Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Detecting compromised accounts caused by phone number recycling on e-commerce platforms: taking Meituan as an example*#

Min GAO†1,2, Shutong CHEN1,2, Yangbo GAO3, Zhenhua ZHANG3, Yu CHEN3, Yupeng LI4, Qiongzan YE1,2, Xin WANG1,2, Yang CHEN†‡1,2

[1]*School of Computer Science, Fudan University, Shanghai 200438, China*

[2]*Shanghai Key Lab of Intelligent Information Processing, Fudan University, Shanghai 200438, China*

[3]*Meituan, Beijing 100005, China*

[4]*Department of Interactive Media, Hong Kong Baptist University, Hong Kong 999077, China*

†E-mail: mgao21@m.fudan.edu.cn; chenyang@fudan.edu.cn

Received Apr. 26, 2023; Revision accepted Oct. 18, 2023; Crosschecked July 24, 2024

Abstract: Phone number recycling (PNR) refers to the event wherein a mobile operator collects a disconnected number and reassigns it to a new owner. It has posed a threat to the reliability of the existing authentication solution for e-commerce platforms. Specifically, a new owner of a reassigned number can access the application account with which the number is associated, and may perform fraudulent activities. Existing solutions that employ a reassigned number database from mobile operators are costly for e-commerce platforms with large-scale users. Thus, alternative solutions that depend on only the information of the applications are imperative. In this work, we study the problem of detecting accounts that have been compromised owing to the reassignment of phone numbers. Our analysis on Meituan's real-world dataset shows that compromised accounts have unique statistical features and temporal patterns. Based on the observations, we propose a novel model called temporal pattern and statistical feature fusion model (TSF) to tackle the problem, which integrates a temporal pattern encoder and a statistical feature encoder to capture behavioral evolutionary interaction and significant operation features. Extensive experiments on the Meituan and IEEE-CIS datasets show that TSF significantly outperforms the baselines, demonstrating its effectiveness in detecting compromised accounts due to reassigned numbers.

Key words: Phone number recycling; Neural networks; E-commerce; Compromised account detection
https://doi.org/10.1631/FITEE.2300291 **CLC number:** TP391

1 Introduction

Phone number recycling (PNR) (Lee and Narayanan, 2021; McDonald et al., 2021) refers to an event wherein a disconnected number is reassigned to another consumer by mobile operators after an aging period. According to the definition provided by the Federal Communications Commission of the

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 62072115, 62202402, 61971145, and 61602122), the Shanghai Science and Technology Innovation Action Plan Project (No. 22510713600), the Guangdong Basic and Applied Basic Research Foundation, China (Nos. 2022A1515011583 and 2023A1515011562), the One-off Tier 2 Start-up Grant (2020/2021) of Hong Kong Baptist University (Ref. RC-OFSGT2/20-21/COMM/002), Startup Grant (Tier 1) for New Academics AY2020/21 of Hong Kong Baptist University and Germany/Hong Kong Joint Research Scheme sponsored by the Research Grants Council of Hong Kong, China, the German Academic Exchange Service of Germany (No. G-HKBU203/22), and Meituan

Electronic supplementary materials: The online version of this article (https://doi.org/10.1631/FITEE.2300291) contains supplementary materials, which are available to authorized users

ORCID: Min GAO, https://orcid.org/0009-0002-9374-1459; Yang CHEN, https://orcid.org/0000-0003-4749-3060

United States (2018), a number is permanently disconnected when "a subscriber permanently has relinquished the number or the provider permanently has reversed its assignment of the number to the subscriber such that the number has been disassociated with the subscriber." In many countries, especially those with large populations, such as China, the U.S., and India (Federal Communications Commission of the United States, 2018), PNR has been proved to be a practical strategy to ease the imbalance between the growing number of mobile subscribers (https://www.gsma.com/r/somic/) and the limited phone numbers. However, this strategy has posed threats to the reliability of the short message service based two-factor authentication (SMS-based 2FA) method, which has been widely adopted by online platforms such as PayPal, Facebook, and Twitter (Mulliner et al., 2013; Dmitrienko et al., 2014). Here, the two factors are a phone number associated with the user account and an SMS one-time password (SMS OTP) sent to the number. This authentication method is used to authorize high-risk operations, e.g., payment and password resetting, and is also used for fast registration and login in many online applications. Fig. 1 demonstrates the process of PNR. Obviously, the compromised account can be controlled by the new owner of the phone number associated with the account. PNR is a serious cause for concern because it has the potential to cause a series of severe privacy breaches, fraud activities (e.g., account takeover (Tao et al., 2018; Doerfler et al., 2019; Kawase et al., 2019; Mirian et al., 2019; Thomas et al., 2021), credit card fraud (Bhattacharyya et al., 2011; Branco et al., 2020; Cheng et al., 2020), and identity theft (Bilge et al., 2009; Tao et al., 2018; Zou et al., 2020; Wang C and Zhu, 2022; Ye et al., 2022)), and property damage to financial institutions and customers (Mobile China, 2017; Li S et al., 2018; Mirian et al., 2019; Thomas et al., 2021). Note that "account takeover" is a form of Internet fraud where a fraudster has unauthorized access to or takes full control of a legitimate account.

Reports from the Federal Communications Commission of the United States (2018) suggest that there are tens of millions of reassigned phone numbers per year and number recycling can lead to a unique form of account takeover. In China, about 50% of the numbers in use are reassigned phone numbers. Lee and Narayanan (2021) claimed that there are about one million recycled phone numbers at Verizon (https://www.verizon.com/) every month. McDonald et al. (2021) revealed the consequences and problems caused by PNR from the perspective of the users, and further explored the challenges faced by online social platforms and mobile operators. According to previous studies (Lee and Narayanan, 2021; McDonald et al., 2021), PNR has induced a series of security and privacy risks for the Internet ecosystem. Communication operators, in particular, are having difficulty in identifying "thoroughly clean" old numbers with the increasing proportion of phone number re-allocations. This difficulty arises from the imbalance between the limited phone numbers and the increasing number of users. Therefore, a unique number for each user can hardly be guaranteed. On the side of users, PNR can lead to user harassment, account access problems, and privacy leakage. McDonald et al. (2021) summarized five types of negative experiences of PNR and analyzed the corresponding consequences and behavior changes.

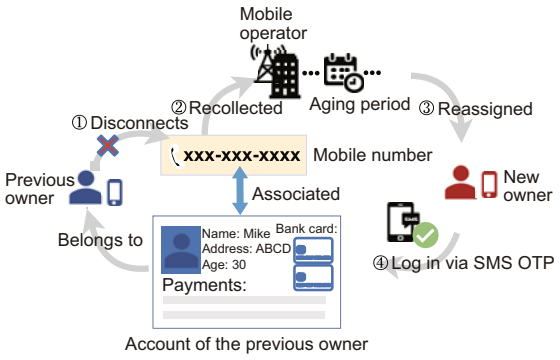

Fig. 1 Workflow of PNR (PNR: phone number recycling; SMS: short message service; OTP: one-time password)

Meituan (Fu et al., 2021; Huang et al., 2021; Ping et al., 2021), which is known as one of the largest online-to-offline e-commerce platforms in China accommodating various third-party services including food, travel, hotel, entertainment, and household services, also faces the issues and risks caused by PNR (https://about.meituan.com/en). The number of customers who have complained of such an issue in a month can reach over 7000. Currently, to guarantee the protection of user accounts and lighten disturbance to users, Meituan has launched some strategies (e.g., risk assessment based on limited

domain expertise and qualitative rules within Meituan) against the PNR issues. Such strategies are difficult to cover complex and time-varying fraud scenarios and patterns, let alone distinguish whether it is an illusion caused by the same user's job relocation or a long silence period. Furthermore, Meituan has about 680 million yearly active users, and thus such strategies require considerable expert audit costs. For the e-commerce platforms, e.g., Meituan, fraud events caused by PNR lead to monetary loss as well as damage to the brand reputation caused by poor user experience. Therefore, to alleviate this pressure, an effective model is needed that would be capable of detecting compromised accounts due to reassigned numbers.

Existing solutions to this problem are based on inquiry or notification services based on databases managed by or aggregated from mobile operators (Mobile China, 2017; Federal Communications Commission of the United States, 2018; Alibaba Cloud, 2023). However, the large customer base and transaction volume, as well as the constantly changing numbers on large-scale e-commerce platforms like Meituan, make it costly to check the status of the associated number by querying mobile operators' databases whenever an authentication occurs. These solutions also introduce the security concern of leaking users' phone number information to external parties. Besides the recently compromised accounts, there are many existing accounts that were taken over long ago and may not be covered by the mobile operators' databases. Thus, an alternative solution that relies on the information of the applications (e.g., the profiles and behavioral data of the accounts) rather than the feedback from external sources is imperative.

To address the abovementioned challenges, we propose the temporal pattern and statistical feature fusion model (TSF), to detect accounts in e-commerce platforms that were compromised owing to PNR. TSF features a novel fusion way using both temporal patterns and statistical features. Specifically, the temporal pattern encoder integrates time-aware long short-term memory (T-LSTM) units (Baytas et al., 2017) and the attention module (Vaswani et al., 2017; Chai et al., 2022) to capture behavioral evolutionary interaction from temporal behavioral history. The statistical feature encoder aims to capture significant statistical features based on user operations by leaf embedding (Friedman, 2001; Ke et al., 2019). Finally, a fusion module uses a self-attention layer to effectively incorporate information from the temporal pattern encoder and the statistical feature encoder to detect whether an account is compromised due to PNR.

We evaluate the performance of TSF using a desensitized PNR dataset provided by Meituan, which totally consists of 123 963 accounts. Our evaluation shows that TSF delivers area under curve (AUC) improvements of 4.95%, 4.48%, 4.66%, and 3.89% over the tree-based model, the fusion of tree-based model, the recurrent neural network (RNN), and fraud detection baselines, respectively. Furthermore, we test the effectiveness of TSF on the IEEE-CIS dataset (Mainali et al., 2022) with an 89.27% AUC (see Section 3 in the supplementary materials).

In summary, our main contributions are as follows:

1. We are the first to formulate the problem of PNR on e-commerce platforms, and design the first end-to-end framework to identify accounts subject to PNR. To characterize the behavioral patterns of a compromised account, we analyze the real-world cases of accounts on Meituan that have been compromised owing to PNR.

