Chapter 1

INTRODUCTION

1.1 Introduction

With the significant increase in the volume, velocity, and variety of user 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.

In online social networks, social bots are social accounts controlled by automated programs that can perform corresponding operations based on a set of procedures. The increasing use of mobile 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 efficiency of collect

collecting and analyzing data from social networks.

1.2 Motivation

Social bots have been used to perform automated analytical services and provide users with improved quality of service. 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. 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. 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 networks. User behavior is the most direct manifestation of user intent, as different users have different habits, preferences, and online behavior like 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 a 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 a user's online behavior. In other words, user behavior is dynamic and its environment is constantly changing i.e., external observable environment like the environment and behavior of application context and the hidden environment in user information. 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. The aim of the paper is to detect malicious social bots on social network platforms in real-time, by proposing the transition probability features between user clickstreams based on the social situation analytics and designing an algorithm for detecting malicious social bots based on spatiotemporal features.

The difference in user behavior can be obtained, for example, by analyzing the image search logs of users to study the search intention of different users, and this approach can facilitate optimization of search engines. The user clickstream data can be used to construct a clickstream graph model to represent user behavior and identify different user groups, in order to detect malicious accounts. There have also been other researches that indicate user intent and abnormal accounts can be determined through behavior analysis, and social situation in facilitating the understanding of users' dynamic behavior. Constructed a new convolutional neural network architecture based on user behavior, search engine content and context information to construct a click model and find out the user's click preferences to improve search quality. Collection of a large amount of user information on the Twitter and YouTube, about 13 million channel activities, analyzing and detecting abnormal behaviors that deviate significantly from large-scale specifications through user behavior in two social networks.

1.3 Problem statement

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. In public opinion about social networks and massive user data can also be mined or disseminated for malicious or nefarious purpose. 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. For example, some fake social bots accounts created to imitate the profile of a normal

user, steal user data and compromise their privacy, disseminate malicious or fake information, malicious comment, promote or advance certain political or ideology agenda and propaganda, and influence the stock market and other societal and economical markets. Such activities can adversely impact the security and stability of social networking platforms. 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.

Chapter 2

LITERATURE SURVEY

2.1 Introduction

In [1] **''A new approach to bot detection: Striking the balance between precision and recall,'' in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining,**

San Francisco, CA, USA, Aug. 2016, by F. Morstatter, L. Wu, T. H. Nazer, K. M. Carley, and H. Liu

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.

In [2] **''Is that social bot behaving unethically?'' Commun. ACM, vol. 60, no. 9, pp. 29–31, Sep. 2017, C. A. De Lima Salge and N. Berente**

Attempting to answer the question posed by the title of this column requires us to reflect on moral goods and moral evils—on laws, duties, and norms, on actions and their consequences. In this Viewpoint, we draw on information systems ethics to present *Bot Ethics*, a procedure the general social media community can use to decide

whether the actions of social bots are unethical. We conclude with a consideration of culpability.

Social bots are computer algorithms in online social networks. They can share messages, upload pictures, and connect with many users on social media. Social bots are more common than people often think. Twitter has approximately 23 million of them, accounting for 8.5% of total users; and Facebook has an estimated 140 million social bots, which are between 5.5%−1.2% total users. Almost 27 million Instagram users (8.2%) are estimated to be social bots. LinkedIn and Tumblr also have significant social bot activity. Sometimes their activity on these networks can be innocuous or even beneficial. For example, *SF QuakeBot* performs a useful service by disseminating information about earthquakes, as they happen, in the San Francisco Bay area. However, in other situations, social bots can behave quite unethically.

