Information Processing and Management 57 (2020) 102337

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Dynamic discovery of favorite locations in spatio-temporal social networks

Xi Xiong[a,c], Fei Xiong[b,*], Jun Zhao[c,*], Shaojie Qiao[d,*], Yuanyuan Li[e], Ying Zhao[f]

[a] School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, China
[b] School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
[c] School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore
[d] School of Software Engineering, Chengdu University of Information Technology, Chengdu 610225, China
[e] West China School of Medicine, Sichuan University, Chengdu 610041, China
[f] School of Public Administration, Sichuan University, Chengdu 610065, China

A R T I C L E I N F O

Keywords:
Location-based social networks
POI recommendation
Heterogeneous networks
Factor graph model
Network embedding

A B S T R A C T

A large volume of data flowing throughout location-based social networks (LBSN) gives support to the recommendation of points-of-interest (POI). One of the major challenges that significantly affects the precision of recommendation is to find dynamic spatio-temporal patterns of visiting behaviors, which can hardly be figured out because of the multiple side factors. To confront this difficulty, we jointly study the effects of users' social relationships, textual reviews, and POIs' geographical proximity in order to excavate complex spatio-temporal patterns of visiting behaviors when the data quality is unreliable for location recommendation in spatio-temporal social networks. We craft a novel framework that recommends any user the POIs with effectiveness. The framework contains two significant techniques: (i) a network embedding method is adopted to learn the vectors of users and POIs in an embedding space of low dimension; (ii) a dynamic factor graph model is proposed to model various factors such as the correlation of vectors in the previous phase. A collection of experiments was carried out on two real large-scale datasets, and the experimental outcomes demonstrate the supremacy of the proposed method over the most advanced baseline algorithms owing to its highly effective and efficient performance of POI recommendation.

1. Introduction

The rapid urban expansion has greatly extended the physical boundary of living areas and formed a large number of POIs (points of interest). Location-based social networks (LBSNs) like Foursquare[1] and Yelp[2] have grown to be widely accepted services for mobile users to establish personal social community, make friends, provide reviews, or raise complaints about various POIs, such as restaurants, hotels, stores and cinemas. LBSNs have created valuable social data for hotspots that are of common interest to the public. Users are allowed to request their most preferred POIs by *POI recommendation* approaches, which automatically capture the interests of users from the behavior records. As a key service of LBSNs, POI recommendation helps users to explore attractive sites, and makes companies aware of the target users to whom the advertisements can be pushed.

* Corresponding authors.
E-mail addresses: flyxiongxi@gmail.com (X. Xiong), xiongf@bjtu.edu.cn (F. Xiong), junzhao@ntu.edu.sg (J. Zhao), sjqiao@cuit.edu.cn (S. Qiao).

[1] https://www.foursquare.com
[2] https://www.yelp.com

https://doi.org/10.1016/j.ipm.2020.102337
Received 8 March 2020; Received in revised form 19 May 2020; Accepted 6 June 2020
0306-4573/

Table 1
Average distance between each POI and its visitors during the specific time period, as well as the percentage of the visits during the period in the whole day in Singapore.

POI (Type)	Monday			Saturday		
	3pm	6pm	9pm	3pm	6pm	9pm
Spize (Food)	0.3km 3.1%	3.7km 18.4%	4.5km 9.3%	2.5km 6.4%	6.1km 16.7%	8.8km 12.9%
ATLAS (Bar)	0.2km 0.8%	3.1km 2.5%	4.1km 13.9%	0.4km 1.3%	6.8km 4.7%	5.0km 15.2%
PS.Cafe (Cafe)	0.8km 12.3%	3.1km 7.2%	0.5km 5.1%	0.6km 9.8%	3.3km 6.9%	0.3km 4.7%
Oceans (Market)	4.5km 4.3%	6.4km 26.2%	–	6.9km 10.6%	8.1km 14.6%	–

1.1. Motivation and challenges

Unlike traditional digital recommender systems that deliver virtual information such as news, music or films (Hu, Li, Shi, Yang, & Shao, 2020), POI recommendation in LBSNs generally involves users' social relationships (e.g., social communities) and the physical features of POIs (e.g., locations) (Xin & Wu, 2020). Although a variety of technical issues, such as the cold start problem (Huang, Shen, & Meng, 2019) and the sparsity of user data (Yin, Zhou, Shao, Wang, & Sadiq, 2016), which affect the POI recommendation performance have been addressed in use of the recently proposed approaches (Wang et al., 2017), there are still many hard problems to solve, and one of them is to find *the dynamic spatio-temporal patterns of visiting behaviors*.

We designed a crawler to extract the user reviews on Foursquare including tips and also photo comments on four popular restaurants, which belong to different categories, respectively in Singapore. Once visiting a POI, some visitors made a review. So it was practical that for each popular POI we selected 100 active visitors who posted reviews. Table 1 illustrates the average distance between each POI and its visitors during a specific time period, as well as the percentage of the visits during the period in the whole day. 3pm and 9pm represent the time periods of 15:00–16:00 and 21:00–22:00, respectively.

From Table 1, we find that users' activities showed intensive temporal cyclic patterns with the implying regularity by hour or by week. A popular restaurant attracted more customers at the dinner time and cost a longer distance during the weekend. He/She was more likely to visit a restaurant near to his/her working place with the colleagues for convenience after the working time (at 6:00pm Mon.) rather than a distant restaurant. A neighboring cafe is also a good choice for meeting with business clients (at 3:00pm Mon.). In the weekend evening (at 9:00pm Sat.), he/she probably moved to a bar for a relax time with her intimate friends.

In addition, current studies illustrate that users' historical visits on POIs have spatial characteristics. A person is more likely to visit POIs near their current locations or near the ones in their historical records (Ye, Yin, Lee, & Lee, 2011). Visited POIs form some spatial clusters, i.e., people prefer exploring around some centers, such as "*home*" and "*office*" at the specific time. We observe the inherent relationships between the attributes at different time or different locations. The time-varying correlated values of an attribute depict the changing of a user's emotion. It is difficult to figure out a consistent spatial-temporal pattern of people's visiting behaviors because people's preference of visiting POIs varies and depends on multiple factors.

1.2. Our contributions

To discover the underlying information and make effective recommendation for POIs, we need to identify the relationship of dynamic factors with fine granularity. The factor graph model (FGM) (Lin et al., 2017), has proved to be useful in the modeling of the dynamic spatio-temporal features of LBSNs. It is one of the most popular models in recommender systems since it sufficiently characterizes the dynamic relations among auxiliary data due to its flexibility in the modeling of time-location-variant systems. Compared with the sequential models based on recurrent neural networks utilizing a series of items with almost identical intervals, the FGM can capture the temporal relations with various periods. To the best of our knowledge, we are the first to study the dynamic correlation between the fine-grained spatio-temporal social relationships in a LBSN via the FGM, which considers the relationships between factors in an effective and explainable manner. One advantage of the proposed framework is that the multiple types of relationships are integrated into a unified model based on a dynamic spatio-temporal social network.

Effective extraction and representation of useful information from the LBSNs is a challenging task due to the sparsity of the relationships among users and POIs. Rather than representing entities in terms of their similarities in the existing studies, we provide a novel approach to learn the representation of spatio-temporal social networks by summarizing critical structural features, such as social relationships between users and geographical proximity between POIs.

It is more reasonable to incorporate the effect of adjacent neighbors for location recommendation rather than the indirect neighbors. In this study, we propose a breadth-first search strategy to represent network nodes in the form of low-dimensional vectors (namely, embedding vectors), which incorporate underlying information such as the static geo-social relationships among nodes. Unlike the uniformed metrics of similarity, these learned embedding vectors with a more compact form are convenient to use and integrate.

Fig. 1 illustrates the basic idea of the proposed model. The relationships among people and POIs are converted into low-dimensional vectors to eliminate the data sparsity and improve the effectiveness of recommendation. Then, several factors including geo-social relationships, time, textual sentiments and geographical distances are involved as the inputs to predict users' tendencies of visiting different POIs.

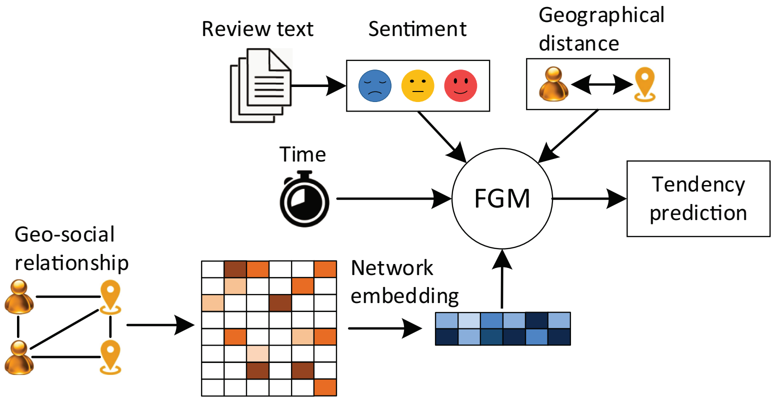

Fig. 1. Overview of the proposed POI recommendation approach.

The contributions of this study include:

1) The effects of users' social relationships, textual reviews, and POIs' geographical proximity are jointly studied in order to excavate complex spatio-temporal patterns of visiting behaviors when the data quality is unreliable for location recommendation in spatio-temporal social networks.
2) An original framework of semi-supervised learning named **Dy**namic **S**patio-**t**empor**al** POI recommendation (DYSTAL) is presented. There are two key components in this framework: (i) a network embedding method learns the vectors of users and POIs in an embedding space with low dimensionality; (ii) a dynamic factor graph model is proposed to model different factors including the correlation of users' vectors and POIs' vectors.
3) The performance of DYSTAL were verified by a collection of experiments based on two given real massive datasets. According to the experimental results, DYSTAL is the most advanced methods owing to its highly effective and efficient performance of POI recommendation.

Following contents are divided into sections: literature review appears in Section 2; some definitions and problem setting is presented in Section 3. POI recommendation framework is posed in Section 4; a network embedding method that integrates various relationships is introduced in Section 5; we give the dynamic factor graph model that models different factors in Section 6; moreover, we display experimental results in Section 7 and conclude our paper in Section 8.

2. Related works

Currently, the effective POI recommendation algorithms on LBSNs generally combine crucial subjective and objective factors such as the impact of social relationships, the spatio-temporal influence and the sentimental effect from textual content.

