Analyzing and Detecting Money-Laundering Accounts in Online Social Networks

Yadong Zhou, Ximi Wang, Junjie Zhang, Peng Zhang, Lili Liu, Huan Jin, and Hongbo Jin

Abstract

Virtual currency in OSNs plays an increasingly important role in supporting various financial activities such as currency exchange, online shopping, and paid games. Users usually purchase virtual currency using real currency. This fact motivates attackers to instrument an army of accounts to collect virtual currency unethically or illegally with no or very low cost and then launder the collected virtual money for massive profit. Such attacks not only introduce significant financial loss of victim users, but also harm the viability of the ecosystem. It is therefore of central importance to detect malicious OSN accounts that engage in laundering virtual currency. To this end, we extensively study the behavior of both malicious and benign accounts based on operation data collected from Tencent QQ, one of the largest OSNs in the world. Then, we devise multi-faceted features that characterize accounts from three aspects: account viability, transaction sequences, and spatial correlation among accounts. Finally, we propose a detection method by integrating these features using a statistical classifier, which can achieve a high detection rate of 94.2 percent at a very low false positive rate of 0.97 percent.

Introduction

Online social networks (OSNs) have started to leverage virtual currency as an effective means to glue financial activities across various platforms such as online shopping, paid online games, and paid online reading. Examples of virtual currency in such OSNs include but are not limited to Tencent Q Coin, Facebook Credits[1], and Amazon Coin. Usually, users purchase virtual money using real currency at a regulated rate; one user can also transfer it to another user via various methods such as recharging their account and sending gifts [1]. These facts enable attackers to gain potentially massive profits through the following steps. First, an attacker can collect virtual currency with zero or low cost. For example, they can compromise and subsequently control a legitimate account or register a huge number of accounts to win gifts (in the form of virtual currency) in online promotion activities. Next, they can instrument accounts under their control to transfer virtual currency to other accounts in return for real currency, with rates that are usually much lower compared to the regulated rate. Attackers usually post advertisements in popular e-commerce websites [2] to attract potential buyers. We call OSN accounts that are used by attackers for the collection and transfer of virtual currency *money-laundering accounts*. Money-laundering accounts have caused a tremendous financial loss for compromised accounts, fundamentally undermined the effectiveness of online promotion activities, and possibly introduced potential conflicts against currency regulations.

Detecting money-laundering accounts in OSNs therefore becomes of essential importance, which, however, is faced with new, significant challenges. First, committing money-laundering activities does not require the use of traditional malicious content such as spam, malicious URLs, or malicious executables. Although spamming might be used by attackers for advertising, neither methods nor the accounts used for spamming are necessarily associated with the money-laundering accounts. Second, money-laundering activities do not rely on social behavior and structures (e.g., "following" or "friend" relationship in popular social networks) to operate. These challenges make existing methods immediately ineffective, since they focus on detecting OSN-based spamming, phishing, and scamming attacks, whose proper operation necessitates malicious content [3, 4], social structures [5], or social behaviors [6].

Detecting money laundering activities in traditional financial transactions has attracted significant research efforts [7]. For example, Dreewski *et al.* [8] designed a system to detect money laundering activities from billings and bank account transactions. Paula *et al.* [9] used the AutoEncoder to classify exporters and detect money laundering activities in exports of goods and products in Brazil. Colladon *et al.* [10] presented predictive models to quantify risk factors of clients involved in the factoring business and proposed a visual analysis method to detect the potential clusters of criminals and prevent money laundering. Different from traditional money laundering detection problems in bank-related activities, account behaviors of laundering virtual currency in OSNs involve bank-related financial activities, online social networks, and virtual recharging and expenditure activities.

The goal of our work is to design an effective method capable of detecting money-laundering accounts. As a means toward this end, we perform an extensive study of behaviors of money-laundering accounts based on data collected from Tencent QQ, one of the largest OSNs in the world with a giant body of reportedly 861 million active users. We have devised multi-faceted features that characterize accounts from three aspects: account viability, transaction sequences, and spatial correlation among accounts. Experimental results have demonstrated that our method can achieve a high detection rate of 94.2 percent with a very low false positive rate of 0.97

Yadong Zhou, Ximi Wang, Peng Zhang, and Lili Liu are with Xi'an Jiaotong University.

Junjie Zhang is with Wright State University.

Huan Jin and Hongbo Jin are with Tencent Technology (Shenzhen) Company Ltd.

[1] Facebook Credits has been replaced by Facebook Game Cards since 2013. After redeeming, Facebook Game Cards also can be used to buy virtual items and send them to other accounts as gifts in online games.

Digital Object Identifier: 10.1109/MNET.2017.1700213

percent. To the best of our knowledge, this work represents the first effort to analyze and detect money-laundering accounts in OSNs integrating virtual currency at this large scale.

