Expert Systems With Applications 000 (2017) 1–11

Contents lists available at ScienceDirect

Expert Systems With Applications

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

A time-aware spatio-textual recommender system

Pavlos Kefalas*, Yannis Manolopoulos

Department of Informatics, Aristotle University, 54124 Thessaloniki, Greece

A R T I C L E I N F O

Article history:
Received 26 August 2016
Revised 8 January 2017
Accepted 9 January 2017
Available online xxx

Keywords:
Recommender systems
Location based services
Collaborative filtering
Review recommendation
Point-of-interest recommendation
Spatio-textual analysis
Temporal analysis

A B S T R A C T

Location-Based Social Networks (LBSNs) allow users to post ratings and reviews and to notify friends of these posts. Several models have been proposed for Point-of-Interest (POI) recommendation that use *explicit* (i.e. ratings, comments) or *implicit* (i.e. statistical scores, views, and user influence) information. However the models so far fail to capture sufficiently user preferences as they change spatially and temporally. We argue that time is a crucial factor because user check-in behavior might be periodic and time dependent, e.g. check-in near work in the mornings and check-in close to home in the evenings. In this paper, we present two novel unified models that provide review and POI recommendations and consider simultaneously the spatial, textual and temporal factors. In particular, the first model provides review recommendations by incorporating into the same unified framework the spatial influence of the users' reviews and the textual influence of the reviews. The second model provides POI recommendations by combining the spatial influence of the users' check-in history and the social influence of the users' reviews into another unified framework. Furthermore, for both models we consider the temporal dimension and measure the impact of time on various time intervals. We evaluate the performance of our models against 10 other methods in terms of *precision* and *recall*. The results indicate that our models outperform the other methods.

1. Introduction

Nowadays, On-line Social Networks (OSNs) give users the opportunity to communicate and share interests with other users. The incorporation of longitude and latitude data triggered new functionalities and introduced the Location-based Social Networks (LBSNs), a subset of OSNs, where users can share geo-tagged information such as check-ins, photos, text etc. In such networks —like for example, Yelp,[1] Google Places,[2] and TripAdvisor[3] users can additionally share ratings and reviews of businesses as shown in Fig. 1. The availability (in large volumes) of *explicit* information such as ratings, comments, social ties, and of *implicit* information such as statistical scores, views, and user influence, raises new challenges in Recommender Systems (RS). In particular, accurate personalized recommendation is hindered by the lack of adequate user information.

To address this issue, previous works aim on combining information derived from multi-modal and heterogeneous explicit and implicit networks. There has been extensive research which primarily focuses on the information that gets derived by the users' interaction with locations over user-location bipartite network ties. However, such approaches are static and fail to capture preference dynamics. Other works examine the evolution of users' preferences by capturing the temporal dynamics of users' check-in behavior. Even though such approaches are dynamic, they fail to capture adequately users' preferences dynamics since the evolution is affected by additional contextual information not taken under consideration. In particular, a newly introduced feature in LBSNs is that users can post reviews on the locations they check-in. The systematic repetition or use of words in the posts indicates a correlation among certain users: teenagers use different vocabulary from the elders, doctors use different terminology than civil engineers. Thus textual evolution should be examined.

In this paper, we present a model that focuses on the user's spatial, temporal, and textual behavior evolution, diversified in time intervals since users potentially change behavior over time (venues they attend or words they use). In the following subsections, we present the background and preliminaries that indicate that these three factors are crucial in recommender systems and that they should all be considered to capture the preference evolution. Moreover, we discuss how the implicit and explicit data that is derived from the users' daily behavior can significantly improve

* Corresponding author.
E-mail addresses: kefalasp@csd.auth.gr (P. Kefalas), manolopo@csd.auth.gr (Y. Manolopoulos).

[1] https://www.yelp.com
[2] http://www.google.co.uk/business/
[3] https://www.tripadvisor.com

http://dx.doi.org/10.1016/j.eswa.2017.01.060
0957-4174/

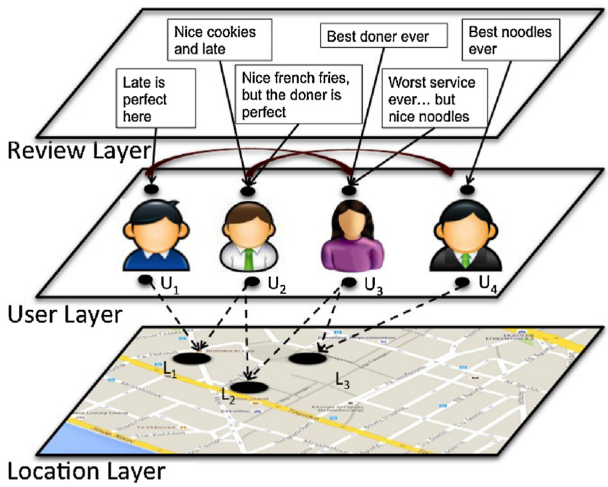


Fig. 1. Toy Example

the personalization of the recommendations, when incorporated as contextual information into a model.

1.1. Spatial-based behavior

Recent research points out that users maintain a fixed daily program in their activities and the locations they check-in (Cho, Myers, & Leskovec, 2011). For example, on weekdays a user checks-in at locations close to work $9-5$, whereas from 6 p.m. until the next morning s/he checks-in at locations close to home. On weekends, this routine changes and users check-in at different locations. It is inferred that users with similar check-in histories will probably share same preferences and interests (Li et al., 2008). Thus spatial proximity should be considered in recommender systems as a user can join existing communities, access services (Ben Nejma, Roose, Gensel, Dalmau, & Ghorbali, 2013), or make new friends located in close distance (Bravo-Torres, López-Nores, Blanco-Fernández, & Pazos-Arias, 2013). We also note that Bravo-Torres, López-Nores, Blanco-Fernández, and Pazos-Arias (2014) pointed that users participate to different networks as they explore the city, composing a Sporadic Social Network (SSN).

1.2. Temporal-based behavior

Several works attempt to predict user behavior by using both explicit and implicit information (Lee, Park, & Park, 2008; Lu, Savas, Tang, & Dhillon, 2010; Vasuki, Natarajan, Lu, Savas, & Dhillon, 2011). However, users tend to change their behavior (Cho et al., 2011) and preferences over time (Koren, 2009) and the previous models fail to capture this evolution. For example, a user may attend events close to the place his favorite band performs on Fridays and at other times he may visit markets for discounts. In both cases, the same user has a different check-in behavior, which should be taken into consideration. The change in preferences according to Lathia, Hailes, Capra, and Amatriain (2010) and Xiong, Chen, Huang, Schneider, and Carbonell (2010) may be due to:

- *New locations exploration:* curiosity leads users to visit new locations contrary to their ordinary choices.
- *User experience:* if a user has a pleasant experience in a POI, then s/he will probably choose the same or a similar venue in the future.
- *Popularity:* users interact with a bias based on popularity irrespectively of their previous history.
- *Social influence:* friends' opinions are important when making decisions. Users tend to examine their friends' evaluations and follow their lead.

Nejma, Roose, Dalmau, and Gensel (2015) points out that the communities in a SSN are short-lived since the involved members diversify with time. For example, users who are constantly on the move are replaced by other users. Thus, the participation of a user into communities evolves spatially and temporally (Smaldone, Han, Shankar, & Iftode, 2008) and the communities should be created ad-hoc, by considering the online presence of a user in the SSN (Srba & Bielikova, 2015). For all these reasons, the temporal dimension should be taken into consideration when establishing a recommender system.

1.3. Textual-based behavior

Reviews, tags, comments and blog contents are used as auxiliary or side information to overcome the sparsity issue and recommend a location or an item (Liu, Fu, Yao, & Xiong, 2013; Liu & Xiong, 2013). Users read online reviews/blogs/forums before attending an event or buying a product. An approach for POI recommendations is based on a lexical analysis of the user's reviews in order to find correlations with reviews of other users (Tang, Tan, & Cheng, 2009). Other approaches focus on rating prediction (Zhang & Varadarajan, 2006), review summarization (Hu & Liu, 2004) etc. In this paper, we propose *review recommendation*, a novel topic in recommender systems. With this term we mean recommendations of reviews that concern proximate POIs, rather than the use of the reviews as side information.

