Temporal People-to-People Recommendation on Social Networks with Sentiment-based Matrix Factorization

Davide Feltoni Gurini, Fabio Gasparetti, Alessandro Micarelli, Giuseppe Sansonetti

Roma Tre University, Via della Vasca Navale 79, 00146 Rome, Italy

Abstract

Nowadays, the exponential advancement of social networks is creating new application areas for recommender systems (RSs). People-to-people RSs aim to exploit user's interests for suggesting relevant people to follow. However, traditional recommenders do not consider that people may share similar interests but might have different feelings or opinions about them. In this paper we propose a novel recommendation engine which relies on the identification of semantic attitudes, that is, sentiment, volume, and objectivity extracted from user-generated content. In order to do this at large-scale on traditional social networks, we devise a three-dimensional matrix factorization, one for each attitude. Potential temporal alterations of users' attitudes are also taken into consideration in the factorization model. Extensive offline experiments on different real world datasets, reveal the benefits of the proposed approach compared with some state-of-the-art techniques.

Keywords: People-to-People Recommendation, Sentiment Analysis, Matrix Factorization

1. Introduction

Microblogging platforms are one of the most versatile and popular technologies on the Internet today. For instance, Twitter sees over 500 million microposts (or *tweets*) published every day on a huge variety of topics, with spikes of more than 100 thousands tweets per second when particular events occur [1]. With the proliferation of user-generated content such as reviews, discussion forums, blogs, and tweets, detecting sentiments and opinions from the Web is becoming an increasingly widespread form of data interpretation. In particular, sentiment analysis aims to understand subjective information, such as opinions, points of view, and feelings expressed by users in the content they generate.

People-to-people recommendation is an important application in these platforms. Almost all the services are capable of recommending interesting users to follow. However, this recommendation task is not easy due to huge graphs of social ties and fast changing contents that must be analyzed. In this scenario, simple people recommendation algorithms based on content similarity and popularity paradigms are usually considered, at the expense of the recommendation accuracy.

In this paper, we propose a novel people-to-people recommender system that takes into account the users' attitudes towards discussed topics. The proposed recommender enables us to leverage users' attitudes such as sentiment, volume, and objectivity extracted from the semantics of tweets, define a *sentiment-volume-objectivity (SVO)* function, and exploit such knowledge to suggest relevant people to follow. The rationale behind this work is that people might have similar interests but different opinions or feelings about them. Therefore, considering the contribution of users' attitudes may yield benefits to people recommendation. For example, two users involved in the discussion about supporting Hillary Clinton for US President are likely to be friends. However, the two users may exhibit the same (both support or oppose Hillary Clinton) or contradictory sentiments (one supports and the other opposes). Therefore, we suppose that the two users are more likely to follow each other in the former case than in the latter.

To handle large-scale social networks, we model this recommendation task using matrix factorization techniques in four steps: (i) build a three-dimensional matrix in which each dimension is represented by a SVO user feature; (ii) learn a latent em-

[1] `https://blog.twitter.com/2013/new-tweets-per-second-record-and-how` (last visited on 20 December 2016)

bedding space from the user-attitudes matrix; (iii) compute the user-user similarity by taking into account the three matrix dimensions; and (iv) recommend to a target user a list of relevant people to follow.

In this work, we address two research questions that arise when approaching the people-to-people recommendation problem:

1. Does content published by users and, in particular, the inferred attitudes, allows for a better identification of potential relationships that exist between them?

2. How does temporal analysis of these attitudes impact the accuracy of the recommendation?

The scientific contributions coming from this paper are: (i) an algorithm for people-to-people recommendation on microblogging platforms that takes advantage of features that represent the users' attitudes on specific topics; (ii) a comparative experimental results of a set of different evaluation metrics, including a range of non-accuracy measures, such as diversity and novelty; (iii) a proof of how the recommendation accuracy can be improved by taking into account the temporal variations of the attitudes expressed by the user; (iv) an evaluation of the proposed algorithm on real world datasets, showing that the considered users' attitudes have unequal correlation with respect to the accuracy of the recommendation, and strongly depend on the topic under consideration.

The rest of the paper is organized as follows: Section 2 introduces the problem formulation. Section 3 describes the recommendation algorithm. Section 4 presents the performed experiments to evaluate the proposed strategy and outlines main results. Section 5 contains a description of some state-of-the-arts approaches. Finally, Sect. 6 reports our conclusions.

2. Problem formulation

In this section, we provide the definition of the people-to-people recommendation problem.

Let $\mathbb{U} = \{u_1, \cdots, u_N\}$ represents the set of users with a valid account on the micro-blogging platform. In our scenario, an adjacency matrix $A^{N \times N}$ represents the explicit ties, where each element $A_{i,j}$ denotes if the user u_i *follows* (or is *friend of*) the user u_j or not, and therefore is usually expressed by a binary value $\{0, 1\}$. Then, let $\overline{\mathbb{U}} = \{u_1, \cdots, u_M\}$ represent the set of candidate users $u_j \in \mathbb{U}$ without an explicit tie with the target user u_i, that is,

$$\overline{\mathbb{U}} = \{\forall u_j \in \mathbb{U} \mid i \neq j \wedge (A_{i,j} = 0 \wedge A_{j,i} = 0)\}$$

Under this setting, the problem can be formulated as follows: given the matrix $A^{N \times N}$, which represents a known set of social relations between N users, define the following function r:

$$r : \mathbb{U} \times \overline{\mathbb{U}} \to [0, 1] \tag{1}$$

such that, given a *target* user u and an adjacency matrix, returns a value between 0 and 1, which expresses the relevance degree of the candidate user u_j for the target user u_i. Based on such value, the system provides the target user with a recommendation list of the top relevant candidates.

First attempts to people-to-people recommendation take advantage of global models and collective classification for the definition of the r function. In other words, they operate on the whole graph of related nodes rather than deriving individual structural and content-based attributes. The problem is therefore seen as the optimization of one global objective function.

Since *link prediction problem* [1, 2] aims at inferring future interactions and missing links on large graphs, various predictors based on the interpersonal social structure (e.g., common neighbors predictor) are also considered for the ranking task.

Our goal is to define the function r by extending the recommendation analysis to relevant information associated with users that can be retrieved by the micro-blogging platform, namely, the timeline consisting of sequences of microposts. In the rest of the paper, we indicate with $\mathbb{T}$ the set of potential microposts that can be published and with $T_u \subset 2^{\mathbb{T}}$ the most recent microposts published by the user u.

Figure 1: Principal steps for the people-to-people recommendation task.

3. The proposed people-to-people recommendation

In this section, we introduce our method for recommendation. A strong correlation exists between the presence of a social tie between two users and the topical similarity of explicit activities of these users in the network [3]. So it is logical investigating the chance of predicting the presence of a tie based on user profile features. The idea behind the proposed approach is that, by taking into account the attitudes, in terms of manifested expressions of favor or disfavor on specific matters, the accuracy of the people-to-people recommender is improved. Multiple steps are demanded to implement the recommendation task, as shown in Fig. 1.

The timeline of users $u_i \in \mathbb{U}$ are first retrieved. A traditional pre-processing of microposts simplifies the identification of relevant features. All characters are converted to lowercase letters and retweet designations (e.g., "RT"), citations, and URLs are removed. Then, text is tokenized into keywords, from which a list of unigram features is created. Traditional stopwords are excluded from the lists.

Micro-blogging services allow users to include metadata tags in the form of keywords followed by the hash symbol #, which are referred to as *hashtags*. By including them in the posts, the author is suggesting them as good candidates in quality of search keys. Popular hashtags often refer to topics that most people are interested in, including breaking events and persistent discussions [4]. For this reason, they are often considered for clustering posts related to specific topics [5, 6].

