IJCSC 0973-7391
International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 9 • Issue 2 pp. 66-72 March 2018 - Sept 2018 www.csjournals.com

Enhanced Security of Twitter Trending With Topic Level Key Factors over Twitter Trends Management

Fatima Rahmath[1], Md Ateeq Ur Rahman[2]
[1]Research Scholar, Dept. of Computer Science & Engineering, SCET, Hyderabad
[2]Professor and Head, Dept. of Computer Science & Engineering, SCET, Hyderabad

Abstract: Twitter trends, a timely updated set of prime terms in Twitter, have the power to have an effect on the general public agenda of the community and have attracted abundant attention. Sadly, within the wrong hands, Twitter trends also can be abused to mislead folks. During this paper, we have a tendency to plan to investigate whether or not Twitter trends are secure from the manipulation of malicious users. We have a tendency to collect over sixty nine million tweets from five million accounts. Mistreatment the collected tweets, we have a tendency to 1st conduct a knowledge analysis and find out proof of Twitter trend manipulation. Then, we have a tendency to study at the subject level and infer the key factors that may confirm whether or not a subject starts trending because of its quality, coverage, transmission, potential coverage, or name. What we discover is that aside from transmission, all of things on top of are closely involving trending. Finally, we have a tendency to any investigate the trending manipulation from the attitude of compromised and pretend accounts and discuss counter measures.

Index Terms: Twitter trend, Social computing, Security.

I. INTRODUCTION

The Internet has subverted the autocratic approach of disseminative news by ancient media like newspapers. on-line trends are completely different from ancient media as a way for data propagation. maybe, Google Hot Trends ranks the most popular searches that have recently old a unforeseen surge in quality. Meanwhile, these trends could attract far more attention than before thanks to their look on Google Hot Trends. additional recently, on-line Social Networking (OSN) like Twitter has inaugurated a replacement era of "We Media." Twitter could be a time period small blogging service. Users broadcast short messages now not than a hundred and forty characters (called tweets) to their followers. Users may also talk over with the others on a range of topics at can. The topics that gain unforeseen quality are hierarchal by Twitter as an inventory of trends (also called trending topics). Twitter and Google trends became a crucial tool for journalists. Twitter particularly is employed to develop stories, track breaking news, and assess however popular opinion is evolving within the breaking story. during this paper, the first queries we have a tendency to arrange to answer square measure whether or not the malicious users will manipulate the Twitter trends and the way they may be ready to do that? Being exposed to time period trending topics, users square measure entitled to possess insight into however those trends really go trending. Moreover, this analysis additionally forged lightweight on the way to enhance an advert promotion campaign by fairly exploitation Twitter trends. to research the chance of manipulating Twitter trends, we'd like to deeply perceive however Twitter trending works. Twitter states that trends are determined by an formula and are continually topics that are straightaway widespread. However, the elaborated trending formula of Twitter is unknown to the general public, and that we haven't any thanks to decide what it specifically is. Instead, we have a tendency to study Twitter trending at the subject level and infer the key factors that may confirm whether or not a subject trends from its quality, coverage, transmission, potential coverage, and name. once distinguishing those key factors that square measure related to the trends, we have a tendency to then investigate the manipulation and countermeasures from the angle of those key factors.

The major contributions of this work square measure as follows:

we have a tendency to demonstrate the proof of the present manipulation of Twitter trends. particularly, using an influence model, we have a tendency to analyze the dynamics of an endogenous hash tag and determine the manipulation from its endogenous diffusion. once additional work the manipulation within the dynamics, we have a tendency to disclose the existence of a suspect spamming infrastructure.

we have a tendency to study Twitter trending at topic level, considering topics' quality, coverage, transmission, potential coverage, and name. The corresponding dynamics for every issue higher than square measure extracted, then Support Vector Machine (SVM) classifier is employed to envision however accurately an element may predict trending. we discover that, aside from transmission, every studied issue is related to trending. we have a tendency to additional illustrate the interaction pattern between malicious accounts and documented accounts, with reference to trending.

we have a tendency to gift the threat of malicious manipulation of Twitter trending, given compromised and pretend accounts within the suspect spamming infrastructure we have a tendency to determined. Then we have a tendency to demonstrate however compromised and pretend accounts may threaten Twitter trending by simulating the manipulation of dynamics as compromised and pretend accounts would do. Corresponding countermeasures square measure then mentioned.

