Journal of Sensors, IoT & Health Sciences
ISSN: 2584-2560
Vol.02(01), March 2024, pp.6-16

Research Article

Enhanced Detection of Social Bots on Online Platforms Using Semi-Supervised K-Means Clustering

Rekha Gangula [1,*] and Rega Sravani [2]
[1]Assistant professor, Department of CSE, Vaagdevi Engineering College, Bollikunta, Warangal, Telangana,506005, India.
[2]Assistant professor, Department of CSE(DS), Vaagdevi Engineering College, Bollikunta, Warangal, Telangana,506005, India. Email: regasravani@gmail.com
[*]Corresponding Author: Rekha Gangula. Email: gangularekha@gmail.com
Received: 24/02/2024; Revised: 05/03/2024; Accepted: 22/03/2024; Published: 31/03/2024.
DOI: https://doi.org/10.69996/ jsihs.2024002

Abstract: Social bots are semi-automatic or automatic computer applications that express human performance in OSN. Social bots are the primary tools utilized by hackers to invade OSNs. The current use of Social bots in communication and voting operations has been highlighted. Twitter and Tumblr have been efficiently used to share information about public sentiment. The developing connection on the Internet has started up avenues for improved cybersecurity threats and perpetuation of an extensive array of cybercrimes occurring in significant financial needs and user data privacy violations. One of the most advanced but critical extensions to the public of malicious software is the bot malware, commonly referred to as botnets. The most current presentation techniques of malicious social bots examine the quantitative characteristics of their behavior. This paper proposed a novel approach to identifying malicious social bots, including feature determination based on the development probability of clickstream progressions and the Semi-supervised K-Means Clustering algorithm for detecting social bots. The proposed method explains the transition probability of user behavior clickstreams and reflects the time feature. The proposed Semi-supervised K-Means Clustering **(**SSKMC**)** algorithm, compared with the traditional detection method based on the quantitative function, improves accuracy by 15% on average. The proposed SSKMC Algorithm can efficiently detect malicious accounts on social bots.

Keywords: Online Social Networks, Botnet, Machine Learning, Social Bots, Semi-Supervised Clustering Algorithm

1.Introduction

Our world has been controlled by online social networks (OSN) like Facebook, Twitter, LinkedIn, etc. They play a key position in our lives as channels of public communication. They provide a platform for their clients to participate, and they have percentage interaction and statistics. Hence, they drive a great network with the value of attracting classified ads. Because of its precious OSN API recognition, it is also an attractive target for exploiting social networking software [1]. A social bot is a computer program for automating consumer activities. These activities can be

- Producing fake, apparently human-made posts to interact with people in a social community,
- Posting flyers and photographs or the reputation of others,
- Adding comments or liking posts,

- Building connections with other accounts.

Therefore, the level of complexity of bots is varied. Social bots can be fictional, like bots that collect data records, weather information, and blog posts and then repost them within the social community. On the other hand, it can also be very complicated, such as infiltration into human conversations. These skills have advantages and disadvantages for OSN users and can be used for harmful or desired intent [2].

- For one reason, bots can be designed with the right intentions. As quoted in the relevant post, it can protect anonymity and automate and execute tasks much faster than humans, such as mechanically submitting information. Weather updates, adding a Wikipedia template to all pages in a specific category, or sending a thank you message to Courtesy of your new followers. It can be designed to be useful as an individual virtual assistant alongside Siri1 or for people-friendly customer service for organizations and chatbots like Microsoft Tay's AI.
- On the other hand, social bots can be designed to engage in malicious sports involving spam, malware spread, spoofing, launching Sybil attacks, etc.

In previous research, several strategies were used to protect the security of social networks over the Internet [3]. User behavior is the most direct manifestation of a consumer's purpose because unique clients have distinctive online behavior, choices, and behavior (for example, click or arrangement method and writing speed). In other words, we can extract and check hidden records of consumer behavior online to create a profile and discover private users. However, we should also be aware of situational factors that can play a role in changing a person's behavior online. In other words, a person's behavior is dynamic, and his environment continuously transforms, i.e., outside of the observable environment (such as the environment and practice) from the application context and the hidden environment in consumer information [4]. We need to collect and examine the social scene of consumer behavior to distinguish between social bots and regular users, find malicious social bots, and reduce harmful social bots [14]. Moreover, it understands the differences between harmful social bots and ordinary users in dynamic behavior.

