Journal of Information Processing Vol.31 643–653 (Sep. 2023)

[DOI: 10.2197/ipsjjip.31.643]

Regular Paper

Automating the Detection of Fraudulent Activities in Online Banking Service

Ichiro Asomura[1,a) Ryo Iijima[1,b) Tatsuya Mori[1,2,3,c)

Received: December 12, 2022, Accepted: June 5, 2023

Abstract: Banks offering Online Banking services need to detect and prevent unauthorized electronic funds transfers to reduce financial crime risk. They monitor online banking transaction histories and use their own methods to detect and prevent unauthorized electronic fund transfers. However, unauthorized electronic fund transfers by criminals have not been eliminated. The average rate of false positives in the transaction monitoring systems installed in Japanese banks is up to 99%, indicating that the monitoring systems are not fully functional. Furthermore, the personnel responsible for fraud detection must manually check a large number of false positives, making it difficult for operators to be productive in their assigned tasks. Based on the above background, we develop a method to detect unauthorized electronic fund transfers and suspicious transactions with high accuracy using machine learning algorithms and evaluate its accuracy. Specifically, a supervised machine learning algorithm is applied to detect fraudulent transactions automatically. We evaluated the proposed method on a large set of online banking transaction data provided by a major Japanese bank for the period March 2019 to May 2020. We demonstrated that our approach could detect fraudulent activity with extremely high accuracy; FPR=0.000 and FNR=0.005 can be achieved for a security policy that minimizes false positives.

Keywords: online banking, machine learning, unauthorized electronic fund transfer, financial crime prevention

1. Introduction

In 1995, Wealth Fargo launched its first Online Banking service [1]. Since then, a large number of banks have now offered online banking services. Internet-only banks, which do not have physical branches, have also emerged. The security threats to Online Banking are diverse, ranging from theft of personal assets through malware and phishing to information leakage from cloud storage and other Internet-based services. Attacks targeting Online Banking provide an incentive for attackers to ingest money directly. For this reason, unauthorized electronic funds transfers are becoming increasingly sophisticated. Attackers combine multiple methods to bypass banks' measures to make unauthorized electronic fund transfers.

Banks that provide Online Banking services have developed various security measures, including using multifactor authentication technology and the provision of software to users free of charge to detect and eliminate malware targeting Online Banking. Despite these efforts, the number of unauthorized electronic fund transfers is still increasing as of 2022 [2].

There have been scattered cases of unauthorized electronic fund transfers in Japan's Online Banking system, where criminals have illegally taken authority over authentication information for login and fund transfers. In 2020, the amount of damage in Japan caused by phishing sites amounted to approximately 1,133 million yen in damages, increasing the need to detect unauthorized electronic fund transfers. According to the MUTUAL EVALUATION REPORT OF JAPAN by the Financial Action Task Force on Money Laundering published in August 2021, the average rate of false positives in the transaction monitoring systems installed in Japanese banks is up to 99%. The report points out that the monitoring systems are not fully functional. As a result, fraud detection staff must manually check a large number of false positives, making it difficult to increase their productivity [3].

Some Japanese banks use rule-based monitoring tools to detect unauthorized electronic funds transfers. Operators detect unauthorized electronic fund transfers by manually inspecting the notifications of the monitoring tool. However, the accuracy and reliability of such manual inspections depend on the operator's knowledge and experience, which is a problem that cannot always be expected to produce stable and highly accurate results.

Based on the above background, this study aims to develop a technique to automatically detect fraudulent activity in Online Banking and evaluate its accuracy. We categorize unauthorized activity into two types, unauthorized fund transfers and unauthorized attempts, which we target for detection. This paper defines unauthorized electronic fund transfer as the unauthorized transfer of deposits in a customer's account to an arbitrary account by a criminal. We define an unauthorized attempt as an operation on a bank account that does not result in an unauthorized transfer, but is suspected to be unauthorized by a person who does not have the authority to transfer funds from the customer's account [4], [5].

The approach to detecting unauthorized electronic fund trans-

[1] Waseda University, Shinjuku, Tokyo 169–8555, Japan
[2] NICT, Koganei, Tokyo 184–8795, Japan
[3] RIKEN AIP, Chuo, Tokyo 103–0027, Japan
[a)] asomura@nsl.cs.waseda.ac.jp
[b)] ryo@nsl.cs.waseda.ac.jp
[c)] mori@nsl.cs.waseda.ac.jp

fers has the limitation that detection is possible only after the successful transfer. Therefore, we expect to detect signs of unauthorized electronic fund transfers by adding unauthorized attempts to the detection target.

Even if the system detects unauthorized activity after a successful unauthorized electronic fund transfer, it is possible to take prompt action after the fact. However, we believe that it is desirable to prevent unauthorized fund transfers from occurring in the first place to prevent account holders from becoming victims of unauthorized fund transfers and to prevent the funds from being transferred to criminals.

To achieve the objective above, we created a dataset from transaction data from online banking transactions operated by one of the largest banks in Japan. The dataset covers 2019/03 to 2020/05 and contains 106,406 users, 28,502,613 login records, 1,896,749 remittance records, and 268 features. We show that it is possible to automatically detect fraudulent activity by applying supervised machine learning to the obtained dataset. The technical challenges of this study are as follows. There have been many previous research cases on the detection of fraudulent activity for credit card transactions [6], [7], [8]. These studies employ methods that analyze trends in an account holder's transaction history and detect activity that deviates from transaction trends as anomalies. Credit card transaction statements contain information on the items transacted and many clues to detect anomalies. In contrast, in the case of Online Banking, there is no concept of an item, and the main clues for detecting anomalies are limited to information related to money transfers. Another significant difference is that Online Banking users have a large imbalance in the data, as many have a tiny number of transactions. For this reason, the methods developed in existing research targeting fraudulent credit card activity cannot be applied directly to Online Banking fraudulent activity detection to achieve a high degree of accuracy. Our idea is to detect unauthorized electronic fund transfers based on the transaction history of Online Banking as a whole rather than focusing on the transaction history of each account holder.

