Augmented Collaborative Filtering for Sparseness Reduction in Personalized POI Recommendation

CHAORAN CUI, Shandong University of Finance and Economics
JIALIE SHEN, Northumbria University
LIQIANG NIE, Shandong University
RICHANG HONG, Hefei University of Technology
JUN MA, Shandong University

As mobile device penetration increases, it has become pervasive for images to be associated with locations in the form of geotags. Geotags bridge the gap between the physical world and the cyberspace, giving rise to new opportunities to extract further insights into user preferences and behaviors. In this article, we aim to exploit geotagged photos from online photo-sharing sites for the purpose of personalized Point-of-Interest (POI) recommendation. Owing to the fact that most users have only very limited travel experiences, data sparseness poses a formidable challenge to personalized POI recommendation. To alleviate data sparseness, we propose to augment current collaborative filtering algorithms along from multiple perspectives. Specifically, hybrid preference cues comprising user-uploaded and user-favored photos are harvested to study users' tastes. Moreover, heterogeneous high-order relationship information is jointly captured from user social networks and POI multimodal contents with hypergraph models. We also build upon the matrix factorization algorithm to integrate the disparate sources of preference and relationship information, and apply our approach to directly optimize user preference rankings. Extensive experiments on a large and publicly accessible dataset well verified the potential of our approach for addressing data sparseness and offering quality recommendations to users, especially for those who have only limited travel experiences.

CCS Concepts: • **Human-centered computing** → **Social media**; • **Information systems** → **Location based services**; **Multimedia information systems;**

Additional Key Words and Phrases: Geotagged photos, hypergraph-based learning, personalized POI recommendation

ACM Reference Format:
Chaoran Cui, Jialie Shen, Liqiang Nie, Richang Hong, and Jun Ma. 2017. Augmented collaborative filtering for sparseness reduction in personalized POI recommendation. ACM Trans. Intell. Syst. Technol. 8, 5, Article 71 (September 2017), 23 pages.
DOI: http://dx.doi.org/10.1145/3086635

This work is supported by the National Natural Science Foundation of China (61671274), China Postdoctoral Science Foundation (2016M592190), and the Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions.
Authors' addresses: C. Cui, School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China; email: crcui@sdufe.edu.cn; J. Shen (corresponding author), Department of Computer and Information Sciences, Northumbria University, Newcastle NE2 1XE, UK; email: jerry.shen@northumbria.ac.uk; R. Hong, Department of Electric Engineering and Information System, Hefei University of Technology, Hefei 230009, China; email: hongrc.hfut@gmail.com; L. Nie and J. Ma, School of Computer Science and Technology, Shandong University, Jinan 250101, China; emails: nieliqiang@gmail.com, majun@sdu.edu.cn.

DOI: http://dx.doi.org/10.1145/3086635

1. INTRODUCTION

Rapid advances in mobile and social media technologies now allow people to interact, create, and share images based on physical location via various photo-sharing platforms such as Flickr.[1] The increasing availability of geotagged photos opens up new opportunities in a wide variety of location-oriented multimedia research and applications, including georeferenced image search [Zhu et al. 2015a, 2016; Cheng and Shen 2016], location estimation [Li et al. 2015b], as well as scene visualization [Rudinac et al. 2013].

As one of the most important digital footprints, user-uploaded geotagged photos record people's actual presences and their activities at specific locations in the physical world [Jiang et al. 2015]. With this insight, recent research efforts [Shen et al. 2014, 2016] have been dedicated to providing online travel guidance services such as personalized Point-of-Interest (POI) recommendation based upon the geotagged photos. In real life, users are often overwhelmed by the countless POIs when arriving in a new city. To minimize the matching gap, personalized POI recommendation works toward routing the most appropriate POIs to the right persons by taking their personal preferences encoded in their historical photos into account.

Recently, the Collaborative Filtering (CF) [Shi et al. 2014] approaches have been proposed and applied to leverage geotagged photos to support effective personalized POI recommendation, with well theoretical underpinnings and great practical success. The philosophy behind CF algorithms is that users who have similar tastes in the past will have common preferences in the future [Huang et al. 2015]. Roughly speaking, prior efforts can be divided into two well-accepted categories: memory-based method and model-based method. The former [Clements et al. 2010; Jiang et al. 2015] estimates the preference of a target user by aggregating other users having similar interests. The latter [Shi et al. 2013; Phan et al. 2014] learns the latent factors of users and POIs, which are further used to predict new preference scores. Despite the encouraging results reported, the performance of these CF algorithms is still far from satisfactory. Such stagnation is mainly caused by the data sparseness problem in personalized POI recommendation. Distinguished from other consumptions like movie watching, traveling can be generally more expensive and requires the availability of relatively large amounts of time and money [Shi et al. 2013]. Taking these factors into account, most users may only select and visit a limited number of POIs. According to our statistics, the density of user-POI visiting matrix based on empirical study is 0.3%, which is much smaller than 1.2% for the traditional movie recommendation dataset.[2] In light of this, memory-based methods are unable to accurately find close users with few visited POIs in common, while current model-based methods may fail to discover reliable latent factors using such insufficient preference data.

In this study, Flickr is leveraged as a rich information source about locations, behaviors, and relationships, from which highly effective personalized POI recommender systems could be developed. Our study mainly serves to provide POI recommendations for the Flickr users, whose personal preferences can be partially sensed from the associated geotagged photos and potential social interactions. Specifically, in order to alleviate data sparseness as aforementioned, we enhance the conventional CF models via jointly taking the following perspectives into account.

Firstly, apart from the cues encoded in user-uploaded photos, we believe that the implicit feedback conveyed by user-favored photos is also beneficial to personalized

[1] http://www.flickr.com/.
[2] http://www.netflixprize.com/.

POI recommendation. Intuitively, bookmarking a photo of a POI as favorite expresses users' preferences on that POI to some extent [Lipczak et al. 2013]. It is thus a natural idea to fuse uploaded and favored photos for better understanding of user interests. However, we argue that such a straightforward fusion strategy may be problematic. The key restriction lies in the discrepant credibility for user uploading and favoring behaviors to indicate user preferences. As mentioned previously, uploading photos for a POI implies the actual travel experiences to the POI, and is usually triggered by the feeling that it is interesting and worth sharing [Kurashima et al. 2010]. In contrast, the motivation of favoring behaviors is rather complicated, spanning from showing the interest in the photo contents, strengthening the friendship with the owner of the photo, to boosting the popularity of the user's own photos [van Zwol et al. 2010]. As a result, favoring behaviors can only be regarded as a signal reflecting users' true preferences with some uncertainty. Toward this end, instead of early fusion, we separately explore user-uploaded and user-favored photos, and then seamlessly sew them up in a principled way.

Secondly, we adapt the CF approach to personalized POI recommendation with a ranking-based matrix factorization technique. Due to the lack of explicit preference ratings on POIs, we leverage the number of a user's uploaded or favored photos for a POI as a proxy to quantify his/her preference degree on that POI. The underlying assumption is that a high uploading or favoring frequency can be interpreted in a way that a user repeatedly confirms his/her preference [Wang et al. 2014; Shi et al. 2013]. Several prior efforts focus on approximating the observed preference scores [Shi et al. 2013; Phan et al. 2014]. In practice, however, users do not care about the specific scores, but about a compact recommendation list of POIs. Inspired by this, we propose to generate the relative ranking of users' preferences on POIs. The matrix factorization technique is applied to learn latent factors of users and POIs from a large number of pairwise preference relations with varying degrees of importance, which are induced via the comparison of two POIs. It is worth noting that, in our approach, both POIs with and without observed preference mutually reinforce the learning process, which helps further reduce the effects of data sparseness.

Lastly, most of the existing methods treat different users and POIs in isolation and ignore their underlying relationships. Most recently, a few studies [Ma 2013] have exploited user social networks together with CF approaches to boost the recommendation accuracy. Although great success has been achieved, these studies only consider the pairwise user similarity and cannot characterize the inherent high-order grouping relations among users. For example, different users usually join the same group, and multiple users commonly follow another contact person. On the other hand, POIs are linked by diverse high-order relationships as well, which reflect their commonalities in certain visual or textual content. For example, various POIs may share close scenery or be associated with similar semantic tags [Zhu et al. 2015b]. In this study, we propose a bi-relational hypergraph representation model. It is able to co-regularize two main hypergraphs by jointly capturing the high-order user relationships and POI relationships. A hypergraph is a generalization of a simple graph, where the edges, called *hyperedges*, are arbitrary non-empty subsets of the vertex set and used to model high-order relations of vertices [Bu et al. 2010]. Smoothness of latent factors on the bi-relational hypergraph representation is leveraged to constrain the matrix factorization algorithm. In this way, heterogeneous high-order relationships can be integrated into our framework simultaneously, serving to alleviate data sparseness and leading to more effective latent factors.

Our study advances several streams of research, and the main contributions can be summarized as follows:

(1) We harvest a set of rich cues including photo uploading records, user favoring behaviors, as well as preference degrees to enhance the personalized POI recommendation performance. The unique nature of the cues allows us to gain a comprehensive understanding of user preferences on various POIs. To the best of our knowledge, we are the first authors to provide a set of such cues.
(2) Our article also contributes literature on georeference social network modeling. We effectively capture and unpack the heterogeneous high-order relationships embodied in user social networks and POI multimodal contents with the hypergraph structures.
(3) Distinguished from the previous approaches, we integrate the disparate sources of preference and relationship information under the matrix factorization framework, and develop our approach toward effective user preference ranking optimization.

The remainder of the article is structured as follows. Section 2 reviews the related work. In Section 3, we introduce the process of POI discovery from geotagged photos. Sections 4 and 5 give a detailed description of our proposed model and the experimental settings, respectively. Experimental results and analysis are reported in Section 6, followed by the conclusion and future work in Section 7.

2. RELATED WORK

In this section, we review the existing literature on POI recommendation that exploits location information in social media networks.

2.1. GPS-based Recommendation

GPS trajectories record the sequences of user-visited POIs and the duration of stay at each POI. Recent work has leveraged this explicit location information for POI recommendation. In one of the earliest efforts, Zheng et al. [2009] adapted the hypertext induced topic search model to recommend popular POIs and classical travel itineraries within a given geospatial region. By extending this work, they further developed personalized recommendation algorithms through mining the correlations between POIs [Zheng and Xie 2011]. In Zheng et al. [2011], the authors proposed to simultaneously recommend potential friends and POIs for individual users. Similarly, Zheng et al. [2010] mined knowledge from GPS trajectories to make joint recommendations of POIs and activities. The major concern regarding the use of GPS trajectory data for recommendation is privacy. Users are unwilling to accept continuous monitoring of their activity trails, which may contain private location information they do not want to disclose.

