International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at : www.ijsrcseit.com

Volume 10, Issue 7, May-June-2024 | Published Online : 20th June 2024

Detecting Malicious Social Bot In Social Media

G. Rajalakshmi

Assistant Professor, Department of Information Technology, Sethu Institute of Technology, India

ABSTRACT

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 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. This method not only analyzes transition probability of user behavior clickstreams but also considers the time feature of behavior. The proposed system introducing Text Classifier with User Wall Filter Algorithm to improving malicious bots detection at social communications.

INTRODUCTION

Online social networks (OSNs) are popular platforms that connect users all over the globe. A botnet represents a group of agents (bots) that are managed and programmed to act in an organized manner. The term social botnet refers to a new generation of botnets that utilize OSNs as command and control (C&C) channels with minimal noise (Burghouwt et al., 2013). 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 analyzing data from social network services. However, 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 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.

In previous work, various methods were used to protect the security of online social network. User behavior is the most direct manifestation of user intent, as different users have different habits, preferences,

and online behavior (e.g., the 1way 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. 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.

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, 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.

LITERATURE SURVEY

1.1 Content-Based Book Recommending Using Learning for Text Categorization

Raymond J. Mooney(1999) Recommender systems improve access to relevant products and information by making personalized suggestions based on previous examples of a user's likes and dislikes. Most existing recommender systems use social filtering methods that base recommendations on other users' preferences. By contrast, content-based methods use information about an item itself to make suggestions. This approach has the advantage of being able to recommended previously unrated items to users with unique interests and to provide explanations for its recommendations. We describe a content- based book recommending system that utilizes information extraction and a machine learning algorithm for text categorization. Initial experimental results demonstrate that this approach can produce accurate recommendations. These experiments are based on ratings from random sampling of items and we discuss problems with previous experiments that employ skewed samples of user selected examples to evaluate performance.

1.2 Machine Learning in Automated Text Categorization FABRIZIO SEBASTIANI(2002) The automated categorization (or classification) of texts into predefined categories has witnessed a booming interest in the last 10 years, due to the increased availability of documents in digital form and the ensuing need to organize them. In the research community the dominant approach to this problem is based on machine learning techniques: a general inductive process automatically builds a classifier by learning, from a set of preclassified documents, the characteristics of the categories. The advantages of this approach over the knowledge engineering approach (consisting in the manual definition of a classifier by domain experts) are a very good effectiveness, considerable savings in terms of expert labor power, and 4 straightforward portability to different domains. This survey discusses the main approaches to text categorization that fall within the machine learning paradigm. We will discuss in detail issues pertaining to three different problems, namely, document representation, classifier construction, and classifier evaluation.

1.3 Extracting Crime Information from Online Newspaper Articles

Rexy Arulanandam (2014) Information extraction is the task of extracting relevant information from unstructured data. This paper aims to 'mine' (or extract) crime information from online newspaper articles and make this information available to the public. Baring few, many countries that possess this information do not make them available to their citizens. So, this paper focuses on automatic extraction of public yet 'hidden' information available in newspaper articles and make it available to the general public. In order to demonstrate the feasibility of such an approach, this paper focuses on one type of crime, the theft crime. This work demonstrates how theft-related information can be extracted from newspaper articles from three different countries. The system employs Named Entity Recognition (NER) algorithms to identify locations in sentences. However, not all the locations reported in the article are crime locations. So, it employs Conditional Random Field (CRF), a machine learning approach to classify whether a sentence in an article is a crime location sentence or not. This work compares the performance of four different NERs in the context of identifying locations and their subsequent impact in classifying a sentence as a 'crime location' sentence. It investigates whether a CRF-based classifier model that is trained to identify crime locations from a set of articles can be used to identify articles from another newspaper in the same country (New Zealand).

1.4 A rule-based message filtering system

Stephen Pollock(1988) Much computerized support for knowledge workers has consisted of tools to handle low-level functions such as distribution, storage, and retrieval of information. However, the higher level processes of making decisions 5 and taking actions with respect to this information have not been supported to the same degree. This paper describes the ISCREEN prototype system for screening text messages. ISCREEN includes a high-level interface for users to define rules, a component that screens text messages, and a conflict detection component that examines rules for inconsistencies. An explanation component uses text generation to answer user queries about past or potential system actions based on Grice's conversational maxims.

1.5 Short Text Classification in Twitter to Improve Information Filtering Bharath Sriram (2010) In micro-blogging services such as Twitter, the users may become overwhelmed by the raw data. One solution to this problem is the classification of short text messages. As short texts do not provide sufficient word occurrences, traditional classification methods such as "Bag- Of-Words" have limitations. To address this problem, we propose to use a small set of domain specific features extracted from the author's profile and text. The proposed approach effectively classifies the text to a predefined set of generic classes such as News, Events, Opinions, Deals, and Private Messages.

