International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 06 | June 2021 www.irjet.net p-ISSN: 2395-0072

ANALYZING RISK ON SOCIAL NETWORK

Surya Acharya[1] , Neetnav Kayasth[2] , Shivam Kanungo[3] , Gurpreet Singh Saini[4] ,

Professor. Reetika Kerketta[5]

[1-4]*Student, Dept. Information Technology, MIT School of Engineering, Maharashtra, India*
[5]*Professor, Dept. Information Technology, MIT School of Engineering, Maharashtra, India*

Abstract - *Despite the tremendous increase in OSN usage, several security and privacy problems remain. In such a circumstance, having a method that can assign a risk score to each OSN user would be quite useful. We offer a risk assessment in this study based on the premise that the more a user's conduct deviates from what may be deemed "typical behavior," the higher the danger risky. We did this while keeping in mind that the OSN community is extremely diverse in terms of observed behaviors. As a result, defining a single standard behavioral model that suits all OSN users' behaviors is impossible. However, we expect similar people to follow similar rules based on similar behavioral models' outcomes. As a result, we advise conducting a risk assessment. structured into two phases: related users are initially grouped together, and then one or more models for normal behavior are built for each identified group. Experiments on a real Facebook dataset indicate that the proposed model outperforms a simplified behavioral-based risk assessment that builds behavioral models over the entire OSN population without going through a group identification step.*

1.INTRODUCTION

ONLINE Social Networks (OSNs) let users to build public or private profiles, stimulate sharing of information and interests with other users, and allow users to communicate with one another. As a result, OSNs are currently utilized by millions of people and have become an integral part of our daily lives. People use OSNs to communicate with family and friends, as well as to share personal information and do business. Over time, users of an OSN form bonds with their friends, coworkers, and other people. These links constitute a social graph, which governs how information is disseminated throughout the social network. Although OSN usage is on the rise — Facebook, for example, now has 1.55 billion monthly active users, 1.31 billion mobile users, and 1.01 billion daily users1 – there are also a slew of security and privacy concerns. One of the main sources of these worries is that OSN users form new interactions with strangers, exposing a large quantity of personal information. Unfortunately, many users are unaware of their exposure, as well as the potentially dangerous effects. As a result, today's social media platforms are vulnerable to a variety of privacy and security threats. These attacks take advantage of OSN infrastructures to acquire and expose personal information about users, for example, by luring them to click on malicious links with the goal of spreading them around the network.

1.1 RISK ASSESSMENT BASED ON USER BEHAVIORS

Our main goal, as stated in the preceding section, is to assign a risk score to a user depending on how he or she behaves in the OSN. More specifically, the key premise is that the more a user's behavior deviates from what may be termed "normal behavior," the more dangerous it is. As a result, we must first construct a user behavioral profile capable of ring those users' actions and contacts that we believe are relevant to risk assessment.

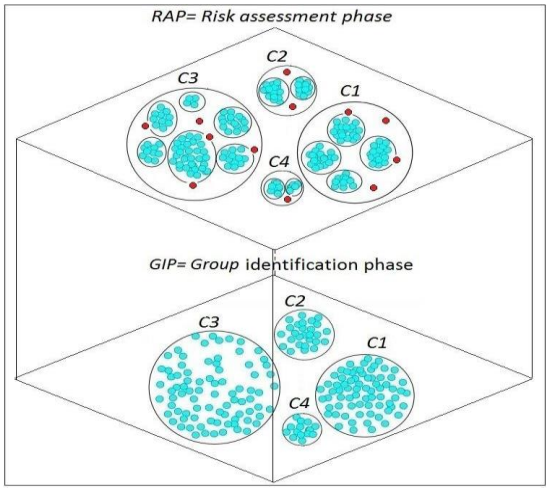

Fig-1: Two phase risk assessment

1.2 RISKY BEHAVIORS IN OSNS

As social networking sites grew in popularity, fraudsters and attackers began to use them to spread malware and perpetrate scams [13]. In general, the activity patterns that underpin assaults differ from those of normal users. The disparity is measured in terms of activity frequency, number, and type. In the following, we illustrate the most notable types of attacks: -

Sybil Attacks: Sybil attacks are one of the most common and effective OSN attacks. On Facebook, for example, over a hundred sociable have been discovered. To carry out a sybil attack, a malicious user must construct many phone identities, also known as sybils, with the goal of legitimizing his or her identity, in order to gain unfairly more authority and influence.

