International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 19
August 2018

Twitter Trends Manipulation: A First Look inside the Security of Twitter Trending

Rama Lakshmi Vemula[1], M.Venkateswara Rao[2], Sk.Khaleelullah[3]
[1]student, M.Tech (Cse), Lingayas Institute Of Management & Technology, A.P., India.
[2]assistant Professor, Dept. Of Computer Science & Engineering, Lingayas Institute Of Management & Technology, A.P., India.
[3]assistant Professor & Hod, Dept. Of Computer Science & Engineering, Lingayas Institute Of Management & Technology, A.P., India.

Abstract *—In this paper, 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. In this paper, we attempt to investigate whether Twitter trends are secure from the manipulation of malicious users. We collect more than 69 million tweets from 5 million accounts. Using the collected tweets, we first conduct a data analysis and discover evidence of Twitter trend manipulation. Then, we study at the topic level and infer the key factors that can determine whether a topic starts trending due to its popularity, coverage, transmission, potential coverage, or reputation. What we find is that except for transmission, all of factors above are closely related to trending. Finally, we further investigate the trending manipulation from the perspective of compromised and fake accounts and discuss countermeasures.*

KEYWORDS — TWEETS, MANIPULATION, SECURITY.

INTRODUCTION

Meanwhile, these trends may attract much more attention than before due to their appearance on Google Hot Trends. More recently, Online Social Networking (OSN) like Twitter has inaugurated a new era of "We Media." Twitter is a real-time micro blogging service. Users broadcast short messages no longer than 140 characters (called tweets) to their followers. Users can also discuss with the others on a variety of topics at will. The topics that gain sudden popularity are ranked by Twitter as a list of trends (also known as trending topics). Twitter and Google trends have become an important tool for journalists. Twitter in particular is used to develop stories, track breaking news, and assess how public opinion is evolving in the breaking story. In this paper, the primary questions we attempt to answer are whether the malicious users can manipulate the Twitter trends and how they might be able to do that? Being exposed to real-time trending topics, users are entitled to have insight into how those trends actually go trending. Moreover, this research also cast light on how to enhance a commercial promotion campaign by reasonably using Twitter trends. To investigate the possibility of manipulating Twitter trends, we need to deeply understand how Twitter trending works. Twitter states that trends are determined by an algorithm and are always topics that are immediately popular. However, the detailed trending algorithm of Twitter is unknown to the

public, and we have no way to find out what it specifically is. Instead, we study Twitter trending at the topic level and infer the key factors that can determine whether a topic trends from its popularity, coverage, transmission, potential coverage, and reputation. After identifying those key factors that are associated with the trends, we then investigate the manipulation and countermeasures from the perspective of these key factors.

The major contributions of this work are as follows: demonstrate the evidence of the existing manipulation of Twitter trends. In particular, employing an influence model, we analyze the dynamics of an endogenous hash tag and identify the manipulation from its endogenous diffusion. After further investigating the manipulation in the dynamics, we disclose the existence of a suspect spamming infrastructure.

study Twitter trending at topic level, considering topics' popularity, coverage, transmission, potential coverage, and reputation. The corresponding dynamics for each factor above are extracted, and then Support Vector Machine (SVM) classifier is used to check how accurately a factor could predict trending. We find that, except for transmission, each studied factor is associated with trending. We further illustrate the interaction pattern between malicious accounts and authenticated accounts, with respect to trending.

present the threat of malicious manipulation of Twitter trending, given compromised and fake accounts in the suspect spamming infrastructure we observed. Then we demonstrate how compromised and fake accounts could threaten Twitter trending by simulating the manipulation of dynamics as compromised and fake accounts would do. Corresponding countermeasures are then discussed.

LITURATURE SURVEY

1) Why We Twitter: Understanding Microblogging Usage and Communities

AUTHORS: A. Java, X. Song, T. Finin, and B. Tseng

Microblogging is a new form of communication in which users can describe their current status in short posts distributed by instant messages, mobile phones, email or the Web. Twitter, a popular microblogging tool has seen a lot of growth since it launched in October, 2006. In this paper, we present our observations of the microblogging phenomena by studying the topological and geographical properties of Twitter's social network. We find that people use microblogging to talk about their daily activities and to seek or share information. Finally, we analyze the user intentions associated at a community level and show how users with similar intentions connect with each other.

2) Social Networks that Matter: Twitter Under the Microscope

AUTHORS: B. Huberman, D. Romero, and F. Wu

Scholars, advertisers and political activists see massive online social networks as a representation of social interactions that can be used to study the propagation of ideas, social bond dynamics and viral marketing, among others. But the linked structures of social networks do not reveal actual interactions among people. Scarcity of attention and the daily rhythms of life and work makes people default to

interacting with those few that matter and that reciprocate their attention. A study of social interactions within Twitter reveals that the driver of usage is a sparse and hidden network of connections underlying the "declared" set of friends and followers.

3) Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter

AUTHORS: G.L. Danah Boyd and S. Golder

Twitter - a microblogging service that enables users to post messages ("tweets") of up to 140 characters - supports a variety of communicative practices; participants use Twitter to converse with individuals, groups, and the public at large, so when conversations emerge, they are often experienced by broader audiences than just the interlocutors. This paper examines the practice of retweeting as a way by which participants can be "in a conversation." While retweeting has become a convention inside Twitter, participants retweet using different styles and for diverse reasons. We highlight how authorship, attribution, and communicative fidelity are negotiated in diverse ways. Using a series of case studies and empirical data, this paper maps out retweeting as a conversational practice.