SYSTEM DESIGN

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. 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 spatio temporal features.

PROPOSED SYSTEM

The Proposed system aim evaluate an automated system, called Filtered Wall(FW), able to filter unwanted messages from OSN user walls. User Wall Filter can classify a variety of different filtering criteria that can be combined and customized according to the user needs. In addition, the system provides the support for user-defined Block lists, that is, lists of users that are temporarily prevented to post any kind of messages on a user wall.

Advantages of Proposed System:

1. Avoid unwanted message from unwanted users.
2. User can able to block a particular user for posting and sending messages.

3.1 COMPARISION WITH EXISTING SYSTEM

In Existing System web server provide very little support to prevent unwanted messages on user walls. For example, Facebook allows users to state who is allowed to insert messages in their walls (i.e., friends, friends of friends, or defined groups of friends). However, no content- based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, etc.

Disadvantages of Existing System:

1. User cannot control the unwanted message from the unknown persons.
2. Unwanted message will come regularly to the users, user will delete the unwanted message.

3.2 DESIGN OF THE PROPOSED SYSTEM

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 8 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.

3.3 ALGORITHMS

4.3.1 TEXT CLASSIFICATION

In content based filtering system content are matched with user profile using information retrieval techniques. Term Frequency and Inverse Document frequency (TF-IDF) document filtered in content based filtering same as text classification. In this documents are classified into relevant and non-relevant categories.

4.3.2 USER-WALL FILTER ALGORITHM: Filtering Process:

To define the language for FR specification, many problems are considered. First issue may arise when the message with different meaning and significance based on who writes it. Filtering rules will be applied ,when a user profile does not hold 9 value for attributes submitted by a FR. This type of situation will be handled by with asking the user to choose whether to block or notify the messages initiating from the profile which does not match with the wall owners FRs, due to missing of attributes. Makers on which a FR applies can be chosen on the premise of a few criteria; a standout amongst the most significant is by forcing conditions on their profile's attribute. For example, conceivable to characterize rules applying just to youthful makers or to inventors with a given religious/political perspective Blacklisting Process:

To improve the effectiveness of the system such information are given to the system through a set of rules called as BL rules. A further segment of our system is a Blacklist (BL) system. It is used to avoid messages to be posted from undesired makers, free from their substance. The user should define his own rules and depending upon that it is illustrated by the system and if that user crosses that limit then he may be blocked or blacklisted. BL is is clearly administered by the by the system, the system has the ability to figure out, who are the users to be added in the BL and choose when user's maintenance in the BL is done. To improve adaptability, such type of data are given to the system through an arrangement of instructions, for this purpose the rules are used are called as BL rules. Maybe, we choose to let the clients themselves, i.e., the wall proprietors to indicate BL principles directing who must be banned from their wall and for to what extent, means that particular user decides which user should be banned from posting on his wall. Consequently, when a user gets blocked then he is no more able to post messages on that particular users wall until that particular user unblocks him

Fig.1 Architectural flow of social media

4.3.3 WORKING PRINCIPLES OF THE PROPOSED SYSTEM

REGISTRATION:

This module is to who want to join the application, they must register and get unique login id and password. The registered users logged into application and establish the communication between users along this application.

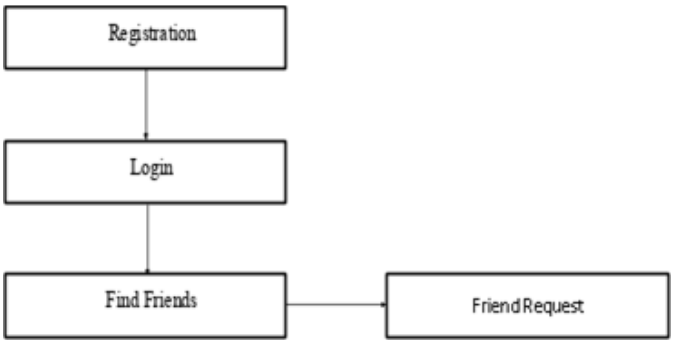

Fig.2 Registration

SHARE COMMUNICATION

The share communication module is used share the communication between users. The each and every communication will be validated by BOT Authentication. The authentication having threshold limit, if the user keyword exceed threshold limit immediately user communication will be blocked through BOT Authentication.