1.4. Motivation

We summarize the limitations of the previous approaches to recommendation strategies, that relate to the users' spatial, temporal, and textual behaviors:

- many methods consider POIs as conventional items and do not capture the geographical proximity influence,
- recommendation strategies that consider textual influence perform lexical analysis over the reviews and use them as side information to provide location/item recommendations ignoring users' preference dynamics, and
- methods that capture temporal dynamics do not treat simultaneously the spatial and textual dimensions.

Thus the need arises for the consideration of a model that combines textual, spatial and temporal influences in recommender systems.

1.5. Contribution and outline

In this paper we present two novel models that combine textual, spatial and temporal influences and provide review recommendation and POI recommendations. To the best of our knowledge, the model for personalized review recommendations is a *novel feature* in recommender systems. In addition, the contributions of our work are as follows:

- For the review recommendations model, we extend the item-based Collaborative filtering (CF) by incorporating the *spatial influence of user reviews*, and the *textual influence among reviews*. For the POI recommendations model, we extend the user-based CF by incorporating the *spatial influence of the user check-in history*, and the *social influence of user reviews*. The two hybrid models combine features from both the collaborative and the content-based filtering to overcome the drawbacks that each approach has separately.

- We consider the *temporal dimension* and for each model we measure the temporal influence at different time intervals.
- We evaluate the performance of our methods against 10 methods in terms of precision and recall for the top-*N* predictions. Regarding the review recommendation model, the experiments indicate that as we extend our model with more dimensions, the results are becoming more personalized and the overall performance is boosted. Also, regarding the POI recommendation model our method outperforms by 20% the state-of-the-art model presented by Yuan, Cong, Ma, Sun, and Thalmann (2013).

This paper is organized as follows. Section 2 summarizes the related studies, whereas Section 3 illustrates our models' structural parts in detail. In Section 3.3 we discuss the incorporation of the temporal dimension to both models. In Section 4 we present the evaluation of our work and finally, Section 5 concludes this paper.

2. Related work

2.1. Collaborative filtering (CF)

CF is a widely used technique in recommender systems, which bases its predictions on other users' behavior (Yang, Guo, Liu, & Steck, 2014). There are many CF methods which make predictions based on the assumption that if users agree on some items, then they most likely agree about other items for which there has been no recorded interaction (Deng, Huang, & Xu, 2014; Lee & Lee, 2015; Liu & Lee, 2010; Sarwar, Karypis, Konstan, & Riedl, 2001). There are two main categories for the CF methods: the *memory-based* methods and the *model-based* methods. The first category is further divided into two main subcategories: the *user-based CF* methods (Breese, Heckerman, & Kadie, 1998; Jin, Chai, & Si, 2004; Wang, de Vries, & Reinders, 2006; Zhao & Shang, 2010) and the *item-based CF* methods (Deshpande & Karypis, 2004; Karypis, 2001; Li, Zhao, Wu, Mao, & Cui, 2015; Pirasteh, Jung, & Hwang, 2014). Here, we focus on memory-based approaches. The user-based approach finds other users with similar rating behavior with the use of a similarity measure (e.g., Pearson correlation, Spearman rank correlation, cosine similarity, etc.). The resulting similarity score is used to compute predictions for new items by weighting it, using users' rating history. The item-based approach correlates items with similar items that the users have rated to make predictions.

2.2. Temporal dimension

Koren introduced the *timeSVD++* algorithm, which captures lasting and transient factors by modeling the user's preference dynamics through the entire time period (Koren, 2009). The goal is to distill the longer-term preferences from noisy patterns with the use of a matrix factorization model. He shows that in an item-item neighborhood model, the essential relations can be extracted by learning how ratings evolve.

Similarly, Zhang et al. presented two models that capture the user's preference dynamics: the *Temporal Matrix Factorization* method (*TMF*) and the *Bayesian Temporal Matrix Factorization* (*BTMF*) method (Zhang, Wang, Yu, Sun, & Lim, 2014). *TMF* maps the user and the item preferences into a joint latent factor with a transition matrix, which captures the user's preference dynamics between time periods. By sampling the rating distribution they update the transition matrix for past and future time periods. *BTMF* extends *TMF* by introducing priors for the hyper parameters to increase the accuracy and deals with the complexity of *TMF*. We remark that the related work focuses on temporal dynamics and do not consider the textual and the spatial dimensions.

In Ding and Li (2005) the authors introduced a user-based CF model that uses a time function. The authors argue that user's behavior evolves between different time periods. Moreover, they assume that the larger the time difference, the smaller the significance of older check-ins. Thus, they introduced a model that weights the check-ins of similar users according to the time difference from the current time period.

2.3. POI recommendation

Recently, LBSNs attracted the attention of the recommender system community (Cao, Cong, & Jensen, 2010; Wang et al., 2013; Hu & Ester, 2013; C. Li et al., 2015; Li, Xu, Chen, & Zong, 2015; Liu, Xu, Liao, & Chen, 2014; Vansteenwegen, Souffriau, Berghe, & Oudheusden, 2011; Yao et al., 2015; Zheng, Cao, Zheng, Xie, & Yang, 2010; Zheng, Zhang, Xie, & Ma, 2009). Ye et al. in Ye, Yin, Lee, and Lee (2011), the authors improved recommendation accuracy using information obtained from *trust-based* relations. They modeled the spatial influence of this network using a Bayesian CF algorithm. They exploited the social influence of user's friends interactions, rather than considering all user interactions, and computed the recommendation score for unvisited POI (Ma, Lyu, & King, 2009). To capture the spatial influence, they assumed that a user is willing to visit a new POI and considered the product of pairs consisting of check-ins and new POI.

In the same direction, Yuan et al. presented another method based on user-based CF which explores the spatial and temporal influence of POIs (Yuan et al., 2013). The main difference between this approach and the one of Ye et al. (2011) is the assumption that the user's willingness to move from her/his current location to a different one is a function of the in-between distance, and it follows the power law distribution. Moreover, they introduced a Bayes probabilistic model to make predictions for unvisited POIs. Finally, they also incorporated the temporal influence of the check-ins to further improve the accuracy of the recommendations.

The previous works focus on GPS datasets (Cao et al., 2010; Liu et al., 2014; Vansteenwegen et al., 2011), consider POIs as conventional items (Wang et al., 2013; Hu & Ester, 2013; C. Li et al., 2015; X. Li et al., 2015; Yao et al., 2015; Zheng et al., 2010; Zheng et al., 2009) and ignore the textual and geographical proximity influence among POIs when predicting an unvisited POI. Additionally, even though they consider the geographical influence, the correlation between two locations is not determined in terms of geographical proximity but in terms of conventional locations ignoring their spatial relationship.

2.4. Review recommendation

Content-based approaches assume that each item is related to a vector of the tokenized words of the review (Balabanović & Shoham, 1997; Esparza, O'Mahony, & Smyth, 2011; Lops, de Gemmis, & Semeraro, 2011). These terms are weighted with the use of Term Frequency/Inverse Document Frequency (TF-IDF) (Salton & Buckley, 1988) and are correlated with users' profiles by aggregating the items/locations profiles with the users' past rating history. The most similar profiles are found using a similarity measure to compute the final predictions. Content-based approaches recommend items/locations but do not consider the spatial and temporal dimensions and do not provide review recommendations. In contrast to these works, our approach performs lexical analysis emphasizing on the impact of the tokenized words of the reviews to correlate users with significant words and our model provides review recommendation by taking into consideration both the spatial and the temporal factors. To the best of our knowledge, our approach is the first attempt to create a model providing personalized review recommendations with combined spatial and textual information.

Table 1
Symbols notations and descriptions.

Symbol	Description
U, R, L, W	Set of users $U=\{u_1,\ldots,u_n\}$, Set of reviews $R=\{r_1,\ldots,r_m\}$, Set of locations $L=\{l_1,\ldots,l_m\}$, Set of words $W=\{w_1,\ldots,w_k\}$
u, v, r, l, w	user $u, v \in U$, review $r \in R$, location $l \in L$, word $w \in W$
$w^{(Textual)}_{(r,r')}$	Textual similarity between reviews r and r'
$w^{(Spatial)}_{(u,v)}$	Similarity between users u and v according to the check-in history
$w^{(Review)}_{(u,v)}$	Similarity between users u and v according to the review history
CL	Check-in matrix of users U over locations L
CR	Review matrix of users U over locations L
RW	Review matrix of reviews R over words W
c_r, c_l	Binary vector of user u over R, Check-in vector of user u over L
$c_{(u,r)}$, $c_{(u,l)}$	Element of c_u, Element of c_l

3. The proposed model

Table 1 presents the main notation used in the sequel, whereas our problem can be formulated as follows.