Data Set

We have collected labeled data from Tencent QQ, a leading online social network in China, which offers a variety of services such as instant messaging, voice chat, online games, and online shopping. These services are glued together using Q coin, the virtual currency distributed and managed by Tencent QQ. Tencent QQ has a giant body of 861 million active accounts with a reported peak of 266 million simultaneously online users. Also, Tencent QQ is one of the leading OSNs that are actively involved in virtual currency based services in the world.

Our data set is composed of 114,891 malicious accounts and 381,523 benign accounts that are active during the first week of August in 2015. In order to label accounts used for money laundering, we follow advertisements of cheap virtual currency in major e-commerce websites and actually purchased virtual currency from sellers, where QQ accounts used by these sellers are labeled as money-laundering accounts. Since an attacker usually controls a large number of malicious accounts for money laundering, we label accounts as malicious if they login from the same IP address used by a confirmed money-laundering account within one day.

Although this labeling process offers us the ground truth, using it as a detection method is practically challenging. First, it requires a considerable amount of investment to engage money-laundering accounts in malicious activities. Second, the IP addresses used to label launder accounts usually will be invalid after a few days, because attackers change the login IP addresses frequently. Therefore, this data labeling process, if used as a detection method, cannot guide OSNs to mitigate their financial loss proactively. For each account, we collect the following activity records. It is worth noting that all these records can be collected from social networks that integrate virtual currency.

- Login activities, which include the account ID, the login date, the login IP address, and the account level.
- Expenditure activities, which include the expenditure account ID, the expenditure date, the expenditure amount, the purchased service, the payment method, and the account ID to receive the service.
- Recharging activities, which include the recharging account ID, the recharging date, the recharging amount, and the payment method.

Behavior Analysis and Feature Extraction

Figure 1 shows a typical process of virtual currency laundering. The first step is to collect virtual currency with zero or extremely low cost. For example, attackers can hack users' accounts (and thus control their virtual currency), exploit the system vulnerabilities, or participate in online promotion activities to win virtual currency for free or at significantly discounted rates [2]. Next, attackers attract potential buyers with considerable discounts, through various ways such as spreading spams and posting advertisements, and then sell the virtual currency in popular e-commerce websites such as eBay or Taobao. Once a buyer commits the purchase (i.e., pays real money to an attacker through the e-commerce websites), their account will receive virtual currency (e.g., as gifts) from one or multiple malicious accounts controlled by an attacker. Since OSNs may investigate an account if it has initiated a large number of transactions in a short period of time, an attacker usually distributes their virtual currency across multiple accounts and uses them alternatively to transfer virtual currency to buyers.

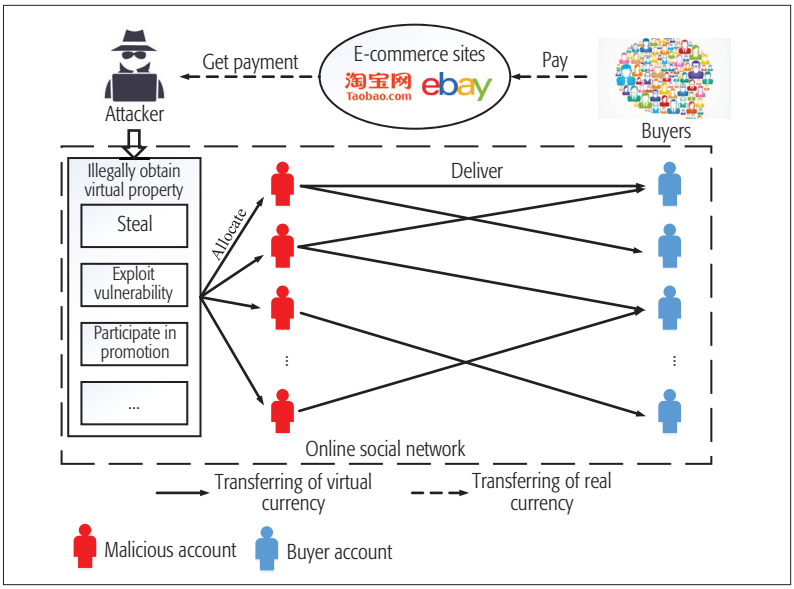

FIGURE 1. Malicious laundering process of virtual currency.

Vitality Features

To avoid detection, attackers usually disguise the anomaly behavior of the malicious accounts. However, some typical behavior patterns are unavoidable to achieve the goal of laundering. We can still design several effective vitality features to distinguish the malicious and benign accounts.

Regular users usually actively use their OSN accounts for various daily activities such as chatting, photo sharing, and finance. In contrast, malicious accounts are mostly driven by transactions for money laundering, which are much less active compared to benign accounts. Therefore, we define the following two features to capture such difference.

