **for** $v \in \mathcal{A}_{u,valid}$ **do**
3: Get the Top 1 activity c_l based on $\Psi_{u,l}$
4: Get the Top 1 activity c_t based on $\Psi_{u,t}$
5: Find local PFRs $R_{u,local} = \{r_u \in R_u | d_{l,r_u} \leq d\}$
6: **if** $R_{u,local}$ is not empty **then**
7: **if** c_l equals user actual activity $v.c$ **then**
8: Augment $M_{spa}(v.t, R_{u,local})$ by 1
9: **end if**
10: **if** c_t equals user actual activity $v.c$ **then**
11: Augment $M_{tem}(v.t, R_{u,local})$ by 1
12: **end if**
13: **end if**
14: **end for**
15: **return** M_{tem} and M_{spa}

Algorithm 3 Context-aware preference fusion

Require: User u's spatial and temporal preference distribution $\Psi_{u,l}$ and $\Psi_{u,t}$, context t and l, PRFs $\mathcal{R}_u$, success rate matrices M_{tem} and M_{spa}

1: Find local PFRs $R_{u,local} = \{r_u \in R_u | d_{l,r_u} \leq d\}$
2: **if** $R_{u,local}$ is not empty **then**
3: Calculate $rate_{spa}$ according to Equation 15
4: Calculate $rate_{tem}$ according to Equation 16
5: **if** $rate_{spa} > rate_{tem}$ **then**
6: $\Psi_{u,l,t} = \Psi_{u,l}$
7: **end if**
8: **if** $rate_{spa} < rate_{tem}$ **then**
9: $\Psi_{u,l,t} = \Psi_{u,t}$
10: **end if**
11: **if** $rate_{spa} = rate_{tem}$ **then**
12: $\Psi_{u,l,t}$ = randomly choose one from $\{\Psi_{u,t}, \Psi_{u,l}\}$
13: **end if**
14: **else**
15: $\Psi_{u,l,t} = \Psi_{u,t}$
16: **end if**
17: **return** $\Psi_{u,l,t}$

TABLE II. DATASET STATISTIC

Dataset	New York (Foursquare)	Tokyo (Foursquare)	New York (Gowalla)
Users	824	1,939	244
Venues	38,336	61,858	9,352
Check-ins	227,428	573,703	85,010
Average number of activity categories per user	38.37	31.39	55.58

VII. EXPERIMENTAL EVALUATION

We evaluate STAP by conducting activity preference inference experiments using three datasets collected from two LBSN services, i.e., Foursquare and Gowalla. In the following, we first present the experiment setting including data collection, evaluation plan and metrics. We then show the impact of parameters on STAP model in order to identify their optimal values. Finally, we present the comparison with baseline approaches in terms of preference inference performance.

A. Experimental Setting

*1) **Data Collection**:* In this work, we use three datasets collected from two LBSN services, i.e., Foursquare and Gowalla, to evaluate our model.

Foursquare Dataset. We use a collection of Foursquare check-ins lasting for about 10 months (from 12 April 2012 to 16 February 2013). We filter out noise and invalid check-ins, and then select active users in two big cities i.e., New York and Tokyo, as experiment dataset. Details about the dataset collection process can be found in our previous work [13]. Venues in Foursquare are classified into 9 root categories and 417 sub-categories[6] at the time of data collection. Based on these sub-categories, we manually merge the similar and infrequent venue categories together, resulting in a total of 251 venue sub-categories. In order to validate that our approach is not dataset dependent, we select check-ins of active users (defined as users who have performed at least three check-ins per week) in two big cities, i.e., New York and Tokyo, as experiment datasets.

Gowalla Dataset. In order to validate that our approach does not depend on the LBSN services, we also conduct experiments using a dataset from another LBSN service Gowalla (from January 2010 to October 2010), which is extracted from the dataset used in [30]. The data filtering and processing step is similar to that of the Foursquare dataset. We select the check-ins in New York as experiment dataset.

The statistics of the selected datasets are shown in Table II. The tag clouds of user activities on the datasets are illustrated in Figure 7. Note that there is no obvious difference between the tag clouds of New York (Foursquare) and New York (Gowalla). We thus only show the results from Foursquare. We observe clearly the cultural differences between the two cities: New York users usually share their activities in bars, gyms, restaurants; while Tokyo users often share their presence at train stations, convenience stores, and Japanese restaurants.