One of the malicious functions of social bots is the power to spread wrong information [5]. For example, the Syrian electronic army infiltrates the Associated Press account on Twitter and declares that the White House is under attack and that Obama has been injured. This fake news leads to panic and a significant loss in the stock market [6-10]. Another malicious ability of social bots is that they are an effective form of advertising. This malicious hobby is called astroturfing: an attempt to falsely influence the real foundations of product, people, and campaign coverage marketing. Another drawback is that bots can purchase fake scores and ratings. For example, there is an effect on bots that serve this cause. Also, it is possible to locate many web pages that offer the wrong service[11-14].

Also, a social bot can be harmful by pretending to be a real man, woman, or business owner, i.e., for example, identity fraud. One of the sinister purposes of impersonation is to work to advance ideologies. With this promotion, attackers can lie to people on networks or create fake identities for a real search. It can be used in malicious sports, such as assaulting Sybil fans and oversized botnets with OSN money owed. Specifically, in this document, we aim to stumble across malicious social bots in real-time social media structures using (1) that provide the features of the opportunity transition between consumer clicks to flow based on an analysis of the social status of things. Moreover, (2) Design an algorithm to detect malicious social bots based on space-time functions.

2.Related Works

Several techniques are proposed in the literature to detect social bots for OSN. We summarize these strategies in a systematic classification and reveal applicable studies for each category of social bot detection.

Barbon et al. (2018**),** Social interactions were raised in settings that influence human behavior and perceptions. OSN customers generate a large amount of content based on social interactions. However, the enormous popularity and simplicity of OSN created an excellent case for malicious activities and placed its reliability at risk. A fully wavelet-based model was developed to locate the automatic dispatch of records in OSN that classified customers as human or dangerous bots. This was due to the spectral patterns being received from consumer text content. Specifically, it started to release by eliminating the use of the Random Forest algorithm in two real sets of Twitter information. Corresponding results showed that the advanced models have an average accuracy of 94. 97% reflected consideration of the exact scenarios of difficulty and diversity.

Zhou et al. (2017), Social media brought step-by-step, cost-effective capabilities with the help of real allowed and digital foreign money to be used. They served as new platforms for hosting a branch of commercial sports that included online promotional activities, where users can earn digital currency as rewards for participating in such events. Both OSNs and business partners participate significantly, while attackers design a fixed amount of money owed to raise foreign digital money from these events, making these activities ineffective and causing significant economic losses. It was essential to detect malicious payments due before net promotional activities proactively and ultimately reduced priorities for the bonus. The authors recommended a unique system, particularly ProGuard, to achieve this goal using systemic integration features that distinguish accounts from three perspectives. These include their general behaviors, reloading patterns, and using their Forex. They have conducted massive experiments primarily based on data collected from Tencent QQ. Experimental results demonstrated that the system could achieve a high detection rate of 96.67% with a deficient false charge of 0.3%.

Ferrara et al. (2016), The Turing test is intended to understand a human's behavior like a machine algorithm. So, the challenge was more relevant than ever in today's social media context, where textual restrictions restricted expressive human electricity, and real incentives abound increased retailers for programs that imitated humans known as social bots. These elusive entities fill the social media ecosystems extensively, often ignoring part of the real population. Bots can be either benign or dangerous, with the goal of persuasion, mutilation, or fraud. They discussed the features of the latest social bots and how their existence can threaten online ecosystems and our society. They then reviewed current efforts to identify social bots on Twitter. The properties associated with content materials, networks, feelings, and time patterns of activity were imitated with the help of bots; however, at the same time, it could help distinguish between artificial behaviors of humans, which results in engineering firms for social manipulation.