Let $\mathbb{C}$ denote the set of all possible concepts. Given a micropost τ, we indicate with $\tau^{(\mathbb{C})}$ the subset of concepts $\mathbb{C}$ that are included in τ, identified by extracting the hashtags in τ. By extension, $T_u^{(\mathbb{C})}$ is the set of concepts that are included in the user u's timeline. The so-obtained representation of microposts is subjected to the SVO analysis (see Sects. 3.1 and 3.2), which aims at determining the user's attitude on each topic. Since determining similarities among users who have limited activity on specific topics is a challenging task, the SVO-based analysis is not performed on concepts not appearing in a timeline above a given frequency threshold (i.e., 10 tweets). This procedure is commonly followed when attitudes expressed by large audiences are explored [7].

Each user's timeline is subjected to a text categorization process based on a Support Vector Machine (SVM) algorithm [8], so that one or more categories belonging to the set $\mathbb{K}$ of all possible macro-categories are associated to the user according to the published content. These macro-categories (namely, *world*, *elections*, *business*, *technology*, *entertainment*, *sports*, *science*, and *health*) are similar to the ones of a popular online news aggregator [9]. The training set is built-up by retrieving titles and snippets of each macro-category on the aggregator over a period of one month. We denote with $T_u^{(\mathbb{K})} \subseteq \mathbb{K}$ the macro-categories assigned to the user u.

When the system returns a ranked list of people to follow, the target user u's latent factors are compared with the ones obtained from all users that have debated similar topics. This latter *candidate set* $\overline{\mathbb{U}}$ is built-up from $\mathbb{U}$ as follows:

$$\overline{\mathbb{U}} = \{u' \in \mathbb{U} \mid T_u^{(\mathbb{C})} \cap T_{u'}^{(\mathbb{C})} \neq \emptyset \wedge T_u^{(\mathbb{K})} \cap T_{u'}^{(\mathbb{K})} \neq \emptyset\} \tag{2}$$

so that, the overlap between u and a candidate u' is extended to the set of macro-categories assigned to each user. Details on the implementation of the r function that assigns a rank to each candidate user can be found in Sects. 3.3 and 3.4, whereas the following two sections detail the identification of users' attitudes.

3.1. Sentiment analysis of microposts

Sentiment analysis or opinion mining is formally defined as the computational study of user's attitudes about an entity expressed in a text [10]. Sentiment analysis is a complex task, hence some assumptions are needed. There are multiple granularity levels of sentiment analysis, as explained in [11]: feature-level, entity-level, sentence-level, document-level. Given the limitations of the micropost length (i.e., 140 characters), we consider sentiment analysis at sentence-level, which corresponds to a whole micropost in our domain. Formally, the goal of our sentiment analysis is to define the following function:

$$sa : \mathbb{T} \rightarrow \{s^{(+)}, s^{(-)}, s^{(\emptyset)}\} \tag{3}$$

where the output is composed of three symbols referring to positive, negative and neutral sentiment expressed by the given micropost.

Several approaches have been proposed for the implementation of this function [10] with an average accuracy from 70% to over 82% by means of techniques based on Naïve Bayes (NB) classification [12], a simple model which provides high performance on text categorization. To solve this sentiment analysis task, we devise a multinomial NB model that takes into account multiple features such as (i) unigram features extracted from each post, (ii) negation cues as proposed in [13], (iii) words polarities using the LIWC dictionary [2], and (iv) a part of speech tagger provided by Stanford University [3]. Furthermore, a feature selection based on the salience and entropy measures has also been considered to improve the accuracy of the classifier by filtering less relevant keywords [14]. Maximum likelihood estimate is finally employed for the parameter estimation, with add-1 smoothing utilized for unseen features.

3.2. SVO-based analysis

User u's attitudes toward a given topic are evaluated from the observable activity and its aspects. In the micro-blogging scenario, we aim at representing attitudes towards each concept $c \in T_u^{(\mathbb{C})}$ through the following three factors: sentiment, volume, and objectivity. *Sentiment* represents a feeling or opinion about a concept expressed by the user, and is obtained as follows:

$$f_u^{(s,c)} = norm\left(\frac{n_u^{(c,+)} - n_u^{(c,-)}}{n_u^{(c,+)} + n_u^{(c,-)}}\right) \tag{4}$$

with

$$n_u^{(c,+)} = \left|\{\forall \tau \in T_u \mid s(\tau) = s^{(+)} \wedge \tau^{(\mathbb{C})} \cap \{c\} \neq \emptyset\}\right| \tag{5}$$

$$n_u^{(c,-)} = \left|\{\forall \tau \in T_u \mid s(\tau) = s^{(-)} \wedge \tau^{(\mathbb{C})} \cap \{c\} \neq \emptyset\}\right| \tag{6}$$

where $n_u^{(c,+)}$ and $n_u^{(c,-)}$ are the sums of the positive and negative posts, respectively, written by the user u regarding the concept c. Since the range of values can vary

[2] `http://liwc.net` (last visited on 20 December 2016)

[3] `http://nlp.stanford.edu/software/tagger.shtml` (last visited on 20 December 2016)

widely, the *norm* function scales the values within the $[0, 1]$ and takes on the following expression:

$$norm(x) = \frac{1}{1 + 10^{-x}} \tag{7}$$

The second attribute is *volume* and indicates how frequently the user discusses a concept, and is defined as follows:

$$f_u^{(v,c)} = \frac{n_u^{(c)}}{n_u} \tag{8}$$

where

$$n_u^{(c)} = \sum_{\tau \in T_u} \left|\tau^{(\mathbb{C})} \cap \{c\}\right|, \quad \text{and} \quad n_u = \sum_{\tau \in T_u} \left|\tau^{(\mathbb{C})}\right| \tag{9}$$

The final contribution is *objectivity*, which expresses how many posts about a concept do not contain any positive or negative attitude. It is defined as follows:

$$f_u^{(o,c)} = \frac{n_u^{(c,\emptyset)}}{n_u^{(c,+)} + n_u^{(c,-)} + n_u^{(c,\emptyset)}}$$

where

$$n_u^{(c,\emptyset)} = \left|\{\forall \tau \in T_u \mid s(\tau) = s^{(\emptyset)} \wedge \tau^{(\mathbb{C})} \cap \{c\} \neq \emptyset\}\right|$$

where $n_u^{(c,\emptyset)}$ is the sum of posts without positive or negative attitudes written by the user u concerning the concept c. We are now able to introduce the SVO vector for the user u and concept c, which takes into account the three factors as follows:

$$\overrightarrow{SVO_u^{(c)}} = [f_u^{(s,c)}, f_u^{(v,c)}, f_u^{(o,c)}] \tag{10}$$

3.3. Matrix factorization model

Matrix Factorization (MF) techniques [15] are employed for learning the latent characteristics of users and concepts, and defining an approximation of the r function (see Eq. 1) by modeling the ranking with inner products in that latent space. The goal is factorizing a 2-dimensional matrix into two matrices $P \in \mathbb{R}^{|\mathbb{U}| \times f}$ and $Q \in \mathbb{R}^{|\mathbb{C}| \times f}$ such that PQ^T approximates the initial matrix, that is, minimizes a loss function between observed and predicted values. Each row q_i represents the association strength between a user and the latent characteristics. Similarly, each row p_j represents the strength between a concept and the latent dimensions. In the case of micro-blogging platforms,

