Peer recommendation based on comments write on social networks

Silvana Aciar[1] y Gabriela Aciar[1]

[1]Universidad Nacional de San Juan, Argentina.
saciar@iinfo.unsj.edu.ar gaby_aciar@yahoo.com.ar

Abstract. Social networks and virtual communities has become a popular communication tool among Internet users. Millions of users share publications about different aspects: educational, personal, cultural, etc. Therefore these social sites are rich sources of information about who can help us solve any problems. In this paper, we focus on using the written comments to recommend a person who can answer a request. An automatic analysis of information using text mining techniques was proposed to select the most suitable users. Experimental evaluations show that the proposed techniques are efficient and perform better than a standard search.

Keywords: Social networks, text mining, recommender systems

1 Introduction

Social networks have become in a few years, a global phenomenon that is expanding every day. Twitter, Facebook, Google+ and LinkedIn have emerged with overwhelming force in society. Social technologies are radically changing the way people communicate and interact.

People often consult the Internet to find answers to your questions. Perform queries on Google, which returns thousands of web pages with possible answers. A person can take a long time to find the right answer among all received pages. This leads to abandon the search or accept an answer that may not be the most appropriate.

Due to overload of information on the Internet, people prefer to ask a friend. The way to communicate is using social media, today a person is part of at least 3 social networks and has more than 100 contacts. In large networks, identify the right person who can answer a specific question is not an easy task for one person, sometimes the right person is not directly related to him.

One way to facilitate interaction and solve the problems of information overload is through the use of recommender systems [1][2]. The purpose recommender systems are to simplify the search process by providing to users information, products or services based on their needs, interests, and preferences. The recommender systems use techniques from several disciplines such as artificial intelligence, information retrieval, data mining and machine learning to identify items of interest for a user in a particular context of recommendations[2][3][4].

This paper proposes to use recommender systems to suggest suitable users to answer queries of people. In the literature there are several studies that recommend users to interact in a social network [5][6][7] none of which used the information posted by users.

In this paper we consider necessary to analyze the content of user's post to get their knowledge about a particular topic, and then suggest users to answer questions based on that value.

Figure 1 shows an example of comments that will be analyzed in this work.

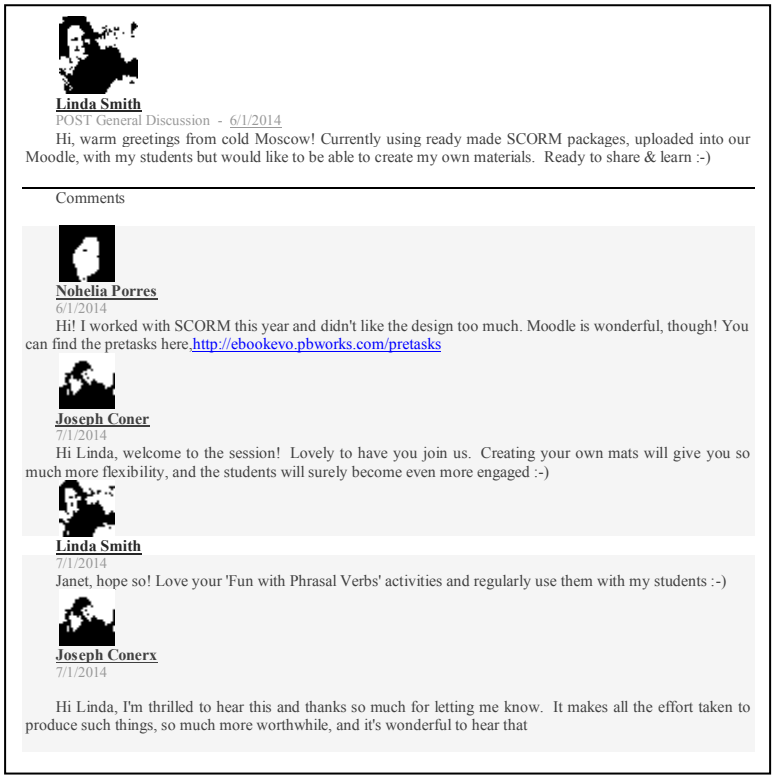

Fig. 1. Sample comments which are analyzed to extract information from users.

