A SECURE TWITTER TRENDING MANIPULATION AND TOPIC CLASSIFICATION

[1]T. SaiPrasad Reddy, [2]Ch. Bhaskar Rao, [3]A. Vamsi Krishna, [4]T. Vamsi Vardhan Reddy
[1]Associate Professor, [2] Assistant Professor, [3] Assistant Professor, [4] Assistant Professor
Dept. of CSE,
Narayana Engineering College, Nellore, A.P, India

Abstract- Twitter trends, a timely updated set of top terms in Twitter, have the ability to affect the public agenda of the community and have attracted much attention. Unfortunately, in the wrong hands, Twitter trends can also be abused to mislead people. Twitter drifts, an opportune refreshed arrangement of best terms in Twitter, can influence people in general motivation of the group and have pulled in much consideration. In this project, we will verify whether Twitter patterns are secure from the control of pernicious clients. We gather more than 69 million tweets from 5 million records. Utilizing the gathered tweets, we initially lead an information investigation to know on what basis data is getting manipulated and find proof of Twitter slant control. In this project, we do research at the subject level data and induce the key factors that can decide if a theme begins slanting because of its notoriety, scope, transmission, potential scope, or notoriety.

Keywords- *Twitter, Slant Control, Notoriety*

I. INTRODUCTION

Twitter is an extremely popular micro blogging site, where users search for timely and social information such as breaking news, posts about celebrities, and trending topics. Users post short text messages called tweets, which are limited by 140 characters in length and can be viewed by user’s followers. Anyone who chooses to have other’s tweets posted on one’s timeline is called a follower. Twit-ter has been used as a medium for real-time information dissemination and it has been used in various brand cam-paigns, elections, and as a news media. Since its launch in 2006, the popularity of its use has been dramatically increasing. As of June 2011, about 200 million tweets are being generated every day [1]. When a new topic becomes popular on Twitter, it is listed as a trending topic, which may take the form of short phrases (e.g., Michael Jackson) or hash tags (e.g., #election). What the Trend provides a regularly updated list of trending topics from Twitter. It is very interesting to know what topics are trending and what people in other parts of the world are interested in. However, a very high percentage of trending topics are hash tags, a name of an individual, or words in other languages and it is often difficult to understand what the trending topics are about. It is therefore important to classify these topics into general categories for easier understanding of topics and better information retrieval.

The trending topic names may or may not be indicative of the kind of information people are tweeting about unless one reads the trend text associated with it. For example, #hap-pyvalentinesday indicates that people are tweeting about Valentine’s Day. A trend named Boone Logan is indicative that tweets are about person named Boone Logan. Anyone who does not follow American Major League Baseball (MLB), however, will not know that the information is regarding Boone Logan, who is a pitcher for the New York Yankees unless a few tweets are read from this trending topic as shown in Figure 1. We find that trend names are not indicative of the information being transmitted or discussed either due to obfuscated names or due to regional or domain contexts. To address this problem, we defined 18 general classes: arts & design, books, business, charity & deals, fashion, food & drink, health, holidays & dates, humor, music, politics, religion, science, sports, technology, tv & movies, other news, and other. Our goal is to aid users searching for information on Twitter to look at only smaller subset of trending topics by classifying topics into general classes (e.g., sports, politics, books) for easier retrieval of information. To classify trending topics into these prede-fined classes, we propose two approaches: the well-known Bag-of-Words text classification, and using social network information.

In this paper, we use supervised learning techniques to classify the twitter trending topics. First, we employ a well-known text classification technique called Naive Bayes (NB) [2]. A document in NB would model as the presence and absence of particular words. A variation of NB is Naive Bayes Multinomial (NBM), which considers the frequency of words and can be denoted as:

$$P(c|d) \propto P(c) \qquad P(tk|c), \qquad (1)$$

$1 \le k \le nd$

where P (c|d) is the probability of a document d being in class c, P (c) is the prior probability of a document occurring in class c, and P (tk|c) is the conditional probability of term tk occurring in a document of class c. A document d in our case is trend definition or tweets related to each trending topic.

Apart from text-based classification, we also incorporate twitter social network information for topic classification. For the latter we make use of topic-specific influential users, which are identified using twitter friend-follower network. The influence rank is calculated per topic using a variant of the Weighted Page Rank algorithm [4]. In general, a tweeter is said to have high influence if the sum of the influence of those following him/her is high. The key idea of the proposed network-based approach is to predict the category of a topic knowing the categories of its similar top-ics. Similar topics are identified using user-similarity metric, which is the cardinality of the intersection of influential users between two topics ti and tj divided by the cardinality of top s influencers of topic ti [3]. We experimented using different classifiers, for example, C5.0 (an improved version of C4.5) [5], k-Nearest Neighbor (kNN) ,Support Vector Machine (SVM) [7], Logistic Regression and ZeroR (the baseline classifier), and found that C5.0 classifier resulted in the best accuracy on our data set. Experimental results show that both our approaches effectively classify trending topics with high accuracy, given that it is a 18-class classification problem.

II. RELATED WORKS

A number of recent papers have addressed the classification of tweets. classified tweets to a predefined set of generic classes such as news, events, opinions, deals, and private messages based on author information and domain-specific features extracted from tweets such as presence of shortening of words and slangs, time-event phrases, opin-ionated words, emphasis on words, currency and percentage signs, “@username” at the beginning of the tweet, and “@username” within the tweet. introduced a wikipedia-based classification technique. The authors classified tweets by mapping message into their most similar Wikipedia pages and calculating semantic distances between messages based on the distances between their closest wikipedia pages. included metadata from external hyperlinks for topic classification on a social media dataset. Whereas all these previous works use the characteristics of tweet texts or meta-information from other information sources, our network-based classifier uses topic-specific social network information to find similar topics, and uses categories of similar topics to categorize the target topic.

