© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 25, 2019

Bot detection on twitter

Alina Amanzholova[1], Aysun Coskun[2], Yasin Akman[3]

[1,2] Gazi University, Faculty of Technology, Department of Computer Engineering
[3] Selcuk University, Huglu Vocational Highschool, Department of Computer Technologies

[1]alina.amanzhol07@gmail.com, [2]aysunc@gazi.edu.tr , [3]yasinakman@selcuk.edu.tr

Abstract. Twitter is one of the most popular social media platforms with 319 million monthly active users who publish 500 million tweets per day. With this popularity, spam accounts are also emerging for phishing on Twitter or spreading malicious software, advertising using shared URLs on tweets, following legitimate users, and attracting their attention, as well as handling trending topics to spread sexual content. In this article, the features of Twitter bot detection are presented. In addition, Twitter bot detection methods and tools are described. The aim of this study is to determine the best method by comparing and comparing the methods used in the literature for the detection of bot in Twitter.

1. Introduction

Bot is an automated running software, which is defined as an abbreviated version of the robot, which is responsible for performing any activity on computers or software. Bots- mediated online manipulation reports have been used in political speech [1], fake news [2], conspiracy theories [3], stock manipulation [4], human health [5], propaganda [6] and in some rare cases [7].

Bots also attracted the attention of the cyber security research community: Sometimes, large group bots, as shown on Twitter, acting behind command-control-style scenes similar to traditional botnets used to distribute cyber attacks and other cyber security threats, it is controlled by the same entity called the master of the bot [8].

Much work on bot detection assumes extensive access to social media data. For example, Wang et al. they used clustering techniques to identify large-scale behavioral anomalies [9], while other authors used controlled learning to analyze all accounts of certain platforms and separate bots from humans [10]. Some have published studies showing the effectiveness of applications such as SybilRank [11] or Facebook Immune System [12]. To prevent limitation of unrestricted data access, other techniques are designed to require smaller user activity instances and fewer tagged bots and human user instances. Examples of such a trend include the classification system proposed by Chu et al. [13], Wang et al. [9] the system based on the source of the crowd he designed, NLP-based detection techniques presented by Clark et al. [14] and BotOrNot [15].

Currently there are 319 million monthly active users on Twitter. [16] Based on research from the University of Southern California and Indiana University, they have up to 15% of them. This means that roughly 48 million accounts are bots, not people. We will talk about data tools that analyze Twitter data to reveal these data and identify partnerships that can associate them. We recommend the best methods of detecting bots in the literature.

 Online ISSN 2603-459X

In this article, first of all, the bot detection features of Twitter are described in detail in Chapter 2. Then, the methods of spambot detection methods on Twitter are described in Chapter 3 and the data tools for detecting bot in Twitter. Finally, the result of the research paper is given.

2. Twitter bot detection features

The features of Twitter bot detection are categorized as follows:

(1) Account-based features, (2) tweet-based features.

Each feature category is discussed in the following subsections.

A. Account-based features

Bot users can be identified by analyzing Twitter accounts that contain the features listed in Table 1. Some of the features such as biography, location, home page and date of creation are useless because they are controlled by the user.

Table 1. Description of Account Based Features

Feature	Description
Username	The unique identifier of the account
Biography	The biography of the account
Location	The location of the account
Statuses Count	Total Account Status
Followers Count	Total Checks of Account
Friends Count	Total Friends of Account
Favorites Count	Total Number of Favorites (favorite) of the account
Listed Count	Total Listed Count of the Account
Default Profile	Default Profile of the Account
Profile Uses Background image	Profile Background Picture
Tweet Count	Total number of tweets the account has
Number of likes	Total tweets of the account
Number of retweets	Total retweets of the account's tweets
Number of moments	The total number of moments the account has

Bot senders tend to publish the same or similar tweets published by one or more controlled accounts [17].

B. Tweet based Features

Bot senders tend to publish a large number of unwanted tweets to take care of normal users. Spammers can be detected by analyzing their tweets. This is necessary to filter spam tweets from legitimate ones and provide users with a spam-free environment that is Twitter's purpose [18]. Each tweet contains the information listed in Table 2.