Identity Clone Attacks: Attacks using a cloned identity. In this form of attack, a malicious user creates profiles that are similar or even identical to victims in an OSN. After

successfully forging a victim, the main purpose is to collect personal information about the victim's acquaintances and to create greater levels of trust with the victim's social circle in preparation for future deceptions.

Compromised account attacks: Attacks on accounts that have been compromised. Legitimate users who have lost entire or partial control of their login credentials have accounts that have been compromised. Accounts can be hacked in a variety of methods, including by exploiting a cross-site scripting vulnerability or by phishing for the user's login credentials.

Creeper Attacks: Creepers are attacking. Creepers are real people who are abusing the OSN's features. They might, for example, send friend requests to many strangers or send out spamming chain letters.

Cyberbullying Attacks: Attacks on cyberbullying are on the rise. In online social networks, cyberbullying has grown frequent. Sexual insults, threats, or repeated harmful communications are used by attackers to torment their victims (typically youngsters and teenagers).

Clickjacking attacks: Clickjacking is a type of cyber-attack. In this type of attack, attackers persuade users to click items that aren't what they planned to click. The attacker can then take control of the user's account by sending spam messages and liking certain items.

2. FEATURE DESCRIPTION

In this section we'll describe the set of features used in out two-phase risk assessment approach.

2.1 Group Identification Features

We recall that the goal of the first clustering is to group users who are predicted to behave similarly. Group identification (GI) features should be those that are highly discriminatory, such as age and gender, as well as those that have an impact on potential users' actions, such as education and nationality. In the actual world, persons with similar backgrounds tend to behave similarly; but, in an OSN, this may be influenced by users' attitudes regarding online social networks, which may differ even among comparable users.

2.2 Behavioral Features

Friendship Rate (FR): Having many friends is not always indicative of risky conduct. However, studies suggest that attackers are more aggressive than average users when it comes to forming new friendships.

$$FR(\overline{u}) = \frac{|Friends(\overline{u})|}{UserLongevity(\overline{u})}$$

Mutual Friendship Rate (MFR): This feature calculates the average number of mutual friends shared by a target user u and all his or her network friends. We chose this functionality because attackers issue friendship requests to many strangers in various assaults (e.g., social bots, sybil attacks). As a result, friendship graphs are sparse, with no mutual friends.

$$MFR(\overline{u}) = \frac{\sum_{\breve{u} \in Friends(\overline{u})} |MutualFriends(\overline{u}, \breve{u})|}{|Friends(\overline{u})|}$$

Friend Mutual Friend Ratio (FMFR): The Friend Mutual Friend Ratio is a measure of how many friends you have in common (FMFR).As an example, consider fraudulent accounts, where attackers have a small community and send a large number of friendship invitations to strangers.

$$FMFR(\overline{u}) = \frac{|FR(\overline{u})|}{MFR(\overline{u})}$$

Comment Rate (CR): Commenting on postings is another typical activity in an OSN. We identify a collection of features to measure the user behaviours connected to comments and posts based on their relevancy. CR calculates the quantity of comments submitted by each target user u in relation to his/her lifetime.

$$CR(\overline{u}) = \frac{|CommentsBy(\overline{u})|}{UserLongevity(\overline{u})}$$

Comments feedback Ratio (CFR): It is designed to identify people that have a low no. of likes in comparison to the number of comments they leave. This is due to the fact that attackers are more aggressive in transmitting messages in some attacks, such as cyberbullying and sybil attacks.

$$CFR(\overline{u}) = \frac{\frac{|CommentsBy(\overline{u})|}{AvgFeedbackOnComment(\overline{u})}}{UserLongevity(\overline{u})}$$