Sankaranarayanan have built a news processing system that identifies the tweets corresponding to late break-ing news. Issues addressed in their work include removing the noise, determining tweet cluster of interest using online methods, and identifying relevant locations associated with the tweets. Yerva et al. classify tweet messages to identify whether they are related to a company or not using company profiles that are generated semi-automatically from external web sources. Whereas all these previous works classify tweets or short text messages into 2 classes, our work classify tweets into 18 general classes such as sports, technology, politics, etc.

Becker explored approaches for distinguishing tweet messages between messages about real-world events and non-event messages. The authors used an online clustering technique to group topically similar tweets together, and computed features that can be used to train a classifier to distinguish between event and non-event clusters.

There has been a lot of research in sentiment classification of short text messages. Go et al. introduced a approach for automatically classifying sentiment of tweets with emoticons using distant supervised learning. Pang etal. classified movie reviews determining whether a review is positive or negative. But none of these classify twitter trending topics.or tweet text is gibberish or if it is in a language other than English, then we classify the topic as other category. The data was labeled by reading topic’s trend definition and few tweets.

We used two annotators to label all topics. In case of disagreement, a third annotator intervened. For the labeling task, a random sample of 1000 topics was selected. From the 1000, we narrowed the data set down to 768 topics for mainly two reasons. First, the topic had no trend definition. Second, the third annotator could not finalize the label. For each of the 768 topics in our dataset, its five most similar topics were also labeled, which are required for the network-based modeling as described in Section III-C2. We ended up manually labeling 3005 topics because some of the similar topics were common to more than one topic. Figure 8 shows the web-interface we deployed for the labeling task.

III. DATA AND METHODS

As shown in Figure 2, the proposed classification system consists of four stages: Data Collection, Labeling, Data Modeling, and Machine Learning. In our experiments, we use two data modeling methods: (1) Text-based data modeling; and (2) Network-based data modeling.

A. Data Collection

The website What the Trend provides a regularly updated list of ten most popular topics called "trending topics" from Twitter. A trending topic may be a breaking news story or it may be about a recently aired TV show. The website also allows thousands of users across the world to define, in a few short sentences, why this term is interesting or important to people, which we refer to as "trend definition" in the paper. The Twitter API3 allows high-throughput near real-time access to various subsets of public Twitter data. We downloaded trending topics and definitions every 30 minutes from What the Trend and all tweets that contain trending topics from Twitter while the topic is trending. All the tweets containing a trending topic constitutes a document. For example, while the topic "superbowl" is trending, we keep downloading all tweets that contain the word "superbowl" from Twitter, and save the tweets in a document called "superbowl". In case a tweet contains more than two trending topics, the tweet is saved in all relevant documents. For example, if a tweet contains two trending topics "superbowl" and "NFL", the same tweet is saved into two documents called "superbowl" and "NFL". From 23000+ trending topics that we have downloaded since February 2010, we randomly selected 768 topics as our dataset.

B. Labeling

We identified 18 classes for topic classification. The classes are art & design, books, charity & deals, fashion, food & drink, health, humor, music, politics, religion, hol-idays & dates, science, sports, technology, business, tv & movies, other news, and other. Since twitter is a primary source of news or information, the news related to political events are classified as politics. If the topic is about news that is not in any of the categories, it is classified as other news. If the trend definition

C. Machine Learning

The 2 datasets constructed as a result of the two ap-proaches in the Data Modeling stage are used as inputs to machine learning stage. We built predictive models using various classification techniques and selected the ones that resulted in the best classification accuracy. The experimental results are discussed in next section.

IV. EXPERIMENTS AND RESULTS

For our experiments, we used popular tools such as WEKA and SPSS modeler WEKA is a widely used machine learning tool that supports various modeling algorithms for data preprocessing, clustering, classification, regression and feature selection. SPSS modeler is another popular data mining software with unique graphical user interface and high prediction accuracy. It is widely used in business marketing, resource planning, medical research, law enforcement and national security. In all experiments, 10-fold cross-validation was used to evaluate the classification accuracy. The ZeroR classifier was used to get a baseline accuracy, which simply predicts the majority class.

V. CONCLUSION

In this paper, we used two different classification schemes for Twitter trending topic classification. Apart from using text-based classification, our key contribution is the use of social network structure rather than using just textual information, which can be often noisy given in the context of social media such as Twitter due the heavy use of Twitter lingo and the limit on the number of characters that users are allowed to generate for their messages. Our results show that network-based classifier performed significantly better than text-based classifier on our dataset. Considering tweets are not as grammatically structured as regular document texts, text-based classification using Naive Bayes Multinomial provides fair results and can be leveraged in cases where we may not be able to perform network-based analysis.

In our future work, we would like to integrate text-based classification using Naive Bayes Multinomial (NBM) and network-based classification. The idea would be to integrate these two classifiers such that if we have all five similar top-ics classified then use network-based classification otherwise use text-based classification. During our experiments we found some topics could fall under more than one category. For example, news about a famous actor's biography would fall under tv & movies and books. Hence, we would also like to explore the use of multiple labels in categorization.

VI.REFERENCES

[1] MarketingGum,http://www.marketinggum.com/twitter-statistics-2011-updated-stats/.

[2] C. D. Manning, P. Raghavan, and H. Schtze, Introduction to Information Retrieval. New York, NY, USA: Cambridge University Press, 2008.

[3] R. Narayanan, "Mining Text for Relationship Extraction and Sentiment Analysis," Ph.D. dissertation, 2010. W. Xing and A. Ghorbani, "Weighted pagerank algorithm," 2004.

[4] J. Quinlan, "Improved use of continuous attributes in c4.5," Arxiv preprint cs/9603103, 1996.