Table 2. Description of Tweet Based Features

Feature	Description
Retweet Count	The number of retweets the tweet has
Reply Count	The number of replies the tweet has
Favorite Count	The number of favorite the tweet has
Number of Hashtags	The number of hashtags the tweet has
Number of URLs	The number of URL's the tweet has
Number of Mentions	The number of mentions the tweet has
Sender	Tweet's sender
Sent date	The date Tweet was sent
Location	Detected location of where Tweet is saved

Bot senders tend to use many hashtags (especially trending ones) to reach more users. In total, Twitter bots are estimated to create approximately 24% of tweets that are on Twitter.

3. Twitter bot detection methods

Below is a characteristic table of the bot detection studies in the literature.

Table 3. Characteristic table of bot detection studies in the literature

Study	Technique	Method	Dataset	Accuracy
Deep neural networks for bot detection. Sneha Kudugunta, Emilio Ferrara (2018)	Deep neural networks	Account-based: Decision tree+ SMOTE Tweet based: LSTM	Cresci et al. dataset (2017) 8386 user accounts and 11,834,866 tweets	99,81% 96,33%

İdentifikatsiya botov v sotsialnyh setyah na baze tehnologiy intellektualnogo analiza dannyh. A.O. Evseeva, R.İ. Gumerova, A.S. Katasev, A.P. Kirpichnikov (2017)	Data mining	Account-based: 1.Neural networks 2.Decision tree 3.Logistic regression Tweet based: -	50 user bot in 200 user accounts	
Empirical Evaluation and New Design for Fighting Evolving Twitter Spammers. C. Yang, R. Harkreader, G. Gu (2011)	Machine learning	Account-based: - Tweet based: 1.Random forest 2.Decision tree 3.Naive Bayes 4.Decorate	500,000 Twitter accounts and more than 14 million tweets	94,7%
Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg? Z.Chu, S. Gianvecchio, H. Wang, S. Member (2012)	Machine learning	Account-based: Depth-First Search (DFS) Tweet based: -	A total of 512.407 user datasets from Twitter	97,6%

Online Human-Bot Interactions: Detection, Estimation, and Characterization. O.Varol, E. Ferrara, C.A. Davis, F. Menczer, A. Flammini	Machine learning	Account-based: K-Means Tweet based: -	A total of 14 milion user datasets from Twitter	94%
Measuring bot and human behavioral dynamics. I.Pozzana, E.Ferrara (2018)	Data mining	Account-based: 1.Decision tree 2.Extra tree 3.Random Forests. 4.k Nearest Neighbors Tweet based: 1.Decision tree 2.Extra tree 3.Random Forests. 4.k Nearest Neighbors	380,000 accounts 16 million tweets	97%

In the Kudugunta and Ferrara study, the bot was detected using SMOTE [19] and LSTM using two techniques an unbalanced data set [20]. Using the SMOTE technique over the account 99,81% and using the LSTM technique over tweet and have achieved high accuracy of 96.33%.

K-Means Clustering Algorithm is one of the most widely used algorithms in the world of data mining. O.Varol et al. have identified normal and bot users over 14 million users in the study. They obtained high 94% accuracy using the K-means algorithm [21].

Neural networks. The use of neural networks to resolve the classification problem involves identifying the input image represented by a feature vector to one or more predetermined classes [22-23]. In the study of Evseeva et al., bot detection was performed by using data mining methods. The decision tree, logistic regression and neural networks methods to compare the percentage of error, as the best method of neural networks have stated [24].

Decision trees are a way of representing rules in a hierarchical, consecutive structure, and each object has a single node that provides a solution. [25].

In Pozzana and Ferrara, the bot was detection using decision tree, extra tree, random forest, k-means algorithms. Thus, a high 94% accuracy was obtained [10].

Logistic regression. Regression algorithms calculate the dependencies between numerical values. The linear regression model uses the best linear approach to present the data obtained. The resulting approach allows to estimate the values of the dependent variables for any value of the independent variables [26].

4. Bot detection tools on Twitter

4.1. Botometer

This is the development of Indiana University and Northwestern University, which helps determine whether an account is a bot. He considers more than a thousand factors to evaluate the service. During the analysis, the Botometer examines the hashtags, the tweet language, the publication time, the text style, the account trackers, and references to other accounts. Score - The probability of whether the account is a bot given as a percentage. The higher this score, the more likely we are to face the bot [27].

