IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

NCAIT 2017

JSS Academy of Technical Education

Vol. 4, Special Issue 8, May 2017

Detecting Online Social Behaviour of Compromised Account

Ambikesh[1], Himansu Singh[2], Kiran B.V[3]

8th Semester B.E, Information Science & Engineering, SJB Institute of Technology, Bengaluru, Karnataka[1,2,3]

Abstract: Compromisation of online social network is a threat for many of us who are a part in OSN. Where many spammers establish and achieve our trust of friends and success in sending malicious spams and try to hack our account. In this paper, our goal is to analysis the social behaviour of such attackers and user, by usage of OSN services. We propose a set of social behaviouralfeatures that can effectively characterize the user social activitieson OSNs. We validate the efficacy of these behavioural features by collecting and analysing real user clickstreams to anOSN website. Based on our measurement study, we devise individual user's social behavioural profile by combining its respectivebehavioural feature metrics. A social behavioural profile accuratelyreflects a user's OSN activity patterns. While an authentic ownerconforms to its account's social behavioural profile involuntarily,it is hard and costly for impostors to feign. We evaluate thecapability of the social behavioural profiles in distinguishingdifferent OSN users, and our experimental results show thesocialbehavioural profiles can accurately differentiate individualOSN users and detect compromised accounts.

Index Terms: Online social behavior, privacy, data analysis,compromised accounts detection.

I. INTRODUCTION

Compromised accountsIn On lineSocialNetworks(OSNs) are more favourablethanSybil accounts to spammers and other malicious OSN attackers. Malicious parties exploit the well-established connections and trust relationships between the legitimate account owners and their friends, and efficiently distribute spam ads, phishing links, or malware, while avoiding being blocked by the service providers. Offline analyses of tweets and Facebook posts reveal that most spam are distributed via compromised accounts, instead of dedicated spam accounts.Recent large-scale account hacking incidents in popular OSNs further evidence this trend.

Unlike dedicated spam or sybil accounts, which are created solely to serve malicious purposes, compromised accounts are originally possessed by benign users, While dedicated malicious accounts can be simply banned or removed upon detection, compromised accounts cannot be handled likewise due to potential negative impact to normal user experience (e.g., those accounts may still be actively used by their legitimate benign owners). Major OSNs today employ IP geo-location logging to battle against account compromisation. However, this approach is known to suffer from low detection granularity and high false positive rate.Previous research on spamming account detection mostly cannot distinguish compromised accounts from sybil accounts, with only one recent study by Egeleetal.features compromised accounts detection. Existing approaches involve account profile analysis and message content analysis(e.g. embedded URL analysis and message clustering). However, account profile analysis is hardly applicable for detecting compromised accounts, because their profiles are the original common users' information which is likely to remain intact by spammers. URL blacklisting has the challenge of timely maintenance and update, and message clustering introduces significant overhead when subjected to a large number of real-time messages.

Instead of analysing user profile contents or message contents, we seek to uncover the behavioural anomaly of compromised accounts by using their legitimate owners' history social activity patterns, which can be observed in a lightweight manner. To better serve users' various social communication needs, OSNs provide a great variety of online features for their users to engage in, such as building connections, sending mes-sages, uploading photos, browsing friends' latest updates, etc.

However, how a user involves in each activity is completely driven by personal interests and social habits. As a result, the interaction patterns with a number of OSN activities tend to be divergent across a large set of users. While a user tends to conform to its social patterns, a hacker of the user account who knows little about the user's behavior habit is likely to diverge from the patterns. To validate the effectiveness of social behavioral profile in detecting account activity anomaly, we apply the social behavioral profile of each user to differentiate clickstreams of its respective user from all other users. We conduct multiple cross-validation experiments, each with varying amount of input data for building social behavioral profiles. Our evaluation results show that social behavioral profile can effectively differentiate individual OSN users with accuracy up to 92.6%, and the more active a user, the more accurate the detection.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