In [3] **‘‘Detecting abnormal behavior in social network Websites by using a process mining technique,’’J. Comput. Sci., M. Sahlabadi, R. C. Muniyandi, and Z. Shukur**

Detecting abnormal user activity in social network websites could prevent from cyber-crime occurrence. The previous research focused on data mining while this research is based on user behavior process. In this study, the first step is defining a normal user behavioral pattern and the second step is detecting abnormal behavior. These two steps are applied on a case study that includes real and syntactic data sets to obtain more tangible results. The chosen technique used to define the pattern is process mining, which is an affordable, complete and noise-free event log. The proposed model discovers a normal behavior by genetic process mining technique and abnormal activities are detected by the fitness function, which is based on Petri Net rules. Although applying genetic mining is time consuming process, it can overcome the risks of noisy data and produces a comprehensive normal model in Petri net representation form.

In [4] **''Detecting social-network bots based on multiscale behavioral analysis,'' in Proc.7th Int. Conf. Emerg. Secur. Inf., Syst. Technol. (SECURWARE), Barcelona, Spain, F. Brito, I. Petiz, P. Salvador, A. Nogueira, and E. Rocha**

Together with the growing predominance of social networking services in human communication, a set of scam attacks perpetrated by bots, have emerged. We are currently witnessing a major increasing in cybercrime attacks against individuals targeting their personal data and financial assets. Most of the current Internet malware threats disseminate themselves using social engineering and, mainly, using social networks, since social networking provides an open field for illicit activities. Social networking sites are always improving their security but this is a constant race behind the leading criminals. Existing botnets can use social networking services to spread themselves, but more importantly, can use social networks to impersonate the owner of the controlled machine in order to obtain valuable personal information or force the person to interact with unwanted individuals or services. One of the better documented examples of social networking services abuse for malicious purposes was Koobface. Being currently the largest social networking service, Facebook is the main vector of attack via social networking services. Current techniques to detect bots within a social network rely on automated algorithms that evaluate social relations. Based on graph-theory techniques, they try to detect unnatural relations in social networks. Another technique used to detect bot activity measures mouse movements and keystrokes produced while interacting in the generation of online contents. In this class of behavioral analysis was applied in blogging activities, but it can also be easily applied to social networks interfaces. The main downside of this approach is that it must rely on software loaded on the client browser, which can be difficult to implement and certainly impossible to generalize to all users due to confidentiality constrains. A viable solution should only rely on ubiquitous statistics that do not compromise the users privacy, namely counting the number of social interactions per time interval.

It is extremely difficult to program a bot to replicate the characteristic human behavior of social interactions over time. Humans actions have an inherent pseudo-periodicity mixed with random (and sometimes chaotic) actions which are almost impossible to emulate/simulate. Nevertheless, this human uniqueness is very

easy to differentiate from other behavioral patterns. Therefore, in this paper, we propose a new methodology that, by jointly analyzing the multiple scales of the users' interactions within a social network, can accurately discriminate the characteristic behaviors of humans and bots within a social network. Consequently, different behavior signatures can be use to accurately detect bots acting with a social network. The proposed methodology applies the concept of multiscalling analysis based on scalograms to the statistical processes that describe the interaction of a user within the social network. Scalograms reveal much information about the nature of non-stationary processes that was previously hidden, so they are applied to a lot of different scientific areas: diagnosis of special events in structural behavior during earthquake excitation, ground motion analysis, transient building response to wind storms, analysis of bridge response due to vortex shedding, among others.

In [5] **''DeT.-K. Huang, M. S. Rahman, H. V. Madhyastha, M. Faloutsos, and B. 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.**

Online social networks (OSNs) have become a popular new vector for distributing malware and spam, which we refer to as socware. Unlike email spam, which is sent by spammers directly to intended victims, socware cascades through OSNs as compromised users spread it to their friends. In this paper, we analyze data from the walls of roughly 3 million Facebook users over five months, with the goal of developing a better understanding of socware cascades.