4.2. BotCheck

The authors of this service are two students who develop a tool to analyze the content of records. If there are many manipulative statements and controversial expressions in the Twitter account, BotChek marks this as a bot "political bot" or "a moderate account for political propaganda" on Twitter.
Tests showed that the results of BotChek and Botometer studies were different. The control of several accounts showed that BotChek has defined the Botometer's accounts as normal accounts, which are defined as bot (50% below the estimate). For the test, they used the account bots, where manual control was performed and evidence of unusual activities [28].

4.3. BotOrNot

This service has existed since 2014, but there are operational disruptions at this stage. The service was created by experts from Bloomington University (Indiana), funded by the US Department of Defense and the National Science Foundation. It can analyze hundreds of parameters, including content, in real time [29]. To use, you must log in with your Twitter account and specify the names of suspicious accounts on Twitter.

5. Conclusion and recommendation

Traditional bot filtering methods cannot detect bot senders on Twitter, because Twitter has unique features from e-mail services and websites. Therefore, a more powerful spam detection approach specifically designed for Twitter is required. To provide a spam-free environment, spammers must be identified and filtered as well as bots. In this article, bot identification was investigated by considering the features of bot detection in Twitter and the approaches proposed in the literature.

According to the results of the study in the literature with the accuracy of 99.81% high data mining methods over the account SMOTE and 96.33% accuracy using LSTM techniques over tweet. Therefore, it is recommended to use these methods in future studies.

Using bot detection tools on Twitter, the user is able to control his account. In this study, it is recommended that the tools used by users with a Twitter account to determine if there are bot users among followers and followers.

It is foreseen that the study will guide the researchers in the field of bot identification.

6. References

[1] A. Bessi , E. Ferrara. (2016). Social bots distort the 2016 us presidential election online discussion, First Monday 21 (11).
[2] A. Badawy, E. Ferrara, K. Lerman. (2018). Analyzing the digital traces of political manipulation: the 2016 Russian interference twitter campaign, arXiv: 1802. 04291.
[3] V. Subrahmanian , A. Azaria , S. Durst , V. Kagan , A. Galstyan , K. Lerman , L. Zhu , E. Ferrara ,

A. Flammini , F. Menczer. (2016). The darpa twitter bot challenge, Computer 49 (6), p. 38–46 .
[4] E. Ferrara. (2015). Manipulation and abuse on social media, ACM SIGWEB Newslett. (Spring). 4.
[5] J. Allem , E. Ferrara. (2016). The importance of debiasing social media data to better understand e-cigarette-related attitudes and behaviors, J. Med. Internet Res. 18 (8).
[6] A. Badawy , E. Ferrara. (2018). The rise of jihadist propaganda on social networks, Journal of Computational Social Science. 1(2). p. 453–470.
[7] B. Monsted , P. Sapiezynski , E. Ferrara , S. Lehmann. (2017). Evidence of complex contagion of information in social media: an experiment using twitter bots, PLoS ONE 12(9): e0184148. https://doi.org/10.1371/journal.pone.0184148 .

