International Journal of Software and Informatics, ISSN 1673-7288
http://www.ijsi.org, ijsi@iscas.ac.cn, +86-10-62661048
IJSI, 2021, 11(4): 453–472, doi: 10.21655/ijsi.1673-7288.00256

International Journal of Software and Informatics

Research Article

Smooth Intervention Model of Individual Interactive Behavior

Xiao Liu (刘霄)[1], Zhaohui Zhang (章昭辉)[1,2], Ziming Wei (魏子明)[1], Pengwei Wang (王鹏伟)[1]

[1] (School of Computer Science and Technology, Donghua University, Shanghai 201620, China)
[2] (Shanghai Engineering Research Center of Network Information Services, Shanghai 201804, China)
Corresponding author: Zhaohui Zhang, zhzhang@dhu.edu.cn

Abstract User feature extraction and identity authentication methods based on interactive behavior are an important method of identity recognition. However, for high-frequency users, the interactive behavior patterns and operating habits are relatively stable, which are easily imitated by fraudsters and make the existing models have a higher misjudgment. The key to solving the above problems is to make the users' behavior change smoothly and distinguishably. This study proposes a smooth intervention model based on an individual interactive behavior system to handle it. Firstly, according to the users' historical web behavior log, the change trend of users' interactive behavior is obtained from multiple dimensions. Then, combined with the stability and deviation of the behavior, the Time-Domain Drift Algorithm (TDDA) is proposed to determine the behavior guidance time of each user. Finally, an intervention model for interactive behavior reconstruction systems is proposed, which superimposes behavior trigger factors on non-critical paths in the system to guide users to generate new interactive behavior habits. Experiments prove that the method proposed in this study could guide the user behavior to change smoothly and produce sufficient distinction to significantly advance the model accuracy in the scenario of behavior camouflage anomaly detection.

Keywords behavioral intervention; interactive behavior; identity theft; identity recognition; Petri net modeling

Citation Liu X, Zhang ZH, Wei ZM, Wang PW. Smooth Intervention Model of Individual Interactive Behavior, *International Journal of Software and Informatics*, 2021, 11(4): 453–472. http://www.ijsi.org/1673-7288/256.htm

The rapid development of e-commerce allows users to enjoy a high-quality life brought by the Internet and also constantly exposes them to different varieties of organized online fraud. According to statistics from CNNIC, more than 30% of Internet users experienced personal information leakage in 2018; more than 25% of users encountered online fraud; 23.8% of users encountered viruses or Trojan attacks, and 19.2% of user accounts or passwords were stolen. For companies in the Internet industry, the gray industry not only interferes with normal

This is the English version of the Chinese article "个体交互行为的平滑干预模型. 软件学报, 2021, 32(6): 1733–1747. doi: 10.13328/j.cnki.jos.006247".
Funding items: Natural Science Foundation of Shanghai (19ZR1401900); High-Tech Field Project of Shanghai Science and Technology Innovation Action Plan (19511101302); National Natural Science Foundation of China (61472004, 61602109)
Received 2020-08-30; Revised 2020-10-26; Accepted 2020-12-19; IJSI published online 2021-12-23

business but also leads to real economic losses. Many start-up Internet companies launch a variety of promotions at the beginning of their market expansion. While these benefits entice many users to change their habits, they can also be the targets of practitioners in the gray industry. Many criminals use illegitimate software and hardware to carry out fraudulent economic activity through false identities and behavior[1]. The cognitive consistency theory holds that the attitudes and behavior of users tend to remain in a balanced and stable state[2]. These common characteristics of users are often stable and can be easily captured by fraudsters, so that they can bypass the monitoring rules, generate large profits, and even gradually form a whole set of gray industrial chains. However, these industries are not perceived by the public because of their concealment and anti-reconnaissance. Therefore, how to effectively and accurately verify user identity has become a major problem in need of solutions.

Nowadays, most identity authentication technologies are based on users' account names and passwords[3]. As long as the users' identity is authenticated within a short time, no matter what the users' real identity is, all the behavior of the users will be regarded as legal ones[4]. However, fraudsters use phishing sites to impersonate banks and other financial service providers to defraud sensitive information such as email addresses and passwords, allowing them to disguise their identities and commit fraud. Therefore, a user with a correct password does not mean that he/she is a legitimate user.

In order to make up for the shortcomings caused by identity authentication based simply on a single username and password, many scholars have also tended to use data feature mining and behavior analysis methods in the field of identity recognition in recent years. For example, Association Rule (AR) mining, hidden Markov decision process, semi-Markov decision process, Bayesian network, Neural Network (NN), Random Forest (RF), and other methods are used for the behavior modeling and prediction of user Web logs. Although efforts are being made to solve the problem of user identification, there are still many difficulties.

(1) The relevant models of machine learning are used, and the training data need to be marked. However, due to the sparsity of fraudulent behavior samples in reality, there will be extreme imbalance in the samples, which makes it difficult for the model to fully learn the characteristics of fraudulent behavior.

(2) Because the age, background and hobbies of users are relatively fixed, the interactive behavior pattern of the system is relatively stable in a certain period of time. With part of the stolen user behavior information, fraudsters make use of this feature to simulate the behavior of normal users and bypass the monitoring rules, so that the identity recognition model has a higher misjudgment of such fraudulent behavior.

(3) At present, the research on Web log data mining of users' interactive behavior and browsing behavior mostly obtains users' browsing preferences based on the feature extraction from the users' massive page browsing records in search engines and online shopping malls. The research is mostly used for the optimization of website structure, Web prediction or recommendation systems. However, in a single system with a single function, when the number of users is large, the similar behavior pattern of multi-user sharing appears[5], which reduces the differentiation of the identity recognition model.

(4) At present, there are still many challenges about how to extract the features of users' interactive behavior to construct user behavior portraits and the abnormal judgment criteria for interactive behavior.

In light of the above problems, a new method of identity recognition is proposed in this paper. Compared with other models, the method proposed in this paper can alleviate the above problems well and has the following advantages.

(1) From the users' point of view, continuously paying attention to the Web usage log of

users and constructing the interactive behavior portrait of users can well avoid the problem of sample imbalance.

(2) Considering the differences among users, a behavioral drift guidance model is proposed. According to the historical interactive behavior records of each user, the stability and deviation of the interactive behavior are comprehensively considered, and the timing of behavioral intervention is determined for each user, which effectively avoids the problem of multi-user sharing similar behavior patterns in a system with low complexity and single function.

(3) The definition of system behavior set is proposed, which is divided into critical behavior set and non-critical behavior set. The method of internal triggering of the system is adopted to superimpose new non-critical behavior processes on the premise of not destroying the running logic of the system. The process of non-mandatory constraint on user behavior enables user behavior to comply with the guidance mechanism, so as to gradually cultivate users to produce new interactive behavior habits and maintain certain behavior differences from the original interactive behavior to counter the existing identity camouflage and fraudulent behavior.

The research in this paper is based on a common inductive bias, namely behavioral stability bias. This deviation is manifested in three aspects: First, the users' interactive behavior will not change significantly in a short period of time; second, when a common account is in the scenario of behavior camouflage or identity theft, the interactive behavior of the fraudster will maintain a high similarity with the legitimate account holder; finally, due to the sparsity of fraudulent behavior, the continuous behavior guidance mechanism will change the interactive behavior habits of ordinary users, while the fraudsters whose identities are disguised will still maintain the interactive behavior of the original scenario.

To sum up, the main contributions of this paper are as follows: First, a method is proposed to depict the users' interactive behavior, which considers the characteristics of users' interactive behavior from multiple dimensions such as system login time and login interval; second, a behavioral drift guidance model is proposed to determine the timing and means of behavior intervention for each user, so that the behavior maintains certain differences and smoothness before and after intervention; third, experiments are designed to verify the feasibility and accuracy of the model in the detection of behavior camouflage anomaly. Therefore, the first section of this paper describes the related work in detail. Section 2 discusses the modeling approach proposed in this paper. Section 3 introduces the data sources, the experimental results and the comparative experiments. Section 4 summarizes the research results of this paper and the prospects for the future.