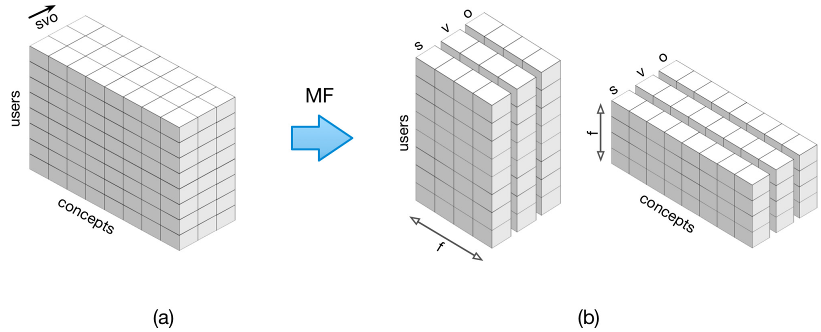

Figure 2: The initial user-concept matrix (a), and the matrices representing the correlation between users, concepts and the latent factors (b).

where the number of users and concepts can be very high, this form of decomposition model allows us to keep bounded the storage requirements by tuning the parameter f (i.e., the number of latent factors) accordingly. In our approach, the SVO-based analysis determines a 3-dimensional vector associated to a pair $(user, concept)$, where the concepts are obtained by analyzing the recent activity on the user's timeline. The observed data forms a ternary relation between users, concepts and SVO features, so we obtain a 3-dimensional sparse matrix $M \in \mathbb{R}^{|\mathbb{U}|\times|\mathbb{C}|\times 3}$, as shown in Fig. 2(a).

Tensor matrix factorization is a generic model framework for recommendations that is able to handle multiple dimensional data taking advantage of the matrix factorization models [16]. Due to multi-dimensional input data, tensor MF seems to be the perfect choice for the dimension reduction task. In our scenario, as proven in Section 4.2, the SVO components representing the user's attitudes have different relevance in the recommendation process according to the category of topics under consideration. For this reason, we decide to perform three MF models, each associated with one of the SVO components, keeping the recommendation process distinct w.r.t each component. For the sake of clarity, we indicate with $P^{(s)}$, $P^{(v)}$ and $P^{(o)}$ the three matrices obtained by the MF model considering the S, V and O component of the SVO vector, respectively, and similarly, we obtain three matrices $Q^{(s)}$, $Q^{(v)}$ and $Q^{(o)}$. Below, we formalize the computation of the only S component, since the other two assume similar formalism. The matrices $P^{(s)}$ and $Q^{(s)}$ are determined by minimizing the regularized squared error:

$$min_{p^{(s)*},q^{(s)*}} \sum_{j=0}^{|\mathbb{U}|} \sum_{i=0}^{|\mathbb{C}|} (M_{i,j}^{(s)} - p_j^{(s)^T} q_i^{(s)})^2 + \lambda(\left\|p_j^{(s)}\right\|^2 \left\|q_i^{(s)}\right\|^2) \tag{11}$$

where $M_{i,j}^{(s)}$ is the (i, j) value considering the s attitude, the regularization factor λ is fixed to 0.1, and the summation is extended only to the concepts on which the user u_i has expressed an attitude, that is, $M_{i,j}^{(s)}$ is known. An iterative approach based on the *alternating least squares* technique with regularization [17] is adopted for ensuring the convergence of the Eq. 16, that is, when either the matrices P and Q are no longer changing or the change is not significant. One of the strengths of this optimization technique is its ability to handle large sparse datasets built up of implicit interactions between users and items. Moreover, parallel implementations suitable for distributed processing frameworks are also available (see, for instance, [18, 19, 20]).
Now, each user $u_i \in \mathbb{U}$ is associated with a vector $q_i^{(s)} \in \mathbb{R}^f$. The rating of the candidate user u_j to be considered for recommendation to u_i is predicted by the cosine similarity measure as follows:

$$r_{i,j}^{(s)} = \frac{q_i^{(s)} \cdot q_j^{(s)}}{\left\|q_i^{(s)}\right\| \left\|q_j^{(s)}\right\|} \tag{12}$$

The contribution of the three components SVO is linearly combined, as follows:

$$r_{i,j;k} = \alpha_k^{(s)} r_{i,j}^{(s)} + \alpha_k^{(v)} r_{i,j}^{(v)} + \alpha_k^{(o)} r_{i,j}^{(o)} \tag{13}$$

where $\alpha_k^{(s)}$, $\alpha_k^{(v)}$, and $\alpha_k^{(o)}$ are three constants in the $[0, 1]$ interval and depend on the macro-category k under examination. Section 4.2 describes the procedure to estimate these parameters.

As mentioned in Section 3, the candidate set of users $\overline{\mathbb{U}}$ consists of the users $u_j \in \mathbb{U}$ that have discussed topics similar to those discussed by the target user u_i. Since the categorization may assign more than one macro-category in K for each pair of users (i, j), multiple $r_{i,j;k}$ values have to be combined. As a result, we select the highest ranking among the considered macro-categories as follows:

$$r_{i,j} = max_{\forall k \in \mathbb{K} | k \in T_{u_i}^{(\mathbb{K})} \wedge k \in T_{u_j}^{(\mathbb{K})}} \; r_{i,j;k} \tag{14}$$

The "Who to follow" functionality in microblogging platforms is often implemented with a list of users that does not depend on the current submitted query or context, as in the case of Twitter [21]. So it seems rational to collect the users that show any form of content-based similarity with the target user, with no regard to a specific macro-category.

As with ratings in collaborative filtering approaches, potential *bias* may exist in terms of both attitudes expressed by users and average perception of debated concepts. Two users might be debating on the same concept, but one being a cynic who expresses often negative attitudes, and the other showing a more enthusiastic disposition. In addition, selected topics on micro-blogging platforms might enjoy strong popularity due to several reasons. In this scenario, the popularity bias usually denotes the tendency for some items to be recommended more frequently [22]. Other forms of bias may generate variations in the attitudes expressed by the user on particular concepts. User's bias corresponds to that tendency of the user to give better or worse ratings than the average.

Koren [23] proved that, by considering user and concept biases in the recommendation, improvements can be obtained because it can allow for the intrinsic difference between users and the between concepts to be represented. MF models face these effects by explicitly taking into account the bias parameters as follows:

$$b_{i,j}^{(s)} = \mu^{(s)} + b_{u_i}^{(s)} + b_{c_j}^{(s)} \tag{15}$$

where the terms b_{u_i} and b_{c_j} represent the observed deviations of user u_i and concept c_j from the average values, and μ is the overall average value of the s dimension. They describe general properties of users and concepts, without accounting for any involved interactions. These bias parameters are summed up with the predicted ranking $p_j^{(s)^T} q_i^{(s)}$ during the minimization phase obtaining:

$$min_{p^{(s)*},q^{(s)*}} \sum_{j=0}^{|\mathbb{U}|} \sum_{i=0}^{|\mathbb{C}|} (M_{i,j}^{(s)} - \mu^{(s)} - b_{u_i}^{(s)} - b_{c_j}^{(s)} - p_j^{(s)^T} q_i^{(s)})^2 + \lambda (b_{u_i}^{(s)2} + b_{c_j}^{(s)2} + \left\| p_j^{(s)} \right\|^2 \left\| q_i^{(s)} \right\|^2) \tag{16}$$

3.4. Temporal analysis of attitudes

User's attitudes constantly change over time, thus tracking the temporal dynamics of user's interests may help improve personalized systems. The proposed MF-based recommendation includes static representations of interests and concepts. A possible solution is to extend the model by considering potential evolution of these two dimensions over time.