2. We propose an approach named TSF to detect compromised accounts due to PNR, where the temporal pattern encoder is designed to capture behavioral evolutionary interaction and the statistical feature encoder is employed to capture significant operation features.

3. Extensive experiments on two real-world datasets show that the proposed TSF outperforms several competitive baselines, demonstrating its effectiveness in detecting compromised accounts by capturing temporal patterns and informative statistical features.

2 Related works

In this section, we first present the existing solutions for detecting compromised accounts via PNR, as explained in Section 2.1. Then we study the security issues with PNR in Section 2.2, followed by a discussion on solutions for detecting compromised accounts in online applications in Section 2.3.

2.1 Existing solutions

To our knowledge, none of the existing fraud detection efforts are specifically tailored for situations such as PNR, except for inquiry or notification services from external mobile operators (Mobile China, 2017; Federal Communications Commission of the United States, 2018; Alibaba Cloud, 2023). For example, Federal Communications Commission's reassigned numbers database (Federal Communications Commission of the United States, 2018), which records and updates all reassigned numbers from operators, allows database users (e.g., an online application) to make an inquiry concerning a phone number, and the input needed for the enquiry would be simply the number and the date of its latest indicated status. These services are usually provided as application programming interfaces (APIs) and charged on a per-call basis.

2.2 Security issues with PNR

Notably, SMS-based 2FA assumes the reliable knowledge of the owner of the phone number associated with the user account. However, in practice, with such an authentication method, the new owner can fully control the recycled phone number, which has posed threats to the reliability of this method (https://recyclednumbers.cs.princeton.edu/). McDonald et al. (2021) have conducted a study of negative experiences and issues caused by PNR with 195 samples. The main findings include problems with spam and account access, harassment, and privacy concerns. Lee and Narayanan (2021) have expressed their concerns about the fact that there are 35 million reassigned phone numbers in the U.S. They have conducted a survey of the security and privacy risks of PNR, and concluded that the security vulnerabilities caused by PNR could give rise to the possibilities for eight different types of attack.

2.3 Detecting compromised accounts

Compromised account detection systems designed for social networks usually rely on the posted contents of users to identify their behavioral changes. Most of them (Viswanath et al., 2014; Egele et al., 2017; Wang C et al., 2020; Wang C and Zhu, 2022) leverage statistical, machine learning, or language models to construct a behavioral profile for each user from his/her historical posts, and classify accounts by how much their new posts are different from the constructed profiles. Recently, deep neural networks are used to detect compromised accounts based on social information and posted contents in an end-to-end way (Karimi et al., 2018; VanDam et al., 2018). There are also graph-based methods (Boshmaf et al., 2015; Cao et al., 2019; Liang et al., 2021; Xu et al., 2021) that leverage users' social graphs to detect fake accounts rather than compromised accounts, based on some observations or assumptions about fraudsters' social distributions like "fraudsters tend to have few connections to legitimate accounts" (Boshmaf et al., 2015; Cao et al., 2019) and deep features offering additional information (Xu et al., 2021). Liang et al. (2021) presented an unsupervised method to detect fake accounts by recognizing the graph structure within the registration stage, which reduces the cost of manually labeling fake accounts.

For e-commerce platforms, these solutions are not suitable as features of social relationships or user-generated content are usually not available in practice. In the work of mobile.de (Kawase et al., 2019), rules capturing multiple abnormal operations were developed and learned by a multi-variate Bernoulli Naive Bayes model. Tao et al. (2018) presented GAS, a graph attention mechanism within a behavioral sequence, to relate user behaviors with the same device, IP address, etc., regardless of their long-term behavioral sequences. A GAS-augmented LSTM was used to detect compromised accounts by learning from behavioral sequences of multiple operations. Graph-based methods model the interactions in a heterogeneous account–device graph (Liu et al., 2018; Wang DX et al., 2019) or a social graph (Liang et al., 2021) to detect malicious accounts or predict credit risk, based on the device aggregation or social aggregation of fraudsters.

Moreover, they do not consider the different fraud patterns in the scenario of PNR, where new owners usually log into the previous owners' account by accident. The new owners are not tailored to detect compromised accounts and behave normally and regularly for an extended period. Only when they find profitable vulnerability, do they engage in fraudulent activities.

However, the following are the reasons for these methods to be unable to tackle our problem: (1) They do not consider the different fraud patterns in the scenario of PNR, where the new owner of a phone

number gains access to the previous owner's account and behaves in a normal and regular manner for a long time. This inconspicuous behavior continues until the new owner identifies a profitable vulnerability and engages in fraudulent activities. Importantly, no authorization is required for these malicious actions while using the phone number legitimately. (2) Methods that solely leverage rules or statistical features are insufficient for our problem (this will be discussed in Section 3.3). (3) Existing sequential methods (Karimi et al., 2018; Yu et al., 2018) that learn from behavioral history incorporate neither the irregular time intervals between a user's operations nor the interaction between different types of behaviors, which are both important for modeling the temporal behavioral patterns (this will be discussed in Section 4). (4) There are no explicit social connections among accounts on Meituan. As a result, the initial assumption of aggregating devices and IP addresses does not hold for compromised accounts by PNR. This is simply because phone numbers are reassigned to new owners independently.

3 Observations from the Meituan dataset and problem definition

This section is organized as follows: Section 3.1 provides a basic description of the used dataset; considering the low likelihood of aggregation among compromised accounts, we focus on the behavioral patterns of the accounts and search for distinctive features that would distinguish compromised accounts from normal ones. Section 3.2 presents an investigation into the temporal behavioral patterns of these two types of accounts. Further, Section 3.3 explores the statistical features that are useful in distinguishing between compromised accounts and normal ones, which gives rise to an intuition for model design. Finally, Section 3.4 presents the problem formulation.

3.1 Dataset

On the Meituan app, users can bind bank cards to their accounts to purchase local life services such as food delivery and hotel booking. Note that the Meituan app typically offers users a variety of payment options, including bank card payment, Meituan Monthly Pay, Alipay, and WeChat Pay. Bank card payment is the default payment method. Moreover, other payment methods are still essentially based on the manner of bank card payment, so here we uniformly use the bank card payment method. Meituan also provides financial products including consumer and business loans for verified accounts with real-name identities. The compromised accounts triggered by PNR have invoked a large number of customer complaints.

We obtain a labeled account dataset containing account behavior records. Here, a record is an entry in the database that logs the information of an operation, e.g., login. It contains a unique account ID, timestamp, type of operation, etc. In this dataset, all compromised accounts are caused by PNR. In practice, compromised accounts can be caused by various attacks, e.g., phishing and credential stuffing. The compromised accounts in our dataset are those caused by PNR only. Each of these accounts has been manually labeled by customer services with the assistance of users and mobile operators. Accounts were compromised between July 2016 and Feb. 2021. For privacy concerns, we have anonymized all user IDs and removed sensitive information, e.g., delivery addresses. The data were stored on a well-protected offline server. We removed accounts with no payment record, and obtained a dataset containing 32 388 compromised accounts and 91 575 sampled normal accounts. To mitigate the class imbalance problem, we have applied down-sampling techniques (Zhang et al., 2011; Wang J et al., 2022) to the normal account class. Here, the compromised accounts refer to the accounts that have been subject to PNR, while the normal accounts denote the accounts that have not experienced PNR. The statistics of the dataset are presented in Table 1.

3.2 Temporal patterns

Each account reveals the owner's unique habits and access environments over a period of time, which appear as temporal patterns in their behavioral sequences. The related habits involve multiple aspects such as the purchased items, the price ranges, payment methods, time period of use, and frequency of use. The access environment refers to a collection of configurations including devices, IP addresses, and global positioning system (GPS) locations when using the Meituan app.

The habits and access environments of an account would change when the account owner

Table 1 Statistics of the Meituan dataset

Type	Number of accounts	Average payment per account	Average login per account	Number of payment records	Number of login records
Compromised	32 388	166	55	5 368 566	1 815 494
Normal	91 575	315	48	28 863 693	4 687 487

changes. Another possible factor in such a change could be alterations in the user's habits and access environments over time, stemming from physical relocation, device switching, and preference shifting; however, this sort of change often happens progressively rather than instantly. To further illustrate this issue, we compare a compromised account with a normal account by visualizing the temporal evolution among the devices that were used to log into the accounts along with other information such as delivery addresses, device names, and cities over time.

Fig. 2a displays the evolutionary processes of a compromised account's devices and associated information. The payment process can provide some sets of entities associated with the account for an indefinite period of time. The changes in these entities over time are clearly visible along the timeline. Note that the phone number has been bound to this account since 2016, and was recycled on Mar. 8, 2019. So, there are significant temporal activities, one corresponding to the period from Dec. 2016 to Jan. 2017, and the other corresponding to the period from Mar. 2019 to Oct. 2019. There is a silent period of more than two years (i.e., from Jan. 2017 to Mar. 2019) during which the account had no activities. The devices and addresses used before the silent period never appear afterward, while the two devices used in 2019 are associated with two new addresses or device names, which implies that they are owned by a new user.

Fig. 2b depicts a normal account's evolution of behavioral sequences. Unlike that of a compromised account, there are continuous behavioral sequences for this normal account, and entities with devices and other information are evolving over time. Therefore, there are no silent periods displayed here. From this observation, we can deduce that a lengthy period of silence is rare for a normal account. It is also important to note that the sequential activities of the normal accounts are both continuous and relevant over time.

In e-commerce, behavioral sequences can often present the evolution process of account behaviors. These results show that temporal patterns provide insights into distinguishing compromised accounts from normal ones. These features fail to uncover overall significant differences between compromised accounts and normal ones. Therefore, we have explored the statistical features of these accounts.