We study socware cascades to understand: (a) their spatio-temporal properties, (b) the underlying motivations and mechanisms, and (c) the social engineering tricks used to con users. First, we identify an evolving trend in which cascades appear to be throttling their rate of growth to evade detection, and thus, lasting longer. Second, our forensic investigation into the infrastructure that supports these cascades shows that, surprisingly, Facebook seems to be inadvertently enabling most cascades; 44% of cascades are disseminated via Facebook applications. At the same time, we observe large groups of synergistic Facebook apps (more than 144 groups of size 5 or more) that collaborate to support multiple cascades. Lastly, we find that hackers rely on two social engineering tricks in equal measure?luring users with free products and

appealing to users' social curiosity?to enable socware cascades. Our findings present several promising avenues towards reducing socware on Facebook, but also highlight associated challenges.

In [6]**‘‘A situational analytic method for user behavior pattern in multimedia social networks,’’ Z. Zhang, R. Sun, X. Wang, and C. Zhao.**

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.

Chapter 3

System Architecture

3.1 Proposed Model:

This section provides an overview for better detection of malicious social bots in online social networks. We analyze user behavior features and identify transition probability features between user click-streams Based on the transition probability features and time interval features, a semi-supervised social bots detection method based on space-time features is proposed.

SocialSitu(t) denotes the situational information at moment t. SocialSitu(t) is a four-tuple SocialSitu(t) = {ID, d, A, E}, where ID refers to the user's identity information (including the group to which the user belongs to and the role of user in the group), d refers to user's wishes at the t time, A refers to user operation corresponding to d at the particular moment (namely, behavior), and E refers to environmental information (e.g., terminal devices, equipment information and location information).

Clickstream is the order of clicks when users visit some websites or use the mobile terminals. The user's click event is a single point of operation. Click-stream is a series of point operations, and it refers to the SocialSitu(t) sequence of user from start point to target achievement. The sequence of the clickstream I = {SocialSitu(1), SocialSitu(2), . . . , SocialSitu(n)} , n ∈ N , SocialSitu(1) refers to the user's first click behavior on the platform. SocialSitu(n) refers to the last click event that the user performs prior to exiting the platform.

The collection of user clickstream consists of many clickstream sequences, the collection of clickstreams, namely Click(s) = {I (1), I (2), . . . , I (m)}, s refers to the user id.

Transition probability between click steams: P (i,j) represents the probability that the click event is j at t + 1 moment when the click event is i at t moment.

Transition probability P (i,j) refers to (1). Here, X (t) = i means the j total number of transitions that P may occur in click status i among all user click events. {X (t + 1) = j, X (t) = i} means the total number of click events when the click event i at t and click event j at t + 1.

$$P_{(i,j)} = \frac{\sum_{k} \{X(t+1) = j, X(t) = i\}}{\sum_{k} X(t) = i}, \quad k \in \mathrm{N} \qquad (1)$$

FEATURE SELECTION BASED ON TRANSITION PROBABILITY OF CLICKSTREAM SEQUENCES

The malicious behavior of social bots refers to a variety of behaviors performed by social bots for a specific purpose. However, the behaviors involved in this paper are not necessarily malicious behaviors, which are related operations that malicious users are most likely to perform for different social network platforms to achieve their goals. For example, social bots may achieve different purposes by performing the main function-related operations in Twitter, such as posting tweets, comments, forwarding tweets and so on. In the social networking platform, we usually determine whether the corresponding behavior is normal or malicious based on the final result of the user behavior. For instance, we determine whether a comment is malicious by analyzing whether the user's comment content contains ads. However, with the constant evolution of social bots, simple text analysis is difficult to detect comments because they can spread the message by posting images or more subtle text. As we all know, social bots achieve different purposes according to the main functions of the platform, and they perform different behaviors in different social networks. Therefore, in this paper, we focus on the operations related to the main functions of the experimental platform. These operations are not necessarily malicious, but are most likely to be performed by malicious social bots to meet different purposes. Malicious social bots search the Internet for information and picture to fill personal information and simulate the human time features in content production and consumption. The