Post-Feedback Ratio (PFR): It can be symptomatic of unsafe conduct if a user has a short lifespan and uploads a lot of items with few likes and comments. Because most types of assaults try to post items either directly or indirectly by compromising legitimate users in the network, we only examine the number of likes and comments on postings when using this feature.

$$PFR(\overline{u}) = \frac{\frac{|PostsBy(\overline{u})|}{AvgFeedbackOnPost(\overline{u})}}{UserLongevity(\overline{u})}$$

Like Rate, Like Propagation (LRLP): This functionality can aid in the detection of creeper, clickjacking, and socware attacks, in which attackers attempt to disseminate a large number of objects in the network at a fast rate.

$$LRLP(\overline{u}) = \frac{\sum_{\forall i \in ItemsLikedBy(\overline{u})} (|LikesOn(i)| + |PostsOn(i)|)}{ItemLongevity(i)} * \frac{|LikesBy(\overline{u})|}{UserLongevity(\overline{u})}$$

Post like post propagation (PLPR): We're interested in modelling the amount of posts created by a target user, as well as the speed at which these messages spread because attackers in some types of attacks have a large number of posts and these postings propagate quickly.

$$PRPP(\overline{u}) = \frac{\sum_{\forall i \in PostedBy(\overline{u})}(|LikesOn(i)| + |PostsOn(i)|)}{ItemLongevity(i)} * \frac{|PostsBy(\overline{u})|}{UserLongevity(\overline{u})}$$

3. TWO PHASES CLUSTERING

We recall that our risk assessment consists of two phases, the first of which aims to organize users according to group identification traits (i.e., the first phase), and the second of which aims to organize users according to behavioral aspects (i.e., the second phase) (i.e., second phase). Regardless of the features taken into account, we employ the same clustering approach in both of these phases. Cluster algorithms are divided into two categories. The best cluster is determined in a deterministic manner using hard clustering techniques (e.g., kmeans), which means that each item is assigned to a unique cluster. Soft clustering (i.e., probabilistic-based clustering) computes the membership probability for each item and each available cluster.

3.1. Probability-based clustering

The values in the user features vector, namely GI features in the first phase and behavioral features in the second, are used to calculate membership likelihood. We employ the most prominent approach, Expectation-Maximization (EM) [7], which uses probability estimates via an iterative procedure to conduct probabilistic clustering. The probability-based clustering for a generic feature vector of user u is described in the following sections. In the first phase, this vector will contain the values of the user's GI features, whereas in the second phase, it will contain the values of the BFs. Let N denote the number of users in the OSN, and u denote the number of characteristics vectors in cluster l.

$$w_l(\vec{u}) = \frac{w_l.p_l(\vec{u}|\theta_l)}{\sum_{i=1}^{K} w_i.p_i(\vec{u}|\theta_i)}$$

3.2. User Risk Score

The goal is to regard users who deviate from regular behaviour as more dangerous, as explained throughout the study. The membership probability generated in the second clustering step truly capture these variations. A high membership probability number indicates that the target user fits well with one of the behaviours that emerged from the group to which he or she belongs.

Definition **-User Risk Score (RS) :** Let N denote the number of users in the OSN, and Ng N denote the number of users in the same cluster g, as determined by the probabilistic based clustering computed in the first phase over the GI feature values of users in N. Let PB(Ng) be the probability-based clustering algorithm that takes a set of users in Ng as input and returns, for each u, a probability-based grouping algorithm based on their BF. provides the highest membership probability, indicated as PCL, for each user u 2 Ng (u). The related risk score RS(u) 2 [0; 1] for a target user u 2 Ng is defined as: RS(u) = 1 PCL (u)

4. EXPERIMENTS:

In this section we'll demonstrate experiments conducted to show the efficiency of two-phase risk assessment.