2.2. Check-In-based Recommendation

Location-Based Social Networks (LBSNs) [Bao et al. 2015; Wang et al. 2015b] have recently become increasingly popular, such as Foursquare.[3] In LBSNs, users check in and share their experiences about POIs with friends. By leveraging check-in frequencies as ratings on POIs, conventional CF algorithms were directly implemented for POI recommendation in Ye et al. [2011] and Cheng et al. [2012]. However, users usually check in at most POIs only once, so it may not be effective to extract user preferences merely with check-in behaviors [Zhang et al. 2015]. In Gao et al. [2012] and Liu et al. [2013], it was assumed that users tend to visit nearby POIs and the probability of visiting a new place decreases as the distance increases. Such geographical influences on check-in behaviors were further exploited to enhance the accuracy of POI recommender systems. Similarly, Lian et al. [2014] discussed the spatial clustering phenomenon in

[3]http://foursquare.com/.

human behaviors in LBSNs, and incorporated it into the matrix factorization model. Wang et al. [2015a] proposed a hybrid predictive model integrating both the regularity and conformity of human mobility, and further elevated the model by learning location profiles from heterogeneous mobility datasets. Based on the premise that users are likely to visit different POIs at different times, several researchers [Yuan et al. 2013; Li et al. 2015a] introduced temporal influences and proposed a series of time-aware POI recommendation algorithms. Additionally, many research efforts [Liu et al. 2016; Cheng et al. 2013; Feng et al. 2015] have been devoted to the task of successive POI recommendation, which aims to predict the next location based on users' past sequential check-in behaviors. It is worth noting that, unlike geotagged photos bearing user trip patterns, check-in data often reflect native residents' preferences on local venues, e.g., restaurants and shops [Liu et al. 2012].

2.3. Geotag-based Recommendation

There has been an emerging research interest in POI recommendation by leveraging information extracted from geotags in photo-sharing sites. In Gao et al. [2010], a travel guidance system was designed to rank POIs using their contents and user-contributed tags. Ji et al. [2011] proposed to incorporate the supervision from both the user confidence and their tagging confidence to refine the popularity ranking of POIs within a given city. Note that the above methods are non-personalized and generate the same recommendation results regardless of the target user. For personalized POI recommendation, Clements et al. [2010] made an initial effort to rerank popular POIs for a user based on preferences of other users with similar interests. Jiang et al. [2015] introduced the Author Topic Model (ATM) to discover user preference topics from photo tags, and measured the similarity between users using their topic distributions. Shi et al. [2013] combined weighted matrix factorization and category-based regularization for non-trivial POI recommendation. Empirical results showed that the proposed method was more appropriate to recommend POIs with infrequent visits. Phan et al. [2014] established the user-POI preference matrix according to photo-taking patterns, and investigated the performance of different matrix factorization models for personalized POI recommendation. In Guo et al. [2014], different relations among users and POIs were embedded into a multipartite graph, and a random walk with restart was performed over the graph to estimate preference scores. Compared to the previous work, our approach specifically addresses the data sparseness problem in the context of personalized POI recommendation via augmenting CF algorithms from different aspects.

3. POI DISCOVERY

As a prerequisite to realize personalized POI recommendation from geotagged photo collections, we need to first discover the geographical locations of popular POIs in a city. As interpreted by the *geographical concentration* assumption [Ji et al. 2011] that most photos are concentrated on several famous POI regions of a given city, this problem can be transformed to a clustering problem of partitioning two-dimensional points. Each point corresponds to a geotagged photo represented by its geographical coordinate. Among various clustering algorithms, mean shift clustering has attracted increasing attention for this problem [Chen et al. 2013], since it neither requires the prior knowledge of the number of clusters, nor constrains the shape of clusters. Therefore, we chose mean shift clustering to discover the geographical locations of POIs.

The basic principle of mean shift clustering is to automatically estimate the modes of an underlying probability distribution over places where people take photos, with a set of observed geotagged photos. Given a kernel function and a bandwidth parameter, the algorithm iteratively shifts each point toward the location corresponding to a mode of the underlying distribution. All points that converge to the same mode constitute a

cluster, and each cluster is considered as an individual POI with high photo density. In our implementation, inspired by the observation that geotagged photos are likely Gaussian distributed in surrounding areas of POIs [Yang et al. 2014], we adopted Gaussian density function as the kernel function. The bandwidth parameter was set to 0.001°, which is roughly the radius of a POI, as mentioned in Chen et al. [2013]. In practice, mean shift clustering can be performed very fast, especially on low-dimensional data. After the clustering process, we filtered out those unpopular POIs by discarding the clusters in which the photos were only taken by a small number of distinct users. Additionally, we computed the standard deviation of the dates on which the photos were taken in a cluster. To avoid the clusters related to events (e.g., festivals) rather than POIs, the clusters with low variability in dates were further eliminated. Finally, we collected the remaining clusters to form the set of discovered POIs.

Note that we did not match the discovered POIs against the places listed in travel guide websites such as Wikitravel[4] and TripAdvisor.[5] The reason is twofold: (1) There are only relatively few places in travel guidance websites, which are subjectively selected by professional editors and may be unable to cater to the diverse needs from real users [Gao et al. 2010]. In contrast, the POIs discovered from geotagged photos represent the hotspots visited during users' actual travel experiences. (2) Experiments in previous work [Crandall et al. 2009] have shown that the POIs discovered via the clustering process correspond to popular landmarks in reality with high accuracy.

4. RECOMMENDATION SCHEME

To formulate our problem, we declare some notations in advance. In particular, we use bold capital letters (e.g., $\mathbf{X}$), bold lowercase letters (e.g., $\mathbf{x}$), and calligraphic capital letters (e.g., $\mathcal{X}$) to denote matrices, vectors, and sets, respectively. We employ non-bold letters (e.g., x) to represent scalars, and Greek letters (e.g., λ) as parameters. If not clarified, all vectors are in column form. Table I summarizes the key notations and definitions used throughout the article.

Our scheme is developed on the foundation of matrix factorization, which is the most successful and widely used recommendation technique due to its advantage in terms of scalability and accuracy [Huang et al. 2015]. Given a collection of users $\mathcal{I}$ and the set of discovered POIs $\mathcal{P}$, the goal is to learn two low-rank matrices $\mathbf{U} \in \mathbb{R}^{d \times n}$ and $\mathbf{V} \in \mathbb{R}^{d \times m}$, with column vectors $\mathbf{u}_u$ and $\mathbf{v}_i$ representing d-dimensional user-specific and POI-specific latent factors of user $u \in \mathcal{I}$ and POI $i \in \mathcal{P}$, respectively. By this means, u and i are simultaneously mapped into a common latent space, and their matching degree can be efficiently calculated by

$$r_{ui} = \mathbf{u}_u^T \mathbf{v}_i + b_i. \tag{1}$$

Here, b_i is a bias term for POI i. Indeed, some POIs are very popular and thus have a high expectation of being preferred by a broad spectrum of users, while other POIs may be less popular and only cater to niche groups. The bias terms associated with POIs are used to take into account these popularity differences.

The architecture of our scheme is illustrated in Figure 1. It consists of three main components: (1) pairwise preference mining, (2) bi-relational hypergraph representation, and (3) ranking-based matrix factorization. In particular, a large number of pairwise preference relations are first mined from the hybrid set of user-uploaded and user-favored photos. After that, a bi-relational hypergraph representation is constructed to jointly capture multiple types of high-order relationships among users and

[4]http://www.wikitravel.org/.
[5]http://www.tripadvisor.com/.

Table I. Summary of Key Notations and Definitions

Notation	Definition
$\mathcal{I}, \mathcal{P}$	User set and POI set
u, v, i, j	User id $u, v \in \mathcal{I}$ and POI id $i, j \in \mathcal{P}$
n, m	Number of users and POIs, i.e., $n = \lvert\mathcal{I}\rvert$ and $m = \lvert\mathcal{P}\rvert$
d	Dimension of latent space
$\mathbf{U}, \mathbf{V}$	Latent factor matrices for users and POIs
$\mathbf{u}_u, \mathbf{v}_i$	Latent factors of u and i
r_{ui}, r_{uj}, r_{ij}^u	Matching degrees between u and i, as well as between u and j, and $r_{ij}^u = r_{ui} - r_{uj}$
$\mathbf{Z}, \mathbf{F}$	User-POI preference matrices built with user-uploaded and user-favored photos
$\mathcal{D}_Z, \mathcal{D}_F$	Sets of pairwise preferences based on $\mathbf{Z}$ and $\mathbf{F}$
c_{ij}^u	Weight assigned to triple (u, i, j)
$\mathcal{G}_\mathcal{I} = (\mathcal{I}, \mathcal{E}_\mathcal{I}, \mathbf{W}_\mathcal{I})$	User hypergraph, where $\mathcal{E}_\mathcal{I}$ denotes the set of user hyperedges, and $\mathbf{W}_\mathcal{I}$ is the weight matrix for user hyperedges
$\mathbf{H}_\mathcal{I}, \mathbf{D}_\mathcal{I}, \mathbf{D}_{\mathcal{E}_\mathcal{I}}, \mathbf{L}_\mathcal{I}$	Auxiliary matrices of $\mathcal{G}_\mathcal{I}$, where $\mathbf{H}_\mathcal{I}$ is the incidence matrix, $\mathbf{D}_\mathcal{I}$ and $\mathbf{D}_{\mathcal{E}_\mathcal{I}}$ are the degree matrices for user vertices and user hyperedges, and $\mathbf{L}_\mathcal{I}$ is the normalized Laplacian
$\mathcal{G}_\mathcal{P} = (\mathcal{P}, \mathcal{E}_\mathcal{P}, \mathbf{W}_\mathcal{P})$	POI hypergraph where $\mathcal{E}_\mathcal{P}$ denotes the set of POI hyperedges, and $\mathbf{W}_\mathcal{P}$ is the weight matrix for POI hyperedges
$\mathbf{H}_\mathcal{P}, \mathbf{D}_\mathcal{P}, \mathbf{D}_{\mathcal{E}_\mathcal{P}}, \mathbf{L}_\mathcal{P}$	Auxiliary matrices of $\mathcal{G}_\mathcal{P}$, where $\mathbf{H}_\mathcal{P}$ is the incidence matrix, $\mathbf{D}_\mathcal{P}$ and $\mathbf{D}_{\mathcal{E}_\mathcal{P}}$ are the degree matrices for POI vertices and POI hyperedges, and $\mathbf{L}_\mathcal{P}$ is the normalized Laplacian
e	Hyperedge index in $\mathcal{E}_\mathcal{I}$ or $\mathcal{E}_\mathcal{P}$
k	Number of nearest neighbors considered when constructing POI hypergraph
$\lambda, \alpha, \mu, \beta, \gamma$	Model hyperparameters

