International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
OPEN ACCESS
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST52411222

Malicious Social Bot Using Twitter Network Analysis in Django

[1]Ms. N. Ezhil Arasi, [2]Dr G Manikandan, [3]Ms. S. Hemalatha, [4]Ms. Vilma Veronica

[1]PG Student, Kings Engineering College, Sriperumbudhur, Tamil Nadu, India
[2]Professor, Kings Engineering College, Sriperumbudhur, Tamil Nadu, India, India
[3]Assistant Professor, Kings Engineering College, Sriperumbudhur, Tamil Nadu, India
[4]Assistant Professor, Kings Engineering College, Sriperumbudhur, Tamil Nadu, India

ARTICLEINFO

Article History:

Accepted: 03 March 2024
Published: 13 March 2024

Publication Issue :
Volume 11, Issue 2
March-April-2024
Page Number :
114-122

ABSTRACT

Malicious social bots generate fake tweets and automate their social relationships either by pretending to be a followers or by creating multiple fake accounts with malicious activities. Moreover, malicious social bots post shortened malicious URLs in the tweets to redirect the requests of online social networking participants to some malicious servers. Hence, distinguishing malicious social bots from legitimate users is one of the most important tasks in the Twitter network. To detect malicious social bots, extracting URL-based features (such as URL redirection, frequency of shared URLs, and spam content in URL) consumes less amount of time in comparison with social graph-based features (which rely on the social interactions of users). Furthermore, malicious social bots cannot easily manipulate URL redirection chains. In this article, learning automata-based malicious social bot detection (LA-MSBD) algorithm is proposed by integrating a trust computation model with URL-based features for identifying trustworthy participants (users) in the Twitter network. The proposed trust computation model contains two parameters, namely, direct trust and indirect trust. Moreover, the direct trust is derived from Bayes' theorem, and the indirect trust is derived from the Dempster– Shafer theory (DST) to determine the trustworthiness of each participant accurately. Finally, we shown the user tweet data in terms of graph visualization of bar chart and pie chart of the system. Experimental results shown the better performance of the system.

INDEX TERMS Agent-based modelling, agent-based social simulation, multi-agent systems, social media, twitter, twitter bot.

I. INTRODUCTION

Malicious social bot is a software program that pretends to be a real user in online social network. Moreover, malicious social bots perform several malicious attacks, such as spread social spam content, generate fake identities, manipulate online ratings, and perform phishing attacks .In Twitter, when a participant (user) wants to share a tweet containing URL(s) with the neighbouring participants (i.e., followers or followers), the participant adapts URL shortened service in order to reduce the length of URL (because a tweet is restricted up to 140 characters). Moreover, a malicious social bot may post shortened phishing URLs in the tweet .They generate fake tweets and automate their social relationships either by pretending like a follower or by creating multiple fake accounts with malicious activities. Moreover, malicious social bots post shortened malicious URLs in the tweet in order to redirect the requests of online social networking participants to some malicious servers. Hence, distinguishing malicious social bots from legitimate users is one of the most important tasks in the Twitter network. When a participant clicks on a shortened phishing URL, the participant's request will be redirected to intermediate URLs associated with malicious servers that, in turn, redirect the user to malicious web pages. Then, the legitimate participant is exposed to an attacker. Malicious social bots generate fake tweets and automate their social relationships either by pretending like a follower or by creating multiple fake accounts with malicious activities. Moreover, malicious social bots post shortened malicious URLs in the tweet in order to redirect the requests of online social networking participants to some malicious servers. Hence, distinguishing malicious social bots from legitimate users is one of the most important tasks in the Twitter network.