Feature 1: The Ratio of Active Days: This feature represents the ratio of active days of an account during the past year. Specifically, if an account is logged in at least once for one day, this day will be labeled as "active" for this account.

Feature 2: Account Level: The OSN assigns a level for each account to characterize its activity, which is usually measured by the total number of active days since the account was registered.

Figures 2a and 2b show that benign accounts are much more active than malicious accounts. Specifically, most malicious accounts (approximately 97 percent) are active for less than 10 percent of total days whereas only a small per-

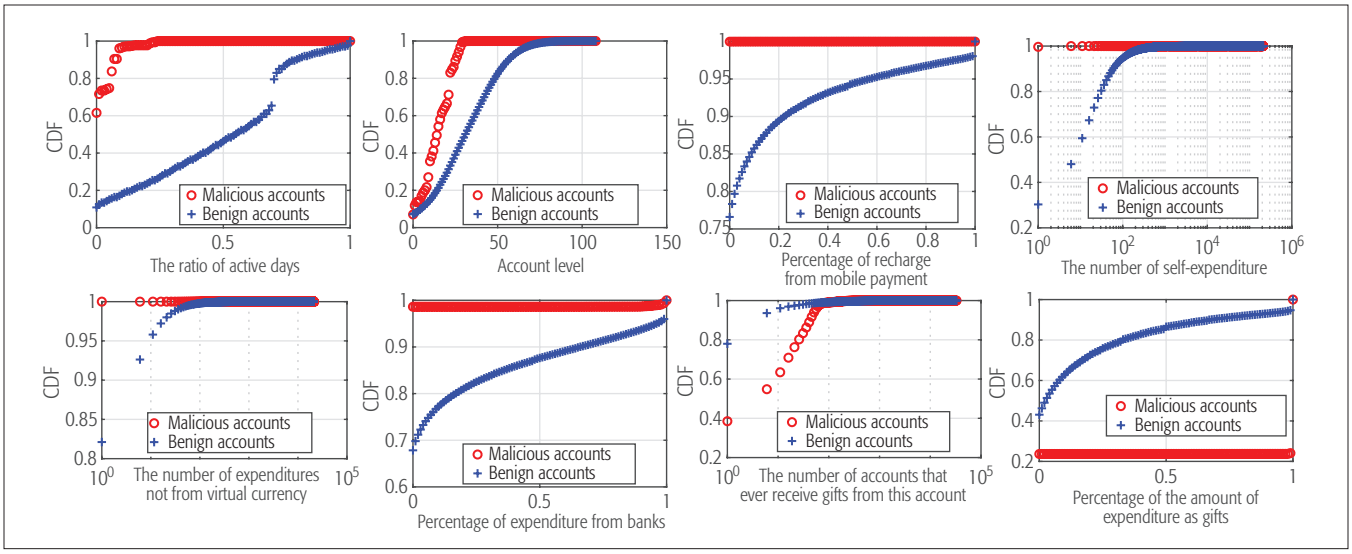

FIGURE 2. Comparisons of vitality features of malicious and benign accounts: a) Feature 1; b) Feature 2; c) Feature 3; d) Feature 4; e) Feature 5; f) Feature 6; g) Feature 7; h) Feature 8.

centage of benign accounts (less than 20 percent) experience the same activity level (i.e., being active for less than 10 percent of total days).

Next, we study the source of virtual currency for benign and laundering accounts. A benign user usually recharges their account via wire transfer (often in the form of mobile payment) and occasionally receives gifts (from friends). Comparatively, money-laundering accounts almost exclusively rely on online promotions to directly collect virtual currency or gifts transferred from other accounts. We therefore introduce the following feature to characterize the currency collection behavior.

Feature 3: Percentage of Recharge from Mobile Payment: This feature represents the percentage of virtual currency recharged through mobile payments (i.e., purchasing virtual currency using mobile online banks).

Figure 2c presents the distribution for this feature, where approximately 24 percent of benign users recharge their accounts via mobile payment, while the vast majorit y of malicious accounts do not use this channel.

As an increasing number of financial functions are integrated into social networks, users conduct a variety of activities such as shopping and gifting. While benign users prefer to engage financial activities with higher diversity, money-laundering accounts only focus on activities relevant to laundering. Therefore, we introduce the following five features to characterize such a difference.

Feature 4: The number of Self-Expenditures: This feature represents the total number of expenditures that an account has committed to itself using virtual currency.

Feature 5: The Number of Expenditures Not from Virtual Currency: This feature characterizes the number of expenditures an account has committed by other methods instead of virtual currency.

Feature 6: Percentage of Expenditure from Banks: A user can associate their bank account with the OSN account. This bank account can be directly used for shopping and gifting in addition to virtual currency in the OSN account. This feature is defined as the percentage of expenditure from associated bank accounts.