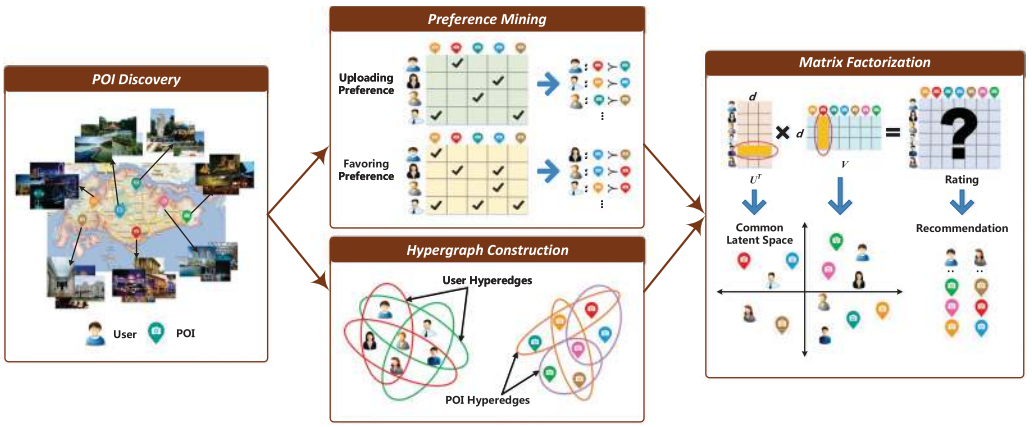

Fig. 1. Schematic illustration of the proposed POI recommendation approach. Best viewed in color.

POIs. Lastly, the matrix factorization algorithm is deployed to exploit different information and optimize relative preference rankings. In the following, we elaborate on each of the components and give a full description of the associated algorithms.

4.1. Pairwise Preference Mining

There are no explicit ratings available in our setting; we thus first turn to extract user preferences from the implicit feedback encoded in their uploaded geotagged photos. A user-POI matrix $\mathbf{Z} = (z_{ui})_{n \times m}$ is collected with its entry z_{ui} representing the preference score of user $u \in \mathcal{I}$ on POI $i \in \mathcal{P}$. To calculate z_{ui}, we count up the number of photos uploaded by u that have geotags matching i. The count value is further normalized to be between 0 and 1 by dividing it by the total number of u's uploaded geotagged photos.

Note that the matrix $\mathbf{Z}$ is sparse due to the lack of uploading interactions between many users and POIs, and each missing element z_{ui} is assigned 0 indicating that u's preference on i has not been observed.

Bookmarking media as favorites is a pervasive behavior in social platforms. It has recently attracted increasing attention because of its potential in reflecting user interests [Lipczak et al. 2013; Cui et al. 2014]. Therefore, we further explore the possibility of extracting user preferences based on their favoring behaviors. To the best of our knowledge, this is the first attempt to leverage the cues encoded in user-favored photos for personalized POI recommendation. Analogous to the procedures described above, we denote an additional user-POI preference matrix by $\mathbf{F} = (f_{ui})_{n\times m}$. The entry f_{ui} represents the proportion of photos favored by u that are related to POI i, and $f_{ui} = 0$ when no user favoring patterns are observed from u for i. Distinguished from the binary implicit feedback in conventional recommender systems [Pan et al. 2015], the user-POI matrices $\mathbf{Z}$ and $\mathbf{F}$ convey richer graded information to discriminate different degrees of user preferences.

Instead of fitting the observed preference scores as previous recommendation methods [Shi et al. 2013; Phan et al. 2014], we model user preference rankings for POIs. For this purpose, pairwise preference relations over two POIs are extracted to reflect the partial order information of user preferences. Formally, a set of pairwise preference relations $\mathcal{D}_Z : \mathcal{I} \times \mathcal{P} \times \mathcal{P}$ is created from the preference matrix $\mathbf{Z}$ by

$$\mathcal{D}_Z = \{(u, i, j) \mid z_{ui} > z_{uj}\}, \tag{2}$$

where each triple (u, i, j) expresses the deduction that user u prefers POI i over POI j. Likewise, the other set of pairwise preference relations $\mathcal{D}_F : \mathcal{I} \times \mathcal{P} \times \mathcal{P}$ can be obtained from the preference matrix $\mathbf{F}$ by

$$\mathcal{D}_F = \{(u, i, j) \mid f_{ui} > f_{uj} \wedge z_{uj} = 0\}. \tag{3}$$

One thing worth noting is that each pairwise preference $(u, i, j) \in \mathcal{D}_F$ is associated with the constraint of $z_{uj} = 0$. For a POI receiving more favoring interactions, we only draw the conclusion that it is preferable to those POIs without user-uploaded photos. As aforementioned, this is due to the fact that favoring behaviors only have relatively lower credibility to signal users' true preferences. In addition, it has been shown that pairwise preference relations have different importance depending on how strong they are [Wang et al. 2014]. For example, assume that the derived preference scores of user u regarding POIs $\{i, j, q\}$ are $\{5, 1, 4\}$, respectively, which reflects the fact that u prefers i over j much more strongly than i over q. Intuitively, pairwise preferences with larger score differences are more credible. Therefore, we further associate each pairwise preference (u, i, j) with a weight c^u_{ij}, which is defined as

$$c^u_{ij} = \begin{cases} \lambda_Z(z_{ui} - z_{uj}), & \text{if } (u, i, j) \in \mathcal{D}_Z; \\ \lambda_F(f_{ui} - f_{uj}), & \text{if } (u, i, j) \in \mathcal{D}_F, \end{cases} \tag{4}$$

where λ_Z and λ_F are the scaling parameters.

4.2. Bi-relational Hypergraph Representation

Relationship information has been demonstrated to be valuable for improving recommendations by multiple research contributions [Ma 2013; Huang et al. 2015]. In this study, a bi-relational hypergraph representation is proposed, which comprises two hypergraphs capturing the high-order user relationships and POI relationships, respectively. In recent literature [Gao et al. 2013, 2015], hypergraph has proven to be a highly effective tool in modeling high-order relationships. A hypergraph is an extension of a simple graph, where a set of vertices are defined as a weighted hyperedge. Let

$\mathcal{G}_{\mathcal{I}} = (\mathcal{I}, \mathcal{E}_{\mathcal{I}}, \mathbf{W}_{\mathcal{I}})$ be the user hypergraph, where $\mathcal{I}$ is the vertex set, $\mathcal{E}_{\mathcal{I}}$ denotes the set of hyperedges connecting different user vertices, and $\mathbf{W}_{\mathcal{I}}$ is a diagonal matrix with the element $w_{\mathcal{I}}(e)$ denoting the weight for hyperedge $e \in \mathcal{E}_{\mathcal{I}}$. The hypergraph $\mathcal{G}_{\mathcal{I}}$ can also be represented by a $|\mathcal{I}| \times |\mathcal{E}_{\mathcal{I}}|$ incidence matrix $\mathbf{H}_{\mathcal{I}}$ with entries as

$$h_{\mathcal{I}}(v, e) = \begin{cases} 1, & \text{if } v \in e; \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

$\mathbf{D}_{\mathcal{I}}$ is the diagonal matrix of vertex degrees. Its element $d_{\mathcal{I}}(v)$ represents the degree of vertex $v \in \mathcal{I}$:

$$d_{\mathcal{I}}(v) = \sum_{e \in \mathcal{E}_{\mathcal{I}}} w_{\mathcal{I}}(e) h_{\mathcal{I}}(v, e). \tag{6}$$

$\mathbf{D}_{\mathcal{E}_{\mathcal{I}}}$ is the diagonal matrix of hyperedge degrees and the element $d_{\mathcal{E}_{\mathcal{I}}}(e)$ indicates the degree of hyperedge $e \in \mathcal{E}_{\mathcal{I}}$:

$$d_{\mathcal{E}_{\mathcal{I}}}(e) = \sum_{v \in \mathcal{I}} h_{\mathcal{I}}(v, e). \tag{7}$$

Similar to the notations introduced above, we use $\mathcal{G}_{\mathcal{P}} = (\mathcal{P}, \mathcal{E}_{\mathcal{P}}, \mathbf{W}_{\mathcal{P}})$ to denote the POI hypergraph, which is accompanied with the auxiliary matrices $\mathbf{H}_{\mathcal{P}}$, $\mathbf{D}_{\mathcal{P}}$, and $\mathbf{D}_{\mathcal{E}_{\mathcal{P}}}$.

In the bi-relational hypergraph representation, the hyperedges in $\mathcal{E}_{\mathcal{I}}$ and $\mathcal{E}_{\mathcal{P}}$ correspond to the heterogeneous relationships among $\mathcal{I}$ and $\mathcal{P}$, respectively. Specifically, the construction of hyperedges is performed as follows:

—*User Hyperedges*. The user hyperedges are constructed to connect different users with their social interactions. We consider two types of user hyperedges by utilizing the grouping memberships and following relationships. In social media sites, users often participate in groups to share common interests. For each group, we build a hyperedge that consists of all vertices corresponding to the users in this group. Additionally, it is generally accepted that the followers of a person would probably have similar tastes. Therefore, we establish a hyperedge containing the users who commonly follow a target person. The weight of a user hyperedge is defined as the average similarity over all pairs of vertices linked with the hyperedge, which is in line with the intuition that a high weight should be assigned to a hyperedge if the users within it are close to each other [Fang et al. 2014]. Similar to previous work [Cui et al. 2014], for two users $u \in \mathcal{I}$ and $v \in \mathcal{I}$, we estimate their similarity by measuring how many groups and persons are co-joined or co-followed by u and v:

$$s(u, v) = \eta \frac{|\mathcal{C}_u^g \cap \mathcal{C}_v^g|}{|\mathcal{C}_u^g \cup \mathcal{C}_v^g|} + (1 - \eta) \frac{|\mathcal{C}_u^c \cap \mathcal{C}_v^c|}{|\mathcal{C}_u^c \cup \mathcal{C}_v^c|}, \tag{8}$$

where $\mathcal{C}_u^g$ is the set of groups u joined, and $\mathcal{C}_u^c$ is the set of persons u followed. η is a parameter adjusting the relative importance of the two factors.