Problem Definition: *"Given a user u at a location l and her/his review history, the goal is to predict: (i) the top-N unread reviews for proximate stores, and (ii) the top-N unvisited proximate POIs, by capturing the temporal, spatial and textual properties."*

Next we present the recommender system that provides review and POIs recommendations. Please notice that both models use contextual pre-filtering of the information to select the most relevant proximate users for the recommendations (Adomavicius & Tuzhilin, 2008).

3.1. Review recommendations

In this section, we introduce the mathematical formulation of our model for personalized review recommendations. Our model incorporates both the *spatial influence of user reviews* (Section 3.1.1), and the *textual influence among the reviews* (Section 3.1.2). The spatial influence of users' reviews (in short, *spatial influence*) represents the impact of proximate users who reviewed similar businesses to the target user, whereas the textual influence among the reviews (in short, *textual influence*) refers to the similarity between reviews.

3.1.1. Incorporating spatial influence

With respect to the spatial influence, our model extends the item-based CF model in two ways: (i) we leverage the proximity factor when computing the similarity of two users, and (ii) we consider the history of proximate user reviews, rather than the entire review history. For the spatial influence, we assume that users are interested only in reviews about proximate businesses, in contrast to distant locations: it is highly unlikely that the target user will read an unread review for a distant store. Thus we use information derived only from: (i) users proximate to the target user, and (ii) reviews made within a range Rg. First, we give the formal definition of the user review range query.

Definition 1 Extract proximate users-reviews history. Given a user-review-location matrix CR, the range query with radius Rg at a given location l is defined as:

$$c_{u,r_{(Rg)}} = \{c_{v,r,l} | c_{v,r,l} \in CR \wedge d(c_{u,r,l}, c_{v,r,l}) \le Rg\} \tag{1}$$

This query returns the entire review history of proximate users at a distance less or equal to the range Rg. Fig. 2 illustrates an example of acquired review history for all users within a range of 2000 m.

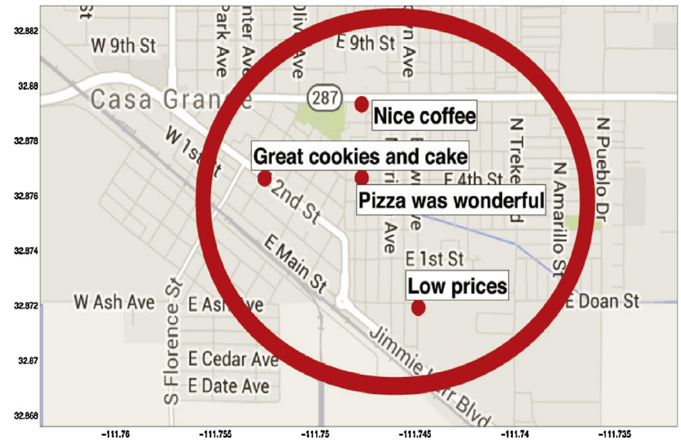


Fig. 2. Proximate users and their review history at a distance Rg = 2000 m

Next, we compute the similarity matrix for proximate users u and v using the formula:

$$w^{(Review)}_{u,v} = \frac{\sum_{r=1}^{m} c_{u,r} \cdot c_{v,r_{(Rg)}}}{\sqrt{\sum_{r=1}^{m} c_{u,r}^2} \cdot \sqrt{\sum_{r=1}^{m} c_{v,r_{(Rg)}}^2}} \tag{2}$$

where $c_{v,r_{(Rg)}}$ are the vectors of proximate users V over the reviews R in range Rg.

Finally, we give the prediction score that incorporates the spatial influence. Namely, given a user u at a location l and her/his review history, the prediction score that the user u would be interested in reading an unread review is as follows:

$$\widehat{c_{u,r}}^{(Review)} = \frac{\sum_v w^{(Review)}_{u,v} \cdot c_{v,r_{(Rg)}}}{\sum_v w_{u,v}{}^{(Review)}}. \tag{3}$$

3.1.2. Incorporating textual influence

With respect to the textual influence, we argue that textual closeness gives a similarity measure. Users who use the same vocabulary, can be classified into the same group: this may be because they belong to the same age group, or because they may have grown at the same geographical region, etc. We tokenize each review and create an adjacency matrix RW of the reviews and the contained words. We compute the similarity between reviews r and r' by the formula:

$$w^{(Textual)}_{r,r'} = \frac{\sum_{w=1}^{k} c_{r,w} \cdot c_{r',w}}{\sqrt{\sum_{w=1}^{k} c_{r,w}^2} \cdot \sqrt{\sum_{w=1}^{k} c_{r',w}^2}} \tag{4}$$

The prediction score that incorporates the textual influence, for a user u at a location l and given her/his review history is:

$$\widehat{c_{u,r}}^{(Textual)} = \frac{\sum_r w^{(Textual)}_{r,r'} \cdot c_{v,r_{(Rg)}}}{\sum_r w_{r,r'}{}^{(Textual)}} \tag{5}$$

where $w^{(Textual)}_{r,r'}$ is the similarity matrix among the reviews based on the words they contain.

3.1.3. Unified model for review recommendations

The unified model for review recommendations uses both the spatial influence and the textual influence of the reviews. Since both scores are in different scale, we map them in the scale $[0-1]$ with the following relations:

$$\begin{aligned} \widehat{c_{u,r}}^{(Review)} &= \frac{\widehat{c_{u,r}}^{(Review)} - min(\widehat{c_{u,r}}^{(Review)})}{max(\widehat{c_{u,r}}^{(Review)}) - min(\widehat{c_{u,r}}^{(Review)})} \\ \widehat{c_{u,r}}^{(Textual)} &= \frac{c_{u,r}{}^{(Textual)} - min(c_{u,r}{}^{(Textual)})}{max(c_{u,r}{}^{(Textual)}) - min(c_{u,r}{}^{(Textual)})} \end{aligned} \tag{6}$$

where $min(\cdot)$ and $max(\cdot)$ are the minimum and the maximum values of these vectors. We combine the previous scores into a single expression as follows:

$$\widehat{c_{u,r}} = a \cdot \widehat{C_{u,r}}^{(Review)} + (1-a) \cdot \widehat{C_{u,r}}^{(Textual)} \tag{7}$$

Thus, our model computes the review scores for all unread reviews and returns the top-N reviews as the recommendation result.

3.2. Points-of-interest recommendations

Here, we present the mathematical formulation of our model that generates personalized POI recommendations (see Section 3.1). Our model incorporates both the spatial influence of the user's check-in history (in short, *spatial influence*) and the social influence of the user's reviews (in short, *social influence*). The former notion is related to the impact of proximate users who behave similarly to the target user, whereas the latter one represents the impact of users who review similar business stores to the target user. In particular, the social influence corresponds to the correlation between the target user and others with regard to the lexical analysis of their reviews. As noted earlier, people who use the same vocabulary may behave similarly and we assume that the vocabulary used in the reviews reveals the social influence and leads to a correlation of users. Our model is based on the assumption that recommendation accuracy is increased if data is used that concerns proximate users, who do check-ins at similar locations combined with the reviews done by users reviewing similar business stores to the ones that the target user does.

3.2.1. Incorporating spatial influence

To take into account the spatial influence, we extend the user-based CF model as in Eq. (5). Thus, given a user u at location l, the prediction score that the user checks-in an unvisited POI is:

$$\widehat{c_{u,l}}^{(Spatial)} = \frac{\sum_v w_{u,v}^{(Spatial)} \cdot c_{v,l_{(Rg)}}}{\sum_v w_{u,v}^{(Spatial)}} \tag{8}$$

where $w_{u,v}^{(Spatial)}$ is the similarity matrix among users based on their check-in history. We claim that the probability of ignoring a POI (for which there exists a recommendation) is inversely related to the distance from the target user. Thus, we use information derived only from users proximate to the target user in a range distance Rg. We derive the similarity among proximate users as:

$$w_{u,v}^{(Spatial)} = \frac{\sum_{l=1}^{m} c_{u,l} \cdot c_{v,l_{(Rg)}}}{\sqrt{\sum_{l=1}^{m} c_{u,l}^2} \cdot \sqrt{\sum_{l=1}^{m} c_{v,l_{(Rg)}}^2}} \tag{9}$$

where $c_{v,l_{(Rg)}}$ are vectors of proximate users v over locations l in range distance Rg. To extract the entire check-in history of these users we use a range query.