[6]https://developer.foursquare.com/docs/venues/categories

(a) New York (Foursquare) (b) Tokyo (Foursquare)

Fig. 7. Tag cloud of activity category (Larger font size implies higher frequency, and vice versa.) We only show the tag cloud of New York (Foursquare) dataset because there is no visual difference between it and the tag cloud of New York (Gowalla) dataset.

*2) **Evaluation Plan**:* In the following experiments, we use the first eight-month check-ins as training dataset to build individual spatial and temporal models. We then use the 9th month check-ins as validation dataset to calculate the success rate of individual models for the context-aware fusion framework. Finally, we use the 10th month check-ins as test dataset for experiments.

*3) **Evaluation Metric**:* Since our application scenario focuses on recommending activities to users, our primary evaluation objective is to see whether a user's interested activity appears at the top of the returned list. Specifically, for each new check-in in the test dataset, we infer the user's activity preference under the given context and compare it with the user's actual activity. Therefore, we use the Top K accuracy ($Accuracy@K$) as the first evaluation metric, which calculates the percentage of the actual activities appearing at the Top K inferred activities in the test dataset. For the test dataset S, the Top K accuracy is calculated as follows:

$$Accuracy@K = \frac{|\{(u,l,c,t)|c \in P_{u,l,t}(K), (u,l,c,t) \in S\}|}{|S|} \tag{17}$$

where $P_{u,l,t}(K)$ is the Top K activities inferred for user u at time t and location l (i.e., the Top K activities in $\Psi_{l,t}$ for user u). Moreover, in activity recommendation scenarios such as AroundMe application (Figure 1), a user may scroll the screen to find her interested activity. Therefore, the actual rank of the desired activity also has impact on user experience. In order to evaluate the overall ranking of the inferred activity preference, we use Average Percentile Rank [9] of the actual activity in the inferred activity list as another metric, which is calculated as follows:

$$AveragePercentileRank = \sum_{(u,l,c,t)\in S} \frac{|\Psi_{l,t}| - rank(c) + 1}{|\Psi_{l,t}|} \tag{18}$$

where $rank(c)$ is the rank of the actual activity c in the inferred activity preference list. The score of Average Percentile Rank is bounded in $(0, 1]$. The high score of Average Percentile Rank implies that the actual user activities appear on the top of the inferred activity preference list, and vice versa.

In addition, from the activity category perspective, some activities may show stronger spatial temporal regularities than others. For example, going to school or office may show stronger spatial temporal regularities than shopping activities. Therefore, in order to study the performance over different

activity categories, we consider our problem as a classification problem with 251 activity categories and calculate precision/recall and the related F1 score for each activity category. For a specific activity category c_i, these metrics are calculated as follows:

$$Precision(c_i) = \frac{|\{(u,l,c_i,t)|P_{u,l,t}(1) = c_i, (u,l,c_i,t) \in S\}|}{|\{(u,l,c,t)|P_{u,l,t}(1) = c_i, (u,l,c,t) \in S\}|} \tag{19}$$

$$Recall(c_i) = \frac{|\{(u,l,c_i,t)|P_{u,l,t}(1) = c_i, (u,l,c_i,t) \in S\}|}{|\{(u,l,c_i,t)|(u,l,c_i,t) \in S\}|} \tag{20}$$

$$F1Score(c_i) = \frac{2 \cdot Precision(c_i) \cdot Recall(c_i)}{Precision(c_i) + Recall(c_i)} \tag{21}$$

B. Impact of Parameters on STAP model

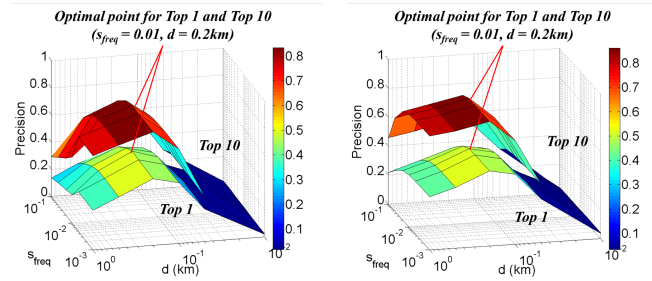
The STAP model separately considers spatial and temporal features of user activity preference. From spatial perspective, a user's Personal Functional Regions are determined by four parameters, i.e., l, d, s_{freq} and $s_{ratio_{PB}}$. From temporal perspective, the latent space dimension (or factorization dimension) determines the number of the features involved in the factorization process. Low dimension may result in unsatisfied performance while the high dimension usually implies high runtime complexity. Note that the impact of parameters on STAP model is similar with the New York (Foursquare) and New York (Gowalla) datasets due to the same geographical constraint. We only report the results with New York (Foursquare) dataset and Tokyo (Foursquare) dataset in this section.