[8] N. Abokhodair , D. Yoo , D.W. McDonald. (2015). Dissecting a social botnet: growth, content and influence in twitter. Presented at the ACM conference on Computer- Supported Cooperative Work and Social Computing. p. 839–851 .
[9] G. Wang , M. Mohanlal , C. Wilson , X. Wang , M. Metzger , H. Zheng , B.Y. Zhao. (2013). Social turing tests: crowdsourcing sybil detection, in: Proc. of the 20th Network & Distributed System Security Symposium (NDSS).
[10] I. Pozzana, E. Ferrara. (2018). Measuring bot and human behavioral Dynamics. arXiv: 1802.04286.
[11] Q. Cao , M. Sirivianos , X. Yang , T. Pregueiro. (2012). Aiding the detection of fake accounts in large scale social online services, in: 9th USENIX Symp on Netw Sys Design
& Implement. p. 197–210.
[12] T. Stein , E. Chen , K. Mangla. (2011). Facebook immune system, in: Proc. of the 4th Workshop on Social Network Systems, ACM. p. 8 .
[13] Z. Chu, S. Gianvecchio, H. Wang, S. Jajodia. (2012). Detecting automation of twitter accounts: are you a human, bot, or cyborg? IEEE Trans. Depend. Secure Comput. 9 (6), p. 811–824.
[14] E. Clark, J. Williams, C. Jones, R. Galbraith, C. Danforth, P. Dodds. (2016). Sifting robotic from organic text: a natural language approach for detecting automation on twitter. J. Comput. Sci. 16 p. 1–7.
[15] C.A. Davis, O. Varol, E. Ferrara, A. Flammini, F. Menczer. (2016). Botornot: a system to evaluate social bots, in: Proceedings of the 25th International Conference Companion on World Wide Web, International World Wide Web Conferences Steering Committee, p. 273– 274 .
[16] M. Newberg, CNBC 20.03.2017. [Online].
http://www.cnbc.com/2017/03/10/nearly-48-million-twitteraccounts-could-be-bots-says- study.html. [Date of access:20.11.2018].
[17] A. Kabakus, R. Kara. (2017). A Survey of Spam Detection Methods on Twitter. (IJACSA) International Journal of Advanced Computer Science and Applications, 8(3).
[18] C. Yang, R. Harkreader, G. Gu. (2013). Empirical Evaluation and New Design for Fighting Evolving Twitter Spammers, IEEE Trans. Inf. Forensics Secur. 8 p.1280–1293.
[19] N.V.Chawla, K.W.Bowyer, L.O.Hall, W. Kegelmeyer. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, s. 321 -357.

[20] S. Kudugunta ve E. Ferrara. (2018) Deep neural networks for bot detection.
Information Sciences, 467, p. 312-322.
[21] O.Varol, E. Ferrara, C.A. Davis, F. Menczer ve A. Flammini. Online Human-
Bot Interactions: Detection, Estimation, and Characterization. arXiv:1703.03107
[22] A.S. Katasev, D.V. Kataseva, A.P. Kirpichnikov. (2015). Neirosetevaya diagnostika anomalnoi setevoi aktivnosti, Vestnik tehnologicheskogo universiteta, 18(6), p. 163-167.
[23] A.S. Katasev, D.V. Kataseva. (2015). Razrabotka neirosetevoi sistemy klassifikaciy elektronnyh pochtovyh soobsheniy, Vestnik Kazanskogo gosudarstvennogo universiteta, 1(25), p.68-78.
[24] A.O. Evseeva, R.İ. Gumerova, A.S. Katasev, A.P. Kirpiçnikov. (2017). İdentifikatsiya botov v sotsialnyh setyah na baze tehnologiy intellektualnogo analiza dannyh. Vestnik

tehnologicheskogo universiteta, 20(5), p. 87-90.
[25] D.P. Zegjda, T.V. Stepanova. (2012). Ocenka effektivnosti ispolzovaniya sredstv zashity dlya neitralizaciy i ustraneniya bot-setei, Problemy informacionnoi bezopasnosti Komputerniye sistemy. 2. p 21-27.
[26] A.S. Katasev. (2013). Formirovaniye bazy znaniy sistemy filtraciy elektronnyh pochtovyh soobsheniy, Nauchno- tehnicheskiy vestnik Povoljya. 5. p. 191-194.
[27] https://www.stopfake.org/kak-identifitsirovat-bota-v-twitter-servisy-i- instrumenty/ (Date of access:rişim tarihi: 02.01.2019)
[28] https://botcheck.me/ (Date of access:02.01.2019) [29] http://botornot.co/ (Date of access:02.01.2019)
Stinson D R 2002 *Cryptography: Theory and Practice Second Edition* CRC Press
Schneier B 1996 *Applied Cryptography Second Edition* (New York: John Wiley & Sons Inc)
Lovine J 2012 *PIC Projects for Non-Programmers* (USA:Elsevier) chapter 7 pp 115-149
Ascii table *http://www.asciitable.com/*
Arduino software *http://www.arduino.cc/en/Main/ Software*
Canberk G and Sagıroglu S 2006 *Bilgi ve Bilgisayar Güvenliği Casus Yazılımlar ve Korunma Yöntemleri* (Ankara: Grafiker Ltd. Şti.)