INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 4.295

(Volume2, Issue6)

Available online at: www.ijariit.com

Analyzing of the Media and Public Tendencies in Twitter

Mr. Ashish Laxman Gaikwad
Department of Computer Engineering
SavitribaiPhule Pune University, Pune.
ashishgaikwad534@gmail.com

Mr. Sumit Suresh Shendkar
Department of Computer Engineering
SavitribaiPhule Pune University, Pune.
sumitamit755@gmail.com

Mr. Mahesh Tukaram Atpadkar
Department of Computer Engineering
SavitribaiPhule Pune University, Pune.
maheshumesh9091@gmail.com

Abstract: Social networking services such as Twitter creates content reflecting a series of talks that appear in the real world events. Twitter is a social networking site that provides service to a large number of users to communicate with each other at same time. That is asymmetrical relationship between friends and followers that provide dramatic structure of interest between Twitter users. A series of Twitter messages called tweets, which are limited to 140 characters, and thus are usually much focused. The basic process is the capture of Twitter tweets that extract most discussed topic in between users. Tweet this Dataset can be processed using standard natural language processing to search for trending stories. Common stories and erosion areso brick is extracting and summarizing information gathered from social networking services. There is the fact of Ways to find common stories that improve the quality of the result. This article proposes an application to detect themes of tendencies of the data that the BNgram Twitter disclosure rules use.

Keywords- Twitter, Topic Detection, BNgram, Social Networks.

I. INTRODUCTION

It has been used social networks on a large scale in Years. As social networking services quickly spread all ages the society. The great discussion of the facts, the interaction of the Communication occurs on social networking sites and this It reflects real-world events and trend issues; Now one day user Is more active to post messages about real-world events in Social networking sites, so the social networking site now becomes For real-time information about Events, where we can effectively detect stories in real time.

It has been become a social networking service in main area, Number of interconnected users rises rapidly fast and effective processes in real time. Twitter It is one of the most social networking sites that become Popular in each to share their thoughts, ideas and express Opinion on real-time events with a single user account. Twitter it has number of users, who post their messages called a tweet. It contains a maximum of 140 characters It is estimated that There 6-7, 000 000 users who use Twitter a total of 134 Millions of times a month. Twitter has open access to publish Information related to breaking news and events About 500 million users and more than 400 million Messages known as tweets; these tweets contain Conversation, thoughts, impacts of events in real time.

Uncovered Trend Stories of the dataset, which contains Series of tweets with respect to the time stamp. Discovering Trend Stories is most useful for last minute television, Social problem analyzer.

The reminder of this document is as follows. We then provide about using Twitter. Then we introduce types of Trends, data sets, related work and Methods to discover twitter trend stories.

II. TWITTER

In this section of Twitter, the details of trends in Twitter and we are using input methods to discover Land of trend stories.

- **Twitter**
  Twitter is the most popular social website, where large number of users share their thoughts, their form of short messages called tweets. The success Twitter is due to two reasons first, the lack of Tweets, which cannot exceed 140

characters Create and share the minimum period of time and second it is transmitting these messages to a large number of users. In a very short time. Twitter has set the syntax for the interaction between them, that the syntax adopted Top Twitter Clients Implemented this too. The standard in the syntax of the interaction includes:

- User mentions: when a user mentions another user in his tweet, a sign is placed before the corresponding user name.
- Reply: When a user wants to direct another user, or respond to a previous tweet, they put the mention @ user name at the beginning of the tweet.
- Retweet: A retweet is a re-share of a tweet posted by another user. Retweets, the new tweet copies the original into it, then the retweet attaches a RT and the @username of the. User who posted the original tweet at the beginning of the retweet.
- Hash tags: It is the same as marking the installations in another social networking service, hash tags included in a tweet to mention another user.

- **Trending Stories**:
  One of the main features of Twitter's homepage is a list of the best terms called trend topics at all times. These terms reflect that they are being discussed more. Twitter focuses on topics that are being discussed much more than usual. Trend themes have attracted great interest not only for the user, mainly for other information consumers like Journalists, researcher of applications in real time and social media. However, no further evidence is known about the algorithm that extracts topics of tendencies.

III. TYPES OF TRENDING TOPICS

Here are the trend topics in the following categories:

News: On many occasions breaking news arrives on Twitter before any news agency. We defined a trend topic can be categorized as news when it gives current information.

Ongoing Events: The current topic is in progress when the information is sent by the community of users tweeting about an ongoing event.

Memes: Also the trend theme is in memes that is published by an individual or a community with Ideas. It can be a funny message that attracts users to re-publish.

Commemoratives: Last type of trend theme that produced by the individual to congratulate the celebrity your birthday or anniversary or any memorable day As Independence Day, Republic Day.