The contributions of this work can be summarized as follows:

- Using the transaction logs of an Online Banking service operated by a Japanese bank, we developed a method to detect fraudulent activity automatically.
- Developed technology capable of detecting fraudulent activity regardless of the volume of an account holder's transaction history
- The proposed method achieved extremely high accuracy in fraudulent activity detection; FPR=0.000 and FNR=0.005 can be achieved for a security policy that minimizes false positives.

2. Backgrounds

This section begins with an overview of unauthorized fund transfers in Online Banking, followed by an overview of typical unauthorized fund transfer methods and examples of unauthorized fund transfer detection in Japanese banks.

2.1 Fraudulent Transfers in Online Banking

Online Banking is a generic term for services that use the Internet to conduct banking transactions. Online Banking customers have the advantage of being able to use services such as transfers and balance inquiries at any time without having to worry about bank hours. In addition, customers can use the service from their PCs or smartphones, eliminating the need to visit a bank teller or ATM. Online Banking provides a security mechanism to identify customers accurately. For example, in addition to the password credentials required to access the system, a second password is often used for money transfers. The second password can be a password table or a one-time password. When verifying a customer's identity, the Online Banking system may ask for knowledge known only to the customer. Examples of knowledge known only to the customer include their pet's name, the elementary school they attended, and their mother's maiden name [9].

Unauthorized fund transfer in Online Banking is the unauthorized transfer of deposits from a customer's account to a criminal's account. As shown in the following subsection, there are multiple unauthorized electronic fund transfer methods. Examples of unauthorized electronic fund transfer methods include phishing via email or SMS and malware or other malware programs to steal authentication information for Online Banking. Many depositors have been victims of unauthorized electronic fund transfers through Online Banking.

According to the "Warning against the surge in online banking fraud" by Japan's National Police Agency, there were 1,872 cases of unauthorized electronic fund transfers due to phishing attacks in 2019, with a total loss of approximately 2.521 billion yen [10].

Between 2019 and 2021, there were 4,190 incidents and the amount of damage was approximately 4,474 million yen [10], [11].

In India, the damage caused by unauthorized electronic fund transfers in 2021 was 160 million rupees [12]. In the United Kingdom, the number of unauthorized fund transfers in 2020 was 73,640, and the amount of damage was 197.3 million pounds [13], far exceeding Japan in both the number of cases and the amount of damage. Unauthorized electronic fund transfers stay a significant threat to banks and customers today. Therefore, countermeasures against the threat of unauthorized electronic fund transfers via Online Banking are required.

2.2 Typical Fraudulent Remittance Methods

Unauthorized electronic fund transfers by attackers can be primarily categorized as follows

- Using phishing sites to steal authentication information
- Intercepting and tampering with communications within the browser using malware
- Stealing credentials left in online storage
- Social Engineering by Vishing

We describe each item in detail below.

Generally, phishing sites aim to illegally obtain user account information, such as IDs and passwords, and personal information, such as credit card numbers [14], [15], [16].

Targeting Online Banking services, phishing sites appear to be a banking service operated by a financial institution. Therefore, they are difficult to distinguish from an actual banking service website at a glance.

Criminals have designed some phishing sites to interactively steal credentials from Online Banking and other authentication information, such as one-time passwords or second PINs for unauthorized electronic funds transfers, and to execute unauthorized transfers in real time [17], [18].

The following are steps of a typical unauthorized electronic fund transfer using a phishing site.

(1) Criminals send phishing emails to victims posing as real banks

(2) By clicking on the URL in the phishing email, victims are directed to a fake login page that elaborately mimics the actual login page.

(3) Victims enter their account information on a fake login page and enter the credentials needed to send money on the page for money transfers.

(4) The attacker ingests the account and authentication information entered by the victim and transfers the victim's funds to an arbitrary account on a simple online banking site.

Malware

Criminals develop malware that is used in financial crimes to steal online banking credentials with the ultimate goal of profiting from unauthorized electronic fund transfers. MITB (Man In the Browser) using malware is one of the attack methods used to achieve unauthorized electronic fund transfers [19], [20]. MITB allows an attacker to rewrite the request process for Online Banking from the browser and steal account information or send money to a payee that the customer does not intend. According to the "Threats in Cyberspace in 2019," published by the National Police Agency, malware caused many unauthorized electronic fund transfers in 2016, with 1,876 incidents and approximately 2.91 billion yen in damages. 2016 was the heyday of unauthorized funds transfers using malware [10], [21].

Online Storage

We interviewed the incident response teams of the banks that provided the data of online banking transactions used in this study on the details of unauthorized electronic fund transfers. Our interviews revealed that the victim's smartphone stored the victim's online banking credentials. We also found a case in which the automatic backup function of the victim's smartphone copied the Online Banking information to the online storage. Since users did not use multifactor authentication to access online storage but only passwords, criminals could steal online banking credentials through a simple password attack. There were no traces of malware infection or access history to phishing sites on the victim's computer or smartphone.