II. Related Works

These short messages tend to mirror a spread of events in real time, creating Twitter notably like minded as a supply of time period event content. during this paper, we have a tendency to explore approaches for analyzing the stream of Twitter messages to tell apart between messages concerning real-world events and non-event messages. Our approach depends on a fashionable family of combination statistics of locally similar message clusters.

Social media sites (e.g., Twitter, Facebook, and YouTube) have emerged as powerful means that of communication for individuals wanting to share and exchange data on a large sort of real-world events. These events vary from fashionable, wide better-known ones (e.g., a concert by a preferred music band) to smaller scale, native events (e.g., a neighborhood assemblage, a protest, or associate accident). Short messages denote on social media sites like Twitter will usually mirror these events as they happen.

For this reason, the content of such social media sites is especially helpful for time period identification of real-world events and their associated user-contributed messages, that is that the downside that we have a tendency to address during this paper. Twitter messages mirror helpful event data for a spread of events of various sorts and scale.

These event messages will give a collection of distinctive views, no matter the event sort (Diakopoulos, Naaman, and KivranSwaine 2010; Yardi and boyd 2010), reflective the points of read of users UN agency have an interest or participate in a happeningEven for planned events (e.g., the 2010 Apple Developers conference), Twitter users typically post messages in anticipation of the event. distinctive events in real time on Twitter could be a difficult downside, because of the heterogeneousness and huge scale of the info. Twitter users post messages with a spread of content sorts, as well as personal updates and varied bits of data (Naaman, Boase, and Lai 2010). whereas a lot of of the content on Twitter isn't regarding any specific real-world event, informative event

messages even so abound. As an extra challenge, Twitter messages, by design, contain very little matter data, and sometimes exhibit calibre (e.g., with typos and ungrammatical sentences).
Several analysis efforts have centered on distinctive events in social media normally, and on Twitter specially (Becker, Naaman, and Gravano 2010; Sakaki, Okazaki, and Matsuo 2010; Sankaranarayanan et al. 2009). Recent work on Twitter has began to method information as a stream, because it is made, however has chiefly centered on distinctive events of a selected sort (e.g., news events (Sankaranarayanan et al. 2009), earthquakes (Sakaki, Okazaki, and Matsuo 2010)). different work identifies the primary Twitter message related to a happening (Petrovic, Osborne, and Lavrenko 2010). ´ Our focus during this work is on on-line identification of realworld event content.
We establish every event—and its associated Twitter messages—using a web agglomeration technique that teams along locally similar tweets (Section three.1). we have a tendency to then reckon revealing options for every cluster to assist verify that clusters correspond to events (Section three.2). we have a tendency to use these options to coach a classifier to tell apart between event and non-event clusters (Section three.3). we have a tendency to validate the effectiveness of our techniques employing a dataset of over a pair of.6 million Twitter messages.

2.1 Existing System

We read variety|the amount|the quantity} of followers and therefore the number of being retweeted as prediction and estimation of influence, severally. it's evident that there exists an outsized gap between the prediction and estimation of influence before the spike, and once the spike, the estimation of influence falls and gets on the point of the prediction of influence. the foremost probably rationalization is that the manipulation before the spike ends up in exceptional retweets and once the spike, the manipulation ends.
Nevertheless, we will more check the accounts that are suspended by Twitter. it's intuitive to link manipulation to malicious accounts. By the time of checking accounts (about two months once travel sample and search stream), 118 accounts are suspended by Twitter. we have a tendency to compare the temporal feature (waiting time) of suspended accounts thereupon of the accounts not being suspended.

