Information Gain of Twitter ‘Social Spambots’

Joaquin Quintana
Applied Computer Science
University of Colorado
Boulder, Colorado, USA
Joaquin.Quintana@Colorado.edu

Jason Weidner
Applied Computer Science
University of Colorado
Boulder, Colorado, USA
Jason.Weidner@colorado.edu

Statement and Motivation (Eventual abstract)

Bots are ubiquitous on social media platforms and researchers have developed many techniques to identify bots and bot swarms. However, new bots are constantly being developed and evade known identification techniques.

Here we aim to use open source Twitter data, gathered by academic researchers, to see if we can successfully identify relatively modern and advanced ‘social spambots’ by the use of a few simple rules. Rule effectiveness will be evaluated by calculating information gain and validated by comparing information gain calculations on a new dataset of traditional spambots and compared to those published in ref 3, which were generated on traditional Twitter spambots. The particular dataset we downloaded was generated to identify so-called ‘social spambots’, a newer type of Twitter bot developed to evade traditional detection.

These social spambots are known to tweet normal content but also work in highly coordinated fashion to promote specific content (such as specific political propaganda during election time). Such bots have the ability to reach and sway the opinion of the masses, which can influence and alter the social fabric of society ref. 1, 2, 3, 4. Thus, a new arms race has begun and it is being fought on social media platforms. Being able to identify and counter social spam bots is going to be an area of research for sometime. Our work here aims to elucidate and validate known techniques for identifying contemporary and obsolete versions of social spambots for identification and classification. The dataset used here is a subset of Cresi compiled by the authors in ref. 1, 2, 3 and summarized in ref 1.

KEYWORDS

Social spam bot, decision tree, clustering, twitter, fake followers

1 Literature Survey

Cresci (2015) reports the creation of several datasets of fake and authentic Twitter accounts. Fake accounts were purchased from third party services (who sell such accounts) and scraped the account data. Authentic accounts were obtained by promoting the research, following up with individual account signups and verifying genuine users by CAPTCHA and evaluation by two sociologists. The paper also evaluates Twitter account data using rules promoted in media as useful for identifying bots. These metrics performed poorly. The paper also shares evaluations of rules promoted by academics to identify bots; these rules perform well.

Cresci (2017) suggests there is a new generation of 'social spambots' different to the traditional spambots present on Twitter. The new bots, the authors claim, evade detection by humans and state-of-the-art detection algorithms published in academic literature. The authors demonstrate that such bots evade detection through a series of analyses. In turn, the authors call for new methods to be developed that can successfully identify 'social spambots'.

Cresci (2017) is the source of the dataset used in our analysis of social spambots.

Cresci (2018) claims there is a recent (>2014) novel species of spambots specifically designed to mimic authentic user behaviour in order to evade state of the art bot identification routines. These bots, according to the authors, are highly coordinated and deployed in a way that they routinely promote similar content. Authors also claim these bots can be identified by using a 'digital DNA' technique. Authors manually translated Twitter account activity (and Tweet content) into sequences of letters representing the type of activity (or content) present. Then the researchers compared the sequence (DNA) to those from known human and known bot accounts, showing that bot activity was markedly different (particularly so for one group of bots compared to another).

2 Proposed work

2.1 Data cleaning

The data is well organized but does contain several columns which are entirely NaN or not useful for computation. Data wrangling to place identified attributes into a table easing analysis and size reduction.

2.2 Data integration

The data is already labeled (bot/authentic) by the academic team in ref. 3. Subsets of the dataset will be used to compute information gain. Functions are being compromised to extract smaller amounts of data from each of the files to perform analysis on.

3 Data set

The dataset used by Cresci in ref. 1, 2 was acquired from Github Bot Repository from Botometer Download datasets zip from Bot Repository.

The dataset is being mined and tracked via github collaboration. Code for the project is being placed in a Jupyter notebook linked to the github account..

4 Evaluation Methods

We will validate our methods by comparing our information gain calculations on traditional spambots to those in Cresci 2015 table 8, as they generated similar results for a different dataset of traditional spambots. We will calculate information gain on a few rules that generated relatively large information gain values in Cresci (2015) as well as rules that generated low gain values. We will then use the validated method definitions to analyse the social spambot dataset. In both cases - traditional and social spambots analyses - we will have to include genuine accounts too.

5 Tools

The group decided to use the well known Python language in a Jupyter notebook for tacking and sharing live code. The Jupter notebook is available on the shared github account and the group is using git to pull and push updates as needed. All team efforts are being tracked via this account.

Pandas was chosen for data cleaning and wrangling due to its speed of processing large data volumes. In our particular dataset, some collections of tweets are 1 GB csv files.

Matplotlib and Seaborn together are used for visualization and Numpy for statistical analysis.

6 **Milestones**

Our goal is to calculate information gain, accuracy and recall for several rules for two different subsets of the data: (1) social spambots and (2) traditional spambots. This will enable readers or other analysts to generate a decision tree for identifying social spambots using simple rules.

Timeline

Task	**Deadline**
List functions required to execute analysis	March 25, 2021
Write function docstrings - id parameters - parameter data types - return data type - description of function	March 25, 2021
Program function definitions	April 8, 2021
Define users to include in analysis (subset of full data set)	April 8, 2021
Execute 50% analysis (analysis on traditional bots)	April 16, 2021
Execute 100% analysis	April 29th, 2021

REFERENCES (no particular order as of now)

1 Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi, and Maurizio Tesconi. 2017. The Paradigm-Shift of Social Spambots. *Proceedings of the 26th International Conference on World Wide Web Companion - WWW '17 Companion* (2017). DOI:http://dx.doi.org/10.1145/3041021.3055135

2 Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi, Maurizio Tesconi, "Social Fingerprinting: Detection of Spambot Groups Through DNA-Inspired Behavioral Modeling," in IEEE Transactions on Dependable and Secure Computing, vol. 15, no. 4, pp. 561-576, 1 July-Aug. 2018, doi: 10.1109/TDSC.2017.2681672.

3 Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi, Maurizio Tesconi, Fame for sale: Efficient detection of fake Twitter followers,Decision Support Systems, Volume 80, 2015, Pages 56-71, ISSN 0167-9236, https://doi.org/10.1016/j.dss.2015.09.003.

4 Stefano Cresci. 2020. A decade of social bot detection. *Communications of the ACM* 63, 10 (2020), 72–83. DOI:http://dx.doi.org/10.1145/3409116

5 Besko and Carley. 2020. TBot-Match: Social Bot Detection with

6 Zhen Guo, Jin-Hee Cho, Ing-Ray Chen, Srijan Sengupta, Michin Hong, and Tanushree Mitra. 2020. Online Social Deception and Its Countermeasures: A Survey. *IEEE Access* 9 (December 2020), 1770–1806. DOI:http://dx.doi.org/10.1109/access.2020.3047337