Vishing

Voice Phishing (Vishing) is a method of deceiving victims through approaches such as direct phone calls to victims. Vishing in Online Banking is used to obtain user account credentials or to induce the user to open a new account for unauthorized electronic fund transfers [22]. The following is a typical Vishing procedure for Online Banking.

- Scammers call victims to get account numbers and PINs that are not Online Banking numbers.
- Based on the information the fraudster obtains from the victim, the fraudster opens an Online Banking account linked to the victim's account.
- Identification is required for users to activate their Online Banking bank accounts. The scammer contacts the victim and has the victim call a toll-free number to verify his or her identity.
- Fraudsters can log in to Online Banking services registered to the victim's account.

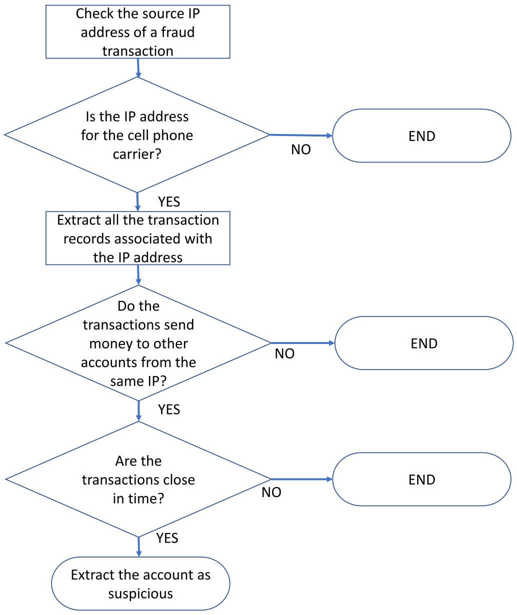
Check the source IP address of a fraud transaction
Is the IP address for the cell phone carrier?
NO
END
YES
Extract all the transaction records associated with the IP address
Do the transactions send money to other accounts from the same IP?
NO
END
YES
Are the transactions close in time?
NO
END
YES
Extract the account as suspicious

Fig. 1 IF-THEN rule for detecting unauthorized debit transfers.

In Japan, there were 16,851 cases of in 2019, with total damage of 31.58 billion yen [11]. Vishing damage is most common in large metropolitan areas. Japan's largest cities, Tokyo, Kanagawa, Chiba, Osaka and Saitama, account for 67% of all Vishing scams. In addition, Vishing victims are concentrated among the elderly (65 years and older), accounting for 78.1% of all vishing fraud victims. There are other methods to trick victims into installing Android malware on their devices by vishing [23]. Additionally, attackers designed malware that allows them to remotely control Android devices for financial fraud. Furthermore, Phishlab reported that the vishing volume increased by 554% in 2021 [24], and vishing is one of the leading methods of financial crime.

2.3 Example of Fraudulent Remittance Detection Operated at an Actual Bank

In the following, we present an example of an unauthorized electronic fund transfer detection method implemented/operated by a Japanese bank and its challenges based on interviews with the bank targeted in this study. That bank used an IF-THEN rule created based on empirical rules to detect unauthorized fraudulent debit transfers and did not adopt machine learning approaches. **Figure 1** shows a flowchart of the IF-THEN rule that was used in the bank for detecting unauthorized debit transfer payment [25]. The rule focuses on the IP address accessed by the victim's account after the fraudulent debit transfer payment occurred. We interviewed the operator who made the rule and they explained to us their rationale for creating that rule. Here is a summary of their insights. When the rule was built, the device and IP address had a specific pattern for the unauthorized debit transfer payments, i.e., they were cell phones and IP addresses were for the mobile carrier. In addition, they found that legitimate customers generally did not access multiple accounts from a single IP address, implying that multiple logins from a single IP address was a sign of fraud transfers. Also, transferring to different customers' accounts in a short period of time was correlated to the fraudulent debit transfer payments. According to the operator, the majority of the fraudulent debit transfer payments detected by the rule were caused by a vulnerability in the identity authentication system. Note that since the bank has fixed the vulnerability, this particular rule is not effective and the operators have built other rules that can detect new patterns of the fraudulent transfers.

The IF-THEN rule to detect unauthorized debit transfers has the following characteristics:

- This detection rule cannot detect the first unauthorized electronic fund transfer.
- IP address information is effective in detecting unauthorized access.
- Detection rules are created based on the operator's personal knowledge and observations.
- The detection rule has a high detection rate for certain unauthorized debit transfer payments (98%), but a low detection rate for other unauthorized transfers.

Challenges

There are two challenges to using a definitive rule like the IF-THEN rule. The first issue is that the detection rate is high for specific patterns of unauthorized fund transfers but not for undefined unauthorized fund transfers. The second issue is that the accuracy of detection of unauthorized electronic fund transfers depends on the operator's knowledge and skill level, since the detection rules are created by relying on the operator's rule of thumb. In addition, when crime trends change, it is necessary to analyze the characteristics of the crimes and create new detection rules.

3. Characteristics of Online Banking Transactions

In this section, we analyze the transaction logs of the online banking services provided by a major Japanese bank to characterize fraudulent activity. We also identify implications for fraudulent activity detection based on the results obtained.

3.1 Labeling of Fraudulent Transfers and Attempts

In addition to actual unauthorized electronic fund transfers, this paper also includes the detection of unauthorized attempts. The banks that provided the dataset considered the following cases as fraudulent attempts.