III. PROPOSED SYSTEM

We use a Kalman filter to come up with the synthesized dynamics. The Kalman filter provides a algorithmic means that to supply the estimation of unknown variables employing a series of measurements ascertained over time, containing noise and different inaccuracies. Since each dynamics are sampled from general dynamics, we are able to estimate incontinuous search dynamics from continuous sample dynamics then treat the calculable search dynamics because the input measurements of the Kalman filter. After that, we have a tendency to generate a syncretized dynamics by integration sample dynamics into search dynamics.demonstrates an example of the Kalman filter for hashtag "oomf." we have a tendency to plot sample dynamics, calculable search dynamics, and also the syncretized dynamics when Kalman filtering. The syncretized dynamics retain the essential options of sample and search dynamics however take away a number of the noise of calculable search dynamics.

IV. MODULES

In this section, we tend to present the proof of Twitter trend manipulation supported an influence model. Existing literature has known 2 necessary factors for topics turning into trends: the endogeneity that captures the propagation impact of the subject within the network and also the exogeneity that represents the thrust external to the network (e.g., the mass media) [9], [30]. First, we'd like to differentiate manipulation from exogenous factors. In general, exogenous factors represent external and bonafide factors, particularly the mass media. However, manipulation is meant either as malice or as a method to AN finish. however it's still not possible to quantify the distinction between them. To avoid the impact of exogenous factors, we elect the hashtags that solely unfold within social networks, like Twitter. Then, we tend to use AN influence model to capture the unfold thanks to the impact of social networks and trace out the proof of manipulation. variety of hashtags continuously flourish in Twitter. a number of them don't correspond to external events (e.g., an earthquake). we tend to decision these endogenous hashtags memes throughout this paper. Most of the memes area unit combos of words or acronyms, that area unit wont to categorical AN feeling or raise an issue. Since the memes don't seem to be related to any external events, the unfold of the memes may be solely thanks to the impact of social networks and manipulation.

The impact of social networks can be captured by the influence model [31], whereas the manipulation of a culture may be thought to be the trouble to drive the culture to trend on the far side the impact of the network. to work out whether or not a hashtag could be a culture, we tend to manually check if the hashtag has been coated by any journalism we will verify our conjecture by investigation the accounts within the highest spike as shown in Fig.3. we tend to collect their friends (i.e., the accounts that they follow) and check whether or not their friends have shown up within the dynamics before, or in different words, whether or not the accounts within the spike be part of the subject once their friends. For the 4,055 accounts within the spike, 63.8% of them be part of the subject once their friends. There are still over 1,000 accounts that don't be part of the subject once their friends. we tend to couldn't merely create any conclusion supported the quantitative relation of the accounts that be part of once their friends as a result of the dynamics is sampled. all the same, we will additional check the accounts that are suspended by Twitter. it's intuitive to link manipulation to malicious accounts. By the time of checking accounts (about a pair of months once creep sample and search stream), 118 accounts are suspended by Twitter. we tend to compare the temporal feature (waiting time) of suspended accounts thereupon of the accounts not being suspended. Waiting time means that the interval from the time once AN account’s friend joins the subject to the time once the account itself joins. Fig.5 depicts the PDF of the waiting time of suspended accounts which of still-active accounts. it's evident that the waiting times of each types of accounts area unit largely among someday, that is analogous to the waiting times of different human activities following power-law distribution. However, the waiting times of these 2 types of accounts have a similar spikes around a hundred hours, implying there exist different malicious accounts that haven't nevertheless been detected by Twitter. we tend to additional check the predecessors of the accounts within the spike, and determine the accounts that have already been suspended by Twitter. we tend to outline descendants of account A as those accounts that follow account A and publish a minimum of one tweet of an explicit topic. we tend to then study the descendant range of the malicious accounts and also the descendant range of their initial generation and second generation, and then forth. Level zero denotes the malicious accounts themselves. Level one is that the initial generation of the malicious accounts.