Impact of social relationships. As the social media develop rapidly, social analysis and mining techniques are receiving great attention. A social community is a group of people closely connecting and sharing akin interests. To clarify the impact of the user opinions, matrix factorization method, proposed by Yang, Lei, Liu, and Li (2017), transformed users into low-dimensional embeddings in terms of trust networks. Some sparse information including social trust relationships among a group of users and rating scores are involved in the model. For the purpose of overcoming the limitations of current embedding-based and path-based approaches for recommendation based on knowledge graphs, Wang et al. (2018) proposed an end-to-end recommendation method to simulate the propagation of users' preferences like the actual ripples propagating on the surface of water. A user's potential interests are automatically and iteratively extended over the set of knowledge entities along links in a knowledge graph. Ying et al. (2018) proposed an algorithm based on the graph convolutional network (GCN), where efficient graph convolutions and random walks are integrated to create embedding vectors of nodes that contain information of both graph structures and node features. Wang, Zhao, Xie, Li, and Guo (2019) introduced an end-to-end model that captures the relationships between items effectively by discovering their associated attributes on a knowledge graph.

Spatio-temporal influence. Users' behaviors have spatial and temporal features which can greatly affect the accuracy of recommendation. Yao (2018) suggested personalized POIs by considering the temporal matching degree between users and POIs. Specifically, they profiled the temporal popularities of POIs and characterized users in terms of the latent regularity. Huang, Ma, Liu, and Sangaiah (2017) created a unified probabilistic generative model for understanding the joint effect of social relationship, time and geographical locations from users' check-in activities, and formulate a POI selection method in reality. This model learns the vectors of users and POIs by a network embedding approach in a shared latent topic space to incorporate the social influence and temporal influence effectively. Zhao et al. (2017) proposed a hierarchical geographical matrix factorization model to utilize the hierarchical structures of both users and POIs for POI recommendation. Then, Zhao et al. (2019) proposed a new spatio-temporal gated network by enhancing long-short term memory network, where spatio-temporal gates were introduced to capture the spatio-temporal relations between successive checkins. Ying et al. (2019) proposed a time-aware metric embedding approach with

Table 2
Notations and their descriptions.

Symbol	Description
$\mathcal{U}, \mathcal{V}$	the set of users and POIs, respectively
$\mathcal{T}, \mathcal{T}'$	the set of time periods in $[1, T_0]$ and (T_0, ∞), respectively
t, t'	a short time period in $\mathcal{T}$ and $\mathcal{T}'$, respectively
G	a spatio-temporal social network
$\mathcal{E}$	a set of edges in G and $\mathcal{E} = \mathcal{E}_{uu} \cup \mathcal{E}_{uu} \cup \mathcal{E}_{uv}$
$\mathcal{Y}^t$	a set of check-in records at t
$\mathcal{Y}_L^t, \mathcal{Y}_U^t$	a set of labeled and unlabeled tendencies at t, respectively
$C^{\mathcal{T}}$	a set of reviews within $\mathcal{T}$
u_i, v_j	a user in $\mathcal{U}$, a POI in $\mathcal{V}$
$\vec{u_i}, \vec{v_j}$	the embedding vector of u_i and v_j, respectively
x_{ij}^t	the correlation between u_i and v_j at t
y_{ij}^t	u_i's visiting tendency to v_j at t
d_{ij}^t	the distance between u_i and v_j at t
s_{ij}^t	u_i's review sentiment for v_j at t
$Z, Z_\alpha, Z_\beta\, Z_\gamma, Z_\delta, Z_\eta$	normalization terms
λ	the learning ratio
$\boldsymbol{\theta}$	the parameter set
$\boldsymbol{S}$	the aggregation of factor functions over all nodes

asymmetric projection for successive POI recommendation.

Sentimental effect in textual content. According to the study by Yin et al. (2016), the textual content (Xiong et al., 2020a) of POIs can incorporate local attributes and users' interests for handling data sparsity in out-of-town recommendation (e.g., when traveling to another city). Ren, Song, Haihong, and Song (2017) performed POI recommendation by their context-based probabilistic matrix factorization method that incorporates multiple factors including social relationships, textual content, geographical locations, popularities of POIs as well as categorical correlation. There were also researchers like Meng, Wang, Liu, and Zhang (2018) who integrated positive and negative emotion of the reviews for recommender systems in a principled and mathematical method from the global and local perspectives.

By analyzing the above approaches, the spatio-temporal factors from the behaviors of users, especially the correlation between visiting records, are considered inappropriately in the case of POI recommendation. Some of them only model the dynamic features of users and POIs based on the time-variant and location-variant data by means of the probabilistic methods, which degrade the accuracy of recommendation. It motivates our upcoming recommendation model which is more accurate and more efficient, in considering the following two points primarily: 1) embedding multiple factors involving the geo-social relationships into POI recommendation, and 2) calculating the dynamic correlation between users' visiting records that include the information of distances and textual sentiments.

3. Problem statement

3.1. Notations and definitions

The POI recommendation problem will be formalized in this section after introducing preliminary definitions and notations. Table 2 lists the key notations used throughout the paper.

Generally speaking, a spatio-temporal social network is typically denoted by a graph where users and POIs can be represented by nodes and their relationships can be represented by edges between nodes. Here, we will give the following important definitions.

Definition 1 (*POI***).** In our model, a POI is a uniquely identified place (e.g., a cafe or a cinema), containing two attributes in the proposed model, the identifier v and the geographical location l_v coordinated by the longitude and the latitude.

In order to share their experience and feeling on travel with friends, users often post reviews or give remarks on POIs. In this study, reviews are treated as the textual content of a POI and the textual information out of a user's reviews are integrated to give accurate recommendation on locations.

Definition 2 (*Textual sentiment***).** The textual sentiments represent the latent sentiments underlying in the textual content of POIs (Li, Xiong, Qiu, Wang, & Xu, 2018; Xiong et al., 2019b). The review posted by user u_i for POI v_j at time t is regarded as belonging to a major sentiment s_{ij}^t.

In this study, we use the tool of *SentiStrength* to annotate each short text by a positive or negative sentimental score. *SentiStrength* has the following merits over other tools: it is helpful to capture the slangs and acronyms in short informal texts. It suits well to the data processing for social media since it is sensitive to various linguistic rules with respect to amplification, booster word, negation, emotion, word spelling. We can compute a polarity score in $[-4, 4]$ by combining the positive and negative sentimental scores as introduced in Xiong et al. (2018).

Definition 3 (*Tendency*). The tendency denotes a probability or a confidence that a user visits a POI. $y_{ij}^t \in [0, 1]$ represents u_i's tendency on v_j at time t. The value of a tendency is regarded as 0 or 1 for a known visiting record, i.e., it is set to 1 if the user visited the POI at least once, otherwise set to 0. The tendency usually has a value $y_{ij}^t \in (0, 1)$ for an unknown visiting record.

Definition 4 (*Spatio-temporal social network*). This network can be denoted by $G = \{\mathcal{U}, \mathcal{V}, \mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}, C^{\mathcal{T}}\}$, where $\mathcal{U}$ represents the user set and $\mathcal{V}$ represents the POI set. The users and POIs have dynamic attributes which change with time. The set of check-in records $\mathcal{Y}^{\mathcal{T}'}$ can be partitioned into two parts: $\mathcal{Y}_L^{\mathcal{T}'}$ and $\mathcal{Y}_U^{\mathcal{T}'}$, which represent the record set with and without tendency, respectively. $C^{\mathcal{T}}$ is the set of textual reviews in accordance with $\mathcal{Y}_L^{\mathcal{T}}$ with the information of posting time.

Definition 5 (*Short time period*). In order to show the dynamic model clearly, we slice the continuous time to a series of short periods. Let t denote one of the periods. The nodes are updated only once during one period.

Only a small part of visiting records contain reviews, thus we utilize $C^{\mathcal{T}}$ and $\mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}$ for training. This study attempts to provide POI recommendation with effectiveness and efficiency. The location recommendation task is formalized as below.

3.2. Problem statement

Let $G = \{\mathcal{U}, \mathcal{V}, \mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}, C^{\mathcal{T}}\}$ be a network, where $\mathcal{T} = [1, T_0]$ and $\mathcal{T}' = (T_0, \infty]$. We must figure out a mapping function r that predicts the unlabeled tendencies for a specific user visiting a set of POIs during the time period $\mathcal{T}'$:

$$r: G = \left\{\mathcal{U}, \mathcal{V}, \mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}, C^{\mathcal{T}}\right\} \rightarrow \mathcal{Y}_U^{\mathcal{T}'} \tag{1}$$

4. Working mechanism

To find the best mapping r presented in Section 3, we craft the dynamic spatio-temporal POI recommendation framework (DYSTAL), illustrated in Fig. 2.

Specifically the proposed framework of POI recommendation is a four-phase implementation:

1) Data collection from famous LBSN portals, including: a) users' profiles and POIs' profiles; b) the relationships of users and POIs; c) the check-in records and the reviews including textual content, time and locations.
2) Calculate the embedding vectors of users and POIs by integrating POIs' geographical proximity, users' social relationships and visiting records.
3) Learn the relationship between the current and the future tendency via a dynamic factor graph model by taking into full consideration several factors including the correlation between each user and each POI, the textual reviews, the check-in records and the distance from a user to a POI. Then, we can work out the probability for each unvisited POI that a user may choose in the future.
4) Top-k POIs of the highest tendencies at later time are recommended to a given user.

5. Geo-social network embedding

In this section, we calculate the vectors of users and POIs according to the friend relationships of users, the geographical relationship of POIs, and the visiting history from users to POIs. The most popular approach of network embedding is DeepWalk (Perozzi, Al-Rfou, & Skiena, 2014), which employs a truncated random walk to learn the embeddings of social nodes, which performs like a depth-first search. This approach can rapidly alleviate the sparsity of nodes' neighborhood by incorporating indirect neighbors. However, the geo-social network with multiple kinds of nodes cannot be applied directly to this method, which requires all nodes to be the same type. A more reasonable strategy is to expand the neighborhood of each node by a breadth-first searching method. In light of the LINE model (Tang et al., 2015), we construct a novel network embedding model, valid for both heterogeneous and homogeneous networks, to learn the embedding vectors of nodes belonging to the same type. Moreover, we also

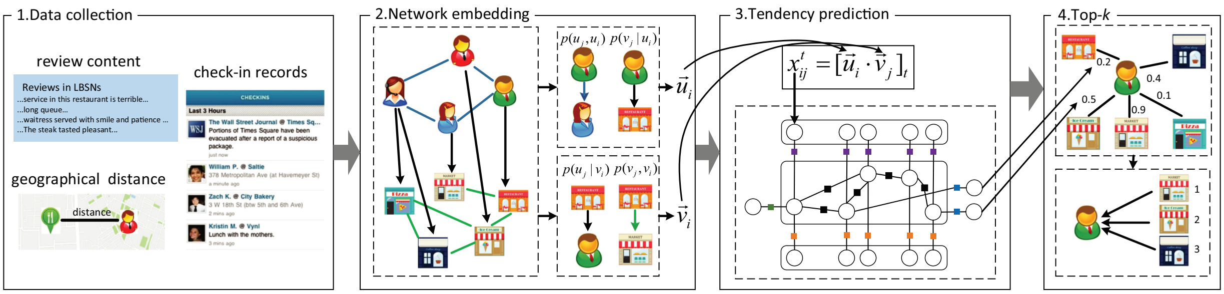

Fig. 2. Working mechanism of the proposed POI recommendation approach.

