Volume: 08 Issue: 03 Pages: 3090-3092 (2016) ISSN: 0975-0290

Celebrity Check – New Friends: Integration of Detection & Removal of Anonymous Identical Celebrity with Best Friend Identification

Mrs.V.Perathu Selvi
Assistant Professor, Department of Computer Science, Francis Xavier Engineering College, Tirunelveli
Email: perathuselvi@gmail.com
Ms.K.Raja Sundari
Assistant Professor, Department of Computer Science, Francis Xavier Engineering College, Tirunelveli
Email: sundari.november@gmail.com
Mrs.P.J.Beslin Pajila
Assistant Professor, Department of Computer Science, Francis Xavier Engineering College, Tirunelveli
Email: beslin.kits@gmail.com

ABSTRACT

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. In the proposed system, we propose identification of same user in different social network sites (SNS) and elimination of fake user account from the SNS. This is achieved via checking screen name, photo, friends list, gender, location, birthday and school / college education and working place. Using these behavioral features of user social behavior, it identifies the fake user. In the modification process, apart from the removal of anonymous accounts we also add on identification Friends based on user's mind set / Interest. We are monitoring Users Interest, Likes posted by the user and Android based mobility pattern analysis. Best friends are identified and security layer is enveloped by monitoring user's behavior pattern. Vulgar worded posts are removed and the user is terminated in case of misbehavior.

Keywords - **Malicious spam, social network sites, security**.

Date of Submission: Dec 02, 2016 Date of Acceptance: Dec 19, 2016

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 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 geolocation 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 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 analyzing user profile contents or message contents, we seek to uncover the behavioral 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.

Volume: 08 Issue: 03 Pages: 3090-3092 (2016) ISSN: 0975-0290

Therefore, as long as an authentic user's social patterns are recorded, checking the compliance of the account's upcoming behaviors with the authentic patterns can detect account compromisation. Even though a user's credential is hacked, a malicious party cannot easily obtain the user's social behavior patterns without the control of the physical machines or the clickstreams. Moreover, considering that for a spammer, who carries very different social interests from those of regular users (e.g., mass spam distribution vs. entertaining with friends), it is very costly to mimic different individual user's social interaction patterns, as it will significantly reduce spamming efficiency. 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 behavioral features that can effectively quantify user differences in online social activities. For each behavioral feature, we deduce a behavioral metric by obtaining a statistical distribution of the value ranges, observed from each user's clickstreams. Moreover, we combine the respective behavioral metrics of each user into a social behavioral profile, which represents a user's social behavior patterns.

II. Literature Survey

People use various social media for different purposes. The information on an individual site is often incomplete. When sources of complementary information are integrated, a better profile of a user can be built to improve online services such as verifying online information. To integrate these sources of information, it is necessary to identify individuals across social media sites. This method aims to address the cross-media user identification problem. (MOBIUS) methodology for finding a mapping among identities of individuals across social media sites. It consists of three key components: the first component identifies users' unique behavioral patterns that lead to information redundancies across sites; the second component constructs features that exploit information redundancies due to these behavioral patterns; and the third component employs machine learning for effective user identification. Here, the cross-media user identification problem is defined and show that MOBIUS is effective in identifying users across social media sites[1].

How much do tagging activities tell about a user? Is it possible to identify people in Delicious based on the tags, which they use in Flickr? In [2], study those questions and investigate whether users can be identified across social tagging systems. It combine two kinds of information: their user ids and their tags. It introduce and compare a variety of approaches to measure the distance between user profiles for identification. With the best performing combination we achieve, depending on the actual settings, accuracies of between 60% and 80%, which demonstrates that the traces of Web 2.0 users can reveal quite much about their identity[2].

The first task any individual faces after joining an online social network (OSN) is locating friends that are present on that particular site. Most OSNs over some variation of a tool that imports email contact lists to facilitate the task of finding one's friends. However, given that OSNs attempt to reconnect individuals with past acquaintances, one might not have access to the email address for a long lost friend. Furthermore, people tend to utilize a number of aliases online, meaning that an email address cannot always be used to reliably find a friend. Thus, new members must still manually search for friends based on a number of biographical attributes, such as gender, age, hometown, etc. It is not clear, however, what attributes are useful for conducting the search. Even after the search has been performed, the person performing the search might be left with a number of candidate profiles. In [3], M. Motoyama and G. Varghese develop a system for searching and matching individuals in OSNs.

Organizations are increasingly mining the personal data users generate as they carry out much of their day- to- day activities online. A range of new business models specifically exploit what users publish on their social network profiles, including services performing background checks and analytics providers who, e.g., associate demographics with consumer behavior. By [4], understand the capabilities of machine learning techniques for linking independent accounts that users maintain on different social networks, based solely on the information people explicitly and publicly provide in their profiles. And perform a large scale study that assesses a range of correlation approaches for matching accounts between five popular social networks: Twitter, Facebook, Google+, Myspace, and Flickr. The results show for instance that by exploiting usernames, real names, locations, and photos, we can robustly identify about 80% of the matching pairs of user accounts between any combination of two social networks among Twitter, Facebook and Google+. This is the first to demonstrate the feasibility of such conceptually simple privacy attacks at large scale, across several major networks, and with such efficiency.