1 Related Work

In recent years, anomaly detection based on users' interactive behavior has gradually gained attention and the existing research is mainly divided into two research directions: modeling the users' operation behavior of computer peripherals and modeling the behavior of system interaction browsing logs. Regarding the analysis of operation behavior of computer peripherals, Roth *et al*.[6] proposed a method to authenticate users by continuous typing. Ma *et al*.[7] proposed a method based on mouse behavior authentication on a dynamic soft keyboard. In the aspect of modeling the browsing log generated by the interaction between the users and the system, Liu *et al*.[8] used the system browsing time and the navigation path to determine the normality of user behavior and the behavior access control methods for the separation of identity authentication and behavior authentication in the open network environment. Zhao *et al*.[9] used the behavioral Markov model to describe the logical behavior of users and proposed an authentication method based on sequence and preference. Zheng *et al*.[10] defined the coefficient of diversity and Logic Graph of BP (LGBP) based on information entropy to characterize the diversity of user

transaction behavior. In addition, a new credit card Fraud Detection System (FDS) based on Behavior Certificate (BC) was proposed[11], which used the users' behavior certificate to identify the users' transactions. Chen *et al*.[12] designed an architecture for the browsing behavior authentication system of mobile terminal APP that comprehensively considered multiple factors, which was suitable for the users who use mobile terminal APP in daily life. Zhong *et al*.[13] proposed an authentication method based on Web browsing behavior. Zhang *et al*.[14] proposed a Web browsing behavior forensics method based on AR mining. Liu *et al*.[15, 16] put forward the concept of security mutual simulation and adopted the formal method for the first time to distinguish two systems with the same function but different security according to the binarization theory. Zhang *et al*.[17] expanded the data records of low-frequency users through the DBSCAN clustering algorithm by migrating the behavior and transaction status of the current transaction group, which enabled a more accurate behavior characterization of low-frequency users.

Reasonable guidance and intervention on users' interactive behavior makes the users change their usage habits, which has been studied by many scholars. According to the Fogg behavior model, Toledo *et al*.[18] established a set of persuasive system modeling framework by combining fuzzy rules and fuzzy reasoning to analyze the motivation and ability level of users triggering learning activities. Hamper *et al*.[19] and Nishiyama *et al*.[20] designed the method and theoretical structure of the application based on the persuasion model according to the characteristics of TTM and FBM of users, and completed the theoretical modeling of the application to promote users to carry out physical fitness exercise in combination with the incentive mechanism. Zhang *et al*.[21] tried to induce the behavior changes through visual and auditory stimuli and automatically adjusted the style of notification icons in the system based on the theory of Behavior Change Support System (BCSS) and the analysis of the current environment state. Although the guidance and intervention of users' interactive behavior has been partially applied and output in some areas of interaction design, relatively complete theoretical systems and practical methods have not yet been formed. Most of the research remains in the stage of analysis and theoretical perfection, and there is a lack of the qualitative and quantitative analysis and the effect verification on the difference and smoothness before and after behavioral changes.

2 Model Method

This section will introduce an intervention model of individual interactive behavior, which mainly includes two parts: The first part is the generation of users' interactive behavior portrait, and the second part is the construction of behavioral drift guidance model. The behavioral drift guidance model is divided into the TDDA and interactive behavior reconstruction model. The users' interactive behavior portraits will be generated from multiple dimensions according to the history of normal system interaction behavior of each user. The behavioral drift guidance model first determines the time of behavior intervention based on the users' historical interactive behavior portraits. Then, according to the obtained intervention time and the internal triggering of the system, the behavior process of the system is changed by superimposing new behavior paths to achieve non-mandatory constraints on users' behavior.

2.1 Generation of users′ interactive behavior portrait

In this paper, it is assumed that the personal access record of users can be recorded and retrieved by the server, and the historical Web interactive behavior log of users can be generated by continuous monitoring. This section will introduce the generation process of the interactive behavior benchmarks of users and generate the interactive behavior portraits of users based on the historical Web interactive log data of users.

Definition 1 (Interactive behavior record)**.** An interactive behavior record i of the user in the system contains m attributes, denoted as

$$i = \{a_1, a_2, a_3, \cdots, a_m | a_1 \in A_1, a_2 \in A_2, a_3 \in A_3, \cdots, a_m \in A_m\}$$

In this paper, the attributes of interaction record include user label, session number, login time, page number, page entry time, page duration, namely the interactive behavior. Given a user label u, the interactive behavior log of the user is the collection of historical interaction records up to the current date and can be denoted as $R_u = \{i_1^u, i_2^u, i_3^u, \cdots, i_n^u\}$, where $n = |R_u|$ is the number of interaction records of the user. The interactive behavior of users records the normal behavior in the log, namely $T_u = (t \in R_u | \mathit{label} = \text{true})$, where $n_{ut} = |T_u|$. For the normal interactive behavior of users, we need to further analyze and process to obtain the interactive behavior portrait of users. The attributes of the users' interactive behavior portraits are defined as follows:

$$\begin{cases} A_1^u = \{a \in A_1 | \exists r \in r_u : a \in r\} \\ A_2^u = \{a \in A_2 | \exists r \in r_u : a \in r\} \\ \quad\quad \vdots \\ A_m^u = \{a \in A_m | \exists r \in r_u : a \in r\} \end{cases}$$

where $A_1^u \subseteq A_1, A_2^u \subseteq A_2, \cdots, A_m^u \subseteq A_m$. Without loss of generality, $A_1^u = \{a_1^i, a_2^i, \cdots, a_m^i\}$ is defined.

Definition 2 (Attribute of system login time)**.** The system login time of user u is defined as an n-tuple of user login probability in n time periods, denoted as $\mathit{LTA}^u = (\mathit{time}_1, \mathit{time}_2, \cdots, \mathit{time}_n)$. In the normal interactive behavior log R_u of user u, A_{time}^u is taken as the time set of the users' login, and $n_{\mathit{time}}^u = |A_{\mathit{time}}^u|$. In order to distinguish the preferences and habits of the login time of different users, we divide the login time into 12 intervals, calculate the time interval corresponding to each element a_i^{time} in A_{time}^u, and label each element. The following subsets can be obtained:

$$\begin{cases} lta_1 & = \{a_i^{\mathit{time}} \in A_{\mathit{time}}^u | 0 \le \log \mathit{intime} < 2\} \\ lta_2 & = \{a_i^{\mathit{time}} \in A_{\mathit{time}}^u | 2 \le \log \mathit{intime} < 4\} \\ lta_3 & = \{a_i^{\mathit{time}} \in A_{\mathit{time}}^u | 4 \le \log \mathit{intime} < 6\} \\ & \vdots \\ lta_{12} & = \{a_i^{\mathit{time}} \in A_{\mathit{time}}^u | 22 \le \log \mathit{intime} < 24\} \end{cases}$$

In this way, $\mathit{time}_1 = \frac{|lta_1|}{n_{\mathit{time}}^u}, \mathit{time}_2 = \frac{|lta_2|}{n_{\mathit{time}}^u}, \cdots, \mathit{time}_{12} = \frac{|lta_{12}|}{n_{\mathit{time}}^u}$ can be calculated to obtain the login time attribute of the users:

$$\mathit{LTA}^u = (\mathit{time}_1, \mathit{time}_2, \mathit{time}_3, \cdots, \mathit{time}_{12})$$