—*POI Hyperedges*. The POI hyperedges are used to model the content relationships among different POIs. We describe POI contents in both visual and textual modalities. Each POI $i \in \mathcal{P}$ is represented by a two-dimensional tuple $[\mathcal{O}_i, \mathbf{t}_i]$, where $\mathcal{O}_i$ is the collection of geotagged photos matching i, and $\mathbf{t}_i$ is a bag-of-word vector generated by aggregating the tags associated with the photos in $\mathcal{O}_i$. $\mathbf{t}_i$ is also weighted using the TF-IDF scheme. The visual similarity between $i \in \mathcal{P}$ and $j \in \mathcal{P}$ is computed as follows:

$$a_v(i, j) = \frac{1}{|\mathcal{O}_i|\,|\mathcal{O}_j|} \sum_{\mathbf{x}_i \in \mathcal{O}_i} \sum_{\mathbf{x}_j \in \mathcal{O}_j} \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\tau^2}\right), \tag{9}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the visual feature vectors of two images belonging to $\mathcal{O}_i$ and $\mathcal{O}_j$, respectively. τ is a deviation parameter set as the average of all pairwise distances between images. Furthermore, we measure the textual similarity between i and j based on the cosine similarity of their bag-of-word vectors:

$$a_t(i,j) = \frac{\mathbf{t}_i \cdot \mathbf{t}_j}{\|\mathbf{t}_i\| \left\|\mathbf{t}_j\right\|}. \tag{10}$$

In each modality, we take each POI as a centroid vertex and form a hyperedge to connect itself and its k-nearest neighbors. For example, POI i serves as the centroid vertex for the hyperedges $e_i^v \in \mathcal{E}_\mathcal{P}$ and $e_i^t \in \mathcal{E}_\mathcal{P}$ in visual and textual modalities, respectively. Their weights are defined as

$$\begin{aligned} w_\mathcal{P}(e_i^v) &= \sum_{j \in e_i^v \setminus i} a_v(i,j), \\ w_\mathcal{P}(e_i^t) &= \sum_{j \in e_i^t \setminus i} a_t(i,j). \end{aligned} \tag{11}$$

Accordingly, the size of each hyperedge is $k+1$, and the total number of POI hyperedges is $2m$.

4.3. Ranking-based Matrix Factorization

With the preference information conveyed by $\mathcal{D}_Z$ and $\mathcal{D}_F$ as well as the relationship information encoded in $\mathcal{G}_\mathcal{I}$ and $\mathcal{G}_\mathcal{P}$, we now develop a matrix factorization model to combine the two information resources. In our approach, we aim to learn the latent factors of users and POIs through optimizing the relative ranking of users' preferences on different POIs. Toward this goal, we adopt the Bayesian Personalized Ranking (BPR) framework [Rendle et al. 2009] as the backbone of our model. Compared with other ranking-based recommendation methods [Li et al. 2015a], BPR works from a strict probabilistic perspective to maximize the posterior probability of parameters given observed pairwise preferences, which is a smoothed version of directly optimizing the well-known ranking measure of Area Under the receiver operating characteristic Curve (AUC). By making the assumption that all pairwise preferences are independent, the likelihood of $\mathcal{D}_Z$ and $\mathcal{D}_F$ over the parameters $\mathbf{U}$ and $\mathbf{V}$ is defined by

$$\mathcal{L} = \left(\prod_{(u,i,j) \in \mathcal{D}_Z} p((u,i,j)|\mathbf{u}_u, \mathbf{v}_i, \mathbf{v}_j)^{c_{ij}^u} \right) \left(\prod_{(u,i,j) \in \mathcal{D}_F} p((u,i,j)|\mathbf{u}_u, \mathbf{v}_i, \mathbf{v}_j)^{c_{ij}^u} \right)^\alpha. \tag{12}$$

Here, $p((u,i,j)|\mathbf{u}_u, \mathbf{v}_i, \mathbf{v}_j)$ is the probability of the pairwise preference (u,i,j) occurring, which is parameterized by the latent factors of user u as well as POIs i and j. The weight c_{ij}^u is introduced into the likelihood function, serving as the number of occurrences of (u,i,j). $\alpha \in [0,1]$ is a discount factor assigned to the pairwise preferences derived from user favoring behaviors, which represents our overall confidence in this uncertain implicit feedback. Mathematically, $p((u,i,j)|\mathbf{u}_u, \mathbf{v}_i, \mathbf{v}_j)$ is approximated by

$$p((u,i,j)|\mathbf{u}_u, \mathbf{v}_i, \mathbf{v}_j) = \sigma(r_{ij}^u), \tag{13}$$

where $r_{ij}^u = r_{ui} - r_{uj} = \mathbf{u}_u^T(\mathbf{v}_i - \mathbf{v}_j) + (b_i - b_j)$ and σ is the logistic sigmoid function, i.e., $\sigma(x) = 1/(1+\exp(-x))$. Furthermore, by placing zero-mean spherical Gaussian priors [Mnih and Salakhutdinov 2007] on the parameters $\mathbf{U}$ and $\mathbf{V}$, maximizing the

log-posterior of $\mathbf{U}$ and $\mathbf{V}$ is equivalent to minimizing the following objective:

$$\Omega = -\sum_{(u,i,j)\in\mathcal{D}_Z} c_{ij}^u \ln\sigma(r_{ij}^u) - \alpha \sum_{(u,i,j)\in\mathcal{D}_F} c_{ij}^u \ln\sigma(r_{ij}^u) + \frac{\beta_U}{2}\|\mathbf{U}\|_F^2 + \frac{\beta_V}{2}\|\mathbf{V}\|_F^2, \quad (14)$$

where $\|\cdot\|_F$ is the Frobenius norm, and β_U and β_V are the regularization parameters.

The above formulation exploits the user-POI preference information but ignores their relationship information, and thus may not result in reliable latent factors for users and POIs, especially when the known preferences are very sparse. Intuitively, users with close social relationships may share common topics of interest. It is expected that this intuition can be preserved under matrix factorization and help to learn better latent factors for users. To this end, we enforce the smoothness constraint of user latent factors with the user hypergraph. A regularizer is designed to allocate similar latent factors to the users together belonging to highly weighted hyperedges. Formally, the regularizer is defined as

$$\Theta_{\mathcal{I}} = \frac{1}{2}\sum_{u,v\in\mathcal{I}}\sum_{e\in\mathcal{E}_{\mathcal{I}}} \frac{w_{\mathcal{I}}(e)h_{\mathcal{I}}(u,e)h_{\mathcal{I}}(v,e)}{d_{\mathcal{E}_{\mathcal{I}}}(e)} \left\| \frac{\mathbf{u}_u}{\sqrt{d_{\mathcal{I}}(u)}} - \frac{\mathbf{u}_v}{\sqrt{d_{\mathcal{I}}(v)}} \right\|_2^2. \quad (15)$$

By introducing $\mathbf{L}_{\mathcal{I}} = \mathbf{I} - \mathbf{D}_{\mathcal{I}}^{-1/2}\mathbf{H}_{\mathcal{I}}\mathbf{W}_{\mathcal{I}}\mathbf{D}_{\mathcal{E}_{\mathcal{I}}}^{-1}\mathbf{H}_{\mathcal{I}}^T\mathbf{D}_{\mathcal{I}}^{-1/2}$ being the normalized Laplacian of $\mathcal{G}_{\mathcal{I}}$, Equation (15) can be rewritten in a concise form:

$$\begin{aligned}
\Theta_{\mathcal{I}} &= \sum_{u,v\in\mathcal{I}}\sum_{e\in\mathcal{E}_{\mathcal{I}}} \frac{w_{\mathcal{I}}(e)h_{\mathcal{I}}(u,e)h_{\mathcal{I}}(v,e)}{d_{\mathcal{E}_{\mathcal{I}}}(e)} \left(\frac{\mathbf{u}_u^T\mathbf{u}_u}{2d_{\mathcal{I}}(u)} + \frac{\mathbf{u}_v^T\mathbf{u}_v}{2d_{\mathcal{I}}(v)} - \frac{\mathbf{u}_u^T\mathbf{u}_v}{\sqrt{d_{\mathcal{I}}(u)}\sqrt{d_{\mathcal{I}}(v)}} \right) \\
&= \sum_{u\in\mathcal{I}} \mathbf{u}_u^T\mathbf{u}_u \sum_{e\in\mathcal{E}_{\mathcal{I}}} \frac{w_{\mathcal{I}}(e)h_{\mathcal{I}}(u,e)}{d_{\mathcal{I}}(u)} \sum_{v\in\mathcal{I}} \frac{h_{\mathcal{I}}(v,e)}{d_{\mathcal{E}_{\mathcal{I}}}(e)} - \sum_{u,v\in\mathcal{I}}\sum_{e\in\mathcal{E}_{\mathcal{I}}} \frac{\mathbf{u}_u^T h_{\mathcal{I}}(u,e)w_{\mathcal{I}}(e)h_{\mathcal{I}}(v,e)\mathbf{u}_v}{\sqrt{d_{\mathcal{I}}(u)}d_{\mathcal{E}_{\mathcal{I}}}(e)\sqrt{d_{\mathcal{I}}(v)}} \\
&= \mathrm{tr}(\mathbf{U}\mathbf{U}^T) - \mathrm{tr}\big(\mathbf{U}\mathbf{D}_{\mathcal{I}}^{-1/2}\mathbf{H}_{\mathcal{I}}\mathbf{W}_{\mathcal{I}}\mathbf{D}_{\mathcal{E}_{\mathcal{I}}}^{-1}\mathbf{H}_{\mathcal{I}}^T\mathbf{D}_{\mathcal{I}}^{-1/2}\mathbf{U}^T\big) = \mathrm{tr}\big(\mathbf{U}\mathbf{L}_{\mathcal{I}}\mathbf{U}^T\big),
\end{aligned} \quad (16)$$

where $\mathrm{tr}(\cdot)$ denotes the trace operator.

We also build upon the assumption that POIs having similar contents might exhibit similar attractions for POI recommendation. In light of this, the proximity of latent factors of similar POIs is required by the following regularizer on the POI hypergraph:

$$\Theta_{\mathcal{P}} = \frac{1}{2}\sum_{i,j\in\mathcal{P}}\sum_{e\in\mathcal{E}_{\mathcal{P}}} \frac{w_{\mathcal{P}}(e)h_{\mathcal{P}}(i,e)h_{\mathcal{P}}(j,e)}{d_{\mathcal{E}_{\mathcal{P}}}(e)} \left\| \frac{\mathbf{v}_i}{\sqrt{d_{\mathcal{P}}(i)}} - \frac{\mathbf{v}_j}{\sqrt{d_{\mathcal{P}}(j)}} \right\|_2^2 = \mathrm{tr}\big(\mathbf{V}\mathbf{L}_{\mathcal{P}}\mathbf{V}^T\big), \quad (17)$$

where $\mathbf{L}_{\mathcal{P}} = \mathbf{I} - \mathbf{D}_{\mathcal{P}}^{-1/2}\mathbf{H}_{\mathcal{P}}\mathbf{W}_{\mathcal{P}}\mathbf{D}_{\mathcal{E}_{\mathcal{P}}}^{-1}\mathbf{H}_{\mathcal{P}}^T\mathbf{D}_{\mathcal{P}}^{-1/2}$ is the normalized Laplacian for $\mathcal{G}_{\mathcal{P}}$.