4.1. Two-Phase vs. One-Phase Risk Assessment

The first and second trials are designed to validate the concept of obtaining model behaviours from user groups. We compare the two-phase risk assessment approach given in this research to a risk assessment that only considers the behavioral features presented without grouping users first. One-phase risk assessment is how we refer to the model. The risk score of a target user u returned by onephase risk assessment is determined as RS(u) = 1- PCL(u), where PCL(u) is defined as the highest of the membership probability values associated with u and returned by one-phase clustering, similar to the two-phase approach. The Fmeasure, detection rate, and false alarm rate are used to assess the performance of our risk estimate method (false positive). The weighted harmonic mean of two measurements, accuracy and detection rate, is the F-measure. The precision is calculated by dividing the number of accurately identified risky users by the number of injected fake and regular users who are identified as risky. The number of accurately recognized dangerous users divided by the total number of injected fraudulent users is the detection rate. Precision and detection rate scores are rarely mentioned separately. Instead, both are merged into a single statistic known as the F-measure. The F-measure has a maximum value of 1 and a minimum value of 0. The rate of false alarms (false positives) is the ratio of the number of normal events to the number of false alarms. Three models were used to create fake users to be injected into the genuine Facebook dataset. The first model

represents risky users with Randomized Features (RF), or users whose
GI feature and BF values are selected at random.

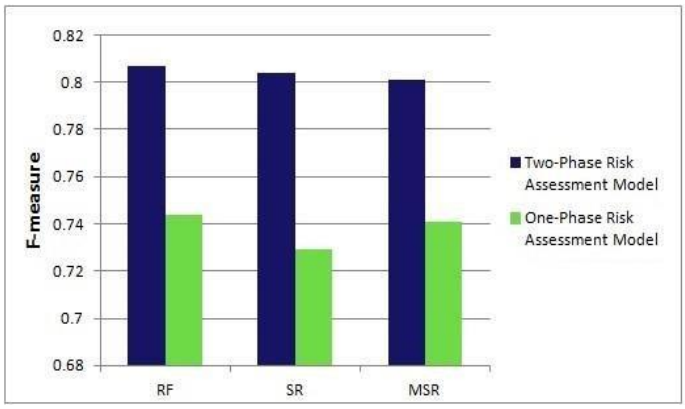

Fig-2: Comparison of F-measure for two-phase vs. one-phase risk assessment

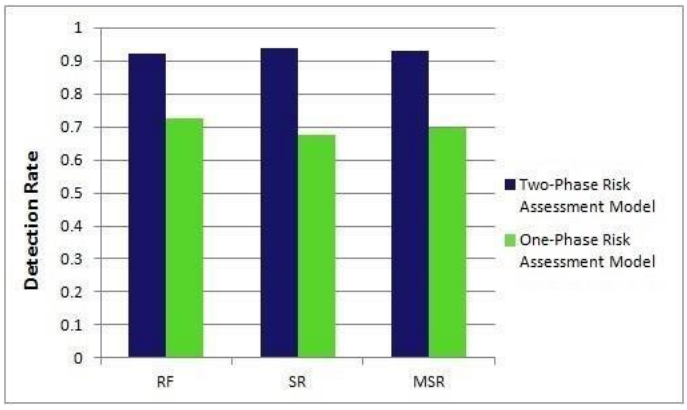

Fig-3: Comparison of Detection Rate for two-phase vs. one-phase risk assessment

More Smarter Risky User (MSU) is a model in which false users are produced with authentic values for profile and friendship features and randomised values for the remaining behavioural features. Friendship characteristics include FR, MFR, and FMFR picked from BFs, as well as number of friends selected from GI features. For these three types of fraudulent users, Figures 2, 3, and 4 show the F-measure value, detection rate, and false alarm rate, respectively. On the basis of F-measure, detection rate, and false positive measure, two-phase risk assessment outperforms one-phase risk assessment, as shown in Figures 2, 3, and 4.

