Abstract.

Social networks enable users to freely communicate with each other and share recent news, ongoing activities or views about different topics. This has made social networks as a viable source of information about users' interests with regards to the current active topics/events. The development of techniques that can automatically detect such topics and model users' interests towards them has become an emerging research area in the recent years. It has the potential to improve the quality of applications that work on a user modeling basis, such as filtering twitter streams, news recommendation, user community detection and retweet prediction, among others.

When processing social posts for the identification and extraction of user interests, traditional keyword-based methods, which are often proposed for processing formal and large documents, are less effective on social posts, due to the short length, noisiness and informality of the content. For addressing these issues, recent works have proposed to utilize external knowledge bases (such as Wikipedia/DBpedia) to link the terms in the posts to the relevant concepts described in those knowledge bases. Since these knowledge bases represent the concepts and their relationships, these links provide a way of inferring underlying semantics of the posts. These works consider each of the semantic concepts, separately as a topic of interest. Therefore, such approaches undermine the fact that a user might not be much interested in a single concept as a broad topic, but are rather interested in a specific topic that requires to be represented through a combination of multiple semantic concepts. In addition, existing work often confine users' interests to a set of predefined semantic concepts, and therefore, they cannot discover emerging topics of interest that are not explicitly included in the knowledge base.

In this thesis, we define a topic of interest as a conjunction of several semantic concepts that are temporally correlated on social network and model user interests over these topics. Therefore, even if a single corresponding semantic concept is not available in the external knowledge base, we construct its semantics by using existing concepts. Based on this definition, our work will concentrate on formulating a multifaceted framework that collectively considers social interactions (network structure of connections), temporal behaviour (chronological order of activities), and semantics of social content (the underlying meaning of user content) in order to perform user interest modeling. This framework consists of three main components to achieve the following goals: (1) Extracting user explicit interests over active topics in a given time interval by analysisng users' textual contents. (2) Infering implicit interests of users over active topics in given time interval by considering the relationship between users and topics, in addition to their explicit interests. (3) Prediction future interests of users over a set of topics which may not have been observed in the past by temporal modeling of user's interest with regards to extracted topics in each time interval.