Feature 7: The Number of Accounts that ever Receive Gifts from this Account: Malicious accounts need to frequently transfer the virtual currency as a gift to the buyer accounts, while a benign user tends to expend the virtual currency themself, and occasionally gives the virtual currency as a gift to their friends. Thus, malicious accounts will have a much larger value of this feature than benign users.

Feature 8: Percentage of the Amount of Expenditures as Gifts: This feature represents the proportion of the amount of expenditures as gifts in all expenditures. After malicious accounts collect virtual currency from the online promotion activities and other vulnerabilities, they will transfer it to other accounts as gifts. We therefore introduce this feature to quantify the percentage of all giving out behavior.

Figures 2d–2h report the distributions for Features 4–8 respectively. Almost all the malicious accounts (more than 99 percent) neither committed for itself using virtual currency nor committed by other methods instead of virtual currency. Comparatively, 61 percent of benign accounts have committed for itself using virtual currency at least once, and 18 percent of benign accounts have committed by other methods instead of virtual currency at least once. The distributions for Features 6–8 are also distinguishable as shown in the figures. We omit the descriptions for brevity.

Sequential Features of Financial Activities

The sequences of financial activities are likely to differ between benign accounts and money-laundering accounts. In order to model the sequential behavior, we use the discrete-time Markov Chain model. Specifically, we record the sequence of three basic financial activities: virtual-currency recharge, self-expenditures, and expenditures as gifts. Each state in the Markov Chain corresponds to one activity and the transition between two states represents a pair of two consecutive financial activities. Hence, the Markov Chain has three states and nine total transitions. Each transition

is associated with the probability of this transition among all observed transitions. Figure 3a illustrates how Markov Chain models are derived from a sequence of financial activities. Specifically, nodes 1′, 2′, and 3′ refer to the three states "virtual-currency exchange," "self-expenditure," and "expenditure as gifts"; P_{ij} denotes the transition probability from state i to state j.

Figure 3b presents the CDF of P_{11}, P_{31}, and P_{33} for malicious accounts (denoted as "MA") and benign accounts (denoted as "BA"), respectively. As shown in the empirical analysis, the values of P_{11} and P_{33} for malicious accounts are much larger than those for benign accounts, which indicates that malicious accounts are more inclined to exchange multiple times continuously (see P_{11}), and expend as gifts multiple times continuously (see P_{33}). The values of P_{31} of malicious accounts are much smaller than those of benign accounts, which implies that benign accounts are more active to recharge virtual currency after expending as gifts compared to malicious accounts. It is worth noting that we omit the other six transition probabilities in the figure for brevity.

Our empirical analysis demonstrates that the sequential behaviors indeed experience significant differences between malicious and benign accounts. Therefore, we define the following features.

Features 9–17: The transition probabilities P_{11}, P_{12}, P_{13}, P_{21}, P_{22}, P_{23}, P_{31}, P_{32}, P_{33}.

Features 18–47: The top 30 most effective subsequences mined from the sequence of financial activities for malicious accounts. To achieve an acceptable time complexity, the PrefixSpan algorithm [11] is used to mine the frequent subsequences of behavior sequences. Then, the effectiveness e of mined subsequence q is measured by Eq. 1. In the equation, f_q denotes the number of times that subsequence q occurs in all the sequences, and N_q^m and N_q^b denote the number of sequences of malicious accounts and benign accounts containing subsequence q, respectively.

$$e_q = f_q \frac{\left| N_q^m - N_q^b \right|}{N_q^m + N_q^b} \tag{1}$$

Spatial Features of Currency Transfer

Each transaction for currency transfer can be characterized as a tuple denoted as <s, t>, where s and t refer to the source and destination account, respectively. For a node s, we identify a set of nodes, to each of which s has transferred virtual currency. We denote this set of nodes as $D(s)$ for this node s. We then build a graph to offer a global overview of the currency transfer behaviors among all accounts. Specifically, we define a weighted undirected graph $G(V, E)$, where V and E indicate the vertex set and the edge set, respectively. Each vertex in V represents an account. An edge, e.g. (i, j), exists between two nodes i and j only if $D(i)$ AND $D(j)$!= null; we further assign the weight to this edge as $|D(i)$ AND $D(j)|$, which quantify the number of common transfer destinations between account i and account j.