Based on the definitions of $\Theta_{\mathcal{I}}$ and $\Theta_{\mathcal{P}}$, we extend our initial formulation in Equation (14) with hypergraph regularization. The new objective function is formulated as

$$\begin{aligned}
\Omega = &-\sum_{(u,i,j)\in\mathcal{D}_Z} c_{ij}^u \ln\sigma(r_{ij}^u) - \alpha \sum_{(u,i,j)\in\mathcal{D}_F} c_{ij}^u \ln\sigma(r_{ij}^u) \\
&+ \frac{\mu_U}{2}\mathrm{tr}\big(\mathbf{U}\mathbf{L}_{\mathcal{I}}\mathbf{U}^T\big) + \frac{\mu_V}{2}\mathrm{tr}\big(\mathbf{V}\mathbf{L}_{\mathcal{P}}\mathbf{V}^T\big) + \frac{\beta_U}{2}\|\mathbf{U}\|_F^2 + \frac{\beta_V}{2}\|\mathbf{V}\|_F^2,
\end{aligned} \quad (18)$$

where μ_U and μ_V are the tradeoff parameters controlling the relative contributions of the regularizers on the user hypergraph and POI hypergraph, respectively. Optimizing Equation (18) will lead to more effective latent factors for users and POIs, which are learned not only from the hybrid preference information derived with user-uploaded

ALGORITHM 1: Stochastic Gradient Descent

Input: Pairwise preferences $\mathcal{D}_Z$ and $\mathcal{D}_F$ and learning rate γ
Output: Latent factor matrices $\mathbf{U}$ and $\mathbf{V}$
1: Initialize $\mathbf{U}$ and $\mathbf{V}$ with random variables
2: **repeat**
3: Randomly draw a triple (u, i, j) from $\mathcal{D}_Z$
4: Update $\theta = \theta - \gamma \nabla_\theta \Omega$ based on Equation (19)
5: Randomly draw a triple (u, i, j) from $\mathcal{D}_F$
6: Update $\theta = \theta - \gamma \nabla_\theta \Omega$ based on Equation (19)
7: **until** convergence

and user-favored photos, but also from the heterogeneous high-order relationship information hidden in user and POI hypergraphs.

4.4. Optimization

Although the objective function in Equation (18) is not convex when simultaneously considering $\mathbf{U}$ and $\mathbf{V}$, a local minimum can be found by performing an alternating gradient descent procedure. In each iteration, we can take a step toward the gradient with respect to $\mathbf{U}$ by fixing $\mathbf{V}$, and vice versa. However, a potential difficulty arises in that the gradients with respect to $\mathbf{U}$ and $\mathbf{V}$ need to be accumulated over all pairwise preferences, leading to an excessive computational cost in each iteration. To solve this problem, we resort to the stochastic gradient descent algorithm. The basic principle is to draw a pairwise preference (u, i, j) randomly from $\mathcal{D}_Z$ or $\mathcal{D}_F$, and only the gradients regarding the parameters involved with (u, i, j) are computed in each iteration. Specifically, we use θ to denote an arbitrary model parameter, and the gradient of Ω with respect to θ given (u, i, j) can be computed by

$$\nabla_\theta \Omega = \begin{cases} -\alpha c_{ij}^u \big(1 - \sigma(r_{ij}^u)\big)(\mathbf{v}_i - \mathbf{v}_j) + \mu_U \sum_{v \in \mathcal{I}} l_{\mathcal{I}}(v, u)\mathbf{u}_v + \beta_U \mathbf{u}_u & \text{if } \theta = \mathbf{u}_u; \\ -\alpha c_{ij}^u \big(1 - \sigma(r_{ij}^u)\big)\mathbf{u}_u + \mu_V \sum_{q \in \mathcal{P}} l_{\mathcal{P}}(q, i)\mathbf{v}_q + \beta_V \mathbf{v}_i & \text{if } \theta = \mathbf{v}_i; \\ \alpha c_{ij}^u \big(1 - \sigma(r_{ij}^u)\big)\mathbf{u}_u + \mu_V \sum_{q \in \mathcal{P}} l_{\mathcal{P}}(q, j)\mathbf{v}_q + \beta_V \mathbf{v}_j & \text{if } \theta = \mathbf{v}_j; \\ -\alpha c_{ij}^u \big(1 - \sigma(r_{ij}^u)\big) & \text{if } \theta = b_i; \\ \alpha c_{ij}^u \big(1 - \sigma(r_{ij}^u)\big) & \text{if } \theta = b_j. \end{cases} \tag{19}$$

Note that $\alpha = 1$ when the triple (u, i, j) is sampled from $\mathcal{D}_Z$. $l_{\mathcal{I}}(v, u)$ and $l_{\mathcal{P}}(q, i)$ are the (v, u)-th and (q, i)-th entries of $\mathbf{L}_{\mathcal{I}}$ and $\mathbf{L}_{\mathcal{P}}$, respectively. At each iteration, model parameters are updated in the opposite direction of their respective gradients with a learning rate γ. The pseudo-code of the entire learning algorithm is presented in Algorithm 1.

Once the model parameters are learned, given a user $u \in \mathcal{I}$, the preference rating of u for a specific POI can be easily estimated by Equation (1). On the basis of this rating, we provide the top-ranked list of POIs as the personalized recommendations for u.

5. EXPERIMENTAL CONFIGURATION

This section introduces the experimental configuration of our performance evaluation. All the methods evaluated in this study have been fully implemented in Python and tested on a server equipped with a 24-core 2.00GHz Intel Xeon processor and 32GB RAM.

Table II. Summary of Basic Information for Each City

City	# of POIs	# of Images	# of Uploaders
London	331	144,724	13,583
Paris	218	110,416	11,251
Rome	93	49,698	6,375
Berlin	97	30,615	4,269
Amsterdam	86	23,920	3,139
Madrid	67	32,739	3,746
Dublin	42	13,607	1,680
Barcelona	106	53,384	5,507
Florence	26	16,533	2,755
Milan	44	15,547	2,464
Venice	74	31,368	4,444
Budapest	32	7,789	1,346
Prague	44	14,095	2,248
Vienna	32	8,798	1,334
Stockholm	34	7,381	1,230
Brussels	23	7,928	1,458
Copenhagen	32	6,730	1,135
Istanbul	33	11,280	1,663
Athens	16	5,789	981
Lisbon	40	13,169	2,080

5.1. Data Collections

Prior studies on personalized POI recommendation from geotagged images carried out experiments mainly based on self-collected datasets. In our evaluation, to ensure the comparability and fairness of empirical results, we adopted a publicly accessible dataset [Mousselly-Sergieh et al. 2014]. The dataset contains 14.1 million geotagged photos crawled from Flickr, which is currently the largest geotagged image dataset. For the purpose of personalized POI recommendation, we conducted experiments on the photos taken in the 20 most visited tourist cities in Europe. Toward this goal, we specified a bounding box[6] corresponding to the geographical region of each city, and retrieved the photos whose coordinates are within the bounding box. As described in Section 3, we applied mean shift clustering to discover the POIs in each city. We ultimately identified 1,470 POIs in the 20 cities from 605,510 geotagged photos uploaded by 45,910 users. Table II summarizes the basic information of each city.

As mentioned before, we took the number of photos uploaded by a user for a specific POI as his/her preference degree. The resulting user-POI preference matrix based on user-uploaded photos consists of 239,342 non-zero entries. Note that a user is considered to have previously visited a POI if there are uploading interactions between them. Figure 2 plots the distribution of the number of visited POIs over users. We can see that the distribution closely follows a power law, which reveals that most of the users have visited only a limited number of POIs. Additionally, we collected the favoring interactions between users and photos from Flickr. The number of a user's favored photos related to a POI also indicates his/her preference on that POI to some extent. The resulting user-POI preference matrix based on user-favored photos contains 270,389 non-zero entries. The main statistics of two preference matrices are displayed in Table III.

In order to construct the user hypergraph, we collected the groups joined by each user, and acquired the list of followees of each user. To reduce the complexity of the user hypergraph, we left out those inactive groups joined by less than 1,000 users,

[6]http://www.mapdevelopers.com/geocode_bounding_box.php.

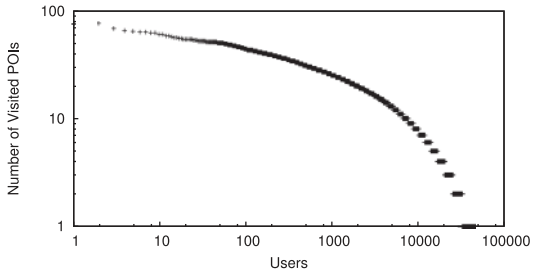

Fig. 2. Log-log plot on the distribution of the number of visited POIs over users. The x-axis represents the users in descending order of the number of visited POIs per user.

Table III. Statistics of User-POI Preference Matrices

Statistics	Uploading	Favoring
Max. # of POIs per user	79	657
Avg. # of POIs per user	5.21	5.89
Max. # of users per POI	3,127	3,898
Avg. # of users per POI	162.82	183.94
Sparseness	99.65%	99.60%

Table IV. Description of Metadata of Users and Photos

Statistics	
# of groups	17,186
# of followees	24,906
# of tags	3,795
Avg. # of groups per user	94.57
Avg. # of followees per user	121.40
Avg. # of tags per photo	4.27

and only retained some famous followees who have at least 100 followers. To construct the POI hypergraph, we used five types of low-level visual features to represent each photo, namely, (1) 64-dimensional color histogram, (2) 144-dimensional color correlogram, (3) 73-dimensional edge direction histogram, (4) 128-dimensional wavelet texture, and (5) 225-dimensional blockwise color moment. These features are easily extracted and characterize photos from different perspectives of color, shape, and texture. We concatenated different features of a photo into a single vector, and separately normalized each dimension of the feature vector into the range [0, 1]. In addition, since the user-generated tags associated with photos are rather noisy, we filtered out those tags appearing less than 50 times, and removed the ones that do not exist in the WordNet. A brief description of the aforementioned metadata is listed in Table IV.