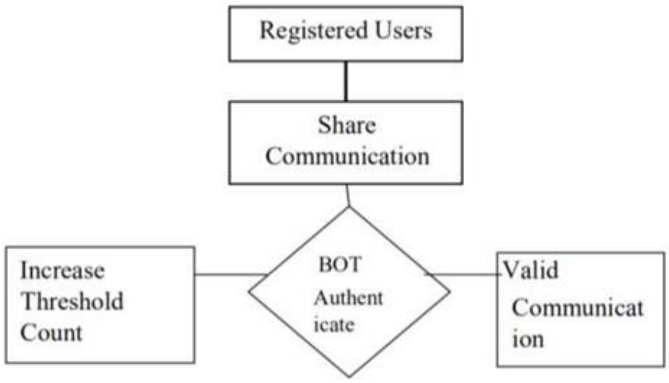

Fig.3 Share Communication

BOT REGISTRATION AND AUTHORIZATION:

BOT Registration and authorization is used to register and authorized user communication for who are administrators of social network. The process of BOT Registration is used to register BOT keyword and also register threshold count. The registered BOT keywords are parsed user communication and create separate count of matched malicious keyword for each and every users. The user count will be reach threshold count, the user communication are blocked through BOT authorization approaches.

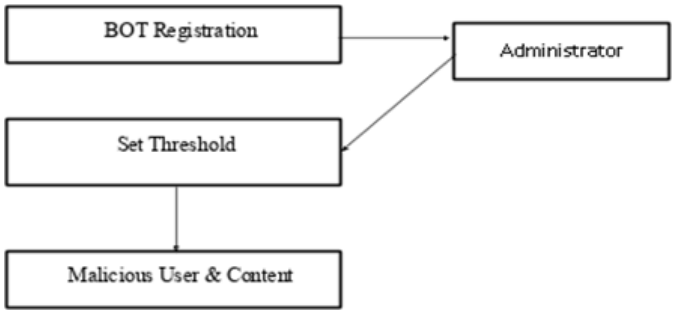

Fig.4 Registration & Authorization

4.3.4 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.

The detailed steps are shown.

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. Normal user account is labeled as 1, and the bots account is labeled. Seed users are classified as the category of clusters.

3) Feature selection: in the spatial dimension: according to the main functions of the .NET 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 matrices of user behavior are constructed, extremely huge data size and sparse matrix can increase the difficulty of data detection.

4) Semi-supervised clustering method: first, the initial centers of two clusters are determined by labeled seed users. Then, unlabeled 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 D 2_Precision_Recall PrecisionCRecall). In the meantime, we use Accuracy as a metric and compare it with the

SVM algorithm to verify the efficiency of the method.

4. RESULTS AND DISCUSSION

Implementation is the state in the system where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that will work efficiently and effectively. The system can be implemented only after thorough testing in done and if found to work according to the specification. If involves careful planning, investigation of the current system and its constraints on implementation, design of methods to achieve the

changeover, an evaluation of changeover methods apart from planning. Two major tasks of preparing the implementation are education, training of the users and testing the systems. System analysis and design efforts will be more for complex systems beings implemented. Based on policies of individuals organization an implementation coordinating committee has been appointed. The implementation process begins with preparing a plan for the implementation system. According to this plan, the other activities are to be carried out. In this plan, discussion has been made regarding the equipment, resources and how to test the activities. Thus a clear plan is preparing for the activities.

Fig. 5 : Home Page

Fig.6 User Registration Page

Fig.7 Admin Login Page

Fig.8 Administrator Service Page

Fig.9 Bot Registration Page

Fig.10 Threshold Setting Page

Fig.11 User Login Page

Fig.12 User Message Posting Page

Fig 13. Timeline Page

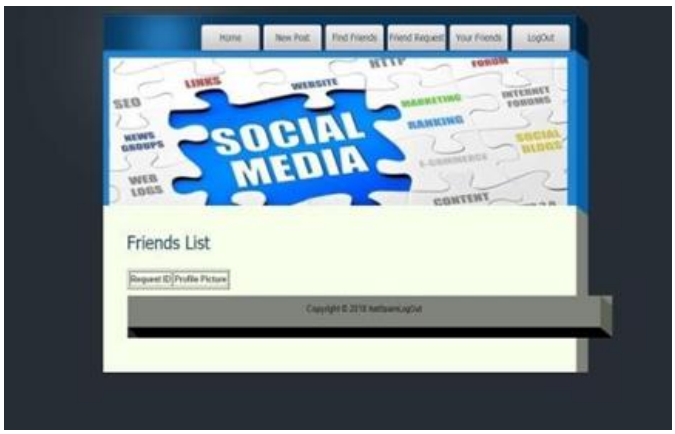

Fig.14 Friend Request Page

Fig.15 Identifying Bot User

Fig.16 Friend Request

CONCLUSION

The system proposed 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.

REFERENCES

[1]. Web Site: WWW.msdn.Microsoft.com & www.a1vbcode.com

[2]. Stephen Walther, ASP.NET Unleashed", Second Edition, Sams Publishing, July 18, 2003