IV. LITRETURE SURVEY

This section describes several techniques in discovering trends topics. In Sensing Trending Topic on twitter, three sets of Twitter data are used to extract the detection of trend topics and is extracted by BNgram Method. In the detection of emerging issues on Twitter based on the evaluation of social terms, we recognize the primary role of twitter and propose a new technique of detection of topics that allows real-time retrieval of the most emerging topic expressed by user communities. Based assets based on the well-known page algorithm. In another work twitters Monitor: Detection of trends on the Twitter flow, they represent Twitter Monitor, a system that provides a meaningful analytic that synthesize an accurate description of each topic using Twitter API, Another work using Twitter API is Twitter Stand: News in Tweets to build A news processing system. In the detection and monitoring of political abuse in social media, we describe a machine learning framework that combines topological characteristics, content-based and multitude of sources using Twitter API.
In Predicting the Political Preference of Twitter Users, they can predict from their interaction with political parties by constructing a prediction model based on a slope of the contextual and behavioral behavior of the models by restoring of a remote supervision approach. In another work beyond trend topics: identifying real-world events on Twitter, explores approaching the stream of Twitter messages to distinguish between messages about real-world events and non-message events sing-level events Cluster based on Temporal, social, -centric. In the next approach, Taking Test Detection from Assessment to Practice avoids the generation of garbage groups, which have returned to different approaches. In social media mining news topics, demonstrate by analyzing tweets corresponding to the events Taken from the word of politics and sport using the BNgram method. Also in the detection of breaking news and Follow on Twitter, propose a method to collect, classify and classify breaking news on Twitter, each group is classified based on factors of popularity and reliability. In recent work, the Twitter Trends real-time ranking uses the Twitter API to get the first trending topic and get the trending theme with text, timestamp, user, and language for each of the underlying tweets.

V. PROBLEM SATEMENT

A. System Model

Describes the system architecture of discover topics of twitter trends.

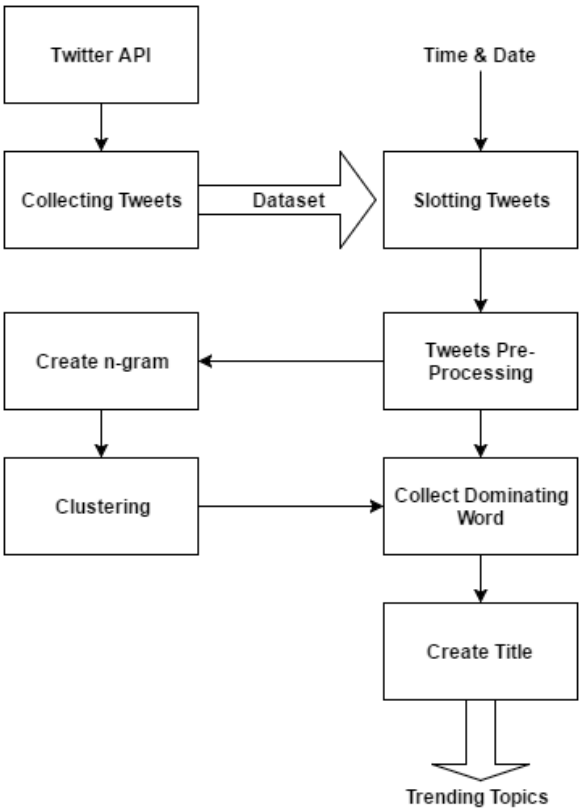

Fig. 1. System Architecture

Here we use the twitter dataset, which contains a series of tweets, between which we select series of tweets of two instances called Slots, and then in the selected tweets we apply the operation of text preprocessing through natural language processing to delete words Nonverbal and nonverbal. We follow the BNgram algorithm, to create n-grams and then organize these n-grams into different clusters using k-means clustering algorithm.
Finally we extract the keyword that having high frequency using this keyword we create the title for the trend story.

B. Algorithm

K-Means Algorithm

K-means clustering is a partitioning method. In these terms they are classified as one of K-groups. The result of this partitioning method is a cluster of K clusters. In K-Means clustering algorithm terms are classified into K clusters; The value of K is defined by the user. First, the centroid of each group is selected and then, according to the centroid, terms having a minimum distance from the given group, are assigned to that particular group. The Euclidean distance is used to calculate the distance of terms from the particular centroid.

