RES MILITARIS Social Science Journal

TWEET-BASED BOT DETECTION USING BIG DATA ANALYTICS

[1]K.Nageswari,[2]Adaniya Hanuman,[3]O.Udhay Kiran,[4]Dr.K.Srikanth

[1,2,3]IV Year Student,[4]Associate Professor

Department of CSE

Visvesvaraya College Of Engineering & Technology,Ibrahimpatnam,Telangana

ABSTRACT

Twitter is one of the most popular micro-blogging social media platforms that has millions of users. Due to its popularity, Twitter has been targeted by different attacks such as spreading rumors, phishing links, and malware. Tweet-based botnets represent a serious threat to users as they can launch large-scale attacks and manipulation campaigns. To deal with these threats, big data analytics techniques, particularly shallow and deep learning techniques have been leveraged in order to accurately distinguish between human accounts and tweet-based bot accounts. In this paper, we discuss existing techniques, and provide a taxonomy that classifies the state- of-the-art of tweet-based bot detection techniques. We also describe the shallow and deep learning techniques for tweet-based bot detection, along with their performance results. Finally, we present and discuss the challenges and open issues in the area of tweet-based bot detection.

I. INTRODUCTION

Nowadays, social media is one of the most popular tools used by people to communicate with one another. It is also largelyused by organizations to reach out to customers. In, it hasbeen reported that there are 3.5 billion active social mediausers globally. Facebook, Twitter, LinkedIn, and other social media networks are used by organizations to improve brand visibility and boost their sales. Twitter is one of the mostpopular social media platforms. It has 340 million active userswho are allowed to communicate at a large scale and sharetheir opinions about different topics.

Twitter could be targeted by various kinds of attacks. For example, a spear phishingattack in July 2020 led to the hijack of high pro_leTwitteraccounts [2]. Also, fraudulent accounts could be created toimpersonate legitimate users and organizations.Twitter can also be exploited by bot net, which is a set ofmalicious accounts that operate under a botmaster, and arecontrolled by software programs rather than human users. The tweet-based social media bots pose serious security risksto Twitter users. These bots are used to spread fake contents, phishing links, and spams. Although they are not used as botsto launch DDOS attacks, they could be utilized as Commandand Control (C&C) infrastructure to coordinate DDOS attack [3], [4]. They are capable of interacting with human accountsto deceive the users and hijack their accounts. These bots arealso used as tools to launch large-scale manipulation campaigns to influence public opinions. According to a study [5],52% of online traffic is generated by botnets, and the rest isproduced by actual users. It is also worthy to note that some bots are found with over 350,000 fake followers. To dealwith the above issues, there is a need to develop detectionsystems that can accurately distinguish between Twitter botaccounts and human accounts. Twitter data represent one of the examples of big data as around 500 million tweets aregenerated every day, i.e., 6,000 tweets every second.

Big data analytics has been widely used in different fields[7]_[11] to process large amount of data, discover hidden patterns, and find correlations among data points.

Artificialintelligence techniques are increasingly leveraged by big dataanalysis. In particular, shallow (conventional) and deep learning techniques have received considerable attention from the academia and industry due to their success in dealingwith heterogeneous and complex data, automatic learningof models, revealing unseen patterns, identifying dependencies,and getting insights from analyzing data.

Artificial intelligence has been extensively used by Twitterto determine tweet recommendations

for users. In fact, deepneural networks are applied on Twitter data to determine the relevant content for users, and hence improve their experience on the platform [12]. Artificial intelligence has played an important role in fighting inappropriate content. In 2017,about 300,000 accounts were suspended and identified with the help of artificial intelligence tools rather than humans.

This review aims at providing an overview of differenttweet-based bot detection methods that use shallow and deep learning techniques to distinguish between human accountsand bot accounts. In particular, the main contributions of thepaper are the following: 1) A taxonomy, which classifies the state-of-the-art onmachine learning techniques for tweet-based botdetection,is presented.2) A comprehensive review is presented on shallow anddeep learning techniques for tweet-based bot detection, which covers the solutions up to year 2020.3) The challenges and open issues related to tweet-based bot detection ar highlighted and discussed.The rest of the paper is structured as follows: Section IIdiscusses the related surveys on tweet-based bot detectiontechniques. Section III presents the state-of-the-art related to deep and shallow learning based detection methods, followedby a discussion and analysis in Section IV.