Definition 3 (Attribute of working time login)**.** The working time login attribute of user u is defined as $\mathit{WTA}^u = (\mathit{isworktime}, \mathit{noworktime})$ to indicate the probability of the users' login time occurring at working time on working days, where working days do not include weekends and statutory holidays. According to each element in the set A_{time}^u, whether it belongs to working time is judged. According to the judgment result, each element is labeled with T and F, which represents the working time login and the non-working time login, so that two subsets are obtained as follows:

$$\begin{cases} wta_1 &= \{a \in A^u_{time} | label = T\} \\ wta_2 &= \{a \in A^u_{time} | label = F\} \end{cases}$$

Thus, $isworktime = \frac{|wta_1|}{n^u_{time}}, noworktime = \frac{|wta_2|}{n^u_{time}}$ can be calculated to obtain the users' working time login attribute:

$$WTA^u = (isworktime, noworktime)$$

Definition 4 (Login interval attribute)**.** The login interval attribute of user u is defined as $LIA^u = (period_1, period_2, \cdots, period_5)$, which indicates the probability that the login time interval of the users occurs in each interval, and reflects the interactive behavior habit of the users logging in the system. According to each element in the set A^u_{time}, the login time interval is calculated in turn and the set A^u_{period} is obtained, where $n^u_{period} = |A^u_{period}|$. The calculation formula for each element in the set A^u_{period} is as follows:

The first quartile Q_1, second quartile Q_2, third quartile Q_3, upper limit $Q_{\max}$, lower limit $Q_{\min}$ of the set A^u_{period} are calculated and the set is divided into five subsets, namely,

$$\begin{cases} lia_1 &= \{a^u_{period} \in A^u_{period} | Q_{\min} \le a^u_{period} < Q_1\} \\ lia_2 &= \{a^u_{period} \in A^u_{period} | Q_1 \le a^u_{period} < Q_2\} \\ lia_3 &= \{a^u_{period} \in A^u_{period} | Q_2 \le a^u_{period} < Q_3\} \\ lia_4 &= \{a^u_{period} \in A^u_{period} | Q_3 \le a^u_{period} < Q_{\max}\} \\ lia_5 &= \{a^u_{period} \in A^u_{period} | a^u_{period} \le Q_{\min}, a^u_{period} \ge Q_{\max}\} \end{cases}$$

Therefore, $period_1 = \frac{|lia_1|}{n^u_{period}}, period_2 = \frac{|lia_2|}{n^u_{period}}, \cdots, period_5 = \frac{|lia_5|}{n^u_{period}}$ can be obtained, namely that the users' login interval attribute is

$$LIA^u = (period_1, period_2, period_3, period_4, period_5)$$

Definition 5 (Key-page dwell time attribute)**.** The key-page dwell time attribute of user u is defined as $KSA^u = (distance_1, distance_2, \cdots, distance_n)$, which represents the probability of the time that the user stays on the carrier page corresponding to the behavior guidance trigger. In the normal interactive behavior log R_u of user u, the residence time set of the key page $a_{\text{page\_no}} = key$ of the user history is taken out, where $n^u_{\text{distance}} = |A^u_{distance}|$. In order to distinguish users' preferences and habits of interaction time on key pages, according to all elements in the set $A^u_{distance}$, the first quartile Q_1, second quartile Q_2, third quartile Q_3, upper limit $Q_{\max}$, and lower limit $Q_{\min}$ of the set are calculated, and the set is divided into five subsets, namely,

$$\begin{cases} ksa_1 &= \{a^u_{distance} \in A^u_{distance} | Q_{\min} \le a^u_{distance} < Q_1\} \\ ksa_2 &= \{a^u_{distance} \in A^u_{distance} | Q_1 \le a^u_{distance} < Q_2\} \\ ksa_3 &= \{a^u_{distance} \in A^u_{distance} | Q_2 \le a^u_{distance} < Q_3\} \\ ksa_4 &= \{a^u_{distance} \in A^u_{distance} | Q_3 \le a^u_{distance} < Q_{\max}\} \\ ksa_5 &= \{a^u_{distance} \in A^u_{distance} | a^u_{distance} \le Q_{\min}, a^u_{distance} \ge Q_{\max}\} \end{cases}$$

Therefore, $distance_1 = \frac{|ksa_1|}{n^u_{distance}}, distance_2 = \frac{|ksa_2|}{n^u_{distance}}, \cdots, distance_5 = \frac{|ksa_5|}{n^u_{distance}}$ can be obtained, namely that the user login interval attribute is $KSA^u = (distance_1, distance_2, distance_3, distance_4, distance_5)$.

Definition 6 (Interactive behavior characteristics)**.** It is assumed that $IBC^u = (LTA^u, WTA^u, LIA^u, KSA^u)$ is the interactive behavior characteristic of user u. According to attributes such as login time, whether the interactive behavior occurs during working time, the time interval of login time, and the residence time of key pages of the system, the interactive behavior of the user is defined as a 24-dimensional feature vector IBC^u to describe the user's interactive behavior:

(1) $LTA^u = (time_1, time_2, \cdots, time_{12})$ represents the attribute of login time, where $time_k$ ($k = 1$–12) respectively indicates the probability of the user's login behavior in each time period.

(2) $WTA^u = (isworktime, noworktime)$ represents the login attribute of working time, where *isworktime* and *noworktime* respectively indicate the probability of users' interaction behavior on working days and non-working days.

(3) $LIA^u = (period_1, period_2, \cdots, period_n)$ represents the login interval attribute, where $period_k$ ($k = 1$–5) indicates the probability that the interval between the user's current login behavior and the last login behavior is within lia_k.

(4) $KSA^u = (distance_1, distance_2, distance_3, distance_4, distance_5)$ represents the page dwell time attribute, where $distance_k (k = 1$–$5)$ indicates the probability that the user stays in the key page of the system within ksa_k.

2.2 Construction of behavioral drift guidance model

Generally speaking, everyone has relatively unchanged behavior habits, which are determined by their character, age, occupation, etc. For example, introverts are more cautious when exposing information; older people are slower to operate; most computer professionals operate faster. These common features are often easily captured by fraudsters. Additionally, everyone's behavior and habits can be changed relatively, but most of the changes come from the conditions imposed by the outside world. In most transaction systems, since system behavior only focuses on business logic and business functions, the user behavior contained in transaction data is realized through the behavior of the transaction system. When the system does not actively interfere with the user, the behavior contained in the data is the behavior formed by the general user itself and the behavior imposed by a user on the system is usually invariant. The user with higher transaction frequency will have more stable behavior. Therefore, when user data is stolen and simulated by fraudsters, the original behavior model of legitimate users will not be able to distinguish the fraudulent behavior. In addition, if the user's transaction frequency is too low, it is equivalent to no user behavior, which makes the fraudster easier to simulate user behavior. Therefore, in order to enable user behavior to actively combat fraudulent behavior and maintain a good interactive experience, we should smoothly drift the original user behavior.

This section proposes the Coefficient of Concentration (*CS*) and Coefficient of Preferences (*CP*) of interactive behavior to quantify the user behavior, proposes a Time-Domain Drift Algorithm (TDDA) based on *CS* and *CP* to determine the timing of trigger factors, and implements the reconstruction access control model for interaction behavior under the action of TDDA.

2.2.1 *TDDA*

In order to make the guide time domain after drift differ from the original login time domain of the user, after fully considering the smoothness of the change, this paper analyzes the record of the users' historical interactive behavior and uses the quantile analysis method and the Inter Quartile Range (*IQR*) to describe the dispersion of different user login time series. Quantile is the variable at each equal position after all the data of the whole are arranged in descending order. The first quartile Q_1, the second quartile Q_2 namely the median, and the third quartile Q_3 of the user login time records are obtained, and then the upper limit $Q_{\max}$ and the lower limit $Q_{\min}$ are solved.

IQR refers to the difference between the third quartile Q_3 and the first quartile Q_1, namely, $IQR = Q_3 - Q_1$. The larger *IQR* presents the greater variability. On the contrary, the smaller *IQR* presents the smaller variability. To reduce the influence of extreme values on the measurement of user behavior stability, we calculate the upper and lower limits of user behavior according to the following equations:

$$\begin{cases} Q_{\max} & = Q_3 + \alpha IQR \\ Q_{\min} & = Q_1 - \alpha IQR \end{cases}$$

where α is the weight of abnormality. If α is larger, more deviation points will be accepted. Conversely, if α is smaller, more deviation points will be excluded.