This leads to twitter network suffering from several vulnerabilities (such as phishing attack). Several approaches have been proposed to detect spam in the Twitter network. These approaches are based on tweet-content features, social relationship features, and user profile features. However, the malicious social bots can manipulate profile features, such as hash tag ratio, follower ratio, URL ratio, and the number of retweets. The malicious social bots can also manipulate tweet-content features, such as sentimental words, emoticons, and most frequent words used in the tweets, by manipulating the content of each tweet. The social relationship- based features are highly robust because the malicious social bots cannot easily manipulate the social interactions of users in the Twitter network. However, extracting social relationshipbased features consumes a huge amount of time due to the massive volume of social network graph. Therefore, identifying the malicious social bots from the legitimate participants is a challenging task in the Twitter network. The existing malicious URL detection approaches are based on DNS information and lexical properties of URLs. The malicious social bots use URL redirections in order to avoid detection. However, for detectors, identification of all malicious social bots is an issue because malicious social bots do not post malicious URLs directly in the tweets. Thus, it is important to identify malicious URLs (i.e., harmful URLs) posted by malicious social bots in Twitter.

Most of the existing approaches are based on supervised learning algorithms, where the model is trained with the labelled data in order to detect malicious bots in OSNs. However, these approaches rely on statistical features instead of analysing the social behaviour of user. More- over, these approaches are not highly robust in detecting the temporal data patterns with noisy data (i.e., where the data is biased with untrustworthy or fake information) because the behaviour of malicious bots changes over time in order to avoid detection. This motivated us to consider one of the reinforcement learning techniques (such as the learning automata (LA) model) to handle

temporal data patterns. In this work, we design a Django based model to detect malicious social bots based on user tweet data of the system. Here, the malicious behaviour of participants is analysed by considering features extracted from the posted URLs (in the tweets), such as URL redirection, frequency of shared URLs, and spam content in URL, to distinguish between legitimate and malicious tweets. The proposed trust computational model contains two parameters, namely, direct trust and indirect trust. The direct trust value is derived from the Bayesian learning (by considering URLbased features) to determine the trustworthiness of tweets posted by each participant. In addition to the direct trust, belief values (i.e., indicators for determining indirect trust) are collected from multiple neighbours of a participant. This is due to the fact that in case the neighbors of a participant are trustworthy, the participant is likely to be trustworthy. Furthermore, Dempster's combination rule aggregates the belief values provided by multiple one-hop neighbouring participants in order to evaluate the indirect trust value of participants in the Twitter network.

In our daily lives, social media has become increasingly crucial. People naturally flock to this medium to read and s hare news, given that billions of users produce and consume information every day. Social media bots are little programmes that can be deployed on social media platforms to perform a variety of useful and destructive functions while encouraging human behaviour. Some social media bots provide helpful services like weather and sports scores. These excellent social media bots are clearly labelled as such, and those who connect with them are aware that they are bots. A huge majority of social media bots, on the other hand, are harmful bots masquerading as human users. Users lose faith in social media platforms' ability to offer accurate news as a result of these bots, since they suspect that the stories at the top of their feeds were "pushed" there by manipulative bots. Because so many individuals are using social media, malevolent users such as bots have begun to manipulate conversations in the direction that their makers desire. These malicious bots have been used for nefarious purposes such as spreading false information about political candidates, inflating celebrities' perceived popularity, deliberately suppressing protestors' and activists' messages, illegally advertising by spamming social media with links to commercial websites, and influencing financial markets in an attempt to manipulate stock prices. Furthermore, these bots have the ability to alter the outcomes of standard social media analysis. Social media bots use a variety of attack strategies, including: Sleeper bots are bots that sleep for lengthy periods of time before waking up to unleash an attack of thousands of postings in a short period of time (perhaps as a spam attack), and then sleep again. Jacking the trend - the use of top trending topics to focus on a certain audience for the purpose of targeting, an attacker employs a watering hole assault to estimate or watch which websites a company frequently visits and infects one or more of them with malware. Click farming or like farming-inflate fame or popularity on a website by like or reposting content via click farms, and hash tag hijacking- use of hash tags to focus an assault (e.g. spam, harmful links) on a specific audience using the same hash tag. In social media, bot detection is a critical duty. Automated accounts are a problem on Twitter, a popular social networking platform. According to certain surveys, roughly 15% of Twitter accounts operate automatically or semi automatically. The peculiarities of Twitter could be one factor that has contributed to the rise in bots. It's also worth noting that a Twitter bot is recognised as a reliable source of information. Although social networking sites have improved our social life, there are still some drawbacks. In online social networks, malicious social bots are a widespread problem. These malevolent social bots are being utilised for a variety of things, including artificially inflating a person's or movement's popularity, influencing elections,

