Towards a Followee Recommender System for Information Seeking Users in Twitter

Marcelo G. Armentano, Daniela Godoy and Analía Amandi

ISISTAN Research Institute, Fac. Cs. Exactas, UNCPBA
Campus Universitario, Paraje Arroyo Seco, Tandil, 7000, Argentina
CONICET, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina
{marmenta, dgodoy, amandi}@exa.unicen.edu.ar

Abstract. Micro-blogging activity taking place in sites such as Twitter gains everyday more importance as a source of real-time information and news spreading medium. Finding relevant information sources among the increasing number of Twitter members is essential for users needing to cope with real-time information. In this paper we study Twitter aiming at generating a set of recommendations to a target user consisting in people who publish tweets that might be interesting to him/her. We evaluate and compare two recommendation approaches: the first selects a set of candidate recommendations using only the network topology and the second exploits the user-generated content available in their tweets. We report the results of a set of controlled experiments with real users carried out to evaluate and compare the performance of both algorithms.

Keywords: Recommender Systems; Micro-blogging Activity; Online Social Networks

1 Introduction

Twitter is a social networking site that has been selected for many users as a means of disseminating (and reading) news and information. There are three main factors for choosing Twitter for this goal. First, unlike many other online social networks such as Facebook, Hi5, Orkut, LinkedIn or MySpace, connections in Twitter are unidirectional. This means that a user decides to "follow" other users with no need of this relation to be accepted or reciprocated. Second, the 140-characters length restriction applied to the messages that users can post in Twitter (which are called tweets) enable users to receive their followees updates in almost any mobile device or to quickly read a bunch of them directly on the Internet or within a desktop application. Finally, any user can easily "retweet" another user's post. In this way the information will be spread out from the author followers to other users' followers. Kwak et al. [11] identified that 77.9% of Twitter's connections are unidirectional and only 22.1% of the relations are reciprocate. Moreover, 67.6% of users are not followed by any of their followees, indicating that these users probably use Twitter as a source of information rather than to keep in touch with friends. Finally, Kwak et al. found that retweets collectively determine the importance of the original tweet expressing a form of collective intelligence. All these facts, in addition to the great explosion in the number of registered users in Twitter[1], make us believe that information-seeking users would benefit from a recommender system able to suggest information sources that they might be interest in following.

In this work we study Twitter from a user modeling perspective. Our goal is to provide recommendations to information seekers about users that publish tweets that might be of their interest. In order to be valuable, the recommended followees should be in the category of information broadcasters, since these users will probably generate content that the target user may be interested in reading.

Unlike traditional recommendation systems, we do not have any explicit information available about the user's interests in the form of ratings on items he/she likes or dislikes. For profiling a Twitter user the structure of the followers/followees network and the tweets published in this network is the only information available. Both are considered in this paper as a means to recommend people either belonging to the user's neighborhood or sharing content-related interests.

In this article we present two recommendation algorithms using two different techniques: a collaborative filtering technique [16] and a content based technique [14]. The first algorithm is based only on the topology of Twitter network. It first explores the connections starting at the target user (the user to whom we wish to recommend new followees) in order to select a set of candidate recommendations and finally it ranks those candidates according to a scoring function. The scoring function we design involves three factors that take into account the most influential properties of the Twitter network, according to previous studies. The second algorithm first creates a vector of terms describing the interests of the target user based on the tweets published by his/her followees. This vector is then used to discover new users that might not belong to the target user neighborhood in spite of being similar to him/her. Since these users are not taken from the connections starting at the target user, in principle they would not be discovered by the topology-based algorithm.

Unlike other works that focus on ranking users according to their influence in the entire network [18,19], the algorithm we propose explores the follower/following relationships of the user up to a certain level, so that more personalized factors are considered in the selection of candidates for recommendation, such as the number of mentions of these candidates. Furthermore, the approach proposed in this work was evaluated with a controlled experiment with real users. From the experiments performed we found that although the average precision tend to be similar for both algorithms, the content-based approach is better at positioning relevant recommendations at the top of the ranking.

The rest of this work is organized as follows. Section 2 describes some aspects about Twitter and discuss how related work is related to our research. Next, in Section 3 we describe our approach to the problem of followee recommendation in Twitter. In Section 4 we present the experiments we performed to validate our proposal. Finally, in Section 5 we discuss the results we obtained and present our conclusions and future work.