5.2. Evaluation Methodology and Metrics

The most direct application of this study is to recommend POIs to a user when he/she is visiting a new city. Therefore, we set up the experiments to simulate this application scenario. To be specific, we selected the users who have visited at least two cities in our dataset. In total, there are 15,483 users meeting this criterion. For each of the selected users, we randomly chose one city that he/she has visited as the "destination city." During training, we ignored the visiting records of selected users regarding POIs in destination cities. A recommendation algorithm was verified by providing the POIs of destination cities for selected users, and comparing the results against their actual visited POIs that are retained in advance. In our experiments, the selected users were split into a validation set and a test set of the same size. The validation set was held out to investigate the impact of parameters, and the test set was used to compare the performance of different methods.

In practice, users may care more about the personalized rankings of POI recommendations than the predicted preference scores. We thus adopted two of the most widely used ranking evaluation metrics, i.e., Precision and Mean Average Precision (MAP), to evaluate the quality of a POI recommendation list for a given user. Precision at the

k-th position is computed as

$$P@k = \frac{1}{k}\sum_{i=1}^{k} rel(i), \tag{20}$$

where $rel(i)$ is a binary indicator, which is equal to 1 if the i-th POI in the recommendation list is truly visited by the user, and zero otherwise. The average value of P@k ($k = 1, 5, 10$) over all users was reported to evaluate the overall performance.

MAP is the mean of the average precision scores over all test users. The Average Precision (AP) is defined as

$$AP = \frac{\sum_{i=1}^{l} rel(i) \times P@i}{\sum_{i=1}^{l} rel(i)}, \tag{21}$$

where l denotes the length of the recommendation list. Obviously, P@k emphasizes the accuracy of the top-k recommendation results, while MAP reflects the quality of the entire recommendation list.

5.3. Competitors

We compared our approach against several state-of-the-art CF methods for personalized POI recommendation. Hyperparameters of these methods were optimized with the validation set. Specifically, the competitors in our experiments are

—**PopRank** [Crandall et al. 2009]: This method ranks POIs of a given city according to their popularity, which are estimated in terms of the number of distinct users having visited them. This is a non-personalized recommendation method. In other words, it recommends the same top-ranked POIs to all users.
—**UCF** [Clements et al. 2010]: This method provides a target user with the POIs preferred by other similar users. The similarity between two users is measured using the Pearson correlation between their visiting histories.
—**ATM** [Jiang et al. 2015]: This method goes a step further than UCF by measuring user-to-user similarities with their preference topic distributions. The author topic model is adopted to discover user topic distributions from photo tags.
—**WMF** [Hu et al. 2008]: This method integrates a weighting scheme into the matrix factorization model. The latent factors of users and POIs are obtained by fitting the observed preference scores from users on POIs. Meanwhile, POIs with more visited times for a user are assigned with larger weights in the learning process.
—**GeoMF** [Lian et al. 2014]: On the basis of WMF, this method further incorporates the geographical information by augmenting the latent factors in the matrix factorization model with activity area vectors of users and influence area vectors of POIs, respectively.
—**BPR** [Rendle et al. 2009]: This method optimizes the relative ranking of users' preferences on different POIs. The latent factors of users and POIs are learned by maximizing their posterior probabilities given observed pairwise preferences.
—**RMF-HG**: This is our proposed method in this article.

We also tried other basic matrix factorization models (e.g., the methods presented in Phan et al. [2014]). However, the performance of these methods is much poorer than that of the baselines listed above. For this reason, they were excluded from the comparative study.

5.4. Parameter Settings

There are several hyperparameters in our model. For the dimension of latent factors d, we performed a grid search over the range of [10, 100] with a step of 10. The results

Table V. Performance Comparison Among Different Methods

Metric	PopRank	UCF	ATM	WMF	GeoMF	BPR	RMF-HG
P@1	0.277	0.268	0.273	0.281	0.276	0.286	**0.306***
P@5	0.182	0.177	0.176	0.183	0.181	0.190	**0.201***
P@10	0.137	0.133	0.134	0.140	0.144	0.144	**0.151***
MAP	0.294	0.288	0.284	0.299	0.295	0.305	**0.323***

Bold typeset indicates the best performance, and * indicates it is statistically significant at $p < 0.05$ compared with the runner-up. This convention is also used in later tables.

demonstrated that there is no significant performance improvement when d is beyond 30. To reduce the computational complexity, we chose $d = 30$ for all the competitors based on matrix factorization. For the scaling parameters λ_Z and λ_F in Equation (4), we empirically chose them based on the ratio of the number of zero entries to that of nonzero ones in the preference matrix. For the adjustment parameter η in Equation (8), we treated the grouping factor and the following factor equally and set $\eta = 0.5$. For the number of nearest neighbors k considered when constructing the POI hypergraph, empirical studies [Fang et al. 2014; Xu et al. 2014] have showed that a relatively small number of k is sufficient to make the graph connected. We thus set $k = 5$ to make the POI hypergraph sparse for computational efficiency. For the discount factor α in Equation (12), we used $\alpha = 0.7$ in the experiments. For the tradeoff parameters μ_U and μ_V in Equation (18), we imposed the simplifying assumption that $\mu_U = \mu_V = \mu$ following the setting in Shi et al. [2013] and Ma [2013], and chose $\mu = 0.003$. The tuning procedures of α and μ will be discussed in detail later. For the regularization parameters β_U and β_V in Equation (18), we simply assume that $\beta_U = \beta_V = \beta$ to reduce the model complexity. The best performance on the validation set was achieved when $\beta = 0.0025$. For the learning rate γ in Algorithm 1, it was set to a small value of 0.001.

6. EXPERIMENT RESULTS AND DISCUSSIONS

In this section, we report a series of experiments conducted to evaluate our approach for personalized POI recommendation. Through these experiments, we try to address the following research questions:

—**RQ1:** Does our approach achieve its goal of enhancing the accuracy of personalized POI recommendation?
—**RQ2:** Does our approach help the users who have only a few travel experiences?
—**RQ3:** Does our approach benefit from the additional implicit feedback harvested from user-favored photos?
—**RQ4:** Does our approach work better when integrated with the heterogenous relationship information?

6.1. Overall Performance

Table V displays the empirical results of different methods for personalized POI recommendation. It is clearly shown that RMF-HG outperforms the other competitors in all evaluation metrics. For example, the maximum relative increases are 14.2% and 13.7% in terms of P@5 and MAP, whereas the minimum gains still reach 5.8% and 5.9%, respectively. To further analyze the results, we conducted a paired t-test [Smucker et al. 2007] to compare the difference between RMF-HG and the other methods, and found that the improvement of RMF-HG is statistically significant at a significance level of 0.05. Therefore, we drew the conclusion that RMF-HG emerges as the most effective recommendation scheme among the competitors. These results provide evidence that the research question **RQ1** can be positively answered. In addition, the following important observations can be made from Table V:

Table VI. Performance Comparison Among Methods on Users with Various Levels of Travel Experiences, in Terms of P@5

Group	PopRank	UCF	ATM	WMF	GeoMF	BPR	RMF-HG
Few	0.149	0.132	0.146	0.150	0.153	0.154	**0.167***
Limited	0.168	0.164	0.166	0.169	0.166	0.177	**0.191***
Medium	0.202	0.205	0.194	0.205	0.199	0.212	**0.221***
Rich	0.244	0.248	0.228	0.246	0.241	0.251	**0.253**

Table VII. Performance Comparison Among Methods on Users with Various Levels of Travel Experiences, in Terms of MAP

Group	PopRank	UCF	ATM	WMF	GeoMF	BPR	RMF-HG
Few	0.284	0.252	0.273	0.287	0.283	0.292	**0.313***
Limited	0.279	0.275	0.275	0.284	0.281	0.293	**0.319***
Medium	0.312	0.317	0.299	0.319	0.315	0.319	**0.330***
Rich	0.327	**0.343**	0.310	0.335	0.329	0.339	0.342

—PopRank achieves a competitive performance only with a simple ranking strategy based on POI popularity. This reflects the fact that visiting mainstream POIs is a general desire of most users when they arrive in a new city.
—The two examples of memory-based CF methods (i.e., UCF and ATM) both fall behind PopRank, and there is no significant performance difference between them. For UCF, a possible reason is that it cannot accurately find similar users when most users have visited only a limited number of POIs as shown in Figure 2. ATM leverages photo tags to identify the topics of user preferences. However, similar to the problem that user intentions are frequently not well described by query words, user preferences may not be well represented by tags as well [Cui et al. 2014].
—GeoMF incorporates the geographical information into WMF, but does not exhibit improved performance. This may be attributed to the fact that the candidate POIs to be recommended for a user are located in the new city and all far away from the other POIs. In addition, dissimilar to the case of choosing local venues, we conjecture that distance may not always be the primary factor for most users to select POIs within a city during travel.
—WMF, GeoMF, BPR, and RMF-HG, which belong to model-based CF methods, obtain better performance than the other contenders. This demonstrates the effectiveness of matrix factorization techniques. In addition, BPR and RMF-HG outperform WMF and GeoMF, which suggests that generating factorizations in a ranking way works better.
—From the results achieved by RMF-HG, we noticed that the relative improvement over the other methods in terms of P@1 and P@5 is more significant than those in terms of P@10. We believe this is a nice property as users are usually more interested in top-ranked recommendations.

6.2. Performance Across Different User Groups

In this subsection, we further study how the competitors behave for users with different levels of travel experiences. The travel experiences of a user were assessed by the number of POIs that user has visited. We categorized the travel experiences into four levels, i.e., few (with only one visited POI), limited (with 2–5 visited POIs), medium (with 6–10 visited POIs), and rich (with more than 10 visited POIs). The users from the test set were divided into different groups according to their respective levels of travel experiences. Out of 7,742 test users, the number of users in the four groups are 1,922, 3,168, 1,372, and 1,280, respectively.