II. LITERATURE SURVEY

In the literature, there exist some previous surveys that discuss and review existing papers published on social bot and spam detection, similar to this work. However, each one has its own limitations and strengths. Therefore, in this section, we briefly describe each survey and summarize it in Table 1. Kabakus and Kara [13] provided a short comparative survey of the research work in the field of Twitter spam detection within the year range of 2009-2015. They described different detection methods within four categories: accountbased, tweet-based, graph- based, and hybrid-based methods. The account-based methods were shown to leverage the user profile's metadata like followers and following count and other derived features such as age of the account. While in graph-based methods, features like distance and strength of connectivity between users were shown to be used for spam detection. However, in tweet-based methods, the survey mainly focused on detecting spam using URL and its derived features, such as length and domain name. To detect a spam user, posted URLs were analyzed and classified as malicious or benign. Besides this, the authors highlighted overlooked features that were argued to improve the spam detection. Another comparative survey was presented by Chakraborty et al. [14] in the field of multiplatform spam user detection. The authors recognized that different platforms, such as e-mails, blogs, or microblogs, require different techniques and features to achieve accurate detection. Therefore, proposed techniques within the year range of 2011-2015 were classified based on the platform that the dataset lies within. A qualitative comparison was conducted for each group of methods under the same platform. Besel et al. [21] observed that the botnet used a URL network shortening services and redirections to obfuscate the actual landing pages. They disclosed that users clicked on these URLs, found the botmaster establishing the Bursty botnet, and registering landing pages on phishing websites. They confirmed that the botmaster is still successful in owning Twitter bot-related services. This study includes a review and insight into Twitter's cyberspace infrastructure, cybercrime operation, and the dark markets. Alothali et al.

[15] summarized recent research work in the field of Twitter social botnet detection. They provided an analytical review of each proposed method with its limitations and advantages. The techniques were classified into three main categories, namely graph-based, machine learning- based, and crowdsourcing based techniques. The crowdsourcing technique uses human intelligence to identify various patterns, which is stated to be the most error prone out of the three techniques. It was also shown that machine learning methods and, more specifically, random forest classifiers are the most commonly used for detecting social bots in Twitter users.

Later presented a comprehensive review focusing on malicious social bots' stealthy manner and

their detection techniques. The author precisely reviewed detection approaches, which are graph-based, machine learning based, and emerging approaches. Besides, the paper reviewed the strengths and weaknesses of these techniques and the means considered by the bots to avoid detection. Consequently, the paper suggested approaches that may enhance the defense procedures against malicious bots. One of the challenges faced in evaluating bot detection approaches is that the ground-truth data is insufficient [17]. Detection techniques were compared with different aspects such as several features, the dataset's size, and the data- crawling operation. The datasets were categorized into synthesized data, crawled from online social networks, and gathered from honey profiles that attract social bots. A detailed review of existing datasets used by researchers was studied along with the results and experimental findings. In the end, the paper highlighted the constraints of the detection approaches and proposed some directions for future work. One of the suggestions was to concentrate on detection methodologies for general purposes. Also, it was suggested to build datasets that have different sets of social bots in order to assist in the generalized evaluation of the detection techniques [17]. Guo et al. [18] presented a survey on Online Social Deception (OSD). OSD is a serious threat in cyberspace, especially for users that are vulnerable to such cyberattacks. Cybercriminals have exploited social network services (SNSs) to conduct risky OSD activities, such as financial fraud, data threats, or sexual/labour violence. Therefore, OSD identifies and implements proactive responses to build credible OSD SNSs. It provided a comprehensive survey of social deceit's multidisciplinary concept focused on various OSD attacks and OSD attack types. Researchers have recently offered several innovative approaches that have vastly increased the efficiency of spam identification. It also offers an opportunity to perform a thorough analysis on Twitter of numerous spam identification methods. Abkenar et al. [19] focused on extensively evaluating the current Twitter spam identification testing techniques. Analysis of the literature review shows that most current approaches depend on algorithms that concentrate on machine learning. Among these algorithms for machine learning, the major differences relate to separate methods of collection of features. Therefore, they suggest a taxonomy focused on multiple approaches and evaluations of functionality collection, namely material analysis, user analysis, tweet analysis, network analysis, and hybrid analysis. Daffa et al. [20] discussed the identification of spam URLs in Twitter by presenting the types of harmful activities, detection avoidance strategies, detection function detection techniques, and their limitations. Via machine learning classification based on different published characteristics, they demonstrated the best results. They used four classifiers on a 10713 consumer dataset of Twitter accounts with 5358 labeled as benign and 5355 labeled as spam along with 17 stable features. The features were content-based and user-based features. The outcome revealed that of the four classifiers, the Random Forest classifier with hybrid feature methods achieved the best estimation with 96.4 percent accuracy. In comparison, J48 classifier obtained 94.5 percent accuracy score. Differently from the above surveys, our review focuses on techniques that employ shallow and deep learning methods for the detection of tweet-based social bots.