(1) When a transaction is suspected to be fraudulent, the support center or branch office calls the customer to confirm whether or not the transaction is a normal transaction by the customer himself/herself.

(2) As a result, it was not an unauthorized electronic fund transfer. However, it was not the customer's own manipulation of Online Banking transactions.

The characteristics of the illegal attempts are as follows.

Table 1 Dataset overview.

Data collection period	2019/3/1~2020/5/31
Features	268
Number of users	106,406
Number of logins	28,502,613
Number of domestic remittance	1,896,749

- Accessing one account from multiple IPs.
- Accessing multiple accounts from a single IP or browser.
- Entering the wrong password multiple times.
- The transfer is for a specific amount.
- The transfer is for a specific amount and is the first transaction.

Table 1 gives an overview of the dataset used in this paper. The dataset contains known unauthorized fund transfers and attempts, each labeled by the bank.

3.2 Characteristics of the Transaction Data

In this study, we obtained our data by collecting logs from multiple servers and applications operated by a bank. It should be noted that the data used for our study was formatted by the bank and not obtained in real-time. The purpose of this study was to assess the feasibility of detecting fraudulent transfers from data stored in system logs and databases. If logs containing targeted data for detection can be obtained in real-time, the proposed method can be applied for real-time detection of fraudulent transfers by using a sliding window approach. We analyze online banking transaction logs collected over 14 months from March 2019 to May 2020 at a major Japanese bank. This data set only includes transactions using Online Banking services and excludes ATM and branch transactions. To protect customer privacy, the data set we analyze does not contain personally identifiable information such as names or account numbers. Online banking operators added fraudulent activity labels to transactions tied to unauthorized fund transfers and attempted fraud. The labels for unauthorized activity do not distinguish between unauthorized fund transfers and unauthorized attempts. Machine learning is used to detect fraudulent activity with binary classification [26]. We present the results of visualizing and analyzing the bank's data from Fig. 2 to Fig. 8, representing either authorized or both unauthorized and authorized activities.

Figure 2 and **Fig. 3** show the number of user logins and the number of remittances made through Online Banking, respectively. Both graphs plot the number per day, indicating that the number of logins and the number of remittances tend to decrease on weekends on a weekly basis. The number of logins is hovering around 40,000 per day and the number of remittances hovers around 3,000 per day. Each month, there is a sharp peak in remittance volume in the last week of the month. Sharp peaks in the last week indicate transfers to other bank accounts after payroll transfers.

Figure 4 presents the number of remittance transactions broken down by time of day. Remittance activity is correlated with human activity and is lowest from 1:00 AM to 5:00 AM. Remittance activity is highest in the morning, and the number of remittance transactions tends to decrease from afternoon to night.

Figures 5, **Fig. 6**, **Fig. 7**, and **Fig. 8** present the complementary

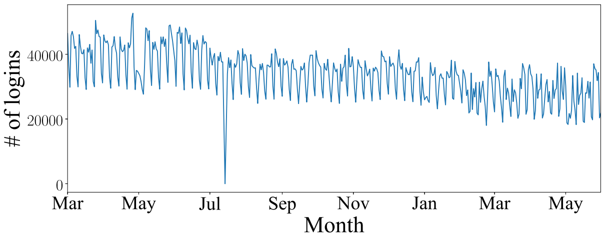

Fig. 2 Number of Online Banking logins for all users (chronological order).

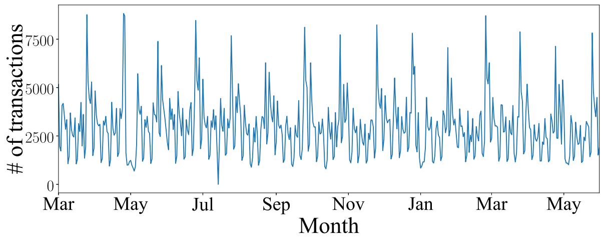

Fig. 3 Number of transfers using Online Banking for all users (chronological order).

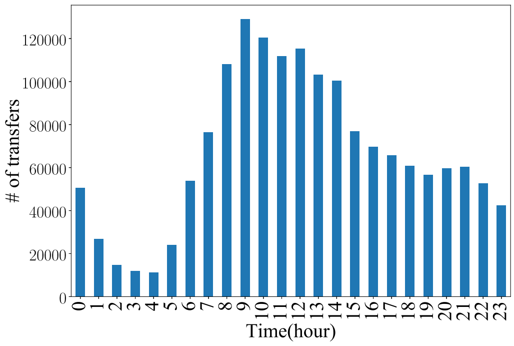

Fig. 4 Number of remittances per hour.

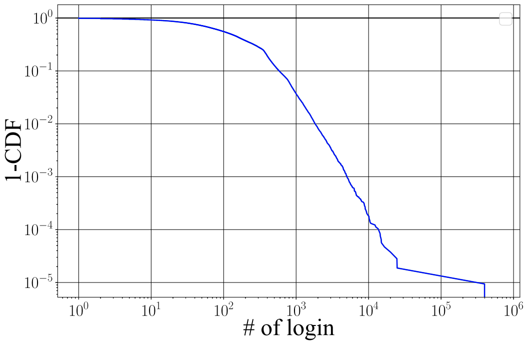

Fig. 5 Log-Log CCDF of the number of logins per user.