---

[1] In 2010 Twitter grew by more than 100 million registered accounts. http://yearinreview.twitter.com/whosnew/. Accessed on April 2011

2 Background and Related Work

Twitter is a social network with micro blogging service that enables users to send and receive messages with a length shorter than 140 characters that are called "tweets" or status updates. Relationships in Twitter are unidirectional: a Twitter user *U* interested in the tweets published by another user registers himself as a "follower". Although user *U* has no need to follow their followers back, it is possible for him/her to obtain the list of users following him/her.

As stated above, the Twitter network is populated with tweets. Tweets can have any (textual) content; however there exist users that only publish tweets about a particular subject, such as sports, movies, music or a about a particular rock band. These users can be considered as information sources or broadcasters. In contrast, many people uses Twitter to get information on particular subject, as a form of RSS reader, registering themselves as followers of their favorite artists, celebrities, bloggers, or TV programs. For this last type of users finding high quality and reliable information sources in the constantly increasing Twitter community becomes a challenging issue.

Several recent research efforts have been dedicated to understand micro-blogging as a novel form of communication and news spreading medium. Java et al. [8] and Krishnamurthy et al. [10] presented a characterization of Twitter users grouping them into three categories. The first correspond to *information sources* or *broadcasters*, which are users that are characterized by having a much larger number of followers than they themselves are following. The second category groups *information seekers*, users who rarely post a tweet authored by themselves but that regularly follows other users. Finally, users categorized as *friends* or *acquaintances* are users that tend to use Twitter as a typical on-line social network and are characterized by reciprocity in their relationships.

The influence of users in Twitter has also been subject of several studies. In [11] it was shown that ranking users by the number of followers and by their PageRank give similar results. However, ranking users by the number of re-tweets indicates a gap between influence inferred from the number of followers and that from the popularity of users' tweets. Coincidentally, a comparison between in-degree, re-tweets and mentions as influence indicators [1] concluded that the first is more related to user popularity. Analyzing spawning re-tweets and mentions, it was found that most influential users hold significant influence over a variety of topics but this influence is gained only through a concentrated effort (such as limiting tweets to a single topic). TwitterRank [18], an extension of PageRank algorithm, tries to find influential twitterers by taking into account the topical similarity between users as well as the link structure. Garcia et al. [5] propose a method to weight popularity and activity of links for ranking users. User recommendation, however, can not be based exclusively on general influence rankings since people get connected for multiple reasons.

While the mentioned studies focus on analyzing micro-blogging usage, other works try to capitalize the massive amount of user-generated content as a novel source of preference and profiling information for recommendation. Chen et al. [3] proposed an approach to recommend interesting URLs coming from information streams such as tweets based on two topic interest models of the target user and a social voting mechanism so that the most popular URLs within the group are recommended. *Buzzer* [15] indexes tweets and recent news appearing in user specified feeds, which are considered as examples of user preferences, to be matched against tweets from the public timeline or from the user Twitter friends for story ranking and recommendation. Esparza et al. [4] address the problem of using real-time opinions of movie fans expressed through the Twitter-like short textual reviews for recommendation. This work assumes that tweets carry on preference-like information that can be used in content-based and collaborative filtering recommendation.

In contrast to the previous works that address the problem of suggesting potentially relevant content from micro-blogging services, we concentrate in recommending interesting people to follow. In this direction, Sun et al. [17] proposes a diffusion-based micro-blogging recommendation framework which identifies a small number of users playing the role of news reporters and recommends them to information seekers during emergency events. Closest to our work are the algorithms for recommending followees in Twitter evaluated and compared using a subset of users in [7]. Multiple profiling strategies were considered according to how users are represented in a content-based approach, a collaborative filtering approach and two hybrid approaches. User profiles are indexed and recommendations generated using a search engine, receiving a ranked-list of relevant Twitter users based on a target user profile or a specific set of query terms. Our work differs from this approach in that we do not require indexing profiles from Twitter users, instead topology-based and content-based algorithms explored the follower/followee network in order to find candidate users to recommend. Furthermore, we consider in the evaluation of our approach the target user assessment about the his/her interest in the provided recommendations.

