RB Journal of Lib & Information Science
(UGC Care Group I Listed Journal)
ISSN: 0972-2750
Vol-12 Issue-11 No.01: 2022

DETECTING MALICIOUS SOCIAL BOTS BASED ON CLICK STREAM SEQUENCES

[1]M.Raghuveer, [2]Dr.V Bapuji,Guide,
[1], Mtech (CSE), Department of Computer Science, VAAGESWARI COLLEGE OF ENGINEERING Thimmapur, Karimnagar, Telangana, INDIA, H.no:20S41D5813
[2]Professor, Department of Computer Science, VAAGESWARI COLLEGE OF ENGINEERING Thimmapur, Karimnagar, Telangana, INDIA, Bapuji.Vala@Gmail.Com

ABSTRACT: With the significant increase in the volume, velocity, and variety of user data (e.g., user generated data) in online social networks, there have been attempted to design new ways of collecting and analyzing such big data. For example, social bots have been used to perform automated analytical services and provide users with improved quality of service. However, malicious social bots have also been used to disseminate false information (e.g., fake news), and this can result in real-world consequences. Therefore, detecting and removing malicious social bots in online social networks is crucial. The most existing detection methods of malicious social bots analyze the quantitative features of their behavior. These features are easily imitated by social bots; thereby resulting in low accuracy of the analysis. A novel method of detecting malicious social bots, including both features selection based on the transition probability of clickstream sequences and semi-supervised clustering, is presented in this paper. This method not only analyzes transition probability of user behavior clickstreams but also considers the time feature of behavior. Findings from our experiments on real online social network platforms demonstrate that the detection accuracy for different types of malicious social bots by the detection method of malicious social bots based on transition probability of user behavior clickstreams increases by an average of 12.8%, in comparison to the detection method based on quantitative analysis of user behavior.

***Keywords** – Online social network, social bots, user behavior, semi-supervised clustering.*

1. INTRODUCTION

In online social networks, social bots are social accounts controlled by automated programs that can perform corresponding operations based on a set of procedures [1]. The increasing use of mobile devices (e.g., Android and iOS devices) also contributed to an increase in the frequency and nature of user interaction via social networks. It is evidenced by the significant volume, velocity and variety of data generated from the large online social network user base. Social bots have been widely deployed to enhance the quality and efficiency of collecting and analyzing data from social network services. For example, the social bot SF QuakeBot [2] is designed to generate earthquake reports in the San Francisco Bay, and it can analyze earthquake related information in social networks in real-time. However, public opinion about social networks and massive user data can also be mined or disseminated for malicious or nefarious purpose [3]. In online social networks, automatic social bots cannot represent the real desires and intentions of normal human beings, so they are usually looked upon malicious ones.