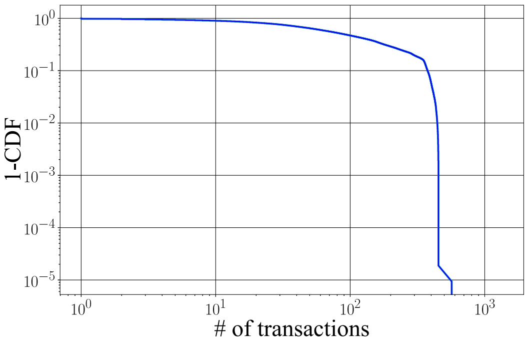

Fig. 6 Log-Log CCDF of the number of transactions per user.

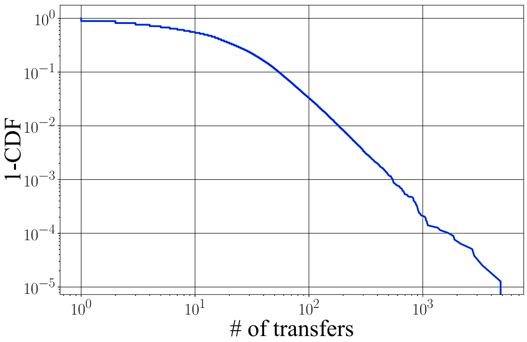

Fig. 7 Log-Log CCDF of the number of transfers per user.

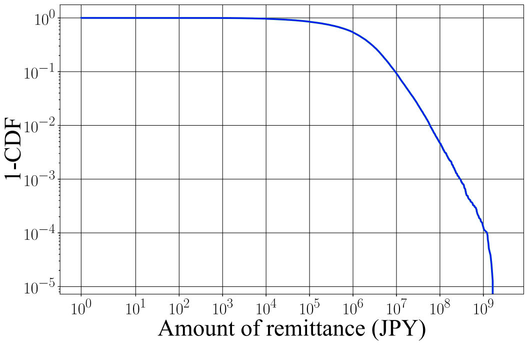

Fig. 8 Log-Log CCDF of the amonut of remittance (JPY) per user.

cumulative distribution (CCDF) of the number of logins, transfers, total transaction value, and total number of transactions per user over the entire period. There is a large bias in the number of user transactions; for example, approximately 60% of the users had fewer than 100 logins, and about 10% of the users had fewer than 10 logins (Fig. 5). Similarly, approximately 60% of the customers transferred less than 10 times (Fig. 7), about half of the customers transferred less than 1 million yen (Fig. 8), and about half of the users had less than 100 transactions (Fig. 6). These facts suggest that most users have few transactions and that an approach that applies anomaly detection to past transaction history is insufficient.

3.3 Implications

In the following, we discuss the implications of fraudulent activity detection methods based on the findings of our measurement study. When creating a model to detect unauthorized electronic fund transfers based on a user's transaction history, it is necessary to create a profile for each user.

A certain amount of transaction history is required to create a profile [6], [7], [8]. If about 100 transactions are needed to create a profile, only half of the users would be applicable, as shown in Fig. 5. The data included in this study are for 14 months, and the conditions would be even more stringent for shorter periods. Therefore, an approach that does not rely on individual transaction history is needed to detect unauthorized electronic fund transfers for users with little transaction history.

Our approach is to create a detection model based on the transaction history of all users of Online Banking.

4. Proposed Method

This section describes our method for detecting fraudulent activity in online banking transaction data using machine learning.

First, we determine the valuable features for applying machine learning models. Next, using the data described in the previous section, we apply several machine learning models and extract the model with the highest detection rate.

4.1 Identifying the Effective Features

The raw data of the online banking transactions we received from the bank contained 268 features. We manually scrutinized these features and found that some of the features did not necessarily contribute to the detection of fraudulent activity. In this study, we extract features most likely to contribute to the detection of unauthorized electronic fund transfers.

Although there are various methods for feature extraction, we adopted an approach that utilizes machine learning models that can output the importance of explanatory variables. Specifically, we selected Random Forests, decision trees, and XGBoost as machine learning models and repeatedly trained and tested these models to exclude features of low importance and features close to the attributes of the label that induce leakage [27], [28], [29]. We obtained the data used for training and evaluation by splitting the data during the training period as shown in **Table 3** (March 1, 2019, to January 15, 2020) at a ratio of 7 : 3 to obtain effective features.

- Split the dataset into training and testing at a ratio of 7 : 3.
- Further, split the training data obtained by dividing the data into a 7 : 3 ratio.
- Train a model with split data for training.
- After training the model, we obtain the feature importance using the test data.

We also verified the learning curve to see if there was any overfitting, which we found to be an excellent indicator of overlearning. As a result, we selected 20 features for our results. (see **Fig. 9**) We interviewed the incident response team of the bank. As a result, we found that unauthorized electronic fund transfers in online banking often involve the use of web browsers with specific language settings. Thus, the rule of thumbs the incident response team has manually discovered also appeared in the extracted features.

4.2 Selection of the Machine-learning Model

We aim to apply several machine learning models to a dataset of unauthorized electronic fund transfers and extract the best machine learning model. The machine learning models we compare are Logistic Regression, k-Nearest Neighbors, Random Forest, Perceptron, Multilayer perceptron, Decision Tree, and XGBoost. We trained and tested these machine learning models on the same data and compared their detection scores, confusion matrix, and learning curves. In the evaluation of the machine learning models, we set the hyperparameters for each model to their default values, as specified by the machine learning framework we used (scikit-learn). It should be noted that, despite our efforts to adjust the hyperparameters, we observed no significant changes in accuracy. We hypothesize that the approach of employing machine learning models for the problems tackled in our research has adapted exceptionally well, resulting in very high accuracy. As a consequence, the impact of the hyperparameters is expected to have diminished. The details of the dataset will be given in the next section. The evaluation procedure for each machine learning model is as follows.

