Detection of Sockpuppets in Social Media

Suman Kalyan Maity
Dept. of CSE
IIT Kharagpur, India

Aishik Chakraborty
Dept. of CSE
IIT Kharagpur, India

Pawan Goyal
Dept. of CSE
IIT Kharagpur, India

Animesh Mukherjee
Dept. of CSE
IIT Kharagpur, India

Abstract

Online deception is a prevalent phenomena in social media. Creation of sockpuppets are one of the ways of online deception. The detection of such accounts are very important and crucial. We study various tweets and profile based features and propose an automated framework to early detect sockpuppet accounts in Twitter. We obtain high accuracy of 90.98% and a high recall of 0.88 in detecting the sockpuppet accounts.

Author Keywords

sockpuppet; social media; classification

ACM Classification Keywords

H.4.m [Information Systems Applications]: Miscellaneous; J.4 [Computer Applications]: [Social and Behavioral Sciences]; K.4.2 [Computers And Society]: [Social Issues]

Introduction

Sockpuppets are identities used for online deception. There can be several uses of creating such sockpuppet accounts on online social media platforms like business promotion, generating favorable book and film reviews etc. to create opinion bias towards an entity. Sockpuppets can also be created during online polls to submit multiple votes in favor of the puppeteer and the sockpuppets, each supporting the puppeteer's views in an argument, attempting to position

CSCW '17 Companion, February 25 - March 01, 2017, Portland, OR, USA
ACM 978-1-4503-4688-7/17/02.
http://dx.doi.org/10.1145/3022198.3026360

the puppeteer as representing majority opinion and sideline opposition voices. There are also instances of state-sponsored sockpuppetry [1] which corresponds to the use of paid Internet propagandists with the intention of swaying online opinion, undermining dissident communities, or changing the perception about what is the dominant view by a country's own government. In Twitter, the prevalent way of sockpuppetry are creation of accounts for the purpose of generating more followers and also for the purpose of conducting mass propaganda through retweets. However, it is difficult to identify such accounts without manual intervention. In this paper, we propose an automated method for detecting sockpuppets in online social media using a combination of tweets and profile based features. There have been several reports on fake followers [2] of the 2016 US presidential election candidates which prompted us to consider 2016 US presidential election as a potential scenario and we shall observe in the paper that indeed such sockpuppets have been created to increase followership of the candidates (by means of fake followers) or to propagate their opinions and views.

Related Work
There has been some previous research in the field of sockpuppetry and online deception. Among these, there has been a work on Wikipedia in [4] where the authors create a machine learning algorithm for detecting sockpuppets using tweet features that take into account the writing style of the culprit and the fake account. The authors also extended their work in [5] where they ran their algorithm using richer features on a larger dataset. There has been a study on sockpuppets on online forums in [8] where the authors identify puppets in the same forum as well as across forums using a scoring mechanism. In [7], the authors use a Wikipedia dataset to identify sockpuppets using non verbal features. Also there have been several works on spammer detection/Sybil detection in Twitter [2, 6, 1, 3].

Dataset and Labelling

We crawled the followers of the two US presidential nominee - Donald Trump and Hillary Clinton (@realDonaldTrump and @HillaryClinton respectively). The data collected for each follower includes the account details of the followers along with the most recent 3200 tweets and retweets from their timeline. Such data was collected for a total of 77186 followers of Donald Trump and 46023 followers of Hillary Clinton. We then randomly sample ~ 3400 followers and manually label them by two of the authors. We found 98 as sockpuppet accounts (agreed upon by both of them) among those set of followers.

Feature Engineering

To automate the detection of sockpuppets, we adopt a feature-based machine learning framework. We found two distinguishing feature category - tweet based features and profile based features which are instrumental for the classification task (to classify whether a twitter account is a sockpuppet account or not)

Tweet features
These set of features are derived from the tweets posted by the account.
Entropy of tweets This feature captures the regularity with which the user tweets. The feature is useful if the sockpuppet is a bot as the bot would want to tweet and retweet at regular intervals and thus will have different entropy than an ordinary user. We find the intervals between two tweets and say we find these intervals to be t_0 , t_1 , t_2 and so on. We find the probabilities Pr[t = t_0], Pr[t = t_1] which is simply the number of times the interval t_0 and t_1 occurs normalized by the total number of intervals and compute the entropy considering the above probabilities.