1) ***Spatial Parameter Setting****:* In order to identify the optimal parameters of Personal Functional Regions for activity preference inferring, we conduct two experiments. First, we show the preference inference accuracy with different parameter combinations of PFRs, i.e., region size d and visiting frequency s_{freq}. Second, by fixing optimal values of the above two parameters, we tune the threshold of preference bias ratio, i.e., $s_{ratio_{PB}}$. Note that we do not study the impact of the PFR centers (i.e., l) in the above experiments because they are automatically determined by the PFR discovery algorithm.

In the first experiment, we set the threshold of preference bias ratio to its lower bound, i.e., $s_{ratio_{PB}} = 0$ according to Proposition 1, which implies that we consider all user frequented regions as PFRs regardless user activity preference bias there. We then plot Top 1 and Top 10 accuracy of activity preference inference by varying s_{freq} within [0.001, 0.005, 0.01, 0.02, 0.05, 0.1], and d within [0.01, 0.05, 0.1, 0.2, 0.5, 1] km.

Figure 8 plots the results using the New York (Foursquare) and Tokyo (Foursquare) datasets. For each dataset, we observe a convex surface for preference inference accuracy. We analyze such results as follows.

- *Region Size* d. The small region size (small d) results in bad performance. In this case, users are hardly influenced by their PFRs because their current location rarely belongs to those small PFRs. In contrast, the large region size also generates unsatisfied results because some noisy activities might be included in large PFRs.
- *Threshold of visiting frequency* s_{freq}. The small s_{freq} implies that some detected PFRs might be areas that user occasionally visited. These PFRs are not necessarily reflecting their habitual behaviors and thus cause noise in preference inference. In contrast, the large s_{freq} implies that only highly frequented regions are considered, which cannot fully capture users' habitual behaviors, either.



(a) New York (Foursquare) (b) Tokyo (Foursquare)

Fig. 8. Parameter tuning of region size d and visiting frequency s_{freq}

The optimal value for these parameters ($d = 0.2$km and $s_{freq} = 0.01$) can be identified in Figure 8, where the Top 1 and Top 10 activity preference inference accuracy achieves the optimal value. An interesting observation is that the 0.2km optimal radius of PFRs is also in agreement with the optimal urban neighborhood radius identified by urban planning community in [31].

In the second experiment, we fix the optimal d and s_{freq} and decrease the threshold of ratios of preference bias $s_{ratio_{PB}}$ from 1 to 0 with the step of 0.1. Figure 9 shows the Top 1, Top 5 and Top 10 inference accuracy. A higher value of $s_{ratio_{PB}}$ implies that we only select PFRs with stronger preference bias. Therefore, some useful PFRs with $ratio_{PB}$ lower than $s_{ratio_{PB}}$ are eliminated so that user spatial activity preference cannot be fully described. With the decreasing threshold $s_{ratio_{PB}}$, more PFRs with relatively lower $ratio_{PB}$ are taken into account. Those PFRs with lower $ratio_{PB}$ has less ability to characterize user spatial activity preference. In the extreme case of $ratio_{PB} = 0$, where users conduct all activities equally, such PFRs cannot help to characterize user spatial activity preference at all. In Figure 9, we observe that there is no further improvement for $s_{ratio_{PB}} <= 0.4$ which indicates that the activity preference inference accuracy converges in terms of $s_{ratio_{PB}}$.

In the following, we set the three main parameters of PFRs as $d = 0.2km$, $s_{freq} = 0.01$ and $s_{ratio_{PB}} = 0.4$ for all three datasets.