Definition 7 (*CS* of interactive behavior)**.** The *CS* of interactive behavior of user u is defined as CS^u, which represents the stability and concentration of user behavior and is the ratio of the range of the observed values Q_1 and Q_3 in the middle of the samples to the range of the upper and lower limits of the samples. The calculation formula is as follows:

$$CS^u = \frac{IQR}{Q_{\max} - Q_{\min}} = \frac{R\left[\frac{3(n+1)}{4}\right] - R\left[\frac{n+1}{4}\right]}{Q_3 + \alpha IQR - (Q_1 - \alpha IQR)}$$

where R is the overall sorted set, and n is the number of elements in the set. Therefore, according to the *CS* values, we can measure the dispersion of each user's interactive behavior events. With the users' login time attribute as an example, if the corresponding *CS* value is larger under its login time attribute, the users' historical login time is more concentrated, and the behavior dispersion is smaller. On the contrary, if the corresponding *CS* value is smaller, the users' login time is more scattered, and the behavior dispersion is higher. On this basis, the changes in the users' *CS* values in different time intervals also reflect the concentration trend of users' interactive behavior. On the basis of measuring the stability and aggregation of users, we should also consider the bias of users' interactive behavior.

Definition 8 (*CP* of interaction behavior)**.** The *CP* of interactive behavior of user u is defined as CP^u, which indicates the preference and bias of user behavior, namely the difference between the sample mean and the second quartile. The calculation formula is as follows:

$$CP = \bar{R} - Q_2 = \frac{1}{n}\sum_{i=1}^{n} r_i - Q_2$$

where $\bar{R}$ is the mean value of the overall set, and r is all the elements in the set R. With user login time as an example, the time for different users to log in the system is related to their own habits and nature of work: If the time for users to log in the system is more inclined to the left side of the mean, $CP < 0$; on the contrary, if the time for users to log in the system is more inclined to the right side of the mean, $CP > 0$. The change of CP^u in different time intervals also reflects the preference trend of users' interactive behavior.

The guidance of user behavior needs to be fully combined with the users' own behavior ability and habits, and does not bring a greater burden to the user experience, so that the users will have a higher degree of acceptance. In this paper, according to users' historical interactive behavior habits, on the basis of CS^u and CP^u, a TDDA of interactive behavior is proposed by fully considering users' historical behavior. Algorithm 1 presents the pseudo-codes of the TDDA for user login behavior, as shown below.

Algorithm 1. TDDA on login behavior

Input: R—The list of user login time log; n—The size of user login time log.
Output: *drift_start*—The guide mechanism start time; *drift_stop*—The guide mechanism end time.

threshold: $= 0.4$;
sort(R);
calculate $Q_1, Q_2, Q_3, \bar{R}$;
$IQR = Q_3 - Q_1$
$CS = IQR/(Q_{\max} - Q_{\min})$
$CP = \bar{R} - Q_2$
if $CS >$ *threshold* **then**
    if $CP > 0$ **then**
        drift_start $= \bar{R}$; *drift_stop* $= Q_3$;
    else
        drift_start $= Q_1$; *drift_stop* $= \bar{R}$;
    end if
else
    drift_start $= Q_1$; *drift_stop* $= Q_3$
end if
return *drift_start*, *drift_stop*;

2.2.2 *Interactive behavior reconstruction model*

In this section, Petri nets are used as a system modeling tool. Firstly, the system behavior set is divided into critical behavior set and non-critical behavior set. The interactive behavior Petri net is defined and the definition of behavior profile for the interactive behavior Petri net is given. On this basis, the Petri net model of interactive behavior reconstruction is proposed, and the effectiveness of the model is verified by an example of an online credit trading system.

As a modeling and analysis tool for concurrent and distributed systems, Petri nets have a strong theoretical support for the analysis of the nature and behavior of the system[22], which is also widely used in the analysis and optimization of business process modeling. In the field of computer software systems, Petri nets can be used for the complete modeling and the analysis of the related structure and property[23, 24]. Besides, because there must be system interaction between the software system and the users, the interaction between the users and the system platform often reflects the users' interest in the service provided by the platform, the degree of attention, and the interactive behavior habit. Therefore, the analysis of user access behavior is undoubtedly of direct and important value in evaluating and optimizing the business process and service setting of the platform.

Definition 9 (System behavior set)**.** $S_A = \{s_1, s_2, \cdots, s_n\}$ is set as the total set of behavior events that can be triggered during the normal operation of the system. The system behavior set is further divided into critical behavior set S_A^* and non-critical behavior set S_A'.

Definition 10 (Critical behavior set)**.** $S_A^* = \{s_1, s_2, \cdots, s_p\}$ is set as the entire set of key behavior events that can be triggered during the normal operation of the system, where $p < n$. The critical behavior set corresponds to the core function page of the system and undertakes the operation of the key functions of the system. The critical behavior process cp^* is the specific arrangement of the elements in the set S_A^*. Since the critical behavior process reflects the operating logic of the core functions of the system, it has a certain business logic sequence, namely $|cp^*| < p!$.

Definition 11 (Non-critical behavior set)**.** $S_A' = \{s_1, s_2, \cdots, s_q\}$ is the total set of non-critical behavior events that can be triggered during the normal operation of the system, where $q < n$. The set of non-critical behaviors corresponds to the secondary function page of the system and plays a supplementary role for the key functions of the system. The non-critical

behavior process cp' is the specific arrangement of the elements in the set S'_A, namely $|cp'| < q!$.

The set of critical behaviors corresponds to the core function pages of the system, such as submitting loan applications, credit information verification, loan agreement signing and other key function pages in the online credit business system. The set of non-critical behaviors often contains other function business pages with low sensitivity, such as bank card information page, credit business browsing, help center, and personal center. As shown in Figure 1, the user behavior set $U_A = \{u_1, u_2, \cdots, u_m\}$ is a set of all the behavioral events that can occur to the user, and $U_A \subseteq S_A$. The operation process of users is the complete arrangement of all elements in the set U_A, and there are $m!$ kinds of user behavior sequences. In the figure, A represents the critical behavior set of the system; B represents the non-critical behavior set of the system; C represents the user behavior set ***Ru***, and $A + B$ represents the system behavior set S_A.

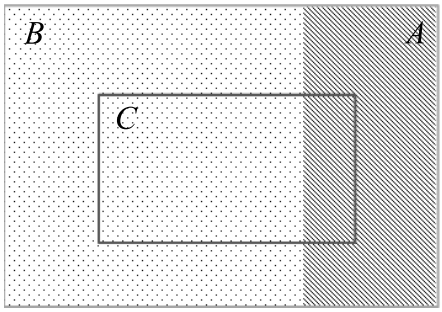

Figure 1 System behavior set

The trigger factor is the inducing factor that prompts the users to make a certain behavior and can be divided into the external trigger and the internal trigger: The external trigger is often determined by the external environment where the user is, while the internal trigger is embedded in the product and system, which is the key to causing behavior change. The internal trigger clearly conveys the next action to the user in a friendly interaction manner, and the user frequently associates the trigger factor with the system usage behavior, so that it can develop new interactive behavior habits. Therefore, we change the system behavior process by internally triggering and superimposing new non-critical behavior processes to non-compulsorily constrain the user behavior process.

