International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 24 Issue 8 – MAY 2017.

DETECTION OF COMPROMISED ACCOUNTS ON SOCIAL NETWORKS BASED ON ANOMALOUS USER BEHAVIOUR

Jince P Kuruvilla [#1],Kavitha P[*2], and Dr.G.Kalpana [*3]

[#] *PG Scholar M-Tech(ISCF), Dr.M.G.R. Educational and Research Institute, Chennai, India*

[*] *Professor, Dept. of CSE, Dr.M.G.R. Educational and Research Institute, Chennai, India*

***Abstract*— Account compromisation is a serious threat to users of Online Social Networks (OSNs). While relentless spammers exploit the established trust relationships between account owners and their friends to efficiently spread malicious spam, timely detection of compromised accounts is quite challenging due to the well-established trust relationship between the service providers, account owners, and their friends. Instead of analyzing 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 messages, 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.**

This paper presents a novel method to detect Account Compromisation as and when it happens by profiling online social behaviors. In sight of the above intuition and reasoning, we first conduct a study on online user social behaviors by collecting and analyzing user clickstreams of a well-known OSN website. Based on our observation of user interaction with different OSN services, we propose several new behavioural features that can effectively quantify user differences in online social activities. For each behavioural feature, we deduce a behavioural metric by obtaining a statistical distribution of the value ranges, observed from each user's clickstreams. Moreover, we combine the respective behavioural metrics of each user into a social behavioural profile, which represents a user's social behavior patterns. To validate the effectiveness of social behavioural profile in detecting account activity anomaly, we apply the social behavioural 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 behavioural profiles. Our evaluation results show that social behavioural profile can effectively differentiate individual OSN users with accuracy up to 98.6%, and the more active a user, the more accurate the detection.

***Index Terms*— Compromised account detection, Online Social Networks, Suspicious account, Online Social Behavior..**

I. INTRODUCTION

Compromised accounts in Online Social Networks (OSNs) are more favorable than Sybil 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 analysis of tweets and Facebook posts reveal that most spam are distributed via compromised accounts, instead of dedicated spam accounts. Account compromisation is a serious threat to usersof Online Social Networks (OSNs). While relentless spammers exploit the established trust relationships between account ownersand their friends to efficiently spread malicious spam, timely detection of compromised accounts is quite challenging due to the well-established trust relationship between the service providers, account owners, and their friends [1].

The social behaviors of OSN users is examined i.e., their usage of OSN servicesand the application in detecting the compromisedaccounts. A set of social behavioral features that can effectively characterize the user social activities on OSNs is proposed. The efficacy of these behavioural features is validated by collecting and analyzing real user clickstreams to anOSN website. The individual user's social behavioral profile is devised by combining its respectivebehavioral feature metrics. A social behavioral profile accurately reflects a user's OSN activity patterns. While an authentic ownerconforms to its account's social behavioral profile involuntarily it is hard and costly for impostors to feign [2].

II. LITERATURE SURVEY

Previous research on spamming account detection mostly cannot distinguish compromised accounts from Sybil accounts, with only one recent study by Egele*et al.* 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. 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. Major OSNs today employ IP geolocation logging to battle against account compromisation. However, this approach is known to suffer from low detection granularity and high false positive rate.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 [3].

As social networking sites have risen in popularity, cyber-criminals started to exploit these sites to spread malware and to carry out scams. Previous work has extensively studied the use of fake (Sybil) accounts that attackers set up to distribute spam messages (mostly messages that contain links to scam pages or drive-by download sites). Fake accounts typically exhibit highly anomalous behavior, and hence are relatively easy to detect. As a response, attackers have started to compromise and abuse legitimate accounts. Compromising legitimate accounts is very effective, as attackers can leverage the trust relationships that the account owners have established in the past. Moreover, compromised accounts are more difficult to clean up because a social network provider cannot simply delete the corresponding profiles. They presented a novel approach to detect compromised user accounts in social networks, and we apply it to two popular social networking sites, Twitter and Facebook. Our approach uses a composition of statistical modeling and anomaly detection to identify accounts that experience a sudden change in behavior [4].

Since behavioral changes can also be due to benign reasons (e.g., a user could switch her preferred client application or post updates at an unusual time), it is necessary to derive a way to distinguish between malicious and legitimate changes. To this end, we look for groups of accounts that all experience similar changes within a short period of time, assuming that these changes are the result of a malicious campaign that is unfolding. We developed a tool, called proposed system, that implements our approach, and we ran it on a large-scale dataset of more than 1.4 billion publicly-available Twitter messages, as well as on a dataset of 106 million Facebook messages. Proposed system was able to identify compromised accounts on both social networks with high precision [5].