The developed graph can effectively profile the behavior of coordinated laundering accounts. Specifically, an attacker usually distributes their virtual currency across multiple money-laundering accounts to reduce the risk of being detected and subsequently banned. As a result, when a buyer purchases virtual currency from the attacker, they usually need to instrument a set of money-laundering accounts to transfer currency to the buyer's account. As this process repeats for a large number of buyers, these money-laundering accounts will share a giant set of destination accounts, forming a fully-connected graph with high weight values for edges. A group of benign accounts may also transfer virtual currency to one or a few accounts (e.g., as birthday gifts) and thus form a fully connected graph, whose edges, however, are likely to have small weights. Since an account may receive gifts from both benign accounts (e.g., friend accounts) and money-laundering accounts, edges that connect benign and money-laundering accounts will also exist. To summarize, the graph is mainly composed of three types of connected subgraphs: subgraphs entirely composed of fully-connected malicious accounts, subgraphs entirely composed of fully-connected benign accounts, and subgraphs composed of both malicious accounts and benign accounts. Figure 4a presents one example of the third type of connected subgraphs. Specifically, malicious accounts A-D and C-E transfer to the same destination accounts respectively, a destination account obtains the virtual currency from both malicious account E and benign account F, and benign accounts F-I transfer to the same destination account. Then, the corresponding graph is a connected graph composed of both malicious accounts and benign accounts.

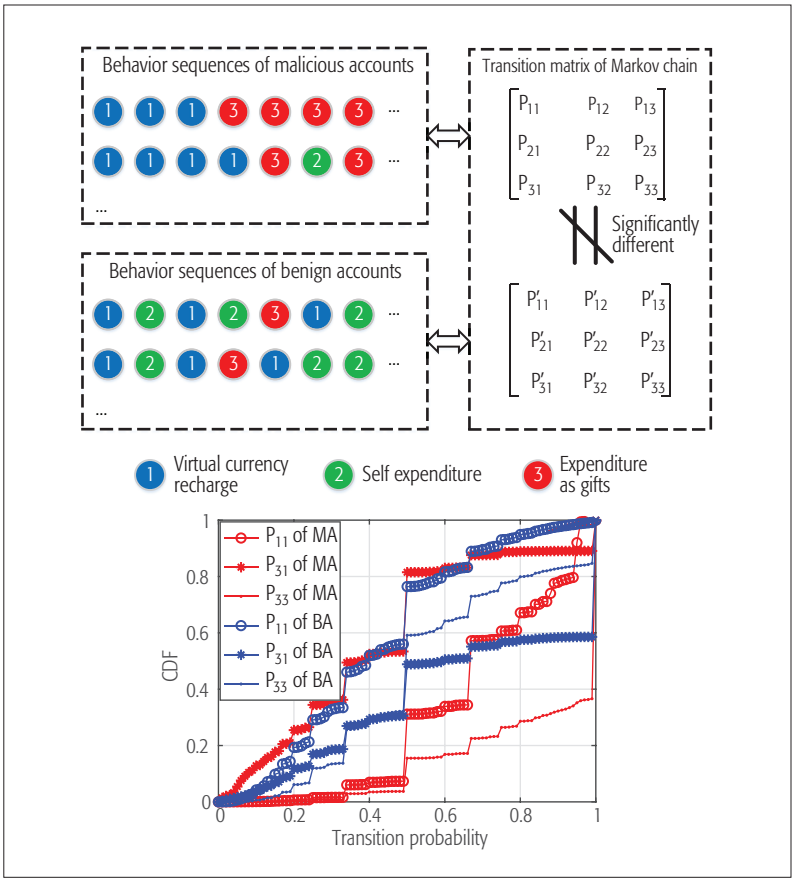

FIGURE 3. Analysis of behavior sequences of accounts; a) Illustration of behavior sequences of accounts; b) Distributions of transition probabilities.

We leverage machine learning techniques to integrate all these features to perform effective detection. Specifically, feature values extracted from labeled malicious and benign users have been employed to train a statistical classifier.

Through analyzing the behaviors of destination accounts, we find that most of the destination accounts as buyers tend to purchase the virtual currency or goods from the launder accounts rather than receiving gifts from benign accounts, and other destination accounts behave in the opposite way. This finding is validated by analyzing the neighbors of malicious accounts and benign accounts in the graph. It is analyzed that 80.1 percent of the neighbors of malicious accounts are malicious and 84.3 percent of the neighbors of benign accounts are benign on average. Thus, the malicious accounts and benign accounts tend to connect with the same type of vertices, and form a community structure in which some densely connected components are composed of the same type of vertices and the connections among the components are sparse. The interpretation of the forming of community structure in transferring related graph is shown in Fig. 4a, and a real illustration of the structure is shown in Fig. 4b, where the red vertex denotes a malicious account and the blue vertex denotes a benign account.

To design the spatial features, we process the behaviors of accounts in the following two steps.

Step 1: Form the graph based on the definition of $G(V, E)$.

Step 2: Detect the densely-connected subgraphs (communities) of the connected subgraphs of G based on a widely-used community detection method, Fast Unfolding [12]. The method is a heuristic method based on modularity optimization, and is capable of dealing with large weighted graphs due to its acceptable time complexity.