NCAIT 2017

JSS Academy of Technical Education
Vol. 4, Special Issue 8, May 2017

II. RELATED WORK

Schneider etal.andBenevenutoet al. measured OSN users' behaviors based on network traffic collected from ISPs. Both works analyze the popularity of OSN services, session length distributions, and user click sequences among OSN services, and discover that browsing accounts for a majority of users' activities. Benevenutoet al. further explored user interactions with friends and other users multiple hops away. While these works primarily emphasize the overall user OSN service usage, and aim to uncover general knowledge on how OSNs are used, this paper studies users' social behavior characteristics for a very different purpose. We investigate the characterization of individual user's social behaviors to detect account usage anomaly. Moreover, we propose several new user behavioral features and perform measurement study at a fine granularity. Viswanathet al. also aim to detect abnormal user behaviors in Facebook, but they soly focus on "like" behaviors to detect spammers.While most previous research on malicious account detection cannot differentiate compromised accounts from spam accounts, Egeleet al.specifically studied the detection of compromised accounts.

By recording a user's message posting features, such as timing, topics and correla-tion with friends, they detected irregular posting behaviors; on the other hand, all messages in a certain duration are clustered based on the content, and the clusters in which most mes-sages are posted by irregular behaviors are classified as from compromised accounts. While they also leveraged certain user behavior features to discern abnormality, we use a different and more complete set of metrics to characterize users' general online social behaviors, instead of solely focusing on message posting behaviors. Additionally, our technique does not rely on deep inspection and classification of message contents and avoids the heavy weight processing.

Wanget al. proposed an approach for sybil account detection by analyzing clickstreams. They differentiated sybil and common users' clicks based on inter-arrival time and click sequence, and found that considering both factors leads to better detection results. Since sybils are specialized fake identi-ties owned by attackers, their clickstream patterns significantly differ from those of normal users. However, for compromised accounts, their clickstreams can be a mix from normal users and spammers, As a result, methods incannot handle compromised accounts well. In contrast, this paper aims to uncover users' social behavior patterns and habits from the clickstreams, with which we can perform accurate and delicate detection on behavioral deviation.

III. USER SOCIAL BEHAVIORS STUDY

In this section, we first propose several social behavior features on OSNs, and describe in detail how they can reflect user social interaction differences. Then, we present a measurement study on user behavior diversity by analyzing real user clickstreams of a well known OSN, Facebook, with respect to our proposed features.

A. Social Behavior Features

We categorize user social behaviors on an OSN into two classes, extroversive behaviors and introversive behaviors. Extroversive behaviors, such as uploading photos and sending messages, result in visible imprints to one or more peer users; introversive behaviors, such as browsing other users' profiles and searching in message inbox, however, do not produce observable effects to other users. While most previous research only focus on the extroversive behaviors, such as public posting [8], we study both classes of behaviors for a more complete understanding and characterization of user social behaviors.

1) Extroversive Behavior Features: Extroversive Behaviorsdirectly reflect how a user interacts with its friends online, and thus they are important for characterizing a user's social behaviors. We specify extroversive behaviors on the following four major aspects.

◦ First Activity:

The first extroversive activity a user engages in after logging in an OSN session can be habitual. Some users often start from commenting on friends' new updates; while some others are more inclined to update their own status first. The first activity feature aims to capture a user's habitual action at the beginning of each OSN session.

◦ Activity Preference:

How often a user engages in each type of extroversive activities relates to their personalities. Some users like to post photos, while some others spend more time respond-ing to friends' posts; some mostly chat with friends via private messages, while some others always communicate by posting on each other's public message boards. Typical OSNs provide a great variety of social activities to satisfy their users' communication needs, for example, commenting, updating status, posting notes, sending messages, sharing posts, inviting others to an event, etc. As a result, this feature can provide a detailed portrayal of a user's social communication preferences.