Each timeline T_u is partitioned into $N_{\Delta t}$ intervals of Δt time span. The SVO-based analysis required for the definition of the matrix M is performed on each of these intervals. Therefore, we obtain multiple matrices, one for each time span, on which we perform the MF. The rationale is that, given two users, if they both have discussed the same topic but at different times, they have to be considered less relevant to each other than users that have discussed same topics at similar times. Formally, each rating function $r_{i,j}$ is dependent on the time slot $t_l \rightarrow t_l + \Delta t$ with $l = [1, \cdots, N_{\Delta t} - 1]$, as well. The final ranking is obtained by averaging the time-dependent ranking as follows:

$$r_{i,j} = \frac{1}{N_{\Delta t}} \sum_{l=1}^{N_{\Delta t}-1} r_{i,j}(t_l) \tag{17}$$

where $r_{i,j}(t_l)$ is evaluated by considering the partition of the users' timeline in the interval $t_l \rightarrow t_l + \Delta t$.

3.5. Computational Complexity

The computation complexity of the approach is driven by the MF process. Indeed, in order to provide up-to-dated recommendations, the MF must be regularly recomputed according to new published content. Instead, the complexity of the SVO-based analysis is determined by the SVM-based categorization of microposts to pre-defined classes (Sect. 3) and NB classification used for the sentiment analysis (Sect. 3.1), which can be trained at once, so we are more interested in the computational requirements after the training step.

A popular algorithm that implements non-negative MF [24] has computational complexity $O(f|\mathbb{C}||\mathbb{U}|)$ per iteration during the training phase. Of course when new information is added to the user-concept matrix, the factorization can be initiated from the P and Q matrices obtained in the previous cycle, speeding up the time required for the completion of the iterative process to the convergence of the Eq. 16. Once the MF is completed, the rating for a candidate user is computed with $O(f|\mathbb{C}||\mathbb{U}|)$ complexity, where f is related to the computation of the cosine similarity (Eq. 1), and $|\mathbb{C}||\mathbb{U}|$ is due to the retrieval of the users whose timelines contain hashtags similar to the ones in the

	Time span	# Tweets	# Users	Lang	Topic
D1:	Jan 2013 → Feb 2013	1,0M	71K	IT	Political Elections
D2:	Sep 2015 → Feb 2015	3,5M	181K	EN	Car Brands
D3:	Dec 2014 → Feb 2015	2,9M	110K	EN	Mobile Phone Brands
D4:	Jan 2015 → Dec 2015	1,2M	99K	IT	Movies
D5:	Jan 2016 → Mar 2016	25,3M	1,1M	IT	Trending Topics

Table 1: Statistics of datasets.

target user's timeline. A pre-processing of the set $\mathbb{C}$ removes from the feature space those hashtags whose micropost frequency is less than some predetermined threshold. The assumption is that rare hashtags are non-informative for the recommendation.

Since TDMF takes into consideration a constant number of partitions of the timeline, the above-mentioned big O notation is still valid but a $N_{\Delta t}$-fold increase exists in the processing time.

As for the SVO-based analysis, the categorization based on the SVM technique shows complexity of $O(|\mathbb{V}|)$, where $\mathbb{V}$ corresponds to the vocabulary of terms that compose $\mathbb{T}$. $O(|\overline{\tau}|)$ is the complexity of the NB classification of a timeline's micropost to one of the three classes $\{s^{(+)}, s^{(-)}, s^{(\emptyset)}\}$, where $|\overline{\tau}|$ is the average length of a post. The two computations are performed for each recent post in the user's timeline T_u, so the SVO-based analysis shows an approximate complexity of $O(|\mathbb{U}||\mathbb{V}|)$ by considering the number of posts and the average length of a posts constant.

4. Evaluation

Experimental tests of the proposed approach were performed on different real-world datasets, obtained by monitoring the traffic produced by users on Twitter. Such data enabled us to realize a comparative analysis of our system with similar approaches proposed in the research literature.

To guarantee a correct statistical significance of the results, the experimental evaluation were carried out taking into account different datasets as shown in Table 1. The considered datasets were gathered as follows:

D1: We filtered from the Twitter stream the hashtags related to politician leaders and Italian parties during the 2013 Italian general election.

D2: A dataset on majors technology brands, among others Samsung, Apple, Nokia, Huawei, LG, Motorola, and Blackberry.

D3: Tweets matching terms related to the automotive landscape, such as Audi, BMW, Ferrari, Jaguar, Mercedes, Toyota, and Porsche.

D4: Corpus of tweets that counts more than 200 movies released in Italy during 2015.

D5: This dataset includes tweets of trending topics automatically suggested by the microblogging platform over a period of 3 months, such as #bruxellesattacks, #oscars, #syriaconflict.

The time period in which each dataset has been collected is splitted in two parts. The initial 70% corresponds to the data for the training set, the subsequent remaining data are used for testing the proposed recommendation system against other benchmarks. A crawler periodically updated the profiles of each user during the whole time period, considering also new followers/following relationships. Each time a social tie is discovered, its timestamp is being associated with the time the crawler found it.

4.1. Benchmark: metrics and comparative algorithms

A wide spectrum of evaluation metrics for RSs exist, most of them focused on their accuracy. While the assessment of such aspect is foundamental, there are limits that emerge due to the discrepancy between the users' perception and the outcome of the metrics [25]. An accurate recommendation, however, is not necessarily perceived as a useful one. If the users in the recommendation list are very similar to the target users, the benefits of the system are limited because good chances are that the users discover them by querying the microblogging service or exploring the neighbors of their personal social network by themselves. A more useful recommender provides accurate and personalized recommendations guaranteeing, at the same time, high levels of novelty and diversity. For this reason, multiple metrics have been considered for better evaluating the optimal trade-off between accuracy, novelty and diversity of the considered recommendation approaches.

4.1.1. Accuracy

The goal of the people-to-people recommendation is to provide the target user with a set of relevant people to follow. In our approach, the output is a set $\mathbb{L}_u$ of potentially relevant users, where the timeline of each u in $\mathbb{L}_u$ is considered for the user profiling.

A straightforward methodology to measure the accuracy of a RS is to assess how many suggestions are relevant to the user. We suppose that user u_i is relevant to u_j if a real *following relationship* exists between them.

Precision is the most used accuracy measure and gives a general idea of the overall performance of the recommendation. Since it is known that users focus their attention on the top ranked items of a list [26], we employed the *Success at Rank K (S@K)* measure that is commonly used for evaluating ranked lists of recommendations. It expresses the mean probability that a relevant user is located in the first K positions of the suggested users set, and is obtained as follows:

$$Acc(u, \mathbb{L}_u) = \frac{1}{|\mathbb{L}_u|} \sum_{u_i \in \mathbb{L}_u} S@K(u_i) \tag{18}$$

where $S@K(u_i)$ is one whether u_i is a relevant user for the target user u, zero otherwise.

4.1.2. Diversity

Diversity generally applies to a set of recommended items, and is related to how different the items are with respect to each other. Diversity is then determined by evaluating the dissimilarity of textual features extracted from users' timeline of the recommended set $\mathbb{L}_u$.

The diversity measure we devised is based on the Latent Dirichlet Allocation (LDA) [27], a generative probabilistic model for collections of discrete data such as text corpora. LDA shapes latent topics as a distribution over the words of the vocabulary, and every given document as a distribution over these topics, which is sampled from Dirichlet distributions. LDA model is often used for dimensionality reduction, where any input document d is assigned to a fixed set of real-valued features, that is, the posterior

Dirichlet parameters $\gamma^*(d)$. If we assume that γ^* is represented by means of a vector, we define the diversity based on LDA as follows:

$$Div(\mathbb{L}_u) = 1 - \left\|\gamma^*(d(T_{\mathbb{L}_u}))\right\| \tag{19}$$

where $d(T_{\mathbb{L}_u})$ represents a text document consisting of the concatenated posts from the users' timeline in $\mathbb{L}_u$. The LDA diversity reaches high values if the combination of users' timeline represent several different latent topic.