Table 2 Comparison of coefficients of determination (CoD) for different machine learning models.

Models	CoD
Random Forest	0.998
Decision Tree	0.998
XGBoost	0.998
LightGBM	0.996
k-Nearest Neighbors	0.971
Multilayer perceptron	0.914
Logistic Regression	0.909
Perceptron	0.819

We obtained the data used for training and evaluation by splitting the data during the training period in Table 3 (March 1, 2019, to January 15, 2020) at a ratio of 7 : 3.

- Split the dataset into training and testing at a ratio of 7 : 3.
- Further, split the training data obtained by dividing the data into a 7 : 3 ratio.
- Train a model with split data for training.
- After training the model, use the data for testing to obtain the coefficient of determination.

Figure 9 shows the results of the coefficient of determination obtained using the test data after training the model. We chose the Random Forest because of its good coefficient of determination results. Random Forest is a machine learning model that creates many decision trees and employs a majority decision (or average) method [27], [30].

5. Peformance Evaluation

In this section, we demonstrate that our proposed method based on machine learning models is effective for fraud detection. First, we evaluate the general performance of the proposed fraud detection method using the ROC curve and the learning curve (Section 5.1). Next, we evaluate the performance of the fraud detection in specific operational scenarios where false positive or false negative is set to 0. These scenarios are chosen according to the security policy in the banking service and have different advantages in reducing false alarms or not missing frauds, respectively. Finally, we compare the performance of the proposed method with that of the IF-THEN rule.

We obtained the data used for training and evaluation by splitting the data during the training period in Table 3 (March 1, 2019, to May 31, 2020) at a ratio of 7 : 3.

- Split the dataset into training and testing at a ratio of 7 : 3.
- Train a model with split data for training.
- After training the model, use the data for testing for performance.

The hardware specifications and types of software used in our experiment are as follows:

- HW: CPU Intel Core i9 11900 / Memory 128GB+256GB swap / GPU RTX 3080
- SW: Ubuntu22.04 lts / scikit-learn 1.2.1 / CUDA Version: 12.0

The following are the times required for the machine learning model to perform learning/prediction during the experiment.

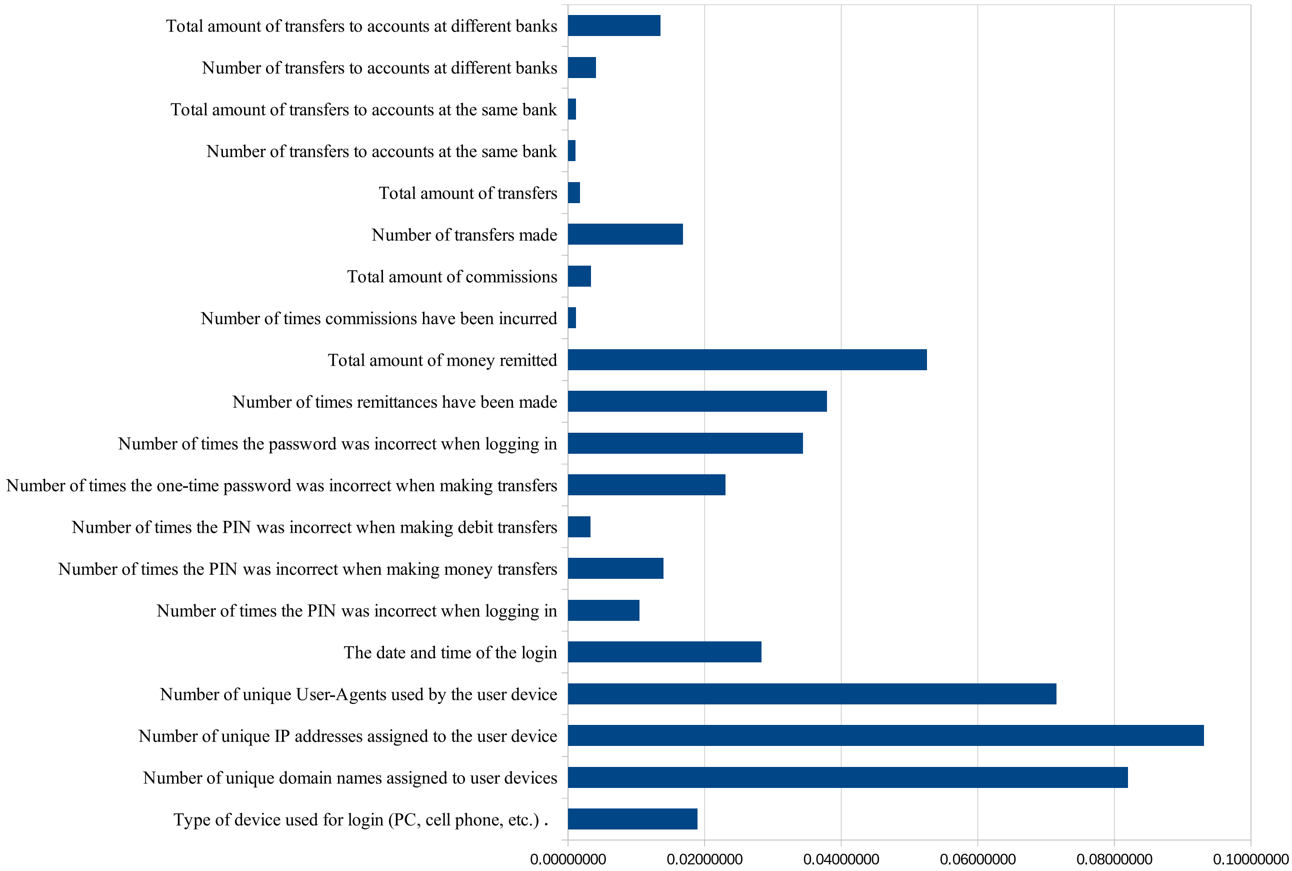

Fig. 9 20 features extracted from the dataset. (All per day)

Table 3 Details of the dataset.

