Leveraging Microblogs for Resource Ranking

Tomáš MAJER*

Slovak University of Technology
Faculty of Informatics and Information Technologies
Ilkovičova 3, 842 16 Bratislava, Slovakia
xmajert@fiit.stuba.sk

Nowadays, the thinking of modern people and basically of the whole society is affected by the Web and many people benefit from the Internet. Its most important service – the Web – no longer consists of only static pages that people browse and search information in. User generated content has become more important than ever, which caused that a new medium for publishing brief posts has emerged – microblogging.

Microblogs are a lightweight form of traditional blogs, where users publish only brief reports called tweets. This phenomenon has become popular mainly thanks to Twitter, which enabled users to communicate with each other in an entirely new way through microblogs. They can share their actual feelings and experiences as well as opinions. The opinions are often linked to various entities, which typically represent personalities, products or events. By aggregating different views of many users on a topic, microblogs are considered a very valuable source of information.

Twitter represents a potential source of valuable user data. It is possible to monitor current users' opinions all around the world. Different world events can be monitored through microblogs [2] such as elections, putting a new product on market, or revolutions in what is very current and extensive user content. This content can be processed and utilized in order to improve access to information, e.g. by improving search. Currently search takes into consideration several criteria, where the most important are the content and structure, format, links with other sites or domain name credibility. By utilizing microblogs and by mining user opinions, it would be possible to obtain data coming directly from the users. Similarly, there are other areas where it is possible to use ranking according to user reviews. For example, product reviews and ratings can be modified by owners or manufactures in e-shops. Thinking of Twitter as of a separate service, it contains un-moderated user data as the user posts are not as biased and more objective.

Twitter is a service combining elements of both social networking and microblogging with certain specifics [1]. While user profiles in social networks are connected bi-directionally, connections on microblogs have only one-way orientation. Users may follow other users in order to track their posts, however, it does not

* Supervisor: Marián Šimko, Institute of Informatics and Software Engineering

Spring 2011 PeWe Workshop, April 1, 2011, pp. 31–32.

necessarily mean that also reverse links exist, i.e., a followed user may choose not to follow ones that follow him. In contrast with traditional blogs, post comments cannot be assigned directly to a post. However, a new post linking to another could be created (referred to as *retweets*). User profiles on Twitter are usually simple and public.

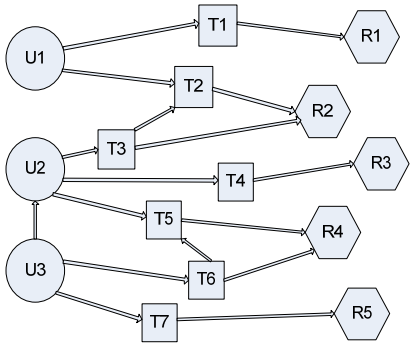

Figure 1: Basic Twitter graph representation.

The Twitter graph contains three types of entities (see Figure 1). The first type denotes users and they are represented by circles on the left side (U1-U3). Squares in the middle represent tweets (T1-T7). Hexagon represents resources (R1-R5) and they are located on the right side of graph. Otherwise there are also more tweets on Twitter which are not connected to pages.

Relations between users mean that a user follows another user. Tweets can be linked to users, another tweets and pages. Relations between tweets mean the tweets are re-tweets of parent tweets. Tweets with included text that contains links to external sites on the Web, create relation between tweets and pages.

We believe the structure of microblog posts represented by a graph (Figure 1) can be leveraged to determine ranking of resources referenced in microblog tweets. We propose a method for computing a resource rank, which we call TweetRank. TweetRank is derived from tweets topology by our method. The principle of the method lies in user ranking estimation and propagating such rank via graph relationships into web resources. By leveraging microblog specifics, we have presented how micro-blogging data can be used to improve web resource ranking traditionally based on resource content or resource links.

Acknowledgement. This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. KEGA 345-032STU-4/2010

References

[1] Huberman, B.A., a D.M. Romero: *Social networks that matter: Twitter under the microscope*. First Monday, 2009: 8.

[2] Dong, A.: *Time is of the essence: improving recency ranking using Twitter data.* In *Proc. of the 19th Int. Conf. on World Wide Web.* ACM Press, 331–340, 2010.