4.1.3. Novelty

There have been several attempts to capture the degree of novelty in single measures [28, 29]. Novel recommendations consist in suggesting items that the user did not know about, and whose description is semantically far from users' interests. Therefore, the measure takes into consideration both the recommended content and the target users' interests. Hijikata et al. [30] use collaborative filtering to derive novel recommendations by explicitly asking users which items they already know. The scale of the domain we are dealing with and the number of users involved do not allow us to follow a similar methodology. The novelty measure assumes high values if the recommended users' timeline include several topics that are not discussed yet by the target user. Therefore, we can define novelty in terms of overlap among topics discussed by the target user u and the suggested users $\mathbb{L}_u$. More formally we define:

$$Nov(u, \mathbb{L}_u) = \frac{1}{|\mathbb{L}_u|} \sum_{u_i \in \mathbb{L}_u} \frac{1}{|T_u^{(\mathbb{C})}|} \sum_{c \in T_u^{(\mathbb{C})}} \left(-\frac{n_{u_i}^{(c)}}{n_{u_i}}\right) \tag{20}$$

4.1.4. Algorithms for comparative evaluation

In order to outline comparative conclusions from the experimental evaluations on the considered datasets, the following people-to-people recommendation approaches have been devised and included in the experimental tests:

R: A baseline recommender that randomly suggests users from the considered dataset.

NP: A non-personalized recommender that always suggests the most popular users in the dataset, that is, the users with the highest number of followers.

CB: The content-based approach proposed in [31] (with the name of S1), which represents each user u through the function $d(T_u)$, that is, the text document consisting of the concatenated posts from the users' timeline. A traditional search engine based on the vector space model with a TF-IDF scoring function and cosine similarity measure [32] returns the users that are more similar to the target one by considering their timeline's content.

CF: It represents each user u_i by the following set:

$$\{\forall u_j \in \mathbb{U} \mid i \neq j \wedge (A_{ij} = 1 \vee A_{ji} = 1)\} \tag{21}$$

that includes any user with an explicit tie with u_i (i.e., followers and followees). The IDs of these users are converted to unique keywords and, similarly to the CB approach, a IR-based search engine returns a ranked link of recommendations. It corresponds to the S7 approach in [31].

CBH: Similarly to CB, each user is represented by the posts included in the timeline, but instead of every keyword, the content is limited to the set of concepts in $T_u^{(\mathbb{C})}$. The frequency of the concept in the user's timeline corresponds to the term frequency.

FoF The Friend-of-Friend recommender is available in popular social network services, such as Facebook and LinkedIn [33, 34]. It relies on the following hypothesis: if many users followed by u subsequently follow a particular person, this latter person is more likely to be suggested to u. The greater the number of u's friend that follow the candidate, the higher is the relative rank in the suggested list. It follows the common neighbor paradigm that makes use of explicit social ties often considered in the link prediction task [35].

MFE: The straightforward recommender based on MF [15] where the items to suggest are the users themselves. Therefore, the training set is composed of ratings $r_{u_i,u_j} \in 0, 1$, which represent the existence of an explicit social tie (i.e., following relationship) that bind the pair of users. The estimated rating between the target user u_i and the generic candidate u_j is obtained by the inner product in the latent factor space, that is:

$$q_i^T p_j$$

The top-ranked candidates of the target user are the ones assessed in the evaluation.

MF: The recommendation approach based on the SVO-based analysis and the MF models introduced in Section 3.3.

TDMF: The previous recommendation approach enhanced with temporal dynamic features, as explained in Section 3.4.

The explicit social ties used by CF, FoF and MFTB approaches are extracted from the training set, whereas the test set is used to assess the performances. Similarly, the timelines considered for the learning process in the CB, CBH, MF and TDMF approaches consist of microposts published in the first split (i.e., training set) of each dataset.

4.2. Experimental results

The evaluation of the accuracy is achieved by comparing our system with some state-of-the-art people-to-people RSs. To perform an offline comparison analysis, an evaluation set has been built. We selected 1,000 random users from each of the dataset already introduced in Section 4, that match the following criteria:

- users that posted at least five tweets
- users with at least ten friends and followers into the dataset (that can be selected for the evaluation test)

This kind of offline analysis suffers from an important weakness: the natural sparsity of datasets derived from social network limits the amount of relevant content that can be evaluated. In this way, selecting exclusively random users without matching the above-mentioned criteria may lead to have no real friends or followers to compare with into the test dataset, and therefore resulting in a zero accuracy for every recommender. On the other hand, offline evaluations are often considered in RS studies because they allow researchers to perform large scale evaluations on thousands of users, different datasets and algorithms at once [36].

In Table 2, we report the results of the comparative analysis in terms of accuracy. All the experimental results were tested for statistical significance through a two-tailed *t-test* with a significance level set to $p < 0.05$.

In terms of accuracy, the outcomes show the substantial benefits obtained with the proposed approaches and confirm our initial hypothesis about the potential combination of sentiment, volume and objectivity to better identify real relationships between users. A traditional MF-based approach that limits its analysis to the explicit ties between users, i.e., **MFE**, does not reach similar accuracies. The results highlight also how the **TDMF** approach obtains the best values among all datasets. This is a relevant achievement that endorses how important is to consider temporal features for the people-to-people recommendation.

Subsequently, we evaluated the performance of the RSs in terms of diversity and novelty. Table 3 summarizes the diversity and novelty obtained on average among all datasets. Approaches that leverage social network information such as **NP**, **CF**, and **FoF** reach high values of novelty, that is, they are able to suggest people that are more likely to discuss topics unknown to the target user. On the contrary, **MF** and **TDMF** techniques, thanks to matrix decomposition and temporal analysis, supply the RS with the ability of suggesting diverse users to follow, that is, a list of recommended users that are different, one from each other.

As for the temporal factor, we analyzed the variation of the accuracy as a function of the extent of the Δt time span. Table 4 shows that datasets D1, D2, D3, and D4 achieve the best accuracy with Δt intervals of 14 days and 21 days, while D5 with Δt of 7 days. Since the latter dataset consists of several fragmented and temporary trending topics, by considering a time span of 7 days, the most relevant topics are better represented. One popular example in the dataset is the news about the 2016 Brussels bombings. By considering a shorter time span, the recommendation is more tailored to users that are interested in the terrorism attack instead of considering people fascinated by the capital of Belgium, its history or cultural events.

In order to understand the behaviour of the users' attitudes, we performed a sen-

Table 2: A comparison of accuracy outcomes among some state-of-the-arts recommender approaches.

RS	D1	D2	D3	D4	D5
R	0.049	0.057	0.024	0.045	0.038
NP	0.146	0.114	0.122	0.111	0.065
CB	0.130	0.118	0.115	0.126	0.111
CF	0.172	0.163	0.161	0.167	0.151
CBH	0.127	0.099	0.105	0.055	0.078
FoF	0.165	0.155	0.159	0.140	0.132
MFE	0.121	0.105	0.111	0.119	0.107
MF*	0.187	0.181	0.178	0.201	0.182
TDMF**	**0.212**	**0.233**	**0.235**	**0.241**	**0.255**

(*) With the best SVO values for each dataset and f = 5

(**) Best Δt for each dataset showed in Table 4 and f = 5

sitivity evaluation of SVO parameters through a large-scale *gradient descent algorithm* [37] with learning rate $\zeta = 0.1$. This evaluation enabled us to observe how the performance could be improved by tuning the weights that define the different contributions of sentiment, volume, and objectivity, based on the nature of topics (on which the users' similarity is computed). In particular, the results in Table 5 highlight how the contribution of sentiment is higher for topics about politics and movies, while the contribution of volume is on average significant for all of the considered topics.