◦ Activity Sequence:

The relative order a user completes multiple extroversive activities. While users have their preferences on different social activities, they may also have habitual patterns when switch from one activity to another. For instance, after commenting on friends' updates, some users often update their own status, while some other users prefer to send messages to or chat with friends instead. Therefore, the action sequence feature reflects a different social behavioral pattern from the activity preference.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
NCAIT 2017
JSS Academy of Technical Education
Vol. 4, Special Issue 8, May 2017

◦ Action Latency:
The speed of actions when a user engages in certain extroversive activities reflects the user's social interaction style. Many activities on OSNs require multiple steps to complete. For example, posting photos involves loading the upload page, selecting one or more photos, uploading, editing (e.g., clipping, decorating, tagging, etc.), previewing and confirmation. The time a user takes to complete each action of a given activity is heavily influenced by the user's social characteristics (e.g., serious vs. casual) and familiarity with the respective activity; but it doesn't directly reflect how fast a user acts due to different content complexity. The action latency feature is proposed to provide more fine-grained and accurate metric.

2) Introversive Behavior Features: Although invisible topeer users, introversive behaviors make up the majority of a user's OSN activity; as studied in previous work. the dominant (i.e., over 90%) user behavior on an OSN is browsing. Through introversive activities users gather and consume social information, which helps them to form ideas and opinions, and eventually, establish social connections and initiate future social communications. Hence, introversive behavior patterns make up an essential part of a user's online social behavioral characteristics. We propose the following four features to portray a user's introversive behavior.

◦ Browsing Preference:
The frequence a user visits various OSN page types depicts its social information preferences. Typical OSNs classify social information into different page types. For instance, profile pages contain personal information of the account owners, i.e., names, photos, interests etc.; the homepage compose of the account owner's friends' latest updates while a group page consists posts or photos shared by group members. Users' preferences on various types of social information naturally differ by their own interests, and the browsing preference feature intends to reflect this differenceby observing users' subjective behaviors.

◦ Visit Duration:
The time a user spends on visiting each webpage depicts another aspect of its social information consumption. Intuitively, users tend to spend less time on information that are "good-to-know", while allocate more time on consuming information that are "important", and their judgments are made based on their own personal interests. For example, some users prefer to stay on their own homepage reading friends' comments and updates, while some others tend to spend more time reading others' profile pages. The visitdurationfeature aims at capturing the social informationconsumption patterns for different users.

B. Facebook Measurement Study
We conduct a measurement study of Facebook users to understand their online social behaviors. In order to observe both extroversive and introversive behaviors from the partic-ipating users, we develop a browser extension to record user activities on Facebook in the form of clickstreams. In the following, we first present our data collection methodology and techniques, and an overview of the collected data set. Then, we detail the measurement results of user behavioral features.

1) Data Collection: We have recruited a total of50 Facebook users for our measurement study—22 are graduate students at universities and the rest are recruited via Amazon Mechanical Turk or Odesk, both of which are popular online crowdsourcing marketplaces. For each user, we collect approximately three weeks of their Facebook activities. To ensure that the recruited users are actually normal Facebook users, we use their first week as "trial" periods, during which we conduct manual review on the collected activity data.

The clickstreams in our dataset are organized in units of "sessions". We denote the start of a session when a user starts to visit Facebook in any window or tab of a browser; the end of a session is denoted when the user closes all windows or tabs that visit Facebook, or navigates away from Facebook in all windows or tabs of the browser. Clickstreams from concurrently opened tabs/windows are grouped into a single session, but are recorded individually (i.e., events from one window/tab are not merged with those from another window/tab). In total, we have collected 2678 sessions.

