

# Recommendations in social networks: an extra feature or an essential need

Milen Chechev, Ivan Koychev

Faculty of Mathematics and Informatics, Sofia University,  
5 James Baucher blvd, 1164 Sofia, Bulgaria

{milen.chechev, koychev}@fmi.uni-sofia.bg

**Abstract.** This paper analyzes user's need of content recommendation at the social network Facebook. It presents results from a survey on real social network's users. The results shows that Facebook users need better interface for news feed browsing. It have to provide better information filtering options, recommendation system and options for manual refinement of the results from it. The collected information from the survey is used to determine features which an application has to provide as social network news feed browser and to receive user's trust. Further some implementation details and faced difficulties are presented.

**Keywords:** recommender system, social recommendation, Facebook.

## 1 Introduction

Recommendation systems and social networks are two of the hottest topics at the information society at the beginning at 21 century. According to the statistics internet is used by 2,4 billion people [1] and social network Facebook already have more than a billion users [2][3]. These statistics indicate the importance of the social networks as irreplaceable part of the online live of the people.

At the beginning Facebook was created as communication environment for university's students, but it quickly expands and nowadays it's used by everyone for photo and information sharing. Users share information about their education, workplace, personal relationships, their activities and opinions on different social topics. Its simple interface and the abilities for information sharing make it attractive and lead its continuous expansion. However the increasing number of users and content at the social network bring some difficulties to the users about managing information sources. According to Robin Dunbar [17] the human have physical limitation of the social connections, which it can maintain. He determines 150 as limit of the maintainable connections per person. The official Facebook statistics [3] shows that the average number of the friends at Facebook for users, which are more than 2 years at the social network, is 300. This large number of friends cause problems in maintenance of the connections and cause the need of classification on the connections to close friends and acquaintances. Further the users are overloaded with information from

their connections and they spend most of their time to filter their social feed and to find information they are interested in. This problem looks similar to the main problem at the information retrieval - finding useful information in a document's set, the success in dealing with it made Google the most popular web search engine. Unlike it at the social networks the users doesn't know what exactly they are interested in. The information need to be automatically filtered and ranked based on explicit and implicit information about the user's preferences.

## 2 Related Work

Last decade recommendation systems are widely developed and improved, because of their commercial value for increasing of the purchases at online stores. To stimulate the research activity at this field, competitions on real user's data was organized: NetflixPrize[4], ECML-PKDD 2011[6], HetRec[14], KDDCup[5,7]. The series started with the NetflixPrize competition which looks for the best recommendation system for internet site for movie rental. Further a trend of changing the task from recommendation at internet store to recommendation at social network is observed. For example: the HitRec competition exploits data from musical social network – lastfm and internet bookmarks – delicious; KDDCup 2012 data are provided from the most popular Chinese social network Transent Weibo[8].

Recommendations at the social network are based on the relations between the users. There is a research which shows that this relations can be modeled as connection based on trust [9,15,16]. The trust at the social network can be two types – explicit and implicit, as an example of trust network with explicit trust is the review's network Epinions [18]. Massa and Avesani [9] study the dissemination of the trust and distrust at the social networks. Their research shows that the recommendation strategies based on trust provide good results even when the information is insufficient for the others recommendation approaches.

Social Network Facebook can be considered as network of trust. The trust is given when user likes publication of another user As this trust can be considered as implicit because the users didn't explicitly specify that they trust the author. Some recommendation options are already integrated at Facebook. For browsing the news feed, Facebook users have two views – “Most Recent news” or “Top News”. First view shows the news sorted on their publication time. It's comfortable for users which often read their news stream otherwise Facebook suggest the view “Top News”. It sort the news based on several factors like how often the user use the social network, what are the relations between the user and the user who has published the news, etc. [10]. The “Top News” view is useful for users which don't spend much time at the social network, but it works as a black box and the users feels uncertain to use it because there isn't an explanation how exactly the ranking is made and they cannot refine it based on their preferences.

Our research shows that Facebook users need new and flexible application, which provides similar functionality to the “Top News” view of the user's stream, but with additional options for manual refinement and filtering. Next section presents the user survey that determines the main functionality to be included in a social network recommending application.

### 3 Survey of user needs

As part of the experiments to determine necessity of additional application for content recommendations at the social network, an online survey was conducted. Participants at the survey were 114 Facebook users with the following demographic characteristics: 70% of participants are between 18 and 30 years, and 30% older than 30 years. 3% of the asked users have secondary education, 30% have BSc degree, 60% - MSc degree and 7% are PhD. 45% from the participants are males and 55% females.