Finally, Figure 3 reports the RS accuracy for the **MF** approach as a function of the latent factor's number f. As can be noted, there are no relevant accuracy improvements by increasing the number of latent factors. This finding motivated us to select a fixed f=5 for all of the aforementioned experimental evaluations. A lower number of latent features decreases a lot the computational resources.

The obtained outcomes pave the way to the hypothesis that a hybrid approach that accurately selects the recommendations from multiple approaches, such as **FoF**, **CF** and **TDMF**, may show benefits to the user. For instance, the approach based on explicit social ties (FoF) outperforms attitudes when the goal is to have high novelty,

Table 3: Results for diversity and novelty metrics

RS	Novelty	Diversity
R	0.21	0.14
NP	**0.29**	0.34
CB	0.14	0.23
CF	**0.39**	**0.55**
CBH	0.11	0.15
FoF	**0.35**	0.29
MFE	0.25	0.28
MF	0.19	**0.45**
TDMF	0.25	**0.47**

Table 4: Results of S@10 for TDMF recommender system while varying the length of Δt time span

Dataset	7gg	14gg	21gg	30gg
D1	0.191	0.202	**0.212**	0.187
D2	0.210	**0.233**	0.221	0.200
D3	0.201	**0.235**	0.18	0.199
D4	0.192	0.205	**0.241**	0.225
D5	**0.255**	0.189	0.188	0.173

Table 5: Sensitivity analysis of sentiment-volume-objectivity parameters for the best obtained values of MF recommender system

Dataset	S@10	S	V	O
D1	0.187	**0.45**	0.45	0.10
D2	0.181	0.20	**0.60**	0.20
D3	0.178	0.30	**0.65**	0.05
D4	0.201	**0.45**	0.45	0.10
D5	0.182	0.20	**0.70**	0.10

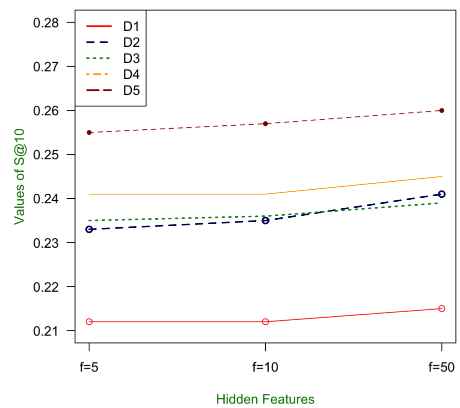

Figure 3: S@10 while varying the number of latent factors f, with the best values of Δt

whereas **MF** and **TDMF** obtain in general better accuracy and diversity on the considered datasets. But a simple linear combination of the outputs would not be optimal. Future work is required to understand what the user is currently expecting from the recommender, promoting items that are not similar to what they have previously liked (i.e., maximizing the diversity), or pursuing higher accuracy, that is, items similar to what users have previously liked.

5. Related work

In this section we describe several works somehow related to the proposed system, especially focusing on people-to-people recommendation.

From the seminal works on link prediction [1, 38], many relevant contributions have been proposed. Freyne *et al.* [39] provide the active user with suggestions about key people to connect to, based on social relationship information coming from different external sources and gathered through the social aggregator Sonar [40]. In [41] techniques that exploit both the user-generated content and the social network structure are proposed for recommending people of potential interest to the target user. Such techniques rely on the Friend-of-Friend (FoF) hypothesis that if many of the target user's friends have a friend in common, this latter could be friend of the target user as well.

This system is one of the baseline approaches that appear in the comparative analysis reported in Section 4. The authors of [42] address the same problem in an enterprise scenario. They aggregate information from different sources in order to profile users, thus being able to identify those who have provided a similar contribution (e.g., co-author papers, patent authorship, etc.). This work is based on the assumption that if two users have generated content on similar topics, they are more likely to appreciate getting in touch with each other than other users. Quercia and Capra propose a mobile application that relies on the users' physical proximity for generating people-to-people recommandations [43]. In [44], a supervised machine-learning approach is proposed to address the link recommendation problem on an enterprise social network. To this end, the authors mine the user-generated content, the social graph, and the company's organizational chart to profile enterprise users. Some work has been focused on the user recommendation problem in social micro-blogging services like Twitter. In particular, the authors of [45] make a comparison between content-based and collaborative filtering approaches for user profiling. To this end, they resort to a classic search engine to index and classify such profiles via the traditional TF-IDF approach of Information Retrieval. Then, the top-k users are suggested to the target user. Their experimental results show the better performance of collaborative filtering approaches compared to those of content-based.

Such findings suggest that the relations between users are more structured, and therefore more relevant for user recommendation task than the noisy microposts. Given the relevance of these approaches, both of them have been implemented and compared with the proposed system (see Sect. 4). In [46] the authors address the same problem through an algorithm which explores the topology of the social graph in Twitter to locate users to recommend to the active user. This approach extends the well-known item-based recommendation nearest neighbor technique [47] to the user recommendation task. However, the works presented in [48] and [49] show that approaches based on matrix factorization provide better performance than those based on neighborhood techiniques. Such consideration, along with the need to operate on large-scale social networks, inspired our approach.

Matrix factorization techniques have been previously considered in the link prediction problem. In [50] the authors combine explicit and latent features and prove the effectiveness on various datasets. Kutty *et al.* [51] propose tensor space models as a potential framework able to include also additional attributes associated with each user. Other works extend the analysis by considering dynamic interactions, that is, the time in which a tie is created, e.g., [52, 53, 54, 51]. The above-cited works have not been explicitly evaluated on popular social network services, such as Twitter or Facebook, and do not take into account user attitudes.

Yang *et al.* [55] extends the check-ins left by the users on location-based services with additional features, such as fine-grained user preferences extracted from opinions expressed in user comments. MF techniques are considered for capturing both social and inter-venue influence based on similarity measures between user comments, geo-distance, categories, reviews, etc. Similarly, in [56] the authors use a three-way tensor model User$\times$Keyword$\times$Venue for personalized location ranking.

Although a large number of contributions have been devoted to the people-to-people recommendation issue - to the best of our knowledge - exploiting sentiment analysis of user-generated contents for purposes of community detection and/or user recommendation has not been deeply investigated. Xu *et al.* [57] transform the sentiment-based community discovery into a correlation clustering problem and propose a random rounding algorithm based on semidefinite programming for its solution. In [58] the authors describe an unsupervised approach based a non-parametric clustering algorithm for detecting hyper-groups of communities, called *hyper-communities*, where users share the same sentiments. In [59], the authors extract users' interests from their microposts and identify some sentiment-based features that express the likelihood of two users establishing a relationship (i.e., following each other or mutually mentioning) between them. They also advance a factor graph model including a sentiment-based version of the cognitive balance theory for predicting potential relationships.

As far as we aware, this is the first work combining sentiment analysis and matrix factorization techniques to assist users in locating interesting people.

6. Conclusion

In this paper, we have described a people-to-people recommendation approach for large-scale social networks. Our work emphasizes the use of user's attitudes such as implicit sentiment, volume and objectivity to improve recommendation performance and matrix factorization models to maximize efficiency and scalability. The experimental results showed the advantage of our approach compared with the state-of-the-art techniques. Taking advantage of implicit sentiment related to the users' timeline, enables us to better identify the relationship of interest between users. The experimental evaluation on different datasets has also proved that the SVO factors are influenced by the topics under discussion. When multiple factors obtained from the user-generated content are taken into consideration, an adequate analysis of their relevance in the recommendation process is required. The same conclusion holds for the time unit considered for the temporal analysis of the expressed users' attitudes.