Schifanella et al. (2015) they were studied and analyzed the search behavior of Yahoo!. The image search dataset is based on the assumption that behavior was conditioned on the query type. The queries are divided into two orthogonal classification taxonomies, and relevant questions are recognized, along with the types of essential items at the intersection of those classifications. The authors studied user search behavior across a range of huge search periods for each type of query, session attribute analysis, query reformulation patterns, click patterns,

and web page display patterns. They detected critical behavioral differences in the types of queries, especially some exploratory questions, while other types correspond to specific searches. They also completed observations using a survey to correlate behavioral variations with customer motivation. The results highlighted the importance of considering query categories to understand consumer behavior on image search platforms better.

Gao et al. (2014). In OSN, spam from friends and colleagues reduces the enjoyment of the Internet and causes harm to users with less security knowledge. The above countermeasures struggled OSN Spam from unique perspectives. Due to the diversity of spam, there was seldom a current approach that could independently detect most OSN spam. This article performed an experimental analysis of the text sample for a wide range of OSN spam. An inspirational result was the general public (63.0%) of spam was accumulated using basic templates. Therefore, extracted models of spam were detected through the current strategies and then matched the messages with the forms to quickly and accurately detect spam. We implement this vision through tangram, an OSN spam filter that performs online scanning of consumer traffic. Tangram automatically splits OSN spam into sections and uses parts to collect templates to remove spam in the future. The experiment showed that the tangram is corrected and could be spurted into creating models to accelerate newly emerging campaigns. Specifically, tangram discovered the maximum spam email based on the regular template at 95.7% premium quality original prices, while it was detected the current template squeeze technology 32.3% more productive. Tangram integration and spam removal achieved a standard 85.4% accuracy rate and 0.33% false-positive rate.

Brito et al. (2013) Online social network services have become one of the primary forms of human communication and interaction. However, the emergence of massive hidden attacks using bots on social media led to a growing need for environmental disclosure methodologies. The bots objective could be similar to the goals of traditional human criminal activity by conducted multiple fraud sellers. The tracking software also acted as neutral entities that create fake (significantly persuasive) profiles or hijack a real personal profile from the use of the affected laptop. It can be challenging to discover media social bots using ordinary human or computer algorithms that evaluate family members themselves. However, bots cannot simulate the reaction function of human behavior over the years. False repetition In addition to random and sometimes chaotic movements, the characteristic of human behavior remains challenging to simulate. However, it can be straightforward to distinguish this area of human experience from particular behavioral patterns. Therefore, behavioral assessment and identification methodologies are vital for the accurate detection of a social bot. In those artworks, we call a new paradigm, while using more than one scale of observation of Jupiter interactions within a social network; it is capable and must be prominent between the unique behaviors of humans and bots in the social community. Consequently, bot unique behavioral patterns could be built from social networks, and traditional human interactions allowed accurate detection of one of the hidden threats of the Internet.

Hwang et al. (2012) The Social Mediator Forum was created to close gaps between theory and practice of studies and the development of social networks. The articles were aimed to foster additional interest in recent insights and stories in the rapidly evolving field of social media, some of which may also influence perspectives and methods in the more established areas of human-computer interaction (HCI). Each forum post contained many small human contributions that represent unique perspectives on a specific topic. Previous installments from

this panel gathered many views on how social media reshape relationships between exceptional stakeholders within national health and government mandates. The ground-breaking article highlights some of the approaches of social bots, programs that work independently on social media, and transform relationships within sites. It also addressed how these differences further affected interpersonal relationships. The recent article in ACM Communications, known as "The Social Life of Robots," said that researchers have commenced exploring the potential of 'social' machines capable of operating with minimal human supervision. This article highlighted recent trends in human and robot interactions within the physical world; this newsletter specializes in human communications in the international virtual world. Authors were explored and expanded the boundaries of designing, implementing, and analyzing the behavior and impact of media social bots online, who have invited some humans from other frontiers to share some of their insights, reviews, and destination expectations for social bots.

3.Proposed Methodology

To sufficiently recognize malicious social bots in online social networks, examines user behavior features, and identify transition possibility features among user clickstreams based on the transition probability features and time period features. Further, this paper proposed a semi-supervised K-Means Clustering Algorithm based on space-time features.