Despite the importance and value of the information introduced by users in a comment, there is no comprehensive mechanism that formalizes in an application:

- The process of selection and retrieval of user comments.
- Information processing to obtain a value representing the knowledge of users on a topic.
- Recommendation of a user based on that value.

Part of the problem lies in the complexity of information extraction from textual and turns that information into recommendations. Figure 2 shows the structure of the proposed recommender system.

The structure of the paper is as follows. Section 2 presents related works about the recommendation of users using information from social networks. In Section 3 we describe in details the mining comments process, which uses text mining techniques. The set of measures used to calculate a recommender value of a user is also briefly explained in this section. Section 4 presents the experiments performed that demonstrates the effectiveness of proposed approach. The example is in a virtual community on the e-learning domain. Finally, the conclusions, the limitations of the

work and the directions for future research are discussed in Section 5.

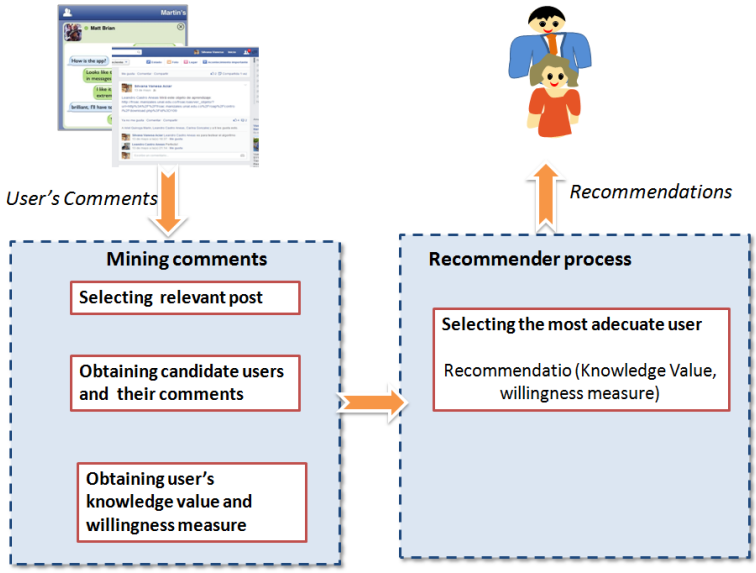

Fig. 2. Process to recommend user based on information from comments.

2 Related works

While in the community of recommender systems has been much progress in the study of methods to make recommendations of products and services [2] [3] [4] [8], there are few work of recommending people. The methods used in recommending users take into account the similarity between the profiles, reputation within a community and network of contacts.

In [9][10] recommendations are made based on the similarity of user profile on a social network. The profile is constructed explicitly. The profile and reputation of individuals in a virtual community are considered to make recommendations in [11][12].

Recommendations are currently being conducted using the network of contacts of people [13][14]. The virtual community is modeled as a graph consisting of nodes representing individuals and links between those nodes that usually represents the distance between people. Once the network structure is obtained, recommendations are made based on the type, organization and network properties [5][6][7][15][16].

From the review, has not found work that uses information from user's comment to recommend people to interact. The content of the comment have to be analyzed in order to decide who can answer a question for those who cannot. The following sections detail the process for recommending users based on comments written in virtual communities.

3 Recommendation of users based on comment posted

The following tasks are required to select and retrieve information from the comments, process that information and use it to know if a user is suitable for recommendation:

- Implementation of text mining methods to obtain information from written comments by users.
- The definition of metrics that allow prioritized users by the degree of their knowledge on a specific topic.
- Developing recommender mechanisms to filter and present the most appropriate users.