2) ***Temporal Parameter Setting****:* We use the non-negative tensor factorization method to infer user temporal activity preference. The *latent space dimension* controls the number of the features involved in the factorization process. In this experiment, we vary the latent space dimension in the order of 8, 16, 32, 64 and 128. Figure 10 reports the comparison

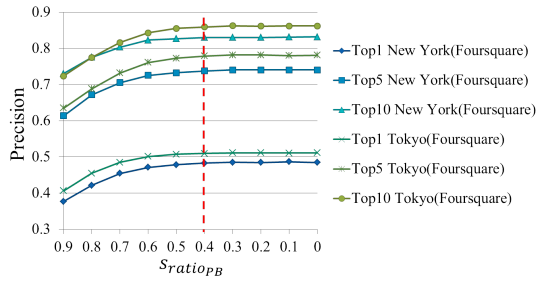


Fig. 9. Parameter tuning of threshold of preference bias ratio $s_{ratio_{PB}}$

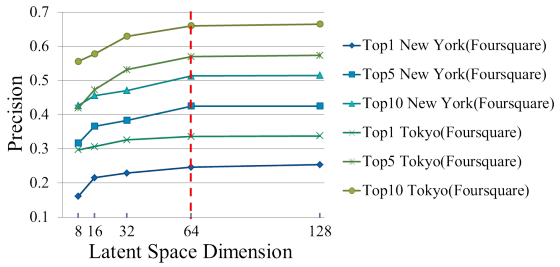


Fig. 10. Parameter tuning of latent space dimension

results. With the increase of the latent space dimension, the inference accuracy also increases. We observe no significant improvement in inference accuracy for dimension higher than 64, which indicates the convergence in terms of latent space dimension. Hence, in the following experiments, the latent space dimension is set to 64.

C. Comparison with Baseline Approaches

To evaluate the activity preference inference accuracy of the STAP model, we compare it with the following baseline approaches:

Sequential pattern mining approaches:

- *Order-K Markov Model (Markov-K)*. People's activities usually follow certain sequential patterns. For example, a user often has coffee after lunch; therefore it would be logical to infer her activity after lunch as having coffee in a coffee shop. Order-K Markov model considers the latest K activities of a user, and searches for the most frequent patterns to predict the next activity. We set K as 1 and 2 in the experiments.

Temporal based approaches:

- *Most Frequent Activity by Time (MFT)*. In general, people seem to conduct the same activity in the same time slot, which is usually regarded as a routine activity. For example, a user may have lunch around 12:00 during the weekdays. In this model, one's temporal activity preference is modeled by the distribution of her historical activity categories in each time slot of a week.
- *High Order Singular Vector Decomposition (HOSVD)*. HOSVD [32] is considered as a baseline for tensor factorization approach. It corresponds to the Tucker decomposition optimized for square-loss.
- *Temporal Model of STAP (Ours_NTF)*. The temporal activity preference model of STAP uses non-negative tensor factorization (NTF) approaches. It corresponds to the Canonical decomposition optimized for square-loss.

Spatial based approaches:

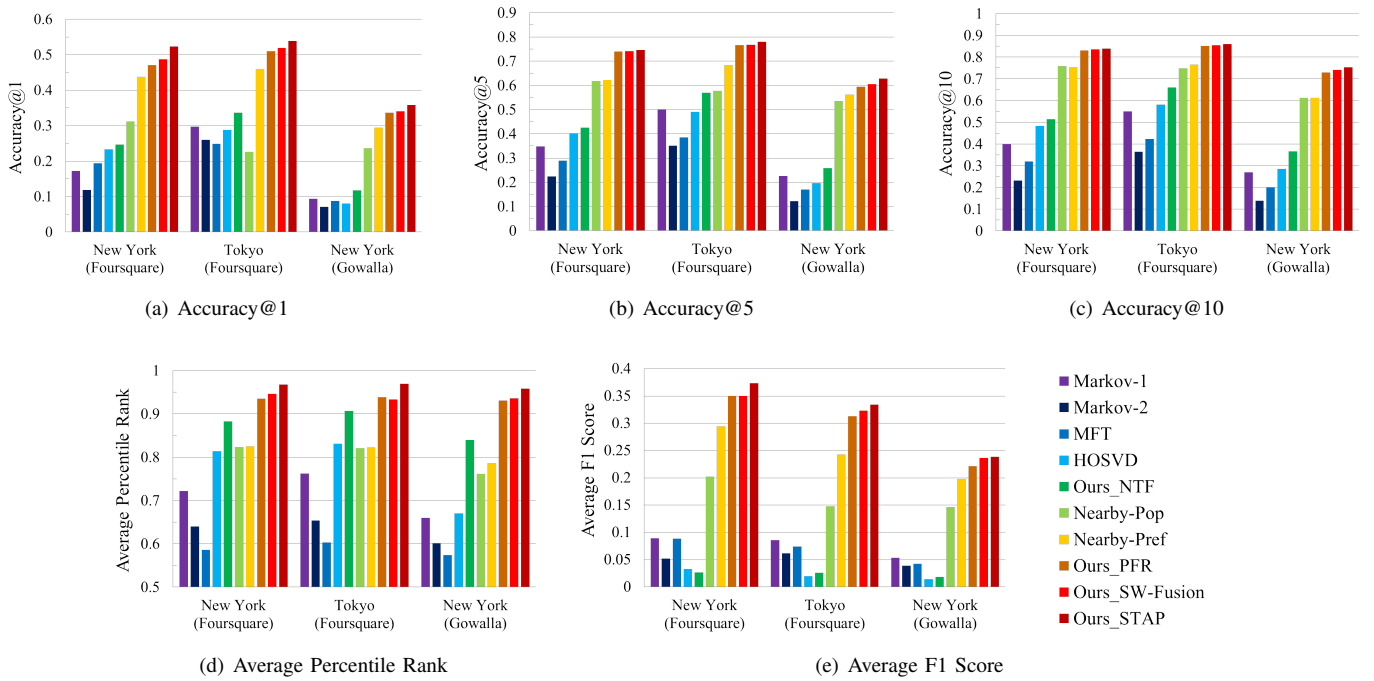
- *Most Popular Activity Around (Nearby-Pop)*. Using one's current location as center, it infers the user's activity preference according to the region's (radius = d) activity popularity. By activity popularity we mean the total number of check-ins for a specific activity category that we observe in the training dataset. It can be regarded as a simple non-personalized functional region based model.
- *Most Preferred Activity Around (Nearby-Pref)*. Using one's current location as center, it infers the user's activity preference according to the user's own activity popularity in the region (radius = d). By one's activity popularity we mean the number of check-ins of the user for a specific activity category that we observe in the training dataset. It can be regarded as a basic personalized model.
- *Spatial Model of STAP (Ours_PFR)*. The spatial activity preference model of STAP uses Personal Functional Regions to capture users' spatial activity preference.

Spatial temporal based approaches:

- *Static Weighted Fusion (Ours_SW-Fusion)*. The spatial and temporal preference distributions are combined using optimized static weight, i.e., $\Psi_{u,l,t} = \alpha\Psi_{u,t} + (1-\alpha)\Psi_{u,l}$. The optimized α is obtained when inference accuracy is maximized by increasing α from 0 to 1 with the step of 0.1, using the validation dataset. We then find $\alpha = 0.3$ with the New York (Foursquare) dataset and New York (Gowalla), and $\alpha = 0.4$ with the Tokyo dataset (Foursquare).
- *Ours_STAP*. The proposed STAP model uses the context-aware fusion framework.

Figure 11 shows the activity preference inference comparison with baselines under different evaluation metrics with the three datasets. We observe that our solution is consistently better than the other baseline approaches. Taking the Top 1 accuracy with the New York (Foursquare) dataset as an example, our solution is 203.54% better than the best sequential pattern mining approach, 124.83% better than the best temporal based approach, and 68.09% better than the best spatial based approach. We also conduct the one-tailed and two-tailed paired t-test over the results. We find that all the p-values are much less than 0.01, which proves that our STAP model is significantly better than the baselines in spatial temporal user activity inference task. In the following, we further analyze the performance of each baseline approach.

First, for sequential pattern mining approaches, both order 1 and order 2 Markov Model obtain unsatisfied results. In LBSNs, users have a choice to share their location information. Therefore, although user activities follow certain sequential patterns in their daily life, user check-ins do not fully contain their daily activities due to privacy concern or lack of time. Moreover, since we consider activities with fine granularity including 251 categories rather than 9 top categories, the large



(a) Accuracy@1 (b) Accuracy@5 (c) Accuracy@10

(d) Average Percentile Rank (e) Average F1 Score

Fig. 11. Comparison with baselines under different evaluation metrics

number of categories further aggravates the sparsity issue in the Markov Model when searching for frequent patterns.

Second, for temporal based approaches, tensor factorization methods, i.e., NTF and HOSVD, can better capture user activity preference than the frequency based approach, i.e., MFT. This observation shows that collaborative filtering can efficiently handle the sparse check-in data for user temporal activity preference inference. Furthermore, the improvement of NTF over HOSVD shows the advantage of considering non-negative constraint. The proposed temporal model using NTF can effectively capture the temporal characteristics of user activity preference, particularly for the users whose activities show strong temporal regularities.