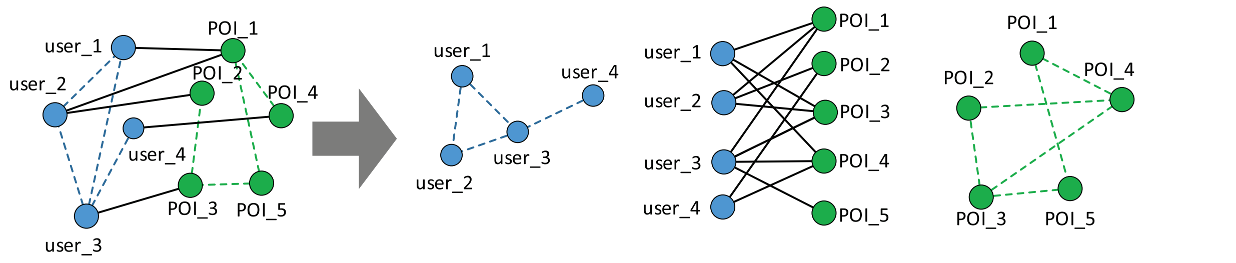

Fig. 3. Illustration of encoding geographical locations, historical reviews and sentiments from reviews into a low-dimensional vector by the graph-based method.

modify this model to recognize the vectors of nodes belonging to two categories, i.e., users and POIs in a LBSN.

As shown in Fig. 3, the embedding vectors are learned by our graph-based approach. What can be extracted from data are three kinds of graphs, two of which are homogeneous subgraphs and the other one as a bipartite subgraph.

5.1. Formalizing the objective functions

Pieces of crucial information co-occur when there is similarity or relevance between users and POIs. Our network embedding method is based on the fact that there will be close co-occurrence of users and POIs in an embedding space of low dimension if they get characterized similarly by geographical locations, user profiles, POI profiles, etc.

As shown in Fig. 3, in order to capture the latent information (the geographical effect of POIs, the patterns of visiting behaviors as well as the influence of users' social relationships), the heterogeneous network (Fig. 3(a)) is partitioned into two homogeneous subgraphs (Fig. 3(b) and (d)) and a bipartite subgraph (Fig. 3(c)). These subgraphs consist of one or two types of nodes. Here defined are these three subgraphs.

Definition 6 (*POI-POI subgraph*). Let $G_{vv} = (\mathcal{V}\bigcup\mathcal{V}, \mathcal{E}_{vv})$ be the POI-POI subgraph, which displays the local geographical proximity between POIs. Geographical clustering approaches (e.g., the DBSCAN algorithm (Xiong et al., 2020b; 2020c)) are used to cluster all of the POIs into several regions. Two POIs have proximity if they belong to the same type (e.g., two restaurants) and are clustered in one region. Every couple of POIs are linked by one edge (v_i, v_j), whose weight $w_{v_iv_j}$ represents the proximity degree between v_i and v_j and is computed by: $w_{v_iv_j} = 1/(1 + e^{d(v_i,v_j)})$, $d(v_1, v_2)$ being the geographical distance from v_1 to v_2. A short distance between two neighboring POIs indicates a strong proximity value in the interval of (0, 1). Generally speaking, users favor a region with a great many POIs identical in type.

Definition 7 (*user-user subgraph*). The user-user subgraph is represented by $G_{uu} = (\mathcal{U}\bigcup\mathcal{U}, \mathcal{E}_{uu})$, which represents the social relationship in a social network, that is, the friendship between two users. Every friend pair are linked by one edge (u_i, u_j), whose weight $w_{u_iu_j}$ represents the intimacy degree between u_i and u_j and is equal to the number of interaction between u_i and u_j within a period.

Definition 8 (*user-POI subgraph*). The user-POI subgraph, represented by $G_{uv} = (\mathcal{U}\bigcup\mathcal{V}, \mathcal{E}_{uv})$, is a bipartite subgraph with $\mathcal{U}$ being a collection of users and $\mathcal{V}$ being a collection of POIs. $\mathcal{E}_{uv}$ denotes the collection all possible edges that bridge the users to the POIs. Formally, $e_{ij} \in \mathcal{E}_{uv}$ if and only if v_j were visited by u_i. w_{ij} is just the number of check-ins or reviews on v_j by u_i.

Our network embedding tries to characterize these subgraphs in a space with a small d as the number of dimensions. In $\mathbb{R}_d$, user i and POI j are expressed by $\vec{u_i}$ and $\vec{v_j}$, the so-called embedding vectors, respectively.

1) For G_{vv}, the joint probability of v_i and v_j are defined according to the following formula:

$$p(v_i, v_j) = \frac{1}{1 + \exp(-\vec{v}_j^T \cdot \vec{v_i})} \tag{2}$$

where $\vec{v_i}$ denotes the embedding vector of v_i.

Eq. (2) is a distribution $p(\cdot, \cdot)$ over $V \times V$ and the empirical probability over all POI pairs is written as $\hat{p}(v_i, v_j) = w_{v_iv_j}/\mathcal{W}$, where $\mathcal{W} = \sum_{(i,j)\in\mathcal{E}_{vv}} w_{v_iv_j}$.

The objective function is defined for the learning of the embedding vectors $\vec{v_i}$ and $\vec{v_j}$, which represent v_i and v_j in a low-dimensional space, respectively.

$$O_{vv} = d(\hat{p}(v_i, v_j), p(v_i, v_j)) \tag{3}$$

where $d(\cdot, \cdot)$ denotes the KL-divergence of two distributions. In order to keep the proximity of two POIs, we minimize the objective function in Eq. (3), which can be transformed to the following equation by ignoring some constants that will not affect the result.

$$O_{vv} = -\sum_{e_{ij}\in\mathcal{E}_{vv}} w_{v_iv_j}\log p(v_i, v_j) \tag{4}$$

In the learning of the optimal $\{\vec{v_i}\}_{i=1\ldots|\mathcal{V}|}$ that minimizes this objective, we represent each v_i as a vector $\vec{v_i}\in\mathbb{R}_d$.

2) Similarly, we calculate the objective function corresponding to the homogeneous subgraph G_{uu} according to the following equation:

$$O_{uu} = -\sum_{e_{ij}\in\mathcal{E}_{uu}} w_{u_iu_j}\log p(u_i, u_j) \tag{5}$$

In the learning of the optimal $\{\vec{u_i}\}_{i=1\ldots|\mathcal{U}|}$ which minimizes this objective, each user u_i is denoted as a vector $\vec{u_i}\in\mathbb{R}_d$.

3) For the bipartite G_{uv}, the conditional probability of a POI v_j ($v_j\in\mathcal{V}$) visited by a user u_i ($u_i\in\mathcal{U}$) is defined according to the following equation:

$$p(v_j|u_i) = \frac{\exp(\vec{v}_j^{\,T}\cdot\vec{u_i})}{\sum_{v_k\in\mathcal{V}}\exp(\vec{v}_k^{\,T}\cdot\vec{u_i})} \tag{6}$$

Eq. (6) defines a conditional distribution $p(\cdot|u_i)$ and its corresponding empirical probability has the form $\hat{p}(v_j|u_i) = w_{u_iv_j}/deg_{u_i}$, where $deg_{u_i} = \sum_{v_k\in\delta(u_i)} w_{u_iv_k}$ and $\delta(u_i)$ is the adjacent nodes of u_i in G_{uv}. The empirical distribution $\hat{p}(\cdot|u_i)$ must approach the conditional distribution $p(\cdot|u_i)$ in order to obtain the weight $w_{u_iv_j}$ over the edge e_{ij}. So the objective function is defined as below for the learning of the embedding vectors $\vec{u_i}$ and $\vec{v_j}$, which represent u_i and v_j, respectively.

$$O_{uv} = \sum_{e_{ij}\in\mathcal{E}_{vu}} \lambda_i d(\hat{p}(v_j|u_i), p(v_j|u_i)) \tag{7}$$

where λ_i is the importance of u_i in G_{uv}, which can be replaced by deg_{u_i}. Eq. (7) is equivalent to the following equation.

$$O_{uv} = -\sum_{e_{ij}\in\mathcal{E}_{uv}} w_{u_iv_j}\log p(v_j|u_i) \tag{8}$$

In the learning of $\{\vec{u_i}\}_{i=1\ldots|\mathcal{U}|}$ and $\{\vec{v_j}\}_{j=1\ldots|\mathcal{V}|}$ that minimize this objective function, we write u_i and v_j as $\vec{u_i}$, $\vec{v_j}\in\mathbb{R}_d$, respectively.

4) Based on the same subgraph G_{uv}, we define the conditional probability of user u_j ($u_j\in\mathcal{U}$) visiting POI v_i ($v_i\in\mathcal{V}$) as follows:

$$p(u_j|v_i) = \frac{\exp(\vec{u}_j^{\,T}\cdot\vec{v_i})}{\sum_{u_k\in\mathcal{U}}\exp(\vec{u}_k^{\,T}\cdot\vec{v_i})} \tag{9}$$

Then, we obtain the objective function according to the following equation:

$$O_{vu} = -\sum_{e_{ij}\in\mathcal{E}_{vu}} w_{v_iu_j}\log p(u_j|v_i) \tag{10}$$

In the learning of $\{\vec{u_j}\}_{j=1\ldots|\mathcal{U}|}$ and $\{\vec{v_i}\}_{i=1\ldots|\mathcal{V}|}$ that minimize this objective function, we write u_j and v_i as $\vec{u_j}$, $\vec{v_i}\in\mathbb{R}_d$, respectively.

