

Haters Gonna Hate: Job-Related Offenses in Twitter

Ricardo Kawase, Patrick Siehndel and Eelco Herder

L3S Research Center

Leibniz Universität Hannover, L3S, Appelstr. 9a, 30167 Hannover, Germany

{kawase, siehndel, herder}@L3s.de

ABSTRACT

In this paper, we aim at finding out which users are likely to publicly demonstrate frustration towards their jobs on the microblogging platform Twitter - we will call these users *haters*¹. We show that the profiles of *haters* have specific characteristics in terms of vocabulary and connections. The implications of these findings may be used for the development of an early alert system that can help users to think twice before they post potentially self-harming content.

Categories and Subject Descriptors

H.5.m. [Information Interfaces and Presentation]: Miscellaneous

General Terms

Human Factors, Verification

Keywords

FireMe!, WebScience, Social Networks, Twitter, User issues, Privacy awareness

1. INTRODUCTION

A 2013 social recruiting survey² provided by the social recruiting platform Jobvite³, claims that 94% of recruiters already use or plan to begin using social networks/social media for recruiting. Top social networks for recruiting are LinkedIn⁴, Facebook⁵ and Twitter⁶, with adoption of respectively 94%, 65% and 55%. Even the president of United

¹For the remainder of this paper, we call the author of an offensive tweet a '*hater*'.

²http://web.jobvite.com/rs/jobvite/images/Jobvite_SocialRecruiting2013.pdf

³<http://recruiting.jobvite.com>

⁴<https://www.linkedin.com>

⁵<https://www.facebook.com>

⁶<https://www.twitter.com>

States, Barack Obama, when giving advice to a high school class, once said: 'Be careful about what you post on Facebook, because in the YouTube age, whatever you do, it will be pulled up again later somewhere in your life'.

In this light, we extend our previous analysis on identifying the unawareness of Twitter users regarding their privacy [2]. We specifically choose to study those users who put their jobs at risk by publicly announcing their discontentment with their works or their bosses. Based on a representative sample, we identify the main features of users that are more likely to intentionally post something that is self-compromising.

We believe that many users could use some assistance when it comes to social network behavior. According to a recent report from the Pew Internet & American Life Project [3], particularly males and young adults have posted content that they regret; not surprisingly, these are also the users with the least restricted privacy settings. However, due to the raising awareness of privacy issues and their implications, more and more users actively manage their privacy settings and prune their profiles.

2. FIREME!

In order to address privacy issues and sensitive information leaks on social networks, we chose to tackle specifically those public updates on Twitter in which users express their disappointment regarding their jobs and bosses. To emphasize the recklessness of some people when posting updates about their working environments, we called our framework FireMe! [2].

In FireMe!⁷, we track every Twitter update in which the author's working environment is mentioned in an inappropriate, negative manner. We chose a set of thirteen sentences to catch a collection of such tweets. For example, sentences like 'I hate my job', 'I hate my boss', 'I have the worst job' and other sentences that include harsh profanity. Note that our goal is not to identify all possible tweets that contain inappropriate work-related content, but to sample a representative subset.

To address the real state of awareness of Twitter users, we built the online FireMe! alert system that warns these users about tweets that may put their jobs at risk - we assume that no boss would be happy to be publicly profaned online or to find out that their employees hate their jobs.

⁷<http://fireme.l3s.uni-hannover.de/>

Table 1: Averages characteristics of *haters*, *lovers* and a set of random users. We excluded outliers with more than 10,000 followers and more than 10,000 tweets. *Speed* is tweets/day.

	Haters	Lovers	Random
Users' profiles			
Followers	209	325	452
Friends	236	301	454
Tweets	3305	3188	4064
Within latest 200 tweets			
Speed	7.0	3.8	2.7
ReTweets	42.9	43.8	47.3
Profanity	14.7	8.3	10.2
Negative Tweets	27.0	20.5	21.9
Neutral Tweets	137.3	129.7	141.8
Positive Tweets	32.1	45.4	39.6

2.0.1 Dataset

Before deploying the FireMe! as an alert system, we first collected as many *haters* as we could during one week, between June 18 until June 26, 2012. In this period, we gathered a total of 21,852 *haters*, which corresponds to almost two reckless tweets per minute. During the same period, we also collected tweets from what we call '*lovers*', people who posted positive updates about their jobs, such as 'I love my job', 'my boss is the best'. We found twice as many *lovers*(44,710) than *haters*.

In addition, we polarized each tweet using sentiment140⁸ natural language processing API, based on the Maximum Entropy classifier[1], a state-of-the-art method for classifying the sentiment of tweets. For each given tweet, the service classifies it as positive, negative or neutral. Finally, we counted the number of profanity words in each tweet.

2.0.2 Data Analysis

A closer look at the collected data (Table 1) reveals interesting characteristics of *haters* in comparison to *lovers*. The first thing to notice is that *lovers* are better connected within the social graph. In our sample, *lovers* have significantly more followers and more friends. On the other hand, *haters* seem to be more active in terms of tweeting speed: they post twice as many tweets per day than *lovers*. As expected, *haters* are less careful regarding profanity in their language: they curse more often than *lovers*. Finally, we verified that *lovers* are generally more positive in their comments.

To illustrate the behavior of the different users, we created tag clouds of the top 50 most used *haters*' words. To avoid bias in the computation, we removed the tweets that were used to identify a user as a *lover* or a *hater*. From the remaining 199 tweets of each user, we removed stop words and computed the frequencies. Figure 1 depicts the tag clouds of *haters*. Tags in blue indicate words that do not appear in the equivalent *lovers*' tag cloud, while tags in red indicate that the word's frequency is higher than in *lovers*' tag cloud. It is interesting to see that even in the *haters*' tag clouds, the most frequent word is 'love'; however, the word appears 15% more frequently in the *lovers*' cloud. Haters



Figure 1: Haters' Tag Cloud

use the word 'hate' 34% more than *lovers*, in addition to a list of expletives.

3. DISCUSSION AND CONCLUSION

In this paper, we investigated which features characterize '*haters*', Twitter users who publicly complain about their jobs. The analysis extends and complements earlier work [2], in which we investigated users' awareness of the potential consequences of negatively loaded, personal tweets. As a basis for analysis, we used a representative subset of 'job haters', which we identified with only a couple of English-language *hate filters* - which were sufficient for identifying over two *haters* per minute.

Our data analysis shows that *haters* tweet more than regular users and are typically less connected than others. However, such aspects turned out not to be useful for pro-actively identifying potentially reckless users, opposing our previous hypothesis [2]. Another characteristic of *haters* is that their tweets are more negatively loaded. Our early experimental results showed that the content and polarity of tweets are the most crucial aspects to identify these users.

Our future work aims at improving the prediction methods with more accurate parameters and exploring the social graph in order to find additional *hating* evidence. In the long run, we aim to build an infrastructure that is able to alert users of their misbehavior, before they send out that one reckless tweet that might cost them their jobs.

Acknowledgments This work was partially funded by the European Commission Seventh Framework Program under grant agreement No.600826 for the ForgetIT project.

4. REFERENCES

- [1] A. Go, R. Bhayani, and L. Huang. Twitter sentiment classification using distant supervision. *Technical report, Stanford*.
- [2] R. Kawase, B. P. Nunes, E. Herder, W. Nejdl, and M. A. Casanova. Who wants to get fired? In *WebSci*, pages 191–194, 2013.
- [3] M. Madden. Privacy management on social media sites. Technical report, Pew Internet and American Life Project, 2012.

⁸www.sentiment140.com