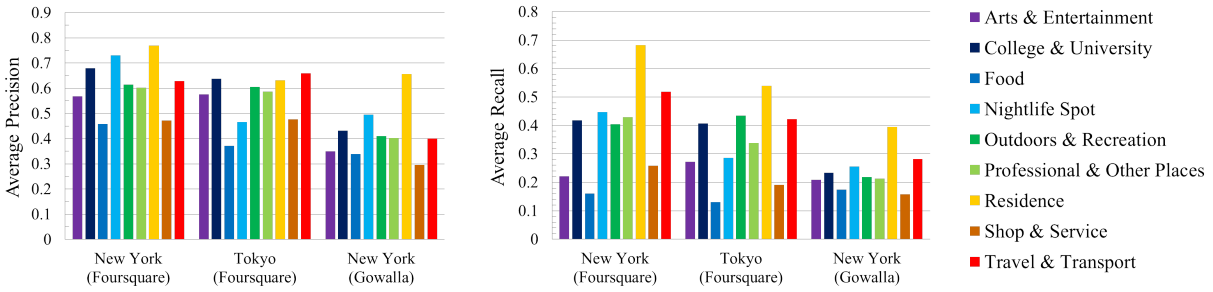
Third, spatial based approaches lead to better performance than temporal based methods. This observation shows that the spatial regularity of user activity in LBSNs is more significant than the temporal regularity. Specifically, Nearby-Pref performs better than Nearby-pop baseline due to the consideration of personal preference. The improvement of PFR over Nearby-Pref shows the advantages of eliminating noisy data in capturing spatial features of user activity preference. In other words, the infrequent activities of a user may not actually reflect her preference. The proposed Personal Functional Region can delicately capture the spatial characteristics of user activity preference, particularly for the users whose activities exhibit obvious spatial specificity.

Finally, compared to the static weighted fusion method SW-Fusion, the context-aware fusion framework achieves the best performance. It takes advantage of both spatial and temporal features under varying contexts. An interesting observation is that the improvement of considering the temporal model from merely considering the spatial model is relatively small, which further shows that the importance of spatial features in modeling user activity preferences in LBSNs. Moreover, by comparing the $Accuracy@1$ of the spatial and temporal models, we find that a large number of activities can be correctly inferred by both spatial and temporal models. For example, there are 20.6% of the activities in the test dataset can be correctly inferred by both models for the Top 1 accuracy with the New York (Foursquare) dataset.

D. Comparison between Different Datasets

By comparing the results obtained from different datasets, we observe some interesting findings.

First, comparing the results on the New York (Foursquare) and Tokyo (Foursquare) datasets, we find that: a) the accuracy difference between temporal based approaches and spatial based approaches is relatively small with the Tokyo dataset (e.g., 0.34 for Top 1 NTF and 0.51 for Top 1 PFR) than that with the New York dataset (e.g., 0.25 for Top 1 NTF and 0.47 for Top 1 PFR). The most possible explanation is that Tokyo users' activities have stronger temporal regularities than those of New York users; b) the improvement of PFR-based approaches over the non-personalized functional region based approach (Nearby-Pop) is larger with the Tokyo dataset than with the New York dataset, particularly for Top 1 activity inference accuracy. Such an improvement may probably be explained by two reasons: 1) higher density of venues implies higher diversity of nearby activities in Tokyo, which causes less difference for top popular activities; 2) Tokyo users have stronger preference bias in their PFRs, resulting in higher accuracy for PFR-based approaches.



(a) Precision (b) Recall

Fig. 12. Comparison between different activity categories using STAP model

Second, comparing the results on the New York (Foursquare) and New York (Gowalla) datasets, we find that our solution consistently achieves better performance than the baselines. This is due to the fact that users in different LBSNs often exhibit similar spatial-temporal preference activity patterns, which enables us to model the user activity preference over different LBSNs. Furthermore, we find that the performance is slightly better with the New York (Foursquare) dataset than that with the New York (Gowalla) dataset. It can probably be explained by the fact that users in New York (Gowalla) dataset show broader activity preference than that of users in New York (Foursquare) dataset, which makes it more difficult in preference inference task. Specifically, Gowalla users indicate activities in more categories (55.58 on average per user) than Foursquare users (38.37 on average per user).