5.2. Joint embedding learning

The embedding phase involves three subgraphs: POI-POI, user-user and user-POI. All the objective functions mentioned before must be considered in order to learn the embeddings for these three subgraphs. An intuitive method is to minimize the sum of all objective functions as follows:

$$O = O_{vv} + O_{uu} + O_{vu} + O_{uv} \tag{11}$$

For every step, a small set of edges are sampled by the negative sampling method (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013), and then input into the asynchronous stochastic gradient algorithm (ASGD) for parameter updating. Provided with the sampled edge $e_{ij}\in\mathcal{E}_{uv}$, the gradient over the embedding vector $\vec{u_i}$ is calculated according to the following equation:

$$\frac{\partial O_{uv}}{\partial\vec{u_i}} = w_{u_iv_j}\cdot\frac{\partial\log p(v_j|u_i)}{\partial\vec{u_i}} \tag{12}$$

With respect to the objectives (4), (5) and (8), the negative sampling method is used by replacing $\vec{u_i}$ and $\vec{v_j}$, and then the subsequent equations can be obtained:

$$\frac{\partial O_{vv}}{\partial\vec{v_i}} = w_{v_iv_j}\cdot\frac{\partial\log p(v_i, v_j)}{\partial\vec{v_i}} \tag{13}$$

$$\frac{\partial O_{uu}}{\partial\vec{u_i}} = w_{u_iu_j}\cdot\frac{\partial\log p(u_i, u_j)}{\partial\vec{u_i}} \tag{14}$$

Input: A POI-POI graph G_{vv}, a user-user graph G_{uu}, a POI-user grahp G_{vu}, a learning factor η and the number of negative samples R.
Output: A POI embedding vector $\vec{v}$ and a user embedding vector $\vec{u}$.

while *not convergence* **do**
    Sample an edge from $\mathcal{E}_{vv}$ and extract R negative edges;
    Sample an edge from $\mathcal{E}_{uv}$ and extract R negative edges;
    Update POI embeddings by Eq. 16 ;
    Sample an edge from $\mathcal{E}_{uu}$ and extract R negative edges;
    Sample an edge from $\mathcal{E}_{vu}$ and extract R negative edges;
    Update user embeddings by Eq. 17 ;
end
return $\vec{v}$, $\vec{u}$;

Algorithm 1. Joint embedding training.

$$\frac{\partial O_{vu}}{\partial \vec{v_i}} = w_{v_i u_j} \cdot \frac{\partial \log p(u_j|v_i)}{\partial \vec{v_i}} \tag{15}$$

Furthermore, the embedding vectors can be updated according to the following formulas:

$$\vec{v_i} = \vec{v_i} - \eta \frac{\partial O}{\partial \vec{v_i}} = \vec{v_i} - \eta \left(\frac{\partial O_{vv}}{\partial \vec{v_i}} + \frac{\partial O_{vu}}{\partial \vec{v_i}}\right) \tag{16}$$

$$\vec{u_i} = \vec{u_i} - \eta \frac{\partial O}{\partial \vec{u_i}} = \vec{u_i} - \eta \left(\frac{\partial O_{uu}}{\partial \vec{u_i}} + \frac{\partial O_{uv}}{\partial \vec{u_i}}\right) \tag{17}$$

where η is the learning rate.

We use a method called *joint embedding training* (detailed in Algorithm 1) to merge all edges in these four sets (i.e., $\mathcal{E}_{vv}$, $\mathcal{E}_{uv}$, $\mathcal{E}_{uu}$, $\mathcal{E}_{vu}$) and performs sampling on these four sets sequentially.

According to Algorithm 1, we sample one positive edge and R negative edges from each set of edges by the negative sampling method, and then update the corresponding embedding vector, respectively.

5.3. Computational complexity analysis

The computational complexity of the optimization with the negative sampling method is $O(t_0 \times (1 + R))$, with the constant R being the number of negative samples and the constant t_0 denoting the time spent on one round of sampling. Generally, the number of steps for optimization is proportional to the number $|\mathcal{E}|$ of edges in the network. Therefore, the entire computational complexity of optimization is $O(t_0 \times R \times |\mathcal{E}|)$. The proposed edge sampling method in Algorithm 1 is efficient because it is linear to the number of edges $|\mathcal{E}|$ and is irrelevant to the number of nodes $|\mathcal{V}|$. Therefore, the overall complexity of the network embedding is $O(|\mathcal{E}|)$.

6. Dynamic prediction of tendency

In this section, we predict the probable POIs via computing the time-varying tendency by taking into full consideration the textual sentiment information, the spatial distances and the learned embeddings of POIs and users in Section 5.

Generally speaking, we have two methods to model this problem. The first way is to use a node in a bipartite graph to represent each user and POI, respectively. For each user, we are attempt to calculate the probability distributions of different relationships from a user to both his/her neighboring POIs in a bipartite graph and his friends in a social network, respectively. However, this intuitive method suffers from the limitation that it is a tough task to compute the correlations between two relationships, whose phase is costly with high computational complexity. An alternative method is to treat each relationship as a node in the graphical model, then the task of relationship mining is transformed to predict the semantic label for each relationship node in the model which greatly reduces the computational complexity. More importantly, it can calculate different correlations between relationships.

In order to infer the probabilities of social relationships, we have three basic intuitions: (1) users with intimate relationships such as close friends or family members may share similar sentiment towards a specific POI in their review content; (2) a user is more likely to pay attention to items near to her current place because a longer distance usually costs more time; (3) users' attitudes may change over time. For instance, a user prefers to different restaurants in a weekday or a weekend based on her different moods.

It is difficult to model the correlations between users and dynamic spatio-temporal factors by conventional classifiers such as SVM. We exploit a factor graph model (FGM) (Lin et al., 2017) to incorporate geographical distances, time and review content for learning and predicting users' tendency to visit different POIs. Compared with the sequential models based on recurrent neural networks utilizing a series of items with almost identical intervals, the FGM can capture the temporal relations with various periods. An FGM is a category of probabilistic graphical model, which represents both undirected and directed graphical structures in an elegant manner, on the emphasis of the factorization of the distribution. The critical point of an FGM is to define correlations by multiple categories of factor functions, and the objective function is generated based on the joint probability of the factor functions.

6.1. Model description

Based on the above fundamental theories, we propose a dynamic factor graph model called Spatio-Temporal Factor Graph Model (ST-FGM), as shown in Fig. 4. The "dynamic" means the model changes with time. Each relationship between two nodes in partially labeled network G is mapped to a relationship node such as y_{12}^t in ST-FGM.

The ST-FGM model can be explained in detail. The relationships in the input are denoted by four types of nodes in Fig. 4. Beneficial from these learned embeddings of POIs and users in Section 5, we use x_{ij}^t to represent the geo-social similarity of a pair of user-POI at time t by the following equation:

$$x_{ij}^t = [\vec{u_i} \cdot \vec{v_j}]_t \tag{18}$$

Assume y_{ij}^t represents u_i's tendency to visiting v_i at time t, s_{ij}^t represents u_i's sentiment on v_i at time t, which can be discovered from review texts, and d_{ij}^t denotes the distance between u_i and v_j at time t. Based on these four types of correlations, we define the following five correlation factor functions:

1) **Tendency correlation factor** $r_1(y_{ij}^t, y_{i'j'}^t)$. It represents the correlation between two tendencies y_{ij}^t and $y_{i'j'}^t$. This factor can be

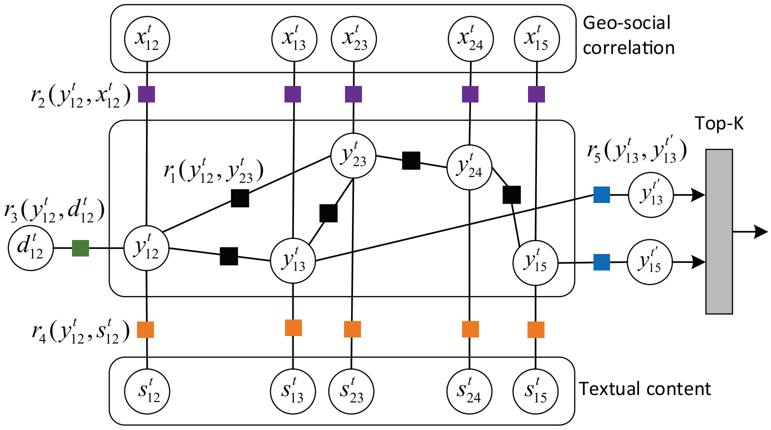

Fig. 4. Graphical representation of the ST-FGM model.

denoted by the following equation.

$$r_1\left(y_{ij}^t, y_{i'j'}^t\right)=\frac{1}{Z_\alpha}\exp\left\{\boldsymbol{\alpha}\cdot\boldsymbol{f}\left(y_{ij}^t, y_{i'j'}^t\right)\right\} \tag{19}$$

where $\boldsymbol{\alpha}$ is a weighting vector. Different pairs of nodes share the same $\boldsymbol{\alpha}$ vector. $\boldsymbol{f}(\cdot, \cdot)$ is the **tendency correlation function** to depict the tendency correlation and will output a $(|\mathcal{U}||\mathcal{E}|)^2$ vector. If two tendencies start at the same user or end at the same POI, a tendency correlation is set to 1 between these two tendencies. For instance, the value of $f(y_{12}^t, y_{13}^t)$ is equal to 1 because both y_{12}^t and y_{13}^t start at u_1. Similarly, the value of $f(y_{42}^t, y_{52}^t)$ is equal to 1 because both y_{42}^t and y_{52}^t ends at v_2. Z_α is a normalization term which is equal to $\sum_{i,j,i',j'}\exp\{\boldsymbol{\alpha}_{ij}\cdot\boldsymbol{f}(y_{ij}^t, y_{i'j'}^t)\}$.

2) **Geo-social correlation factor** $r_2(y_{ij}^t, x_{ij}^t)$. x_{ij}^t contains the geo-social information of u_i and v_j. This correlation denotes the dependency relationship between y_{ij}^t and x_{ij}^t. This factor can be denoted by the following equation:

$$r_2(y_{ij}^t, x_{ij}^t) = \frac{1}{Z_\beta}\exp\{\boldsymbol{\beta_{ij}}\cdot\boldsymbol{m}(x_{ij}^t)\} \tag{20}$$

where $\boldsymbol{\beta_{ij}}$ is a weighting vector w.r.t. v_i and v_j. $\boldsymbol{m}(\cdot)$ is the vector of indicator functions. $m(x_{ij}^t)$ is set to 1 when x_{ij}^t is larger than the threshold which is equal to the average value of x_{ij}^t. Z_β is a normalization term.