Finally, the problem of helping users to find and to connect with people on-line to take advantage of their friend relationships has been also studied in the context of social networks. For example, SONAR [6] recommends related people in the context of enterprises by aggregating information about relationships as reflected in different sources within a organization, such as organizational chart relationships, co-authorship of papers, patents, projects and others. [12] presented different methods for link prediction based on node neighborhoods and on the ensemble of all paths. These methods were evaluated using co-authorship networks. Authors found that there is indeed useful information contained in the network topology alone. Chen et al. [2] compared relationship-based and content-based algorithms in making people recommendations, finding that the first ones are better at finding known contacts whereas the second ones are stronger at discovering new friends. Weighted minimum-message ratio (WMR) [13] is a graph-based algorithm which generates a personalized list of friends in a social network built according to the observed interaction among members. Unlike these algorithms that gathered social networks in enclosed domains from structured data (such as interactions, co-authorship relations, etc.), we proposed two algorithms to take advantage of the massive, unstructured, dynamic and inherently noisy user-generated content from Twitter.

3 Followees Recommendations on Twitter

We have designed two different algorithms for followee recommendation on Twitter. The first algorithm is only based on the topology of the followers/followees network and suggests users that are neighboring the target user up to some distance. The second algorithm is content-based and aims at suggesting users that may not be in the neighborhood of the target user, but whose tweets may be interesting to him/her.

3.1 Topology-based recommender

The general idea behind this algorithm is to suggest users that are in the neighborhood of the target user and that can be potential followees. A user's neighborhood is determined from the follower/followee relations in the social network. We apply the following heuristic to obtain the list of candidate users for recommendation:

1. Starting with the target user u_T, obtain the list of users he/she follows, let's call this list S, i.e. $S(u_T) = \bigcup_{\forall f \in followees(u_T)} f$.
2. For each element in S get its followers, let's call the union of all these lists L, i.e. $L(u_T) = \bigcup_{\forall s \in S} followers(s)$.
3. For each element in L obtain its followees, let's call the union of all these lists T, i.e.$T(u_T) = \bigcup_{\forall l \in L} followees(l)$.
4. Exclude from T those users that the target user is already following. Let's call the resulting list of candidates R, $R = T - S$.

Each element in R is a possible user to recommend to the target user. Notice that each element can appear more than once in R, depending on the number of times that each user appears in the the followees or followers lists obtained at steps 2 and 3 above.

The rationale behind this heuristic procedure is that the target user is an information seeker that has already identified some interesting users acting as information sources, which are his/her followees. Other people that also follows some of the users in this group (i.e. is subscribe to some of the same information sources) have interests in common with the target user and might have discover other relevant information sources in the same topics, which are in turn their followees.

Finally, we give each unique user $u_c \in R$ a score given by the Equation 1:

$$score(u_c) = \frac{occurrences(u_c, R)}{|R|} \times \frac{|\, followers(u_c)\,|}{|\, followees(u_c)\,|} \times \frac{|\, mentions(u_c)\,|}{M} \tag{1}$$

The first term corresponds to the number of occurrences of the user in the final list of $|R|$ candidates for recommendations. The number of occurrences of a user u_c in this final list is an indicator of the amount of (indirect) neighbors that also have u_c as a (direct) connection itself.

The second term is the relation between the number of followers a user has with respect to the number of users that he/she follows. Since we seek for information sources to be recommended, we assume that this kind of users will have many followers and that they will follow few people. In [11] it has been shown that the rankings of users that can be obtained by number of followers and by PageRank are very similar. We opted to use this factor as an estimator of the "importance" of a given user because the number of followers is a metric by far more easy to obtain that the PageRank score in a network with an order of almost 2 billion social relations. Cha et al. [1] also support the fact that the number of followers, along with both retweets and mentions, are factors that represents a user's influence on Twitter. They found that while the number of followers is an indicator of a user's popularity, retweets and mentions represent other important factors such as engaging the audience with valuable content.

For the reason expressed above, we finally add a factor that considers the number of times that a user has been mentioned in the social network in recent posts. According to Kwak et al. [11] ranking Twitter users by the number of retweets shows the rise of micro-blogging as an alternative communication media. In other words, retweets are considered the feature that has made Twitter a new medium of information dissemination. Hence, we consider mentions of a user instead of retweets because mentions are a broader concept that includes retweets. The most recent mentions to a user can be easily obtained through Twitter's Query API, up to a maximum of M mentions. Currently M is set to 100.