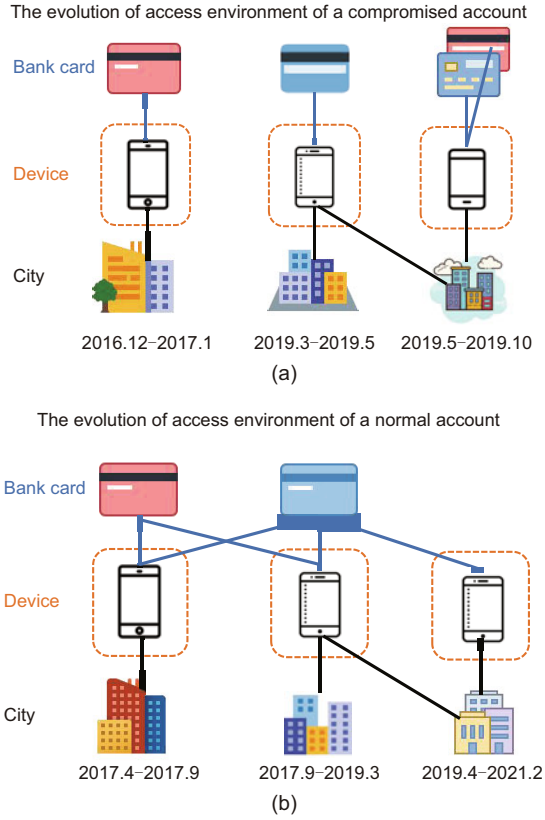

Fig. 2 Temporal patterns involving devices and other entities for the compromised accounts (a) and normal accounts (b)

3.3 Statistical features

Here, we aim to ascertain what makes compromised accounts different from normal ones. Based on previous knowledge of fraud detection (Hu et al., 2019) and expert knowledge of Meituan's business logic, we divide these behavior records into four types: (1) payment, (2) login, (3) bank card binding,

and (4) payment password (a six-digit password set by a user to authorize his/her payment) manipulation. In e-commerce, the payment process is a significant user behavior and also a key opportunity for fraudsters to profit. Login is a prerequisite for other behaviors. On Meituan, a user is allowed to add a bank card to his/her account only if the cardholder's identity is consistent with the account's real-name identity. Therefore, a compromised account may have abnormal bank card binding attempts, if the previous owner had already bound his/her bank card or real-name identity. Additionally, the new phone number owner is unaware of the identity or payment password of the previous owner, leading to abnormal password reset attempts.

Together with technical experts from Meituan, we finally retain 33 informative statistical features. The differences in four representative features have been illustrated in Figs. 3a–3d, which explain the distributions of behaviors between compromised and normal accounts.

Binding behaviors. According to accounts' behavioral data, we study the average time interval of bank card bindings and the number of failed bank card binding attempts between these two types of accounts. Not surprisingly, the results show that compromised accounts may bind bank cards more frequently (Fig. 3a). The reason for this is that the new owner may need to try multiple payment passwords to bind his/her cards properly. Similarly, there are more failed bank card binding attempts for compromised accounts, as shown in Fig. 3b. Compromised accounts, for the purpose of whether easy payment (unintentionally) or fraudulent transactions (intentionally), often attempt to bind their bank cards multiple times and persist after each failure, indicating a high frequency of card binding attempts.

Login behaviors. Based on accounts' behavioral data, we study the distributions of the number of logins via SMS OTP on a new device and the number of times of failure for each of the accounts. Figs. 3c and 3d show that compromised accounts are characterized by a greater number of logins via SMS OTP on a new device and higher possibilities to fail at login compared with normal accounts. After the phone number is reassigned, the new owner must log in by SMS on a device different from the previous owner's. This may also involve logging in from a new location or with a new IP address. The anomaly in the login environment may trigger the account protection

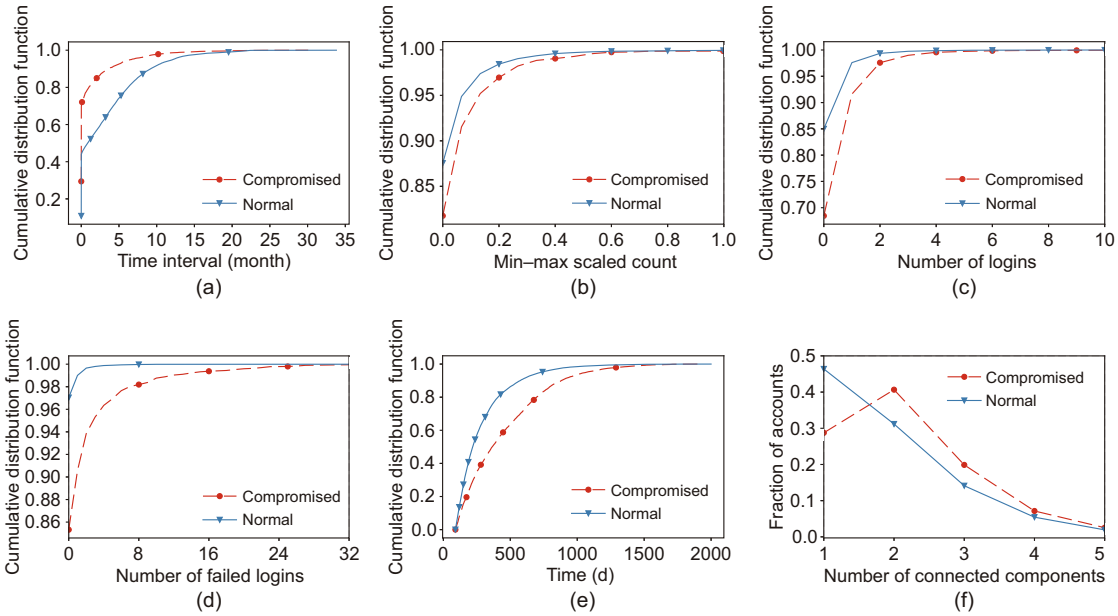

Fig. 3 Differences of statistic features and temporal patterns between compromised and normal accounts on Meituan: (a) average time interval of bank card bindings for each account; (b) number of failed bank card binding attempts ("min–max scaled count" means the ratio of the number of failed bindings to the maximum of 99% accounts); (c) number of logins on a new device by SMS OTP; (d) number of failed logins; (e) maximum length of silent periods; (f) number of connected components in an account's behavioral sequences (SMS: short message service; OTP: one-time password)

system and lead to unsuccessful login attempts.

Generally, our findings suggest that normal accounts may experience silent periods, which are usually shorter than those of compromised accounts (Fig. 3e). When a number is reassigned, the evolution of access environments before and after could be completely different. However, these different temporal patterns differ by personally identifiable attributes linked to different users. Inspired by these observations, we also define accounts' behavioral sequences by the notion of the number of connected components (Bonaccorsi et al., 2020), which characterizes continuity and correlation among accounts' behavior sequences. As shown in Fig. 3f, the ratio of normal accounts that have connected sequences (47%) is much higher than the corresponding ratio (28%) for the compromised accounts.

We examine behavioral differences between compromised and normal accounts with Welch's t-test (Welch, 1951) to ascertain whether our above observations are reliable. We also calculate the corresponding p-values to investigate the difference between compromised accounts and normal ones. By performing Welch's t-test, p is derived for each characteristic with a value less than 0.001. This verifies the significance of the selected characteristics in distinguishing compromised accounts from normal accounts.

3.4 Problem definition

In light of the above data exploration, we extract statistic features and temporal patterns of accounts from our dataset. Table 2 contains the main notations used in this paper. Formally, we integrate the behavior and evolutionary information

Table 2 Notations and descriptions

Notation	Description
$\boldsymbol{S}$	The behavior and evolutionary information of each account
$\boldsymbol{X}^{\mathrm{p}}$	Payment behaviors for all accounts
$\boldsymbol{X}^{\mathrm{l}}$	Login behaviors for all accounts
$\boldsymbol{f}$	Statistical features for all accounts
$\boldsymbol{X}_i^{\mathrm{p}}$	Payment behaviors for account i
$\boldsymbol{X}_i^{\mathrm{l}}$	Login behaviors for account i
$\boldsymbol{x}_{ik}^{\mathrm{p}}$	The k^{th} payment sequence of account i
$\boldsymbol{t}_{ik}^{\mathrm{p}}$	The corresponding timestamps of the k^{th} payment sequence of account i
L_{p}	The length of the corresponding timestamps of the payment sequences
$\boldsymbol{x}_{ik}^{\mathrm{l}}$	The k^{th} login sequence of account i
$\boldsymbol{t}_{ik}^{\mathrm{l}}$	The corresponding timestamps of the k^{th} login sequence of account i
L_{l}	The length of the corresponding timestamps of the login sequences
$\boldsymbol{f}_i$	Statistical features for account i
$\hat{\boldsymbol{Y}}$	The predicted labels for all accounts
$\boldsymbol{Y}$	The true labels for all accounts
$g(\cdot)$	A non-increasing function of time interval
$\boldsymbol{\Delta}_t$	Time intervals of behaviors including payment and login
δ	A threshold for measuring the correlation of two consecutive series
$\boldsymbol{H}^{\mathrm{p}}$	The encode vector of the payment behaviors
$\boldsymbol{H}^{\mathrm{l}}$	The encode vector of the login behaviors
$\mathbf{PE}_i^{\mathrm{p}}$	The matrix with each positional encoding of the payment input position of account i
$\mathbf{PE}_i^{\mathrm{l}}$	The matrix with each positional encoding of the login input position of account i
d_{p}	The dimension of payment initial representations
d_{l}	The dimension of login initial representations
$\boldsymbol{Q}_i$	The query parameter matrix of the input payment and login vectors for account i
$\boldsymbol{K}_i$	The key parameter matrix of the input payment and login vectors for account i
$\boldsymbol{V}_i$	The value parameter matrix of the input payment and login vectors for account i
d_{k}	The dimension of $\boldsymbol{Q}$ and $\boldsymbol{K}$
d_{v}	The dimension of $\boldsymbol{V}$
$\boldsymbol{H}^{\mathrm{inter}}$	The inter-relationship encoding vector
$\boldsymbol{H}_i^{\mathrm{feature}}$	The statistical feature encoding vector for account i
$\boldsymbol{W}_{\mathrm{a}}$	The corresponding weight of a single-layer neural network
$\boldsymbol{b}_{\mathrm{a}}$	The corresponding bias of a single-layer neural network
$\boldsymbol{H}_i^{\mathrm{all}}$	The sum of the rows of the output matrix of the self-attention module

of each account $\boldsymbol{S} = \{\boldsymbol{X}^{\mathrm{p}}, \boldsymbol{X}^{\mathrm{l}}, \boldsymbol{f}\}$, where $\boldsymbol{X}^{\mathrm{p}} = \{\boldsymbol{X}_i^{\mathrm{p}} | i = 1, 2, \ldots, n\}$, $\boldsymbol{X}^{\mathrm{l}} = \{\boldsymbol{X}_i^{\mathrm{l}} | i = 1, 2, \ldots, n\}$, and $\boldsymbol{f} = \{\boldsymbol{f}_i | i = 1, 2, \ldots, n\}$ denote payment behaviors, login behaviors, and statistic features of each account, respectively. Here, "p" and "l" refer to the payment and login events, respectively, and n is the number of accounts. Labels are assigned to each account to indicate if it is a compromised account ($y_i = 1$) or a normal one ($y_i = 0$).