References

[1] D. Liben-Nowell, J. Kleinberg, The link prediction problem for social networks, in: Proceedings of the Twelfth International Conference on Information and Knowledge Management, CIKM '03, ACM, New York, NY, USA, 2003, pp. 556–559. doi:10.1145/956863.956972.

[2] P. Wang, B. Xu, Y. Wu, X. Zhou, Link prediction in social networks: the state-of-the-art, Science China Information Sciences 58 (1) (2015) 1–38. doi:10.1007/s11432-014-5237-y.
URL http://dx.doi.org/10.1007/s11432-014-5237-y

[3] L. M. Aiello, A. Barrat, R. Schifanella, C. Cattuto, B. Markines, F. Menczer, Friendship prediction and homophily in social media, ACM Trans. Web 6 (2) (2012) 9:1–9:33. doi:10.1145/2180861.2180866.
URL http://doi.acm.org/10.1145/2180861.2180866

[4] A. Cui, M. Zhang, Y. Liu, S. Ma, K. Zhang, Discover breaking events with popular hashtags in twitter, in: Proceedings of the 21st ACM International Conference

on Information and Knowledge Management, CIKM '12, ACM, New York, NY, USA, 2012, pp. 1794–1798. doi:10.1145/2396761.2398519.

[5] J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, J. Sperling, Twitterstand: News in tweets, in: Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '09, ACM, New York, NY, USA, 2009, pp. 42–51. doi:10.1145/1653771.1653781.

[6] J. Costa, C. Silva, M. Antunes, B. Ribeiro, Adaptive and Natural Computing Algorithms: 11th International Conference, ICANNGA 2013, Lausanne, Switzerland, April 4-6, 2013. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, Ch. Defining Semantic Meta-hashtags for Twitter Classification, pp. 226–235. doi:10.1007/978-3-642-37213-1_24.

[7] M. R. Levy, S. V. E. N. Windahl, AUDIENCE ACTIVITY AND GRATIFICATIONS: A Conceptual Clarification and Exploration, Communication Research 11 (1) (1984) 51–78. doi:10.1177/009365084011001003.

[8] T. Joachims, Text categorization with suport vector machines: Learning with many relevant features, in: Proceedings of the 10th European Conference on Machine Learning, ECML '98, Springer-Verlag, London, UK, UK, 1998, pp. 137–142.

[9] Google, Google news, last visited on 20 December 2016.
URL https://news.google.com

[10] B. Liu, Sentiment analysis and subjectivity, Handbook of Natural Language Processing, (2010) 627–666.

[11] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau, Sentiment analysis of twitter data, in: Proceedings of the Workshop on Languages in Social Media, LSM '11, Association for Computational Linguistics, Stroudsburg, PA, USA, 2011, pp. 30–38.

[12] A. Go, R. Bhayani, L. Huang, Twitter sentiment classification using distant supervision, Processing (2009) 1–6.

[13] I. G. Councill, R. McDonald, L. Velikovich, What's great and what's not: learning to classify the scope of negation for improved sentiment analysis, in: Proceedings of the Workshop on Negation and Speculation in Natural Language Processing, NeSp-NLP '10, Association for Computational Linguistics, Stroudsburg, PA, USA, 2010, pp. 51–59.

[14] A. Pak, P. Paroubek, Twitter as a corpus for sentiment analysis and opinion mining, in: N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, D. Tapias (Eds.), Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), European Language Resources Association (ELRA), Valletta, Malta, 2010.

[15] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (8) (2009) 30–37. `doi:10.1109/MC.2009.263`.

[16] A. Karatzoglou, X. Amatriain, L. Baltrunas, N. Oliver, Multiverse recommendation: N-dimensional tensor factorization for context-aware collaborative filtering, in: Proceedings of the Fourth ACM Conference on Recommender Systems, RecSys '10, ACM, New York, NY, USA, 2010, pp. 79–86. `doi:10.1145/1864708.1864727`.

[17] Y. Hu, Y. Koren, C. Volinsky, Collaborative filtering for implicit feedback datasets, in: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, ICDM '08, IEEE Computer Society, Washington, DC, USA, 2008, pp. 263–272. `doi:10.1109/ICDM.2008.22`.

[18] Y. Zhou, D. Wilkinson, R. Schreiber, R. Pan, Large-scale parallel collaborative filtering for the netflix prize, in: Proceedings of the 4th International Conference on Algorithmic Aspects in Information and Management, AAIM '08, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 337–348. `doi:10.1007/978-3-540-68880-8_32`.

[19] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. J. Franklin, S. Shenker, I. Stoica, Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing, in: Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation, NSDI'12, USENIX Association, Berkeley, CA, USA, 2012, pp. 2–2.

[20] M. Winlaw, M. B. Hynes, A. Caterini, H. D. Sterck, Algorithmic acceleration of parallel ALS for collaborative filtering: Speeding up distributed big data recommendation in spark, in: 21st IEEE International Conference on Parallel and Distributed Systems, ICPADS 2015, Melbourne, Australia, December 14-17, 2015, IEEE, 2015, pp. 682–691. doi:10.1109/ICPADS.2015.91.

[21] Twitter Inc., About twitter's account suggestions, last visited on 20 December 2016.
URL https://support.twitter.com/articles/227220

[22] O. Celma, P. Cano, From hits to niches?: Or how popular artists can bias music recommendation and discovery, in: Proceedings of the 2Nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition, NETFLIX '08, ACM, New York, NY, USA, 2008, pp. 5:1–5:8. doi:10.1145/1722149.1722154.

[23] Y. Koren, Factorization meets the neighborhood: A multifaceted collaborative filtering model, in: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '08, ACM, New York, NY, USA, 2008, pp. 426–434. doi:10.1145/1401890.1401944.
URL http://doi.acm.org/10.1145/1401890.1401944

[24] D. D. Lee, H. S. Seung, Learning the parts of objects by non-negative matrix factorization., Nature 401 (6755) (1999) 788–791. doi:10.1038/44565.
URL http://dx.doi.org/10.1038/44565

[25] S. M. McNee, J. Riedl, J. Konstan, Accurate is not always good: How accuracy metrics have hurt recommender systems, in: Extended Abstracts of the 2006 ACM Conference on Human Factors in Computing Systems (CHI 2006), 2006.

[26] T. Joachims, L. Granka, B. Pan, H. Hembrooke, G. Gay, Accurately interpreting clickthrough data as implicit feedback, in: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '05, ACM, New York, NY, USA, 2005, pp. 154–161. doi:10.1145/1076034.1076063.

[27] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.

[28] T. Murakami, K. Mori, R. Orihara, Metrics for evaluating the serendipity of recommendation lists, in: Proceedings of the 2007 Conference on New Frontiers in Artificial Intelligence, JSAI'07, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 40–46.

[29] M. de Gemmis, P. Lops, G. Semeraro, C. Musto, An investigation on the serendipity problem in recommender systems, Inf. Process. Manage. 51 (5) (2015) 695–717. doi:10.1016/j.ipm.2015.06.008.

[30] Y. Hijikata, T. Shimizu, S. Nishida, Discovery-oriented collaborative filtering for improving user satisfaction, in: Proceedings of the 14th International Conference on Intelligent User Interfaces, IUI '09, ACM, New York, NY, USA, 2009, pp. 67–76. doi:10.1145/1502650.1502663.