2) Feature Measurements: We first conduct a systematicstudy of services and webpages on Facebook. Based on request URL, we categorize 29 different types of extroversive activities that a user can conduct to interact with peer users; we also classify 9 types of Facebook webpages containing different kinds of social information, which users can browse privately (i.e., the introversive activities). With the mapping between the clickstream information and the user behaviors, we analyze each user's clickstreams to extract the corresponding behavior patterns. We present the combined measurement results of each behavior feature for all users to show the value space, and finally we use an example to illustrate user behavior diversities.
The user browsing preference distribution we can see that visits to the "homepage" and "profile" account for 86% of all user browsing actions.the large user diversities manifest on all types of webpages.presents the distribution of webpage visit duration. With a heavy-taileddistribution, over 90% of visits last less than 600 seconds. Users tend to have highly divergent behaviors for visit duration in the 0 to 3 minute range. We study the request latency using the same technique for measuring action latency, and observe similar results.with 90% of inter-request delays less than 10 seconds, the latencies for request sending during webpage browsing are generally slightly larger than those for engaging in extroversive activities. User diver-gence is the most obvious in the 0 to 9 second range.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
NCAIT 2017
JSS Academy of Technical Education
Vol. 4, Special Issue 8, May 2017

Our measurement study shows that we can discern user online social behavior characteristics by analyzing their click-streams. The results confirm that given a large number of social activities, individual OSN users tend to have diverse behavior patterns. We illustrate the diversity with an example. We randomly pick two users from our data set and present the most significant factors of each user's behavioral features, side-by-side,

IV. PROFILING SOCIAL BEHAVIORS

In this section, we first detail the formation of a user social behavioral profile using our proposed behavioral features. Based on our Facebook measurement study, we quantify Facebook user behavior patterns into a set of eight fine-grained metrics that correspond to the eight social behavioral features. The social behavior profile of an individual user can thus be built by combining the respective social behavioral metrics. Then, we describe the application of social behavior profiles in differentiating users and detecting compromised accounts.

A. Facebook User Behavioral Profile

In order to quantify user social behavior patterns on a specific OSN, we must first convert the social behavioral features into concrete metrics. We apply our knowledge gained in the Facebook measurement study, and devise a quantification scheme for each behavioral feature as follows.

◦ The first activity metric is defined as a 29-element vector, with each element corresponds to an extroversive activity on Facebook. The value of each element is the empirical probability a user engages in the associated activity as the first extroversive activity in a browser session.

◦ The activity preference metric is also a 29-element vector, similar to the first activity metric. The value of each element is the empirical probability a user engages in the associated activity throughout a browser session.

◦ The activity sequence metric is defined as a 29 × 29-element vector. If we conceptually arrange the vector as a 29-by-29 matrix, each cell of the matrix represents a transi-tion between two Facebook extroversive activities $a_1 \rightarrow a_2$, whose indices are reflected by the row or column number of the cell. The value of each cell is the probability of a user to transit to activity a_2 from activity a_1.

◦ The action latency metric is defined as an 11-element vector, and it records the empirical probability distribution of delays between consecutive HTTP requests while a user performs extroversive activities. We define the initial duration as zero, the first ten elements as one-second-wide bins, and element eleven as an infinite-time-width bin.

◦ The browsing preference metric is defined as a 9-element vector. Each element corresponds to a type of webpage on the Facebook website. The value of each element is the empirical probability a user visits the associated webpage throughout a browser session.

◦ The visit duration metric is defined as a 3 × 15-element vector, and each group of 15 elements records the empirical probability distribution of the duration a user visits homepages, profile pages or application page, respectively.[2] For each 15-element vector, we define the initial duration as zero, the first ten elements as 30-second-wide bins, the following four elements as 60-second-wide bins, and the fifteenth element as an infinite-time-width bin.

◦ The request latency metric is also a threefold 11-element vector, and each group of 11 elements records the empirical probability distribution of delays between consecutive HTTP requests during a user's visits to homepages or profile pages or application pages, respectively. Similar to the action speed metric, the initial duration is zero, element one through ten are one-second-wide bins, and element eleven is an infinite-time-width bin.