The questionnaire which all participants answered aims to test the following hypothesis:

- There is a necessity of new application that helps Facebook users and recommends interesting news according to their preferences.
- Users are generally skeptical to recommendation systems that work as black boxes. They want an explanation about the recommendations and options for manual tuning of the system.
- To be more attractive a recommendation application has to provide functionality for information filtering and easily browsing of the content.

Users were asked about the time they spend at the social network. The answers shows that 25% of the users spend just several minutes per day, 47% claim that they spend more than a half hour per day and 28% more than an hour per day. Then users were asked series of questions about their concrete activity at the network. The question “What percent of your time at the social network you spend for an activity?” was asked for the different activities at the social network – reading news feed, browsing friends photos, chat with friends, playing games, etc.

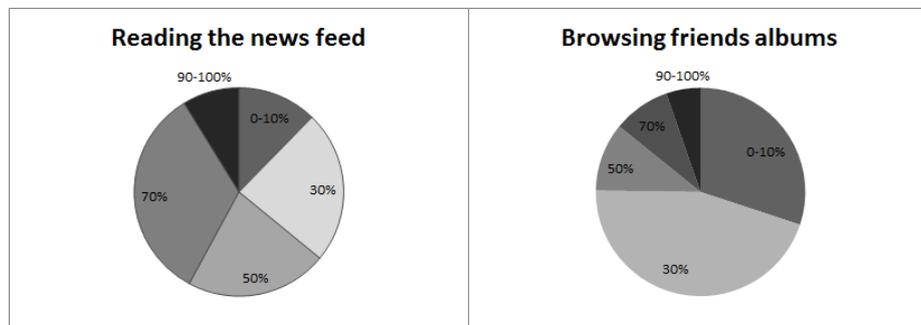
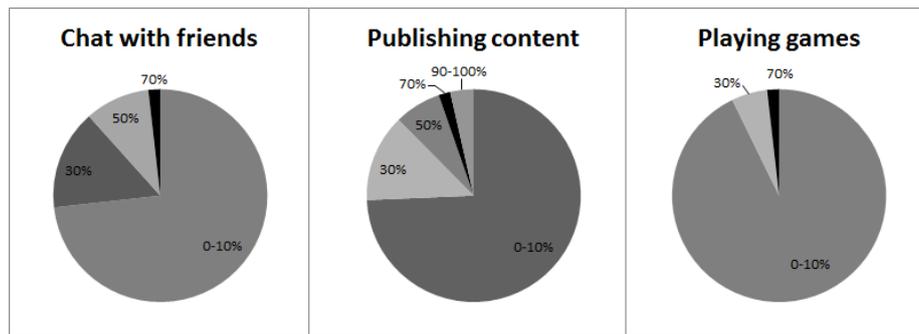


Fig. 1. Main activities at the social network

Users determine news feed reading as the activity, which is the most time consuming, at their interaction with the social network. Answers show that 62% of the people spend more than a half of their time, at the social network, for reading the news feed. Second place for most time consuming activity is hold from the activity – browsing friend’s photo albums. 70% from the users claims that they spend more than 30% from their time for that activity. Detailed distribution of the percentages spend by the users for the activities – reading news feed and browsing friends photo albums can be

seen on fig. 1. The statistics shows that most of the users spend less than 20% for all other activities at the social network. Detail distributions on the percentages spend for the activities – chat with friends, publishing content and playing games is presented on fig.2.



**Fig. 2.** Additional activities at the social network

Next set of questions tried to determine user's needs and what are the new features they require from the social network. 54% of the users answer that they need better search options at the social network. They want to be able to search for the news at their stream in similar way as they can search at Google. 52% of the users wants to have an options to set manually the priority of the different information sources, as only 35% considered automatically selected priority like something they need. Furthermore 54% of the users answer that system that automatically create their user profile is acceptable only if there are options for manual refinement of the automatically selected source's weights. Only 25% of the users consider as acceptable the system to continue drift the weights based on their activity after they manually refine their profile. In addition users report high interest on different filtering options – filter read/unread news, filter by date, by user activity, by publication type, etc.

Users report very low activity for adding new applications. Only 13% of the users claim that they have added at least one application to their profiles at the last month. Nevertheless 69% report willingness to add new application if it gives them possibilities for easy filtering of their news feed and options for automatic and manual calibration of the priority of the different information sources.