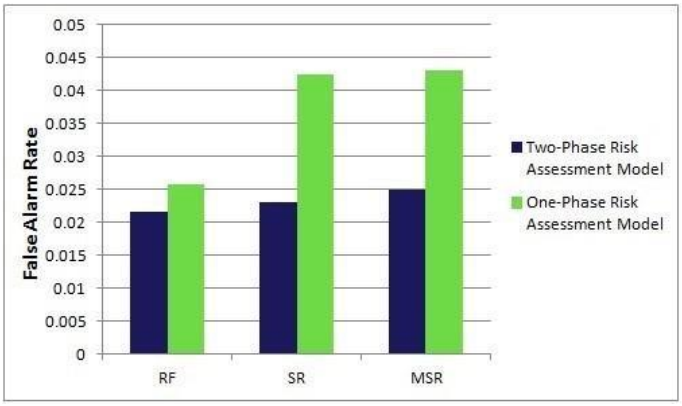

Fig. 4: Comparison of False Alarm Rate for two-phase vs. one-phase risk assessment

4.2. Risk assessment vs. Attacks

The goal of this experiment is to demonstrate how our users' risk scores can be used to identify high-risk users. We did this by injecting fake users into the real dataset to imitate the behavioral patterns of each sort of attack mentioned in Section 2. We conducted this experiment with the smallest dataset, FB75, because users in this dataset have also posted information, making it possible to conduct it. Create ten distinct test datasets for ten various types of attacks (for example, sybils (dense graph), sybils (sparse graph), socware, and so on). We insert 55 real users and 15 bogus users into each one. We create normal users in each of the ten datasets so that each of their feature values is chosen at random within the relevant standard deviation in real FB75. Fake users, on the other hand, are customized. As a result, we randomized the value of these anomalous characteristics in the test set outside the corresponding standard deviation in real FB75. We then test our algorithm to see if these unusual users can be identified. Experiments have been conducted for both the two-phase and one-phase risk assessments. We also looked at the effectiveness by taking into account numerous different values for the threshold. As a result, we received different F-measures, detection rates, and false positive rates for each type of attack. Figure 5 illustrates the F-measure for Social bots or sybils (dense friendship graph) assaults, with the F-measure value on the Y-axis and the threshold value on the X-axis.

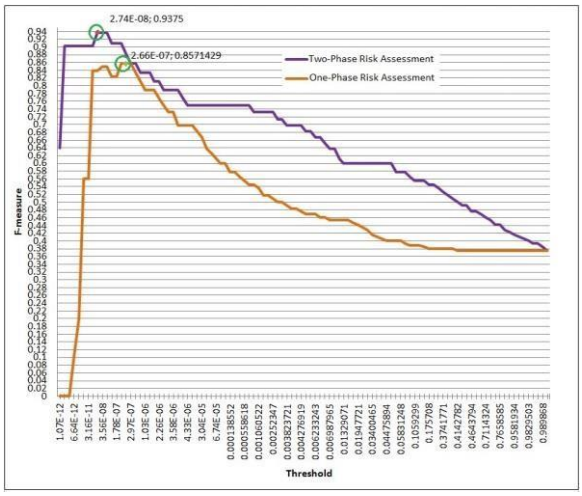

Fig. 5: F-measure for social bots or sybils (dense friendship graph)