◦ Thebrowsingsequencemetric is defined as a9 × 9-element vector. Similar to the activity sequence metric, we conceptually arrange the vector as a 9-by-9 matrix, and each cell of the matrix represents a transition between browsing two types of Facebook webpages $p_1 \rightarrow p_2$, whose indices are reflected by the row or columnnumber of the cell. The value of each cell is the probability of the user to switch to type p_2 after browsing page type p_1.

With concrete behavioral metrics in hand, we build a Facebook user's social behavioral profile by first combining their social behavior metrics into an 8-vector tuple, then normalizing each vector so that the sum of all elements in a vector equals to one. In particular, the visit duration and request latency vectors are multiplied by a factor of 1/3; the activity sequence vectors and the browsing sequence vectorsare multiplied by 1/29 and 1/9, respectively, while all other metrics are unchanged. Table II lists the sample of a Facebook user social behavioral profile.

B. Differentiating Users

The social behavioral profile depicts various aspects of a user's online social behavior patterns, and it enables us to quantitatively describe the differences in distinct user social behaviors. In the following, we first describe how to com-pare social behavioral profiles by calculating their difference. Then, we discuss the application of social behavioral profile comparison to distinguishing different users and detecting compromised accounts.

2) Applying Profile Comparison: To apply the profile com-parison technique for differentiating users, we must further introduce another concept, self variance, in addition to the profile difference.

With two or more distinct pieces of behavioral data (i.e., clickstreams) collected from the same user, the social behavioral profiles built from each piece of behavioral data are not identical. The reasons for the differences are twofold. First, human behaviors are intrinsically non-

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

NCAIT 2017

JSS Academy of Technical Education
Vol. 4, Special Issue 8, May 2017

deterministic, thereforea small amount of variation is expected even for the same activity performed by the same user. Second, because the social behavioral profile is built on top of statistical obser-vations, errors always exist for a finite amount of samples.

3) Detecting Compromised Accounts: Together with theself variance, we can apply profile comparison to distinguish different users and detect compromised accounts.
After building a user's behavior profile and variance during a training phase, we can decide whether the user's account is compromised. While the method illustrated before can be employed to fulfill the task, we adjust the method by personalizing the computation of difference to each user's behavior profile.

User consistency on behaviour features differs from one to one. The personalized weight on each feature in the training phase enlarges the distance in user differentiation. Heavy-weighted behavior features that a user behaves more consistently on play more important roles in detecting impostors than light-weighted features. If an unknown behavior profile belongs to U , it is likely that its distance on heavy-weighted features are smaller than that on light-weighted features. For an impostor's profile that does not hold this pattern, it is highly likely that the distance to U on heavy-weighted features is also large, which results in comparatively larger difference.

As it is possible that a user's behaviour patterns change over time, the behaviour profile needs to be updated periodically to accurately portray its patterns. While some online habits remain, a user's behaviour may evolve over time. To capture the change, the training phase can be repeated using a user's latest clickstream to update a user's behaviour profile including feature weights.

In addition, when there are introduction of new services, new behaviour features may need to be extracted. At the same time, multiple existing behaviour features may also experience significant changes, which could be large enough to produce false alarms. This increased false alarm rate cannot be limited by weighing potential harmfulness. The training phase also needs to be repeated in this scenario. New training data collection is required, and it may take some time for the detector to work accurately again.
4) Incomplete Behaviour Profiles: Dependent upon user Activeness in OSNs, the completeness of a user's behaviour profile varies. The incomplete behaviour profiles should be specially processed while calculating the difference, considering the lack of sample activities from which metric vectors are built.

When some feature vectors are not available, they are not considered while calculating the difference; in this scenario the final difference will be normalized. For instance, if a user's extroversive activity metric vectors are not available due to the reason that it does not conduct extroversive activities, its difference to another behaviour profile only counts into the distances on the four extroversive activity vectors; Furthermore, when there are rare sample activities to build a metric vector, it is taken as N/A. For example, if there are only 5 extroversive activities in a clickstream, the activitypreferencevector built from them can hardly be representativeof the user's behaviour pattern. Hence, a threshold of the minimum number of sample activities should be assigned to guarantee the quality of metric vectors. Those vectors built from a lower-than-threshold number of sample activities are taken as N/A.