According to the data characteristics of users' interactive behavior log and the definition of Petri net, Petri nets of system behavior and users' interactive behavior are defined as follows:

Definition 12 (Petri net of users' interactive behavior)**.** The Petri net of the system behavior for the online system of a certain platform is assumed as $PN = (S; T; F)$, and then, the Petri net of interactive behavior of the users in the system is $IPN = (IS; IT; IF)$, where

(1) $IS \subseteq S$ is the library element corresponding to the relevant input and output executed by the users;

(2) $IT \subseteq T$ is the transition element corresponding to the users' interactive behavior;

(3) $S \cap T = \varnothing$;

(4) $IF \subseteq F$ is the flow relationship between the transition elements in the system Petri net *PN*, namely $IF \subseteq (IS \times IT) \cup (IT \times IS)$.

Definition 13 (System behavior profile)**.** It is assumed that the Petri net of system behavior is $PN = (S; T; F)$, and the set $MB_I = \{\Rightarrow, !, \forall\}$ is the behavior profile of Petri net $PN = (S; T; F)$. For any given transition pair $(t_1, t_2) \in (T \times T)$, one of the following relationships can be satisfied:

(1) Sequence relationship $\Rightarrow$: if $\tau(t_1, t_2) = \{\Rightarrow\}$, then $t_1 \succ t_2$ and $t_2 \not\succ t_1$;

(2) Parallel relationship !: If $\tau(t_1, t_2) = \{!\}$, then $t_1 \not\succ t_2$ and $t_2 \not\succ t_1$;

(3) Circular relationship $\forall$: If $\tau(t_1, t_2) = \{\forall\}$, then $t_1 \succ t_2$ and $t_2 \succ t_1$.

As shown in Figure 2, in the Petri net of system behavior, T_2 and T_4 are parallel and denoted as $T_2!T_4$. T_6, T_7 and T_8 occur sequentially and have a strict sequence relationship, which is denoted as $T_6 \Rightarrow T_7 \Rightarrow T_8$. T_9 and T_{10} have a circular relationship denoted as $T_9 \forall T_{10}$. With the three kinds of behavior profiles of the system, reconstruction of the interactive behavior under different behavior profiles is proposed and shown in Figure 3–Figure 5.

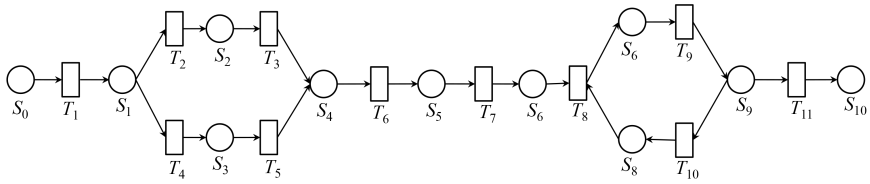

Figure 2 Example of system behavior Petri net

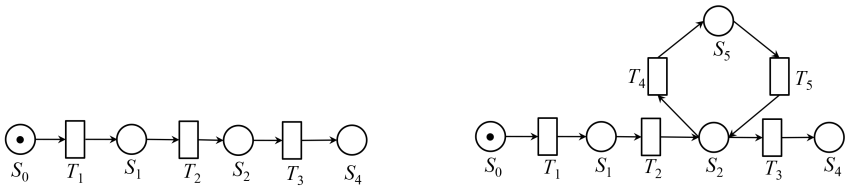

Figure 3 Reconstruction of sequence relationship

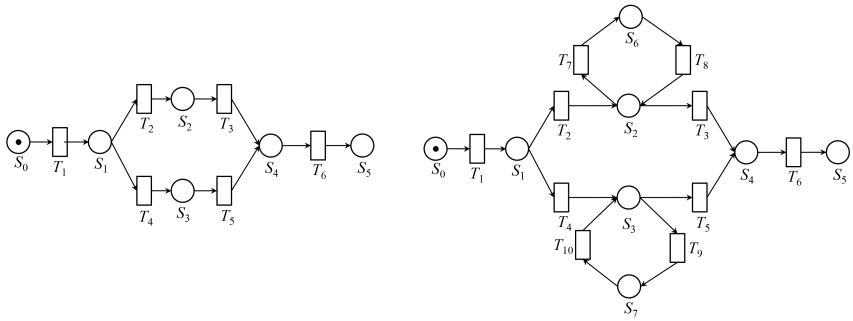

Figure 4 Reconstruction of parallel relationship

The construction of an interactive behavior reconstruction model is to enable legitimate user behavior to smoothly change and actively fight fraudulent behavior, without destroying the original interaction logic and user experience of the system. Therefore, a new non-critical behavior loop structure is superimposed on the profile of the basic interactive behavior, which not only keeps the basic business logic and processes of the system unaffected but also forms a new set of system behaviors. Because the purpose of the fraudsters in the operating system is very clear, in order to obtain illegal benefits as soon as possible, the fraudsters tend to keep a higher priority in the pages of the critical behavior set in the system, while keeping a low interest in the non-critical behavior set. On the other hand, the operation behavior of normal users keeps a balanced priority between the two sets. Therefore, with the above interactive

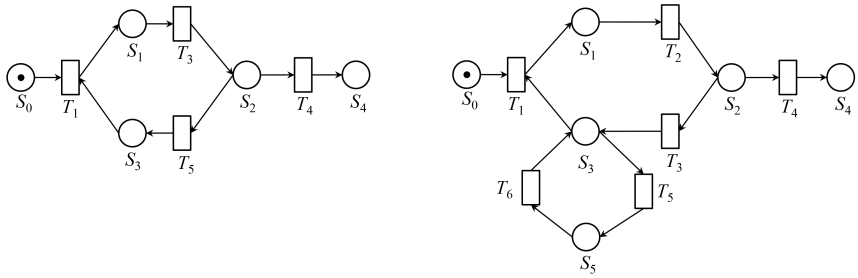

Figure 5 Reconstruction of circular relationship

behavior reconstruction method, the interactive behavior habits of the normal users can be changed to generate new interactive behavior habits under the guidance of continuous behavior reconstruction and gradually antagonize the fraudulent behavior of the fraudsters through the change of their own behavior.

The system used in this paper is an online credit trading system built by the laboratory, and the Petri net model of its business process is shown in Figure 6.

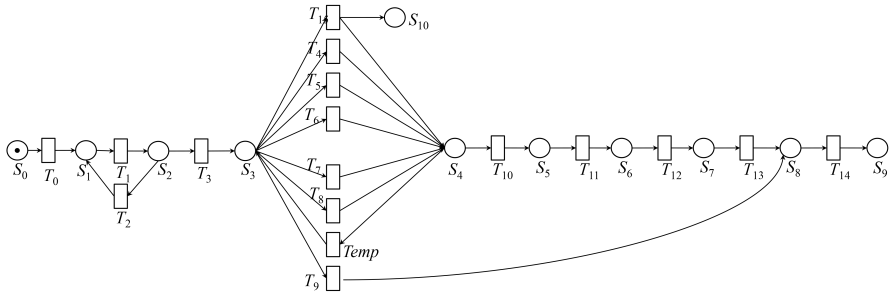

Figure 6 Petri net model for a credit trading system

The platform is divided into client and server, and its main business operations are shown in the description of the transition identification in Table 1, mainly including registering, logging in, browsing personal center, modifying bank card information, browsing credit projects, evaluating the line of credit, repaying loans, submitting loan applications, and other functions.

In this paper, the process of interactive behavior reconstruction and the corresponding control module are added to the system. The transition of reconstruction is the trigger condition of the control module in the system, as shown in Figure 7. After the user successfully logs in the home page of the system to trigger T_3, the control module T_{16} of the interactive behavior reconstruction will be activated at the same time. The activation condition of the interactive behavior reconstruction depends on the output result of the TDDA mentioned in the previous section, namely, whether T_{17} can be executed is judged. If the activation condition is met, T_{18} is activated. If the user activates T_4 while operating the system, the reconstruction and superimposition processes T_{21} and T_{22} of the interactive behavior will be activated to generate a new behavior process.