Definition 2 Extract proximate users-POIs history. Given check-in matrix CL, the range query with radius Rg at a given location l, is defined as:

$$c_{u,l_{(Rg)}} = \{c_{v,l} | c_{v,l} \in CL \wedge d(c_{u,l}, c_{v,l}) \leq Rg\} \tag{10}$$

This query returns the entire check-ins history of the proximate users at a distance less or equal to range Rg. Fig. 3 illustrates an example of acquired check-in history of all users within a range of 2000 m.

3.2.2. Incorporating social influence

Regarding the social influence, our model extends the user-based CF model in two ways: (i) it leverages the proximity factor when computing the similarity of two users, and (ii) it explores the reviews about the proximate business stores, rather than the entire review history of all business stores. Given a user u and her/his location l, the prediction score that the user would be interested in an unread review is:

Fig. 3. Proximate users and their check-in history in distance Rg = 2000 m

$$\widehat{c_{u,r}}^{(Social)} = \frac{\sum w_{u,v}^{(Social)} \cdot c_{v,r_{(Rg)}}}{\sum w_{u,v}^{(Social)}}. \tag{11}$$

$w_{u,v}^{(Review)}$ is the similarity matrix of users u and v according to their review history, where we assume that a user would be interested only in reviews of proximate business stores and extract reviews of proximate stores at a distance Rg:

$$w_{u,v}^{(Social)} = \frac{\sum_{r=1}^{k} c_{u,r} \cdot c_{u,r_{(Rg)}}}{\sqrt{\sum_{r=1}^{k} c_{u,r}^2} \cdot \sqrt{\sum_{r=1}^{k} c_{v,r_{(Rg)}}^2}} \tag{12}$$

where $c_{u,r_{(Rg)}}$ is as in Eq. (10) and returns the entire review history of proximate users whose distance is less or equal to the given range Rg. We point out that in our approach, each word in a review is regarded as a node and thus the similarity of the reviews according to the lexical analysis of the words is emphasized. In the approach discussed in Section 3.1.1 the review was regarded as a node and thus the emphasis was on the spatial proximity of the reviews.

3.2.3. Unified model for POIs recommendations

As in Section 3.1.3, we map Eqs. (8) and (11) in the scale $[0-1]$ as follows:

$$\begin{aligned} \widehat{C_{u,l}}^{(Spatial)} &= \frac{\widehat{c_{u,l}}^{(Spatial)} - min(\widehat{c_{u,l}}^{(Spatial)})}{max(\widehat{c_{u,l}}^{(Spatial)}) - min(\widehat{c_{u,l}}^{(Spatial)})} \\ \widehat{C_{u,r}}^{(Social)} &= \frac{\widehat{c_{u,r}}^{(Social)} - min(\widehat{C_{u,r}}^{(Social)})}{max(\widehat{c_{u,r}}^{(Social)}) - min(\widehat{c_{u,r}}^{(Social)})}. \end{aligned} \tag{13}$$

We amalgamate Eq. (13) into a single formula that computes the check-in scores for unvisited POIs and recommends the top-N POIs:

$$\widehat{c_{u,l}} = a \cdot \widehat{C_{u,l}^{(Spatial)}} + (1-a) \cdot \widehat{C_{u,r}^{(Social)}} \tag{14}$$

3.3. Incorporating temporal influence

As mentioned, user's daily schedules show a periodic behavior depending on time (Cho et al., 2011). This daily/weekly/ monthly routine indicates that users tend to check-in locations at a certain region during the same time interval. We divide the dataset according to time intervals so that there is enough data collected for each time interval. For example, in our experiments (see next

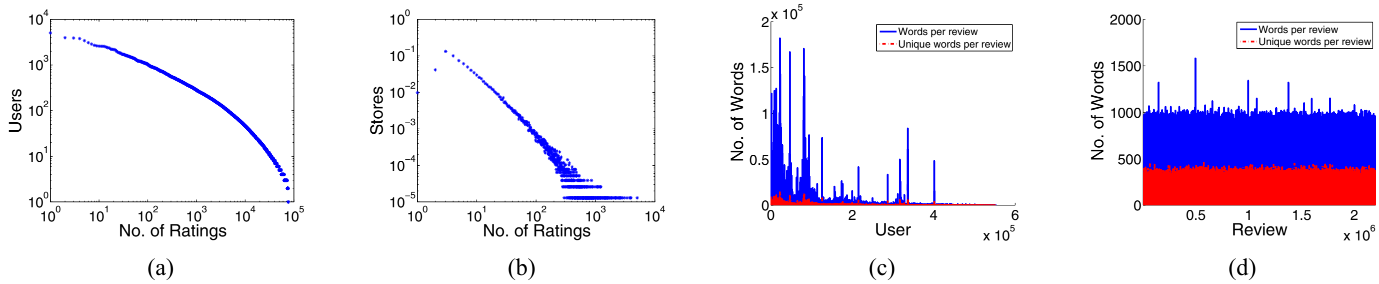


Fig. 4. Distribution diagrams for Yelp! dataset [(a) Number or ratings per user, (b) Number of ratings per business store, (c) Number of words (blue line) and number of unique words (red line) of all the reviews of each user, (d) Number of words (blue line) and the number of unique words (red line) of each review]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

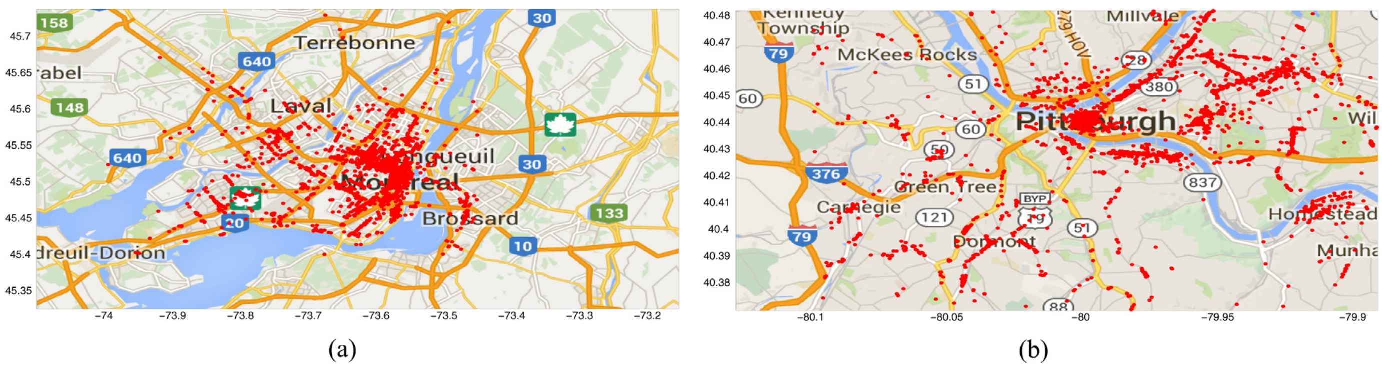


Fig. 5. Distribution check-ins at (a) Montreal, and (b) Pittsburg.

subsection) we decided to split the dataset into monthly time intervals because of the sparsity of the 11 years collected dataset. If the dataset is denser, weekly or even daily time intervals could be valid choices. The dataset used must give the exact date and time for each check-in (i.e. 19-1-2015 18:50) for the accurate incorporation of the temporal influence. Review recommendations and POI recommendations for all time intervals are obtained via the application of Eqs. (7) and (14). The resulting models are denoted by $USTT_r$ (for review recommendations) and $USTT_c$ (for POI recommendations). We point out that the models above indicate trends according to time intervals. For example, a review may be related to a restaurant that employs discounts on a particular day of the week; similarly, a review may be related to a POI in a festival that lasts for three weeks. In such cases, the importance of the review is higher during this particular time interval than any other time interval.