The results from the survey confirms the preliminary formulate hypothesis. There is a need of new application that helps users in their everyday use of the social network (69% of the requested users answered that they are ready to use such kind of application). The users are generally skeptical to the recommendation systems (only 25% of them recognize as acceptable their priorities to be automatically calibrated by recommendation system) and more than a half of the users (54%) want to be able to change manually their user profile. Despite user's skeptical opinion about automatic recommendation system, the need of better tool for filtering the content at the news feed determine good chances of fast dissemination. Users are ready to use application that support their everyday news feed activity and this gives the application opportunity of collecting information about more users which will cause and better recommendations.

## 4 Available Information and Its Accessibility

Facebook provides for developers program interface - graph API and FQL queries, which are used to access information about users and their friends, according to strict access policy [11].

Collected data for the purposes of the experiments is in the following categories:

- Objects liked by the user
- Comments by the user
- Shared pages
- Published photos and statuses
- User's news feed
- News published by the user's friends

The access to the friend's data is restricted and even when the application have access to see what user's friends publish to the social network, it cannot access explicitly what user's friends like. Despite this data can be accessed implicitly from the detailed information about the objects accessible by the application.

There are several difficulties during the process of gathering data from Facebook most of them are due to limitations which Facebook have on its program API. The following limitations are noticed:

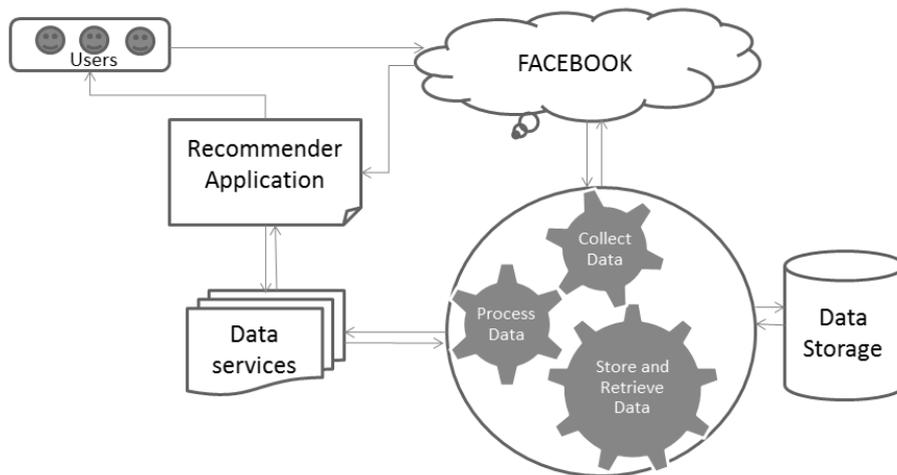
- The maximum number of requests to the Facebook API is 600 for the period of 10 min. This limit restricts an application of retrieving detailed information for each accessible object on regular bases.
- The access to the historical data is limited to:
  - Last 100 liked objects per user.
  - News feed for the last 2 days.
  - Max 420 publications from the user feed.
  - Without additional query there is an access to maximum 5 likes and comments per content.

Gathered data shows the following statistics:

- There are around 200 new contents which are published at the user's feed per 24 hours. This content collects around 400 likes.
- Most of the examined users have more than 300 friends.
- The execution time for execution of a query to the Facebook API depends on the network connectivity and current workload of the social network, as in the provided experiments 50% of the queries took more than 15 sec for response. The time consumption determines the use of asynchronous calls and processing of the results at different threads.

## 5 Implementation

Multilayer architecture is selected for the new Facebook news feed recommender application. Presentation layer include user interface providing the required functionality and which is called by the user from as an embedded environment at the social network. The application layer contains the business services which gather information from the social network process it and send it to the data layer where it is stored. (see fig 3)



**Fig. 3.** Architecture of the recommendation application

When user open the application for first time it ask about permission to access user's basic information and news feed. Once the permissions are granted it starts immediately to process the data as the average time that the application needs to provide results to the user is around half a minute. Meanwhile a welcome page is displayed which provide information about the application in textual and video format.

After the data is processed and the application can provide results, the functionality for reading and filtering of the news feed become active. The system offer search functionality and all gathered objects are included at the search results. The user profile browsing and calibration remains inactive until all the needed information is processed.

After the user profile is created it is used for ranking of the news feed and the user receives access to browse and calibrate it. The offered calibration activities are – changing of the preferences to a particular friend, changing the preferences to a particular keyword, adding or removing a keyword. All changing values are stored at the database and are later used by the system to recalculate the user's profile.