Let $\boldsymbol{S}_i = \left(\boldsymbol{X}_i^{\mathrm{p}}, \boldsymbol{X}_i^{\mathrm{l}}, \boldsymbol{f}_i\right)$ be a sample, i.e., the features of account i ($i = 1, 2, \ldots, n$). Specifically,

$$\boldsymbol{X}_i^{\mathrm{p}} = \left\{ \left(\boldsymbol{x}_{i1}^{\mathrm{p}}, \boldsymbol{t}_{i1}^{\mathrm{p}}\right), \left(\boldsymbol{x}_{i2}^{\mathrm{p}}, \boldsymbol{t}_{i2}^{\mathrm{p}}\right), \ldots, \left(\boldsymbol{x}_{iL_{\mathrm{p}}}^{\mathrm{p}}, \boldsymbol{t}_{iL_{\mathrm{p}}}^{\mathrm{p}}\right) \right\}, \quad (1)$$

where $\boldsymbol{x}_{ik}^{\mathrm{p}}$, $\boldsymbol{t}_{ik}^{\mathrm{p}}$, and L_{p} stand for the payment sequence, the corresponding timestamps of the account, and the length of the corresponding timestamps, respectively. Similarly,

$$\boldsymbol{X}_i^{\mathrm{l}} = \left\{ (\boldsymbol{x}_{i1}^{\mathrm{l}}, \boldsymbol{t}_{i1}^{\mathrm{l}}), (\boldsymbol{x}_{i2}^{\mathrm{l}}, \boldsymbol{t}_{i2}^{\mathrm{l}}), \ldots, (\boldsymbol{x}_{iL_{\mathrm{l}}}^{\mathrm{l}}, \boldsymbol{t}_{iL_{\mathrm{l}}}^{\mathrm{l}}) \right\}, \quad (2)$$

where $\boldsymbol{x}_{ik}^{\mathrm{l}}$, $\boldsymbol{t}_{ik}^{\mathrm{l}}$, and L_{l} indicate the login sequence, the corresponding timestamps of the account, and the length of the corresponding timestamps, respectively.

Both sequences contain records with information about the access environment and the time of day. Additionally, other information pertaining to payment, such as orders and bank cards, is included together in records $\boldsymbol{X}_i^{\mathrm{p}}$. $\boldsymbol{f}_i$ is a vector whose elements are all real numbers, each representing the value of a statistical feature extracted from account i's multiple operations mentioned in Section 3.1. The problem is formulated as follows:

Definition 1 (Detecting compromised accounts via PNR) Given account behavior information and account temporal records $\boldsymbol{S} = \{\boldsymbol{X}^{\mathrm{p}}, \boldsymbol{X}^{\mathrm{l}}, \boldsymbol{f}\}$, and the label vector $\boldsymbol{Y}$, where $\boldsymbol{X}^{\mathrm{p}}$, $\boldsymbol{X}^{\mathrm{l}}$, $\boldsymbol{f}$, and $\hat{\boldsymbol{Y}}$ denote payment behaviors, login behaviors, statistical features, and predicted labels for all accounts, respectively, we intend to infer label y_i for a given sample $\boldsymbol{S}_i \in \boldsymbol{S}$, i.e., to identify the compromised accounts from normal ones.

4 Design of TSF

In this section, based on insights gained from Section 3, we propose TSF, a model that incorporates temporal patterns and statistical features to detect compromised accounts. Observations in Section 3 lead to the following conclusions:

1. Temporal patterns assist with the detection of compromised accounts via PNR.

2. Informative statistical features facilitate the detection of compromised accounts triggered by PNR.

3. Statistical features and temporal patterns of the same account are interdependent.

Enlightened by the observed results, we design a model called TSF to detect compromised accounts via PNR in Section 4.1. Detailed descriptions of key building blocks of TSF are expounded in Sections 4.2–4.4.

4.1 Model architecture

The framework of the proposed model, TSF, is shown in Fig. 4. A sequential encoder based on an attention-augmented RNN is used to process payment and login sequences (as temporal pattern encoder on the left-hand part in Fig. 4 with the attention module), and a pre-trained gradient boosting decision tree (GBDT) model embeds statistical features of each account (as statistical feature encoder on the right-hand part in Fig. 4). All encoding vectors are fused with the self-attention module (as fusion module on the upper part in Fig. 4), and a classification layer yields the detection results.

4.2 Temporal pattern encoder

The application of LSTM (Hochreiter and Schmidhuber, 1997) in the processing of sequential data has shown success in a variety of fields. However, traditional LSTM units, designed with the assumption that the time interval between consecutive elements is a constant, fail to capture effective temporal information for cases with irregular time intervals. For our case, it is critical to consider irregular time intervals because: (1) the frequency of operations is one characteristic of an account; (2) abnormal or malicious behavior often occurs continuously or repeatedly within a short time; and (3) long periods of silence are a strong indicator of possible ownership change. Thus, we propose temporal pattern encoder based on T-LSTM (Baytas et al., 2017) to discover the interrelationships between multiple behavioral sequences and the temporal relationships within individual behavioral sequences. Keeping the three gate functions of LSTM unchanged, T-LSTM uses a neural network to project the previous cell

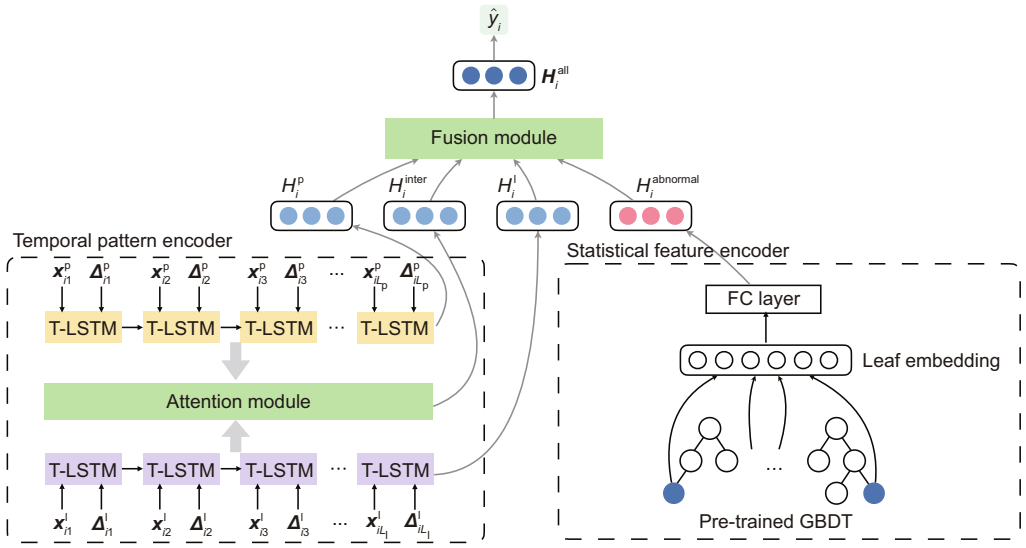

Fig. 4 Framework of TSF (TSF: the temporal pattern and statistical feature fusion model; T-LSTM: time-aware long short-term memory; FC: fully-connected; GBDT: gradient boosting decision tree)

state to a subspace to represent the short-term memory, and discounts it with a non-increasing function $g(\cdot)$ of time interval $\boldsymbol{\Delta}_t$. Additionally, the temporal pattern encoder can be any among the off-the-shelf recurrent neural networks (Hochreiter and Schmidhuber, 1997; Cho et al., 2014; Keren and Schuller, 2016) that are subject to the temporal patterns of the sequential data. This suggests that our model is general and flexible to various applications with different temporal patterns.

In our case, normal accounts may show similar payment temporal patterns on a daily or weekly basis, which is often consistent with the lifestyle and consumption habits of the owner. However, the compromised accounts tend to carry out a series of continuous and related fraud activities in a short period of time (between a few minutes and several hours), which is manifested as a surge of fraud objective-oriented behavior in a short period. Based on the above common situation in e-commerce platforms, we make specific modifications to the function $g(\cdot)$ to adapt to our case. To capture the short-term frequent operations and the corresponding owner's consumption patterns, we expect that our $g(\cdot)$ decays between a few seconds and a few days, but the decrease should not be significant. Meanwhile, since $\boldsymbol{\Delta}_t$ might have large numerical values, the value of g should not decrease too rapidly with the increase of $\boldsymbol{\Delta}_t$. Therefore, we choose to measure $\boldsymbol{\Delta}_t$ as minutes, and design $g(\cdot)$ as a non-increasing piecewise function. Here, for the payment behaviors, we have $\boldsymbol{\Delta}_t = \left[\Delta t^{\mathrm{p}}_{i1}, \Delta t^{\mathrm{p}}_{i2}, \ldots, \Delta t^{\mathrm{p}}_{iL_{\mathrm{p}}}\right]$, where i denotes the account i. Similarly, for the login behaviors, we have $\boldsymbol{\Delta}_t = \left[\Delta t^{\mathrm{l}}_{i1}, \Delta t^{\mathrm{l}}_{i2}, \ldots, \Delta t^{\mathrm{l}}_{iL_{\mathrm{l}}}\right]$. We simplify the length of each timestamp as Δt, and can measure the importance of account behaviors by

$$g(\Delta t) = \begin{cases} 1, & \text{if } 0 \le \Delta t \le \delta, \\ \dfrac{1}{\ln\left(\mathrm{e} + \Delta t - \delta\right)}, & \text{if } \Delta t > \delta, \end{cases} \tag{3}$$

where δ is a threshold below which the two consecutive events are considered to be highly correlated; there is thus no need to decay the short-term memory.