Normalized retweets count This is simply the (total number of retweets/ total number of tweets). Thus this feature tells us what fraction of tweets are retweets. This feature is useful as many sockpuppets tend to retweet the person they follow.

Profile based features
These set of features are derived from the profile information.
Verified or not This binary feature tells us whether the profile is verified or not.

[1] https://en.wikipedia.org/wiki/State-sponsored_Internet_propaganda

[2] http://53eig.ht/23BLdPh
http://bit.ly/1fFcmcZ
http://bit.ly/2cDrk23
http://bit.ly/1rL4zQI

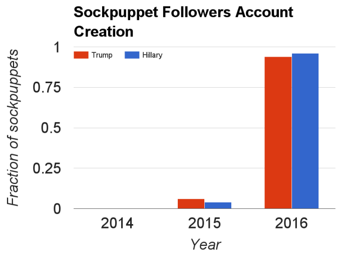

Figure 1: Account creation years for Trump and Hillary's sockpuppet followers.

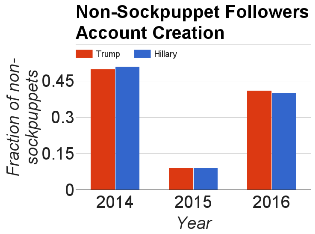

Figure 2: Account creation years for Trump and Hillary's non-sockpuppet followers.

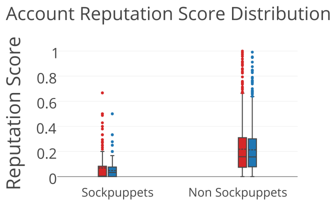

Figure 3: Comparison of reputation of different accounts.

Description is present or not This binary feature tells us if the profile description is present or not.

Location given or not This binary feature tells us if the location information is available or not.

Followers count This feature takes into account the number of followers of the user.

Friends count This feature takes into account the number of friends (accounts the user follows) of the user.

Reputation score This feature takes into account the reputation score of the user. The reputation score is calculated as: reputation score = no. followers/(no. followers + no. friends) Thus, celebrities who tend to have very large followership but who follow less people, i.e., having $followers >> friends$ have a reputation score close to 1. As we shall see later that sockpuppets have lower reputation scores than normal users.

Status count This feature tells us about the number of tweets the user has tweeted on his own.

Profile creation date This is a binary feature that tells us if the profile was created on/after 2015. This feature is relevant as most of sockpuppets, as we shall see, have been created after the announcement of the US presidential elections in 2015.

Classification

We formulate the detection of sockpuppets as a binary classification problem. In this setting of the problem, we are given a user profile as input and based on the features described above, we predict whether the user belongs to the class 'sockpuppet' or not. As we have a class imbalance, to mitigate this problem, we further select equal no of user accounts from the classes while training and testing. We use various classifiers like Support Vector Machines, Logistic regression, Random Forest etc. Table 1 shows the various classification techniques we employed and the evaluation results. SVM classifier performs the best as we obtain 90.98% accuracy with high recall value of 0.88.

Table 1: Classification Results

Classifier	Accuracy	Precision	Recall	F1-Score
SVM	90.98%	0.56	0.88	0.68
Logistic Regression	80.32%	0.52	0.39	0.45
Random Forest	88.52%	0.65	0.57	0.61

Analysis

This section presents some analysis of the tweeting behavior and other profile properties of sockpuppets in comparison to normal users. The analysis is done by running the above classifier on the entire dataset and then extracting relevant results from there. About **4.2%** of Trump's followers and **4.87%** of Hillary's followers have been detected as sockpuppets. Also we present a comparison between the sockpuppets who are followers of Hillary vs sockpuppets who are followers of Trump. In Figure 1, we observe that there were no sockpuppets who followed Trump before the year 2014. We also see that the number of sockpuppets is on the rise with a sharp increase in 2016.