MODULE DESCRIPTION

WEB-API Module:

We obtain a sample stream via Twitter's Streaming API. We define the 15 most frequent hash tags in the sample stream as sample trends. Sample trends are retrieved from the sample stream every 30 minutes The Internet has subverted the autocratic way of disseminating news by traditional media like newspapers. Online trends are different from traditional media as a method for information propagation. For instance, Google Hot Trends ranks the hottest searches that have recently experienced a sudden surge in popularity. Meanwhile, these trends may attract much more attention than before due to their appearance on Google Hot Trends. More recently, Online Social Networking (OSN) like Twitter has inaugurated a new era of "We Media." Twitter is a real-time micro blogging service. Users broadcast short messages no longer than 140 characters (called tweets) to their followers. Users can also discuss with the others on a variety of topics at will. The topics that gain sudden popularity are ranked by Twitter as a list of trends (also known as trending topics). Twitter and Google trends have become an important tool for journalists. Twitter in particular is used to develop stories, track breaking news, and assess how public opinion is evolving in the breaking story.

Twitter Using Module:

Primary questions we attempt to answer are whether the malicious users can manipulate the Twitter trends and how they might be able to do that? Being exposed to real-time trending topics, users are entitled to have insight into how those trends actually go trending. Moreover, this research also cast light on how to enhance a commercial promotion campaign by reasonably using Twitter trends. To investigate the possibility of manipulating Twitter trends, we need to deeply understand how trending works twitter. Twitter states that trends are determined by an algorithm and are always topics that are immediately popular. However, the detailed trending algorithm of Twitter is unknown to the public, and we have no way to find out what it specifically is. Instead, we study Twitter trending at the topic level and infer the key factors that

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 19
August 2018

can determine whether a topic trends from its popularity, coverage, transmission, potential coverage, and reputation.

Twitter Searching Module.

We define the dynamics of a topic as the variation of the topic against time with respect to a specific frequency feature, such as tweet number or account number. For a certain topic, we obtain its dynamics through its sample stream and search stream independently. Sample dynamics represent how the topic evolves in the sample stream, while search dynamics reflect the evolution of the topic in the search stream.

Dynamic Searching Module.

First, we need to distinguish manipulation from exogenous factors. In general, exogenous factors represent external and legitimate factors, especially the mass media. However, manipulation is intended either as malice or as a means to an end. But it is still impossible to quantify the difference between them. To avoid the impact of exogenous factors, we choose the hash tags that only spread inside social networks, like Twitter. Then, we employ an influence model to capture the spread due to the effect of social networks and trace out the evidence of manipulation.

Graph Module.

Interval from the time when an account's friend joins the topic to the time when the account itself joins. Depicts the PDF of the waiting time of suspended accounts and that of still-active accounts. It is evident that the waiting times of both kinds of accounts are mostly within one day, which is similar to the waiting times of other human activities following power-law distribution. However, the waiting times of those two kinds of accounts have the same spikes around 100 hours, implying there exist other malicious accounts that have not yet been detected by Twitter.

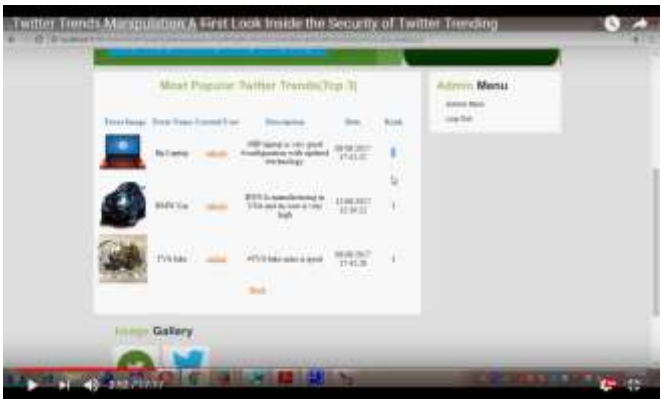
SAMPLE SCREENS

ADMIN LOGIN PAGE

CREATE TWEETS FILTERS PAGE

TWEETS RELATIONSHIPS PAGE

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 19
August 2018



MOST POPULAR TWITTERTRENTS PAGE

TWEET TOPIC WITH NUMBER OF NEHATIVEWORDS PAGE

CONCLUSION

In this paper, proposed the datasets we collected via Twitter API, we first evidence the manipulation of Twitter trending and observe a suspect spamming infrastructure. Then, we employ the SVM classifier to explore how accurately five different factors at the topic level (popularity, coverage, transmission, potential coverage, and reputation) could predict the trending. We observe that, except for transmission, the other factors are all closely related to Twitter trending. We further investigate the interacting patterns between authenticated accounts and malicious accounts. Finally, we present the threat posed by compromised and fake 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.

[11] Morstatter, F., Ave, S. M., and Carley, K. M., Is the Sample Good Enough? Comparing Data from Twitters Streaming API with Twitters Firehose, AAAI 2013.