manipulating financial markets, amplifying phishing attempts, spreading spam, and suppressing free expression. As a result, detecting these bots in online social networks is critical. Nefarious social bots create phoney tweets and automate their social relationships by impersonating a follower or creating many fake accounts that are used for malicious purposes. Malicious social bots broadcast shortened malicious URLs in tweets in order to reroute online social networking users' requests to malicious sites.

II. EXISTING SYSTEM

Bots have made an impact on a variety of social media platforms. Twitter has been hit particularly hard, with bots accounting for a sizable amount of its user base. These bots have been used for nefarious purposes like distributing fake information about politicians and increasing celebrities' perceived popularity. These bots have the ability to alter the outcomes of standard social media analysis. Malicious social bots have also been employed to spread incorrect information (for example, emailing fraudulent urls), which can have real-world effects. To detect such hostile behaviours, the suggested systems employ machine learning methods such as Naive Bayes and RF.

III.PROPOSED SYSTEM

The proposed framework consists of three components: data collection, feature extraction, and LA model. To collect tweets posted by participants (users), the tweets can be crawled using Twitter Streaming APIs. The data collection component consists of three subcomponents reading tweets from Twitter streaming, collecting tweets, and URLs. Moreover, the collected tweets and collected URLs are stored in a repository. The feature extraction consists of two sub components: expanding shortened URLs and extracting feature set. Whenever a feature extraction component obtains a shortened URL from the repository, it is converted into a long URL using URL shortened services. For each URL, we extract several features that are based on the lexical properties of URLs along with the features of URL redirection. Moreover, the trust model determines the probability of a tweet containing any malicious information. Finally, after evaluating the malicious behavior of a series of tweets posted by a participant, we classify tweets as malicious and legitimate tweets. However, malicious tweets are likely to be posted by malicious social bots. This helps in distinguishing malicious social bots from benign participants.

IV. Literature Review

[1] Lingam, G et al., (2020). “Particle swarm optimization on deep reinforcement learning for detecting social spam bots and spam-influential users in twitter network”. In online social networks (OSNs), detection of malicious social bots is an important research challenge to provide legitimacy of user profiles and trustworthy service ratings. Further, spam-influential users must be minimal to control the fake information-spread in OSNs. Learning from example patterns using supervised learning may not provide accurate results in cases where existing data items are biased and bot behaviour dynamically changes over a period of time. Moreover, deep reinforcement learning provides improved learning by repeated interactions with the environment. However, a typical deep reinforcement leaning algorithm converges slower to find an optimal sequence of actions to reach out a goal state. In this article, we design a particle swarm optimization (PSO) based deep Q-learning algorithm for detecting social spam bots by integrating PSO with Q-value function. In addition, spam-influential users are identified using the proposed spam influence minimization model and it helps in restricting the flow of illegitimate tweets in Twitter network. Further, an influential community detection algorithm has been proposed to reduce the spreading of spam content through influential communities in Twitter network. Experimental

results illustrate the efficacy of our proposed algorithms by considering two Twitter datasets and performance metrics such as precision, recall, and modularity.

[2] Rout, R. R et al., (2020, October). "Social botnet community detection: a novel approach based on behavioural similarity in twitter network using deep learning". Detecting social bots and identifying social botnet communities are extremely important in online social networks (OSNs). In this paper, we first construct a weighted signed Twitter network graph based on the behavioural similarity and trust values between the participants (i.e., OSN accounts) as weighted edges. The behavioural similarity is analysed from the viewpoints of tweet-content similarity, shared URL similarity, interest similarity, and social interaction similarity for identifying similar types of behaviour (malicious or not) among the participants in the Twitter network; whereas the participant's trust value is determined by a random walk model. Next, we design two algorithms - Social Botnet Community Detection (SBCD) and Deep Auto encoder based SBCD (called DA-SBCD) - where the former detects social botnet communities of social bots with malicious behavioural similarity, while the latter reconstructs and detects social botnet communities more accurately in presence of different types of malicious activities. Finally, we evaluate the performance of proposed algorithms with the help of two Twitter datasets. Experimental results demonstrate the efficacy of our algorithms with better performance than existing schemes in terms of normalized mutual information (NMI), precision, recall and F-measure. More precisely, the DA-SBCD algorithm achieves about 90% precision and exhibits up to 8% improvement on NMI.