Following the above two steps, the graph G will be divided into many communities, each account will belong to a community, and each community will be composed of almost the same type (malicious or benign) of accounts. We present the features of each account (vertex) below.

Features of General Attributes of Vertex in Graph

- *Feature 48 — Degree:* The number of connected edges of the vertex.
- *Feature 49 — Weighted degree:* The summation of the weights of connected edges of the vertex.
- *Feature 50-51 — Average/variance of weights of edges:* The average/variance of the weights of connected edges of the vertex.
- *Feature 52 — Core number:* The highest order of a core that contains the vertex [13]. This feature represents the influence of the vertex in the graph and can be calculated with the time complexity of $O(m)$, where m is the number of edges in the graph.

Features of Community Attributes of Vertex in Graph

- *Feature 53 — Percentage of the number of expenditures as gifts in the community:* This feature is equal to $\Sigma_{i=1}^{N} P_i / \Sigma_{i=1}^{N} C_i$, where N denotes the number of vertices in the community, P_i denotes the number of expenditures as gifts of vertex i, and C_i denotes the number of all the expenditures of vertex i. Malicious accounts tend to mostly expend virtual currency as gifts, so the greater this feature of a community, the greater the likelihood that it would be malicious.
- *Feature 54 — Normalization of the number of destination accounts in the community:* This feature is equal to $\Sigma_{i=1}^{N} U_i / \Sigma_{i=1}^{N} P_i$, where U_i denotes the number of transferring destination accounts of vertex i, and the other variables are the same as above formula. Malicious accounts tend to transfer virtual currency as gifts to a large number of buyer accounts, thus this feature represents how likely the vertices in the community would be malicious.

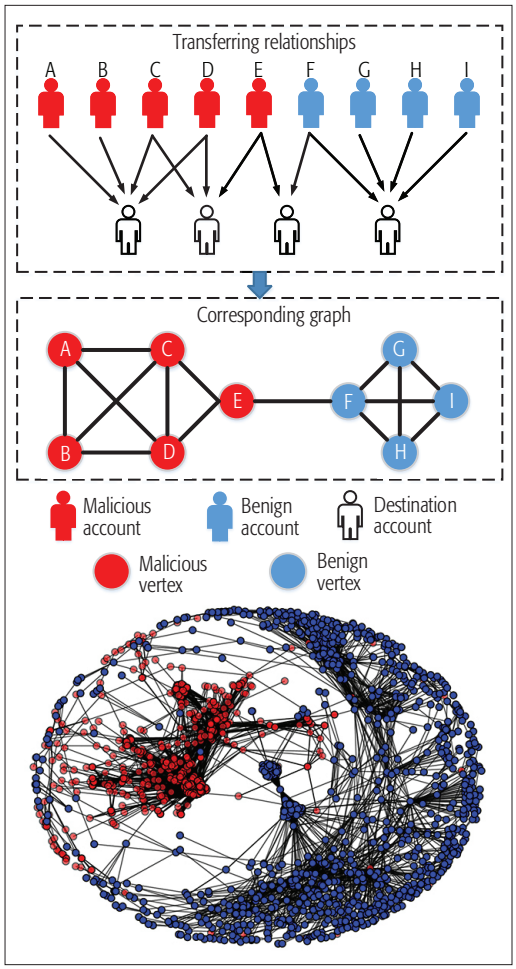

FIGURE 4. Analysis of transferring related graph: a) Sketch map of transferring relationship; b) Illustration of part of the graph structure.

Detection and Evaluation

We leverage machine learning techniques to integrate all these features to perform effective detection. Specifically, feature values extracted from labeled malicious and benign users have been employed to train a statistical classifier. After an unknown user is represented by a vector of feature values, the classifier can automatically evaluate the maliciousness of this user. A variety of

statistical classifiers could be employed in our system to perform detection.

In order to evaluate the effectiveness of the proposed detection method, we use a total number of 496,414 accounts, of which 114,891 are malicious and 381,523 are benign. Without the loss of generality, we use Support Vector Machine (SVM), Random Forest (RM), and Logistic Regression (LR) [14] as the statistical classifier, where the SVM classifier was trained with a Gaussian Kernel and the RF classifier was trained with 3000 trees. We use three metrics to quantify the effectiveness of our method: detection rate (same definition as the true positive rate), false positive rate (FPR), and the area under the ROC curve (AUC) [15]. Specifically, AUC is a widely-used measure of the quality of the statistical classifier. It is defined as the probability that a randomly chosen sample of malicious accounts will have a higher estimated probability of belonging to malicious accounts than that of benign accounts. Since AUC is cutoff-independent and the values of AUC range from 0.5 (no predictive ability) to 1.0 (perfect predictive ability), a higher AUC of a classifier indicates better prediction performance, irrespective of the cutoff selection.