Data	Period	#legit transactions	#fraud transactions
All	2019/3/1–2020/5/31	15,013,620	360,032
Train	2019/3/1–2020/1/15	10,509,534	254,722
Test	2020/1/16–2020/5/31	4,504,086	105,310

- Learning CPU times: 22 min 21 s
- Prediction CPU times: 9.38 s

5.1 Generic Performance

Based on the experimental results of comparing machine learning models shown in Section 4.2, we adopt Random Forest (RF) as our machine learning model. In the following, we evaluate the fraudulent transaction detection performance obtained by applying RF. We split the dataset into a training set and a test set in a 7 : 3 ratio. Table 3 shows a breakdown of the data we used.

Since most bank transactions are normal transactions, there is a problem of label bias [2], [31]. To address label bias, we used SMOTE (Synthetic Minority Oversampling Technique) to train our model after oversampling the data. SMOTE uses the K-nearest neighbors algorithm to increase the minority data, thus balancing the number of labels in the data [31], [32], [33], [34], [35]

We used SMOTE, which works with scikit-learn, to oversample the imbalanced data. The parameters used for SMOTE are as follows. “sampling_strategy=‘auto’, k_neighbors=5, random_state=71”

The data before and after oversampling are as follows.
0: normal electronic fund transfers.

Table 4 The data before and after oversampling

	data before oversampling.
0	10,509,534
1	245,722
	data after oversampling.
0	10,509,534
1	10,509,534

1: unauthorized electronic fund transfers, unauthorized attempts.

Under the setup described above, we evaluated the performance of fraud detection. The results are presented below.

Detection Performance **Fig. 10** is the ROC curve obtained by varying the detection threshold. From the figure, we can see that the proposed method can achieve both low FPR and high TPR, and can achieve extremely high accuracy. In this case, the AUC was 0.998.

Validation of Overlearning Since our approach can achieve extremely high accuracy, there is a risk of overlearning. Therefore, we verify the existence of overlearning by analyzing the learning curve. **Figure 11** and **Fig. 12** present the learning curve for the samples from the period of March 1, 2019 to January 15, 2020 in Table 3.

We can see that both the training score and the test score increase as the number of samples increases and that they converge to high accuracy. In other words, it was shown that overlearning did not occur.

5.2 Performance under the Specific Security Policies

In the following, we study the accuracy of fraudulent transac-

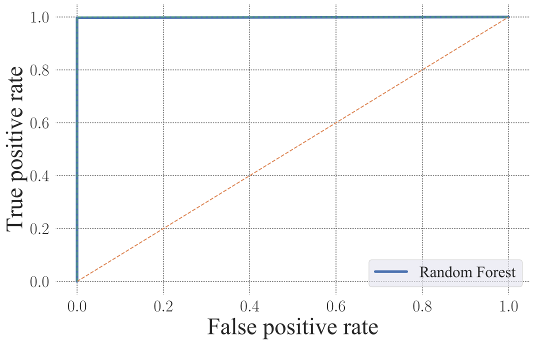

Fig. 10 ROC curve (AUC = 0.998).

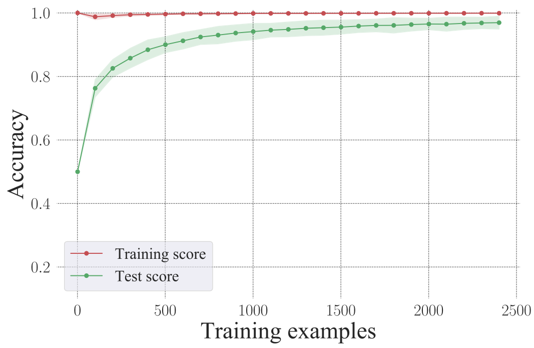

Fig. 11 Learning curve shows the samples ranging from 0 to 2,500

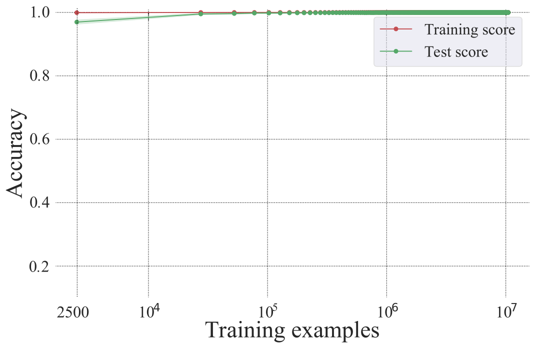

Fig. 12 Learning curve represents the samples from the period of March 1, 2019 to January 15, 2020 in Table 3, and starts from 2,500.

Table 5 Confusion matrix (zero false positive policy).