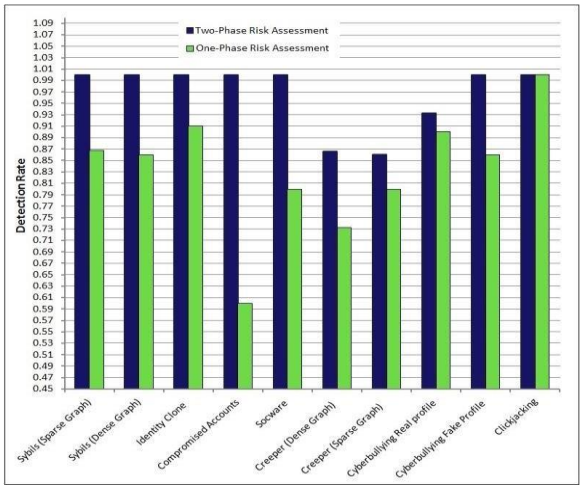

Fig. 6: Detection rate for all types of attacks

5. RELATED WORK

Researchers have begun to suggest several detection and mitigation ways to handle the rising problem of finding harmful activities on social networks. Several methods for detecting assaults and false accounts in an OSN have recently been proposed. These are based on graphs or behaviors. The expansion of the OSN graph is assumed in graph-based sybil detection techniques. Researchers have developed algorithms for sybils identification using various graph analysis approaches, such as sybilGuard, sybil-Limit, sybilinfer, and SumUp. Recent research has shown, however, that these assumptions may not always be correct. Sybils do, in fact, mix well with the rest of OSN graphs, and the majority of OSN graphs are not fastmixing. supervised learning techniques are used in most current behavior-based algorithms for detecting anomalous users in OSNs .For instance, in , the proposed system trains a classifier by extracting four features, such as accepted incoming requests, accepted outgoing requests, invitation frequency, and clustering coefficient, to recognize sybils. Using a small manually labelled dataset of genuine and anomalous users, the authors presented a supervised technique to detect compromised account attacks in. To detect spam and malware, employed classifiers, respectively.

6. CONCLUSION

We suggested a two-phase risk assessment approach in this paper that may assign a risk score to each OSN user. This risk assessment is based on the user's behaviour, with the premise that the more the conduct deviates from what is deemed "normal," the more the user should be regarded risky. Experiments on a genuine Facebook dataset demonstrate the efficacy of our solution. We intend to expand on this work in a number of ways. The extension of the proposed two-phase risk assessment to allow for continuous monitoring and estimation of risk scores is an exciting future project. Furthermore, we intend to adapt the risk assessment model so that it may be used in Decentralized Online Social Networks, which are characterised by the lack of a central authority. To acquire user features, this will necessitate looking into decentralised data mining methods.

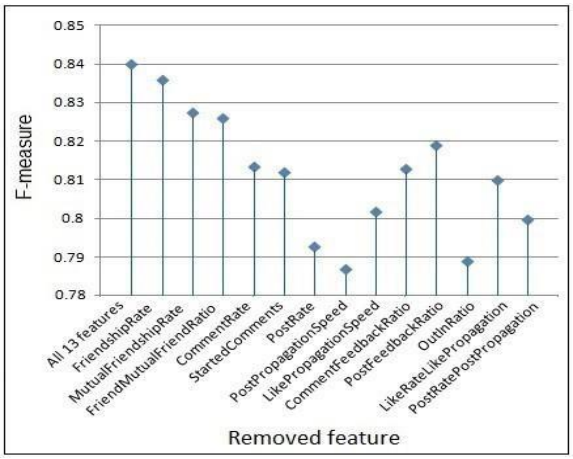

Fig-7: The best F-measure by removing one Behavioural Feature at a time

ACKNOWLEDGEMENT

We would like to thank the anonymous reviewers and the editor for their insightful remarks and ideas, as well as the opportunity to improve our paper.

REFERENCES:

[1] Cuneyt Gurcan Akcora, Barbara Carminati, and Elena Ferrari. Privacy in socialnetworks how risky is your social graph? In Data Engineering (ICDE), 2012 IEE 28th International Conference on, pages 9–19. IEEE, 2012.

[2] Christa SC Asterhan and Tammy Eisenmann. Online and faceto- face discussions in the classroom: A study on the experiences of'active'and'silent'students. In Proceedings of the 9th international conference on Computer supported collaborativelearning-Volume 1, pages 132–136. International Society of the Learning Sciences,2009.

[3] Leyla Bilge, Thorsten Strufe, Davide Balzarotti, and Engin Kirda. All yourcontacts are belong to us automated identity theft attacks on social network. InProceedings of the 18th international conference on World wide web, pages 551–560.ACM, 2009.

[4] Yazan Boshmaf, Konstantin Beznosov, and Matei Ripeanu. Graphbased sybildetection in social and information systems. In Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on, pages 466–473. IEEE, 2013.