We perform 10-fold cross-validation to evaluate the detection performance of each selected statistical classifier based on all features, using metrics including DR, FPR, and AUC. The results are presented in Table 1. Both Support Vector Machine and Random Forest can achieve high detection rates, high AUC values, and very low false positive rates. These results demonstrate that the features we extract can effectively differentiate between malicious accounts and benign accounts.

We evaluate the effectiveness of our method when using features from one aspect or two aspects. Table 1 presents the results when SVM is adopted as the statistical classifier. The experimental results demonstrate that features from each aspect show great promise in effectively detecting malicious accounts; features of two aspects show better performance compared to features from one aspect; the integration of features from all three aspects show the best performance. This implies high robustness of the proposed method. Specifically, if features of one aspect are evaded by attackers, remaining features can still accomplish high detection accuracy.

Classifiers	Features	FPR	Detection rate	AUC
SVM	**All features**	**0.97%**	**94.2%**	**0.966**
RF	All features	0.22%	92.3%	0.960
LR	All features	4.56%	90.2%	0.928
SVM	Vitality features	3.0%	86.9%	0.920
SVM	**Sequential features**	**3.83%**	**93.3%**	**0.947**
SVM	Spatial features	2.4 %	91.6 %	0.946
SVM	Vitality + sequential features	1.47%	92.9%	0.957
SVM	Vitality + spatial features	1.64%	93.7%	0.961
SVM	**Sequential + spatial features**	**1.38%**	**94.0%**	**0.963**

TABLE 1. Performance analysis of the detection method.

On the scalability of the proposed detection method, although some of the vitality features may not be suitable for all the social networks (e.g., Feature 3 — percentage of recharge from mobile payment, because of that not all the social networks support mobile payment), the sequential and spatial features can be extracted in almost all the social networks integrating virtual currency, and are effective enough to detect the malicious accounts according to the performance analysis shown in Table 1. Therefore, other social networks can also adopt and extend the proposed method to detect the money-laundering accounts.

We also analyze the contribution of each single feature using information gain, where a higher value of information gain indicates more significant contribution. The rank of each feature based on information gain is shown in Table 2, where the top 20 features consist of seven spatial features, eight sequential features, and five vitality features. This indicates that all three aspects are useful for detection.

Rank #	Feature #	Feature type	Information gain	Rank #	Feature #	Feature type	Information gain
1	54	Spatial	0.54447	11	10	Sequential	0.25410
2	53	Spatial	0.54418	12	1	Vitality	0.24621
3	52	Spatial	0.53718	13	9	Sequential	0.24010
4	48	Spatial	0.52206	14	7	Vitality	0.22923
5	49	Spatial	0.51513	15	11	Sequential	0.22550
6	50	Spatial	0.43176	16	2	Vitality	0.22220
7	4	Vitality	0.40336	17	51	Spatial	0.21190
8	8	Vitality	0.38941	18	37	Sequential	0.19719
9	22	Sequential	0.35341	19	17	Sequential	0.19460
10	23	Sequential	0.26119	20	24	Sequential	0.19226

TABLE 2. Top 20 features ranked by information gain.

We designed a collection of 54 features to systematically characterize the behaviors of benign accounts and malicious accounts. Experimental results based on labeled data collected from Tencent QQ, a global leading OSN, demonstrated that the proposed method achieved high detection rates and very low false positive rates.

Conclusions

This article presents the analysis and detection method of money-laundering accounts in OSNs. We analyzed and compared the behavior of both malicious and benign accounts from three perspectives: the account viability, the transaction sequences, and spatial correlation among accounts. We designed a collection of 54 features to systematically characterize the behavior of benign accounts and malicious accounts. Experimental results based on labeled data collected from Tencent QQ, a global leading OSN, demonstrated that the proposed method achieved high detection rates and very low false positive rates.

Acknowledgment

The research presented in this article is supported in part by the National Key Research and Development Program of China (no. 2017YFB0801703, 2016YFB0800100), Tencent Technology (Shenzhen) Company Ltd., the National Natural Science Foundation (61572397, 61502383, 61672425, 61772412, 61402357, 61702407, 61375040, 31500340), the State Grid Corporation of China (DZ71-16-030), and the Fund of China National Aeronautical Radio Electronics Research Institute (PM-12210-2016-001).