Table: 2.1 summery of Exiting surveys

Tweet-Based Bot Detection

Although the detection of social bots is a challenging task, there are some works that analyzed the characteristics and behavior of bots [14], [15], [22] and offered various features that are recurrent in the majority of works. For

example, verified accounts are guaranteed to be human users. Moreover, the ratio of followers to following and the age of the account are considered discriminative characteristics in detecting bots since bots generally mass-follow and have short life span [16]. The following features are mainly used by tweet-based bot detection techniques to distinguish between tweet-based bots and humans accounts [23]: • ID: It represents the unique identifier of the tweet. • User: It represents the user who posted the tweet. • Created_at: It indicates the UTC time when the tweet is created. • Text Tweet: It refers to the body of the tweet. • Length of Tweet: It gives the number of characters in the tweet. • #Hashtags: It indicates the number of hashtags in the tweet. • #URLs: It indicates the number of URLs in the tweet • in reply to status id: If the tweet is a reply, this feature represents the original tweet's ID. • in reply to user id: If the tweet is a reply, this feature represents the author of the original tweet. • Coordinates: It represents the geographic location of the tweet. • Favorite Count: It indicates how many times the tweet has been liked by Twitter users. • Retweet Count: It is the number of times the tweet has been retweeted • Reply Count: It is the number of times the tweet has been replied to. • Favorited : a boolean feature, which holds true when the tweet is liked by the authenticating user. • Retweeted: a boolean feature, which holds true when the tweet is retweeted by the authenticating user. • Possibly_sensitive: a boolean feature, which holds true when the tweet contains a link. A. TAXONOMY In this section, we describe machine learning techniques used for tweet-based bot detection. As shown in Fig. 1, the stateof-the-art techniques are classified into two major categories: shallow learning- based detection and deep learning-based detection. According to the learning approach, the shallow detection techniques are further classified into three subcategories: supervised learning, semi-supervised learning, and unsupervised learning. In supervised learning, the learning model is trained with labeled data, so it can predict the output of the new data. An unsupervised learning technique builds the model from unlabelled data. It aims to find structures and patterns within the data itself. The semi-supervised learning techniques use both labeled and unlabeled data to train the model. On the other hand, deep learning-based detection techniques are further classified into two subcategories: generative architecture based techniques and discriminative architecture based techniques. If we have input data x and we want to classify them into labels y, a generative model learns the joint probability distribution p(x, y). On the other hand, the discriminative model learns the conditional probability distribution p(y/x). The deep generative architecture is formed by combining a generative model and a deep neural network. It is generally associated with unsupervised learning. The deep discriminative architecture adopts supervised learning, and is built by combining a discriminative model and a deep neural network to compute and optimize p(y/x). The detailed discussion on each category is given in the rest of the section. B. DEEP LEARNING- BASED DETECTION METHODS Recently, deep neural networks have gained noticeable attention from researchers in different fields ranging from computer vision to language processing. It has proven its effectiveness in terms of textual classification. It can process structured data like sentences and automatically produce discriminant features, thus relinquishing handcrafting features, which is expensive and requires extensive knowledge of the data. Therefore, since the overall performance of a classifier relies heavily on the quality of its data, deep neural networks such as Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) were employed as a feature extractor or classifier for many language processing problems, one of which is the tweet-based bot detection. However, neural networks require a certain form of input, preferably structured data, but most importantly, a numerical vector representing the data. There are several pre-trained word embedding models for that purpose, such as the popular Word2Vec model. Therefore, as a preliminary step, all text tweets are converted into a form accepted by the network using trained models. The Long Short-Term Memory (LSTM) model is the most popular model for language processing and

classification. It is an improved version of the RNN vanilla model that can maintain a memory of the past input for a longer period, preferable for long text input. Hence, most work mentioned in this section employs a variation of LSTM. For example, Kudugunta and Ferrara [24] recognized the limitation of utilizing either tweet metadata or tweet text as a single input. Therefore, a Contextual LSTM was proposed that takes both features for improved bot detection. They used the public dataset Cresci-2017 to train the model to reduce the exhibited imbalance. The Synthetic Minority Oversampling Technique (SMOTE) was used to fill the minority class without fully synthetic data that might affect the performance. Training data of 8,386 users' tweets were tokenized and loaded into the Glove word embedding model and fed to the model for feature extraction. Besides, tweet metadata such as retweet and reply count were concatenated with the tweet text's features before classifying in the dense layer. This yielded better performance than using tweet metadata only. To prove the strategy's superiority, the model was tested using single and combined features resulting in 96% for both precision and accuracy favoring the proposed method. Wei and Nguyen [25] proposed a bidirectional RNN to identify bot accounts in Twitter by utilizing the LSTMmodel.