The varied thresholds of sample activity are assigned to different feature vectors. For browsing preference vector, it is possible that 15 page browsing activities are able to derive a comparatively representative vector; but for browsing sequence metric, 15 browsing transitions can hardly demonstrate illustrative transition probabilities. Hence, when a sample threshold is assigned, it is applicable to all features except for browsing sequence and activity sequence, whose thresholdsare two times of the assigned threshold.

V. EVALUATION

We first verify that behavioral profile can accurately portray a user's behavior pattern. Next, we validate the feasibility of employing behavioral profiles to distinguish different users, which can be used to detect compromised accounts.

A. Difference vs. Variance
We demonstrate that compared to a user's behavior variance, its behavioral profile difference from others is more significant. That is, a user's behavioral profile can accurately characterize its behavior pattern. We compute each sample user's behavior variance and behavioral profile differences between its own and other users'. For each sample user, we equally partition its clickstream into four complementary parts by session, and four behavioral profiles are built accordingly; its weights on each feature are calculated and used for difference calculation. A 4-fold cross-validation is conducted to calculate its average behavior variance and the average difference from the others' behavioral profiles.

Figure 4 shows each user's behavior variance and the average behavioral profile difference to others'. Note that only those users who have sufficient social activities, referred as "valid" users, are included in the figure. In particular, the behavioral profile of each valid user must have more than or equal to 4 non-empty feature vectors, to ensure that the behavioral profile is complete enough to represent the user's behavior patterns. Each feature should be derived from more than or equal to 10 social activities. For those users whose social profiles do not meet the requirement above are excluded.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

NCAIT 2017

JSS Academy of Technical Education

Vol. 4, Special Issue 8, May 2017

As the figure shows, each user's self variance is obviously lower than its average difference from other users. This coincides with our intuition that a user's behavior variance is usually within a certain range; comparatively complete behavioral profiles can portray users' behavior patterns. More importantly, it is possible to take advantage of the difference between behavioral profiles to discern a user.In addition, we compare the results between the student users and the online-recruited users. The difference between average behavior variance of students and that of online-recruited users is only 6%. Moreover, using the difference between self-variance and average difference from other users as a metric, we observe that there is no distinction between the student users and the online-recruited users. Thus, neither of the two group of sample users are biased in evaluation.

B. Detection Accuracy

Here we further evaluate the accuracy of using social behavioral profiles to differentiate online users. We conduct three sets of experiments by varying training data size, feature quality, and profile completeness, respectively, to evaluate their impacts upon the detection accuracy. Both false positive and false negative are considered as inaccurate detection.

1) Input Size vs. Accuracy: Intuitively, the more trainingdata are given to build a user's behavioral profile, the better the profile reflects its behavior pattern; hence the profile difference demonstrates the dissimilarity between two user behaviors more accurately.

Cross-validation is used to make sure that each part of data are used for both training and validation, and the result is not derived from biased data. Furthermore, we only consider users whose behavioral profiles consist of more than or equal to 4

2) Feature Quality vs. Accuracy: We adjust the threshold ofthe number of sample activities to explore whether the feature vector quality affects the detection accuracy.

When building a user's behavioural profile, the number of sample activities that derive a feature vector determines whether the feature vector represents a user's behaviour accurately. By assigning a threshold to the number of sample activities, we can take control over the quality of feature vectors. We designate those vectors derived from insufficient activities to be N/A. Intuitively, higher sample activity threshold results in feature vectors with higher quality, which reduces the noise of behaviour variance introduced by rare sample activities. Thus, the difference between users can be discerned more accurately.

3) Profile Completeness vs. Accuracy: Due to the lack ofcertain activities, some behavioural feature vectors can be N/A. For instance, when one never conducts extroversive activities in its clickstream, at least four of its feature vectors are N/A, which makes its profile incomplete. By adjusting the least number of non-empty features vectors, the completeness of selected behavioural profiles can be guaranteed.