4. Experimental evaluation

4.1. Data sets

We performed our experiments using a real world dataset acquired from Yelp! challenge[4] (Liu, Shang, Wang, Ren, & Han, 2015; McAuley & Leskovec, 2013). This dataset contains information of user ratings and reviews for businesses of 10 cities: Charlotte, Edinburgh, Karlsruhe, Las Vegas, Madison, Montreal, Phoenix, Pittsburgh, Urbana-Champaign and Waterloo. The dataset characteristics are presented in Table 2.

Fig. 4(a) presents the number of the ratings per user in log-log scale, whereas Fig. 4(b) presents the number of the ratings per business in log-log scale as well. Further, Fig. 4c) illustrates the number of words and the number of unique words per user's review of all reviews. Finally, Fig. 4(d) illustrates the number of

Table 2
Datasets specifications.

Dataset	Yelp! dataset
Users	552,338
Businesses	77,078
Rating	2,225,204
Reviews	2,225,204
Period	12 Oct. 2004–24 Dec. 2015

words and the number of unique words per each review individually. As expected, the dataset follows a power law distribution; i.e. there are few business stores/users with many ratings, whereas there are lots of stores/users with few ratings. Also, there are few users that use a large number of words for their reviews, whereas the majority of users use few words for their reviews. Lastly, the number of words of each review individually, follows the normal distribution.

Fig. 5 illustrates the location distribution of Yelp dataset in terms of geographical latitude and longitude. In particular, Fig. 5(a) and (b) depict the check-ins performed in Montreal and Pittsburgh, respectively. Notice that in both figures the reviews for POIs are denser in the center of the city and sparse at the suburbs.

4.2. Compared algorithms

Here, we examine the accuracy performance of our models against 10 methods. More specifically, we divide the compared methods into two groups of experiments, one for each task.

- For the review recommendation we compare our method denoted $USTT_r$ against:
  1. U_r that is the baseline item-based CF model on the reviews (Karypis, 2001),
  2. US_r that is the base line U_r model related to the spatial influence of user reviews (Section 3.1.1),

[4] https://www.yelp.com/dataset_challenge

Table 3
Features supported per model.

	Algorithm	Spatial	Textual	Temporal
Review Rec.	U_r	–	–	–
	US_r	✓	–	–
	UT_r	–	✓	–
	UST_r	✓	✓	–
	$USTT_r$	✓	✓	✓
POI Rec.	U_c	–	–	–
	US_c	✓	–	–
	UT_c	–	✓	–
	UST_c	✓	✓	–
	UTF	✓	–	✓
	$UTE+SE$	✓	–	✓
	$USTT_c$	✓	✓	✓

3. UT_r that is the base line U_r related to the textual influence of the reviews (Section 3.1.2), and
4. UST_r that combines both the US_r and the UT_r models (Section 3.1.3).

- For the POI recommendation task we compare our unified method denoted as $USTT_c$ against:
  1. U_c that is the baseline user-based CF model on the entire user check-in history (Karypis, 2001),
  2. US_c that is the base line U_c related to the spatial influence of the user check-in history (Section 3.2.1),
  3. UT_c that is the base line U_c related to the user review influence respectively (Section 3.2.2),
  4. UST_c that combines both the US_c and the UT_c models (Section 3.2.3),
  5. UTF is a variation of the user-based CF model using a time function (Ding & Li, 2005). The distinction of the base line model is that it weights the check-ins of similar users according to the gaps between the time periods from the current time period. Thus, the larger the time interval from the current time period, the less usefulness of the particular check-in.
  6. $UTE+SE$, another variation of the user-based CF technique which explores the spatial and temporal influences of POIs (Yuan et al., 2013). This method uses a Bayes probabilistic model to make predictions of unvisited POIs. In addition, it extends this spatial model by incorporating the temporal influence of the check-ins to further improve the recommendation accuracy.

Table 3 presents the features supported by each model for the review and recommendation tasks. We point out that even though some of the models for POI recommendations use spatial or the temporal information (e.g. UTF, $UTE+SE$) none of them uses the information from the textual network.

4.3. Evaluation protocol

For the review recommendation task, we divide the target review into two sets: (i) the training set E_R^T is treated as known information, and (ii) the probe set E_R^P is used for testing. Therefore, for a target review we generate the recommendations based only on the reviews in E_R^T. Similarly, for the POIs recommendation task, we divide the check-ins of each target user into two sets: (i) the training set E_L^T, and, (ii) the probe set E_L^P. Thus, we generate the recommendations based only on the POIs in E_L^T.

Each experiment has been repeated 30 times, where each time a different training set is randomly selected. The presented measurements, based on two-tailed t-test, are statistically significant at 0.05 level. All algorithms have the task to predict either the top-N reviews to be read or the top-N unvisited proximate POIs in the probe sets. We use the precision/recall metrics for review and POI recommendations (Davis & Goadrich, 2006):

- *Precision*: Ratio of the number of relevant entities in the top-N list over N.

$$Precision = \frac{Relevant \cap Retrieved}{Retrieved} \tag{15}$$

- *Recall*: Ratio of the number of relevant entities in the top-N list over the total number of relevant entities.

$$Recall = \frac{Relevant \cap Retrieved}{Relevant} \tag{16}$$

4.4. Temporal analysis

We examine the impact of time intervals for the $USTT_r$ and $USTT_c$ models. In particular, we evaluate the influence of the time dimension in terms of precision and recall for both models. Fig. 6(a) presents the number of the reviews performed in a range of 0.5, 1, and 2 km of the entire dataset. Analogously, Fig. 6(b)–(e) present the number of reviews performed in the same ranges but for varying time intervals: from 2 and 3 weeks, to 1 and 2 months, respectively. It can be seen how many reviews correspond at each case. We note that the number of reviews increases as the time interval increases. For each time interval we measured the performance of the $USTT_r$ and $USTT_c$ models, in terms of precision and recall, to determine the *best* time interval and range distance for the specific dataset.

Regarding the review recommendation task, we notice that the precision decreases smoothly and that the recall increases with increasing time intervals. As shown in Fig. 7, both metrics achieve lower values after the time interval of one month, which means that the textual influence is not as important for longer time periods. Fig. 7(a) depicts the performance of $USTT_r$ in terms of precision for top-N values for different range distances. The ranges are set to 0.5, 1, and 2 km per each time interval to determine the willingness of users to read a review about a business store in close or longer distances. The experiments performed for different time intervals show that users are willing to read reviews about proximate stores in a range of 2 km. Based on this finding for the review recommendation task, we set the range distance at 2 km and the time interval at 1 month. Moreover, comparing the results for time intervals of 1 month, 2 and 3 weeks at a range of 2 km for the top-5 recommendations, we have an average improvement of 15% and 7%, respectively. Notice that the performance drops for longer time intervals. The same conclusions hold with respect to precision and recall for different parameters (e.g. range and time intervals) as shown in Fig. 7.

We also perform a temporal analysis for the POI recommendation task (see Fig. 8). In particular, we examine the performance of the $USTT_c$ model in terms of precision and recall for different ranges and time intervals. We focus on the willingness of users to attend proximate or more distant POIs. Reviews in a range of 2 km within a 1 month period seem to influence users the most. We note that in terms of precision, there is a small deviation in the performance of the method when different parameter values are used. For example, for the top-5 recommendations and the 2 km range, the 1 month interval compared to the 2 and 3 week time intervals shows an improvement of 12% and 8%, respectively. The textual influence of longer time intervals in the same ranges is less helpful as there are many more reviews (see Fig. 6(e)) that are not relevant to user preferences. Thus, when we expand the time interval, we incorporate reviews that are not relevant. The findings of the temporal analysis indicate that proximity of the textual and temporal dimensions are closely related to user behaviors.