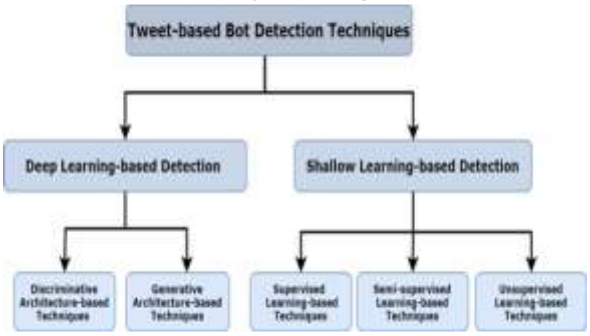

Fig :1. Lstm

Table: 2.2 summary of deep learning based detection methods

TABLE 1. Summary of deep learning based detection methods.

Ref	Dataset	Tweets	Training Set	Testing Set	Pre-processing	Features	Classifier	Architecture/ Approach	Accuracy	FP	TP	Precision
[24]	Cresci-2017	11.4 m	8,386	N/A	GloVE, SMOTE	Tweet & account Metadata	LSTM	Discriminative /Supervised	96%	N/A	N/A	96%
[25]	Cresci-2017	11.4 m	[illegible]	[illegible]	GloVe	Tweet Text&Tweet & Account Metadata	BiLSTM	Discriminative /Unsupervised	97%	[illegible]	94%	[illegible]
[26]	Own	[illegible]	[illegible]	[illegible]	Word embedding	Tweet Metadata	Autoencoder LSTM	Generative unsupervised	[illegible]	N/A	N/A	[illegible]
[27]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Tweet Text & metadata	CNN + LSTM	Discriminative/ Supervised	N/A	N/A	N/A	[illegible]
[28]	CLEF-2019	N/A	[illegible]	[illegible]	Word embedding	Tweet Text & metadata	CNN	Discriminative/ Unsupervised	[illegible]	N/A	N/A	[illegible]
[29]	Twitter	[illegible]	[illegible]	[illegible]	Spam detection	IDs, screen name, location	Bayesian classification	Supervised	N/A	N/A	N/A	[illegible]
[30]	[illegible]	[illegible]	N/A	N/A	Social bot detection	Tweet-based, user profile-based and social graph based attributes	Deep Q-Learning (DQL)	Unsupervised	[illegible]	N/A	N/A	[illegible]
[31]	[illegible]	[illegible]	N/A	N/A	N/A	[illegible]	[illegible]	Unsupervised	[illegible]	N/A	N/A	N/A
[32]	Bots and Gender Profiling 2019	[illegible]	[illegible]	[illegible]	[illegible]	BiLSTM	[illegible]	Unsupervised	[illegible]	N/A	N/A	N/A
[28]	CLEF 2019	[illegible]	[illegible]	[illegible]	[illegible]	N/A	Convolutional neural network	Unsupervised	N/A	N/A	N/A	[illegible]
[33]	[illegible] dataset	[illegible]	N/A	N/A	N/A	N/A	Deep neural network	Discriminative/ Unsupervised and Semi-supervised	[illegible]	N/A	N/A	[illegible]

III. MODULES

3.1.1 SERVICE PROVIDER

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Train and Test Data Sets,View Trained and Tested Accuracy in Bar Chart,View Trained and Tested Accuracy Results,View Predicted Tweet Type Details,Find Tweet Type Ratio on Data Sets, Download Trained Data Sets, View Tweet Type Ratio Results, View All Remote Users..

3.1.2 View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

3.1.3 Remote User

Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like POST TWEETS DATA SETS, PREDICT TWEET TYPE, VIEW YOUR PROFILE.

