Neural Networks 96 (2017) 11–21

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

User emotion for modeling retweeting behaviors

Jinpeng Chen [a,*], Yu Liu [b], Ming Zou [b]

[a] School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[b] State Key Laboratory of Software Development Environment, Beihang University, Beijing 100191, China

ARTICLE INFO

Article history:
Received 31 October 2016
Received in revised form 23 April 2017
Accepted 21 August 2017
Available online 8 September 2017

Keywords:
Retweeting behavior
Emotion
Social network
Twitter
Factor analysis

ABSTRACT

Twitter and other microblogs have rapidly become a significant mean of information propagation in today's web. Understanding the main factors that make certain pieces of information spread quickly in these platforms has emerged as a popular topic. Therefore, as a simple yet powerful way of disseminating useful information, retweeting has attracted much interest. Existing methods for retweets have been conducted for analyzing the social network structure, or understanding the retweeting mechanism. However, little attention is paid to whether users' emotion will affect users' retweeting behavior. In this paper, we study the user emotion problem in a large social network. Particularly, we consider users' retweet behaviors and focus on investigating whether users with a certain emotional status will retweet the tweet corresponding with users' current mood from their friends. In order to achieve this goal, we propose a retweeting prediction framework. First, we construct a model of emotion detection via considering two kinds of emotional signals; second, we extract possible retweeted friends and tweets; third, based on the first two steps, we obtain Top-N retweets using Learn-to-Rank method. Experiments are performed on two real-world datasets, the Twitter network and Obama–McCain Debate dataset, with comprehensive measurements. Experimental results demonstrate that our retweeting prediction framework has substantial advantages over commonly used retweeting prediction approaches in predicting retweeting behaviors. Consider Precision in Twitter network as an example. For the Top-N stage, our method can, on average, increase by 15.2% and 11.2% in relation to Tweet(+SV) and User(+ED), respectively. We find that emotion is a vital feature which affects retweetability.

1. Introduction

Social network has emerged as a vital tool for information dissemination, search and marketing. Meanwhile, with the increasing of people's requirements, more and more social networks are appearing, like Twitter,[1] Facebook,[2] and Sina Weibo.[3] Taking Twitter as an example, through Twitter, users can post tweets whose length is up to 140 characters, which tells your family or friends anything in your daily life, such as what you are doing, what you are thinking, or what is happening around you. So far, the number of Twitter users has climbed to 288 million (Urabe, Rzepka, & Araki, 2013) and the number of tweets published on Twitter every day is over 65 million.[4]

As a result of the rapidly increasing number of tweets, more and more researchers pay attention to mining people's emotions expressed in tweets (He, Zheng, & Bao, 2014; Jiang, Yu, & Zhou, 2011). As Bollen, Pepe, and Mao (2011) said, users argued that the events in the social, political and cultural fields did have a significant influence on users' mood. Usually, different emotions may be expressed by users towards different topics, where users may be happy with some aspects of an entity but unhappy with other aspects. Thus, by analyzing the semantic orientation of tweets posted by users, we can obtain the current sentiment of users, like positive, negative or neutral. Nowadays, some web sites constructed on the Internet have already been emerging in order to provide a Twitter sentiment search service, like Tweetfeel,[5] Twendz,[6] and Twitter Sentiment.[7] However, most of the current research or services focus on sentiment analysis itself, such as sentiment classification (Jiang et al., 2011; Lin & He, 2009), inferring individual emotional states (Kim, Yoo, & Lim, 2013; Tang, Zhang, & Sun, 2012), and disclosing the emotions in tweets (Zhao, Dong, & Wu, 2012). Few research works involve how emotions affect

* Corresponding author.
E-mail address: jpchen@bupt.edu.cn (J. Chen).

1 https://twitter.com/
2 https://www.facebook.com/
3 The most popular Chinese microblogging service,
4 http://en.wikipedia.org/wiki/Twitter
5 http://www.tweetfeel.com/
6 http://twendz.waggeneredstrom.com/
7 http://twittersentiment.appspot.com/

http://dx.doi.org/10.1016/j.neunet.2017.08.006
0893-6080/

people's behaviors, like retweeting behaviors, which motivates our work.

Here, retweeting behavior is a strong indication of the direction of information flow in the Twitter social graph in that people explicitly identify the source, and is also a social practice that occurs in Twitter so as to quickly share a piece of information (Recuero, Araujo, & Zago, 2011). In a certain extent, retweet action indicates that the original tweet contains valuable information (Suh, Hong, & Pirolli, 2010). Also, retweets have been a measure of the tweet's popularity and influence (Kwak, Lee, & Park, 2010). Retweeting behavior is an important feature for personalized Tweet recommendation (Chen, Chen, & Zheng, 2012) and learning to rank of Tweets (Duan, Jiang, Qin, & Zhou, 2010). Understanding retweeting mechanism and predicting retweeting behavior is an important and essential step for various social network applications, such as user behavior analysis, business intelligence, and popular event prediction (Liang, Jiang, Yin, Wang, Tan, & Bai, 2016; Zhang, Gong, Guo, & Huang, 2015).

Currently, there exist many studies about retweets. One important branch of these studies is retweeting the prediction model. It contains two main points: one is which tweet will be retweeted (Chen et al., 2012), the other one is finding retweeters in Twitter (Luo, Osborne, & Tang, 2013). However, few works pay attention to which tweet a certain user will repost, which also motivates our research work.

