

Singapore Management University
Institutional Knowledge at Singapore Management University

Research Collection School Of Information Systems

School of Information Systems

7-2012

Detecting Anomalous Twitter Users by Extreme Group Behaviors

Hanbo DAI

Singapore Management University

Ee-peng LIM

Singapore Management University, eplim@smu.edu.sg

Feida ZHU

Singapore Management University, fdzhu@smu.edu.sg

Hwee Hwa PANG

Singapore Management University, hhpang@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

 Part of the [Databases and Information Systems Commons](#), and the [Social Media Commons](#)

Citation

DAI, Hanbo; Ee-peng LIM; ZHU, Feida; and Hwee Hwa PANG. Detecting Anomalous Twitter Users by Extreme Group Behaviors. (2012). *Proceedings of the ACM International Conference on Net Science (NetSci)*. 1-2. Research Collection School Of Information Systems.

Available at: https://ink.library.smu.edu.sg/sis_research/2909

This Conference Proceeding Article is brought to you for free and open access by the School of Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Detecting Anomalous Twitter Users by Extreme Group Behaviors

by Hanbo DAI, Ee-Peng LIM, Feida ZHU and Hweehwa PANG

Singapore Management University

Twitter has enjoyed tremendous popularity in the recent years. To help categorizing and search tweets, Twitter users assign hashtags to their tweets. Given that hashtag assignment is the primary way to semantically categorizing and search tweets, it is highly susceptible to abuse by spammers and other anomalous users [1]. Popular hashtags such as #Obama and #ladygaga could be hijacked by having them added to unrelated tweets with the intent of misleading many other users or promoting specific agenda to the users. The users performing this act are known as the **hashtag hijackers**. As the hijackers usually abuse common sets of hashtags, they demonstrate common extreme group behaviors which can be used for detection.

In this research work, we aim to detect the hashtag hijacker groups in Twitter by their shared extreme behaviors in hashtag assignment. We model a group of such hashtag hijackers as an **extreme ranked anomalous collection (ERAC)**. The ERAC theory determines the anomalousness of a group of users by (a) a good number of them using some common hashtag very frequently; and (b) a good number of them demonstrating the same behavior using other hashtags. Detecting hashtag hijackers for large Twitter datasets using the ERAC theory can incur very costly computation. We instead can adopt a more efficient heuristic algorithm [2].

Our Twitter dataset consists of 21,666 users and their 231,803 tweets between September and November 2011. The tweets are collected from Twitter Streaming API and each tweet carries at least one of Singapore related hashtags¹. Using the ERAC theory, we are able to detect several hashtag hijacker groups and the corresponding abused hashtags as shown in the list below. The ten most anomalous hijacker groups can be categorized into pornography, job, travel and politics related depending on the content they promote. We derive some interesting observations about these hijacker groups. Firstly, each group of hijackers often post identical tweets containing the *same* URL. Secondly, the group size can be as large as 17 suggesting that the hijackers are highly coordinated and may be controlled by the same person or organization. Finally, most of hijacker groups promote some pornographic sites.

Group	Size	Category	Remarks
1	7	Pornography	The users posted identical tweets that contain the URL of a Brazilian pornography website. Every tweet from them contains #Japan, #Singapore, # Philippines, #Thailand, #Malaysia, #Hongkong, and #Indonesia. Most of the accounts have been suspended by Twitter.
2	5	Pornography	Similar to group 1 except that the hashtags are slightly different (#korea, #japan, #thailand, #hongkong, #guam, #taiwan, #philippines, #brunei, #indonesia, #malaysia, and #vietnam).
3	7	Pornography	Same as group 1 except that the hashtags are slightly different (#korea, #japan, #thailand, #hongkong, #guam, #taiwan, #philippines, #brunei, #indonesia, #malaysia, and #vietnam)
4	17	Job	The users posted identical tweets with the same URL pointing to the same job-hunting related website. The hashtag involved are #career, #job, and #jobs.
5	4	Pornography	Same as group 1 except that the hashtags are slightly different (#korea, #indonesia, #malaysia, : #philippines, #thailand, #hongkong, #guam, #brunei, #taiwan, #japan, #cambodia, and #china)
6	5	Pornography	Same as group 1 except that the hashtags are slightly different (#nepal, #indonesia, #malaysia, #taiwan, #japan, #philippines, #thailand, #hongkong, #korea, and #cambodia)
7	5	Pornography	Same as group 1 except that the hashtags are slightly different (#guam, #philippines, #thailand, #hongkong, #taiwan, #japan, #indonesia, #malaysia, #fiji, #china, #brunei, and #korea)
8	3	Travel	The users posted identical tweets followed by the same URL pointing to webpages promoting travel guidance publications. The abused hashtags are #urban, #cities, #kids, #city, #local, #newyork, #asia, #london, #hongkong, and #digital.
9	2	Politics	All their tweets contains messages "Occupy MANIFESTO FEARLESS HUMAN-ISM" followed by the same URL. The abused hashtags are #hyderabad, #hanoi, #occupy, #slovenia, #slovakia, #riyadh, #sierraleone, #baghdad, #dhaka, and #seychelles.
10	8	Job	Similar to group 4.

¹ #ndrsg, #sg, #sgedu, #sgselection, #sgselections, #sgpolitics, #sgpresident, #singapore, and #sosingaporean

References:

- [1] Chris Grier, Kurt Thomas, Vern Paxson, Michael Zhang: @spam: The Underground on 140 Characters or Less. ACM Conference on Computer and Communications Security (CCS), 2010.
- [2] Hanbo Dai, Feida Zhu, Ee-Peng Lim, Hwee Hwa Pang: Detecting Extreme Rank Anomalous Collections. SIAM International Conference on Data Mining (SDM), 2012.

Acknowledgements:

This research/project is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.