IV. OUTPUT SCREENS

Home page

Fig.2. Home page

Admin pages

Fig.8.2 Service provider login page

Service provider home page

Fig .3 Service provider home page

List of Algorithms

Fig .4 List of Algorithms

Accuracy of Algorithms

Fig .5 Accuracy of Algorithms

Line chat og Algorithms

Fig .6 Line chat og Algorithms

Pie chat of Algorithms

Fig .7 Pie chat of Algorithms

Prediction details

Fig .8 Prediction details

Ratio Details

Fig .9 Ratio Details

Ratio in line chart

Fig .10 Ratio in line chart

V. CONCLUSION

Twitter is one of the most popular social media platforms that allows connecting people and helps organizations reaching out to customers. Tweet-based botnet can compromise Twitter and create malicious accounts to launch large-scale attacks and manipulation campaigns. In this review, we have focused on big data analytics, especially shallow and deep learning to fight against tweet-based botnets, and to accurately distinguish between human accounts and tweet-based bot accounts. We have discussed related surveys, and have also provided a taxonomy that classifies the state- of-the-art tweet-based bot detection techniques up to 2020. In addition, the shallow and deep learning techniques are described for tweet-based botdetection, along with their performance results. Finally, we presented and discussed the open issues and future research challenges.

FUTURE SCOPE

The future scope of tweet-based bot detection using big data analytics involves leveraging advanced machine learning algorithms, natural language processing (NLP), and deep learning techniques to enhance accuracy. Additionally, integrating real-time data analysis, anomaly detection, and behavioural modelling will be crucial to stay ahead of evolving bot strategies. Collaborative efforts among researchers, industry, and social media platforms will likely play a pivotal role in developing robust and adaptive bot detection systems. As the sophistication of bots increases, continuous refinement of algorithms and the incorporation of ethical considerations in bot detection will be essential for maintaining the integrity of online communication

REFERENCES

[1] M. Mohsin. (2020). 10 Social Media Statistics You Need to Know in 2021. [Online]. Available: https://www.oberlo.com/blog/social-mediamarketing- statistics.

[2] I. Arghire. (2020). Twitter Hack: 24 Hours From Phishing Employ- ees to Hijacking Accounts. https://www.securityweek.com/twitter-hack- 24-hours-phishing-employees-hijacking-accounts

[3] The Rise of Social Media Botnets. Accessed: Feb. 21, 2021. [Online]. Available: https://www.darkreading.com/attacks-breaches/the-rise-ofsocial- media-botnets/a/d-id/1321177

[4] M. Imran, M. H. Durad, F. A. Khan, and A. Derhab, ``Toward an optimal solution against denial of service attacks in software de_ned networks," Future Gener. Comput. Syst., vol. 92, pp. 444_453, Mar. 2019.

[5] M. S. Savell. (2018). Protect Your Company's Reputation From Threats by Social Bots. [Online]. Available: https://zignallabs.com/blog/protect-yourcompanys- reputation-from-threats-by-social-bots/

[6] S. Aslam. (2021). Twitter by the Numbers: Stats, Demographics & Fun Facts. [Online]. Available: https://www.omnicoreagency.com/twitterstatistics/

[7] A. Aldweesh, A. Derhab, and A. Z. Emam, ``Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues," Knowl.-Based Syst., vol. 189, Feb. 2020, Art. no. 105124.

[8] S. Mahdavifar and A. A. Ghorbani, ``Application of deep learning to cybersecurity: A survey," Neurocomputing, vol. 347, pp. 149_176, Jun. 2019.

[9] E. B. Karbab, M. Debbabi, A. Derhab, and D. Mouheb, ``MalDozer: Automatic framework for Android malware detection using deep learning," Digit. Invest., vol. 24, pp. S48_S59, Mar. 2018.

[10] F. A. Khan, A. Gumaei, A. Derhab, and A. Hussain, ``A novel twostage deep learning model for ef_cient network intrusion detection," IEEE Access, vol. 7, pp. 30373_30385, 2019.

[11] A. Derhab, A. Aldweesh, A. Z. Emam, and F. A. Khan, ``Intrusion detection system for Internet of Things based on temporal convolution neural network and ef_cient feature engineering," Wireless Commun. Mobile Comput., vol. 2020, pp. 1_16, Dec. 2020.
[12] B. Marr. (2020). How Twitter Uses Big Data and Arti_cial Intelligence (AI). [Online]. Available: https://www.bernardmarr.com/default.asp?contentID=1373
[13] A. T. Kabakus and R. Kara, ``A survey of spam detection methods on Twitter," Int. J. Adv. Comput. Sci. Appl., vol. 8, no. 3, pp. 29_38, 2017.
[14] M. Chakraborty, S. Pal, R. Pramanik, and C. R. Chowdary, ``Recent developments in social spam detection and combating techniques: A survey,"
Inf. Process. Manage., vol. 52, no. 6, pp. 1053_1073, Nov. 2016.
[15] E. Alothali, N. Zaki, E. A. Mohamed, and H. Alashwal, ``Detecting social bots on Twitter: A literature review," in Proc. Int. Conf. Innov. Inf. Technol. (IIT), Nov. 2018, pp. 175_180.