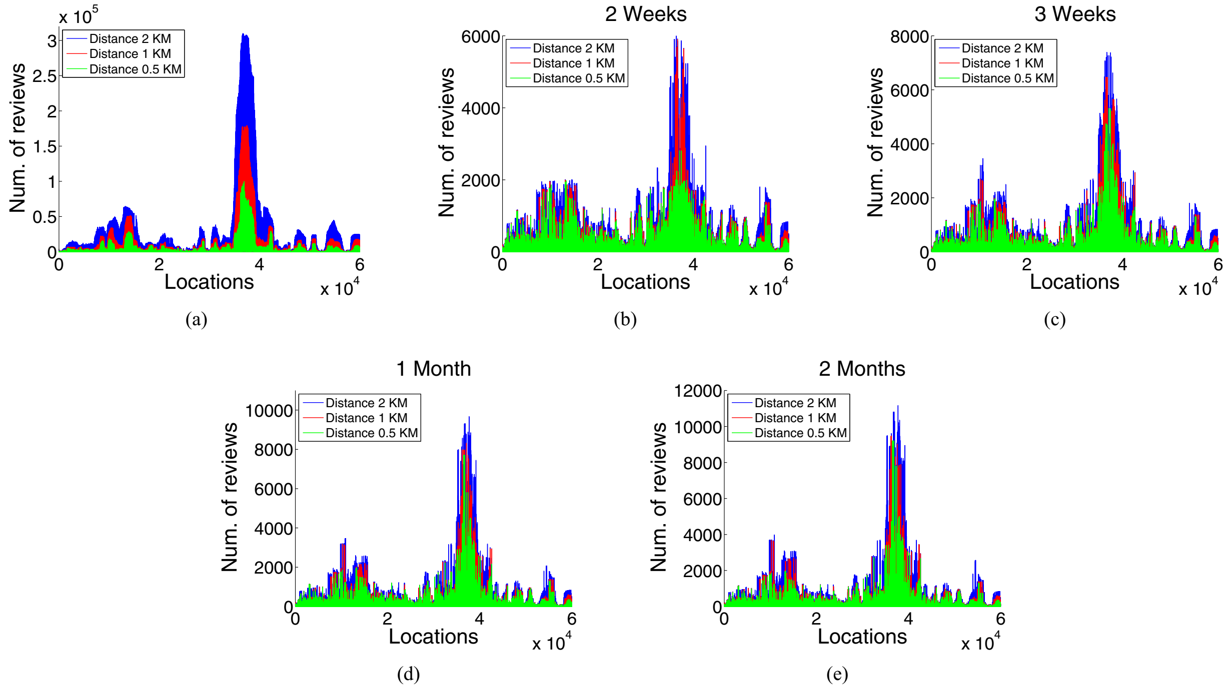


Fig. 6. Distribution diagrams for Yelp! dataset [(a) Number of reviews per location, (b) Number of reviews per location for two weeks interval, (c) Number of reviews per location for three weeks interval, (d) Number of reviews per location for one month interval, and (e) Number of reviews per location for two month interval] in range 0.5, 1, and 2 Km.

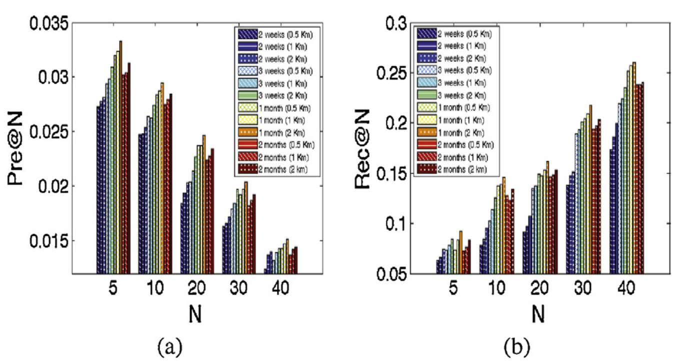


Fig. 7. Sensitivity analysis of $USTT_r$ method for the review recommendations in terms of: (a) precision, and (b) recall for the top-N predictions at 2,3,4,8 week time intervals.

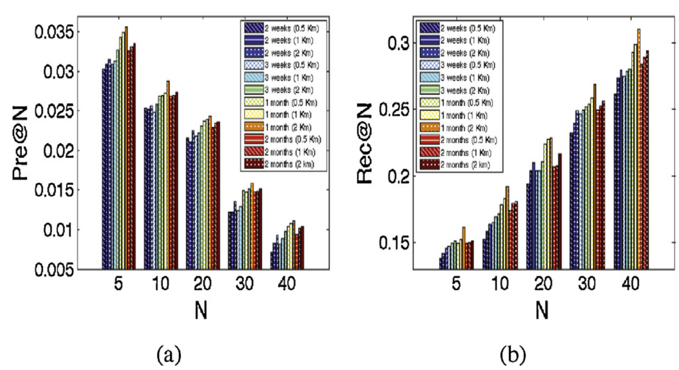


Fig. 8. Sensitivity analysis of $USTT_c$ method for the POI recommendations in terms of: (a) precision, and (b) recall for the top-N predictions at 2,3,4,8 week time intervals.

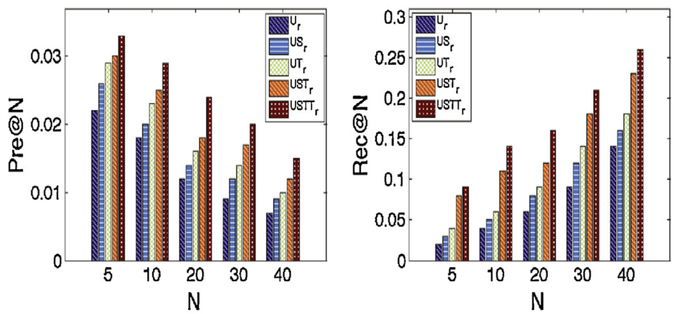
4.5. Sensitivity analysis

In this section, we perform a sensitivity analysis in terms of precision and recall by varying the top-N recommendations as depicted in Fig. 9. We set the time interval equal to 1 month and the range equal to 2 km. As expected, the precision of all models decays with increasing number of recommended reviews or POIs. This is reasonable since precision drops as we increase the number of top-N recommendations, whereas at the same time recall increases.

For the review recommendation task, notice that our method $USTT_r$ outperforms the other methods. Also, notice that US_r outperforms the base line model by 16% in terms of precision for the top-5, whereas UT_r further improves this model by 10%, as shown in Fig. 9(a). Moreover, the model that combines both spatial influence and the textual influence increases the precision by 6% and 14%, respectively. Finally, $USTT_r$ outperforms the second best model by 15%, since it exploits the influence of both factors at each time interval separately. This result indicates that time is an essential factor during review recommendations. Simply expressed, words are more important during a particular time interval rather than taking all time intervals together.

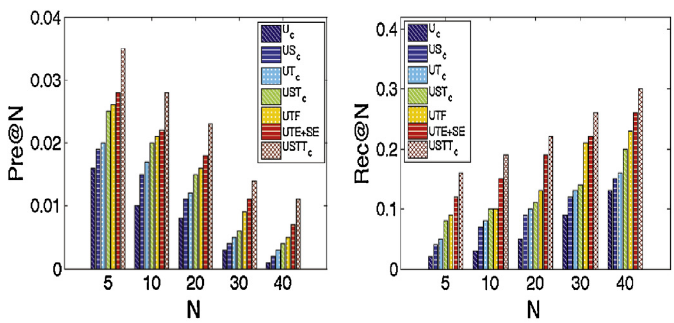
As shown in Fig. 9(b), $USTT_r$ outperforms the second best method by 11% in terms of recall, that is the ratio of number of the relevant recommendations of the top-5 to the total number of relevant ones. Moreover, the improvement over the baseline method U_r is 34%, since it does not consider any of the spatial influence, the textual influence, or the time dimension during the recommendation process. Other findings show that methods considering the spatial and the textual influence, as well as the time dimension always exhibits better results for all measures than methods that consider these dimensions separately. Finally, notice that as long as we enhance the dimensions, we succeed higher values of precision and recall.

For the POI recommendation task, the performance of $USTT_c$ is significantly higher than all other methods, as illustrated in Fig. 10. In particular, compared to the state-of-the-art methods $UTE + SE$



(a) (b)

Fig. 9. Sensitivity analysis of all methods for the review recommendations in terms of: (a) precision, and (b) recall for the top-N predictions.



(a) (b)

Fig. 10. Sensitivity analysis of all methods for the POI recommendations in terms of (a) precision and (b) recall for the top-N predictions.

Table 4
Average % improvement in terms of precision of the top-N recommendations for different time intervals against $UTE + SE$ model.