Overall, active users can be distinguished more accurately by their behavioural profiles compared to inactive users. The more types of activities a user conducts, the more complete its behaviour profile can be. And the more activities a user conduct, the more sample activities can be obtained within certain duration, leading to more accurate behavioural profile. On the other hand, as compromised accounts are usually manipulated to become active to spread spam, there will be a sudden change of behaviour when an inactive user account is compromised. Thus, we can still detect the compromisation of an inactive user account, even without its accurate and complete behaviour profile.

VI. CONCLUSION

In this paper, we propose to build a social behaviour profile for individual OSN users to characterize their behavioural patterns. Our approach takes into account both extroversive and extroversive behaviours. Based on the characterized social behavioural profiles, we are able to distinguish a user's from others, which can be easily employed for compromised account detection. Specifically, we introduce eight behavioural features to portray a user's social behaviours, which include both its extroversive posting and introversive browsing activities. A user's statistical distributions of those feature values comprise its behavioural profile. While users' behaviour profiles diverge, individual user's activities are highly likely to conform to its behavioural profile. This fact is thus employed to detect a compromised account, since impostors' social behaviours can hardly conform to the authentic user's behavioural profile. Our evaluation on sample Facebook users indicate that we can achieve high detection accuracy when behavioural profiles are built in a complete and accurate fashion. Portray a user's social behaviours, which include both its extroversive posting and introversive browsing activities. A user's statistical distributions of those feature values comprise its behavioural profile. While users' behaviour profiles diverge, individual user's activities are highly likely to conform to its behavioural profile. This fact is thus employed to detect a compromised account, since impostors' social behaviours can hardly conform to the authentic user's behavioural profile. Our evaluation on sample Facebook users indicate that we can achieve high detection accuracy when behavioural profiles are built in a complete and accurate fashion.

REFERENCES

[1] 250,000 Twitter Accounts Hacked. [Online]. Available:http://www.cnn.com/2013/02/01/tech/social-media/twitter-hacked, accessed Sep. 2013.

[2] 50,000 Facebook Accounts Hacked. [Online]. Available:http://www.ktsm.com/news/thousands-of-facebook-accounts-hacked, accessed Sep. 2013.

[3] Detecting Suspicious Account Activity. [Online]. Available:http://googleonlinesecurity.blogspot.com/2010/03/detecting-suspicious-account-activity.html, accessed Sep. 2013.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

NCAIT 2017

JSS Academy of Technical Education

Vol. 4, Special Issue 8, May 2017

[4] Facebook Tracks the Location of Logins for Better Security. [Online].Available: http://www.zdnet.com/blog/weblife/facebook-adds-better-security-tracks-the-location-of-your-logins/2010, accessed Sep. 2013.

[5] Y. Bachrach, M. Kosinski, T. Graepel, P. Kohli, and D. Stillwell, "Personality and patterns of Facebook usage," in Proc. 3rd Annu. ACMWeb Sci. Conf. (WebSci), Evanston, IL, USA, 2012, pp. 24–32.

[6] F. Benevenuto, T. Rodrigues, M. Cha, and V. Almeida, "Characterizing user behavior in online social networks," in Proc. 9th ACM SIGCOMMConf. Internet Meas. Conf. (IMC), Chicago, IL, USA, 2009, pp. 49–62.

[7] Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, "Aiding the detection of fake accounts in large scale social online services," in Proc. 9thUSENIX Conf. Netw. Syst. Design Implement. (NSDI), San Jose, CA,USA, 2012, p. 15.

[8] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "COMPA: Detecting compromised accounts on social networks," in Proc. Symp. Netw.Distrib. Syst. Secur. (NDSS), San Diego, CA, USA, 2013.