3.2 Content-based recommender

Information seekers are characterized for posting few tweets themselves, but follow people that generate content more actively. It is assumed that users actively select their followees expecting that their tweets will be interesting to them. Then, in order to develop a content-based followees recommender algorithm, we assumed that the interests of the target user can be described by the content of the tweets published by the users he/she follows. Let $tweets(u_f) = \{t_1, t_2, \cdots, t_k\}$ be the set of tweets published by user u_f, $profile_{base}(u_f)$ the term vector built from $tweets(u_f)$, and $followees(u_T) = \{f_1, f_2, \cdots, f_l\}$ the followees of user u_T. Then the profile of a user u_T is defined as the union of term vectors of his/her followees:

$$profile(u_T) = \bigcup_{\forall u_f \in followees(u_T)} profile_{base}(u_f)$$

In order to search for candidate recommendations, this algorithm does not take candidate users from the topology of the social network. Instead, it aims at discovering new users that may not be connected to the target user by a short path in the graph but appear in an information stream provided by Twitter which is known as *public timeline*. This stream contains the collection of the most recently published tweets, and is fed by all accounts that are not configured to be private. The public timeline can be considered as the current flow of information in Twitter, and is a good source to obtain active users in the social network.

The content-based algorithm we designed works as follows:

1. Obtain the authors of the most recent publications that appear in Twitter's public timeline, $U = \{u_1, u_2, \ldots, u_m\}$.

2. For each user $u_C \in U$, build $profile_{base}(u_C)$. That is, we build the term vector corresponding to each u_C .
3. For each user $u_C \in U$, compute

$$sim(u_C, u_T) = max_{\forall i: f_i \in followees(u_T)} sim_{cos}\left[profile_{base}(f_i), profile_{base}(u_C)\right]$$

   Where sim_{cos} is simply the cosine similarity between the two vectors.
   If $sim(u_C, u_T) > \gamma$, add u_C to the list of recommendations ordered by similarity.
4. Repeat steps 1 to 4 until the desired number of recommendations is obtained.

In order to build the term vectors associated to users, we first detect the language of the tweets[2] and then we apply the corresponding stop-word and stemming filters. We use a term frequency weighting scheme in the term vectors.

We use a similarity threshold of $\gamma = 0.1$ to consider a user relevant for recommendation. This threshold was set very low so that the desired number of recommendations could be obtained in a reasonable time. However, it can be adjusted according to the recommender application. For example, if recommendations can be calculated off-line the threshold can be set to a higher value, likely improving the precision of recommendations, at expense of some additional calculation time.

4 Experimental evaluation

4.1 Experiment setup

In order to evaluate the proposed algorithms, we have carried out a preliminary experiment using a group of 26 users. These users, 20 males and 6 females, were in the last years of their course of studies and were students of a "Recommender Systems" course dictated at our university as an elective course during 2010. The students selected for the experiment were volunteers familiarized with Twitter. We asked these users to create a new Twitter account[3] and to follow at least 20 Twitter users who publish information or news about a set of particular subjects of their interest. The general interests expressed by users ranged among diverse subjects. Some users only concentrated on one particular subject while others distributed their followees among several topics. Then we provided these users with a desktop tool that allowed them to login to Twitter and ask for recommendations using both methods (topology-based and content-based). The tool offered the logged user 20 recommended users and we asked them to explicitly evaluate whether the recommendations were relevant or not according to the same topical criteria they have chosen to select their followees as information sources.

For each recommendation in the resulting ranking the application showed the user name, description, profile picture and a link to the home page of the corresponding account. This link could be used to read the tweets published by the recommended user in the case that the information provided by the application was not enough to determine the student's interest in the recommendation. The question we asked students

[2] We are currently working with English and Spanish.