Upon completing the function $g(\cdot)$ design, the encoding vector of the payment behaviors $\boldsymbol{H}^{\mathrm{p}}$ and the encoding vector of the login behaviors $\boldsymbol{H}^{\mathrm{l}}$ can be derived from the sum of all the hidden states of each time step of T-LSTM. Motivated by He et al. (2020), we employ the inter-attention mechanism to learn the temporal relationship between login and payment sequences. The input vectors for the inter-attention mechanism can be calculated by summarizing the positional embedding with the initial representations as follows:

$$\boldsymbol{E}^{\mathrm{p}}_i = \boldsymbol{X}^{\mathrm{p}}_i + \mathbf{PE}^{\mathrm{p}}_i, \tag{4}$$

where $\boldsymbol{X}^{\mathrm{p}}_i \in \mathbb{R}^{L_{\mathrm{p}} \times d_{\mathrm{p}}}$ and $\mathbf{PE}^{\mathrm{p}}_i \in \mathbb{R}^{L_{\mathrm{p}} \times d_{\mathrm{p}}}$ denote the payment behaviors and the matrix with each positional encoding of the payment input position of

account i, respectively. In this context, we define L_{p} and d_{p} as the length of the payment timestamps and the dimension of payment initial representations, respectively. Meanwhile,

$$\boldsymbol{E}_i^{\mathrm{l}} = \boldsymbol{X}_i^{\mathrm{l}} + \mathbf{PE}_i^{\mathrm{l}}, \tag{5}$$

where $\boldsymbol{X}_i^{\mathrm{l}} \in \mathbb{R}^{L_{\mathrm{l}} \times d_{\mathrm{l}}}$ and $\mathbf{PE}_i^{\mathrm{l}} \in \mathbb{R}^{L_{\mathrm{l}} \times d_{\mathrm{l}}}$ indicate the login behaviors and the matrix with corresponding positional encoding for each login input position of account i, respectively. The length of the timestamps of login and the dimension of its initial representations are, respectively, denoted by L_{l} and d_{l}.

Generally, an attention mechanism maps a query and a set of keys to an output, which is a weighted sum of values. For our inter-attention module, the queries, keys, and values are derived from the linear projection of the input payment vectors and login vectors: $\boldsymbol{Q}_i = \boldsymbol{E}_i^{\mathrm{p}} \boldsymbol{W}^{\mathrm{Q}} \in \mathbb{R}^{L_{\mathrm{p}} \times d_{\mathrm{k}}}$, $\boldsymbol{K}_i = \boldsymbol{E}_i^{\mathrm{l}} \boldsymbol{W}^{\mathrm{K}} \in \mathbb{R}^{L_{\mathrm{l}} \times d_{\mathrm{k}}}$, and $\boldsymbol{V}_i = \boldsymbol{E}_i^{\mathrm{l}} \boldsymbol{W}^{\mathrm{V}} \in \mathbb{R}^{L_{\mathrm{l}} \times d_{\mathrm{v}}}$. According to the scaled dot-product attention function, each payment attends to different logins, and the relationship between them, denoted by $\boldsymbol{Q}'$, can be calculated by

$$\boldsymbol{Q}' = \mathrm{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_{\mathrm{k}}}}\right)\boldsymbol{V} \in \mathbb{R}^{L_{\mathrm{p}} \times d_{\mathrm{v}}}, \tag{6}$$

where d_{k} is the dimension of $\boldsymbol{K}$ and $\boldsymbol{Q}$, and d_{v} denotes the dimension of $\boldsymbol{V}$, which is set to be equal to d_{model}. Finally, we summarize all the rows in $\boldsymbol{Q}'$ to obtain the inter-relationship encoding vector as

$$\boldsymbol{H}^{\mathrm{inter}} = \sum_{r=1}^{L_{\mathrm{p}}} \boldsymbol{Q}'[r] \in \mathbb{R}^{d_{\mathrm{v}}}, \tag{7}$$

where r stands for each row in $\boldsymbol{Q}'$.

4.3 Statistical feature encoder

Based on the analysis in Section 3.3, it has been observed that some anomaly behaviors exhibited by compromised accounts are related to the design of some functions of some e-commerce platforms, including the payment password reset process. In some fraud cases, combining expertise in designing rules and statistical features can provide more accurate and efficient detection of compromised accounts. Similar to the approaches adopted by Ke et al. (2019) and Yao et al. (2020), we develop a statistical feature encoder, which employs a GBDT model to classify accounts based on statistical features, and we further exploit a single-layer neural network to obtain the statistical feature encoder by transforming leaf embeddings of GBDT.

GBDT (Friedman, 2001) is an ensemble model that builds a group of decision trees as weak learners in a boosting way. It is very popular in production environments for many applications such as click prediction (Ling et al., 2017) and fraud detection (Cao et al., 2019), owing to its effectiveness, efficiency, and interpretability. Existing studies (Ke et al., 2019; Yao et al., 2020) on incorporating knowledge from GBDT into neural networks learn an embedding matrix for one decision tree or a group of decision trees, where each row of the matrix is an embedding of the corresponding leaf node. Different from them, we directly take the leaf index as the representation of each leaf node, to reduce the number of learnable parameters. Specifically, our GBDT model is trained with all $\boldsymbol{f}_i$ (illustrated in Section 3.4) as input. For each account, the leaf embedding is calculated by concatenating all leaf indices of N decision trees as $\boldsymbol{D}_i = [l_1, l_2, \ldots, l_N]$. Finally, we use a single-layer fully-connected neural network to combine all the information from each decision tree. Thus, the statistical feature encoding vector for each account can be calculated according to $\boldsymbol{H}_i^{\mathrm{feature}} = \tanh(\boldsymbol{W}_{\mathrm{a}}\boldsymbol{D}_i + \boldsymbol{b}_{\mathrm{a}})$, where $\boldsymbol{W}_{\mathrm{a}}$ and $\boldsymbol{b}_{\mathrm{a}}$ are the corresponding weight and bias within the single layer, respectively.

4.4 Fusion module

The information from all the aforementioned encoding vectors should be incorporated to predict the label of an account. However, for different cases, the temporal patterns and the abnormality of accounts' operations may contribute differently to the final prediction. For instance, some accounts might encounter payment authorization difficulties due to the new owner's lack of knowledge of the payment password. Consequently, these accounts could lack sufficient behavioral history to capture or compare the temporal patterns of the previous and new owners, but exhibit an obvious abnormality. Therefore, instead of fusing the encoding vectors in a mean pooling or concatenating way, we leverage a self-attention module to automatically assign weights to the encoding vectors.

Similar to the inter-attention module, the self-attention module uses the scaled dot-product

attention mechanism. The queries, keys, and values in this module are the linear projection of the stacked matrix of all encoding vectors. After applying the attention function in Eq. (6), we derive $\boldsymbol{H}_i^{\text{all}}$ as the sum of all the rows of the output matrix.

Finally, $\boldsymbol{H}_i^{\text{all}}$ is fed into a fully-connected layer and then a softmax function is applied to obtain the prediction result $\hat{y}_i = \text{softmax}(\boldsymbol{W}_{\text{c}}\boldsymbol{H}_i^{\text{all}} + \boldsymbol{b}_{\text{c}})$, where $\boldsymbol{W}_{\text{c}}$ and $\boldsymbol{b}_{\text{c}}$ are the learnable weight matrix and bias matrix applied to transform $\boldsymbol{H}_i^{\text{all}}$ to a two-dimensional subspace, respectively.

The training process for the proposed model is outlined as follows. The GBDT model is initially trained with the training set and then yields leaf embeddings for all samples. The loss of detection of the compromised accounts is defined as a cross-entropy loss. Subsequently, our model is trained to minimize the cross-entropy loss function, which is specifically defined as

$$\mathcal{L} = -\sum_{i=1}^{n} \big(y_i \ln \hat{y}_i + (1 - y_i) \ln (1 - \hat{y}_i) \big), \qquad (8)$$

where n is the number of accounts, and y_i and $\hat{y}_i$ indicate the true and predicted labels for each account in our task, respectively.

5 Performance evaluation

In this section, we first describe dataset information, baseline methods, evaluation metrics, and experimental details (Section 5.1; Sections 1 and 2 in the supplementary materials). Then we present the performance comparison over baselines and TSF variants for the Meituan dataset (Section 5.2) and IEEE-CIS dataset (Section 3 in the supplementary materials). Furthermore, we conduct a case study (Section 5.3) and introduce feature importance analysis (Section 5.4) for the Meituan dataset to enhance the interpretability of TSF. Finally, we discuss the scalability and some potential applications of our model (Section 5.5).

5.1 Experimental setup

5.1.1 Dataset information

We use two datasets, the Meituan dataset and IEEE-CIS dataset (https://www.kaggle.com/competitions/ieee-fraud-detection/data), to evaluate the performance of the proposed TSF. For the Meituan dataset introduced in Section 3.1, we use the extracted 33 statistical features (mentioned in Section 3.3) as the input of the GBDT model. The details of the 33 features can be found in Section 4 in the supplementary materials. The payment and login records have 19 and 11 features, respectively. These features include access environments, payment details, and login methods. Each value of a categorical attribute such as device, delivery address, and GPS location is represented by a positive integer within an account. For example, a set of four devices of an account is represented as $\{1, 2, 3, 4\}$. We randomly split this dataset into training, validation, and test sets with a ratio of 3:1:1. The IEEE-CIS dataset (Mainali et al., 2022; Nti and Somanathan, 2024) is a representative real-world e-commerce transaction fraud dataset available on Kaggle. Detailed descriptions of the IEEE-CIS dataset can be found in Section 1 in the supplementary materials.

5.1.2 Baseline methods

No existing methods are tailored for the detection of compromised accounts via PNR. Since there are no significant graph features such as clustering and connections among accounts, we carefully select six representative baseline models of machine learning and deep neural networks, including LightGBM (Ke et al., 2017), GBDT embedding+NN (Ke et al., 2019), LSTM (Hochreiter and Schmidhuber, 1997), GRU (Cho et al., 2014), DeepScan (Gong et al., 2018), and Al-Qurishi et al. (2018)'s. The specific setups for these baseline methods can be found in Section 2 in the supplementary materials.