3) **Geographical correlation factor** $r_3(y_{ij}^t, d_{ij}^t)$. The previous study (Yin et al., 2016) has showed that one is reluctant to visit POIs that are far away from her. The correlation denotes the dependency between y_{ij}^t and d_{ij}^t. This factor can be denoted by the following equation:

$$r_3(y_{ij}^t, d_{ij}^t) = \frac{1}{Z_\gamma}\exp\{\boldsymbol{\gamma_{ij}}\cdot\boldsymbol{h}(d_{ij}^t)\} \tag{21}$$

where γ_{ij} is a weighting vector w.r.t. v_i and v_j. $\boldsymbol{h}(\cdot)$ is the vector of indicator functions. $h(d_{ij}^t)$ is set to 1 when d_{ij}^t is larger than a constant value. Z_γ is a normalization term.

4) **Sentimental correlation factor** $r_4(y_{ij}^t, s_{ij}^t)$. The review content usually contains latent sentiment (Xiong et al., 2018). The correlation denotes the dependency between y_{ij}^t and s_{ij}^t. This factor can be denoted by the following equation:

$$r_4(y_{ij}^t, s_{ij}^t) = \frac{1}{Z_\delta}\exp\{\boldsymbol{\delta_{ij}}\cdot\boldsymbol{g}(s_{ij}^t)\} \tag{22}$$

where $\boldsymbol{\delta_{ij}}$ is a weighting vector w.r.t. v_i and v_j. $\boldsymbol{g}(\cdot)$ is the vector of indicator functions. $g(s_{ij}^t)$ is set to 1 when s_{ij}^t is larger than the constant threshold. The SentiStrength tool is utilized to calculate the sentimental score of a short text, which ranges from -4 to 4. Therefore, we define the sentimental threshold to 0 to only detect the positive sentiment. Z_δ is a normalization term.

5) **Temporal correlation factor** $r_5(y_{ij}^t, y_{ij}^{t'})$. Some recent researchers have verified that there is a strong dependency between one's current tendency to visit POIs and the tendency in the past (Wang et al., 2015). This correlation denotes the dependency between y_{ij}^t and $y_{ij}^{t'}$. This factor can be denoted as follows:

$$r_5(y_{ij}^t, y_{ij}^{t'}) = \frac{1}{Z_\eta}\exp\{\boldsymbol{\eta_{tt'}}\cdot\boldsymbol{l}(y_{ij}^t, y_{ij}^{t'})\} \tag{23}$$

where $\boldsymbol{\eta_{tt'}}$ is a weighting vector w.r.t. t and t'. $\boldsymbol{l}(\cdot, \cdot)$ is the **temporal correlation function** to depict the temporal correlation and will output a vector. Since y_{ij}^t has a value in [0,1], $l(y_{ij}^t, y_{ij}^{t'})$ is set to 1 when both y_{ij}^t and $y_{ij}^{t'}$ are equal to 1. Z_η is a normalization term.

All parameters in the above functions are randomly initialized. Based on these factors, we define the joint distribution of the model over the spatio-temporal social network $G = \{\mathcal{U}, \mathcal{V}, \mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}, \mathcal{C}^{\mathcal{T}}\}$. The joint distribution is calculated by multiplying these five kinds of correlation factor functions.

$$\begin{aligned} p(\mathcal{Y}_U^{\mathcal{T}'}|G) &= \frac{1}{Z}\prod_t \prod_{i,j} \prod_{i',j'} r_1\left(y_{ij}^t, y_{i'j'}^t\right) \\ &\cdot \prod_t \prod_{i,j} r_2\left(y_{ij}^t, x_{ij}^t\right) \cdot \prod_t \prod_{i,j} r_3\left(y_{ij}^t, d_{ij}^t\right) \\ &\cdot \prod_t \prod_{i,j} r_4\left(y_{ij}^t, s_{ij}^t\right) \cdot \prod_t \prod_{t'} \prod_{i,j} r_5\left(y_{ij}^t, y_{ij}^{t'}\right) \\ &= \frac{1}{Z}\exp\{\theta^T \boldsymbol{S}\} \end{aligned} \tag{24}$$

where $Z = Z_\alpha Z_\beta Z_\gamma Z_\delta Z_\eta$ is the normalization term, $\boldsymbol{S}$ is the aggregation of factor functions over all nodes, and $\theta = \{\alpha, \boldsymbol{\beta_{ij}}, \boldsymbol{\gamma_{ij}}, \boldsymbol{\delta_{ij}}, \boldsymbol{\eta_{tt'}}\}$ denotes all the parameters.

6.2. Model learning

The challenge for learning the ST-FGM model is that only part of visiting records contains review content, i.e., some users prefer to checking-in rather than writing reviews. To calculate the parameters Z and $\boldsymbol{\theta}$, it is necessary to sum up the likelihood of possible states for all nodes (including nodes with no values). To address this problem, we use the records with reviews to infer the unknown possible visits, represented by $p(\mathcal{Y}_U^{\mathcal{T}'}|G)$, which means a prediction of $\mathcal{Y}$ inferred from the partially known records. Therefore, we employ a concise representation for the conditional probability in the following equations. For instance, $p(\mathcal{Y}|G)$ is a concise version of $p(\mathcal{Y}_U^{\mathcal{T}'}|G)$. Then we define the log-likelihood objective function $O(\theta)$ as the follows:

$$\begin{aligned} &O(\theta) = \log p(\mathcal{Y}|G) = \log \textstyle\sum_{\mathcal{Y}|\mathcal{Y}_L} \frac{1}{Z} p(\mathcal{Y}_L|G) \\ &= \log \textstyle\sum_{\mathcal{Y}|\mathcal{Y}_L} \frac{1}{Z}\exp\{\theta^T \boldsymbol{S}\} \\ &= \log \textstyle\sum_{\mathcal{Y}|\mathcal{Y}_L} \exp\{\theta^T \boldsymbol{S}\} - \log Z \\ &= \log \textstyle\sum_{\mathcal{Y}|\mathcal{Y}_L} \exp\{\theta^T \boldsymbol{S}\} - \log \textstyle\sum_{\mathcal{Y}} \exp\{\theta^T \boldsymbol{S}\} \end{aligned} \tag{25}$$

where $\mathcal{Y}|\mathcal{Y}_L$ is short for $\mathcal{Y}_U^{\mathcal{T}'}|\mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}$, which means the probability of tendencies after T_0 under the condition of $\mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}$. Then, the gradient of $\boldsymbol{\theta}$ can be calculated as follows:

$$\begin{aligned} \nabla_\theta &= \frac{\partial O(\theta)}{\partial \theta} \\ &= \frac{\partial(\log \sum_{\mathcal{Y}|\mathcal{Y}_L} \exp\{\theta^T \boldsymbol{S}\} - \log \sum_{\mathcal{Y}} \exp\{\theta^T \boldsymbol{S}\})}{\partial \theta} \\ &= \frac{\sum_{\mathcal{Y}|\mathcal{Y}_L} \exp\{\theta^T \boldsymbol{S}\}\cdot \boldsymbol{S}}{\sum_{\mathcal{Y}|\mathcal{Y}_L} \exp\{\theta^T \boldsymbol{S}\}} - \frac{\sum_{\mathcal{Y}} \exp\{\theta^T \boldsymbol{S}\}\cdot \boldsymbol{S}}{\sum_{\mathcal{Y}} \exp\{\theta^T \boldsymbol{S}\}} \\ &= E_{p_\theta(\mathcal{Y}|\mathcal{Y}_L, G)}\boldsymbol{S} - E_{p_\theta(\mathcal{Y}|G)}\boldsymbol{S} \end{aligned} \tag{26}$$

Another challenge is that the graphical structure in ST-FGM may be arbitrary and contain loops and it is a difficult task to directly calculate the second expectation value $E_{p_\theta(\mathcal{Y}|G)}\boldsymbol{S}$. In this study, we employ the Loopy Belief Propagation (LBP) (Murphy, Weiss, & Jordan, 1999), an approximate algorithm, to address this problem. Here, we estimate the marginal probability $p(y_i|\boldsymbol{\theta})$ by LBP. Based on the marginal probabilities, the gradient is computed by summing up all relationship nodes. It is worthwhile to note that we need to conduct the LBP twice to estimate $p_\theta(y_i|\mathcal{Y}_L, G)$ and $p_{\boldsymbol{\theta}}(y_i|G)$, respectively. The learning algorithm of ST-FGM is shown in Algorithm 2.

In Algorithm 2, we initiate the parameters in $\boldsymbol{\theta}$ with random values (line 1). In each iteration, we estimate $p_\theta(y_i|\mathcal{Y}_L, G)$ and $p_{\boldsymbol{\theta}}(y_i|G)$ by the LBP method, respectively (lines 3–4). Then, update $\boldsymbol{\theta}$ via the gradient algorithm (lines 5–8).

6.3. Inferring recommendation list

Based on the learned parameters in $\boldsymbol{\theta}$, we utilize the LBP algorithm to maximize the joint probability according to Eq. (24):

$$\mathcal{Y}^* = \arg\max_{\mathcal{Y}|\mathcal{Y}_L} p(\mathcal{Y}|G) \tag{27}$$

Then, we calculate the marginal probability of each tendency node $p_{ij} = p\,(y_{ij}^{t'}|\mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}, G)$. For user $u_i \in \mathcal{U}$, we provide a recommendation list composed of k POIs with largest values of $p_{ij} * y_{ij}^{t'}$.

6.4. Computational complexity analysis

The intrinsic difficulty of ST-FGM is to calculate the normalization factor Z by summing up all possible nodes in Fig. 4. Then, the computational complexity of ST-FGM equals $O(|\mathcal{U}|^2)$.

