Chapter 6
Exploiting Social Media Features for Automated Tag Recommendation

Rabeeh Ayaz Abbasi
Quaid-i-Azam University, Pakistan

ABSTRACT

In today's social media platforms, when users upload or share their media (photos, videos, bookmarks, etc.), they often annotate it with keywords (called tags). Annotating the media helps in retrieving and browsing resources, and also allows the users to search and browse annotated media. In many social media platforms like Flickr or YouTube, users have to manually annotate their resources, which is inconvenient and time consuming. Tag recommendation is the process of suggesting relevant tags for a given resource, and a tag recommender is a system that recommends the tags. A tag recommender system is important for social media platforms to help users in annotating their resources. Many of the existing tag recommendation methods exploit only the tagging information (Jaschke et al., 2007, Marinho & Schmidt-Thieme, 2008, Sigurbjornsson & van Zwol, 2008). However, many social media platforms support other media features like geographical coordinates. These features can be exploited for improving tag recommendation. In this chapter, a comparison of three types of social media features for tag recommendation is presented and evaluated. The features presented in this chapter include geographical-coordinates, low-level image descriptors, and tags.

DOI: 10.4018/978-1-4666-0894-8.ch006

1. RELATED WORK

A tag recommendation system is used to assist users in tagging resources. These systems have been discussed in various research works over the last few years. Researchers have come up with frameworks which allow the comparison of different tag recommendation methods. (Jaschke et al., 2009) presents a tag recommendation framework for their system Bibsonomy. The framework allows the evaluation of different tag recommendation algorithms. The framework is though limited to the tag recommender systems which only use the tagging information.

(Jaschke et al., 2007) compared two algorithms, FolkRank and Collaborative Filtering (Goldberg et al., 1992) for tag recommendation. FolkRank is based on PageRank (Brin and Page, 1998). It uses random walk techniques on the graph of users, tags, and resources and assumes that popular users, tags, and resources can reinforce each other. In collaborative filtering, similarity between users and tags and between users and resources is used to recommend tags. Their experiments based on the datasets from delicious (http://www.delicious.com/), last.fm (http://www.last.fm/), and Bibsonomy (http://www.bibsonomy.org/) show that the FolkRank algorithm outperforms other methods. The tag recommendation methods as proposed by (Jaschke et al., 2007) depend mainly on the tagging information and do not consider the features (like geographical coordinates or low-level image features) available in rich media (like photos or videos). The tag recommendation methods proposed by (Jaschke et al., 2007) suggest tags for already partially tagged resources.

(Sigurbjornsson and van Zwol, 2008) presents a tag recommendation system which evaluates different similarity measures, tag aggregation methods and ranking strategies. Given a photo and some initial tags, candidate tags are derived for each of the given tag. The candidate tags are retrieved based on the tag co-occurrence information. All of the candidate tags are then merged and ranked. A final list of tags is then presented to the user. As in the work presented by (Jaschke et al., 2007), the methods proposed by (Sigurbjornsson and van Zwol, 2008) lack a tag recommendation strategy for newly uploaded resources. Although the experiments were performed on Flickr (photos) dataset, the methods do not consider the available rich media features.

Nowadays, the state-of-the-art imaging devices provide photos together with the geographical coordinates (*geo-tags*) stating precisely where they have been acquired. Therefore, more and more researchers make use of this additional information. (Cristani et al., 2008) exploits geographical coordinates for improving visualization of images on a map. (Kennedy et al., 2007) and (Kennedy and Naaman, 2008) use low-level image features and geographical coordinates to identify the landmarks of a city. (Moellic et al., 2008) presents a system which combines tags and low-level image features for clustering images. They suggest that clustering images can enhance the browsing and visualization of the images.

In addition to the features available in rich media, some researchers have used external data sources for recommending tags. (Heymann et al., 2008) predicts tags by using information available in the resource content, anchor text, and already available tags. Given a set of objects, and a set of tags applied to those objects by users, their approach predicts whether a given tag could/should be applied to a particular object. (Heymann et al., 2008) formulates the problem of tag recommendation into a supervised learning problem. For each tag to be recommended, they train a binary Support Vector Machine (SVM) classifier which predicts the association of a resource with the tag. Their approach is limited to a set of tags that can be recommended and may not be applied in a generic large scale scenario. Resource features like titles of webpages have also been exploited by other researchers (Lipczak, 2008). (Lipczak, 2008) suggests a tag recommendation system which extracts the tags from the resource title. The