Further we observe from Figure 1 that Hillary's increase in sockpuppets is larger. From Figure 2, we can infer that Trump has more normal follower accounts created in 2016 compared to Hillary.

In the feature description section, we had already defined the concept of account reputation. As we can see from the box plots in Figure 3 that on average the reputation of normal followers of both Trump and Hillary are higher than the reputation of the followers that are sockpuppets.

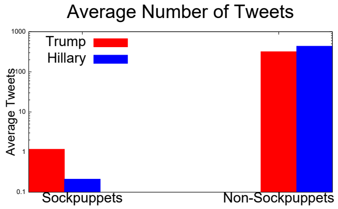

Figure 4: Tweeting behavior of sockpuppets vs normal users.

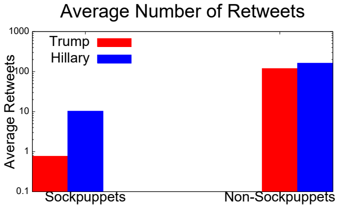

Figure 5: Retweeting behavior of sockpuppets vs normal users.

Figure 4 suggests that on average, the sockpuppets tend to tweet less than the normal users. Similarly, Figure 5 seems to suggest that the normal users retweet more than the sockpuppets. Figure 4 and 5 also seem to suggest that the sockpuppets in our dataset are mostly used to increase followership and less used as propaganda accounts.

Subscription Network. We created a network using the sockpuppet followers of Trump and Hillary as nodes and an edge was drawn if one follower followed the other. Incidentally for both Trump and Hillary, there were no edges in the graph. Thus we can conclude here that sockpuppets here, do not follow each other.

Conclusion and Future Work

We proposed an automated classification framework to detect sockpuppets in Twitter. We also see some trends regarding the tweeting behavior and retweeting behavior of sockpuppets and normal users. While there were no sockpuppets in the followers of the contestants before 2014, their number seemed to be increasing as the 2016 elections were approaching.

As a future work, we would like to analyze a larger dataset and obtain some more differences (possibly social) between sockpuppets and normal users. For instance we would like to understand how sockpuppets interact among each other, mention or make friendship with each other. Also, as we are continuously monitoring the followers of the two US presidential candidates, it would be interesting to investigate if the number of sockpuppets decrease after the elections are over.

REFERENCES

1. Fabricio Benevenuto, Gabriel Magno, Tiago Rodrigues, and Virgilio Almeida. 2010. Detecting spammers on twitter. In *CEAS '10*, Vol. 6. 12.
2. Zi Chu, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. 2010. Who is Tweeting on Twitter: Human, Bot, or Cyborg?. In *ACSAC '10*. ACM, 21–30.
3. Saptarshi Ghosh, Bimal Viswanath, Farshad Kooti, Naveen Kumar Sharma, Gautam Korlam, Fabricio Benevenuto, Niloy Ganguly, and Krishna Phani Gummadi. 2012. Understanding and combating link farming in the twitter social network. In *WWW '12*. ACM, 61–70.
4. Thamar Solorio, Ragib Hasan, and Mainul Mizan. 2013a. A case study of sockpuppet detection in wikipedia. In *LASM at NAACL HLT '13*. 59–68.
5. Thamar Solorio, Ragib Hasan, and Mainul Mizan. 2013b. Sockpuppet detection in Wikipedia: A corpus of real-world deceptive writing for linking identities. *arXiv preprint arXiv:1310.6772* (2013).
6. Kurt Thomas, Chris Grier, Dawn Song, and Vern Paxson. 2011. Suspended Accounts in Retrospect: An Analysis of Twitter Spam. In *IMC '11*. 243–258.
7. Michail Tsikerdekis and Sherali Zeadally. 2014. Multiple account identity deception detection in social media using nonverbal behavior. *IEEE Transactions on Information Forensics and Security* 9, 8 (2014), 1311–1321.
8. Xueling Zheng, Yiu Ming Lai, Kam-Pui Chow, Lucas CK Hui, and Siu-Ming Yiu. 2011. Sockpuppet detection in online discussion forums. In *IIH-MSP '11*. IEEE, 374–377.