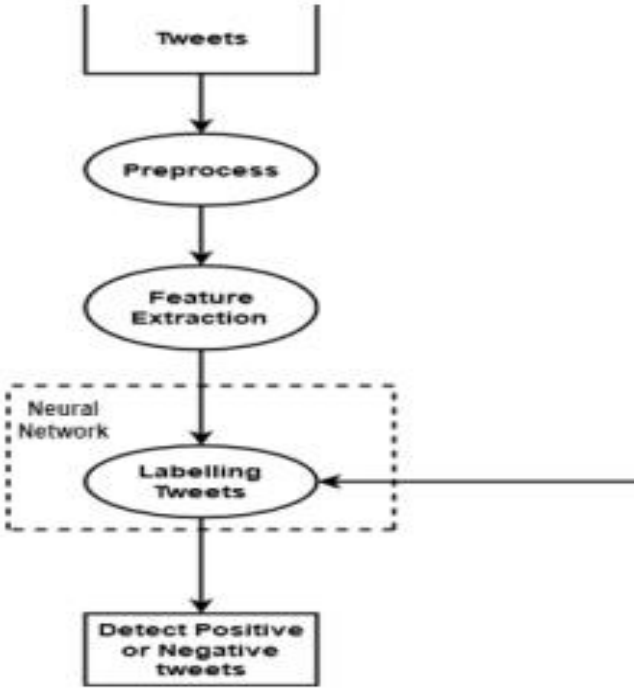
[3] Guo, Q et al., (2021). "Social bots detection via fusing Bert and graph convolutional networks". The online social media ecosystem is becoming more and more confused because of more and more fake information and the social media of malicious users' fake content; at the same time, unspeakable pain has been brought to mankind. Social robot detection uses supervised classification based on artificial feature extraction. However, user privacy is also involved in using these methods, and the hidden feature information is also ignored, such as semi-supervised algorithms with low utilization rates and graph features. In this work, we symmetrically combine BERT and GCN (Graph Convolutional Network, GCN) and propose a novel model that combines large scale pretraining and transductive learning for social robot detection, BGSRD. BGSRD constructs a heterogeneous graph over the dataset and represents Twitter as nodes using BERT representations. Corpus learning via text graph convolution network is a single text graph, which is mainly built for corpus-based on word co-occurrence and document word relationship. BERT and GCN modules can be jointly trained in BGSRD to achieve the best of merit, training data and unlabelled test data can spread label influence through graph convolution and can be carried out in the large-scale pre-training of massive raw data and the transduction learning of joint learning representation. The experiment shows that a better performance can also be achieved by BGSRD on a wide range of social robot detection datasets.

[4] Heidari, M et al., (2022). "Online user profiling to detect social bots on twitter". Social media platforms can expose influential trends in many aspects of everyday life. However, the movements they represent can be contaminated by disinformation. Social bots are one of the significant sources of disinformation in social media. Social bots can pose serious cyber threats to society and public opinion. This research aims to develop machine learning models to detect bots based on the extracted user's profile from a Tweet's text. Online users' profile shows the user's personal information, such as age, gender, education, and personality. In this work, the user's profile is constructed based on the user's online

posts. This work's main contribution is three-fold: First, we aim to improve bot detection through machine learning models based on the user's personal information generated by the user's online comments. When comparing two online posts, the similarity of personal information makes it difficult to differentiate a bot from a human user. However, this research turns personal information similarity among two online posts into an advantage for the new bot detection model. The new proposed model for bot detection creates user profiles based on personal information such as age, personality, gender, education from users' online posts and introduces a machine learning model to detect social bots with high prediction accuracy based on personal information. Second, create a new public data set that shows the user's profile for more than 6900 Twitter accounts in the Cresci 2017 data set.