Online Social Networks (OSNs) are extremely popular among Internet users. Unfortunately, in the wrong hands, they are also effective tools for executing spam campaigns. We present an online spam filtering system that can be deployed as a component of the OSN platform to inspect messages generated by users in real-time. We propose to reconstruct spam messages into campaigns for classification rather than examine them individually. Although campaign identification has been used for offline spam analysis, we apply this technique to aid the online spam detection problem with sufficiently low overhead. Accordingly, the system adopts a set of novel features that effectively distinguish spam campaigns. It drops messages classified as "spam" before they reach the intended recipients, thus protecting them from various kinds of fraud. They evaluated the system using 187 million wall posts collected from Facebook and 17 million tweets collected from Twitter. In different parameter settings, the true positive rate reaches 80.9% while the false positive rate reaches 0.19% in the best case. In addition it stays accurate for more than 9 months after the initial training phase. Once deployed, it can constantly secure the OSNs without the need for frequent re-training. Finally, tested on a server machine with eight cores (Xeon E5520 2.2 GHz) and 16GB memory, the system achieves an average throughput of 1580 messages/sec and an average processinglatency of 21.5ms on the Facebook dataset [6].

With over 2.5 hours a day spent browsing websites online and with over a billion pages, identifying and detecting web spam is an important problem. Although large corpora of legitimate web pages are available to researchers, the same cannot be said about web spam or spam web pages. We introduce the Webb Spam Corpus 2011 - a corpus of approximately 330,000 spam web pages - which we make available to researchers in the fight against spam. By having a standard corpus available, researchers can collaborate better on developing and reporting results of spam filtering techniques. The corpus contains web pages crawled from links found in over 6.3 million spam emails. We analyze multiple aspects of this corpus including redirection, HTTP headers and web page content. We also provide insights into changes in web spam since the last Webb Spam Corpus was released in 2006. These insights include: 1) spammers manipulate social media in spreading spam 2) HTTP headers also change over time (e.g. hosting 'P addresses of web spam appear in more 'P ranges) 3) Web spam content has evolved but the majority of content is still scam [7].

Social networks are popular platforms for interaction, communication and collaboration between friends. Researchers have recently proposed an emerging class of applications that leverage relationships from social networks to improve security and performance in applications such as email, web browsing and overlay routing. While these applications often cite social network connectivity statistics to support their designs, researchers in psychology and sociology have repeatedly cast doubt on the practice of inferring meaningful relationships from social network connections alone. This leads to the question: Are social links valid indicators of real user interaction? If not, then how can we quantify these factors to form a more accurate model for evaluating socially-enhanced application [8].

We address this question through a detailed study of user interactions in the Facebook social network. We propose the use of interaction graphs to impart meaning to online social links by quantifying user interactions. We analyze interaction graphs derived from Facebook user traces and show that they exhibit significantly lower levels of the "small-world" properties shown in their social graph counterparts. This means that these graphs have fewer "super nodes" with

extremely high degree, and overall network diameter increases significantly as a result. To quantify the impact of our observations, we use both types of graphs to validate two well-known social-based applications (RE and Sybil Guard). The results reveal new insights into both systems, and confirm our hypothesis that studies of social applications should use real indicators of user interactions in lieu of social graphs [9]. Conventional network security solutions are performed on network-layer packets using statistical measures. These types of traffic analysis may not catch stealthy attacks carried out by today's malware. We aim to develop a host-based security tool that identifies suspicious outbound network connections through analyzing the user's surfing activities. Specifically, our solution for Web applications predicts user's network connections by analyzing Web content; unpredicted traffic is further investigated with the user's help. We describe our method and implementation as well as the experimental results in evaluating its efficiency and effectiveness. We describe how our studies can be applied to detecting bot infection [10].

In order to assess the workload of our host-based traffic-analysis tool, we also perform a large-scale characterization study on 500 university-users' wireless network traces for 4-month period. We study both the statistical and temporal patterns of individuals' web usage behaviors from collected wireless network traces. Users are classified into different profiles based on their web usage patterns. Our results show that users have regularities in their Web activities and the expected workload of our traffic-analysis solution is low [11].