5.1.3 Evaluation metrics

Similar to previous studies (Yu et al., 2018; Gong et al., 2023), we adopt AUC as a metric of performance. It is commonly used in fraud detection as it is robust to the imbalance of positive and negative samples (Liu et al., 2018; Dou et al., 2020). Meanwhile, we aim at detecting as many compromised accounts as possible while reducing the disturbance to normal accounts caused by misclassification, that is, to achieve a high recall rate while maintaining a high precision similar to Li A et al. (2019). Therefore, Recall@T%Precision, which refers to the recall when the precision is T%, is an essential metric to differentiate the performances of TSF and the baselines.

For the Meituan dataset, we set T to 85, 90, and 95, separately. For the IEEE-CIS dataset, we set T to 80, 85, and 90, separately.

5.1.4 Implementation details

We train TSF in a mini-batch way. For the Meituan dataset, we test the lengths of payment and login sequences L_p and L_l in the range of [20, 200], and finally set them as 100 and 50, respectively. Sequences with a shorter length are padded with all zero vectors. The sequential model aims to capture the possible change before and after an account is compromised. If the sequences are longer than L_p (resp. L_l), we take the first and last sub-sequences of $L_\mathrm{p}/2$ (resp. $L_\mathrm{l}/2$) length and concatenate them. We train our model using Adam (Kingma and Ba, 2015) at a learning rate of 0.001 with 30 epochs. For the IEEE-CIS dataset, we take the same strategy on the behavioral data for each account and set the lengths of two input sequences as 5. We train our model using Adam (Kingma and Ba, 2015) at a learning rate of 0.001 with 15 epochs for the IEEE-CIS dataset. For a fair comparison, we train and test TSF and all baselines with the same settings five times, and take the average of the five tests for each metric as results.

5.2 Performance of TSF for the Meituan dataset

5.2.1 Performance comparison

We present the results of all methods for the Meituan dataset in Table 3, and the corresponding precision–recall curves are illustrated in Fig. 5.

First, one can observe that the proposed TSF outperforms all the baseline models, with a significant improvement in all metrics, particularly for the Recall@T%Precision in Table 3. While maintaining a high precision rate of 95%, our model is still able to detect up to 73.58% of compromised accounts, which demonstrates the effectiveness and usability of our model.

Second, while the sequential models (i.e., GRU and LSTM) focus on learning the temporal relationships in users' behavioral sequences, the ensemble models (i.e., LightGBM, Al-Qurishi et al. (2018)'s, and GBDT embedding+NN) use hand-crafted rules that may capture the anomalies in users' multiple operations to classify accounts. Although the two kinds of models have similar performances in AUC, the ensemble models can maintain a higher recall when improving the precision, which can avoid interruption to normal users. According to Table 3, the sequential models can achieve a good recall when the precision is relatively low (85%); however, as the precision increases, the recall suffers from a sharp drop. Although DeepScan leverages sequential models and ensemble models, its performance is subject to degradation because it does not consider the importance of these features. In our scenario, a well-developed deep learning model is an optimal solution.

Overall, TSF outperforms the baselines with a 3.89%–4.95% AUC increase and a 16.55%–50.49% recall increase under three fixed precision thresholds

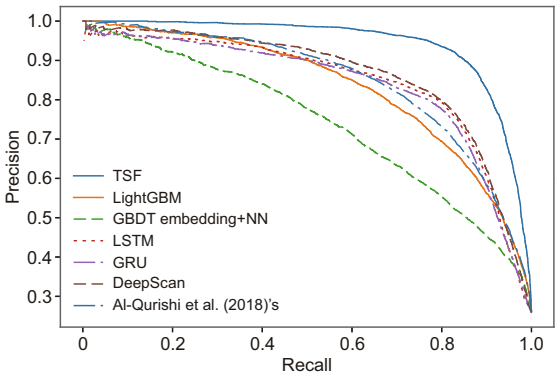

Fig. 5 Precision–recall curves of TSF and the baselines for the Meituan dataset

Table 3 Performances of TSF and baselines for the Meituan dataset

Model	AUC	R@85%P	R@90%P	R@95%P
LightGBM (Ke et al., 2017)	0.9195	0.6006	0.5035	0.3329
GBDT embedding+NN (Ke et al., 2019)	0.9242	0.6387	0.5208	0.3370
GRU (Cho et al., 2014)	0.9195	0.6625	0.4998	0.2309
LSTM (Hochreiter and Schmidhuber, 1997)	0.9224	0.6810	0.5245	0.2758
DeepScan (Gong et al., 2018)	0.9301	0.7170	0.5990	0.3840
Al-Qurishi et al. (2018)'s	0.9264	0.6512	0.5450	0.3781
TSF	**0.9690**	**0.8825**	**0.8372**	**0.7358**

The best results are in bold. AUC: area under curve; R@T%P: Recall@T%Precision. T: 85, 90, or 95

for the Meituan dataset. Furthermore, we have performed McNemar's test (McNemar, 1947) on the results, validating that our model is significantly different from other methods with $p < 0.01$.

5.2.2 Ablation study

To explore the importance of each module in our model, we perform ablation studies by removing one specific module at a time. Here, we evaluate the detection performance of four variants for the Meituan dataset: (1) TSF (w/o TPE)—TSF without the whole temporal pattern encoder, i.e., the statistical feature encoder in Fig. 4; (2) TSF (w/o SFE)—TSF without pre-trained GBDT or the single-layer neural network to transform leaf embedding, i.e., the temporal pattern encoder in Fig. 4; (3) TSF (w/o IA)—TSF without the inter-attention mechanism to align payment behaviors with login behaviors; and (4) TSF (w/o SA)—TSF without the self-attention mechanism but using concatenation to fuse information.

As shown in Table 4 and Fig. 6, the performance degradation indicates that all modules of our model contribute to the task of detecting compromised accounts via PNR for the Meituan dataset. The following are the findings of the ablation study:

Temporal pattern encoder. Removing the temporal pattern encoder has the most severe impact on the performance of our model. In this case, the AUC drops by 0.0487, and the recall decreases by up to 0.4214. The resulting model also performs much worse than TSF (w/o SFE), with a recall reduced by 0.2065. This again shows that solely using handcrafted features to detect compromised accounts is ineffective as the new owner of an account compromised by PNR usually has regular and stable payment behaviors. However, compromised accounts always have different temporal patterns owing to user changes. A temporal pattern model capturing the evolution information of accounts can be useful for detecting compromised accounts.

Statistical feature encoder. After removing the statistical feature encoder, the AUC score drops by 2.21% to 94.69%, and the recall decreases drastically with the improvement of precision. Our results highlight the importance of specifying rules to capture significant behaviors, and through leaf embedding, we show how the knowledge from a pre-trained GBDT model can be applied to detection. Notably, TSF (w/o SFE) still significantly outperforms the LSTM and GRU models with a recall increase by 11.57%–29.84%. This indicates that our temporal pattern encoder can better learn the temporal information of user behaviors by considering the effect of irregular time intervals and the interactions between two sequences.

Inter-attention module. The performance of our model decreases slightly after removing the inter-attention module, which aims to capture the interactions between payment events and certain login events. The result implies that the inter-attention module also contributes to the detection performance of TSF as well as that the correlation between two sequences is essential.

Self-attention module. In our experiments, the

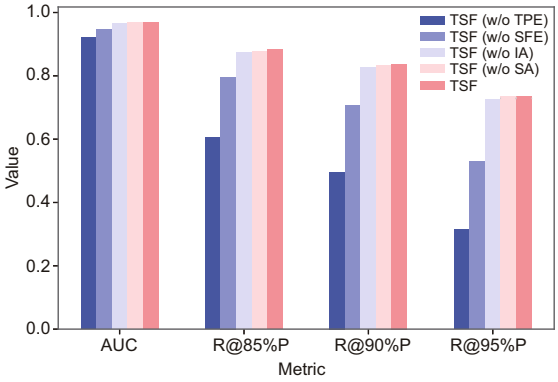

Fig. 6 Ablation study for the Meituan dataset

Table 4 Results of the ablation study for the Meituan dataset

Model	AUC	R@85%P	R@90%P	R@95%P
TSF (w/o TPE)	0.9203 (−4.87%)	0.6047 (−27.78%)	0.4962 (−34.10%)	0.3144 (−42.14%)
TSF (w/o SFE)	0.9469 (−2.21%)	0.7967 (−8.58%)	0.7068 (−13.04%)	0.5293 (−20.65%)
TSF (w/o IA)	0.9657 (−0.33%)	0.8746 (−0.79%)	0.8267 (−1.05%)	0.7241 (−1.17%)
TSF (w/o SA)	0.9649 (−0.41%)	0.8717 (−1.08%)	0.8259 (−1.13%)	0.7064 (−2.94%)
TSF	**0.9690**	**0.8825**	**0.8372**	**0.7358**

The best results are in bold. The values in parentheses indicate the ratio of metric value decrease compared with that of TSF. AUC: area under curve; R@T%P: Recall@T%Precision. T: 85, 90, or 95

model with the self-attention module consistently outperforms TSF (w/o SA) with the recall improved by up to 2.94%. Notably, this module can interpret the detection results by demonstrating the importance of different representations to the prediction.

5.3 Case study for the Meituan dataset

To further investigate the decision-making process of our model, we select two representative compromised accounts from the Meituan dataset and perform a case study. The self-attention mechanism of the proposed model enables it to capture the relationships and dependencies among the input data, and thus provides an explanation for the decision-making process of our model regarding the data. Thus, we focus on the ability of the self-attention module to fuse information from the Meituan dataset. In our model, the self-attention module assigns different weights to the four encoding vectors (i.e., $\boldsymbol{H}^{\mathrm{p}}$, $\boldsymbol{H}^{\mathrm{l}}$, $\boldsymbol{H}^{\mathrm{inter}}$, and $\boldsymbol{H}^{\mathrm{feature}}$) for detection. Their attention weights are listed in Table 5.

Table 5 Case study for the Meituan dataset

Case	Attention weight			
	Payment	Login	Inter-attention	Statistical features
1	0.5791	0.1075	0.2141	0.0993
2	0.4526	0.1362	0.0347	0.3765

In case 1, this account has no record of bank card binding or manipulating payment password attempts. The recent payment statistics are within the distribution range of most accounts. Thus, it is hard to distinguish it from normal accounts through statistical features. In the payment sequence, there is a noticeable change in access environments and habits before and after a silent period of about 1400 d. Following the silent period, the device information, delivery address, IP address, and frequently used business all change. Moreover, analyzing login behaviors can provide valuable information for payment. Specifically, after the silent period, the account is suddenly logged in with a new device, which is also used in subsequent payment via SMS. Multiple login attempts fail before finally succeeding. Results from Table 5 demonstrate how the self-attention module assigns greater weights to the payment and the interrelationship encoding vector.