[5] Yazan Boshmaf, Dionysios Logothetis, Georgos Siganos, Jorge Ler´ıa, Jose Lorenzo, Matei Ripeanu, and Konstantin Beznosov. ´Integro: Leveraging victim prediction for robust fake account detection in osns. In Proc. of NDSS, 2015.

[6] Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu. The socialbot network- when bots socialize for fame and money. In Proceedings of the 27th Annual Computer Security Applications Conference, pages 93–102. ACM, 2011.

[7] Paul S Bradley, Usama Fayyad, and Cory Reina. Scaling em (expectationmaximization) clustering to large databases. Technical report, Technical Report MSRTR-98-35, Microsoft Research Redmond, 1998.

[8] Qiang Cao, Michael Sirivianos, Xiaowei Yang, and Tiago Pregueiro. Aiding the detection of fake accounts in large scale social online services. [9] George Danezis and Prateek Mittal. Sybilinfer- detecting sybil nodes using social networks. In NDSS, 2009.

[9] Vacha Dave, Saikat Guha, and Yin Zhang. Measuring and fingerprinting click-spam in ad networks

[10] Vacha Dave, Saikat Guha, and Yin Zhang. Viceroi catching clickspam in search ad networks. In Proceedings of the 2013 ACM SIGSAC conference on Computer and communications security, pages 765–776. ACM, 2013.

[11] Arthur P Dempster, NanMLaird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society. Series B(Methodological), pages 1–38, 1977.

[12] Manuel Egele, Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna. Compa- detecting compromised accounts on social networks. In NDSS, 2013.

[13] Mohammad Reza Faghani and Hossein Saidi. Malware propagation in online social networks. In Malicious and Unwanted Software (MALWARE), 2009 4th International Conference on, pages 8–14. IEEE, 2009.

[14] Michael Fire, Roy Goldschmidt, and Yuval Elovici. Online social networks- threats and solutions. 2013.

[15] Carlos A Freitas, Fabrcio Benevenuto, Saptarshi Ghosh, and Adriano Veloso. Reverse engineering socialbot infiltration strategies in twitter. arXiv preprint arXiv:1405.4927, 2014.

[16] Hongyu Gao, Jun Hu, Christo Wilson, Zhichun Li, Yan Chen, and Ben Y Zhao. Detecting and characterizing social spam campaigns. In Proceedings of the 10th ACM SIGCOMM conference on Internet measurement, pages 35–47. ACM, 2010. [18] Chris Grier, Kurt Thomas, Vern Paxson, and Michael Zhang. Spam: the underground on 140 characters or less. In Proceedings of the 17th ACM conference on Computer and communications security, pages 27–37. ACM, 2010. [19] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In Proceedings of the 2005 ACM workshop on Privacy in the electronic society, pages 71–80. ACM, 2005.

[17] Ting-Kai Huang, Md Sazzadur Rahman, Harsha V Madhyastha, Michalis Faloutsos, and Bruno Ribeiro. An analysis of socware cascades in online social networks. In Proceedings of the 22nd international conference on World Wide Web, pages 619–630. International World Wide Web Conferences Steering Committee, 2013.

[18] Lei Jin, Xuelian Long, Hassan Takabi, and James BD Joshi. Sybil attacks vs identity clone attacks in online social networks. ISA, 2012.

[19] Lei Jin, Hassan Takabi, and James BD Joshi. Analysing security and privacy issues of using e-mail address as identity. International Journal of Information Privacy, Security and Integrity, 1(1):34–58, 2011. Kyumin Lee, James Caverlee, and Steve Webb. Uncovering social spammers- social honeypots+ machine learning. In Proceedings of the 33rd international ACM SIGIR conference on Research.

BIOGRAPHIES

Surya Acharya
Student,
MIT School of Engineering,
Pune

Neetnav Kayasth
Student,
MIT School of Engineering,
Pune

Shivam Kanungo
Student,
MIT School of Engineering,
Pune

Gurpreet Singh Saini
Student,
MIT School of Engineering,
Pune

Prof. Reetika Kerketta
Professor,
MIT School of Engineering,
Pune