Instance matching targets the extraction, integration and matching of instances referring to the same real-world entity. In [5], K. Cortis, S. Scerri, I. Rivera, and S. Handschuh present a weighted ontology-based user profile resolution technique which targets the discovery of multiple online profiles that refer to the same person identity. The elaborate technique takes into account profile similarities at both the syntactic and semantic levels, employing text analytics on top of open data knowledge to improve its performance. A two-staged evaluation of the technique performs various experiments to determine the best out of alternative approaches. These results are then considered in an improved algorithm, which is evaluated by real users, based on their real social network data. Here, a profile matching precision rate of 0.816 is obtained. The presented Social Semantic Web technique has a number of

Volume: 08 Issue: 03 Pages: 3090-3092 (2016) ISSN: 0975-0290

useful applications, such as detection of untrusted known persons behind anonymous profiles, and information sharing management across multiple social networks.

In the existing system, 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. Disadvantages of the existing system are unreliable, less security, less effective

III. Proposed System

In the proposed system, we propose identification of same user in different social network sites (SNS) and elimination of fake user account from the SNS. This is achieved via checking screen name, photo, friends list, gender, location, birthday and school / college education and working place. Using these behavioral features of user social behavior, it identifies the fake user.

In the modification process**,** apart from the removal of anonymous accounts we also add on identification Friends based on user's mind set / Interest. We are monitoring Users Interest, Likes posted by the user and Android based mobility pattern analysis. Best friends are identified and security layer is enveloped by monitoring user's behavior pattern. Vulgar worded posts are removed and the user is terminated in case of misbehavior. Advantages of the proposed system are Reliable, High Security, More Effective. The algorithm used in the proposed system is classification algorithm.

IV. Conclusion & Future Enhancement

This study addressed the problem of user identification across SMN platforms and offered an innovative solution. As a key aspect of SMN, network structure is of paramount importance and helps resolve de-anonymization user identification tasks. Therefore, we proposed a uniform network structure-based user identification solution. We also developed a novel friend relationship-based algorithm called FRUI. To improve the efficiency of FRUI, we described two propositions and addressed the complexity. Finally, we verified our algorithm in both synthetic networks and ground truth networks. In this paper, we propose to build a social behavior profile for individual OSN users to characterize their behavioral patterns. Our approach takes into account both extroversive and introversive behaviors. Based on the characterized social behavioral profiles, we are able to distinguish users from others, which can be easily employed for compromised account detection. Specifically, we introduce eight behavioral 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 behavioral profile.

The future work can be four-fold. First, we would like to evaluate our system on large-scale field experiments. Second, we intend to implement the life style extraction using LDA and the iterative matrix-vector multiplication method in user impact ranking incrementally, so that Friend book would be scalable to large-scale systems. Third, the similarity threshold used for the friend-matching graph is fixed in our current prototype of Friend book. It would be interesting to explore the adaption of the threshold for each edge and see whether it can better represent the similarity relationship on the friend matching graph.

References

[1] R. Zafarani and H. Liu, "Connecting users across social media sites: a behavioral-modeling approach," in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2013, pp. 41–49.
[2] T. Iofciu, P. Fankhauser, F. Abel, and K. Bischoff, "Identifying users across social tagging systems," in Proc. 5th Int. AAAI Conf. Weblogs Social Media, 2011, pp. 522–525.
[3] M. Motoyama and G. Varghese, "I seek you: searching and matching individuals in social networks," in Proc. 11th Int. Workshop Web Inf. Data Manage., 2009, pp. 67–75.
[4] O. Goga, D. Perito, H. Lei, R. Teixeira, and R. Sommer, "Large-scale correlation of accounts across social networks," University of California at Berkeley, Berkeley, California, Tech. Rep. TR-13-002, 2013.
[5] K. Cortis, S. Scerri, I. Rivera, and S. Handschuh, "An ontologybased technique for online profile resolution," in Proc. 5th Int. Conf. Social Informat., 2013, pp. 284–298.
[6] S. Tan, Y. Li, H. Sun, Z. Guan, X. Yan, J. Bu, C. Chen, and X. He, "Interpreting the public sentiment variations on twitter," IEEE Trans. Knowl. Data Eng., vol. 26, no. 5, pp. 1158–1170, May 2014.
[7] Wikipedia. (2014). Twitter [Online]. Available: http://en.wikipedia. org/wiki/Twitter
[8] Xinhuanet. (2014). Sina Microblog Achieves over 500 Million Users [Online]. Available: http://news.xinhuanet.com/tech/ 2012–02/29/c_122769084.htm
[9] D. Perito, C. Castelluccia, M. A. Kaafar, and P. Manils, "How unique and traceable are usernames?" in Proc. 11th Int. Conf. Privacy Enhancing Technol., 2011, pp. 1–17.
[10] J. Liu, F. Zhang, X. Song, Y. I. Song, C. Y. Lin, and H. W. Hon, "What's in a name?: An unsupervised approach to link users across communities," in Proc. 6th ACM Int. Conf. Web Search Data Mining, 2013, pp. 495–504.
[11] A. Acquisti, R. Gross, and F. Stutzman, "Privacy in the age of augmented reality," in Proc. Nat. Acad. Sci., 2011, pp. 36–53, Available: https://www.usenix.org/legacy/events/sec11/tech/slides/ acquisti.pdf.