3.1Bot Detection in Online Social Networks

In online social networks, bot detection has leading technologies. It is mainly based on content inspection and entirely depends on a network graph and a mixed set. Initially, it tries to find that bots that are mostly restricted in the direction of the relationship between different clients. The quality of relationships has become one of the critical indicators for detecting malicious or false hobbies. This type of detection method focuses on content examination and management administration. Acceptance control, coupled with the scrutiny of content materials, becomes robust rather than fake profiles. Another way to facilitate the detection of bots is to divide the network into smaller non-public subnets. These smaller networks are more aware of their relationships with members. Therefore, the growth of the false relationship became difficult in one of these cases. It is believed that if there are a large number of Sybil nodes, they are smaller edge groups looking for such short edges on the web. A unique "random walk" type is suggested to reduce the low quotient (the leading edges) between an honest node and a benign Sybil node. Sybil nodes can also benefit from legality by authenticating some benign nodes, but doing so on a large scale, automated, can be difficult.

3.2 Traditional K-Means Clustering Algorithm

K -means the clustering algorithm has long records of improvement. However, it is also a great research topic for many scholars. The basic idea of the traditional K-means clustering algorithm is as follows. First, randomly select the primary cluster center K. The K is the parameter set by users. It is the expected cluster number. Every data object within a set of statistics is assigned to the nearest cluster center—the data objects attached to the same cluster center will form a group. Then, it will be updated according to the data object in each group. The assign and updates will be repeated until the cluster center is not modified. In general, the algorithm should preferably use the iterative approach to improve clustering by helping to move the clustering centers continuously.

The basic procedure of the traditional K-means clustering algorithm is shown below in Figure 1.

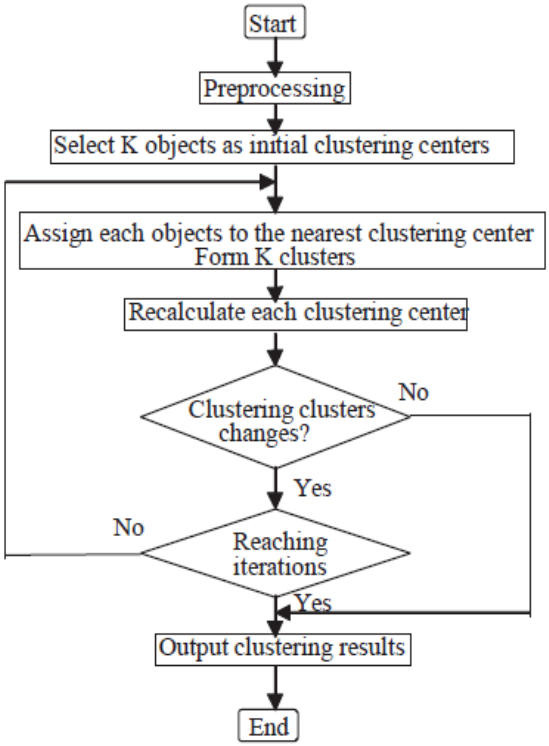

Figure 1: Basic structure for Traditional K-means clustering algorithm

3.3 Proposed System Architecture

The proposed work illustrates the various stages of data cleaning, data preprocessing, feature selection, statistic set, and sequence of operations performed after obtaining the user's clickstream data set. The detailed steps of the proposed work are shown in Figure 2