The process will be done through an application that given a user request Q, it searches user comments containing Q obtaining set of relevant post. A set of metrics for obtaining recommending users from the set of relevant post have been defined. In next sections are explained the process to obtain such measures.

3.1 Selection of relevant post and identification of candidate users

Given a user request Q of a user and Q is composed of keywords such that $Q = (q_1, q_2, ..., q_n)$, the first step is to find all the post containing Q. The search for relevant post is a simple search where those publications containing Q chains (q_i) are only selected. The selected post constitute a collection of relevant publications $Pr = \{p_1, p_2, p_3, .. p_n\}$ where each publication can have an associated set of comment and users who wrote them. Figure 3 shows the structure of the data obtained from relevant publications.

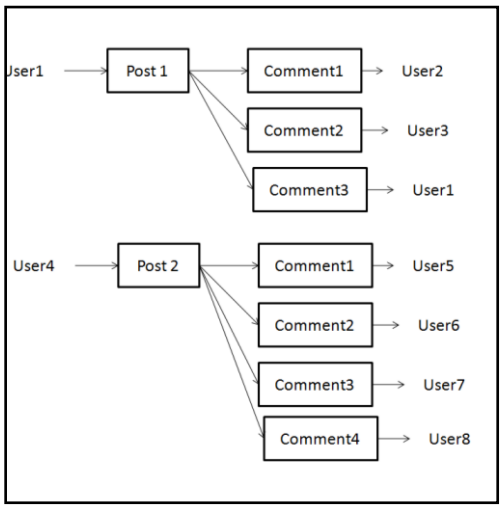

Fig. 3. Information obtained from relevant post

Users who made comments and those who made the publication constitute a set of candidate users $UC = \{u_1, u_2, u_3, ..., u_n\}$. For each candidate user, comments that were made and the times they interacted in the virtual community are obtained.

3.2 Analyzing user comments

A user can answer a request Q if he has knowledge about Q and he is available to answer questions. We have defined a set of metrics to obtain such information. The **User Information about a Topic (UIT)** measure is computed with information written on comments while the **User Interaction (UI)** and **Willingness to Respond (WR)** measures are computed using information from interactions in the virtual community.

User Information about a Topic (UIT) is defined as the sum of weights (w_i) of the words q_i multiplied by the times C that these words appear together in user comments. It is calculated using equation 1.

$$UIT = \sum_{i=1}^{n} w_i * C \quad \textbf{(1)}$$

w_i is obtained by means of the measure tf-idf (Term Frequency-Inverse Document Frequency) widely used in Information Retrieval. In our proposal it means that the q_i terms that appear most frequently written in a user's comments, but less in the other reviews, is more probable that it be most used by the user. Equations 2, 3 and 4 are employed to compute the measure.

$$w_i = f_i * idf_i \quad \textbf{(2)}$$

$$f_i = \frac{frec_i}{MaxFrec} \quad \textbf{(3)}$$

$$idf_i = \log \frac{N}{n_i} \quad \textbf{(4)}$$

f_j is the normalized term frequency of q_i in the comments made by a user. Idf_i represents the inverted term frequency of q_i. $frec_i$ is the frequency of term q_i in all comments. $MaxFrec$ represents the maximum frequency of all frequencies of words in the comments. N is the number of comments and n_i is the number of comments containing the term q_i.

Interactions User (IU): is defined as the times that a user has posted or has commented NCU over the total number of post and comments on the virtual community TNC, as indicated by Equation 5.

$$IU = \frac{NCU}{TNC} \quad \textbf{(5)}$$

User's knowledge on the topic (UKT): it is obtained based on the user information on the topic UIT multiplied by the number of interactions of the user in the virtual network IU. Equation 6 is used to implement this measure.

$$UKT(u) = UIT * IU \quad \textbf{(6)}$$

Willingness to Respond (WR) is defined as the probability that a user responds to a request based on past behavior. It is probably that a person who has frequently answered questions in the past, now answer a question. Equation 7 is used to calculate the willingness of reply.

$$WR(u)\frac{RPOS+1}{(RPOS+1)+(RNEG+1)} \quad (7)$$

Where *RPOS* is the number of times that the user has responded request and RNEG is the total number of time that he has not responded.