[9] H. Gao, Y. Chen, K. Lee, D. Palsetia, and A. Choudhary, "Towards online spam filtering in social networks," in Proc. Symp. Netw. Distrib.Syst. Secur. (NDSS), San Diego, CA, USA, 2012.

[10] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao, "Detecting and characterizing social spam campaigns," in Proc. 10th ACMSIGCOMM Conf. Internet Meas. (IMC), Melbourne, VIC, Australia,2010, pp. 35–47.

[11] K.-I. Goh and A.-L. Barabási, "Burstiness and memory in complex systems," Europhys. Lett., vol. 81, no. 4, p. 48002, 2008.

[12] C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@spam: The under-ground on 140 characters or less," in Proc. 17th ACM Conf. Comput.Commun. Secur. (CCS), Chicago, IL, USA, 2010, pp. 27–37

[13] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: Social honeypots + machine learning," in Proc. 33rd Int. ACM SIGIRConf. Res. Develop. Inf. Retr. (SIGIR), Geneva, Switzerland, 2010, pp435–442.

[14] C. Ross, E. S. Orr, M. Sisic, J. M. Arseneault, M. G. Simmering, and R. R. Orr, "Personality and motivations associated with Facebook use," Comput. Human Behavior, vol. 25, no. 2, pp. 578–586, 2009.

[15] F. Schneider, A. Feldmann, B. Krishnamurthy, and W. Willinger, "Under-standing online social network usage from a network perspective," in Proc. 9th ACM SIGCOMM Conf. Internet Meas. Conf. (IMC), Chicago,IL, USA, 2009, pp. 35–48.

[16] J. Song, S. Lee, and J. Kim, "Spam filtering in Twitter using sender-receiver relationship," in Proc. 14th Int. Conf. Recent Adv. IntrusionDetection (RAID), Menlo Park, CA, USA, 2011, pp. 301–317.

[17] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in Proc. 26th Annu. Comput. Secur. Appl. Conf. (ACSAC), Austin, TX, USA, 2010, pp. 1–9.

[18] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, "Design and evaluation of a real-time URL spam filtering service," in Proc.IEEESymp. Secur. Privacy (S&P), Oakland, CA, USA, May 2011,pp447–462.

[19] K. Thomas, D. McCoy, C. Grier, A. Kolcz, and V. Paxon, "Trafficking fraudulent accounts: The role of the underground market in Twitter spam and abuse," in Proc. 22nd USENIX Secur. Symp., Washington, DC, USA, 2013, pp. 195–210.

[20] B. Viswanathet al., "Towards detecting anomalous user behavior in online social networks," in Proc. 23rd USENIX Secur. Symp., San Diego, CA, USA, Aug. 2014, pp. 223–238.

[21] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove, "An analysis of social network-based Sybil defenses," in Proc. ACM SIGCOMMConf. (SIGCOMM), New Delhi, India, 2010, pp. 363–374.

[24] C. Wilson, B. Boe, A. Sala, K. P. N. Puttaswamy, and B. Y. Zhao, "User interactions in social networks and their implications," in Proc.4th ACM Eur. Conf. Comput. Syst. (EuroSys), Nuremberg, Germany,2009, pp. 205–218.

[25] Y. Xieet al., "Innocent by association: Early recognition of legitimate users," in Proc. ACM Conf. Comput. Commun. Secur. (CCS), Raleigh, NC, USA, 2012, pp. 353–364.

[26] H. Xiong, P. Malhotra, D. Stefan, C. Wu, and D. Yao, "User-assisted host-based detection of outbound malware traffic," in Proc. 11th Int.Conf. Inf. Commun. Secur. (ICICS), Beijing, China, 2009, pp. 293–307.

[27] C. Yang, R. Harkreader, J. Zhang, S. Shin, and G. Gu, "Analyzing spammers' social networks for fun and profit: A case study of cyber criminal ecosystem on Twitter," in Proc. 21st Int. Conf. World WideWeb (WWW), Lyon, France, 2012, pp. 71–80.