C. Mathematical Model

 Input Sets:

$Ds=\{Dsi, i < 0 < n\}$ -Set of Twitter Datasets
Where, n=Number of Twitter Datasets
$St=\{Stj , j < 0 < n\}$ -Set of Text Streams
Where, n=Number of Text Streams

 Processing Sets:

$Tw=\{Twi, i < 0 < n\}$ -Set of Tweets
Where, n=Number of Tweets
$Gr=\{Grj , j < 0 < n\}$ -Set of n-grams
Where, n=Number of n-grams
$Cs=\{Csk, k < 0 < n\}$ -Set of Clusters
Where, n=Number of Clusters

 Output Set:

Rt={Rti, i < 0 < n} -Set of Ranked tweets
Where, n=Number of Ranked tweets
Ur={Urj , j < 0 < n} -Set of Users
Where, n=Number of Users

Let S be the systems we can mathematically represent S using Set of Theory as,
S={Ds, St, Tw, Gr, Cs, Rt, Ur}

D. BNgram

Frequency Term Document frequency, or tf-idf, has been used to index the document since it was first introduced. But here we want to find the term that appears more time period more than others. We select the terms with a frequency of documents of frequency of documents of high temporal frequency, comparing the most recent messages m with the messages m previous and how many terms are repeated. We assume the most recent messages as a slot. After standard text preprocessing, we index all the terms of these messages. For each term, we calculate the frequency of the document for a set of messages using dfti defined as the number of messages in a set of i containing the term t.

$$df\text{-}idf_{ti}=(df_{ti}+1).(1/\log (df_{(ti-1)}+1)+1) \quad (1)$$

This produces a list of terms that can be sorted by their df-idft score. To maintain some word order information, we define terms as an n-gram, that is, n-word sequence. We then organize these tweets into groups called clusters. Each cluster Define a topic as a list of n-grams, we call this process to find bursty n-grams "BNgrams".

E. Topic Ranking

To maximize the usability of the topic range result of a large number of topics. Therefore, we want to classify the results by relevance. Here we use maximum n-gram, in this method we classify the topics according to the maximum value of df-idft of its constituent n-grams.

CONCLUSION

The goal of our work is to provide a solution to discover trend stories using BNgram and take an overview of the various methods that have emerged in recent years. This analysis shows that different technologies used throughout the paper take different ways to detect trending topic for various purposes. Despite the application of this method along with pre-processing to discover the stories of Twitter trends by extracting the truth on the ground. And this system is showing the history of the trends discussed in the particular time period using the efficient way.

FUTURE WORK

In future work, we extract the subject of trends and its initiator, that is, identify which user has the first tweet. It helps us find myth stories and their initiator.

REFERENCES

1) LucaMariaAiello,Sensing Trending Topics in Twitter, IEEE Transaction on Multimedia, Vol. 15 No.6, Oct. 2013 pp. 1268-1282.
2) Complete.com- http://siteanalytics.compete.com/twitter.com Site Profile for twitter.com Retr. July 1,2009
3) M. Cataldi, L. Di Caro, and C. Schifanella, “Emerging topic detection on Twitter based on temporal and social terms evaluation,” in Proc.
4) MDMKDD: 10th Int. Workshop Multimedia Data Mining, New York, NY, USA, 2010, pp. 4:1–4:10, ACM.. 271–350.
5) M. Mathioudakis and N. Koudas, “Twitter monitor: Trend detection over the Twitter stream,” in Proc. SIGMOD: Int. Conf. Management of
6) Data, New York, NY, USA, 2010, pp. 1155–1158, ACM.
7) J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, andJ.
8) Sperling, “Twitter stand: News in tweets,” in Proc. GIS: 17th ACM Int.
9) Conf. Advances in Geographic Information Systems, New York, NY, USA, 2009, pp. 42–51.
10) J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini,and F.
11) Menczer, “Detecting and tracking political abuse in socialmedia,” in Proc. ICSWM: 5th Int. AAAI Conf. Weblogs and Social. Media, 2011.
12) M. D. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer, “Predicting the political alignment of twitter users,” in Proc.SocialCom: 3rd IEEE Int. Conf. Social Computing, Boston, MA, USA,Oct. 2011.
13) H. Becker, M. Naaman, and L. Gravano, “Beyond trending topics:Real-world event identification on Twitter,” in Proc. ICWSM: 5th Int.AAAI Conf. Weblogs and Social Media, 2011.
14) James Allan,Stephen Harding and Devid Fisher, Taking Topic Detection From Evaluation to practice, Carlos Martin,

David Corney, AyseGoker and Andrew MacFarlane, Miniing Newsworthy Topic from Social Media.
15) JjS. Phuvipadawat and T.Murata, "Breaking news detection and tracking in Twitter," in Proc.Web Intelligence and Intelligent Agent Technology, IEEE/WIC/ACM Int. Conf., 2010, vol. 3, pp. 120–114.