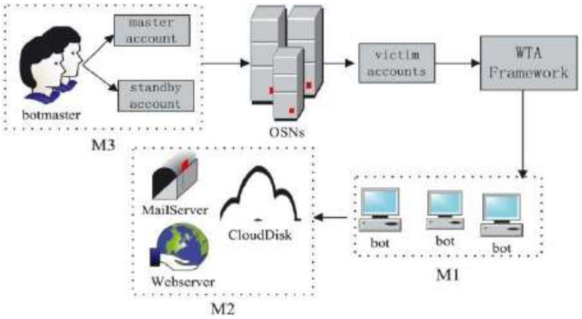

Fig.1: Example

For example, some fake social bots accounts created to imitate the profile of a normal user, steal user data and compromise their privacy [4], disseminate malicious or fake information [5], [6], malicious comment, promote or advance certain political or ideology agenda and propaganda [7], and influence the stock market and other societal and economical markets [8]. Such activities can adversely impact the security and stability of social networking platforms. In previous research, various methods were used to protect the security of online social network [9]–[11]. User behavior is the most direct manifestation of user intent, as different users have different habits, preferences, and online behavior (e.g., the way one clicks or types, as well as the speed of typing). In other words, we may be able to mine and analyze information hidden in user's online behavior to profile and identify different users. However, we also need to be conscious of situational factors that may play a role in changing user's online behavior. In other words, user behavior is dynamic and its environment is constantly changing – i.e., external observable environment (e.g., environment and behavior) of application context and the hidden environment in user information [12]. In order to distinguish social bots from normal users accurately, detect malicious social bots, and reduce the harm of malicious social bots, we need to acquire and analyze social situation of user behavior and compare and understand the differences of malicious social bots and normal users in dynamic behavior. Specifically, in this paper, we aim to detect malicious social bots on social network platforms in real-time, by (1) proposing the transition probability features between user clickstreams based on the social situation analytics; and (2) designing an algorithm for detecting malicious social bots based on spatiotemporal features.

2. LITERATURE REVIEW

A new approach to bot detection: Striking the balance between precision and recall:

The presence of bots has been felt in many aspects of social media. Twitter, one example of social media, has especially felt the impact, with bots accounting for a large portion of its users. These bots have been used for malicious tasks such as spreading false information about political candidates and inflating the perceived popularity of celebrities. Furthermore, these bots can change the results of common analyses performed on social media. It is important that researchers and practitioners have tools in their arsenal to remove them. Approaches exist to remove bots, however they focus on precision to evaluate their model at the cost of recall. This means that while these approaches are almost always correct in the bots they delete, they ultimately delete very few, thus many bots remain. We propose a model which increases the recall in detecting bots, allowing a researcher to delete more bots. We evaluate our model on two real-world social media datasets and show that our detection algorithm removes more bots from a dataset than current approaches.

ProGuard: Detecting malicious accounts in social network- based online promotions:

Online social networks (OSNs) gradually integrate financial capabilities by enabling the usage of real and virtual currency. They serve as new platforms to host a variety of business activities, such as online promotion events, where users can possibly get virtual currency as rewards by participating in such events. Both OSNs and business partners are significantly concerned when attackers instrument a set of accounts to collect virtual currency from these events, which make these events ineffective and result in significant financial loss. It becomes of great importance to proactively detecting these malicious accounts before the online promotion activities and subsequently decreases their priority to be rewarded. In this paper, we propose a novel system, namely ProGuard, to accomplish this objective by systematically integrating features that characterize accounts from three perspectives including their general behaviors, their recharging patterns, and the usage of their currency. We have performed extensive experiments based on data collected from the Tencent QQ, a global leading OSN with built-in financial management activities. Experimental results have demonstrated that our system can accomplish a high detection rate of 96.67% at a very low false positive rate of 0.3%.

Efficient compressed ciphertext length scheme using multi-authority CP-ABE for hierarchical attributes:

In an attribute-based encryption, the user is identified with help of some attributes and their functions for encryption and decryption of the data. The current techniques based on attribute-based encryption have found that if user's access structure includes a considerable amount of attribute information labeled as Don't Care, then the encryption pairing operation has low calculation efficiency and ciphertext information redundancy. In this paper, we have proposed a hierarchical multi-authority attribute-based encryption on prime order groups to tackle these problems. Our encryption technique has a polycentric attribute authorization system based on an AND gate access structure, with a unified attribute index established by each attribute authority throughout the system, to form a binary tree, i.e., attribute access tree. The state value of the parent node can be determined by the state of its child node in an attribute access tree. The attribute-based encryption established in this manner is theoretically proven to effectively decrease the calculation amount for decryption and compress the redundant information in the ciphertext as much as possible. Our encryption technique has a theoretical and practical significance in the system of “large universe”constructions.