In case 2, this account shows frequent unsuccessful attempts to link a bank card. Furthermore, there is a sudden increase in the payment amount from around 30 CNY to a few hundred CNY, which implies that there might be some anomalies in the user's activities. Notably, there is also a switch in the access environment during the payment sequence of the account. Therefore, attention weights for the abnormality encoding vector and payment behavior encoding vector are higher.

Additionally, we notice that the average attention weights of the four encoding vectors of all compromised accounts are 0.5479, 0.1407, 0.1015, and 0.2099, respectively. Therefore, to detect compromised accounts, it is crucial to consider the temporal information of behavioral sequences, especially the payment sequence, which is more important than that of hand-crafted features. Our analysis indicates that the self-attention module can take advantage of various aspects of accounts and also allows for clear interpretation of the detection results.

5.4 Statistical feature importance analysis for the Meituan dataset

In the Meituan dataset, four representative events of each account have been analyzed in detail, including the payment, login, binding of the bank card, and manipulation of the passwords. The statistical features are listed in Table S3 in the supplementary materials. These statistical features are also used as the input of the statistical feature encoder of TSF (Section 4.3). We calculate the importance score of each feature as the total decrease in the cross-entropy objective function of all the splits using that feature. The 10 most important features are listed in Table 6.

5.5 Discussion

The performance of the proposed TSF model has been validated in the above experimental results. We would like to discuss the scalability of the proposed model and some potential applications.

First, model scalability is critical for practical applications on large-scale e-commerce platforms. We discuss the two core components, namely the LSTM model (used in Section 4.2) and the GBDT technique (used in Section 4.3), included in our model. The LSTM model is a type of well-established

Table 6 Top 10 features and their importance scores

Feature	Importance score
Number of card binding records	668 173.95
Number of logins via SMS in 90 d	261 422.25
Number of successful card bindings in 90 d	194 851.78
Total payment amount in 6 months	182 175.04
Number of failed payment records in 30 d	167 175.48
Maximum length of silent periods	122 230.82
Number of payment records in 30 d	103 628.88
Number of failed card bindings in 30 d	76 628.77
Number of payment records in 90 d	59 196.52
Average amount per payment in 30 d	56 217.99

SMS: short message service

temporal model that has been widely used in a variety of domains and it has been demonstrated with good scalability on large-scale datasets (Greff et al., 2017). Moreover, GBDT is a popular method in machine learning due to its excellent performance and scalability (Ke et al., 2017). In summary, these two components will increase the overall scalability of our model.

Second, we discuss some potential applications of our model. In the context of e-commerce, TSF can identify unusual indicatives of fraudulent activities by analyzing sequential patterns of users' activities, such as login activities, password resets, and transaction histories. Therefore, TSF could detect other financial frauds such as account takeover, credit card fraud, and financial transaction fraud by recognizing deviations from typical user behavior patterns. Similarly, our model can be leveraged for in-depth user behavior analysis in e-commerce platforms. It can help companies gain insights into customer preferences, purchasing habits, and product interests over time. For example, TSF could identify patterns of user behavior and provide recommendations to optimize the user experience. This enables more effective personalized marketing strategies and product recommendations. Furthermore, the adaptability of the proposed model could be extended to the healthcare domain. It could be applied to analyze the sequences of patient data, such as electronic health records. TSF is capable of helping identify patient health trends over time, predict disease outbreaks, and personalize treatment plans based on the sequential data of patient symptoms and medical history.

In summary, our model has the potential for practical applications in various domains such as e-commerce, online social networks, and healthcare. It should also be noted that the experimental performance depends on the specific practical scenarios and data characteristics.

6 Conclusions

In this paper, we have identified the vulnerability of SMS-based 2FA, a widely adopted authentication method on e-commerce platforms. We revealed that PNR is the primary cause of this vulnerability. We focused on the problem of detecting compromised accounts via PNR, a widely prevalent problem in e-commerce platforms. By analyzing the statistical behaviors and the evolution process of accounts on a real-world dataset provided by Meituan, we observed that the compromised accounts and the normal ones have distinctive behaviors and evolution patterns. Inspired by our observations, we proposed the TSF model to detect the compromised accounts from the normal ones. Extensive experiments on two real-world datasets (the Meituan dataset and the IEEE-CIS dataset) demonstrated that our model significantly outperforms some carefully selected baselines. The effectiveness of our model in detecting compromised accounts has been demonstrated. A case study and the analysis of feature importance conducted on the Meituan dataset provided explanations of the proposed TSF model.

In the future, we attempt to construct the association graphs (Gao et al., 2023) between accounts and entities (such as devices and IP addresses) before and after the silence periods. With these association graphs, we will define the strength of the multi-relations and use graph neural networks to distinguish whether the user using the account is the same person, i.e., whether the account has been released twice. We also plan to expand our approach to distinguishing more fine-grained types like lost, stolen, and other abnormally lost numbers for compromised accounts. Besides, we will evaluate the scalability of our model once we obtain new larger-scale real-world datasets. We plan to deploy TSF to Meituan's production pipeline to detect compromised accounts via PNR and other factors like identity theft. Furthermore, we will continue to develop more comprehensive strategies like data augmentation and transfer learning techniques, for further addressing class imbalance challenges. We will focus on new techniques and approaches to further improve the scalability of

our model and ensure its stability and performance in real-world applications.

Contributors

All the authors contributed to the study conception and design. Min GAO, Shutong CHEN, Yangbo GAO, Zhenhua ZHANG, Yu CHEN, and Yang CHEN proposed the motivation of the study. Min GAO, Shutong CHEN, and Qiongzan YE performed the experiments. Min GAO drafted the paper. All the authors commented on previous versions of the paper. Min GAO, Yupeng LI, Xin WANG, and Yang CHEN revised the paper. All the authors read and approved the final version of the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Alibaba Cloud, 2023. Phone Number Verification Service (in Chinese). https://www.alibabacloud.com/product/verify?spm=a3c0i.23458820.2359477120.2.2e137d3frQSEAI [Accessed on Mar. 25, 2023].

Al-Qurishi M, Hossain MS, Alrubaian M, et al., 2018. Leveraging analysis of user behavior to identify malicious activities in large-scale social networks. *IEEE Trans Ind Inform*, 14(2):799-813. https://doi.org/10.1109/TII.2017.2753202

Baytas IM, Xiao C, Zhang X, et al., 2017. Patient subtyping via time-aware LSTM networks. Proc 23rd ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining, p.65-74. https://doi.org/10.1145/3097983.3097997

Bhattacharyya S, Jha S, Tharakunnel K, et al., 2011. Data mining for credit card fraud: a comparative study. *Dec Support Syst*, 50(3):602-613. https://doi.org/10.1016/j.dss.2010.08.008

Bilge L, Strufe T, Balzarotti D, et al., 2009. All your contacts are belong to us: automated identity theft attacks on social networks. Proc 18th Int Conf on World Wide Web, p.551-560. https://doi.org/10.1145/1526709.1526784

Bonaccorsi G, Pierri F, Cinelli M, et al., 2020. Economic and social consequences of human mobility restrictions under COVID-19. *Proc Natl Acad Sci USA*, 117(27):15530-15535. https://doi.org/10.1073/pnas.2007658117

Boshmaf Y, Logothetis D, Siganos G, et al., 2015. Integro: leveraging victim prediction for robust fake account detection in OSNs. Proc 22nd Network and Distributed System Security Symp, p.8-11.

Branco B, Abreu P, Gomes AS, et al., 2020. Interleaved sequence RNNs for fraud detection. Proc 26th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining, p.3101-3109. https://doi.org/10.1145/3394486.3403361

Cao SS, Yang XX, Chen C, et al., 2019. TitAnt: online real-time transaction fraud detection in ant financial. *Proc VLDB Endowment*, 12(12):2082-2093. https://doi.org/10.14778/3352063.3352126

Chai YD, Zhou YH, Li WF, et al., 2022. An explainable multi-modal hierarchical attention model for developing phishing threat intelligence. *IEEE Trans Depend Sec Comput*, 19(2):790-803. https://doi.org/10.1109/TDSC.2021.3119323

Cheng DW, Xiang S, Shang CC, et al., 2020. Spatio-temporal attention-based neural network for credit card fraud detection. Proc 34th AAAI Conf on Artificial Intelligence, p.362-369. https://doi.org/10.1609/aaai.v34i01.5371

Cho K, van Merriënboer B, Gulcehre C, et al., 2014. Learning phrase representations using RNN encoder–decoder for statistical machine translation. Proc Conf on Empirical Methods in Natural Language Processing, p.1724-1734. https://doi.org/10.3115/v1/D14-1179

Dmitrienko A, Liebchen C, Rossow C, et al., 2014. On the (in)security of mobile two-factor authentication. Proc 18th Int Conf on Financial Cryptography and Data Security, p.365-383.

Doerfler P, Thomas K, Marincenko M, et al., 2019. Evaluating login challenges as a defense against account takeover. Proc World Wide Web Conf, p.372-382. https://doi.org/10.1145/3308558.3313481

Dou YT, Liu ZW, Sun L, et al., 2020. Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. Proc 29th ACM Int Conf on Information & Knowledge Management, p.315-324. https://doi.org/10.1145/3340531.3411903

Egele M, Stringhini G, Kruegel C, et al., 2017. Towards detecting compromised accounts on social networks. *IEEE Trans Depend Sec Comput*, 14(4):447-460. https://doi.org/10.1109/TDSC.2015.2479616

Federal Communications Commission of the United States, 2018. Reassigned Numbers Database. https://www.fcc.gov/reassigned-numbers-database [Accessed on Apr. 1, 2023].