[3] We asked students to create a new Twitter account so that they did not need to reveal their real account and the people they follow

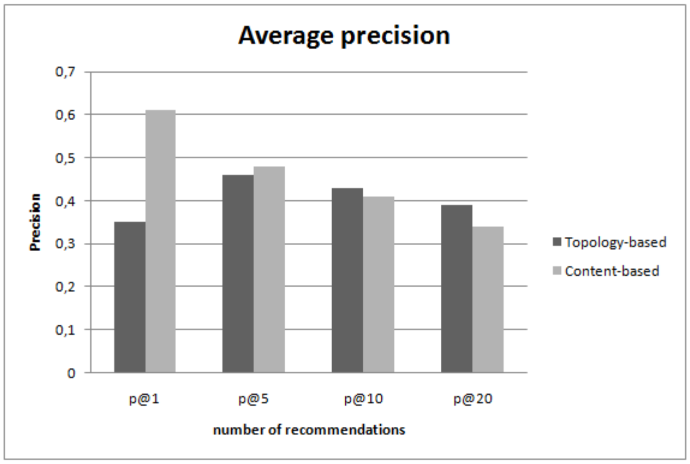

Fig. 1. Average precision for both recommendation algorithms

to ask themselves to determine whether a recommendation was relevant or not was "Would you have followed this recommended user in the first place (when selecting which users to follow in the first part of the experiment), if you had know this account?" For example, if a given student was interest in technology and he/she had not discovered the account @TechCrunch during his/her first selection of followees, that would be an interest recommendation because @TechCrunch tweets about news on technology.

4.2 Results

The quality of lists of top-N followee recommendations generated by each algorithm was first evaluated in terms of their overall precision. *Precision* can be defined as the number of relevant recommendations over the number of recommendations presented to the user and it can be also computed at different positions in the ranking. For example, P@5 ("*precision at five*") is defined as the percentage of relevant recommendations among the first five, averaged over all runs. Figure 1 shows the precision obtained for both algorithms at four different positions of the ranking: P@1, P@5, P@10 and P@20.

In general, both algorithms perform similarly at different positions of the ranking, with the exception of P@1 for which the content-based approach clearly outperforms the topology-based algorithm (61% of relevant recommendations for the content-based algorithm against 35% of relevant recommendations for the topology based-algorithm). For P@5 we obtained 48% of relevant recommendations for the content-based algorithm and 46% of relevant recommendations for the topology based algorithm. At this point we should point out that although we report precision up to 20 recommendations, recommender systems generally present to users shorter recommendations lists aiming at helping them to focus on the most relevant results. In these small lists the content-based algorithm reached good levels of precision, recommending mostly interesting users.

For recommendations lists longer than 5, performance decreases and we can observe that the topology based algorithm tends to give better results than the content-based algorithm. However the difference in performance of both algorithms is always lower than 5%. Due to the reduced number of users who participate in the experiment, we performed the Student's t-test of significance on the results obtained. The Student's t-test looks at the average difference between the performance scores of two algorithms, normalized by the standard deviation of the score difference. For this test we obtain that only the difference in precision at the first position of the ranking (P@1) is statistically significant.

Although precision measure gives a general idea of the overall performance of the presented algorithms, it is also very important to consider the position of relevant recommendations in the ranking presented to the user. Since it is known that users focus their attention on items at the top of a list of recommendations [9], if relevant recommendations appear at the top of the ranking using one algorithm and at the bottom of the ranking using the other, the first algorithm will be perceived by users as performing better even though their general precision might be similar.

Discounted cumulative gain (DCG) is a measure of effectiveness used to evaluate ranked lists of recommendations. DCG measures the usefulness, or gain, of a document based on its position in the result list using a graded relevance scale of documents in a list of recommendations,. The gain is accumulated from the top of the result list to the bottom with the gain of each result discounted at lower ranks. The premise of DCG is that highly relevant documents appearing lower in a list should be penalized as the graded relevance value is reduced logarithmically proportional to the position of the result. The DCG accumulated at a particular rank position k is defined as $DCG@k = rel_1 + \sum_{i=2}^{k} \frac{rel_i}{log_2 i}$. DCG is often normalized using an *ideal DCG* that is computed by sorting documents of a result list by relevance. Figure 2 shows the normalized DCG obtained for both algorithms at four different positions of the ranking: nDCG@1, nDCG@5, nDCG@10 and nDCG@20.

Success at rank k (S@k) is another metric commonly used to evaluate ranked lists of recommendations. The success at rank k is defined as the probability of finding a good recommendation among the top k recommended users. In other words, S@k is the percentage of runs in which there was at least one relevant user among the first k recommended users. Figure 3 shows the results we obtained for this metric for values of k ranging from 1 to 6.

