Cyber – Security Project Documentation

Topic – Detecting Malicious Social Bots based on click stream sequence

Submitted by

1) **P. Sai Tarrun , 2nd year**

Computer Science and Engineering

Gitam University, Visakhapatnam,
Andhra Pradesh

2) **E.V.K Srikar , 2nd year ECE**
Electronics and communication Engineering

Gitam University, Visakhapatnam

Andhra Pradesh

Table of Content

1)Abstract

With the significant and rapid increase in the volume and velocity of data in social network, there have been significant changes in the methodology of collecting and analysing data.

The current social networks are using various designs for collecting large data. Social bots are one of the current ways used to collect user information. However malicious social bots have also been used to disseminate false information and this can result in real-world consequences.

Therefore, detecting social bots is one of the important tasks to be done. A novel method of detecting malicious social bots featuring both selections based on transition probability of click stream sequences and semi-supervised clustering.

2)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. 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 analysing data from social network services.

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.

User behaviour is the most direct manifestation of user intent, as different users have different habits, preferences, and online behaviour (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 analyse information hidden in user's online behaviour 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 behaviour. In other words, user behaviour is dynamic and its environment is constantly changing – i.e., external observable environment (e.g., environment and behaviour) 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 analyse social situation of user behaviour and compare and understand the differences of malicious social bots and normal users in dynamic behaviour

Social bots

A **social bot** (also: ***socialbot*** or ***socbot***) is an agent that communicates more or less autonomously on social media, often with the task of influencing the course of discussion and/or the opinions of its readers. It is related to chatbots but mostly only uses rather simple interactions or no reactivity at all. The messages (e.g. tweets) it distributes are mostly either very simple, or prefabricated (by humans), and it often operates in groups and various configurations of partial human control (hybrid). It usually targets advocating certain ideas, supporting campaigns, or aggregating other sources either by acting as a "follower" and/or gathering followers itself. In this very limited respect, social bots can be said to have passed the Turing test If the expectation is that behind every social media profile there should be a human, social bots always use fake accounts. This is not different from other social media API uses.

3)Procedure

3.1)Behavioural Analysis of malicious social bots

- Malicious users in social network platforms are likely to exhibit behaviour patterns that different from normal users, because their goals in maximizing their own needs and purposes (e.g., promote a certain product or certain political beliefs or ideology).
- User behaviour analysis is not only helpful in gaining an in-depth understanding of user intent, but it is also important to the detection of malicious social bots' accounts in online social networks.
- User behaviour likely change under different situations.

USING SITUATION ANALYTICS IN SOFTWARE SERVICE REQUIREMENT ANALYSIS:

- Chang proposed this Idea which can facilitate the analysis of any change in user's requirements.
- Such an analysis is useful to understand the dynamic needs of a software service environment.

USING A FRAMEWORK DESIGNED FOR TO DICOVER THE USER BEHAVIOUR PATTERN ON ONLINE SOCIAL NETWORKS:

- Zhang *et al.* presented a framework to the discovery of user behaviour pattern in multimedia video recommendation services on online social networks.
- Their framework is based on social context and analyses the changes in user need for different social situations.
- Such user behaviour data can be obtained if we have access to the user's logs or user's clickstreams (e.g., recorded by social network platforms).
- The difference in user behaviour can be obtained, for example, by analysing the image search logs of users to study the search intention of different users, and this approach can facilitate optimization of search engines.

CONSTRUCTING A CLICKSTREAM GRAPH FROM USED SS CLICKSTREAM DATA:

- Wang *et al.* used SS clickstream data to construct a clickstream graph model to represent user behaviour 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 behaviour analysis, and social situation in facilitating the understanding of users' dynamic behaviour.

USING A CLICK MODEL BY CONSTRUCTING A CONVOLUTIONAL NUERAL NETWORK:

- Liu *et al.* constructed a new convolutional neural network architecture based on user behaviour, search engine content and context information to construct a click model and find out the user's click preferences to improve search quality.

ANALYSING ABNORMAL BEHAVIOURS THAT DIFFER FROM LARGE SCALE SPECIFICATIONS:

- Al-Qureshi *et al.* collected a large amount of user information on the Twitter and YouTube, about 13 million channel activities, analysing and detecting abnormal behaviours that deviate significantly from large-scale specifications through user behaviour in two social networks.

3.2) Social Bots Detection:

The first generation of bots could sometimes be distinguished from real users by their often superhuman capacities to post messages around the clock (and at massive rates). Later developments have succeeded in imprinting more "human" activity and behavioural patterns in the agent. To unambiguously detect social bots as what they are, a variety of criteria must be applied together using pattern detection techniques, some of which are:

cartoon figures as user pictures

- sometimes also random real user pictures are captured (identity fraud)
- reposting rate
- temporal patterns
- sentiment expression
- followers-to-friends ratio
- length of user names
- variability in (re)posted messages
-

Botometer (formerly BotOrNot) is a public Web service that checks the activity of a Twitter account and gives it a score based on how likely the account is to be a bot. The system leverages over a thousand features. An active method that worked well in detecting early spam bots was to set up honeypot accounts where obvious

nonsensical content was posted and then dumbly reposted (retweeted) by bots. However, recent studies show that bots evolve quickly and detection methods have to be updated constantly, because otherwise they may get useless after a few years.