User behavior is ultimately manifested through system behavior. Therefore, in order to change user behavior, we should change the set of system behaviors to persuade and constrain the set of user behaviors. On the basis of determining the occurrence time of the trigger factors of the user login behavior through the TDDA, it is necessary to map the realization path of the trigger factors to the system behavior sequence by reconstructing the system behavior. Interactive behavior reconstruction is to superimpose new trigger factors to realize the path on the premise

of keeping the basic business logic of the trading system (namely the critical system behavior path) unchanged, thereby forming a new set of system behaviors.

Table 1 Corresponding names of transition indicator

Transition	Description
T_0	Registration
T_1	Login
T_2	Login failure
T_3	Access to homepage
T_4	Browsing of personal center
T_5	Check and modification of bank card information
T_6	Browsing of credit projects
T_7	Evaluation of the line of credit
T_8	Check of loan details
T_9	Repaying loans
T_{10}	Submission of loan applications
T_{11}	Signature of credit terms
T_{12}	Verification of personal information
T_{13}	Loans from platform
T_{14}	Update and restoration of lines
T_{15}	Logout

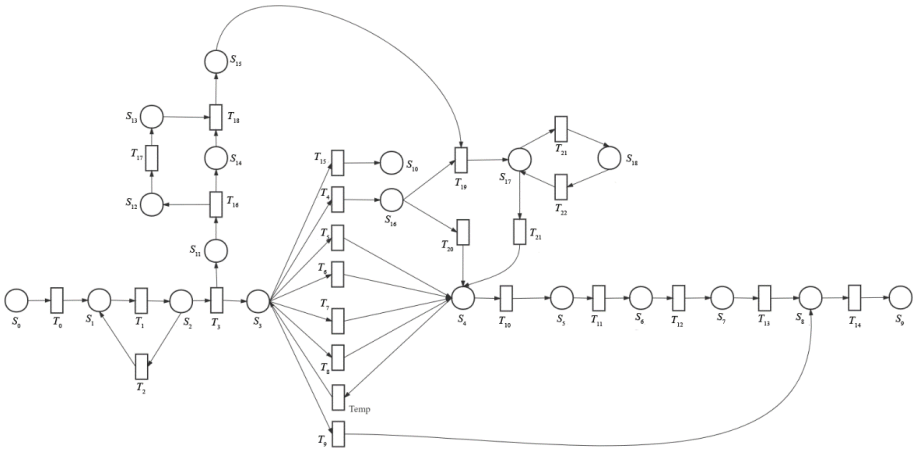

Figure 7 Petri net model for the interactive behavior reconstruction of the credit trading system

When users interact with each other under the guidance of periodically repeated TDDA and interactive behavior reconstruction model, new user behavior will gradually be formed. Because of the sparsity and clear purpose of the fraudulent behavior, the continuous behavior guidance mechanism will not be able to change the fraudsters' interactive behavior habits, so that the fraudsters with disguised identities still maintain the interactive behavior in the original scenario, which is different from the guided legitimate user behavior.

3 Experimental Results

In this section, we will verify the effectiveness of the persuasive interactive behavior-guided identity recognition method proposed in this paper through experiments. Firstly, the data set used in this experiment is introduced and then the stability and differentiation of the behavior-guided experiment are explained, as well as the effect in detecting the behavior anomaly of identity camouflage recognition.

3.1 Experimental data set

At present, there are few studies on behavior guidance and behavior change, and most of the studies stay in the stage of analysis and theoretical improvement. The effect of previous papers is mostly based on analytical conclusions, with a lack of qualitative and quantitative indicators to measure the difference of changes. In addition, the public data set before and after behavior change in continuous time periods has not been found in the survey. Therefore, the research data of this paper come from the online credit trading system built by the laboratory. According to the task flow of the credit product design, the system collects the users' operation data in the page through SDK on each page of the system, and records the sequence data of the users' operation page, including user name, users' mobile phone number, session ID, operation page number, page name, time to enter the page (timestamp) and time to leave the page (timestamp). This data set continuously collects 1,017 pieces of system login behavior and 16,741 page operation data of five users between September 2018 and April 2019. Among them, the records from September 2018 to December 2018 are the users' interactive behavior data in the original scenario of the system. From January to April 2019, according to the TDDA proposed in this paper, the behavior guidance mechanism is superimposed on the "personal center page" of the non-key behavior page of the system, and the data collected during this period is used as the record of users' interactive behavior after behavior guidance.

Because this paper studies the identity recognition in the behavior camouflage scenario, the identity camouflage fraud is highly similar to the normal interactive behavior. Because of the sparsity of black samples in the real transaction environment, we randomly mark 20% of the total interactive behavior data of each user's behavior record in the original data set as the black sample of the user's identity camouflage fraud, and the rest of the interaction records are taken as the white sample of the user. After the behavior guidance, the camouflage fraud is carried out, and these black samples of identity camouflage fraud are added to the corresponding user data set according to the timeline.

3.2 Contrast experiment

3.2.1 *Behavioral drift guidance model*

Figure 8–Figure 12 respectively reflect the time and frequency changes of users logging in the system before and after the behavioral drift guidance model is applied to the system.

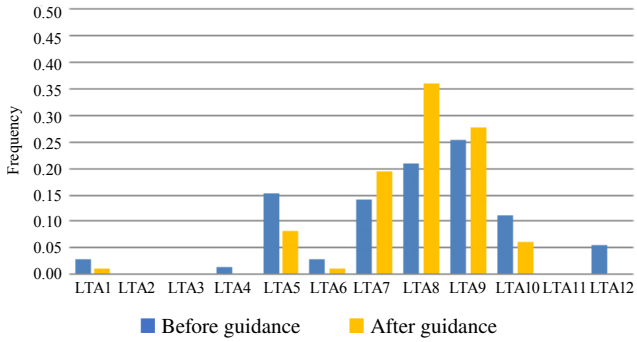

Figure 8 Frequency change of U1

In order to verify the change difference in the users' interactive behavior before and after the guidance, we introduce an evaluation index, Population Stability Index (*PSI*), to measure the change difference of user login probability distribution before and after the guidance. In the field of risk control modeling, PSI is often used to verify the distribution of samples in each

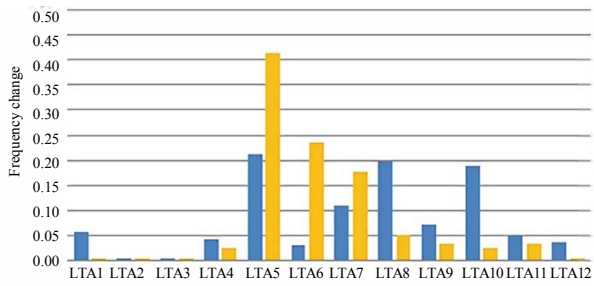

Figure 9 Frequency change of U2

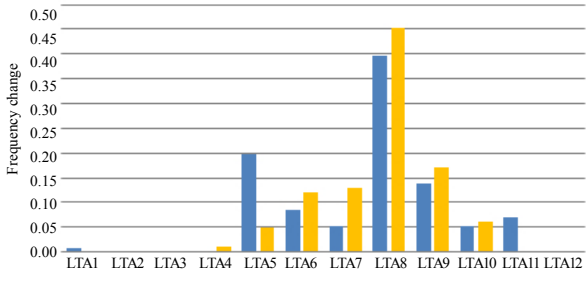

Figure 10 Frequency change of U3

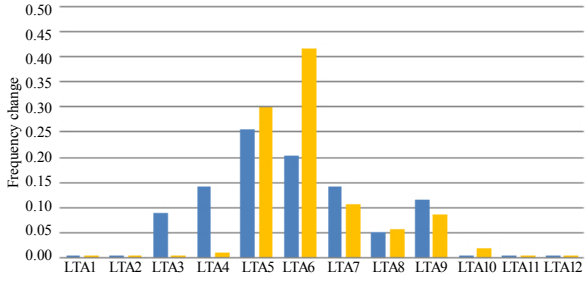

Figure 11 Frequency change of U4

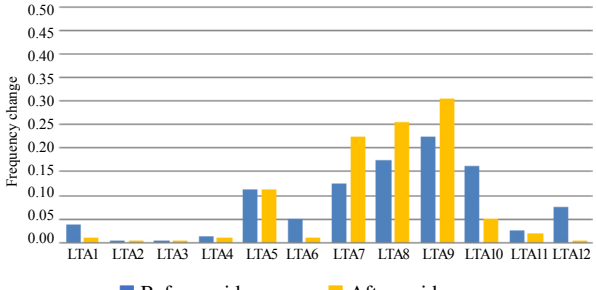

Figure 12 Frequency change of U5

score segment and the distribution stability of modeling samples, judge the deviation between the samples in the modeling period and the current samples, and measure the stability and adaptability of user groups. The calculation formula is as follows:

$$PSI = \sum_{i=1}^{n}(A_i - E_i) \times \ln\left(\frac{A_i}{E_i}\right)$$

where A_i is the actual proportion and E_i is the expected proportion. If the *PSI* value is smaller, the difference between the two distributions is smaller, which means that the two distributions are more stable. Generally speaking, the differences reflected by *PSI* are shown in Table 2.

Table 2 Meaning of *PSI*

PSI	Stability	Suggestions
0–0.1	Relatively stable	Few or no changes
0.1–0.25	Slightly instable	With certain changes, the subsequent changes should be paid attention to
>0.25	Instable	With obvious changes, the modal analysis is needed

The *PSI* value corresponding to the login time probability of each user before and after guidance is respectively calculated and listed in Table 3.

Table 3 *PSI* value of each user

User	Before guidance	Before and after guidance	After guidance
1	0.105303	0.466226	0.146624
2	0.169659	1.556132	0.166411
3	0.124616	0.628511	0.167358
4	0.109631	0.973336	0.100913
5	0.142602	0.658965	0.165745

It can be seen from Table 3 that

- Before the behavioral drift guidance model is applied to the system, the changes of *PSI* values corresponding to the probability distribution of login time of each user are less than 0.17, most less than 0.15, indicating that the login behavior of users remains relatively stable without additional intervention of system behavior. Therefore, behavior camouflage fraud is possible.
- After the incentive mechanism of drift guidance model is applied to the system, the *PSI* corresponding to the probability distribution of users' login behavior has changed in varying degrees before and after the guidance, and the *PSI* of each user is greater than 0.45, namely that under the incentive mechanism, compared with that in the original scenario, the probability distribution of user login time has changed obviously.
- After the continuous application of behavior-guided incentive mechanism, the *PSI* of each user tends to remain stable at about 0.16, which is slightly lower than that of the original behavior.

The experimental results show that after the incentive mechanism of the behavioral drift guidance model is applied to the system, the login time distribution of users has changed obviously, and under the guidance of the incentive mechanism proposed in this paper, the new login behavior pattern generated by the change of user behavior is relatively stable. Compared with the original login time distribution, the new login behavior pattern not only maintains a certain degree of differentiation but also ensures the smoothness of behavior changes.

3.2.2 *Identity recognition of camouflage anomaly*

The method of user identity recognition adopted in this paper models the users' interactive behavior for the 24-dimensional vectors merged by LTA^u, WTA^u, LIA^u and KSA^u based on

system interactive behavior, and the UR model proposed in previous research[25] is used to detect the identity camouflage anomaly of each user on two data sets before and after guidance.

Since the purpose of this paper is to identify abnormal behavior, the confusion matrix is slightly modified, with emphasis on the discrimination of abnormal interaction behavior. The modified confusion matrixes include the following parts: *TP* (True Positive) is the number of real abnormal behavior judged as abnormal behavior; *FP* (False Positive) is the number of real normal behavior judged as abnormal behavior; *TN* (True Negative) is the number of real normal behavior judged as normal behavior, and *FN* (False Negative) is the number of real abnormal behavior judged as normal behavior.

In order to make the comparison results more convincing, this experiment uses several common indicators in the field of fraud identification as the evaluation indicators of this paper. There are four evaluation indicators, namely accuracy, recall, precision, and $F1$-value. The calculation method is shown in Table 4: The accuracy represents the percentage of the number of correct judgments in the total number of behavior; the precision is the ratio of the real abnormal behavior judged by the model to the abnormal behavior judged by the model; the recall is the percentage of the real number of fraudulent transactions judged by the model in all abnormal behavior; the $F1$-value is a comprehensive index to measure the performance of the model, which is the harmonic mean of accuracy and recall.

Table 4 Indicator calculation method

Indicator name	Calculation method
Accuracy	$(TP + TN)/(TP + TN + FP + FN)$
Precision	$TP/(TP + FP)$
Recall	$TP/(TP + FN)$
$F1$-value	$2 \times Precision \times Recall/(Precision + Recall)$

It can be seen from the experimental results in Figure 13–Figure16 that in the detection experiment of users' identity camouflage anomaly using the UR model, the indexes after behavior guidance with TDDA are higher than those without TDDA. The details are as follows:

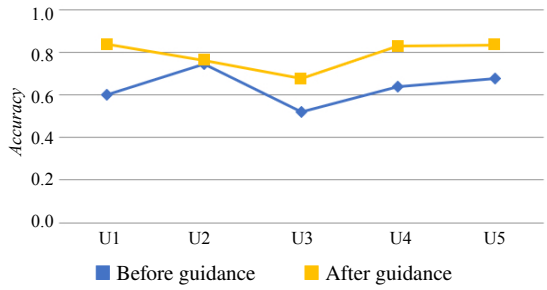

Figure 13 *Accuracy* of the OM model

- The accuracy results are shown in Figure 13, and the index increases by 15.12% on average, indicating that the TDDA guidance mechanism in camouflage behavior recognition can more accurately judge users' normal behavior and camouflage abnormal behavior.
- The precision is shown in Figure 14. Although there are fluctuations, the index is still 14.10% higher on average, indicating that the TDDA guidance mechanism is better at judging abnormal behavior in the vast majority of users.
- The recall is shown in Figure 15, with an average increase of 10.87%, indicating that after the TDDA behavior guidance mechanism is adopted, abnormal behavior is accurately judged, with fewer misjudgments about normal transactions.

- The $F1$-value is shown in Figure 16, and the $F1$-value reflects the overall performance of the model. It can be seen that after the TDDA behavior guidance mechanism is applied, the $F1$-value of the model increases by 28.76% on average. In other words, with the TDDA behavior guidance mechanism, the overall performance of the model is better than that in the original scenario without behavior guidance.

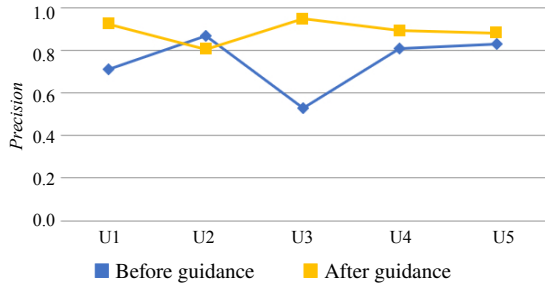

Figure 14 *Precision* of the OM model

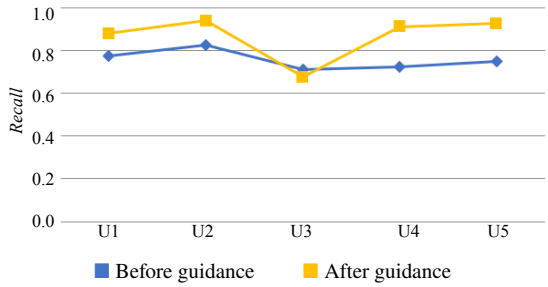

Figure 15 *Recall* of the OM model

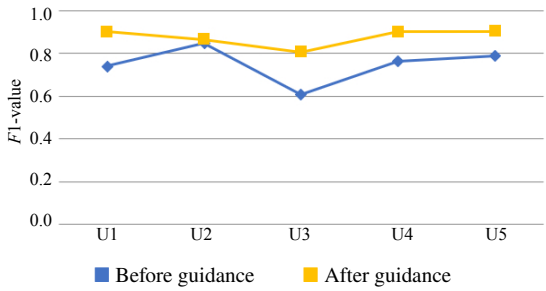

Figure 16 $F1$-value of the OM model

Through the analysis of the experimental results, it can be seen that after the TDDA behavior guidance mechanism is adopted, the detection result of behavior camouflage anomaly has a better overall performance than that in the original scenario. The main reasons are as follows: First, from the users' historical interactive behavior, based on the stability and preference of the user behavior, the intervention model of interactive behavior reconstruction system proposed in this paper can smoothly guide the change of user behavior to have a certain behavior differentiation before and after guidance, which provides a new solution for identity camouflage fraud detection. Second, the outlier analysis method of 1.5IQR is used to characterize the behavior stability and *CP*, which can better avoid the interference of outliers with the smoothness of the guidance