From the previous research, major factors affecting the retweeting behaviors are categorized into the following: (1) content and contextual features of the tweets, such as URLs and hashtags; (2) users' features, such as users' interests, the age of the account, and social influence (Zhang, Liu, & Tang, 2013). Unluckily, emotions can never be considered as a vital feature of users to affect retweetability. Intuitively, users with good mood are likely to follow the positive tweets and ones with bad mood would like to follow the negative tweets. In this paper, we investigate how different emotions in the context of online social media affect users' retweeting behaviors. To the best of our knowledge, this is the first attempt towards studying that human's emotions somehow influence retweeting behaviors in a large scale and real world setting.

The main contributions of this research can be summarized as follows:

- We study the problem of the retweeting behavior prediction in Twitter social network.
- We present a novel retweeting prediction framework consisting of three stages: emotion detection, capturing the possible retweets, and finding Top-N retweets.
- We propose a emotion detection model to generate users' mood by leveraging two kinds of emotional signals, emoticons and word-level emotions.
- Experiments on real dataset indicate that our proposed method can predict retweeting behavior with high accuracy compared with the state-of-the-art prediction methods.

The remainder of this paper is organized as follows. We first introduce the related work in Section 2. We next present several necessary definitions and denote the task of retweeting behavior prediction in Section 3. We further show a novel three-stage prediction framework to predict the users' retweeting behavior in Section 4. We report our experiments and results in Section 5, and conclude the study in Section 6.

2. Related work

In this section, we review two categories of related work: studies on emotion mining and modeling retweeting behaviors in social networks.

Emotion Mining. Emotion mining has recently emerged as a popular way to help better understand a user's online behavior and benefit a set of applications relevant to online campaign and recommender systems (Gao, Mahmud, & Chen, 2014; Mohammad, Zhu, Kiritchenko, & Martin, 2015; Rangel & Rosso, 2016). Early works mainly focused on how to understand users' emotions in social media, such as qualitative and quantitative analysis of how an individual's emotional state can be inferred from her/his historic emotion log (Tang et al., 2012).

In Aoki and Uchida (2011), authors proposed a method by which emotional vectors of emoticons can be automatically generated via using emotional words that co-occur with emoticons derived from many weblog articles. In He et al. (2014), they first investigated emotion entrainment in the context of online social media and then showed a framework which can model entrainment phenomenon. Jiang et al. (2011) addressed target-dependent sentiment classification of tweets. Specifically, given a query, the sentiments of the tweets can be classified as positive, negative or neutral according to whether they contained positive, negative or neutral sentiments about that query. Similarly, Shen, Kuo, and Yeh (2013) tried to solve a clinical problem, depression detection. A two-stage supervised learning framework was proposed to identify potential depression candidates based on their write-ups. In Urabe et al. (2013), they described an emoticon recommendation system by which users can express their feelings with their input. In Calais Guerra, Veloso, and Meira (2011), authors handled the task of real time sentiment analysis as a relational learning task via using the network structure formed from the mutual endorsements among social network users. In Hu, Tang, and Tang (2013), they investigated whether social relations can help sentiment analysis and presented a sociological approach to addressing noisy and short texts for sentiment classification. Zhao et al. (2012) built an online sentiment analysis system for Chinese tweets in Weibo, which generated sentiment labels for tweets by employing the emoticons, and constructed an incremental learning Naive Bayes classifier for the categorization of four types of sentiments. In Hu, Tang, and Gao (2013), authors detected a new problem of interpreting emotional signals for unsupervised sentiment analysis. Li, Wang, and Hovy (2014) tried to answer the long-lasting question about mood-weather correlation by showing a framework that harnessed Twitter data. In Guillory, Spiegel, and Drislane (2011), they emphasized the importance of the intensity of specific emotions, in other words, emotions impacted text-based communication, and pointed out that the process of emotion contagion in this context was more complex than previously thought.

Another line similar to our research problem is how to use emotion to guide users' behavior (Calais Guerra et al., 2011; Gao et al., 2014; Kramer, 2012; Mohammad et al., 2015; Yang, Jia, & Zhang, 2014). In Yang et al. (2014), they presented a novel emotion learning method, that is, they modeled the comments and visual features of images and bridged these two pieces of information via learning a latent space. In Gao et al. (2014), they proposed a computational model to infer a user's attitude according to sentiment, opinion and likelihood of taking an action towards controversial topics in social network. In order to recommend mood-specific movies, Calais Guerra et al. (2011) put forward a novel mood-specific movie similarity concept. To be specific, they used a joint matrix factorization model to factorize both the user-movie rating matrix and the mood-specific movie similarity matrix. In this work, we also try to extract users' emotions from social media, like Twitter, but our aim is to detect whether users' emotions will affect their retweeting behaviors.

Modeling Retweeting Behaviors. In recent years, retweeting behavior has attracted much interest because of the importance of retweet practice (Bi & Cho, 2016; Starbird & Palen, 2012; Zeng, Luo, & Chen, 2013). However, most of the research works focused on how to utilize retweet behavior to analyze the phenomenon in social network, such as cascade prediction (Cheng, Adamic, & Dow,