Botnets become widespread in wired and wireless networks. In particular, bots in a botnet are able to cooperate towards a common malicious purpose. In recent years, social bots have become very popular in social networks, and they can imitate human activities in social networks. They are also programmed to work together to fulfil the prescribed tasks. For example, in order to imitate the features of human users successfully, social bots may 'crawl' for words and pictures from online social networks to complete fabricated user profiles and so on.

Semi-social bots between humans and social bots have also reportedly emerged in social networks, which are highly complex social bots that bear the characteristics of human behaviour and social bot behaviour. The automated procedure for semi-social bot is generally activated by humans, and the subsequent actions are automatically performed by social bots. This process further increases the uncertainty of the operation time of social bots . Social bots are generally more intelligent and they can more easily imitate human behaviour, and they cannot be easily detected.

In existing literature, social bots are generally detected using machine learning-based approaches, such as BotOrNot released by the Twitter in 2014. In BotOrNot, the random forest model is used in both training and analysis by using historical social information of normal users and social bots accounts. Based on six features (i.e. network, user, making friends, time, content and emotion), this model distinguished normal users from social bots.

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. A time act model, namely, Act-M, was constructed focusing on the timing of user behaviour activities , which can be used to accurately determine the interval between different behaviours of social media users to accurately detect malicious users. There have been focused on detecting semi-social bots too. The supervised learning method can be effective in detecting social bots, however 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 labelling 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.

We also observe that social bots usually have similar features and same purpose. Unsupervised clustering algorithms can classify users into different clusters based on the similarity of users. To identify potential malicious social bots in online social networks in real-time, we analyse the social situation behaviour of users in online social networks. We also evaluate user behaviour features and select the transition

probability of user behaviour on the basis of general behaviour characteristics. We then analyse and classify situation aware user behaviours 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' clear applications.

3.3) Proposed method for detection of malicious social bots

FEATURE SELECTION BASED ON TRANSITION PROBABILITY OF CLICKSTREAM SEQUENCES:

The malicious behaviour of social bots refers to a variety of behaviours performed by social bots for a specific purpose. However, the behaviours involved in this are not necessarily malicious behaviours, 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 behaviour is normal or malicious based on the final result of the user behaviour. For instance, we determine whether a comment is malicious by analysing 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 behaviours 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.

The clickstream sequences can reflect the dynamic changes of the user behaviour, while also hiding the important behaviour features of the user. We get more information on the click behaviour in three ways, namely:

(1) In terms of user behaviour 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 behaviour features selection, we extend user behaviour features from the single click behaviour to the linear features of clickstream sequences, which can better reflect user intent in special situations.

(3) In the dimension of user behaviour features, we add temporal dimension features to the spatial dimension of user behaviour features, and analyse user behaviour features in multiple dimensions, which make user behaviour features more robust.

3.4) 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 behaviour and social bot's behaviour based on features of temporal and spatial dimensions. Based on the constrained seed K-means algorithm, we set the sample mean square error threshold to determine the number of iterations, then obtain the social bots detection algorithm

3.5) MALICIOUS SOCIAL BOTS DETECTION

Data set cleaning and screening, data feature processing, data classification, and a series of operations were conducted after acquiring clickstream data set of the user

1) Data cleaning: data that are clicked less must be cleaned to remove wrong data, obtain accurate transition probability between clickstreams, and avoid the error of transition probability caused by fewer data.

2) Data processing: some data are selected randomly from the normal user set and social bots set to the label.

3) Feature selection: in the spatial dimension: according to the main functions of the CyVOD platform, we select the transition probability features related to the playback function: P(play, play) , P(play, like) , P(play, feedback) , P(play, comment) , P(play, share) and P(play, more) ; in the time dimension: we can get the inter-arrival times (IATs). Because if all transition probability matrixes of user behaviour are constructed, extremely huge data size and sparse matrix can increase the difficulty of data detection.

4) Semi-supervised clustering method: first, the initial centres of two clusters are determined by labelled seed users. Then, unlabelled data are used to iterate and optimize the clustering results constantly.

5) Obtain the normal user set and social bots set: the normal user set and social bots set can be finally obtained by detecting.

6) Result evaluation: we evaluate results based on three different metrics: Precision, Recall, and F1 Score (F1 is the harmonic average of Precision and Recall, $F1 = 2 \cdot$ Precision /Recall Precision+Recall). In the meantime, we use Accuracy as a metric and compare it 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.

The corresponding user data from the CyVOD web platform and Android client are collected. To protect user privacy, users are assigned a unique and anonymous id. 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 analysed by acquiring the corresponding click times of different categories of malicious social bots and transition probability features between clickstreams. Three categories of features are selected to train the proposed semi-supervised detection method and detect its classification. Precision of detection methods based on different features for different types of malicious social bots. To verify the effectiveness of the proposed features, different types of malicious social bots are modelled using three categories of features to obtain precise detection. 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.

This method can effectively detect malicious accounts on social platforms. Finally, the malicious social bot's 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.

4) Conclusion

The above methods are a novel method to accurately detect malicious social bots in online social networks. The Transition probability between user clickstreams based on the social situation analytics can be used to detect malicious social bots .

5)References

1) https://en.wikipedia.org/wiki/Social_bot#:~:text=A%20social%20bot%20(also%3A%20socialbot,or%20no%20reactivity%20at%20all.
2) https://searchcustomerexperience.techtarget.com/definition/clickstream-analysis-clickstream-analytics
3) https://personal.utdallas.edu/~jjue/cs6352/markov/node3.html
4) https://en.wikipedia.org/wiki/K-means_clustering