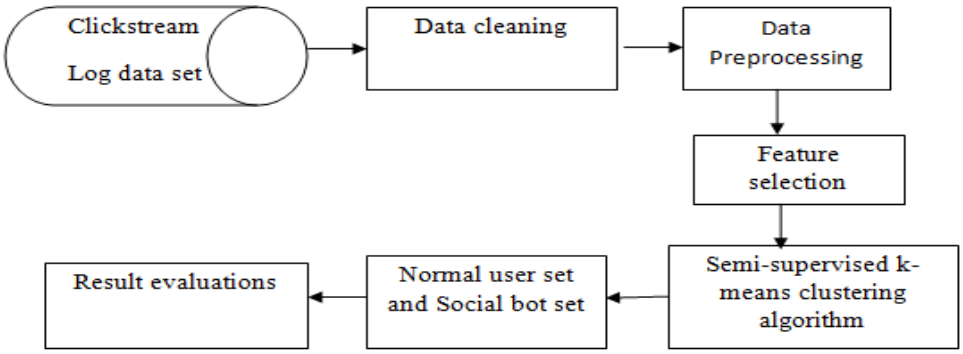

Figure 2: Proposed method architecture

1) Data cleaning

Clickstream data should be cleaned to remove the wrong facts, receive the distinct possibility of passing between click flows, and avoid errors due to less information. Collecting statistics from one company and making them useful to others is an enormous different challenge. Accumulated data may be in a disorganized design, and there may be many blank values, invalid

data values, and un-relevant information in different ways. Cleaning all facts cha,nging them with approximate details, and filling in negative points with a few fixed alternative costs are the necessary steps in preprocessing information. Even accumulated records can contain garbage values. It doesn't have to have an original design that should be in any format. This process is performed to maintain the meaning of the statistics and similar processing. The data should be stored in a prepared layout.

2) Data processing

Some data is randomly chosen from the regular user group, and social bots are identified on the tag. The regular user account is called 1, and the social bot account is called -1. Seed users are described as a cluster category.

3) Feature selection

In feature selection, examine the difficulty of clustering observations using a probably large set of features. One might expect that the real underlying clusters existing in the data oppose only with deference to a small portion of the features, and will be desired if one clusters the observations utilizing the complete set of features.

4) Semi-supervised K-Means Clustering Algorithm (SSKMC)

SSKMC uses a small number of specified data to support and determines the clustering of unlabeled data. This paper investigates the method of labeled data to generate and optimize initial cluster centers for the k-means algorithm.

Algorithm:

$\boldsymbol{Input}$: $Data\ set\ X = \{x1, x2, \ldots\ldots, xN\}\ and\ its\ external\ index$

$XL = \{xL1, xL2, , \ldots\ldots, xLlable\}$

$\boldsymbol{Output}$: $Clustering\ result\ C = \{C_1, C_2, \ldots\ldots, C_k\}\ and\ initial$

$clustering\ centers\ \ c_i = c_1, c_2, \ldots\ldots, c_k.$

$\boldsymbol{BEGIN}$

$K_L = Number\ of\ Distinct\ Label(X_L);$

$Select\ K_L\ objects\ with\ different\ labels\ from\ X_L;$

$\boldsymbol{if}(K_L \geq \sqrt{N})$

{

$K = K_L;$

$\boldsymbol{do}$ {

$\boldsymbol{for\ each}\ x_i \in X$

$if(x_i \in X_L)$

$Assign\ x_i\ to\ the\ cluster\ whose\ initial\ cluster\ center\ has\ the\ same\ label;$

$\boldsymbol{else}$

$Assign\ x_i\ to\ the\ closest\ cluster;$

$update\ clusters;$

$\}\ until\ converges$

}

$else$

$for\ (k = K_L + 1; k \leq \sqrt{N}; K++)$

{

$Select\ the\ farthest\ object\ from\ the\ k-1\ objects\ as\ the\ next\ initial\ cluster\ center,$
$do\ \{$
$for\ each\ x_i \in X$
$if\ (x_i \in X_L)$
$Assign\ x_i\ to\ the\ cluster\ whose\ initial\ cluster\ center\ has\ the\ same\ label;$
$else$
$Assign\ x_i\ to\ the\ closest\ cluster;$
$Update\ cluster;$
$\}until\ converges;$
$J_k = \sum_{i=}^{k} \sum_{j=1}^{n_i} d(C_i, x_j)$
$\}$
$Select\ \ the\ minimum\ J_k\ an\ K = k;$
$\}$
$\boldsymbol{Output}\ K\ clusters\ and\ K\ initial\ cluster\ centers$

5) Obtain the normal user set and social bots set

The average user set and social bots set can be finally obtained by detecting.

6) Result evaluation

The results evaluation is based on three different metrics of Precision, Recall, and F1 Score. In the meantime, use accuracy metric to compare with the SVM algorithm to verify the efficiency of the method. Accuracy is the ratio of the number of samples correctly classified by the classifier to the total number of samples.