Input: A spatio-temporal social network $G = \{\mathcal{U}, \mathcal{V}, \mathcal{Y}_L^{\mathcal{T}+\mathcal{T}'}, \mathcal{C}^{\mathcal{T}}\}$, the learning ratio λ

Output: The learned parameter $\boldsymbol{\theta} = \{\alpha, \beta_{ij}, \gamma_{ij}, \delta_{ij}, \eta_{tt'}\}$

Initiate $\boldsymbol{\theta}$ randomly;

while *not convergence* **do**

  Compute $E_{p_{\boldsymbol{\theta}}(\mathcal{Y}|\mathcal{Y}_L, G)}\boldsymbol{S}$ by LBP;

  Compute $E_{p_{\boldsymbol{\theta}}(\mathcal{Y}|G)}\boldsymbol{S}$ by LBP;

  Compute the gradient of $\boldsymbol{\theta}$ by Eq. 26:

$$\nabla_{\boldsymbol{\theta}} = E_{p_{\boldsymbol{\theta}}(\mathcal{Y}|\mathcal{Y}_L, G)}\boldsymbol{S} - E_{p_{\boldsymbol{\theta}}(\mathcal{Y}|G)}\boldsymbol{S};$$

  Update $\boldsymbol{\theta}$ with λ:

$$\boldsymbol{\theta} = \boldsymbol{\theta} + \lambda \cdot \nabla_{\boldsymbol{\theta}};$$

end

return $\boldsymbol{\theta}$;

Algorithm 2. The learning algorithm of ST-FGM.

7. Experiments

In this section, a series of experiments are conducted to evaluate the performance of the proposed DYSTAL framework on real datasets. We first introduce the experimental settings and then compare our experimental results with the baseline approaches. All experiments were carried out in the environment with Java 10, Windows 10 as well as a PC with a Core (i7 8700) CPU and 16GB RAM.

7.1. Experimental settings

7.1.1. Datasets

The following two real datasets are used in the experiments:

`Foursquare dataset.` Foursquare is a fashionable LBSN and can provide several location-based services, e.g., presenting online reviews for POIs and sharing photos. This dataset was collected from 74,250 users located in Singapore, and involves users' profiles, review content as well as POIs' profiles such as the location information in terms of the latitude and the longitude.

`Yelp Challenge dataset.` [3] The dataset involves 1,326,101 users and 174,567 POIs from ten cities in four countries. In order to concentrate on POIs in a specific city, we use the data from Las Vegas of USA, which contains about 26,809 POIs and 1,605,396 review records. Each record contains textual content and meta data such as the timestamps and the locations when reviews were posted. The city of the POIs the user most visited is regarded as his/her living place, which is not marked in the dataset. We extracted 337,084 users located in Las Vegas in the dataset.

In order to clean up the data and remove the outliers that seldomly appear, we reserve the users that posted more than ten reviews and each user posted reviews for at least five different POIs. We also constrain that a POI is given at least ten reviews. Due to the absence of some location information such as people's homes and working places, we first specified these locations by the following method: 1) clustered the POIs by means of *the DBSCAN algorithm* (Xiong et al., 2019a), and 2) obtained the homes and the working places by calculating the central points during the working time and the weekend.

7.1.2. Baseline approaches

We compare the proposed DYSTAL approach with the following state-of-the-art POI recommendation techniques.

TRUSTMF Yang et al. (2017) transforms users into low-dimensional embeddings by the matrix factorization method in terms of the trust networks, which facilitates the understanding of the influence by users' opinions. Some sparse information including the rating scores and the social trust relationships among a group of users are involved in the model.

LSARS Wang et al. (2017) is a latent probabilistic generative model which emulates users' check-in activities for both home-town and out-of-town scenarios by considering user interest drift and crowd sentiment. It learns individual's interests relevant to locations and sentiments from the profiles of POIs and reviews of users. However, the geographical influence is not considered when generating the review contents in LSARS.

MMBE Huang et al. (2017) is a unified probabilistic generative model to discover the joint effect of social relationship, time and geographical locations from users' check-in activities, and formulate a POI selection method. This model learns the vectors of users and POIs by a network embedding approach in a shared latent topic space via incorporating the social influence and temporal influence effectively.

HEREC Shi, Hu, Zhao, and Yu (2017) is a novel model for POI recommendation in heterogeneous networks by embedding several kinds of factors. The model produces informative sequential nodes for network embedding by a random walk approach. The nodes are encoded by different fusion functions and incorporated into a new matrix factorization model, which are used for rating prediction.

TGSC-PMF Ren et al. (2017) is a context-based probabilistic matrix factorization method for recommending POIs that incorporate multiple factors including geographical locations, social relationship, textual content, popularity of POIs as well as categorical correlation.

PINSAGE Ying et al. (2018) is a GCN-based algorithm, which integrates efficient random walks and graph convolutions to create embedding vectors of nodes that comprise both graph structure and node feature information.

KGCN Wang et al. (2019) is an end-to-end model that captures the relationships between items effectively by discovering their associated attributes on a knowledge graph.

The features used by the above POI recommendation methods are shown in Table 3, where the first row contains the names of factors including geographical influence (denoted by Geo), temporal effect (denoted by Temp), users' preferences (denoted by Pref), social relationship (denoted by Soc), users' behaviors (denoted by Beh), textual content (denoted by Tex) as well as network embedding (denoted by Emb). Users' preferences are the static information shown directly on the websites rather than the dynamic information inferred by meta-data such as historical visiting records.

7.1.3. Evaluation metrics

As advised by the previous literatures (Huang et al., 2017), we exploit four widely-used metrics including *precision* and *recall, F1 score* and *hit rate* to measure the effectiveness and efficiency of different POI recommendation methods. In the POI recommendation

[3] https://www.yelp.com/dataset/challenge

Table 3
Features of different recommendation methods.

	Geo	Temp	Pref	Soc	Beh	Tex	Emb
TRUSTMF			•	•	•		
LSARS			•		•	•	
MMBE	•	•	•	•	•		•
HEREC			•	•			•
TGSC-PMF	•		•	•	•	•	
PINSAGE				•	•	•	•
KGCN		•			•		•
DYSTAL	•	•		•	•	•	•

task, precision measures the proportion of ground-truth POIs in the recommended POIs, and recall is equal to the ratio of ground-truth POIs to all the actually visited POIs. F1 score is calculated as the harmonic mean of the precision and the recall.

Supposing $\mathcal{V}_u$ and $\mathcal{R}_u$ denote the visited POIs and recommended POIs w.r.t. user u. The metrics of precision, recall and F1 score are defined as follows:

$$Precision = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{|\mathcal{V}_u \cap \mathcal{R}_u|}{|\mathcal{R}_u|} \tag{28}$$

$$Recall = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{|\mathcal{V}_u \cap \mathcal{R}_u|}{|\mathcal{V}_u|} \tag{29}$$

$$F1 - -score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{30}$$

Hit rate represents the fraction of hits in accurate prediction (Hu & Ester, 2013), where one hit denotes that the recommendation list w.r.t. a specific user involves at least one POI that the user has visited. Hit rate is defined as given below:

$$Hit - -rate = \frac{|hit|}{|D_{test}|} \tag{31}$$

where $|hit|$ and $|D_{test}|$ denote the number of hits in the testing set and the number of all testing cases, respectively. Specially for a single testing case, $|hit|$ is equal to 1 if the ground-truth POI appears in the top-k list, or 0 if not.

Given a collection of recommendation records, we partition each dataset into ten portions and use the 10-fold cross validation method to evaluate. Thus, the average precision and recall over all users can be adopt for the purpose of exact evaluation.

7.2. Effectiveness analysis of location recommendation

The comparison between DYSTAL and the baseline methods with well-tuned parameters is presented in this section. Figs. 5 and 6 demonstrate the performance of POI recommendation with different k values in `the Foursquare dataset` and `the Yelp dataset`, respectively. In experiments, k is specified to {1,5,10,20}, for a larger k value is meaningless for location recommendation and therefore always out of consideration in the top-k recommendation.

As shown in Figs. 5 and 6, comparing to other baseline methods, DYSTAL has 0.2% to 13.0% higher effectiveness. The reason is four-fold:

1) DYSTAL and MMBE incorporate different categories of factors, especially the spatio-temporal information, which shows that both time and locations are very useful for POI recommendation. Extracting spatio-temporal information can effectively improve the performance of recommendation. These two models and HEREC also have another common phase, i.e., network embedding, which

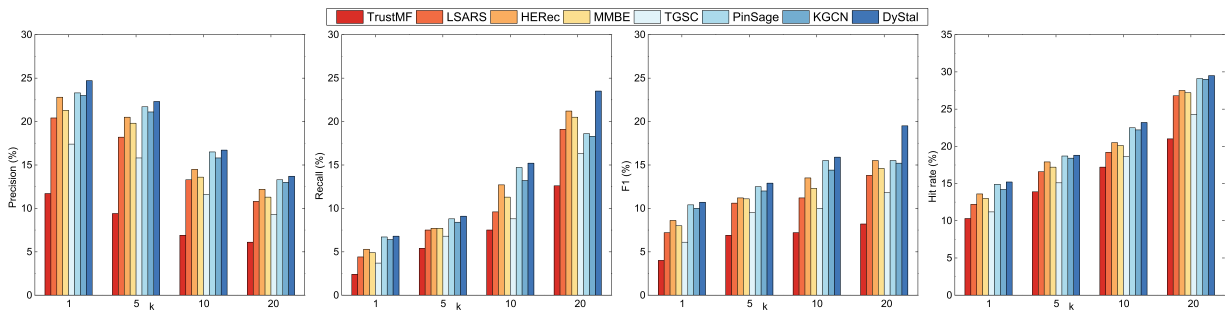

Fig. 5. Performance comparison on `the Foursquare dataset`.

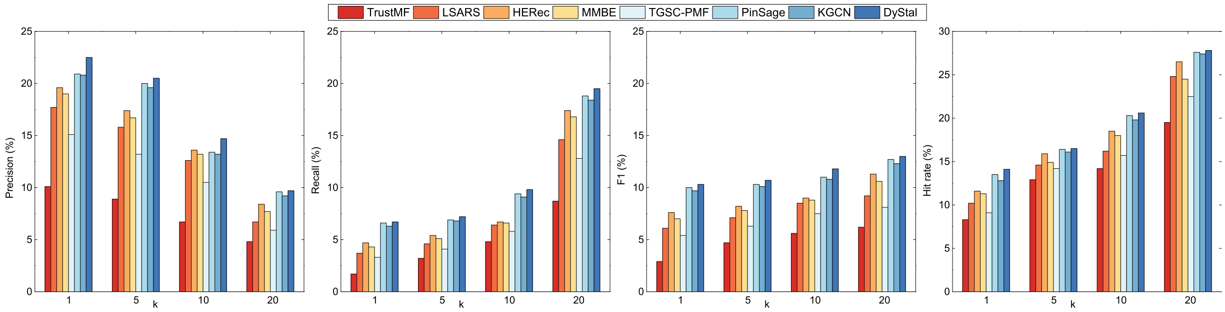

Fig. 6. Performance comparison on the Yelp dataset.

overcomes the problem of data sparsity. In addition, we find that the methods with network embedding including HEREC, MMBE, PINSAGE, KGCN and DYSTAL outperform these methods including TRUSTMF, LSARS and TGSC-PMF without network embedding.