[5] Pham, P et al., (2022). "Bot2Vec: A general approach of intra-community oriented representation learning for bot detection in different types of social networks". Recently, due to the rapid growth of online social networks (OSNs) such as Facebook, Twitter, Weibo, etc. the number of machine accounts/social bots that mimic human users has increased. Along with the development of artificial intelligence (AI), social bots are designed to become smarter and more sophisticated in their efforts at replicating the normal behaviors of human accounts. Constructing reliable and effective bot detection mechanisms is this considered crucial to keep OSNs clean and safe for users. Despite the rapid development of social bot detection platforms, recent state-of-the-art systems still encounter challenges which are related to the model's generalization (and whether it can be adaptable for multiple types of OSNs) as well as the great efforts needed for feature engineering. In this paper, we propose a novel approach of applying network representation learning (NRL) to bot/spammer detection, called Bot2Vec. Our proposed Bot2Vec model is designed to automatically preserve both local neighbourhood relations and the intra-community structure of user nodes while learning the representation of given OSNs, without using any extra features based on the user's profile. By applying the intra-community random walk strategy, Bot2Vec promises to achieve better user node embedding outputs than recent state-of-the-art network embedding baselines for bot detection tasks. Extensive experiments on two different types of real-word social networks (Twitter and Tagged) demonstrate the effectiveness of our proposed model.

V. SYSTEM ARCHITECTURE



Literature Review

Research by Ikeda et al. also use as a base the SIR model to generate an Agent-Based Information Diffusion Model to evaluate data from the previously mentioned 2011 East Japan earthquake. To make the model more robust, they introduced the idea of diversity and multiplexing of infor- mation paths [10]. Similarly, research conducted by Okada et al. also provide a more robust version of the SIR model, with

data from the same event [11]. The results showed by the latter were compared with real Twitter data to evaluate the susceptible, infected and recovered users using metrics such as number of retweets. Similar to other related works, the rumor propagation is the focus, rather than the origin of that false information, which is something we address in our research.

Kundu et al. approach the information propagation model in their work by developing a novel fuzzy relative willingness model. The diffusion model was able to successfully utilize the external influence factor, as well as the susceptibility of individual nodes to quantify human willingness [12]. While the objective of the paper differs from our research, the implementation of the external influence factor contributed to our research as we applied a similar function as well.

Ross et al. validate the concept of the spiral of silence in their research paper [13] through an agent-based model. The spiral of silence explains how the influence of surrounding negative opinions can affect the spread of positive opinions in a social media environment. The way the simulation was built differs from ours, but it provided insight into how users tend to react while consuming negative media. Wang et al. study information entropy, which incorporates new types of variables to the simulation, including the degrees of trusts agents can set between each other [14]. This work, similar to others, centers on rumor spreading models on social media, but focuses on the interaction between two given agents by adding weights between each node to represent trust.

Research by Yan et al. dives deep into the concept of how *retweets* and *quote-tweets* influence the behavior of users in social networks by using the concepts of game theory and developing a reward mechanism [15]. The objective of the research is to study agent behavior given goals, such as a higher number of retweets. This work does not center on rumor spreading or malicious agents, but it proved to be insightful in providing more information about the retweet cycle, which is the means by which information spreads on social media sites, such as Twitter.

Research highlighting bot behavior includes work by Carley [16] outlining the BEND framework as a way of iden- tifying misinformation maneuvers in social media environ- ments. Other studies related to bot behavior can be reviewed in Cresci et al. paper [17] where they compare human behav- ior in social media similar to DNA sequences and identify bot behavior based on a predictable set of actions. They compare the actions of several Twitter human users and bot users, which they call their digital DNA, and try to identify similarities between them.