3.3 Selecting the most suitable user

The most suitable and reliable user are chosen to make the recommendations. A selection algorithm has been defined to make the choice automatically. The algorithm is composed by 3 elements, a set *(U)* of candidate users, a selection function *Selection(UKT(u),WR(u))* to obtain the most relevant and reliable user which uses the values of user's knowledge about a topic *UKT(u)* and willingness to respond WDR(u) as parameters and a solution set *(F)* containing the users selected *($F \subset U$)*. With every step the algorithm chooses a user of *U*, let us call it *u*. Next it checks if the $u \subset F$ can lead to a solution; if it cannot, it eliminates *u* from the set *U*, includes the user in *F* and goes back to choose another. If the users run out, it has finished; if not, it continues.

Algorithm to select relevant users

```
Algorithm (U: Set of candidates users)
    F: =∅  ;
      while (U <>∅) do
        if Selection(UKT(u),WR(u))  > threshold then
          F := F∪u;
            Eliminate (U, u)
      end if
    end while
    return F;
```

The selection function has as parameters the information of the user about a topic UKT(u) and the availability to respond WR(u). This function is obtained through equation:

$$selection(UKT(u),WR(u)) = UKT(u) * WR(u) \quad (8)$$

The lists of relevant users are ordered from highest to lowest selection value and the top of the list are recommended.

4. Experimental results

The domain of education was chosen to carry out the experiments and evaluate the feasibility of the proposal. A user recommender system has been implemented in the Moodle platform where post made in forums for teachers and students are analyzed. Mooldle forums allow users to post information about a topic through comments. A post on a Moodle forum includes a section containing the user name of the writer who can be real or fictitious. A section describing the content of the post written by the user and a comment section where other users can comment the post. For the experiments have been analyzed data from the participation of 22 students and teachers in several forums of Moodle during 2013. Data collection for testing consists of: 22 users, 45 posts and 134 comments. The recommender system implemented can be seen in Figure 4.

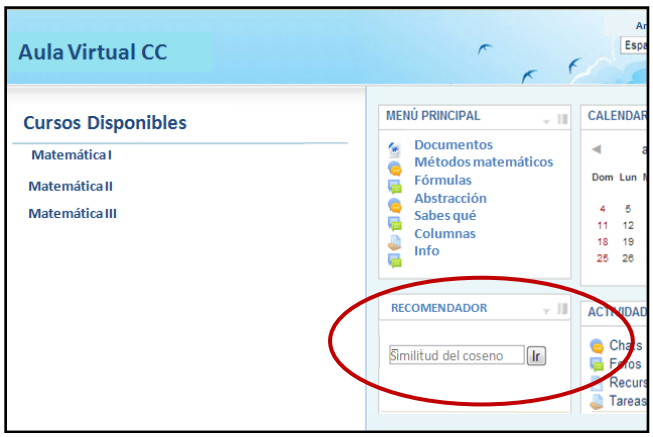

Fig. 4. User recommender system implemented on Moodle platform

Two experiments were performed using the same data set.

Experiment 1 Recommendation based on UKT and WR: The recommender system used the proposed method in this paper to suggest users. 10 users requested by users using the recommender systems. They introduced request about several subjects.

The recommender analyzed forums to obtain candidate users. For each candidate user, the system performances the metrics defined in Section 2.1 with information from their comments. The most suitable user is selected applying the proposed algorithm. When a person has been recommended, the system solicits to user the evaluation of people's behavior. Therefore, the system presents to user a form like is shown Figure 5. The form asks the user to answer yes if the person recommended answered your request. This information is used to compute the *WR* measure presented in section 3.