References

[1] Y. Wang and S. D. Mainwaring, "Human-Currency Interaction: Learning from Virtual Currency use in China," *Proc. SIGCHI Conf. Human Factors in Computing Systems*, ACM, 2008, pp. 25–28.
[2] Y. Zhou *et al.*, "ProGuard: Detecting Malicious Accounts in Social-Network-Based Online Promotions," *IEEE Access*, vol. 5, 2017, pp. 1990–99.
[3] F. Wu *et al.*, "Social Spammer and Spam Message Co-Detection in Microblogging with Social Context Regularization," *Proc. 24th ACM Int'l. Conf. Information and Knowledge Management*, ACM, 2015, pp. 1601–10.
[4] L. Wu *et al.*, "Adaptive Spammer Detection with Sparse Group Modeling," *Proc. 11th Int'l. AAAI Conf. Web and Social Media*, AAAI, 2017, pp. 319–26.
[5] S. Fakhraei *et al.*, "Collective Spammer Detection in Evolving Multi-Relational Social Networks," *Proc. 21st ACM SIGKDD Int'l. Conf. Knowledge Discovery and Data Mining*, ACM, 2015, pp. 1769–78.
[6] F. Hao *et al.*, "Robust Spammer Detection in Microblogs: Leveraging User Carefulness," *ACM Trans. Intelligent Systems and Technology*, vol. 8, no. 6, 2017, pp. 83:1–31.
[7] G. K. Palshikar, "Detecting Frauds and Money Laundering: A Tutorial," *Proc. Int'l. Conf. Big Data Analytics*, Springer, 2014, pp. 145–60.
[8] R. Dreewski, J. Sepielak and W. Filipkowski, "The Application of Social Network Analysis Algorithms in a System Supporting Money Laundering Detection," *Information Sciences*, vol. 295, 2015, pp. 18–32.
[9] E. L. Paula *et al.*, "Deep Learning Anomaly Detection as Support Fraud Investigation in Brazilian Exports and Anti-Money Laundering," *2016 15th IEEE Int'l. Conf. Machine Learning and Applications (ICMLA)*, Anaheim, CA, 2016, pp. 954–60.
[10] A. F. Colladon and E. Remondi, "Using Social Network Analysis to Prevent Money Laundering," *Expert Systems with Applications*, vol. 67, 2017, pp. 49–58.
[11] J. Pei *et al.*, "Mining Sequential Patterns by Pattern-Growth: The PrefixSpan Approach," *IEEE Trans. Knowledge and Data Engineering*, vol. 16, no. 11, 2004, pp. 1424–40.
[12] M. E. J. Newman, "Communities, Modules and Large-Scale Structure in Networks," *Nature Physics*, vol. 8, no. 1, 2012, pp. 25–31.
[13] R. Li *et al.*, "Finding Influential Communities in Massive Networks," *The VLDB Journal*, 2017.
[14] S. Rogers, and M. Girolami, *A First Course in Machine Learning*, CRC Press, 2016.
[15] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Elsevier, 2011.

Biographies

Yadong Zhou (ydzhou@xjtu.edu.cn) is an assistant professor in the Department of Automation at Xi'an Jiaotong University. He received his B.S. and Ph.D. degrees in control science and engineering from Xi'an Jiaotong University, China, in 2004 and 2011, respectively. He was a postdoctoral researcher at The Chinese University of Hong Kong in 2014. His research focuses on data analysis and mining, network science and its applications, and network security.

Ximi Wang (wangximi@stu.xjtu.edu.cn) is a master student in the Department of Automation at Xi'an Jiaotong University. She received her B.S. degrees in computer science from Xi'an Jiaotong University, China, in 2015. Her research focuses on data analysis and data mining.

Junjie Zhang (junjie.zhang@wright.edu) is an assistant professor in the Department of Computer Science and Engineering at Wright State University. He received his Ph.D. in computer science from Georgia Institute of Technology in 2012. He also received his M.S. in systems engineering and B.S. in computer science from Xi'an Jiaotong University, China, in 2006 and 2003, respectively. His current research focuses on network security and cyber-physical system security.

Peng Zhang (p-zhang@xjtu.edu.cn) is an associate professor in the Department of Computer Science and Technology, Xi'an Jiaotong University, China. He received his Ph.D. degree in computer science from Tsinghua University, China, in 2013. His research interests include network security, privacy, and software-defined networks.

Lili Liu (llliu@sei.xjtu.edu.cn) is a master student in the Department of Automation at Xi'an Jiaotong University. She received her B.S. degrees in control science and engineering from Xi'an Jiaotong University, China, in 2014. Her research focuses on data analysis and data mining.

Huan Jin (rayjin@tencent.com) is a senior engineer at Tencent, Inc. He received his B.S. and M.S. degrees in computer science and technology from Xi'an University of Technology, China, in 2006 and 2009, respectively. His current research focuses on online payments and risk control.

Hongbo Jin (kingboyjin@tencent.com) is a senior engineer at Tencent, Inc. He received his M.S. degrees in control theory and engineering from Chongqing University, China, in 2006. He also received his B.S. in control theory and engineering from Chongqing University of Posts and Telecommunications, China, in 2003. His current research focuses on online payments and risk control.