Behavior enhanced deep bot detection in social media:

Social bots are regarded as the most common kind of malwares in social platform. They can produce fake messages, spread rumours, and even manipulate public opinions. Recently, massive social bots are created and widely spread in social platform, they bring negative effects to public and netizen security. Bot detection aims to distinguish bots from human and it catches more and more attentions in recent years. In this paper, we propose a behavior enhanced deep model (BeDM) for bot detection. The proposed model regards user content as temporal text data instead of plain text to extract latent temporal patterns. Moreover, BeDM fuses content information and behavior information using deep learning method. To the best of our knowledge, this is the first trial that applies deep neural network in bot detection. Experiments on real world dataset collected from Twitter also demonstrate the effectiveness of our proposed model.

A situational analytic method for user behavior pattern in multimedia social networks:

The past decade has witnessed the emergence and progress of multimedia social networks (MSNs), which have explosively and tremendously increased to penetrate every corner of our lives, leisure and work. Moreover, mobile Internet and mobile terminals enable users to access to MSNs at anytime, anywhere, on behalf of any identity, including role and group. Therefore, the interaction behaviors between users and MSNs are becoming more comprehensive and complicated. This paper primarily extended and enriched the situation analytics framework for the specific social domain, named as SocialSitu, and further proposed a novel algorithm for users’ intention serialization analysis based on classic Generalized Sequential Pattern (GSP). We leveraged the huge volume of user behaviors records to explore the frequent sequence mode that is necessary to predict user intention. Our experiment selected two general kinds of intentions: playing and sharing of multimedia, which are the most common in MSNs, based on the intention serialization algorithm under different minimum support threshold (Min_Support). By using the users’ microscopic behaviors analysis on intentions, we found that the optimal behavior patterns of each user under the Min_Support, and a user’s behavior patterns are different due to his/her identity variations in a large volume of sessions data.

3. METHODOLOGY

Morstatter *et al.* proposed a heuristic-type supervised BoostOR model with increasing recall rate to detect malicious bots, which using the proportion of tweets forwarded to the published tweets on the Twitter, the mean length of tweets, URL, and forwarding interval.

Wang *et al.* constructed a semi-supervised clickstream similarity graph model for user behavior to detect abnormal accounts in Renren. According to the social interactions between users of the Twitter user to identify the active, passive and inactive users, a supervised machine learning method was proposed to identify social bots on the basis of age, location and other static features of active, passive, and inactive users in the Twitter, as well as interacting person, interaction content, interaction theme, and some dynamic characteristics.

Disadvantages:

Annotation and training for large amounts of data are required in supervised learning. Tagging data requires time, manpower, and is generally unsuitable for the big data social networking environment. In other words, such an approach is generally ill-suited for real-time detection of malicious social bots on social networking platforms.

Unsupervised learning, on the other hand, it does not require manual labeling of data. However, unsupervised learning approaches are sensitive to initial values and can only classify different results. It is not possible to determine which cluster is normal and which cluster is abnormal.

In this paper, we aim to detect malicious social bots on social network platforms in real-time, by (1) proposing the transition probability features between user clickstreams based on the social situation analytics; and (2) designing an algorithm for detecting malicious social bots based on spatiotemporal features.

In order to better detect malicious social bots in online social networks, we analyze user behavior features and identify transition probability features between user clickstreams Based on the transition probability features and time interval features, a semi-supervised social bots detection method based on space-time features is proposed.

Advantages:
We then analyze and classify situation aware user behaviors in social networks using our proposed semi supervised clustering detection method. This allows us to promptly detect malicious social bots using only a small number of tagged users.

To identify potential malicious social bots in online social networks in real-time, we analyze the social situation behavior of users in online social networks. We also evaluate user behavior features and select the transition probability of user behavior on the basis of general behavior characteristics. We then analyze and classify situation aware user behaviors in social networks using our proposed semi supervised clustering detection method. This allows us to promptly detect malicious social bots using only a small number of tagged users.