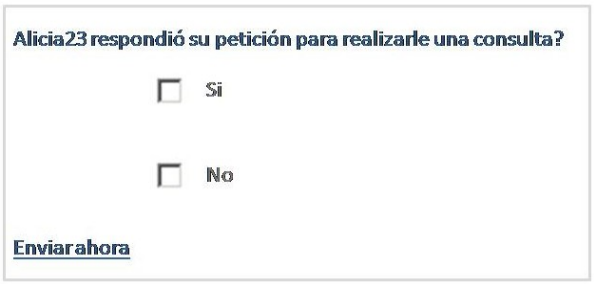

Fig. 5. Form to obtain feedback of the user about the recommendation. User has to answer yes if the recommended person responds his requests.

Experiment 2 Recommendation based on keyword search: In this experiment the recommender system perform a basic search to recommend users. Others 10 users have been requested recommendation. They introduced request about several subjects. The system searches for the keywords in the comments and it recommends users who mentioned more times the key words in your comments.When a person has been recommended, the system solicits to user the evaluation of people's behavior. Therefore, the system presents to user a form like is shown Figure 5. The form asks the user to answer yes if the person recommended answered your request. This information is used to compute the *WR* measure presented in section 3.

4.1 Performance Metric

In order to evaluate the result of recommendations made in both experiments the accuracy rate calculated by equation 9 is used. The evaluation is based on good recommendations over all recommendations, in which a good recommendation is defined as: given a request, a person who will already answer the question is recommended; this information is retrieval from the feedback given by the user after contact with the recommended person (see Figure 5)

$$AccurracyRate = \frac{NG}{N} \tag{9}$$

Where NG is the number of the users who are correctly recommended (recommended person who was responded the question); N is the total number of the recommended person.

Both experiments were performed 10 times with different users each time. Experimental results, as shown in Figure 6 and Table 1, display that the proposed method has a higher accuracy.

Table 1. Evaluation proposed recommendation method. Rate accuracy resulting of the experimentation.

Times	Accuracy Rate Recommendations based on UKT and WR	Accuracy Rate Recommendations based Keywords Search
1	0,90	0,26
2	0,68	0,46
3	0,64	0,24
4	0,74	0,32
5	0,76	0,38
6	0,66	0,44
7	0,92	0,42
8	0,94	0,38
9	0,68	0,38
10	0,78	0,26

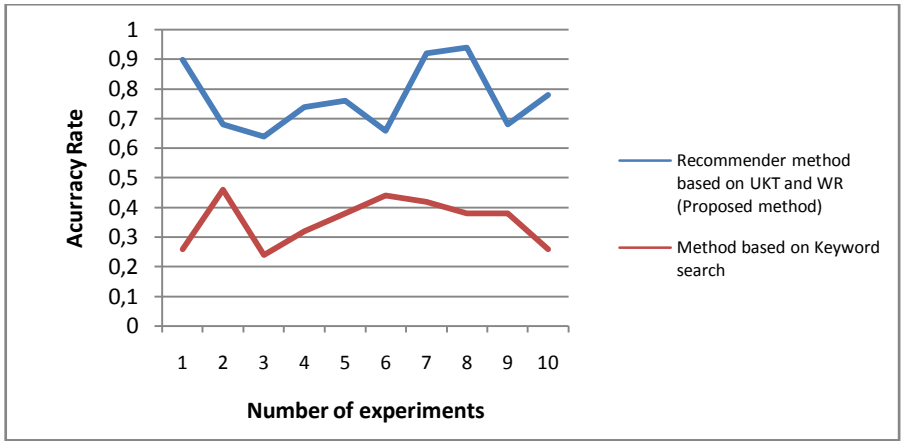

Fig. 6. Accuracy of our proposed method against accuracy of a baseline method.

5. Conclusions

The large amount of information available nowadays makes the process of detecting user for interact more and more difficult. Recommender systems have been used to make this task easier, but the use of these systems does not guarantee that the recommended person be the most suitable. We propose a recommender system that uses information from user's comments to obtain the most suitable to interact with other user. We have defined a set of metrics that allow us know if the user has relevant information or not about a subject to answer a request. Observing the results of the experiments carried out in this paper, we note that the recommendations are better if we use such knowledge. In our future work, we intend to evaluate our approach with open social networks such as Facebook and Google+.