Spreading of the application at the social network is crucial for his evaluation. There are two different ways to advertise the application – explicit and implicit. The explicit way is via invitations send to the friends of the users which already used the application. The implicit way is via link to the application when user shares content through the application interface.

## 6 Recommendation approaches

The system uses both explicit and implicit information about user's preferences and processes it at three recommendation approaches:

- Collaborative recommendations – it use as source of information user's object marked as "liked" and the recommendation is made based on which users have preferences similar to ours. The selected algorithm for collaborative filtering is Slope One [13]. It takes as input parameter user-object table which contains 1 if the object is liked from the user or 0 if it is not selected as liked.
- Content-based recommendations – the content which is liked from the user is used to create user profile, which is later compared to the new contents. Their rank is calculated based on the similarity to the profile. The user profile is created from the keywords and tags contain at objects published, liked or shared by the user. Both user profile and content are represented as N-dimensional keyword vectors extracted from the text content. Vector Space model [12] is used as the weight of the keywords is calculated with TF/IDF metric and similarity between them is calculated with cosine similarity metric.
- Social recommendations – it's based on the social interaction between the user and its friends. The approach differs from collaborative recommendation as it use only the friend's "liked" objects. Comments, likes and shares are used to determine the relationship between the users. Relationship with the users who publish the content and users which comment and liked the content is used to determine how interesting it will be for the user.

## 7 Conclusions and Future Work

Presented research shows that social network's users spend most of their time looking for interesting information. Results clearly show the necessity of additional application that helps users at their everyday interaction at the social network and order their news feed based on their preferences. Authors already have developed an application which offers filtering options, automated and manual settings on the different sources of content. Next steps are beta testing and spreading at the social network.

### Acknowledgments

This work was supported by the European Social Fund through the Human Resource Development Operational Program under contract BG051PO001-3.3.06-0052 (2012/2014).

### References

1. Internet Usage Statistics, <http://www.internetworldstats.com/stats.htm> (visited June 2013)

2. Kiss, J., Facebook hits 1 billion users a month, The Guardian, 4<sup>th</sup> October, 2012
3. One billion – key metrics, <http://newsroom.fb.com/download-media/4227> (visited June 2013)
4. Bennett, J., Lanning, S.: The Netflix Prize. KDD Cup and Workshop, 2007
5. Dror, G., Koenigstein, N., Koren, Y., and Weimer, M.: The Yahoo! Music Dataset and
6. Fantulin, N., Bosnjak, M., Znidarsic, M., Grcar, M.: ECML-PKDD 2011 Discovery Challenge Overview", 2011
7. Chechev, M.: "Recommender systems challenges. KDD Cup 2011" Days of the Science, V. Turnovo, 2011.
8. KDD Cup 2012, <http://www.kddcup2012.org/> (visited June 2013)
9. Massa, P., Avesani, P.: Trust-aware collaborative filtering for recommender systems. In: Proceedings of the International Conference on Cooperative Information Systems, CoopIS, pp. 492–508, 2004
10. What's the Difference between Top News and Most Recent?, <http://www.facebook.com/blog/blog.php?post=414305122130> (visited June 2013)
11. Facebook Graph Api, <https://developers.facebook.com/docs/reference/apis/> (visited June 2013)
12. Salton, G., Wong, A., Yang, C.: A vector space model for automatic indexing, Communications of the ACM, 1975
13. Lemire, D. и Maclachlan, A. Slope One Predictors for Online Rating-Based Collaborative Filtering. SIAM Data Mining Conference. 2005.
14. Cantador, I., Brusilovsky, P., Kuik, T.: 2nd Ws. on Information Heterogeneity and Fusion in Recommender Systems (HetRec 2011), In Proc. Of the 5th ACM conf. on Recommender systems. RecSys 2011. Chicago, IL, USA, 2011
15. O'Donovan, J., Smyth, B.: Is trust robust?: An analysis of trust-based recommendation. In: IUI'06: Proceedings of the 11th international conference on Intelligent user interfaces, pp. 101–108. ACM, New York, NY, USA (2006)
16. Fuguo, Z., Shenghua, X.: Analysis of trust-based e-commerce recommender systems under recommendation attacks. In: ISDPE '07: Proceedings of the The First International Symposium on Data, Privacy, and E-Commerce, pp. 385–390. IEEE Computer Society, Washington, DC, USA, 2007
17. Dunbar, R.: How many friends does one person need? Dunbar's number and other evolutionary quirks, Faber and Faber, 2010.
18. Epinions.com: Product Reviews and Consumer Reports, <http://epinions.com> (visited June 2013)