2) It is worth to notice that although MMBE incorporates more factors such as users' preferences, but it still has a worse performance than our DYSTAL method. This can be explained by the reason that a user's static preferences such as the age and the gender contain limited information, much less than the information inferred dynamically from the meta-data such as historical visiting records. Sentiment is one of the most important factors for POI recommendation, which can be discovered from textual content such as reviews. The growth of k leads to the invariant of $\mathcal{V}_u$ and the increasing of $\mathcal{R}_u$. As shown in Eqs. (28) and (29), the precision is declined and the recall is growing accordingly.
3) All methods show better performance in the Foursquare dataset than the Yelp dataset. This is because the review records in the Yelp dataset are more sparse and less than that of the Foursquare dataset.
4) The GCN networks have been an effective model that greatly improves the performance of recommender systems. However, making these methods practical and scalable to web-scale recommendation tasks with sparse items and sparse relationships remains a challenge. Some nodes in the network are not trained effectively in an end-to-end manner due to the limited data. Another important reason is that GCN is a neural network that is unexplainable and cannot help to discover the underlying features and relationships.

7.3. Impact of different factors

In order to evaluate the impact of each factor in the DYSTAL framework, we designed five variant models based on DYSTAL that does not consider the effects of social relationship, geographical relationship, temporal variation and textual sentiments, as shown below.

DYSTAL -social is the first simplified version in which two users are connected with each other in Fig. 3(b). Thus all users have the same social influence that can be omitted.

DYSTAL -geo is the second simplified version that neglects the geographical influence, which implies each user and each POI is located in the same place with the same distance between them in Fig. 3(c) and (d). It also implies that d_{ij}^{t}) and the geographical correlation factor in Fig. 4 is abandoned.

DYSTAL -geo-social is the third simplified version that omits the social and geographical influence simultaneously, which implies that x_{ij}^{t}) and the geo-social correlation factor in Fig. 4 is abandoned.

DYSTAL -time is the fourth simplified version that eliminates the temporal effect, i.e., $y_{ij}^{t'}$ in Fig. 4 is abandoned. We infer the unknown y_{ij}^{t} and perform top-k recommendation in $\mathcal{Y}^{\mathcal{T}}$.

DYSTAL -sentiment is the fifth simplified version that omits the sentimental effect of reviews, i.e., s_{ij}^{t} and the sentimental correlation factor in Fig. 4 is abandoned.

In order to explore the benefits of incorporating social relationship, geographical locations, time and textual sentiments into the DYSTAL model, respectively, we compare DYSTAL with the five variations. The comparison results under $k = 10$ are given in Table 4.

As shown in Table 4, the gap between DYSTAL and any its variants could present the benefit from each factor. DYSTAL outperforms its five variant versions with the gaps ranging from 1.1% to 8.7% on the Foursquare dataset and ranging from 0.5% to 8.3% on the Yelp dataset, respectively. Each factor makes a different contribution to the improvement of the effectiveness of recommendation. According to the significance of these four factors in these five variants, it is interesting to find that their impact are ranked as sentiment > time > social relationship > geographical effect. We can conclude that textual sentiment is more significant in affecting the accuracy of POI recommendation than other factors.

7.4. Sensitivity analysis of parameters

It is important to tune the parameters in order to acquire good performance. In this section, we analyze the sensitivity of the parameters on DYSTAL's performance. The parameters includes: 1) *the local clustering coefficiency* (denoted by cc), which is a measurement of the degree to which nodes in part of a network tend to cluster together; 2) *the time span of the training set* (the parameter $\mathcal{T}$ in Table 2); 3) *the distance threshold* (denoted by d_0); 4) *the number of embedding dimensions* (denoted by D). Fig. 7 shows the hit rate of DYSTAL with different parameters on these two datasets. Each sub-figure denotes the effect of one parameter while other parameters

Table 4
Accuracy comparison with the five variant models.

	Precision	Recall	F1	Hit rate
DYSTAL	16.7	15.2	15.9	23.2
DYSTAL-social	14.8	14.1	14.4	20.1
DYSTAL-geo	15.3	14.7	15.0	21.5
DYSTAL-geo-social	11.1	10.6	10.8	14.4
DYSTAL-time	12.5	11.3	11.9	16.2
DYSTAL-sentiment	9.4	8.8	9.1	14.5
The Yelp dataset (%)				
	Precision	Recall	F1	Hit rate
DYSTAL	14.7	9.8	11.8	20.6
DYSTAL-social	13.1	9.3	10.9	17.0
DYSTAL-geo	13.7	9.4	11.1	17.3
DYSTAL-geo-social	9.7	8.2	8.9	13.8
DYSTAL-time	10.9	8.9	9.8	14.6
DYSTAL-sentiment	8.2	7.6	7.9	12.3

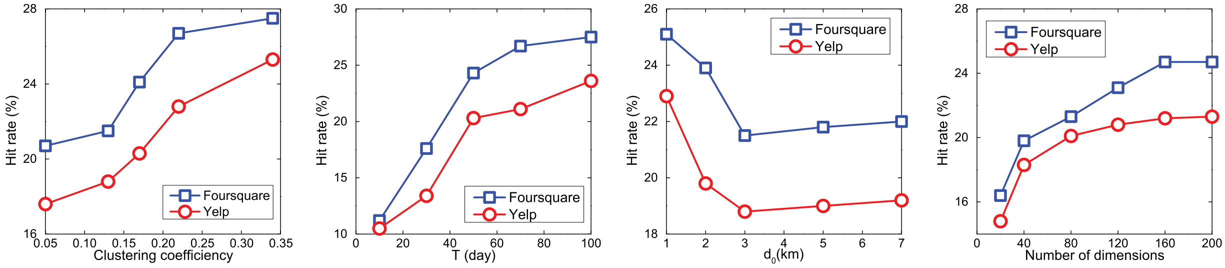

Fig. 7. The hit rate under different parameters.

are fixed, i.e., $cc = 0.13$, $\mathcal{T} = 30$ days, $d_0 = 3$km, $D = 120$.

Beyond these two datasets, we observe the similar results that the hit rate of DYSTAL is sensitive to these four parameters, which shows the different tendencies of the hit rate. 1) In Fig. 7(a), a user can only acquires limited recommendations from a small cluster which seems unreliable. Thus, a smaller value of hit rate corresponds to a smaller clustering coefficiency. 2) In Fig. 7(b), the foursquare dataset requires a longer time of period, which means more data are exploited for recommendation and the hit rate is increased. 3) In Fig. 7(c), the hit rate is decreased and reaches to the lowest value when $d_0 = 3$km, then it increases slowly. The reasons canbe explained by the reason: a user is familiar with the places nearby which has been frequently visited. The decision depends on his/her moods rather than his hobbies. When $d_0 > 3$km, some users appeal to the information in LBSNs, such as rating and reviews, in order to obtain useful recommendations. 4) In Fig. 7(d), the hit rate increases rapidly and then becomes stable. This is because the network embedded into more dimensions contains more information. However, When the number of dimensions is large enough, the embedding vectors contain all information but lead to a high complexity. 5) DYSTAL shows better performance in the Foursquare dataset than in the Yelp dataset, because the Yelp data is more sparse, which will affect the accuracy of recommendation.

To acquire an acceptable trade off between effectiveness and efficiency in the phase of model training, we set $cc = 0.13$, $\mathcal{T} = 30$ days, $d_0 = 3$km and $D = 120$ as default values on these two datasets. It is worthwhile to note that we will make the similar parameter settings in the following experiments.

7.5. Efficiency comparison of different models

In this section, we analyze the computational complexity of DYSTAL and compare it with the baseline methods.

DYSTAL consists of two major phases: the network embedding and the ST-FGM model, whose complexity has been analyzed in Sections 5.3 and 6.4. The entire complexity is $O(|\mathcal{E}| + |\mathcal{U}|^2)$), which can be simplified to $O(|\mathcal{U}|^2)$).

Starting from the above discussion, we further evaluate the efficiency of the proposed model by comparing it with the baseline methods and variant models. In order to observe the results of different networks with similar numbers of users, we extract another dataset containing 76,370 users from the Yelp dataset. The average training time of these approaches are given in Fig. 8.

As shown in Fig. 8(a), most models have comparative training time except PINSAGE and KGCN. This is because the two models are the graph neural networks with billions of parameters to be trained. The joint embedding method in our proposed model can shorten the training time. Furthermore, the original Yelp dataset with 337,084 users contains more links and reviews than the other two datasets, which greatly increase the training time. We also find that LSARS, TGSC-PMF, PINSAGE and DYSTAL spend more time on training because the phase of textual analysis is time-consuming.

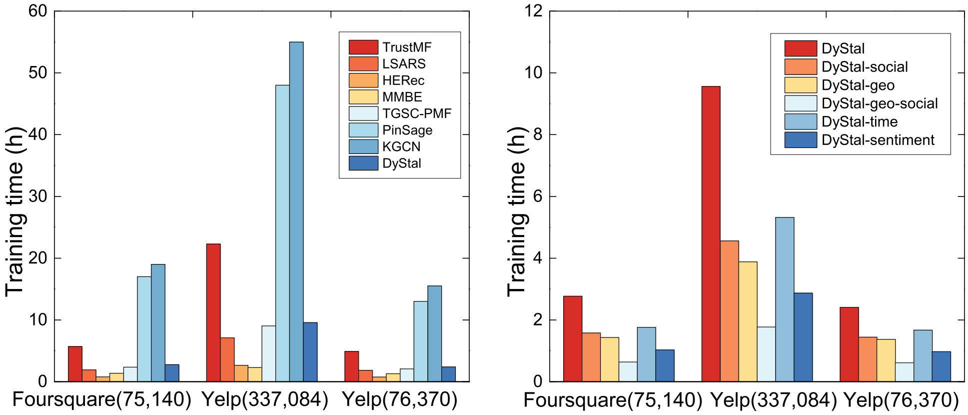

Fig. 8. Training time comparison of different algorithms and variant models.

Fig. 8 (b) compares the training time of different variant models. The factor of sentiment costs the least amount of time when removing it from DYSTAL because textual analysis is time-consuming. The social factor and the geographical factor can affect the phases of network embedding and ST-FGM and have a great impact on the time cost of training.