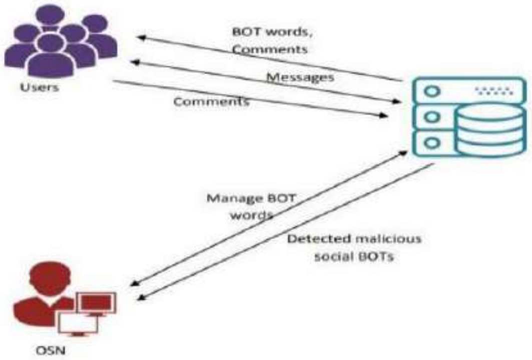

Fig.2: Architecture

Detect malicious social bots on social network platforms in real-time, by (1) proposing the transition probability features between user click streams based on the social situation analytics; and (2) designing an algorithm for detecting malicious social bots based on spatiotemporal features. In order to better detect malicious social bots in online social networks, we analyze user behavior features and identify transition probability features between user clickstreams Based on the transition probability features and time interval features, a semi-supervised social bots detection method based on space-time features is proposed.

4. IMPLEMENTATION

4.1 Modules:

- **OSN Server**

In this module, the OSN Server has to login by using valid user name and password. After login successful he can do some operations such as view all user details and authorize them, list of all friends requests and response, View all posts like images and messages user, view all Similar group users like doctors, Engineers, Business Man, etc., OSN Server can add some BOTS words to the database and view the all words added by him and based on that negative words admin can find all users behavior and also produce chart for that behavior words.

- **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.

- **User**

In this module, there are n numbers of users are present. User should register with group option before doing some operations. After registration successful he has to wait for admin to authorize him and after admin authorized user can login by using authorized user name and password. Login successful he will do some operations like view profile details, Search friends based on keyword or friends name, view the friend requests, post message with image to all friends. Find posts of friends and comment on those posts.

Users can also view all his friends and delete those who don't want, view all group posts like doctor or engineer etc.,

- **Viewing Profile Details**

In this module, the user can see their own profile details, such as their address, email, mobile number, profile Image.

- **Search Friends, Request, and View Friend Requests, View all Friend Details**

In this, the user search for other users by their names, send requests and view friend requests from other users. User can see all his friend details with their images and personnel details.

5. EXPERIMENTAL RESULTS

Detecting Malicious Social Bots Based on Clickstream Sequences

Fig.3: URL Link

Detecting Malicious Social Bots Based on Clickstream Sequences
HOME
USER
OSN
DETECTING MALICIOUS SOCIAL BOTS BASED ON CLICKSTREAM SEQUENCES
SOCIAL MEDIA

Fig.4: Home Page

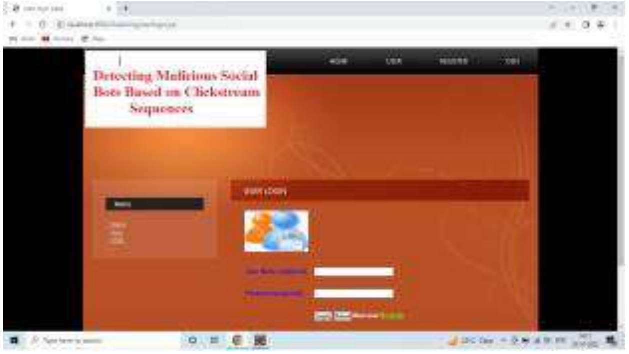
Detecting Malicious Social Bots Based on Clickstream Sequences
USER LOGIN

Fig.5: User login page

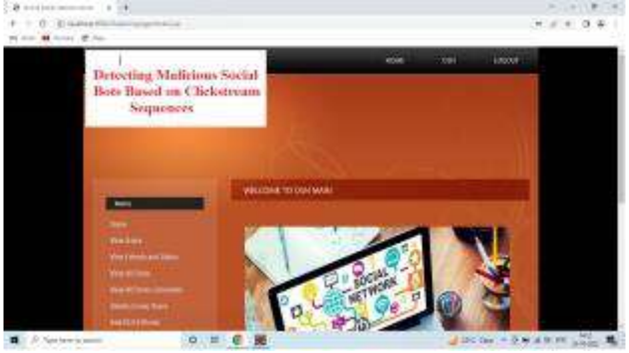
Detecting Malicious Social Bots Based on Clickstream Sequences
WELCOME TO OSN MAIN
SOCIAL NETWORK