III. METHODOLOGY

In this paper, we study the social behaviors of OSN users, i.e., their usage of OSN services, and the application of detecting the compromised accounts. In particular, we propose a set of social behavioral features that can effectively characterize the user social activities on OSNs. We validate the efficacy of these behavioral features by collecting and analyzing real user click streams to an OSN website. Based on our measurement study, we devise individual user's social behavioral profile by combining its respective behavioral feature metrics. A social behavioral profile accurately reflects a user's OSN activity patterns. While an authentic owner conforms to its account's social behavioral profile involuntarily, it is hard and costly for impostors to feign. We evaluate the capability of the social behavioral profiles in distinguishing different OSN users, and our experimental results show the social behavioral profiles can accurately differentiate individual OSN users and detect compromised accounts.

A. Click stream Method

A click stream is the recording of the portions of the monitor a user clicks on while browsing the web or using another software application. As the user clicks anywhere in the webpage or application, the activity is recorded on a client or inside the web server, as well as possibly the web browser, router, proxy server or ad server. Click stream analysis is beneficial for web activity examination, software testing, market investigation, and for validating employee productivity.

An authentic user's social patterns are recorded, checking the compliance of the account's forthcoming behaviors with the authentic patterns can detect compromisation of accounts. Even though a user's credential is hacked, a malicious party cannot easily obtain the social behavioral patterns of the user without the control of the physical devices or the click streams. We present a measurement study on user behavior diversity by analyzing real user click streams of Social Network, say Facebook with respect to our proposed features. 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 participating users, we develop a browser extension to record user activities on Facebook in the form of click streams. The click streams 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. We discern the user *action latency* by first grouping click streams belonging to each user activity type, and then measuring the inter-arrival time of consecutive HTTP requests within each group of click streams.

B. Profiling social behaviors

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.

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.

C. 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 compare social behavioral profiles by calculating their difference. With two or more distinct pieces of behavioral data (i.e., click streams) 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-deterministic; therefore a 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 observations, errors always exist for a finite amount of samples.

D. Detecting Compromised Accounts

Together with the self 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.

Giving a weight on each feature is to portray a user's degree of consistency on different behavior features, which is also difficult to feign. User consistency on behavior 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 is 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 on heavy weighted features is also large, which results in comparatively larger difference.

As it is possible that a user's behavior patterns change over time, the behavior profile needs to be updated periodically to accurately portray its patterns. While some online habits remain, a user's behavior may evolve over time. To capture the change, the training phase can be repeated using a user's latest click stream to update a user's behavior profile including feature weights. 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 thresholds are two times of the assigned threshold.

IV. RESULTS AND ANALYSIS

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. Our experiments indicated that a small number of popular applications resulted in a large number of false positives. Therefore, we removed the six most popular applications, including Mafia Wars from our dataset. Note that these six applications resulted in groups spread over the whole dataset. Thus, we think it is appropriate for a social network administrator to white-list applications at a rate of roughly three instances per year.

Our system flagged 33 messages as violating their user's profile. The reason proposed system did not flag these accounts in the first place is that the clusters generated by these messages were too small to be evaluated, given the API limit we mentioned before. If we did not have such a limit, proposed system would have correctly flagged them. Seven more messages contained URLs that were similar to those in the 33 messages. Even though these compromised accounts did not violate their behavioral profiles, they would have been detected by proposed system, because they would have been grouped together with other messages that were detected as violating their behavioral profiles.

Overall, active users can be distinguished more accurately by their behavioral profiles compared to inactive users. The more types of activities a user conducts, the more complete its behavior profile can be. And the more activities a user conduct, the more sample activities can be obtained within certain duration, leading to more accurate behavioral profile. On the other hand, as compromised accounts are usually manipulated to become active to spread spam, there will be a sudden change of behavior 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 behavior profile.

V. CONCLUSION

A social behavioral profile for individual OSN users to characterize their behavioural patterns is proposed and built. The approach takes into account both extroversive and introversive behaviors. Based on the characterized social behavioural profiles, we are able to distinguish a user from others, which can be easily employed for compromised account detection. Specifically, we introduce eight behavioural features to portray a user's social behaviors, 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' behavioural 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 behaviors can hardly conform to the authentic user's behavioural profile. Our evaluation on sample Facebook a user indicates that we can achieve high detection accuracy when behavioural profiles are built in a complete and accurate fashion.

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] 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 Wide Web (WWW), Lyon, France, 2012, pp. 71–80.

[5] Y. Bachrach, M. Kosinski, T. Graepel, P. Kohli, and D. Stillwell, "Personality and patterns of Facebook usage," in Proc. 3rd Annu. ACM Web 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 SIGCOMM Conf. 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, "proposed system: 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 ACM SIGCOMM 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 underground 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 SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR), Geneva, Switzerland, 2010, pp. 435–442.

[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, "Understanding 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.