4.Results and Discussions

To show the effectiveness of the proposed method, various malicious social bots designed using three categories of capabilities to precise detection. The precision of detecting distinctive forms of malicious social bots through the use of 3 types of features shown in Figure 3. The recall of identifying unique types of malicious social bots by using three groups of features shown in Figure 4. Here found that (1) the accuracy of the semi-controlled aggregation approach to detect the same form of malicious social bots that entirely dependent on transport, opportunity capabilities, and combined skills are higher than the semi-supervised aggregation method wholly based on the quantitative characteristic (2) about social bot programs. Easy malware, the benefit of the transfer opportunity feature, and the combined feature can effectively stumble upon funds due to malicious social bots. However, the application of diverse functions may have more significant effects.

An experimental bottom line indicates that the accuracy of a semi-supervised and fully-assembled method of aggregation for the detection of malicious social bots with aggregate malicious functions can be as excessive as 93.1%, and the rate of forgetfulness is 97—F1 score is 95.2%. Compared to the semi-supervised package approach with quantitative functions, our method can reveal the malicious debt of social bots in social systems over the Internet, as it should. To confirm the accuracy of the technique, the issuance of the auxiliary vector system depends entirely on the transfer opportunity, the semi-supervised clustering technique based on mixed features. The semi-supervised clustering algorithm fully supervised on the transition opportunity and the semi-supervised approach the quantitative feature-based clustering technique is associated with the same set of facts.

The precision of exposure of the various methods of malicious social bots is illustrated in

Figure 3. The experiment demonstrates that the proposed semi-supervised k-means clustering method is entirely based on the ability to switch between clickstream flows through people who can multiply bots efficiently. The contrast between different technologies shows that the precision and recall of the method for detecting malicious social bots based on the chance of transmission can reach 95% or more.

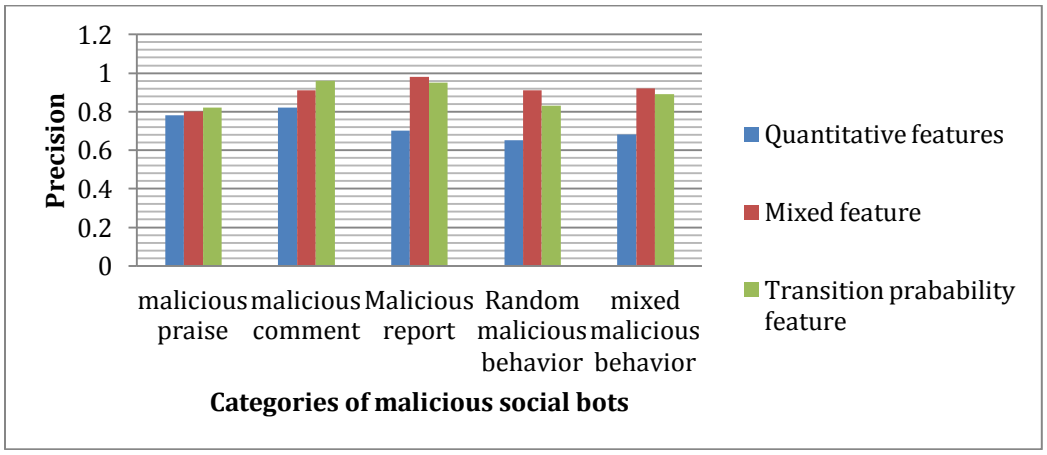

Figure 3: Precision of detection methods on various features for various types of malicious social bots

Compared with the traditional detection method based on the quantitative feature, accuracy is improved by 15 % on average. The method can effectively detect malicious accounts on social platforms. Finally, the malicious social bots detection program was deployed and run on the CyVOD platform. In the background user information list, malicious accounts of social bots are marked in red for convenience in addressing malicious social bots. Figure 4 shows Recall of detection methods on various features.