Once showing the suspected manipulation of Twitter trends, we tend to proceed to infer the key factors of Twitter trending. during this section, we tend to initial syncretize sample dynamics and search dynamics to supply the syncretized dynamics. With the syncretized dynamics, we tend to then infer the key factors that interest trending mistreatment the SVM classification methodology Since sample dynamics and search dynamics area unit obtained from freelance streams, syncretizing sample dynamics and search dynamics might integrate the knowledge from each. Sample dynamics is continuous however could be a smaller portion of general dynamics, whereas search dynamics is discontinuous and consists of a bigger portion of general dynamics. we tend to use a Kalman filter to get the synthesized dynamics. The Kalman filter provides a algorithmic means that to supply the estimation of unknown variables employing a series of measurements determined over time, containing noise and different inaccuracies. Since each dynamics area unit sampled from general dynamics, we will estimate incontinuous search dynamics from continuous sample dynamics so treat the calculable search dynamics because the input measurements of the Kalman filter. After that, we tend to generate a syncretized dynamics by integration sample dynamics into search dynamics. Fig.7 demonstrates AN example of the Kalman filter for hashtag "oomf." we tend to plot sample dynamics, calculable search dynamics, and also the syncretized dynamics once Kalman filtering. The syncretized dynamics retain the essential options of sample and search dynamics however take away a number of the noise of calculable search dynamics.

The trending formula processes a stream of tweets and produces trends for users. From the user's perspective, the trending formula is meant to dig out the foremost every phase corresponds to a binary sign, that indicates whether or not the subject trends or not at the top of the phase. we elect Support Vector Machines (SVMs) as our classifier to work out however accurately an element might perform the binary classification. SVMs are wide wont to address many various classification issues, as well as written digit recognition [32], beholding [33], text classification [34], and image retrieval [35]. the essential purpose of SVMs during a binary classification drawback, is to map the feature vectors into a high dimensional area and realize the optimum hyperplane that represents the most important separation or margin between 2 categories. we tend to get d-dimensional feature vectors by scheming the statistics of the segments (e.g., mean and customary deviation) and acquire corresponding category labels supported the binary signs mentioned higher than.

4.1 Module Description:

In this project, we have three modules.

- WEB-API Module.
- Twitter Using Module.
- Twitter Trending Searching Module.
- Dynamic Searching Module.
- Graph Module.

WEB-API Module:

We acquire a sample stream via Twitter's Streaming API. we have a tendency to outline the fifteen most frequent hash tags within the sample stream as sample trends. Sample trends area unit retrieved from the sample stream each half-hour the web has subverted the autocratic manner of disseminative news by ancient media like newspapers. on-line trends area unit completely different from ancient media as a way for info propagation. as an instance, Google Hot Trends ranks the most well liked searches that have recently fully fledged a explosive surge in quality. Meanwhile, these trends might attract rather more attention than before because of their look on Google Hot Trends. a lot of recently, on-line Social Networking (OSN) like Twitter has inaugurated a replacement era of "We Media." Twitter may be a period of time small blogging service. Users broadcast short messages not than a hundred and forty characters (called tweets) to their followers. Users also can check with the others on a range of topics at can. The topics that gain explosive quality area unit hierarchic by Twitter as a listing of trends (also referred to as trending topics). Twitter and Google trends became a crucial tool for journalists. Twitter specifically is employed to develop stories, track breaking news, and assess however vox populi is evolving within the breaking story.

Twitter Using Module:
Primary queries we tend to decide to answer are whether or not the malicious users will manipulate the Twitter trends and the way they could be ready to do that? Being exposed to period of time trending topics, users are entitled to own insight into however those trends really go trending. Moreover, this analysis conjointly solid light-weight on the way to enhance an ad promotion campaign by moderately mistreatment Twitter trends. to analyze the chance of manipulating Twitter trends, we want to deeply perceive however trending works twitter. Twitter states that trends are determined by an algorithmic rule and are perpetually topics that are like a shot standard. However, the elaborate trending algorithmic rule of Twitter is unknown to the general public, and that we haven't any thanks to determine what it specifically is. Instead, we tend to study Twitter trending at the subject level and infer the key factors that may verify whether or not a subject trends from its quality, coverage, transmission, potential coverage, and name.