Fig.6: OSN Home Page

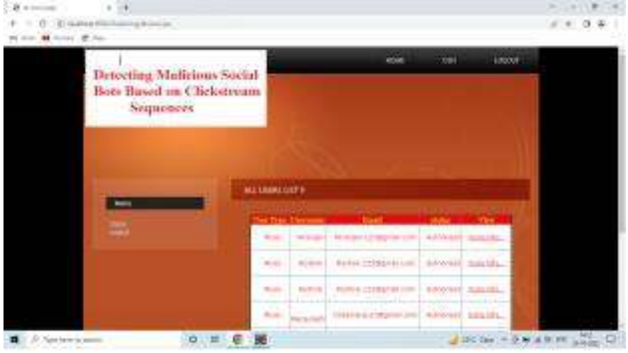
Detecting Malicious Social Bots Based on Clickstream Sequences
ALL USERS LIST'S

Fig.7: User details

6. CONCLUSION

We proposed a novel method to accurately detect malicious social bots in online social networks. Experiments showed that transition probability between user click streams based on the social situation analytics can be used to detect malicious social bots in online social platforms accurately.

7. FUTURE WORK

In future research, additional behaviors of malicious social bots will be further considered and the proposed detection approach will be extended and optimized to identify specific intentions and purposes of a broader range of malicious social bots.

REFERENCES

[1] F. Morstatter, L. Wu, T. H. Nazer, K. M. Carley, and H. Liu, ``A newapproach to bot detection: Striking the balance between precision andrecall," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal.Mining*,San Francisco, CA, USA, Aug. 2016, pp. 533_540.

[2] C. A. De Lima Salge and N. Berente, ``Is that social bot behaving unethically?"*Commun. ACM*, vol. 60, no. 9, pp. 29_31, Sep. 2017.

[3] M. Sahlabadi, R. C. Muniyandi, and Z. Shukur, ``Detecting abnormalbehavior in social networkWebsites by using a process mining technique,"*J. Comput. Sci.*, vol. 10, no. 3, pp. 393_402, 2014.

[4] F. Brito, I. Petiz, P. Salvador, A. Nogueira, and E. Rocha, ``Detectingsocial-network bots based on multiscale behavioral analysis," in *Proc.7th Int. Conf. Emerg. Secur.Inf., Syst. Technol. (SECURWARE)*, Barcelona,Spain, 2013, pp. 81_85.

[5] T.-K. Huang, M. S. Rahman, H. V. Madhyastha, M. Faloutsos, andB. Ribeiro, ``An analysis of socware cascades in online social networks,"in *Proc. 22nd Int. Conf. World Wide Web*, Rio de Janeiro, Brazil, 2013,pp. 619_630.

[6] H. Gao*et al.*, ``Spam ain't as diverse as it seems: Throttling OSN spam withtemplates underneath," in *Proc. 30th ACSAC*, New Orleans, LA, USA,2014, pp. 76_85.

[7] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, ``The rise ofsocial bots," *Commun. ACM*, vol. 59, no. 7, pp. 96_104, Jul. 2016.

[8] T. Hwang, I. Pearce, and M. Nanis, ``Socialbots: Voices from the fronts,"*Interactions*, vol. 19, no. 2, pp. 38_45, Mar. 2012.

[9] Y. Zhou *et al.*, ``*ProGuard*: Detecting malicious accounts in socialnetwork-based online promotions," *IEEE Access*, vol. 5, pp. 1990_1999,2017.

[10] Z. Zhang, C. Li, B. B. Gupta, and D. Niu, ``Ef_cient compressedciphertext length scheme using multi-authority CP-ABE for hierarchicalattributes," *IEEE Access*, vol. 6, pp. 38273_38284, 2018. doi:10.1109/ACCESS.2018.2854600.