E. Comparison between Different Activity Categories

Due to the fact that users' behaviors in some activity categories often show stronger spatial temporal regularities than that in other categories, we investigate such difference by calculating the precision and recall for individual categories using STAP model. Rather than exhaustively listing all the results for the 251 categories, we present the average precision and recall for each of the 9 root-categories in Foursquare, i.e., Arts & Entertainment, College & University, Food, Great Outdoors, Nightlife Spot, Professional & Other Places, Residence, Shop & Service, Travel & Transport. Figure 12 presents the results with the three datasets.

First, in all the three datasets, we observe that categories in Residence and College & University yield good precision and recall, which implies that users in LBSNs exhibit strong spatial temporal regularities in activities like going home and going to school. Intuitively, these activities are usually conducted at regular times and places. We also observe that categories in Shop & Service and Food show relatively lower precision and recall, which implies that users in LBSNs have relatively flexible temporal and spatial preference for shopping or going to a restaurant.

Second, there are also some activity categories yield different results with different datasets. Specifically, with both New York (Foursquare) and New York (Gowalla) datasets, nightlife activities exhibit high precision and recall, which implies that New York users tend to enjoy their nightlife at regular time and places. In addition, transportation related activities show higher spatial temporal regularities with Tokyo (Foursquare) dataset, which is probably due to the fact that Tokyo users often check in when they are on their daily commute.

VIII. Conclusion

Understanding user spatial temporal activity preference can benefit users by providing them with customized location based services. However, it is difficult to directly tackle such four dimensional data, i.e., user-location-time-activity quadruples, which usually suffers from data sparsity problem. This paper presents STAP, a spatial temporal activity preference model. To reduce the problem complexity, STAP separately considers the spatial and temporal features of user activities by introducing the notion of spatial specificity and temporal correlation. First, spatial specificity suggests that users usually conduct certain specific activities in their frequented areas. We define Personal Functional Regions to quantitatively measure one's preference bias in her frequented regions and use them to infer spatial activity preference. Second, temporal correlation suggests that users with the similar lifestyle tend to have similar activity preference at the similar time. We resort to tensor factorization techniques to collaboratively build temporal activity preference from the sparse check-in data. Finally, we propose a context-aware fusion framework to make best use of the advantage of both features in activity preference inference. We experimentally evaluate STAP using three datasets collected from two LBSNs, i.e., Foursquare and Gowalla. The experiment results show that the STAP model achieves consistently good performance with all three datasets and outperforms various baseline approaches, which verifies the generality and advantages of our solution in modeling spatial-temporal activity preference with sparse check-in data.

In the future, we plan to broaden this work in several directions. First, we plan to study different geographical representations of PFRs in order to better characterize user spatial activity preference. Second, we intend to take a step forward to build a POI recommendation application based on the inferred user activity preference. Third, we plan to further explore new ways of accommodating the accumulated user activity data from LBSNs and enabling effective, scalable and personalized location based services.

APPENDIX A
PROOF OF PROPOSITION 1

In one frequented region r of a user u, the number of user visited location categories $|C_{u,r}|$ is less than or equal to that of the existing categories $|C_l^d|$, i.e., $|C_{u,r}| \leq |C_l^d|$. The maximum entropy of a variable x is when x follows the uniform distribution. Thus, we have

$$H_{max}(C_{u,r}) = \log_2 |C_{u,r}| \leq H_{max}(|C_l^d|) = \log_2 |C_l^d| \quad (22)$$

The entropy of user u's actual visit distribution $H(\psi_{u,r})$ is less than or equal to that of uniform visit distribution $H_{max}(C_{u,r})$. Thus, we obtain

$$H(\psi_{u,r}) \leq H_{max}(C_{u,r}) \leq H_{max}(|C_l^d|) \quad (23)$$

Since entropy is non-negative, we have

$$0 \leq \frac{H(\psi_{u,r})}{H_{max}(|C_l^d|)} \leq 1 \quad (24)$$

ACKNOWLEDGMENT

The authors would like to thank the editors and reviewers for their valuable comments and suggestions.

REFERENCES

[1] D. Lian and X. Xie, "Collaborative activity recognition via check-in history," in *Proc. LBSN*, 2011, pp. 45–48.

[2] Y. Jihang, Z. Zhe, , and C. Hong, "What's your next move: user activity prediction in location-based social networks," in *Proc. SDM*, 2013, pp. 171–179.