[31] J. Hannon, M. Bennett, B. Smyth, Recommending twitter users to follow using content and collaborative filtering approaches, RecSys'10 : Proceedings of the 4th ACM Conference on Recommender Systems, 26-30 (10) (2010) 8.

[32] R. Baeza-Yates, B. Ribeiro-Neto, Modern Information Retrieval, 2nd Edition, Addison-Wesley Publishing Company, USA, 2008.

[33] J. Chen, W. Geyer, C. Dugan, M. Muller, I. Guy, Make new friends, but keep the old: recommending people on social networking sites, in: Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI '09, ACM, New York, NY, USA, 2009, pp. 201–210. doi:10.1145/1518701.1518735.

[34] G. Groh, S. Birnkammerer, V. Köllhofer, Recommender Systems for the Social Web, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, Ch. Social Recommender Systems, pp. 3–42. doi:10.1007/978-3-642-25694-3_1.

[35] D. Liben-Nowell, J. Kleinberg, The link-prediction problem for social networks, J. Am. Soc. Inf. Sci. Technol. 58 (7) (2007) 1019–1031. doi:10.1002/asi.v58:7.
URL http://dx.doi.org/10.1002/asi.v58:7

[36] J. L. Herlocker, J. A. Konstan, L. G. Terveen, J. T. Riedl, Evaluating collaborative filtering recommender systems, ACM Trans. Inf. Syst. 22 (1) (2004) 5–53. doi:10.1145/963770.963772.

[37] T. Zhang, Solving large scale linear prediction problems using stochastic gradient descent algorithms, in: Proceedings of the twenty-first international conference on Machine learning, ACM, 2004, p. 116.

[38] M. A. Hasan, V. Chaoji, S. Salem, M. Zaki, Link prediction using supervised learning, in: In Proc. of SDM 06 workshop on Link Analysis, Counterterrorism and Security, 2006.

[39] J. Freyne, M. Jacovi, I. Guy, W. Geyer, Increasing engagement through early recommender intervention, in: Proceedings of the Third ACM Conference on Recommender Systems, RecSys '09, ACM, New York, NY, USA, 2009, pp. 85–92. doi:10.1145/1639714.1639730.

[40] I. Guy, M. Jacovi, E. Shahar, N. Meshulam, V. Soroka, S. Farrell, Harvesting with sonar: The value of aggregating social network information, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08, ACM, New York, NY, USA, 2008, pp. 1017–1026. doi:10.1145/1357054.1357212.

[41] J. Chen, W. Geyer, C. Dugan, M. Muller, I. Guy, Make new friends, but keep the old: Recommending people on social networking sites, in: Proceedings of the

SIGCHI Conference on Human Factors in Computing Systems, CHI '09, ACM, New York, NY, USA, 2009, pp. 201–210. doi:10.1145/1518701.1518735.

[42] I. Guy, I. Ronen, E. Wilcox, Do you know?: Recommending people to invite into your social network, in: Proceedings of the 14th International Conference on Intelligent User Interfaces, IUI '09, ACM, New York, NY, USA, 2009, pp. 77–86. doi:10.1145/1502650.1502664.

[43] D. Quercia, L. Capra, Friendsensing: Recommending friends using mobile phones, in: Proceedings of the Third ACM Conference on Recommender Systems, RecSys '09, ACM, New York, NY, USA, 2009, pp. 273–276. doi:10.1145/1639714.1639766.

[44] J. Zhang, Y. Lv, P. Yu, Enterprise social link recommendation, in: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, CIKM '15, ACM, New York, NY, USA, 2015, pp. 841–850. doi:10.1145/2806416.2806549.

[45] J. Hannon, M. Bennett, B. Smyth, Recommending twitter users to follow using content and collaborative filtering approaches, in: Proceedings of the Fourth ACM Conference on Recommender Systems, RecSys '10, ACM, New York, NY, USA, 2010, pp. 199–206. doi:10.1145/1864708.1864746.

[46] M. G. Armentano, D. Godoy, A. Amandi, Topology-based recommendation of users in micro-blogging communities, Journal of Computer Science and Technology 27 (3) (2012) 624–634. doi:10.1007/s11390-012-1249-5.

[47] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms, in: Proceedings of the 10th International Conference on World Wide Web, WWW '01, ACM, New York, NY, USA, 2001, pp. 285–295. doi:10.1145/371920.372071.

[48] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (8) (2009) 30–37. doi:10.1109/MC.2009.263.

[49] G. Zhao, M. L. Lee, W. Hsu, W. Chen, H. Hu, Community-based user recommendation in uni-directional social networks, in: Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management, CIKM '13, ACM, New York, NY, USA, 2013, pp. 189–198. doi:10.1145/2505515.2505533.

[50] A. K. Menon, C. Elkan, Link prediction via matrix factorization, in: Proceedings of the 2011 European Conference on Machine Learning and Knowledge Discovery in Databases - Volume Part II, ECML PKDD'11, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 437–452.
URL http://dl.acm.org/citation.cfm?id=2034117.2034146

[51] S. Kutty, L. Chen, R. Nayak, A people-to-people recommendation system using tensor space models, in: Proceedings of the 27th Annual ACM Symposium on Applied Computing, SAC '12, ACM, New York, NY, USA, 2012, pp. 187–192. doi:10.1145/2245276.2245312.
URL http://doi.acm.org/10.1145/2245276.2245312

[52] E. Acar, D. M. Dunlavy, T. G. Kolda, Link prediction on evolving data using matrix and tensor factorizations, in: 2009 IEEE International Conference on Data Mining Workshops, 2009, pp. 262–269. doi:10.1109/ICDMW.2009.54.

[53] J. Guo, H. Guo, Multi-features link prediction based on matrix, in: 2010 International Conference On Computer Design and Applications, Vol. 1, 2010, pp. V1–357–V1–361. doi:10.1109/ICCDA.2010.5540852.

[54] S. Gao, L. Denoyer, P. Gallinari, Temporal link prediction by integrating content and structure information, in: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, CIKM '11, ACM, New York, NY, USA, 2011, pp. 1169–1174. doi:10.1145/2063576.2063744.
URL http://doi.acm.org/10.1145/2063576.2063744

[55] D. Yang, D. Zhang, Z. Yu, Z. Wang, A sentiment-enhanced personalized location recommendation system, in: Proceedings of the 24th ACM Conference on Hyper-

text and Social Media, HT '13, ACM, New York, NY, USA, 2013, pp. 119–128. doi:10.1145/2481492.2481505.

[56] D. Yang, D. Zhang, Z. Yu, Z. Yu, Fine-grained preference-aware location search leveraging crowdsourced digital footprints from lbsns, in: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '13, ACM, New York, NY, USA, 2013, pp. 479–488. doi:10.1145/2493432.2493464.
URL http://doi.acm.org/10.1145/2493432.2493464

[57] K. Xu, J. Li, S. S. Liao, Sentiment community detection in social networks, in: Proc. of the 2011 iConference, ACM, New York, NY, USA, 2011, pp. 804–805. doi:10.1145/1940761.1940913.

[58] T. Nguyen, D. Q. Phung, B. Adams, S. Venkatesh, A sentiment-aware approach to community formation in social media., in: J. G. Breslin, N. B. Ellison, J. G. Shanahan, Z. Tufekci (Eds.), ICWSM, The AAAI Press, 2012.

[59] G. Yuan, P. K. Murukannaiah, Z. Zhang, M. P. Singh, Exploiting sentiment homophily for link prediction, in: Proceedings of the 8th ACM Conference on Recommender Systems, RecSys '14, ACM, New York, NY, USA, 2014, pp. 17–24. doi:10.1145/2645710.2645734.