Friedman JH, 2001. Greedy function approximation: a gradient boosting machine. *Ann Statist*, 29(5):1189-1232. https://doi.org/10.1214/AOS/1013203451

Fu YY, Zhang M, Xu X, et al., 2021. Partial feature selection and alignment for multi-source domain adaptation. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.16654-16663. https://doi.org/10.1109/CVPR46437.2021.01638

Gao M, Li Z, Li RC, et al., 2023. EasyGraph: a multifunctional, cross-platform, and effective library for interdisciplinary network analysis. *Patterns*, 4(10):100839. https://doi.org/10.1016/j.patter.2023.100839

Gong QY, Chen Y, He XL, et al., 2018. DeepScan: exploiting deep learning for malicious account detection in location-based social networks. *IEEE Commun Mag*, 56(11):21-27. https://doi.org/10.1109/MCOM.2018.1700575

Gong QY, Liu YS, Zhang JY, et al., 2023. Detecting malicious accounts in online developer communities using deep learning. *IEEE Trans Knowl Data Eng*, 35(10):10633-10649. https://doi.org/10.1109/TKDE.2023.3237838

Greff K, Srivastava RK, Koutník J, et al., 2017. LSTM: a search space Odyssey. *IEEE Trans Neur Netw Learn Syst*, 28(10):2222-2232. https://doi.org/10.1109/TNNLS.2016.2582924

He Y, Wang C, Li N, et al., 2020. Attention and memory-augmented networks for dual-view sequential learning. Proc 26th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining, p.125-134. https://doi.org/10.1145/3394486.3403055

Hochreiter S, Schmidhuber J, 1997. Long short-term memory. *Neur Comput*, 9(8):1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hu BB, Zhang ZQ, Shi C, et al., 2019. Cash-out user detection based on attributed heterogeneous information network with a hierarchical attention mechanism. Proc 33rd AAAI Conf on Artificial Intelligence, p.946-953. https://doi.org/10.1609/aaai.v33i01.3301946

Huang JQ, Hu K, Tang QT, et al., 2021. Deep position-wise interaction network for CTR prediction. Proc 44th Int ACM SIGIR Conf on Research and Development in Information Retrieval, p.1885-1889. https://doi.org/10.1145/3404835.3463117

Karimi H, VanDam C, Ye LY, et al., 2018. End-to-end compromised account detection. Proc IEEE/ACM Int Conf on Advances in Social Networks Analysis and Mining, p.314-321. https://doi.org/10.1109/ASONAM.2018.8508296

Kawase R, Diana F, Czeladka M, et al., 2019. Internet fraud: the case of account takeover in online marketplace. Proc 30th ACM Conf on Hypertext and Social Media, p.181-190. https://doi.org/10.1145/3342220.3343651

Ke GL, Meng Q, Finley T, et al., 2017. LightGBM: a highly efficient gradient boosting decision tree. Proc 31st Int Conf on Neural Information Processing Systems, p.3149-3157.

Ke GL, Xu ZH, Zhang J, et al., 2019. DeepGBM: a deep learning framework distilled by GBDT for online prediction tasks. Proc 25th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining, p.384-394. https://doi.org/10.1145/3292500.3330858

Keren G, Schuller B, 2016. Convolutional RNN: an enhanced model for extracting features from sequential data. Proc Int Joint Conf on Neural Networks, p.3412-3419. https://doi.org/10.1109/IJCNN.2016.7727636

Kingma DP, Ba J, 2015. Adam: a method for stochastic optimization. Proc 3rd Int Conf on Learning Representations.

Lee K, Narayanan A, 2021. Security and privacy risks of number recycling at mobile carriers in the United States. Proc APWG Symp on Electronic Crime Research, p.1-17. https://doi.org/10.1109/eCrime54498.2021.9738792

Li A, Qin Z, Liu RS, et al., 2019. Spam review detection with graph convolutional networks. Proc 28th ACM Int Conf on Information and Knowledge Management, p.2703-2711. https://doi.org/10.1145/3357384.3357820

Li S, Liu K, Meng R, 2018. Research and design of interface for reassigned mobile numbers. Proc IEEE 18th Int Conf on Communication Technology, p.1311-1314. https://doi.org/10.1109/ICCT.2018.8599932

Liang T, Zeng GX, Zhong QW, et al., 2021. Credit risk and limits forecasting in e-commerce consumer lending service via multi-view-aware mixture-of-experts nets. Proc 14th ACM Int Conf on Web Search and Data Mining, p.229-237. https://doi.org/10.1145/3437963.3441743

Ling XL, Deng WW, Gu C, et al., 2017. Model ensemble for click prediction in Bing search ads. Proc 26th Int Conf on World Wide Web Companion, p.689-698. https://doi.org/10.1145/3041021.3054192

Liu ZQ, Chen CC, Yang XX, et al., 2018. Heterogeneous graph neural networks for malicious account detection. Proc 27th ACM Int Conf on Information and Knowledge Management, p.2077-2085. https://doi.org/10.1145/3269206.3272010

Mainali P, Psychoula I, Petitcolas FAP, 2022. ExMo: explainable AI model using inverse frequency decision rules. Proc 3rd Int Conf on Human-Computer Interaction, p.179-198. https://doi.org/10.1007/978-3-031-05643-7_12

McDonald A, Sugatan C, Guberek T, et al., 2021. The annoying, the disturbing, and the weird: challenges with phone numbers as identifiers and phone number recycling. Proc CHI Conf on Human Factors in Computing Systems, Article 559. https://doi.org/10.1145/3411764.3445085

McNemar Q, 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153-157. https://doi.org/10.1007/BF02295996

Mirian A, DeBlasio J, Savage S, et al., 2019. Hack for hire: exploring the emerging market for account hijacking. Proc World Wide Web Conf, p.1279-1289. https://doi.org/10.1145/3308558.3313489

Mobile China, 2017. Mobile Authentication: Capitalising on China's Identity Market. https://www.gsma.com/solutions-and-impact/technologies/mobile-identity/gsma_resources/mobile-authentication-capitalising-chinas-identity-market [Accessed on Mar. 1, 2023].

Mulliner C, Borgaonkar R, Stewin P, et al., 2013. SMS-based one-time passwords: attacks and defense. Proc 10th Int Conf on Detection of Intrusions and Malware, and Vulnerability Assessment, p.150-159. https://doi.org/10.1007/978-3-642-39235-1_9

Nti IK, Somanathan AR, 2024. A scalable RF-XGBoost framework for financial fraud mitigation. *IEEE Trans Comput Soc Syst*, 11(2):1556-1563. https://doi.org/10.1109/TCSS.2022.3209827

Ping YK, Gao C, Liu TC, et al., 2021. User consumption intention prediction in Meituan. Proc 27th ACM SIGKDD Conf on Knowledge Discovery & Data Mining, p.3472-3482. https://doi.org/10.1145/3447548.3467178

Tao JL, Wang H, Xiong T, 2018. Selective graph attention networks for account takeover detection. Proc IEEE Int Conf on Data Mining Workshops, p.49-54. https://doi.org/10.1109/ICDMW.2018.00015

Thomas K, Akhawe D, Bailey M, et al., 2021. SoK: hate, harassment, and the changing landscape of online abuse. Proc IEEE Symp on Security and Privacy, p.247-267. https://doi.org/10.1109/SP40001.2021.00028

VanDam C, Tan PN, Tang JL, et al., 2018. CADET: a multi-view learning framework for compromised account detection on Twitter. Proc IEEE/ACM Int Conf on Advances in Social Networks Analysis and Mining, p.471-478. https://doi.org/10.1109/ASONAM.2018.8508654

Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention is all you need. Proc 31st Int Conf on Neural Information Processing Systems, p.6000-6010.

Viswanath B, Bashir MA, Crovella M, et al., 2014. Towards detecting anomalous user behavior in online social networks. Proc 23rd USENIX Security Symp, p.223-238.

Wang C, Zhu HY, 2022. Representing fine-grained co-occurrences for behavior-based fraud detection in online payment services. *IEEE Trans Depend Sec Comput*, 19(1):301-315. https://doi.org/10.1109/TDSC.2020.2991872

Wang C, Wang CQ, Zhu HY, et al., 2020. LAW: learning automatic windows for online payment fraud detection. *IEEE Trans Depend Sec Comput*, 18(5):2122-2135. https://doi.org/10.1109/TDSC.2020.3037784

Wang DX, Lin JB, Cui P, et al., 2019. A semi-supervised graph attentive network for financial fraud detection. Proc IEEE Int Conf on Data Mining, p.598-607. https://doi.org/10.1109/ICDM.2019.00070

Wang J, Zou JH, Wang HY, 2022. Sampling with replacement vs Poisson sampling: a comparative study in optimal subsampling. *IEEE Trans Inform Theory*, 68(10):6605-6630. https://doi.org/10.1109/TIT.2022.3176955

Welch BL, 1951. On the comparison of several mean values: an alternative approach. *Biometrika*, 38(3-4):330-336. https://doi.org/10.2307/2332579

Xu T, Goossen G, Cevahir HK, et al., 2021. Deep entity classification: abusive account detection for online social networks. Proc 30th USENIX Security Symp, p.4097-4114.

Yao TJ, Li Q, Liang SS, et al., 2020. BotSpot: a hybrid learning framework to uncover bot install fraud in mobile advertising. Proc 29th ACM Int Conf on Information & Knowledge Management, p.2901-2908. https://doi.org/10.1145/3340531.3412690

Ye QZ, Gao YB, Zhang ZH, et al., 2022. Modeling access environment and behavior sequence for financial identity theft detection in E-commerce services. Proc Int Joint Conf on Neural Networks, p.1-8. https://doi.org/10.1109/IJCNN55064.2022.9892383

Yu JF, Qiu MH, Jiang J, et al., 2018. Modelling domain relationships for transfer learning on retrieval-based question answering systems in E-commerce. Proc 11th ACM Int Conf on Web Search and Data Mining, p.682-690. https://doi.org/10.1145/3159652.3159685

Zhang YB, Zhao DB, Zhang J, et al., 2011. Interpolation-dependent image downsampling. *IEEE Trans Image Process*, 20(11):3291-3296. https://doi.org/10.1109/TIP.2011.2158226

Zou YX, Roundy K, Tamersoy A, et al., 2020. Examining the adoption and abandonment of security, privacy, and identity theft protection practices. Proc CHI Conf on Human Factors in Computing Systems, p.1-15. https://doi.org/10.1145/3313831.3376570

List of supplementary materials