user's profile picture and other personal data features, likes, comments, and some quantitative features are easily imitated by malicious social bots. Thus, the detection efficiency is also gradually reduced. To explore robust features, user behavior features should be deeply analyzed and expanded. The clickstream sequences can reflect the dynamic changes of the user behavior, while also hiding the important behavior features of the user. We get more information on the click behavior in three ways, namely: (1) In terms of user behavior data acquisition, we employ user clickstream sequences under situation aware environments, rather than simply click events. Social situation analytics can be used to acquire the external observable environment of applied scenarios and the hidden environment of user information in time. (2) In terms of user behavior features selection, we extend user behavior features from the single click behavior to the linear features of clickstream sequences, which can better reflect user intent in special situations. (3) In the dimension of user behavior features, we add temporal dimension features to the spatial dimension of user behavior features, and analyze user behavior features in multiple dimensions, which make user behavior features more robust. The differences between different users can be described by sequence analysis on user clickstream behavior. The transition probability between clickstreams is an important hidden feature in user clickstream sequences, which can reflect the user behavior habits and preferences in different situations. Compared with the quantitative feature, the transition probability features are more robust and not easily imitated. The malicious social bots can imitate the quantitative features of normal users by setting the number of related behaviors (e.g., the number of likes, the number of comments, the number of friends, etc.), and the user cannot observe the transition probability features on the online social network platform. It is also difficult for malicious social bots to mimic the features of normal users. Meanwhile, the transition probability between user click streams also cannot be obtained directly by querying the data in the database or the user's click stream log. Based on the function provided by the experimental platform, we find that the main function of the experimental platform is playing the video and the malicious social bot often use playback-related operations such as comments and likes to achieve their goals. The paper mainly focused on the following behaviors, such as playing, liking, sharing,

commenting, sharing and reporting, as well as their combinations based on user's click events and clickstream sequences. By analyzing clickstream sequences, we use the window sliding method to obtain the number of transitions between specific click events. At the same time, we choose the inter arrival times (IATs) between user-specific click events based on the time feature in the user's click behavior. Based on the playing behavior, we can get the difference between the normal user and social bot in the time and space dimension. For instance, the normal user prefers clicking the like or comment button after some time of watching the video, rather than clicking the comment button as soon as you click the play button.

CLASSIFICATION ALGORITHM OF MALICIOUS SOCIAL BOTS

Real-time detection of malicious social bots in online social platforms can detect and block social bots in a timely manner. We propose the detection method of malicious social bots based on semi-supervised clustering method, which can reduce the time of artificial marking, and the detection program can run periodically in the background of the website. Simultaneously, we choose the hybrid feature of transition probability features and time feature can be used to increase the robustness of the features, thus improving the accuracy of detection. In the meantime, the user's transition probability features and inter-arrival times can be obtained. We can analysis user behavior and social bots behavior based on features of temporal and spatial dimensions. Based on the constrained seed K-means algorithm [24], we set the sample mean square error threshold to determine the number of iterations, then obtain the social bots detection algorithm. The detection algorithm for malicious social bots is described in Algorithm 1.

