

EXPLORATORY SEARCH ON SOCIAL NETWORKING SITES: HOW USER FEEDBACK AND MULTI-PERSPECTIVE MICROBLOG ANALYSIS CAN BE EMPLOYED ON TWITTER

Michal ŽILINČÍK, Pavol NAVRAT, Gabriela KOSKOVÁ

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia*

michal@zilincik.eu, {navrat,koskova}@fiit.stuba.sk

Abstract. We propose a new method for supporting the exploratory search on Twitter as a particular contemporary web based social networking service. Our method copes with several challenges: brevity of microblogs (tweets), limited number of available ratings and need to process the recommendations online. In order to cope with the first challenge, the representation of microblogs is enriched by information from the referenced links, topic summarization and affect analysis. Small number of available ratings is increased by interpreting implicit feedback trained by *feedback model* during browsing. Recommendations are produced by *preference model* that models user's preferences over tweets. The evaluation shows promising results despite a strong constraint by navigating in the space of brief pieces of information, making recommendations based only on small number of ratings, and by optimizing the models to process in real time.

1. Motivation

Social networking services provide platforms to express opinions, announce news, share interesting facts, communicate or share content such as articles or photographs [5]. The content is heterogeneous in nature and is very limited in length (usually up to 140 characters). These facts pose a significant issue in any deeper analysis related to information retrieval tasks. To provide an exploratory search [4] within the information space of social

networks supported by a particular social networking service, it is necessary to enrich the brief textual content to better represent the microblogs [2].

We focused on the social networking service Twitter, where microblogs are called tweets. Twitter provides metadata for tweets in the form of simple characteristics, which are often used for the analysis. But these metadata are not sufficient to describe user's interests or preferences. Therefore further analysis is necessary. We combine the content with the metadata and use this information to learn user's preferences.

We apply the principles of text summarization inspired by experiment named *TweetMotif* [3] in order to provide the *topic summarization*, i.e. to extract the most common phrases from set of tweets. Based on occurrences of the extracted phrases, tweets are categorized into topics. We further learn the *feedback model* for modeling and interpreting implicit feedback in order not to burden the users by requiring too much explicit feedback. Rated tweets are used to train the *preference model*, which models user's preferences over tweets. This model is used for recommending interesting tweets to the user.

2. Results

Our aim was to allow the user to focus on the content and explore related topics. We had to choose appropriate features that are able to describe content from multiple perspectives. We had to keep in mind the requirement for a relatively short processing time. Besides the need to make recommendations in real time, we face two other challenges. One is brevity of tweets. We enriched the content by titles of referenced links, used tweet characteristics provided by Twitter and employed existing methods for affect estimation and topic summarization. Another challenge is the small number of rated tweets in training set for training the *preference model* that makes recommendations for the user. That is why we focused also at obtaining implicit feedback and proposed the *feedback model*. If successful *feedback model* is built, it can increase the amount of known user's opinions of content and subsequently enable to the *preference model* to make better recommendations.

In the experimental evaluation, we showed that it is possible to learn how to interpret implicit feedback even if the observed items are short messages optionally containing URLs. Experimental evaluation also showed that the proposed method reached overall success rate of 68.02%. We believe the evaluation shows that this kind of approach is appropriate and suitable for further research.

The *preference model* was evaluated in several alternatives and with only one exception it proved useful. We showed that model can provide recommendation to users with success rate of over 61% based on 19 values describing only 40 tweets that user read from which not every tweet is explicitly rated. Even though the success rate may not be high enough to guarantee an immediate success in a commercial application, we conjecture that our proposed method is capable of making a genuine improvement to an exploratory search on Twitter.

Acknowledgement: This work was partially supported by the Slovak Research and Development Agency under the contract No. APVV- 0208-10.

References

to other papers publishing the results that are summarized here

- [1] Zilincik, M., Navrat, P., Koskova, G.: Exploratory Search on Twitter Utilizing User Feedback and Multi-perspective Microblog Analysis. *PLOS ONE*, (2013), 8(11): e78857. doi:10.1371/journal.pone.0078857.

Other references

- [2] Kuzar, T.: Clustering on Social Web. *Information Sciences and Technologies Bulletin of the ACM Slovakia*, (2013), 5(1): 34-42.
- [3] O'Connor, B., Krieger, M., Ahn, D.: TweetMotif: Exploratory Search and Topic Summarization for Twitter. In *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*. (2010), pp. 384-385.
- [4] Tvarozek, M.: Exploratory Search in the Adaptive Social Semantic Web. *Information Sciences and Technologies Bulletin of the ACM Slovakia*. (2011), 3(1): 42-51.
- [5] Zhang, Q., Zeng, D., Wang FY., Wang, T.: Understanding Crowd-Powered Search Groups: A Social Network Perspective, *PLoS ONE*. (2012), 7(6): e39749. doi:10.1371/journal.pone.0039749.