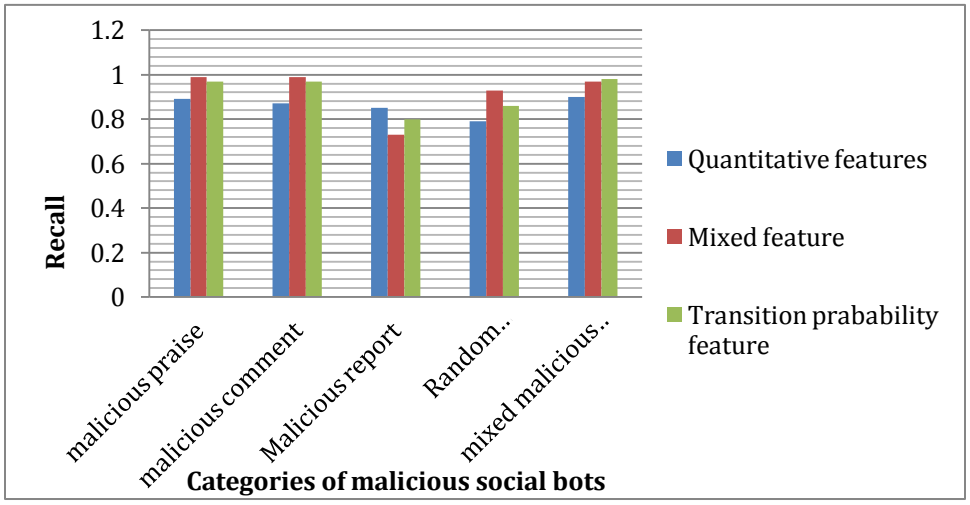

Figure 4: Recall of detection methods on various features

Table1: Profile analysis results

	Accuracy (%)	Misclassification Rate (%)	True positive rate (%)
Social spam bot	95.77	4.23	96.81
Traditional spam bot	96.25	3.75	97.13

Fake followers	100	0	100
NBC News Russian bots	99.87	0.13	98.91
Total	97.75	2.25	98.98

We will show our classification model performance when we use various kinds of bots in our data set. This consists of a total of 16,649 Twitter accounts. Our model got 97.75%, with a rating of 2.25% and real positive rates of 98.98. Here is a graph summarizing our classification results while using support vector processing in profile information. Table 1 shows profile analysis results**.**

As shown in the figure 5, A subset of Russian bots and real user accounts is used for text analysis. It was 90% accurate, and therefore the misclassification rate is 10%. The true positive rate for these results is 100%. Figure 5 shows the results of profile Analysis.