Algorithm 1 The Detection Algorithm for Malicious Social Bots

Input: The log set of users' click event: DS, cluster number $k = 2$, a small number of labeled samples: $S = \bigcup_{j=1}^{k} S_j$, global threshold τ
Output: Normal user set, Social bots set
SocialBotsDection (DS, S)
1: Begin
2: for $s \leftarrow 1$ to n // s refers to the users id
3: $C_s = \{I(1), I(2), \ldots, I(m)\}$ // generate the user's intent sequence sets Cs
4: According to formula (1), Calculate the transition probability $P_{(play,like)}$, $P_{(play,feedback)}$, $P_{(play,comment)}$, $P_{(play,share)}$, $P_{(play,more)}$, $P_{(play,paly)}$
5: $\mathrm{IAT}(s) = \frac{\sum T((t-1)=play)-T(t)}{N(play)}$ // calculate the inter-arrival times
6: $x_S = \{P_{(play,like)}, P_{(play,feedback)}, P_{(play,comment)}, P_{(play,share)}, P_{(play,more)}, P_{(play,paly)}, \mathrm{IAT}(s)\}$ // generate the sets of transition probabilities and time feature
7: endfor
8: for $j \leftarrow 1$ to k
9: $\mu_j = \frac{1}{S_j} \sum_{x \in S_j} x_s$ // initialize the cluster center
10: endfor
11: repeat
12: $C_j = \varnothing \ (1 \le j \le k)$
13: for $j = 1, 2$
14: for all $x \in S_j$
15: $C_j = C_j \cup \{x_s\}$
16: endfor
17: endfor
18: for all $x_s \in D \backslash S$
19: Calculate the distance from the sample x_s to mean vectors $\mu_s (1 \le j \le k)$: $d_{sj} = \|x_s - \mu_s\|_2$
20: Find out the nearest cluster to sample x_s: $r = \arg\min_{j \in \{1,2\}} d_{sj}$
21: $C_r = C_r \cup \{x_s\}$ // divide sample x_s into the corresponding cluster
22: $M = \frac{1}{n} \sum_{s=1}^{n} (x_s - \mu_s)^2$ // calculate mean square error
23: $M_t = M$
24: endfor
25: until $M - M_t < \tau$
26: End

Fig. 3.1 Acquisition process of user clickstream behavior

Chapter 4

Experiments

4.1 Result Analysis

DATA COLLECTION

The experimental platform in this study is the online media social network platform CyVOD [25]. CyVOD is an Internet plus technology information service application platform that integrates science and technology policy, scientific and technological achievements, and technology and social interaction. The CyVOD platform comprises the website platform and Android and iOS applications. On CyVOD, the user clickstream behavior is obtained by a data burying point, and user clickstream data is collected server-side. In the realistic environment, for your own website, you can use the buried technology to get the corresponding data; for other websites, you need to get the data by working with the website or by calling the corresponding API (if provided). The acquisition of user's action in our own website is shown in Figure 1. You can pass the corresponding buried data to the server and record it in the server through the code when you manipulate some controls of the UI layer. In the real social network platform, many platforms use the burying technology to obtain the user's behavior data. In the research, many scholars choose to cooperate with the social platform or call the corresponding API of the social platform to obtain data.

EXPERIMENTAL DESIGN

A total of 1500 malicious social bots accounts on the CyVOD platform are assigned different tasks, including malicious social bots that perform a single task, malicious social bots that coordinate to perform tasks, and malicious social bots that perform mixed tasks. For example, a user can perform two or more actions in the actions of liking, comment, sharing and so on. The social bot for malicious likes, the value of

the P (play,like) (the transition probability of ‘‘the current click event is and the next click event is liking’’) would be high and the value of other transition probability features would be small or zero. The mixed social bot, the values of six transition probability features maybe average, which looks like normal user. In this experiment, four malicious social bots that perform different specific purposes and two malicious social bots with mixed behavior are set up. Malicious social bots are classified as shown in Table 1. We designed an Android application called ‘‘SocialBot’’ to simulate the behavior of social bots. According to the functional characteristics of the experimental platform, we designed such seven categories of social bot that perform different tasks. The social bot program can be activated automatically or manually by clicking these buttons.