Period	top-N				
	5	10	20	30	40
2 weeks	9.7	14.4	20.1	22.4	24.7
3 weeks	13.2	18.6	22.4	26.3	28.4
1 month	20.2	23.4	26.1	30.6	36.1
2 months	12.1	18.7	22.8	26.6	28.9

and *UTF*, our method $USTT_c$ prevails respectively by 20% and 26% in terms of precision for the top-5 recommended POIs. The same holds with respect to recall, where $USTT_c$ outperforms the previous methods by 25% and 44%, respectively. Notice that $UTE + SE$ considers only the spatial and the temporal influence but misses the user review influence. Similarly, *UTF* considers only the temporal dimension through a time function that weights the importance of the check-ins but ignores the influence of user reviews. Compared to UST_c which examines both the spatial and the textual influence, the improvement is 29% and 47% in terms of precision and recall, respectively. Again, the results point that proximity and time dimensions further improve the accuracy of the recommendations. In particular, users are highly influenced by proximate users and reviews about proximate locations before reaching a POI. Also, the influence of the particular time period is more important than all-time intervals together. Thus, considering only the proximate check-ins and reviews during particular time intervals, we can see a significant improvement of the accuracy of the recommendations.

Table 4 presents the average improvement of $USTT_c$ over the $UTE + SE$ model (Yuan et al., 2013) in terms of precision for the top-N recommendations at different time intervals. While time interval increases from 2 weeks to 1 month, the improvement of precision increases as well, for different N values. Notice that even when we set the time interval to 2 months, our method prevails the $UTE + SE$ model, but the precision value is gradually decreasing compared to the higher values of precision of 1 month time interval.

Conclusively, among all methods compared for both tasks, $USTT_r$ and $USTT_c$ always achieve the best results with respect to precision and recall at different N values for Yelp dataset.

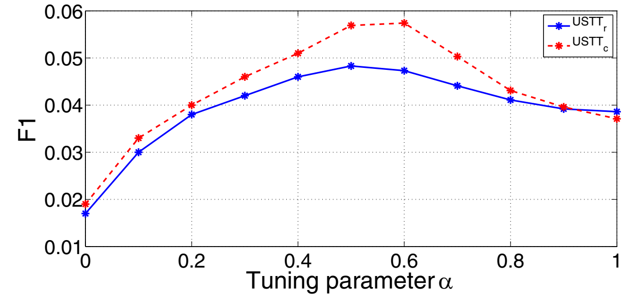


Fig. 11. Tuning Parameter α

4.6. Tuning parameter α

In our experiments we have also used the $F1$ metric, which evaluates accuracy as an amalgamation of precision and recall, as shown in Eq. (17). Since $F1$ is a weighted average of these metrics, the closer the values of $F1$ to 100%, the higher the contribution of particular structural components of the method.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{17}$$

Both our unified models, $USTT_r$ and $USTT_c$, use a weight parameter α to balance the influence of the two structural components of Eqs. (7) and (14). In particular, for $USTT_r$ that provides review recommendation the parameter α controls the weights of the spatial influence of user reviews $\widehat{C_{u,r}}^{(Review)}$ and the textual influence among the reviews $\widehat{C_{u,r}}^{(Textual)}$. Similarly, for $USTT_c$ that provides POI recommendation the parameter α controls the weights of the spatial influence $\widehat{C_{u,l}^{(Spatial)}}$ and social influence $\widehat{C_{u,r}^{(Review)}}$ regarding similar users' reviews. Fig. 11 shows the value of $F1$ as a function of the tuning parameter α value for the two methods $USTT_r$ and $USTT_c$ tested on the Yelp dataset. It is revealed that $F1$ is optimized for values of α in the range $0.5 - 0.6$.

5. Conclusions

Nowadays, the availability of user check-ins and reviews in large volumes has given the opportunity for more accurate recommendations. Several models have been proposed in the literature to improve the recommendation accuracy; however, they have the following drawbacks: (i) models that handle POIs as conventional items do not capture the influence of the geographical proximity, (ii) models that incorporate textual influence as side information ignore the spatial proximity of these reviews, and (iii) models that capture temporal dynamics ignore both the spatial and the textual dimensions.

With this paper we introduce two novel unified models for POI and review recommendations to overcome the above problems by combining the spatial, the textual, and the temporal dimensions. To the best of our knowledge, this is the first work emphasizing in review recommendations (i.e. recommending the entire review posted by other users). Regarding the review recommendations task, we explore the spatial influence of user reviews and the textual influence of the reviews into one unified model. Then, we further extend this model to consider the temporal dimension by examining the influence of both factors in different time intervals.

Similarly, regarding POI recommendations, we explore the spatial influence of user check-in history and the social influence of user reviews into one unified model. Further, we extend this model to consider the temporal dimension by examining the influence of both factors in different time intervals.

Both models extend the item-based CF and the user-based CF model by leveraging the proximity factor when computing the similarity among the users, and by considering only the reviews/check-ins that these proximate users have made, rather than the entire review/check-in history of all users, respectively.

We have evaluated the accuracy performance of our models to measure the improvement of both types of recommendations. With respect to precision, $USTT_r$ demonstrated a significantly improved accuracy against models that do not consider the temporal dimension. Moreover, $USTT_c$ outperforms the state-of-the-art $UTE + SE$ model, which explores the spatial and the temporal influence of POIs, by 20%, since it takes simultaneously under consideration the spatial, the textual and the temporal dimensions. The improvement is due to exploration of the textual influence of the reviews proximity to the user current location in each time interval separately. Also, for the same reasons our method prevails in terms of recall over the previous methods by 11% and 25%, respectively. Findings of our work indicate that users' are influenced by others whose behavior is similar to them; thus, users who employ the same vocabulary can be classified into the same group with the target user. Also, experiments result in that user check-in behavior is different daily, weekly or monthly. Moreover, users who check-in proximate locations during the same time period may also have the same interests.

Furthermore, the textual correlation with the locations and category of the business stores opens future directions towards extending both models by enhancing the multimodal information about user-item interactions with side information about the business stores. Moreover, we intent to measure both our models online with real time data. Finally, we will create a friend recommendation model with respect to the temporal proximity and the textual closeness of the users'.

Acknowledgements

Thanks are due to our students Konstantinos Dalakouras and Iraklis Bekiaris, who created a mobile application based on the material of this paper. Their app helped in having a better intuition on the subject of time-aware spatio-textual recommenders.

References

Adomavicius, G., & Tuzhilin, A. (2008). Context-aware recommender systems. In *Proceedings of the acm conference on recommender systems (recsys)* (pp. 335–336).

Balabanović, M., & Shoham, Y. (1997). Fab: Content-based, collaborative recommendation. *Communications of the ACM, 40*(3), 66–72.

Ben Nejma, G., Roose, P., Gensel, J., Dalmau, M., & Ghorbali, M. A. (2013). Taldea: A tool for fostering spontaneous communities. In *Proceedings 13th icwe workshops: Composableweb, qwe, mdwe, dmssw, emotions, cse, ssn, and phd symposium (revised selected papers)* (pp. 293–301).

Bravo-Torres, J. F., López-Nores, M., Blanco-Fernández, Y., & Pazos-Arias, J. J. (2013). Leveraging short-lived social networks in vehicular environments. In *Proceedings 2nd international conference on future generation communication technologies (fgct)* (pp. 196–200).

Bravo-Torres, J. F., López-Nores, M., Blanco-Fernández, Y., & Pazos-Arias, J. J. (2014). A platform to exploit short-lived relationships among mobile users: A case of collective immersive learning. In *Proceedings 20th international conference on information & software technologies (icist)* (pp. 384–395).

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings 14th conference on uncertainty in artificial intelligence (uai)* (pp. 43–52).

Cao, X., Cong, G., & Jensen, C. S. (2010). Mining significant semantic locations from GPS data. *Proceedings VLDB, 3*(1-2), 1009–1020.

Cho, E., Myers, S. A., & Leskovec, J. (2011). Friendship and mobility: User movement in location-based social networks. In *Proceedings 17th acm international conference on knowledge discovery & data mining (kdd)* (pp. 1082–1090).

Davis, J., & Goadrich, M. (2006). The relationship between precision-recall and roc curves. In *Proceedings 23rd international conference on machine learning (icml)* (pp. 233–240).

Deng, S., Huang, L., & Xu, G. (2014). Social network-based service recommendation with trust enhancement. *Expert Systems with Applications, 41*(18), 8075–8084.

Deshpande, M., & Karypis, G. (2004). Item-based top-*n* recommendation algorithms. *ACM Transactions on Information Systems, 22*(1), 143–177.

Ding, Y., & Li, X. (2005). Time weight collaborative filtering. In *Proceedings 14th acm international conference on information & knowledge management (cikm)* (pp. 485–492).