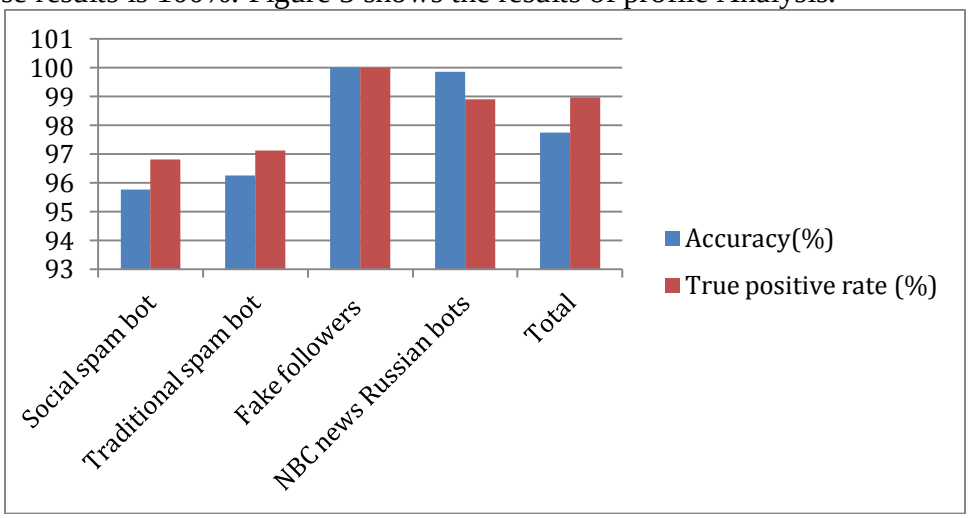

Figure 5: The results of profile Analysis

5.Conclusion

Social networks are powerful media that connect hundreds of thousands of people around the world. Therefore, it can also be attractive to social bots. Because the possible harm of social bots programs consists of identity fraud, surfing, content polluter, incorrect information posting, etc., it may be necessary for bots and humans to become popular to avoid unknown situations entirely based on wrong assumptions. We have suggested a unique approach to detecting social malware bots in social media over the Internet. Here, a semi-supervised k-means clustering algorithm used to identify the social bots in online social networks. The precision and recall of two measures are taken to perform the experiments. The Experimental results have confirmed that the ability to switch between clickstream based on a social scenario analysis can be used to identify malicious social bots on OSNs through the Internet accurately. And the proposed detection method can be expanded and improved to special designs and precise determinations of a broader set of malicious social bots.

Acknowledgement: Not Applicable.
Funding Statement: The author(s) received no specific funding for this study.
Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

[1] S. K. Decade and A. M. Bagade, "A review on detecting automation on Twitter accounts," pp. 69-72, 2015.

[2] Sreedhar Bhukya, D. Srinivasarao and Khasim Saheb, "Environmental Monitoring with Wireless Sensor Network for Energy Aware Routing and Localization," *Journal of Sensors, IoT & Health Sciences*, vol.1, no.1, pp.27-39, 2023.

[3] Freitas, F.Benevenuto, S.Ghosh and A. Veloso, "Reverse engineering social bot infiltration strategies in twitter," *2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Paris, France, pp. 25-32, 2015.

[4] S. Jr Barbon, R.A Igawa, R.Guido, "Detection of the human, legitimate bot, and malicious bot in online social networks based on wavelets," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol.14, no.1s, pp.1-17, 2018.

[5] Y. Zhou, D.W. Kim, J. Zhang, L.Liu, H. Jin et al., "ProGuard: Detecting malicious accounts in a social network- based online promotions,"*IEEE Access*, vol. 5, pp. 1990-1999, 2017.

[6] C. K. Chang, "Situation analytics: A foundation for a new software engineering paradigm," in *Computer*, vol. 49, no. 1, pp. 24-33, 2016.

[7] Sreedhhar Bhukya, K. VinayKumar and N.C. Santosh, "A Novel Dynamic Novel Growth model for Mobile Social Networks," *Journal of Computer Allied Intelligence*, vol.2, no.1, pp.46-53, 2024.

[8] Ferrara, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol.57, no.07, pp. 96-104, 2016.

[9] R.A.Schifanella, A. Jaimes and C.W. Chung, "A large-scale study of user image search behavior on the Web," *International Conference on Human Factors in Computing Systems*, pp. 985-994, 2015.

[10] H.Gao, Y.Yang, K.Bu, Y.Chen, Downey et al., "Spam ain't as diverse as it seems: Throttling OSN spam with templates underneath," 30th Annual Computer Security Applications Conference, ACSAC 2014, New Orleans, United States, pp. 76-85, 2014.

[11] F. Brito, A.Nogueira and E.Rocha, "Detecting social-network bots based on multiscale behavioral analysis," *International Conference on Emerging Security Information, Systems and Technologies*, pp. 81-85, 2013.

[12] T. Hwang, I. Pearce and M. Nanis, ``Socialbots: Voices from the fronts," *Interactions*, vol. 19, no. 2, pp. 38-45, 2012.

[13] G.Rekha, V.Murali Mohan and M.Ranjeeth Kumar, "Network intrusion detection system for Internet of Things based on enhanced flower pollination algorithm and ensemble classifier," *Concurrency and Computation practice and Experience*, vol.34, no.21, 2022.

[14] G.Rekha, V. Murali Mohan and M.Ranjeeth kumar, "Hybridization of Bottlenose Dolphin Optimization and Artificial Fish Swarm Algorithm with Efficient Classifier for Detecting the Network Intrusion in Internet of Things (IoT)," *International Journal of Intelligent Systems and Applications in Engineering*, vol.12, no.6s, pp.220–232, 2024.