References

1. Konstan, Joseph and Riedl, John . Recommender systems: from algorithms to user experience. User Modeling and User-Adapted Interaction. April 2012, Volume 22, Issue 1-2, pp 101-123
2. Deuk Hee Park, Hyea Kyeong Kim, Il Young Choi, Jae Kyeong Kim, A literature review and classification of recommender systems research, Expert Systems with Applications, Volume 39, Issue 11, 1 September 2012, Pages 10059-10072, ISSN 0957-4174.
3. J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, Recommender systems survey, Knowledge-Based Systems, Volume 46, July 2013, Pages 109-132, ISSN 0950-7051.
4. Mohammad Yahya H. Al-Shamri, Power coefficient as a similarity measure for memory-based collaborative recommender systems, Expert Systems with Applications, Volume 41, Issue 13, 1 October 2014, Pages 5680-5688, ISSN 0957-4174.
5. Ioannis Konstas, Vassilios Stathopoulos, and Joemon M. Jose. 2009. On social networks and collaborative recommendation. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval (SIGIR '09). ACM, New York, NY, USA, 195-202.
6. Mohsen Jamali and Martin Ester. 2010. A matrix factorization technique with trust propagation for recommendation in social networks. In Proceedings of the fourth ACM conference on Recommender systems (RecSys '10). ACM, New York, NY, USA, 135-142.
7. Xiwang Yang, Harald Steck, and Yong Liu. 2012. Circle-based recommendation in online social networks. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '12). ACM, New York, NY, USA, 1267-1275.
8. Cristian E. Briguez, Maximiliano C.D. Budán, Cristhian A.D. Deagustini, Ana G. Maguitman, Marcela Capobianco, Guillermo R. Simari, Argument-based mixed recommenders and their application to movie suggestion, Expert Systems with Applications, Volume 41, Issue 14, 15 October 2014, Pages 6467-6482, ISSN 0957-4174.
9. Henry Kautz, Bart Selman, and Mehul Shah. 1997. Referral Web: combining social networks and collaborative filtering. Commun. ACM 40, 3 (March 1997), 63-65. DOI=10.1145/245108.245123
10. Kang Zhao, Xi Wang, Mo Yu, Bo Gao. User recommendation in reciprocal and bipartite social networks -- a case study of online dating. IEEE Intelligent Systems 2014.
11. Aciar G. and Aciar S. Recomendador de usuarios en una platafoma colaborativa. 14th Argentine Symposium on Artificial Intelligence, ASAI 2013. 42 JAIIO - ASAI 2013 - ISSN: 1850-2784 - Page 1-11.
12. Josep M. Pujol, Ramon Sangüesa, and Jordi Delgado. 2002. Extracting reputation in multi agent systems by means of social network topology. In Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1 (AAMAS '02). ACM, New York, NY, USA, 467-474.
13. Silva, N.B.; Ing-Ren Tsang; Cavalcanti, G. D C; Ing-Jyh Tsang, "A graph-based friend recommendation system using Genetic Algorithm," Evolutionary Computation (CEC), 2010 IEEE Congress on , vol., no., pp.1,7, 18-23 July 2010
14. Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Signed networks in social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10). ACM, New York, NY, USA, 1361-1370. DOI=10.1145/1753326.1753532 http://doi.acm.org/10.1145/1753326.1753532
15. Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Predicting positive and negative links in online social networks. In Proceedings of the 19th international conference on World wide web (WWW '10). ACM, New York, NY, USA, 641-650
16. Ziqi Wang; Yuwei Tan; Ming Zhang, "Graph-Based Recommendation on Social Networks," Web Conference (APWEB), 2010 12th International Asia-Pacific , vol., no., pp.116,122, 6-8 April 2010.