The cost of training time is mainly related with different parameters. From Fig. 8, we observe the similar training time in the two datasets of which the number of users and the number of POIs are comparable. Fig. 9 shows the change of the training time under two parameters. More edges within a local part of a social network lead to a higher clustering coefficiency, which greatly increases the training time. However, the negative sampling method can reduce the complexity of the algorithm and accelerate the training. In addition, the negative sampling method alleviates the influence of distance threshold, which is used to define the neighbors of a POI.

8. Conclusion

In order to address the problems of identifying dynamic spatio-temporal patterns of visiting behaviors based on unreliable data, we propose a new framework named Dynamic Spatio-temporal POI recommendation (DYSTAL) for recommending POIs to users in an effective manner. The DYSTAL framework is highlighted by two indispensable techniques: (1) a network embedding method learns the vectors of users and POIs in a low-dimensional embedding space; (2) a dynamic factor graph model integrates different factors including the correlation of vectors that obtained in the previous phase. Then for evaluating the effectiveness of DYSTAL, we designed the experiments on two large-scale LBSN datasets. Our model dominates the state-of-the-art baseline approaches in the effectiveness while guaranteeing the desirable runtime performance of POI recommendation.

An inevitable limitation of our model is that the preferences of users and POIs are not fully considered due to the time complexity. Although the preferences have been verified to be significant factors in many previous literatures, we discover in this work the dynamic spatio-temporal patterns and textual sentiments play a more significant role than the preferences. In future work, we would incorporate all these factors into the model and refine the scenarios to further improve the POI recommendation performance.

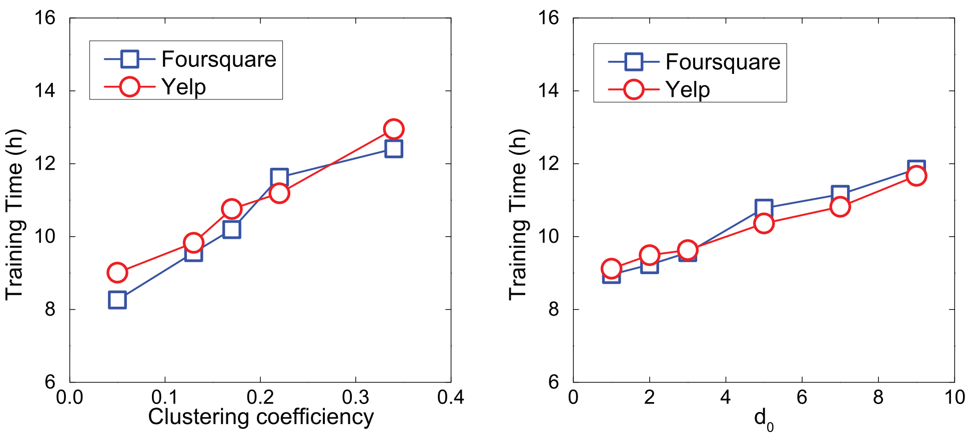

Fig. 9. Training time under different parameters.

CRediT authorship contribution statement

Xi Xiong: Writing - original draft, Methodology, Conceptualization, Methodology, Writing - original draft. **Fei Xiong:** Data curation, Software. **Jun Zhao:** Supervision. **Shaojie Qiao:** Supervision. **Yuanyuan Li:** Visualization, Validation. **Ying Zhao:** Investigation.

Acknowledgments

The work was partially supported by the National Natural Science Foundation of China under Grant nos. 81901389 and 61872033; the National Key R&D Program of China under Grant no. 2018YFC0831605; the China Postdoctoral Science Foundation under Grant no. 2019M653400; the Sichuan Science and Technology Program under Grant no. 2019YFS0236.

References

Hu, B., & Ester, M. (2013). *Spatial topic modeling in online social media for location recommendation. Proceedings of the 7th ACM conference on recommender systems.* ACM25–32.

Hu, L., Li, C., Shi, C., Yang, C., & Shao, C. (2020). Graph neural news recommendation with long-term and short-term interest modeling. *Information Processing and Management, 57*(2), 102142.

Huang, H., Shen, H., & Meng, Z. (2019). Item diversified recommendation based on influence diffusion. *Information Processing and Management, 56*(3), 939–954.

Huang, L., Ma, Y., Liu, Y., & Sangaiah, A. K. (2017). Multi-modal bayesian embedding for point-of-interest recommendation on location-based cyber-physical-social networks. *Future Generation Computer Systems*.

Li, Y., Xiong, X., Qiu, C., Wang, Q., & Xu, J. (2018). Mining latent information in PTSD psychometrics with fuzziness for effective diagnoses. *Scientific Reports, 8*(1), 16266.

Lin, H., Jia, J., Qiu, J., Zhang, Y., Shen, G., Xie, L., ... Chua, T. S. (2017). Detecting stress based on social interactions in social networks. *IEEE Transactions on Knowledge and Data Engineering, PP*(99), 1.

Meng, X., Wang, S., Liu, H., & Zhang, Y. (2018). *Exploiting emotion on reviews for recommender systems. AAAI*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems, 26*, 3111–3119.

Murphy, K. P., Weiss, Y., & Jordan, M. I. (1999). *Loopy belief propagation for approximate inference: an empirical study. Proceedings of the fifteenth conference on uncertainty in artificial intelligence.* 2073849: Morgan Kaufmann Publishers Inc467–475.

Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). Deepwalk:online learning of social representations, (pp. 701–710).

Ren, X., Song, M., Haihong, E., & Song, J. (2017). Context-aware probabilistic matrix factorization modeling for point-of-interest recommendation. *Neurocomputing, 241*(C), 38–55.

Shi, C., Hu, B., Zhao, W. X., & Yu, P. S. (2017). Heterogeneous information network embedding for recommendation. *IEEE Transactions on Knowledge and Data Engineering*.

Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., & Mei, Q. (2015). LINE: Large-scale information network embedding,. 2(2), 1067–1077.

Wang, H., Fu, Y., Wang, Q., Yin, H., Du, C., & Xiong, H. (2017). *A location-sentiment-aware recommender system for both home-town and out-of-town users. Proceedings of the 23rd ACM SIGKDDinternational conference on knowledge discovery and data mining.* ACM1135–1143.

Wang, H., Zhang, F., Wang, J., Zhao, M., Li, W., Xie, X., & Guo, M. (2018). *RippleNet: Propagating user preferences on the knowledge graph for recommender systems. 27th international conference on information and knowledge management*417–426.

Wang, H., Zhao, M., Xie, X., Li, W., & Guo, M. (2019). *Knowledge graph convolutional networks for recommender systems. The world wide web conference*3307??3313.

Wang, X., Jia, J., Tang, J., Wu, B., Cai, L., & Xie, L. (2015). Modeling emotion influence in image social networks. *IEEE Transactions on Affective Computing, 6*(3), 286–297.

Xin, M., & Wu, L. (2020). Using multi-features to partition users for friends recommendation in location based social network. *Information Processing and Management, 57*(1), 102125.

Xiong, X., Li, Y., Qiao, S., Han, N., Wu, Y., Peng, J., & Li, B. (2018). An emotional contagion model for heterogeneous social media with multiple behaviors. *Physica A: Statistical Mechanics and its Applications, 490*, 185–202.

Xiong, X., Li, Y., Zhang, R., Bu, Z., Li, G., & Ju, S. (Li, Zhang, Bu, Li, Ju, 2020a). DGI: Recognition of textual entailment via dynamic gate matching. *Knowledge-Based Systems, 194*, 105544.

Xiong, X., Qiao, S., Han, N., Xiong, F., Bu, Z., Li, R.-H., ... Yuan, G. (Qiao, Han, Xiong, Bu, Li, Yue, Yuan, 2020b). Where to go: An effective point-of-interest recommendation framework for heterogeneous social networks. *Neurocomputing, 373*, 56–69.

Xiong, X., Qiao, S., Li, Y., Han, N., Yuan, G., & Zhang, Y. (Qiao, Li, Han, Yuan, Zhang, 2020c). A point-of-interest suggestion algorithm in multi-source geo-social networks. *Engineering Applications of Artificial Intelligence, 88*, 103374.

Xiong, X., Qiao, S., Li, Y., Xiong, F., He, L., & Han, N. (Qiao, Li, Xiong, He, Han, 2019a). Affective impression: Sentiment-awareness POI suggestion via embedding in heterogeneous lbsns. *IEEE Transactions on Affective Computing,* 1.

Xiong, X., Qiao, S., Li, Y., Zhang, H., Huang, P., Han, N., & Li, R.-H. (Qiao, Li, Zhang, Huang, Han, Li, 2019b). ADPDF: A hybrid attribute discrimination method for psychometric data with fuzziness. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49*(1), 265–278.

Yang, B., Lei, Y., Liu, J., & Li, W. (2017). Social collaborative filtering by trust. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(8), 1633–1647.

Yao, Z. (2018). *Exploiting human mobility patterns for point-of-interest recommendation. ACM international conference on web search and data mining.* ACM757–758.

Ye, M., Yin, P., Lee, W.-C., & Lee, D.-L. (2011). *Exploiting geographical influence for collaborative point-of-interest recommendation. Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval.* ACM325–334.

Yin, H., Zhou, X., Shao, Y., Wang, H., & Sadiq, S. (2016). Joint modeling of user check-in behaviors for point-of-interest recommendation. *ACM Transactions on Information Systems, 35*(2).

Ying, H., Wu, J., Xu, G., Liu, Y., Liang, T., Zhang, X., & Xiong, H. (2019). Time-aware metric embedding with asymmetric projection for successive POI re-commendation. *World Wide Web, 22*(5), 2209–2224.

Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W. L., & Leskovec, J. (2018). *Graph convolutional neural networks for web-scale recommender systems. Proceedings of the 24th ACM SIGKDDinternational conference on knowledge discovery and data mining.* ACM974–983.

Zhao, P., Xu, X., Liu, Y., Zhou, Z., Zheng, K., Sheng, V. S., & Xiong, H. (2017). *Exploiting hierarchical structures for POI recommendation. 2017 IEEE international conference on data mining.* IEEE655–664.

Zhao, P., Zhu, H., Liu, Y., Xu, J., Li, Z., Zhuang, F., ... Zhou, X. (2019). *Where to go next: A spatio-temporal gated network for next POI recommendation. Proceedings of the AAAI conference on artificial intelligence33. Proceedings of the AAAI conference on artificial intelligence* 5877–5884.