[3] P. Fabio, A. Xueli, K. Fahim, , and I. Hiroki, "Discovering and predicting user routines by differential analysis of social network traces," in *Proc. WoWMoM*, 2013, pp. 1–9.

[4] N. Anastasios, M. Cecilia, and F.-M. Enrique, "Exploiting foursquare and cellular data to infer user activity in urban environments," in *Proc. MDM*, 2013, pp. 167–176.

[5] Z. Wang, D. Zhang, D. Yang, Z. Yu, and X. Zhou, "Discovering and profiling overlapping communities in location based social networks," *IEEE Trans. on SMC: System*, vol. 44, no. 4, pp. 499–509, April 2014.

[6] H. Gao, J. Tang, and H. Liu, "Exploring social-historical ties on location-based social networks," in *Proc. ICWSM*, 2012, pp. 114–121.

[7] D. Zhang, B. Guo, and Z. Yu, "The emergence of social and community intelligence," *IEEE Computer*, vol. 44, no. 7, pp. 21–28, 2011.

[8] J. Chang and E. Sun, "Location3: How users share and respond to location-based data on social networking sites," in *Proc. ICWSM*, 2011, pp. 74–80.

[9] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "Mining user mobility features for next place prediction in location-based services," in *Proc. ICDM*, 2012, pp. 1038–1043.

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

[11] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "A random walk around the city: New venue recommendation in location-based social networks," in *Proc. SocialCom*, 2012, pp. 144–153.

[12] C. Cheng, H. Yang, I. King, and M. Lyu, "Fused matrix factorization with geographical and social influence in location-based social networks," in *Proc. AAAI*, 2012, pp. 17–23.

[13] D. Yang, D. Zhang, Z. Yu, and Z. Wang, "A sentiment-enhanced personalized location recommendation system," in *Proc. HT*, 2013, pp. 119–128.

[14] C. Biancalana, F. Gasparetti, A. Micarelli, and G. Sansonetti, "An approach to social recommendation for context-aware mobile services," *ACM Trans. on Intelligent Systems and Technology*, vol. 4, no. 1, p. 10, 2013.

[15] L. Huang, Q. Li, and Y. Yue, "Activity identification from gps trajectories using spatial temporal pois' attractiveness," in *Proc. LBSN*, 2010, pp. 27–30.

[16] J. Antikainen, "The concept of Functional Urban Area," *Informationen zur Raumentwicklung*, pp. 447–456, 2005.

[17] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and pois," in *Proc. KDD*, 2012, pp. 186–194.

[18] 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. WSDM*, 2013, pp. 375–384.

[19] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, "Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering," in *Proc. RecSys*, 2010, pp. 79–86.

[20] D. Yang, D. Zhang, Z. Yu, and Z. Yu, "Fine-grained preference-aware location search leveraging crowdsourced digital footprints from lbsns," in *Proc. UbiComp*, 2013, pp. 479–488.

[21] B. Hidasi and D. Tikk, "Fast als-based tensor factorization for context-aware recommendation from implicit feedback," in *Proc. ECML PKDD*, 2012.

[22] P. Symeonidis, A. Papadimitriou, Y. Manolopoulos, P. Senkul, and I. Toroslu, "Geo-social recommendations based on incremental tensor reduction and local path traversal," in *Proc. LBSN*, 2011, pp. 89–96.

[23] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with gps history data," in *Proc. WWW*, 2010, pp. 1029–1038.

[24] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Towards mobile intelligence: Learning from gps history data for collaborative recommendation," *Artificial Intelligence*, vol. 184, pp. 17–37, 2012.

[25] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "An empirical study of geographic user activity patterns in foursquare," in *Proc. ICWSM*, 2011, pp. 570–573.

[26] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui, "Exploring millions of footprints in location sharing services," in *Proc. ICWSM*, 2011, pp. 81–88.

[27] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[28] L. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sep. 1966.

[29] J. Kim and H. Park, "Fast nonnegative tensor factorization with an active-set-like method," *High-Performance Scientific Computing*, pp. 311–326, 2012.

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

[31] M. Mehaffy, S. Porta, Y. Rofè, and N. Salingaros, "Urban nuclei and the geometry of streets: The 'emergent neighborhoods' model," *Urban Design International*, vol. 15, no. 1, pp. 22–46, 2010.

[32] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, Mar. 2000.