model. Therefore, the adoption of the TDDA behavior guidance mechanism can distinguish between normal behavior and camouflaged fraud without changing the original fraud detection method, so that the accuracy and precision of the model are significantly improved compared with those in the original scenario, and the model has a better overall performance in judging camouflage behavior.

4 Conclusions

In this paper, a smooth intervention model of individual interaction behavior is proposed. Starting from the historical interactive behavior of users, this model considers the differences between users, analyzes the stability and *CP* of users' interactive behavior, and proposes the TDDA of interactive behavior to determine the timing of behavior intervention for each user. In addition, the system implementation method of interactive behavior reconstruction is proposed. Petri nets are used to model the business system, and the behavior process reconstruction method under different system behavior profiles is given. On the premise of not destroying the basic business logic of the system, the behavior of legitimate users can change smoothly and has a certain degree of differentiation from the original behavior characteristics. Experiments have proved that in the behavioral fraud detection scenario, the user behavior simulated by the fraudster is invalidated, so that the accuracy and precision of the detection model are improved by more than 10% compared with those in the original scenario. This paper proves the effectiveness of the proposed behavior intervention method and provides a new solution and perspective for fraud detection in identity camouflage scenarios. In the next step, we will continue to focus on how to measure the balance between behavioral drift and good interaction experience, and how to demonstrate the effectiveness of the strategy from a formal perspective.

References

[1] China Information and Communication Research Institute. Mobile Digital Finance and Electronic Commerce Anti-fraud White Paper (in Chinese).

[2] Aiken LR. Attitude and Behavior. Beijing: China Light Industry Press, 2008 (in Chinese).

[3] Nenadic A, Zhang N, Barton S. A security protocol for certified e-goods delivery. Proc. of the Int'l Conf. on Information Technology: Coding and Computing (ITCC 2004). 2004. 22–28.

[4] Zhong J, Yan C, Yu W. *et al*. A kind of identity authentication method based on browsing behaviors. Proc. of the 2014 7th Int'l Symp. on Computational Intelligence and Design. 2014. 279–284.

[5] Zhao P, Yan C, Jiang C. Authenticating Web user's identity through browsing sequences modeling. Proc. of the 2016 IEEE 16th Int'l Conf. on Data Mining Workshops (ICDMW). 2016. 335–342.

[6] Roth J, Liu X, Metaxas D. On continuous user authentication via typing behavior. IEEE Trans. on Image Processing, 2014, 23(10): 4611–4624.

[7] Ma L, Yan C, Zhao P, *et al*. A kind of mouse behavior authentication method on dynamic soft keyboard. Proc. of the 2016 IEEE Int'l Conf. on Systems, Man, and Cybernetics (SMC). IEEE. 2016.

[8] Liu C, He J. Access control to Web pages based on user browsing behavior. Proc. of the 2017 IEEE 9th Int'l Conf. on Communication Software and Networks (ICCSN). 2017. 1016–1020.

[9] Zhao P, Yan C, Jiang C. Authenticating Web user's identity through browsing sequences modeling. Proc. of the 2016 IEEE 16th Int'l Conf. on Data Mining Workshops (ICDMW). 2016. 335–342.

[10] Zheng L, Liu G, Yan C, *et al*. Transaction fraud detection based on total order relation and behavior diversity. IEEE Trans. on Computational Social Systems, 2018: 796–806.

[11] Zheng L, *et al*. A new credit card fraud detecting method based on behavior certificate. Proc. of the 2018 IEEE 15th Int'l Conf. on Networking, Sensing and Control (ICNSC). 2018. 1–6.

[12] Chen D, Ding Z, Yan C, *et al*. A behavioral authentication method for mobile based on browsing behaviors. Proc. of the Institute of Electrical and Electronics Engineers Inc. 2019.

[13] Zhong J, Yan C, Yu W, *et al*. A kind of identity authentication method based on browsing behaviors. Proc. of the 2014 7th Int'l Symp. on Computational Intelligence and Design. 2014. 279–284.

[14] Zhang Y, Chen G. A forensics method of Web browsing behavior based on association rule mining. Proc. of the 2014 2nd Int'l Conf. on Systems and Informatics (ICSAI 2014). 2014. 927–932.

[15] Liu GJ, Jiang CJ. Behavioral equivalence of security-oriented interactive systems. IEICE Trans. on Information and Systems, 2016, E99-D: 2061–2068.

[16] Liu GJ, Jiang CJ. Secure bisimulation for interactive systems. Proc. of the 15th ICA3PP. LNCS 9530. 2015. 625–639.

[17] Zhang Z, Chen L, Liu Q, *et al*. A fraud detection method for low-frequency Trans. IEEE Access, 2020: 25210–25220.

[18] Toledo FPD, Devincenzi S, Kwecko V, *et al*. A framework for modeling persuasive technologies based on the fogg behavior model. Proc. of the 2018 IEEE Frontiers Education Conf. (FIE). 2018. 1–5.

[19] Hamper A, Wendt J, Zagel C, *et al*, Behavior change support for physical activity promotion: A theoretical view on mobile health and fitness applications. Proc. of the 2016 49th Hawaii Int'l Conf. on System Sciences (HICSS). 2016. 3349–3358.

[20] Nishiyama Y, Okoshi T, Yonezawa T, *et al*. Toward health exercise behavior change for teams using lifelog sharing models. IEEE Journal of Biomedical and Health Informatics, 2016, 775–786.

[21] Zhang Z, Arakawa Y, Oinas-kukkonen H. Design of behavior change environment with interactive signage having active talk function. Proc. of the 2019 IEEE Int'l Conf. on Pervasive Computing and Communications Workshops (PerCom Workshops). Kyoto. 2019. 796–801.

[22] Zhang ZH, Cui J. An agile perception method for behavior abnormality large-scale network service systems. Ji Suan Ji Xue Bao/Chinese Journal of Computers, 2017, 40(2): 505–519.

[23] Pan L, Ma B, Wang Y. The similarity calculation of E-commerce user behaviors with Petri net. 2017.

[24] Weidlich M. Behavioural profiles: A relational approach to behaviour consistency. Journal of Biological Chemistry, 2011, 269(36): 22847–22852.

[25] Chen L, Zhang ZH, Liu QW, *et al*. A method for online transaction fraud detection based on individual behavior. Proc. of the ACM Turing Celebration Conf.-China. ACM, 2019. 119.

Xiao Liu, master. His research interest is detection of interaction behavior anomaly.

Ziming Wei, master. His research interest is detection of interaction behavior anomaly.

Zhaohui Zhang, Ph.D., professor, Ph.D. supervisor, CCF professional member. His research interests include big data and behavior analysis.

Pengwei Wang, Ph.D., associate professor, CCF professional member. His research interests include cloud computing, edge computing, service computing, and data mining and analysis.