Table VIII. Performance Comparison Among Different Models

Metric	BPR	VAR-F	VAR-U	VAR-P	VAR-SG	RMF-HG
P@1	0.286	0.296	0.290	0.294	0.300	**0.306***
P@5	0.190	0.199	0.191	0.193	0.199	**0.201**
P@10	0.144	0.149	0.146	0.146	**0.153**	0.151
MAP	0.305	0.316	0.310	0.313	0.316	**0.323***

Tables VI and VII summarize the performance of different methods for different user groups in terms of P@5 and MAP, respectively. As expected, the proposed approach RMF-HG still outperforms its counterparts with statistically significant improvement in most cases. Specifically, RMF-HG achieves a higher relative improvement for those users with few or limited travel experiences, reaching at least 8.4% and 7.9% for the two groups of users in terms of P@5, as well as 7.2% and 8.9% in terms of MAP. These results indicate that RMF-HG is particularly beneficial for users who have only a few travel experiences, and confirm the positive answer to the research question **RQ2**. Moreover, it should be noted that more than two-thirds of the test users belong to the few or limited groups. It is thus believed that RMF-HG is more practical in real recommendation scenarios. When recommending for users with rich travel experiences, all competitors obtain obvious performance gains, and the difference between the performance of leading methods is rather small. The observation implies that users with more travel experiences will benefit more from the recommendation service. It also underlines the fact that the amount of historical records has a significant impact on the accuracy of personalized POI recommendation.

6.3. Performance of Approach Variants

In the proposed scheme, we leverage the additional implicit feedback conveyed by user-favored photos, and capture the relationships among users and POIs via hypergraph models. To investigate the efficacy of each component, we design the following variants of our original approach:

—**VAR-F:** This method leverages the additional implicit feedback from user favoring behaviors, but do not incorporate the relationship information in the recommendation process. It is equivalent to RMF-HG (refer to Equation (18)) in the case that the parameter $\mu_U = \mu_V = 0$.
—**VAR-U:** This method only incorporates the user relationship information, and ignores the cues encoded in user-favored photos. In other words, it is equivalent to RMF-HG when the parameters $\alpha = 0$ and $\mu_V = 0$.
—**VAR-P:** It is similar to VAR-U, but exploits the POI relationship information instead of the user relationship information. That is, $\alpha = 0$ and $\mu_U = 0$.
—**VAR-SG:** This is a simplified version of RMF-HG adopting the simple graph rather than the hypergraph for modeling relationships among users and POIs, respectively.

Experiments were conducted to compare these variant methods against RMF-HG. BPR was also introduced as a comparative reference.

Table VIII presents the comparison results. As can be seen, on the one hand, compared with RMF-HG, all variants suffer some performance degradation in most evaluation metrics; on the other hand, in contrast to BPR, these methods still enjoy certain performance gains, leading to, for example, up to 4.9% relative improvement in terms of P@1. As a whole, these results imply that the recommendation accuracy can be enhanced by each way of harvesting the implicit feedback from user-favored photos and integrating the heterogeneous relationship information. We can thereby

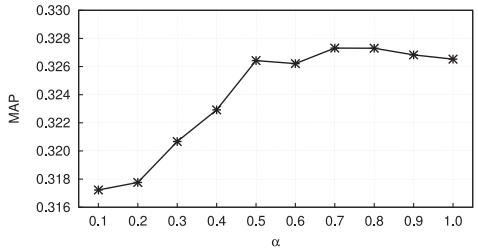

Fig. 3. Impact of the discount factor α regarding MAP performance.

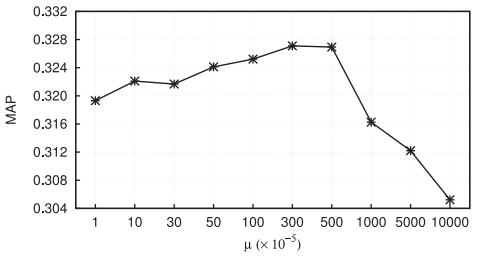

Fig. 4. Impact of the tradeoff parameter μ regarding MAP performance.

give positive answers to both the research questions **RQ3** and **RQ4**. In addition, we observed the following points from Table VIII:

—VAR-F is superior to VAR-U and VAR-P, which reveals that the way of using user-favored photos contributes more to alleviating data sparseness.
—VAR-P achieves better performance than VAR-U, indicating that POI relationship information offers more help than user relationship information. This may be attributed to the fact that the commonalities of POIs can be well estimated with their textual and visual contents, whereas the overlap of potential interests of users is more difficult to be discovered even though they share a few groups or followees in common.
—VAR-SG gets comparable performance to RMF-HG in P@5 and P@10, but loses a lot in P@1 and MAP, respectively. Such results point clearly to the importance of exploiting the hypergraph model to capture high-order relationships among users and POIs.

6.4. Parameter Tuning

In this section, we utilized the validation set to investigate the impact of two important parameters on our approach. Firstly, we analyzed the impact of the parameter α. α is a discount factor, indicating the confidence in the preferences derived with user-favored photos. Figure 3 illustrates how the change of α affects the performance in terms of MAP. We can see that the performance curve starts from a low MAP value, and gradually goes up when increasing α. It peaks when α is 0.7 or 0.8, and keeps relatively steady with larger values of α. This observation again verifies the potential of user favoring behaviors in the recommendation task, and again supports the positive answer to the research question **RQ3**.

We next examined the impact of the parameter μ. μ controls how much influence the relationship information has on the learning of latent factors. Intuitively, if we use a small value of μ, we mainly depend on users' own preferences in making recommendations. On the contrary, if we use a large value of μ, the learning process will be dominated by the relationships among users and POIs. Figure 4 shows how the performance varies with different values of μ. It can be observed that the value of μ impacts MAP performance significantly. As μ increases, the performance goes up at first, but when μ is beyond a certain threshold ($\mu = 0.005$ in our case), the performance experiences a sharp degradation with further increase of μ. This phenomenon coincides with our intuition, and indicates the necessity of considering both the preference information and the relationship information. From the results, the positive answer to the research question **RQ4** is underpinned.

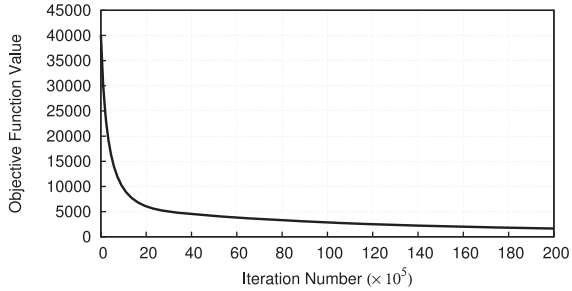

Fig. 5. Convergence process of the iterative optimization.

6.5. Efficiency Analysis

In order to understand the practical utility of our approach, we further analyzed the efficiency of our learning algorithm. Our analysis leaves out the operations for data preparation, which are independent of the proposed framework.

During the training process, the complexity of computing gradients in Equation (19) is $O((n + 2m)d)$. Since $d \ll n, m$, the overall complexity of one iteration in Algorithm 1 can be regarded as $O(n + 2m)$, which is linear in the sum of the number of users and POIs. In the experiments, our Python implementation of the algorithm took approximately 1.87 milliseconds per iteration. Figure 5 shows the convergence process of the iterative optimization during training, which is measured by the objective function value over a subset of examples in the validation set. It shows that the algorithm generally converges within 10^7 iterations. Moreover, it is worth noting that the learning process can be speeded up by adopting more efficient sampling strategies [Rendle and Freudenthaler 2014]. In sum, the above analysis verifies that the training computations are tractable and able to scale up for large-scale datasets.

Once the training process is finished, our approach can efficiently predict the preference degree of a user for a POI within $O(d)$ time. According to the elapsed time measured during testing, we found that our approach took an average of 0.13 milliseconds to produce the personalized recommendation list of POIs for a user. This means that our trained model can be used interactively by users without any perceived delay.

7. CONCLUSION AND FUTURE WORK

Geotagged image content services constitute one of the fastest-growing applications on the Web today. In this article, we have investigated the challenge of personalized POI recommendation based on geotagged photos, and addressed the problem of data sparseness via strengthening conventional CF algorithms from different aspects. Our approach is able to harvest the hybrid implicit feedback from user-uploaded and user-favored photos, and characterize the high-order relatedness among users and POIs. A matrix factorization model is developed to integrate disparate information into a unified framework, and it is capable of predicting user preferences in a ranking way. Extensive experiments are conducted on a large-scale public dataset in comparison with state-of-the-art methods. The results have verified the effectiveness of our approach, especially in providing POI recommendations for users who have visited only a few POIs in the past.

Our study also opens up two promising directions. Firstly, we plan to explore other information to identify user interests. Although our current study has verified that the number of user-uploaded and user-favored photos indeed reflects user preference degrees for POIs, other factors could also be explored to better characterize user interests. For example, we shall consider the amount of time that a user spent at a POI,

which could be estimated based on the temporal span of photos related to that POI. Secondly, we intend to adapt our approach to group recommendation [Shi et al. 2014]. Our current study recommends POIs to individuals. In reality, however, people often travel in groups, for example, with family members or friends. Under this situation, the quality of a given POI recommendation does not depend on the opinion of only one user, but on the group as a whole. We hope that this study will generate further interest in the emerging literature on user behaviors interacting with geotagged image contents and, more broadly, in mobile commerce and mobile social networks.

REFERENCES

Jie Bao, Yu Zheng, David Wilkie, and Mohamed Mokbel. 2015. Recommendations in location-based social networks: A survey. *Geoinformatica* 19, 3 (2015), 525–565.

Jiajun Bu, Shulong Tan, Chun Chen, Can Wang, Hao Wu, Lijun Zhang, and Xiaofei He. 2010. Music recommendation by unified hypergraph: Combining social media information and music content. In *Proc. 18th ACM Int. Conf. Multimedia*. 391–400.

Yan-Ying Chen, An-Jung Cheng, and W. H. Hsu. 2013. Travel recommendation by mining people attributes and travel group types from community-contributed photos. *IEEE Trans. Multimedia* 15, 6 (2013), 1283–1295.

Chen Cheng, Haiqin Yang, Irwin King, and Michael R. Lyu. 2012. Fused matrix factorization with geographical and social influence in location-based social networks. In *Proc. 26th AAAI Conf. Artificial Intelligence*.

Chen Cheng, Haiqin Yang, Michael R. Lyu, and Irwin King. 2013. Where you like to go next: Successive point-of-interest recommendation. In *Proc. 23rd Int. Joint Conf. Artificial Intelligence*, Vol. 13. 2605–2611.

Zhiyong Cheng and Jialie Shen. 2016. On very large scale test collection for landmark image search benchmarking. *Signal Process.* 124 (2016), 13–26.

Maarten Clements, Pavel Serdyukov, Arjen P. De Vries, and Marcel J. T. Reinders. 2010. Using Flickr geotags to predict user travel behaviour. In *Proc. 33rd Int. ACM SIGIR Conf. Research and Development in Information Retrieval*. 851–852.

David J. Crandall, Lars Backstrom, Daniel Huttenlocher, and Jon Kleinberg. 2009. Mapping the world's photos. In *Proc. 18th Int. Conf. World Wide Web*. 761–770.