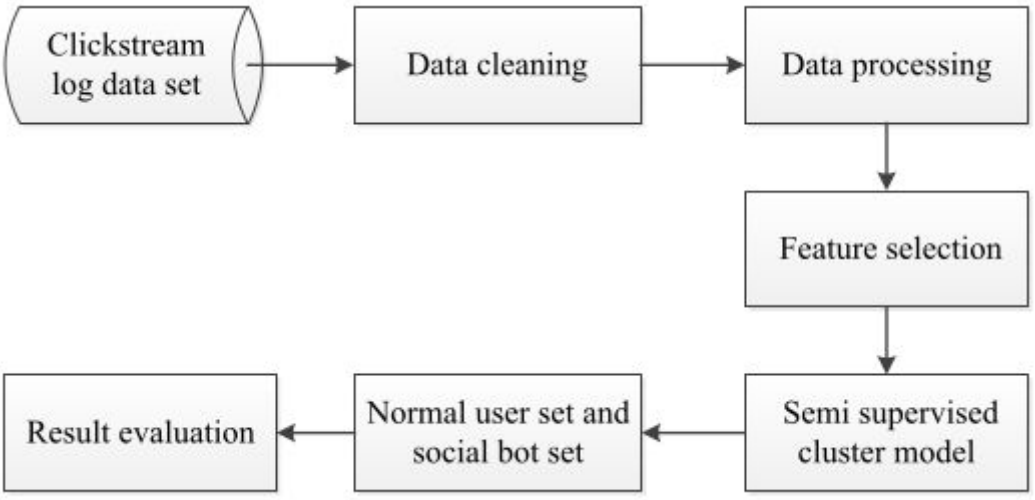

Fig. 4.1 Experimental procedure

In this paper, the corresponding user data from the CyVOD web platform and Android client are collected. The iOS client is discarded because of its lower amount of data compared with the other two. To protect user privacy, users are assigned a unique and anonymous id. The main function of the platform is audio and video playback. The experimental features are selected around the main features of the platform. The accuracy of the detection method for malicious social bots proposed is compared and analyzed by acquiring the corresponding click times of different categories of malicious social bots and transition probability features between click-streams. In this paper, three categories of features are selected to train the

proposed semi-supervised detection method and detect its classification. The comparison of three categories

of features is shown in Table 4.1.

Categories	Corresponding Characteristics
Quantitative feature	Times of play, number of comments, point of praise, number of report, and times of sharing
Transition probability feature	$P_{(play,\ play)}$, $P_{(play,\ like)}$, $P_{(play,\ feedback)}$, $P_{(play,\ comment)}$, $P_{(play,\ share)}$, $P_{(play,\ more)}$
Mixed feature	$P_{(play,\ play)}$, $P_{(play,\ like)}$, $P_{(play,\ feedback)}$, $P_{(play,\ comment)}$, $P_{(play,\ share)}$, $P_{(play,\ more)}$, IATs (inter-arrival times)

Table 4.1. Comparison of three categories of characteristics

To verify the effectiveness of the proposed features, different types of malicious social bots are modeled using three categories of features to obtain precise detection. The precision of detection for different types of malicious social bots by using three categories of features is shown in Figure 3. The recall of detection for different types of malicious social bots by using three categories of features is shown in Figure 4. We find that (1) the precision of the semi-supervised clustering method for the detection of the same type of malicious social bots based on transition the probability features and mixed features is higher than that of the semi-supervised clustering method based on the quantitative feature; (2) for simple malicious social bots, the application of transition probability feature and mixed feature can effectively detect malicious social bots accounts. However, for malicious social bots with mixed feature and random behavior, the application of mixed feature can obtain better results. The experimental

result shows that the precision of a semi-supervised clustering method based on mixed features for detecting malicious social bots with mixed malicious feature can be as high as 93.1%, the recall rate is 97.5%, and the F 1 Score is 95.2%. Compared with the semi-supervised clustering method with quantitative features, our method can detect malicious social bots accounts in online social platforms more accurately.

To verify the accuracy of the method, the support vector machine model based on transition probability, the semi-supervised clustering method based on mixed feature, the semi-supervised clustering method based on transition probability, and the semi-supervised clustering method based on quantitative feature are established in the same data set. The detection accuracy of different methods for malicious social bots are shown in Figure 5. The experiment proves that the proposed semi-supervised clustering method based on transition probability between user clickstreams can effectively detect malicious social bots in online social platforms. The comparison between different methods shows that the precision and accuracy of the detection method of malicious social bots based on transition probability can reach 95% or higher. Compared with the traditional detection method based on the quantitative feature, accuracy is improved by 12.8% 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.

Conclusion

The paper proposes a novel method to accurately detect malicious social bots in online social networks. Experiments showed that transition probability between user clickstreams based on the social situation analytics can be used to detect malicious social bots in online social platforms accurately. 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.