Esparza, S. G., O'Mahony, M. P., & Smyth, B. (2011). Effective product recommendation using the real-time web. In *Proceedings 13th sgai international conference on innovative techniques & applications of artificial intelligence (ai)* (pp. 5–18).

Wang, H., Terrovitis, M., & N. , M. (2013). Location recommendation in location-based social networks using user check-in data. In *Proceedings acm conference on international conference on advances in geographic information systems (sigspatial)*.

Hu, B., & Ester, M. (2013). Spatial topic modeling in online social media for location recommendation. In *Proceedings 7th acm conference on recommender systems (recsys)* (pp. 25–32).

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings 10th acm international conference on knowledge discovery & data mining (kdd)* (pp. 168–177).

Jin, R., Chai, J. Y., & Si, L. (2004). An automatic weighting scheme for collaborative filtering. In *Proceedings 27th international acm conference on research & development in information retrieval (sigir)* (pp. 337–344).

Karypis, G. (2001). Evaluation of item-based top-*n* recommendation algorithms. In *Proceedings 10th international conference on information & knowledge management (cikm)* (pp. 247–254).

Koren, Y. (2009). Collaborative filtering with temporal dynamics. In *Proceedings 15th acm international conference on knowledge discovery & data mining (kdd)* (pp. 447–456).

Lathia, N., Hailes, S., Capra, L., & Amatriain, X. (2010). Temporal diversity in recommender systems. In *Proceedings 33rd international acm conference on research & development in information retrieval (sigir)* (pp. 210–217).

Lee, K., & Lee, K. (2015). Escaping your comfort zone: A graph-based recommender system for finding novel recommendations among relevant items. *Expert Systems with Applications, 42*(10), 4851–4858.

Lee, T. Q., Park, Y., & Park, Y.-T. (2008). A time-based approach to effective recommender systems using implicit feedback. *Expert Systems with Applications, 34*(4), 3055–3062.

Li, C., Zhao, P., Wu, J., Mao, J., & Cui, Z. (2015). An item-based collaborative filtering framework based on preferences of global users. In *Proceedings 4th international conference on computer engineering & networks (cenet): 335* (pp. 1113–1120).

Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., & Ma, W.-Y. (2008). Mining user similarity based on location history. In *Proceedings 16th acm sigspatial international conference on advances in geographic information systems (gis)*.

Li, X., Xu, G., Chen, E., & Zong, Y. (2015). Learning recency based comparative choice towards point-of-interest recommendation. *Expert Systems with Applications, 42*(9), 4274–4283.

Liu, B., Fu, Y., Yao, Z., & Xiong, H. (2013). Learning geographical preferences for point-of-interest recommendation. In *Proceedings 19th acm international conference on knowledge discovery & data mining (kdd)* (pp. 1043–1051).

Liu, B., & Xiong, H. (2013). Point-of-interest recommendation in location based social networks with topic and location awareness. In *Proceedings siam international conference on data mining (sdm)* (pp. 396–404).

Liu, F., & Lee, H. J. (2010). Use of social network information to enhance collaborative filtering performance. *Expert Systems with Applications, 37*(7), 4772–4778.

Liu, J., Shang, J., Wang, C., Ren, X., & Han, J. (2015). Mining quality phrases from massive text corpora. In *Proceedings acm international conference on management of data (sigmod)* (pp. 1729–1744).

Liu, L., Xu, J., Liao, S. S., & Chen, H. (2014). A real-time personalized route recommendation system for self-drive tourists based on vehicle to vehicle communication. *Expert Systems with Applications, 41*(7), 3409–3417.

Lops, P., de Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In *Recommender systems handbook* (pp. 73–105). Springer.

Lu, Z., Savas, B., Tang, W., & Dhillon, I. S. (2010). Supervised link prediction using multiple sources. In *Proceedings 10th ieee international conference on data mining (icdm)* (pp. 923–928).

Ma, H., Lyu, M. R., & King, I. (2009). Learning to recommend with trust and distrust relationships. In *Proceedings 3rd acm conference on recommender systems (recsys)* (pp. 189–196).

McAuley, J., & Leskovec, J. (2013). Hidden factors and hidden topics: Understanding rating dimensions with review text. In *Proceedings 7th acm conference on recommender systems (recsys)* (pp. 165–172).

Nejma, G. B., Roose, P., Dalmau, M., & Gensel, J. (2015). Service discovery for spontaneous communities in pervasive environments. In *Proceedings 16th international conference on web information systems engineering (wise), part ii* (pp. 337–347).

Pirasteh, P., Jung, J. J., & Hwang, D. (2014). Item-based collaborative filtering with attribute correlation: A case study on movie recommendation. In *Proceedings 6th asian conference intelligent information & database systems (aciids), part ii* (pp. 245–252).

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management, 24*(5), 513–523.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings 10th international conference on world wide web (www)* (pp. 285–295).

Smaldone, S., Han, L., Shankar, P., & Iftode, L. (2008). Roadspeak: Enabling voice chat on roadways using vehicular social networks. In *Proceedings 1st workshop on social network systems (socialnets)* (pp. 43–48).

Srba, I., & Bielikova, M. (2015). Dynamic group formation as an approach to collaborative learning support. *IEEE Transactions on Learning Technologies, 8*(2), 173–186.

Tang, H., Tan, S., & Cheng, X. (2009). A survey on sentiment detection of reviews. *Expert Systems with Applications, 36*(7), 10760–10773.

Vansteenwegen, P., Souffriau, W., Berghe, G. V., & Oudheusden, D. V. (2011). The city trip planner: An expert system for tourists. *Expert Systems with Applications, 38*(6), 6540–6546.

Vasuki, V., Natarajan, N., Lu, Z., Savas, B., & Dhillon, I. (2011). Scalable affiliation recommendation using auxiliary networks. *ACM Transaction on Intelligent Systems & Technology, 3*(1).

Wang, J., de Vries, A. P., & Reinders, M. J. T. (2006). Unifying user-based and item-based collaborative filtering approaches by similarity fusion. In *Proceedings 29th international acm conference on research & development in information retrieval (sigir)* (pp. 501–508).

Xiong, L., Chen, X., Huang, T.-K., Schneider, J., & Carbonell, J. G. (2010). Temporal collaborative filtering with bayesian probabilistic tensor factorization. In *Proceedings 10th siam international conference on data mining (sdm)* (pp. 211–222).

Yang, X., Guo, Y., Liu, Y., & Steck, H. (2014). A survey of collaborative filtering based social recommender systems. *Computer Communications, 41*, 1–10.

Yao, L., Sheng, Q. Z., Qin, Y., Wang, X., Shemshadi, A., & He, Q. (2015). Context-aware point-of-interest recommendation using tensor factorization with social regularization. In *Proceedings 38th international acm conference on research & development in information retrieval (sigir)* (pp. 1007–1010).

Ye, M., Yin, P., Lee, W.-C., & Lee, D.-L. (2011). Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings 34th international acm conference on research & development in information retrieval (sigir)* (pp. 325–334).

Yuan, Q., Cong, G., Ma, Z., Sun, A., & Thalmann, N. M. (2013). Time-aware point-of-interest recommendation. In *Proceedings 36th international acm conference on research & development in information retrieval (sigir)* (pp. 363–372).

Zhang, C., Wang, K., Yu, H., Sun, J., & Lim, E. (2014). Latent factor transition for dynamic collaborative filtering. In *Proceedings 14th siam international conference on data mining (sdm)* (pp. 452–460).

Zhang, Z., & Varadarajan, B. (2006). Utility scoring of product reviews. In *Proceedings 15th acm international conference on information & knowledge management (cikm)* (pp. 51–57).

Zhao, Z. D., & Shang, M. S. (2010). User-based collaborative-filtering recommendation algorithms on Hadoop. In *Proceedings 3rd international conference on knowledge discovery & data mining (kdd)* (pp. 478–481).

Zheng, V. W., Cao, B., Zheng, Y., Xie, X., & Yang, Q. (2010). Collaborative filtering meets mobile recommendation: A user-centered approach. In *Proceedings 24th aaai conference on artificial intelligence (aaai)* (pp. 236–241).

Zheng, Y., Zhang, L., Xie, X., & Ma, W.-Y. (2009). Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings 18th international conference on world wide web (www)* (pp. 791–800).