Peng Cui, Shao-Wei Liu, Wen-Wu Zhu, Huan-Bo Luan, Tat-Seng Chua, and Shi-Qiang Yang. 2014. Social-sensed image search. *ACM Trans. Inform. Syst.* 32, 2 (2014), 8.

Quan Fang, Jitao Sang, Changsheng Xu, and Yong Rui. 2014. Topic-sensitive influencer mining in interest-based social media networks via hypergraph learning. *IEEE Trans. Multimedia* 16, 3 (2014), 796–812.

Shanshan Feng, Xutao Li, Yifeng Zeng, Gao Cong, Yeow Meng Chee, and Quan Yuan. 2015. Personalized ranking metric embedding for next new POI recommendation. In *Proc. 25th Int. Joint Conf. Artificial Intelligence*. 2069–2075.

Huiji Gao, Jiliang Tang, and Huan Liu. 2012. gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks. In *Proc. 21st ACM Int. Conf. Information and Knowledge Management*. 1582–1586.

Yue Gao, Jinhui Tang, Richang Hong, Qionghai Dai, Tat-Seng Chua, and Ramesh Jain. 2010. W2Go: A travel guidance system by automatic landmark ranking. In *Proc. 18th ACM Int. Conf. Multimedia*. 123–132.

Yue Gao, Meng Wang, Zheng-Jun Zha, Jialie Shen, Xuelong Li, and Xindong Wu. 2013. Visual-textual joint relevance learning for tag-based social image search. *IEEE Trans. Image Process.* 22, 1 (2013), 363–376.

Yue Gao, Sicheng Zhao, Yang Yang, and Tat-Seng Chua. 2015. Multimedia social event detection in microblog. In *MultiMedia Modeling*. Springer, 269–281.

Long Guo, Jie Shao, Kian Lee Tan, and Yang Yang. 2014. WhereToGo: Personalized travel recommendation for individuals and groups. In *Proc. 15th IEEE Int. Conf. Mobile Data Management*, Vol. 1. 49–58.

Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *Proc. 8th IEEE Int. Conf. Data Mining*. 263–272.

Shanshan Huang, Jun Ma, Peizhe Cheng, and Shuaiqiang Wang. 2015. A hybrid multigroup coclustering recommendation framework based on information fusion. *ACM Trans. Intell. Syst. Technol.* 6, 2 (2015), 27.

Rongrong Ji, Yue Gao, Bineng Zhong, Hongxun Yao, and Qi Tian. 2011. Mining Flickr landmarks by modeling reconstruction sparsity. *ACM Trans. Multimedia Comput.* 7, 1 (2011), 31.

Shuhui Jiang, Xueming Qian, Jialie Shen, Yun Fu, and Tao Mei. 2015. Author topic model-based collaborative filtering for personalized POI recommendations. *IEEE Trans. Multimedia* 17, 6 (2015), 907–918.

Takeshi Kurashima, Tomoharu Iwata, Go Irie, and Ko Fujimura. 2010. Travel route recommendation using geotags in photo sharing sites. In *Proc. 19th ACM Int. Conf. Information and Knowledge Management*. 579–588.

Xutao Li, Gao Cong, Xiao-Li Li, Tuan-Anh Nguyen Pham, and Shonali Krishnaswamy. 2015a. Rank-GeoFM: A ranking based geographical factorization method for point of interest recommendation. In *Proc. 38th Int. ACM SIGIR Conf. Research and Development in Information Retrieval*. 433–442.

Xinchao Li, Martha Larson, and Alan Hanjalic. 2015b. Global-scale location prediction for social images using geo-visual ranking. *IEEE Trans. Multimedia* 17, 5 (2015), 674–686.

Defu Lian, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. 2014. GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation. In *Proc. 20th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*. ACM, 831–840.

Marek Lipczak, Michele Trevisiol, and Alejandro Jaimes. 2013. Analyzing favorite behavior in Flickr. In *Multimedia Modeling*. Springer, 535–545.

Bin Liu, Yanjie Fu, Zijun Yao, and Hui Xiong. 2013. Learning geographical preferences for point-of-interest recommendation. In *Proc. 19th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*. 1043–1051.

Jiajun Liu, Zi Huang, Lei Chen, Heng Tao Shen, and Zhixian Yan. 2012. Discovering areas of interest with geo-tagged images and check-ins. In *Proc. 20th ACM Int. Conf. Multimedia*. 589–598.

Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. 2016. Predicting the next location: A recurrent model with spatial and temporal contexts. In *Proc. 30th AAAI Conf. Artificial Intelligence*. 194–200.

Hao Ma. 2013. An experimental study on implicit social recommendation. In *Proc. 36th Int. ACM SIGIR Conf. Research and Development in Information Retrieval*. 73–82.

Andriy Mnih and Ruslan Salakhutdinov. 2007. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*. 1257–1264.

Hatem Mousselly-Sergieh, Daniel Watzinger, Bastian Huber, Mario Döller, Elöd Egyed-Zsigmond, and Harald Kosch. 2014. World-wide scale geotagged image dataset for automatic image annotation and reverse geotagging. In *Proc. 5th ACM Multimedia Systems Conf.* 47–52.

Weike Pan, Hao Zhong, Congfu Xu, and Zhong Ming. 2015. Adaptive Bayesian personalized ranking for heterogeneous implicit feedbacks. *Knowl.-Based Syst.* 73 (2015), 173–180.

Thomas Phan, Jiayu Zhou, Shiyu Chang, Junling Hu, and Juhan Lee. 2014. Collaborative recommendation of photo-taking geolocations. In *Proc. 3rd ACM Multimedia Workshop Geotagging and Its Applications in Multimedia*. 11–16.

Steffen Rendle and Christoph Freudenthaler. 2014. Improving pairwise learning for item recommendation from implicit feedback. In *Proc. 7th ACM Int. Conf. Web Search and Data Mining*. 273–282.

Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proc. 25th Conf. Uncertainty in Artificial Intelligence*. 452–461.

Stevan Rudinac, Alan Hanjalic, and Matt Larson. 2013. Generating visual summaries of geographic areas using community-contributed images. *IEEE Trans. Multimedia* 15, 4 (2013), 921–932.

Junge Shen, Zhiyong Cheng, Jialie Shen, Tao Mei, and Xinbo Gao. 2014. The evolution of research on multimedia travel guide search and recommender systems. In *MultiMedia Modeling*. Springer, 227–238.

Junge Shen, Jialie Shen, Tao Mei, and Xinbo Gao. 2016. Landmark reranking for smart travel guide systems by combining and analyzing diverse media. *IEEE Trans. Syst., Man, Cybern., Syst.* 46, 11 (2016), 1492–1504.

Yue Shi, Martha Larson, and Alan Hanjalic. 2014. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *ACM Comput. Surv.* 47, 1 (2014), 3.

Yue Shi, Pavel Serdyukov, Alan Hanjalic, and Martha Larson. 2013. Nontrivial landmark recommendation using geotagged photos. *ACM Trans. Intell. Syst. Technol.* 4, 3 (2013), 47.

Mark D. Smucker, James Allan, and Ben Carterette. 2007. A comparison of statistical significance tests for information retrieval evaluation. In *Proc. 16th ACM Int. Conf. Information and Knowledge Management*. 623–632.

Roelof van Zwol, Adam Rae, and Lluis Garcia Pueyo. 2010. Prediction of favourite photos using social, visual, and textual signals. In *Proc. 18th ACM Int. Conf. Multimedia*. 1015–1018.

Shuaiqiang Wang, Jiankai Sun, Byron J Gao, and Jun Ma. 2014. VSRank: A novel framework for ranking-based collaborative filtering. *ACM Trans. Intell. Syst. Technol.* 5, 3 (2014), 51.

Xiangyu Wang, Yi-Liang Zhao, Liqiang Nie, Yue Gao, Weizhi Nie, Zheng-Jun Zha, and Tat-Seng Chua. 2015b. Semantic-based location recommendation with multimodal venue semantics. *IEEE Trans. Multimedia* 17, 3 (2015), 409–419.

Yingzi Wang, Nicholas Jing Yuan, Defu Lian, Linli Xu, Xing Xie, Enhong Chen, and Yong Rui. 2015a. Regularity and conformity: Location prediction using heterogeneous mobility data. In *Proc. 21st ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*. ACM, 1275–1284.

Jiejun Xu, Vignesh Jagadeesh, and B. S. Manjunath. 2014. Multi-label learning with fused multimodal bi-relational graph. *IEEE Trans. Multimedia* 16, 2 (2014), 403–412.

Yiyang Yang, Zhiguo Gong, and others. 2014. Identifying points of interest using heterogeneous features. *ACM Trans. Intell. Syst. Technol.* 5, 4 (2014), 68.

Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proc. 34th Int. ACM SIGIR Conf. Research and Development in Information Retrieval*. 325–334.

Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat Thalmann. 2013. Time-aware point-of-interest recommendation. In *Proc. 36th Int. ACM SIGIR Conf. Research and Development in Information Retrieval*. 363–372.

Jia-Dong Zhang, Chi-Yin Chow, and Yu Zheng. 2015. ORec: An opinion-based point-of-interest recommendation framework. In *Proc. 24th ACM Int. Conf. Information and Knowledge Management*. 1641–1650.

Vincent W. Zheng, Yu Zheng, Xing Xie, and Qiang Yang. 2010. Collaborative location and activity recommendations with GPS history data. In *Proc. 19th Int. Conf. World Wide Web*. 1029–1038.

Yu Zheng and Xing Xie. 2011. Learning travel recommendations from user-generated GPS traces. *ACM Trans. Intell. Syst. Technol.* 2, 1 (2011), 2.

Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie, and Wei-Ying Ma. 2011. Recommending friends and locations based on individual location history. *ACM Trans. Web* 5, 1 (2011), 5.

Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *Proc. 18th Int. Conf. World Wide Web*. 791–800.

Lei Zhu, Jialie Shen, Hai Jin, Liang Xie, and Ran Zheng. 2015a. Landmark classification with hierarchical multi-modal exemplar feature. *IEEE Trans. Multimedia* 17, 7 (2015), 981–993.

Lei Zhu, Jialie Shen, Hai Jin, Ran Zheng, and Liang Xie. 2015b. Content-based visual landmark search via multimodal hypergraph learning. *IEEE Trans. Cybern.* 45, 12 (2015), 2756–2769.

Lei Zhu, Jialie Shen, Xiaobai Liu, Liang Xie, and Liqiang Nie. 2016. Learning compact visual representation with canonical views for robust mobile landmark search. In *Proc. 25th Int. Joint Conf. Artificial Intelligence*. 3959–3965.

Received January 2016; revised February 2017; accepted April 2017