Later work by Beskow and Carley highlights two bot mis- information maneuvers: backing and bridging. The former focused on bots interacting with agents with high influence (larger number of edges) to spread false information, while the latter focuses on bridging two communities, with the bots

Datasets Processing

Data in the form of raw tweets is acquired by using the python library “twee stream” which provides a package for simple twitter streaming API. This API allows two modes of accessing tweets: Sample Stream and Filter Stream. Sample Stream simply delivers a small, random sample of all the tweets streaming at a real time. Filter Stream delivers tweet which match a certain criteria.

- It can filter the delivered tweets according to three criteria: Specific keyword(s) to track/search for in the tweets Specific Twitter user(s) according to their user-id’s
- Tweets originating from specific location(s) (only for geo-tagged tweets). A programmer can specify any single one of these filtering criteria or

a multiple combination of these. But for our purpose we have no such restriction and will thus stick to the Sample Stream mode.

Data Pre-processing:

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data preprocessing is required tasks for cleaning the data and making it suitable for a machine learning model and deep learning which also increases the accuracy and efficiency of a machine learning and deep learning model.

Retrieve words: Twitter is not just an extended source of news. The twitter data can be using the consumer key, consumer secret, access token, access token secret. Twitter allows the usage of their API via an oauth2 authorization framework

1. Feature Extraction:

The feature extraction technique plays an important role. The features are the main parameter that is involved for classification of User Tweet. Texture extraction is determined as the example of information or course of action of the structure with random interval.

2. Prediction module:

When the reasons behind a model's outcomes are as important as the outcomes themselves, Prediction Explanations can uncover the factors that most contribute to those outcomes. This shows the output of the algorithm from the testing case. Final Prediction shown the user tweet data based on different plot of graph shown in the Django Framework of the system.

VI. CONCLUSION

This article presents an LA-MSBD algorithm by integrating a trust computational model with a set of URL-based features for MSBD. In addition, we evaluate the trustworthiness of tweets (posted by each participant) by using the Bayesian learning and DST. Moreover, the proposed LA-MSBD algorithm executes a finite set of learning actions to update action probability value (i.e., probability of a participant posting malicious URLs in the tweets). The proposed LA-MSBD algorithm achieves the advantages of incremental learning. The need for new, low-cost Bot detection systems is evident given the frequency of detecting malicious bots on social media sites such as Twitter. We suggested a DL algorithm for detecting tweets or URLs that are potentially fraudulent or damaging to users. So far, we have downloaded and installed all of the software that is required for the planned system. The dataset was obtained from the Kaggle website, and the preparation stage was completed. The features of pre-processed data will be extracted in the next phase, and the method will be implemented, with a model saved that can be used to categories the data.

VII. REFERENCES

[1]. Lingam, G., Rout, R. R., Somayajulu, D. V., & Ghosh, S. K. (2020). Particle swarm optimization on deep reinforcement learning for detecting social spam bots and spam-influential users in twitter network. IEEE Systems Journal, 15(2), 2281-2292.

[2]. Lingam, G., Rout, R. R., Somayajulu, D. V., & Das, S. K. (2020, October). Social botnet community detection: a novel approach based on behavioral similarity in twitter network using deep learning. In Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (pp. 708-718).

[3]. Guo, Q., Xie, H., Li, Y., Ma, W., & Zhang, C. (2021). Social bots detection via fusing bert and graph convolutional networks. Symmetry, 14(1), 30.

[4]. Heidari, M., Jones Jr, J. H., & Uzuner, O. (2022). Online user profiling to detect social bots on twitter. arXiv preprint arXiv:2203.05966.

[5]. Pham, P., Nguyen, L. T., Vo, B., & Yun, U. (2022). Bot2Vec: A general approach of intra-community oriented representation learning for bot detection in different types of social networks. Information Systems, 103, 101771.

Cite this article as :

Ms. N. Ezhil Arasi, Dr G Manikandan, Ms. S. Hemalatha, Ms. Vilma Veronica, "Malicious Social Bot Using Twitter Network Analysis in Django", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 11 Issue 2, pp. 114-122, March-April 2024. Available at doi : https://doi.org/10.32628/IJSRST52411222
Journal URL : https://ijsrst.com/IJSRST52411222