S@1 is equivalent to P@1 by definition. Then, we can see that the content-based algorithm always positions relevant users earlier in the ranking than the topology-based algorithm. Indeed, all users in the experiment found a relevant recommendation before position 4 in the ranking using the content-based algorithm. For the topology-based algorithm, most users found a relevant recommendation before position 5 except for one user that found the first relevant recommendation at rank 6.

5 Discussion and Conclusions

In this article we presented two simple but effective algorithms for recommending users in the Twitter social network. The first algorithm models a given user from his/her con-

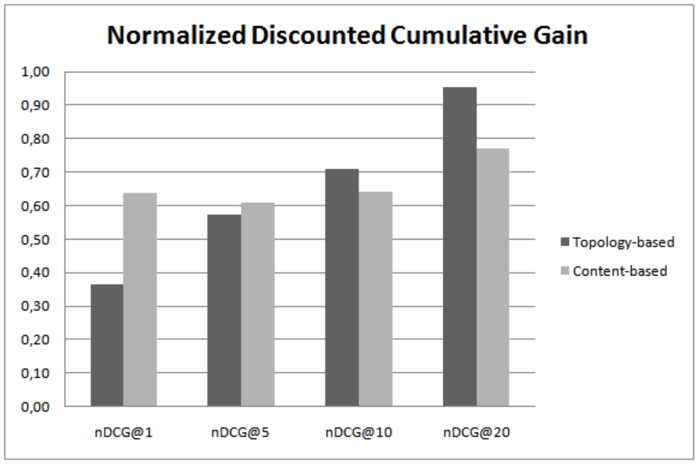

Fig. 2. Normalized discounted cumulative gain for both recommendation algorithms

nections in the social graph whereas the second algorithm models users using the content of the tweets published by his/her followees. We evaluated the proposed algorithms with real users and found that they work fairly similar in finding users that might result interesting for the target user to start following.

From the experiments presented we can conclude that although the average precision tend to be similar for both algorithms, if we consider the position on the recommendations in the ranking the content-based approach is better at giving good recommendations. We believe that results obtained with the content-based algorithm can be improved by setting a higher threshold for the similarity measure used for filtering the term vectors representing users. However, this will increase the response time of the algorithm since users are taken randomly from Twitter's public timeline.

Among the advantages of the topology-based algorithm, on the other hand, we can mention that recommendations can be found quickly based on a simple analysis of the network structure, without considering the content of the tweets posted by the candidate user.

Although the results reported seems promising, we are planning to repeat the experiment this year in order to involve more users in the experiment and obtain more statistical significance about the two proposed algorithms.

A natural extension in which we are currently working on is a hybrid algorithm that combines the best of both algorithms presented in this paper. This hybrid algorithm filters the candidate recommendations found with the topology-based method with a content-based analysis of the tweets posted by the candidate users. We are also very optimistic about the potential improvements that could be obtained with this hybrid approach.

As a possible limitation of our approach, we can mention that, we assumed that the target user is an *information-seeker* user, according to the categorization proposed by [8]. However, users may play different (and multiples) roles of information source,

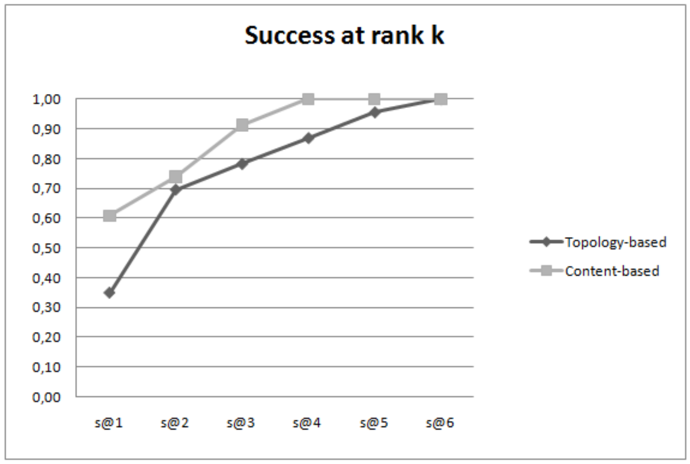

Fig. 3. Success at rank *k* for both recommendation algorithms

information seeker or friends in different communities. This is a challenging factor to consider that we leave for future investigation.