Twitter Trending Searching Module:
We outline the dynamics of a subject because the variation of the subject against the clock with reference to a selected frequency feature, similar to tweet range or account range. For an exact topic, we have a tendency to acquire its dynamics through its sample stream and search stream severally. Sample dynamics represent however the subject evolves within the sample stream, whereas search dynamics mirror the evolution of the subject within the search stream.

Dynamic Searching Module:
First, we want to tell apart manipulation from exogenous factors. In general, exogenous factors represent external and bonafide factors, particularly the mass media. However, manipulation is meant either as malice or as a method to an finish. however it's still not possible to quantify the distinction between them. To avoid the impact of exogenous factors, we decide the hash tags that solely unfold within social networks, like Twitter. Then, we tend to use an influence model to capture the unfold because of the result of social networks and trace out the proof of manipulation.

Graph Module
Interval from the time once an account's friend joins the subject to the time once the account itself joins. Depicts the PDF of the waiting time of suspended accounts which of still-active accounts. it's evident that the waiting times of each sorts of accounts square measure principally at intervals at some point, that is comparable to the waiting times of different human activities following power-law distribution. However, the waiting times of these 2 sorts of accounts have constant spikes around a hundred hours, implying there exist different malicious accounts that haven't nonetheless been detected by Twitter.

V. CONCLUSION

With the datasets we tend to collected via Twitter API, we first proof the manipulation of Twitter trending and observe a suspect spamming infrastructure. Then, we tend to use the SVM classifier to explore however accurately 5 various factors at the subject level (popularity, coverage, transmission, potential coverage, and reputation) might predict the trending. we tend to observe that, aside from transmission, the opposite factors ar all closely relating to Twitter trending. we tend to any investigate the interacting patterns between echt accounts and malicious accounts. Finally, we tend to present the threat posed by compromised and faux accounts to Twitter trending and discuss the corresponding countermeasures against trending manipulation.

References

[1] Wall Street Journal (Inside a Twitter Robot Factory), http://online.wsj.com
[2] Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. Detecting influenza epidemics using search engine query data. Nature, 457(7232), 1012-4.
[3] Nikolov, S. Trend or No Trend: A Novel Nonparametric Method for Classifying Time Series (Doctoral dissertation, Massachusetts Institute of Technology).
[4] Just, M., Crigler, A., Metaxas, P., and Mustafaraj, E. It's Trending on Twitter-An Analysis of the Twitter Manipulations in the Massachusetts 2010 Special Senate Election. In APSA 2012 Annual Meeting Paper.
[5] Ratkiewicz, J., Conover, M., and Meiss, M. Detecting and tracking the spread of astroturf memes in microblog streams. 5th International Conference on Weblogs and Social Media, 2010.
[6] Becker, H., Naaman, M., and Gravano, L. Beyond trending topics: Real-world event identification on twitter. ICWSM 2011.
[7] Zubiaga, A., Spina, D., and Martinez, R. Classifying Trending Topics: A Typology of Conversation Triggers on Twitter. CIKM 2011.
[8] Agarwal, M. K., Ramamritham, K., and Bhide, M. Identifying Real World Events in Highly Dynamic Environments. VLDB 2012.
[9] Naaman, M., Becker, H., and Gravano, L. Hip and trendy: Characterizing emerging trends on Twitter. Journal of the American Society for Information Science and Technology, 62(5), 902-918.
[10] Lee, K., Palsetia, D., Narayanan, R., Patwary, M. M. A., Agrawal, A., and Choudhary, A. Twitter Trending Topic Classification. 2011 IEEE 11th International Conference on Data Mining Workshops, 251-258.