The experiments presented make us feel optimistic about the potential of a followee recommender system for Twitter using the methods described in this article or a combination of them. This work is the first step towards exploring the great potentials of this new platform to build recommendation systems.

References

1. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.: Measuring user influence in Twitter: The million follower fallacy. In: Proc. of the 4th Int. Conf. on Weblogs and Social Media (ICWSM'10). Washington DC, USA (2010)
2. Chen, J., Geyer, W., Dugan, C., Muller, M., Guy, I.: Make new friends, but keep the old: recommending people on social networking sites. In: Proc. of the 27th Int. Conf. on Human Factors in Computing Systems. pp. 201–210 (2009)
3. Chen, J., Nairn, R., Nelson, L., Bernstein, M., Chi, E.: Short and tweet: experiments on recommending content from information streams. In: Proc. of the 28th Int. Conf. on Human Factors in Computing Systems (CHI'10). pp. 1185–1194 (2010)
4. Esparza, S.G., O'Mahony, M.P., Smyth, B.: On the real-time web as a source of recommendation knowledge. In: Proc. of the 4th ACM Conf. on Recommender Systems (RecSys'10). pp. 305–308 (2010)
5. Garcia, R., Amatriain, X.: Weighted content based methods for recommending connections in online social networks. In: Workshop on Recommender Systems and the Social Web. pp. 68–71. Barcelona, Spain (2010)
6. Guy, I., Ronen, I., Wilcox, E.: Do you know?: recommending people to invite into your social network. In: Proc. of the 13th Int. Conf. on Intelligent User Interfaces (IUI'09). pp. 77–86 (2009)
7. Hannon, J., Bennett, M., Smyth, B.: Recommending Twitter users to follow using content and collaborative filtering approaches. In: Proc. of the 4th ACM Conf. on Recommender Systems (RecSys'10). pp. 199–206 (2010)

8. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: Proc. of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis. pp. 56–65 (2007)
9. Joachims, T., Granka, L., Pan, B., Hembrooke, H., Gay, G.: Accurately interpreting click-through data as implicit feedback. In: Proc. of the 28th Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR'05). pp. 154–161 (2005)
10. Krishnamurthy, B., Gill, P., Arlitt, M.: A few chirps about Twitter. In: Proc. of the 1st Workshop on Online Social Networks (WOSP'08). pp. 19–24 (2008)
11. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media? In: Proc. of the 19th Int. Conf. on World Wide Web (WWW'10). pp. 591–600 (2010)
12. Liben-Nowell, D., Kleinberg, J.: The link prediction problem for social networks. In: Proc. of the 12th Int. Conf. on Information and knowledge management. pp. 556–559. CIKM '03, ACM, New York, NY, USA (2003)
13. Lo, S., Lin, C.: WMR–A graph-based algorithm for friend recommendation. In: Proc. of the 2006 IEEE/WIC/ACM Int. Conf. on Web Intelligence (WI'06). pp. 121–128. Washington, DC, USA (2006)
14. Pazzani, M.J., Billsus, D.: Content-based recommendation systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The adaptive Web, pp. 325–341. Springer-Verlag (2007)
15. Phelan, O., McCarthy, K., Smyth, B.: Using Twitter to recommend real-time topical news. In: Proc. of the 3rd ACM Conf. on Recommender Systems (RecSys'09). pp. 385–388 (2009)
16. Schafer, J., Frankowski, D., Herlocker, J., Sen, S.: Collaborative Filtering Recommender Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, LNCS, vol. 4321, chap. 9, pp. 291–324. Springer Berlin Heidelberg, Berlin, Heidelberg (2007)
17. Sun, A.R., Cheng, J., Zeng, D.D.: A novel recommendation framework for micro-blogging based on information diffusion. In: Proc. of the 19th Workshop on Information Technologies and Systems (2009)
18. Weng, J., Lim, E.P., Jiang, J., He, Q.: TwitterRank: finding topic-sensitive influential twitterers. In: Proc. of the 3rd ACM Int. Conf. on Web Search and Data Mining (WSDM'10). pp. 261–270. New York, NY, USA (2010)
19. Yamaguchi, Y., Takahashi, T., Amagasa, T., Kitagawa, H.: TURank: Twitter user ranking based on user-tweet graph analysis. In: Web Information Systems Engineering. LNCS, vol. 6488, pp. 240–253. Hong Kong, China (2010)