		Predicted labels	
		Fraud	Legit
Actual labels	Fraud	104,817	493
	Legit	1	4,504,085

Table 6 Confusion matrix (zero false negative policy).

		Predicted labels	
		Fraud	Legit
Actual labels	Fraud	105,310	0
	Legit	2,075	4,502,011

Table 7 Performance comparison between IF-THEN and ML-based method: Policy 1 (zero false positives) and Policy 2 (zero false negatives).

	IF-THEN	Policy 1	Policy 2
# false positives	3	0	22
# false negatives	129	7	0

tion detection when two extreme security policies are imposed.

(Policy 1) Zero false positives

Table 5 presents the confusion matrix of the detection results for the test data when the threshold is set so that the false positives in the training data are zero. Of approximately 4.5 million normal transactions, only one was false positive, but it can be seen that an extremely low false positive was achieved and that the false negative in this case was also low. In fact, the FNR was 0.004. Manual inspection revealed that one false positive was a suspicious transaction, that is, it needed to be checked by the operator. It is useful that the proposed method was able to detect such a case.

(Policy 2) Zero false negatives

Table 6 presents the confusion matrix of the detection results for the test data when the threshold is set so that the false negatives in the training data are zero. In the test data that measured transactions over 14 months, there were zero false negatives. In other words, the machine learning model was able to capture all fraudulent transactions. On the other hand, 2,075 false positives occurred. The FPR at this time was 0.019. This false positive corresponds to 7.4 false alarms per day, and if false alarms of this magnitude are acceptable, it is a practical policy.

5.3 Comparison with the Fraud Detection Method Implemented at the Bank

In the following, we compare the performance of our proposed ML-based approach with the fraud detection method used by the bank to detect fraudulent account transfers. Specifically, we used the IF-THEN rule shown in Fig. 1. The bank used the IF-THEN rule for a certain period of time in the past and the rule was valid for data from May to June 2019. To match the period when the IF-THEN rule was in effect, we train our machine learning model on data from March to April 2019 and test it on data from May to June 2019. Because the IF-THEN rule we are targeting did not include the detection of fraudulent attempts, we excluded fraudulent attempts in the performance comparison experiments.

Table 7 presents the results. Although the IF-THEN rule had low false positives, it had a fairly large number of false negatives (FNR $>$ 0.5), which implies that operators were required to deploy other approaches such as manual inspection, to detect these fraud transactions. In contrast, our ML-based approach successfully kept both false positives and false negatives very low and was able to adjust their balance according to policy.

As mentioned in Section 1, some Japanese banks use rule-based monitoring tools to detect unauthorized electronic fund transfers. However, the accuracy and reliability of manual inspections by operators depend on their knowledge and experience, which may not always produce stable and highly accurate results. As we have demonstrated, our ML-based approach can achieve quite low error rates while successfully automating the process of building the model to detect unauthorized electronic fund transfers.

6. Discussion

This section discusses the limitations, future works, and ethical considerations of this study.

6.1 Limitations

In this study, we evaluated our proposed methodology using a dataset provided by a major Japanese bank. Due to the bank's contract, the data are not publicly available, so other researchers cannot conduct replication experiments using this data set. Although the machine learning model built in this study was effective in detecting fraudulent activity that occurred during the period for which we measured the data, it is not clear whether it will be effective in the future when there are changes in fraudulent transfer methods. To respond to changes in fraudulent remittance methods, it is necessary to continuously analyze transaction data and update features and models that are effective in fraud detection. We believe that the framework and methodology proposed in this study will be useful in such cases.

6.2 Future Research Direcition

In this study, we propose a method to detect fraudulent activity by applying machine learning to bank transaction data and show that it can detect fraudulent remittances with high accuracy. As mentioned above, future changes in fraudulent remittance methods will require new sources of information that are effective in detecting fraud. One promising source of information is access history on online banking services. For example, criminals tend to check certain information, such as balance status and account holder attributes, when making fraudulent transfers. Information on such access history patterns could make fraud detection more robust. The development and evaluation of fraud detection methods using such information sources is a future issue.

6.3 Ethical Considerations

The transaction data we used in this study are privacy protected by removing or anonymizing personally identifiable information by the bank operator who provided us with the data. For example, raw IP addresses are not necessary for counting the number of unique IP addresses assigned to a user's device, which we adopted as an effective feature for detecting fraudulent money transfers. Therefore, we protect user privacy by anonymizing IP addresses to random numbers.

In this study, we present the IF-THEN rule used by a real bank to detect fraudulent transactions and evaluate its performance. In general, disclosing such rules is equivalent to disclosing information that is beneficial to the attacker, and, therefore, appropriate ethical considerations are necessary. The IF-THEN rule we used was created to detect attacks on vulnerabilities targeted by almost all attackers as of spring 2019. Today, the IF-THEN rule has become practically useless and of no value to attackers, as the bank has fixed that vulnerability.

7. Related Work

As mentioned in Section 1, most research papers on fraudulent activity detection have focused on credit card transactions [6], [7], [8]. On the other hand, there are very few research cases on detecting fraudulent activity targeting online banking. To the best of our knowledge, only two studies use authentic transaction logs accumulated in financial institutions to detect fraudulent activity; Carminati et al. and Sonal et al. In the following, we present the summary of each study. Carminati et al. applied an anomaly detection method based on semi-supervised Learning to detect unauthorized electronic fund transfers in online banking and showed that the method could detect unauthorized fund transfers with up to 98% accuracy [36]. They proposed a method to create a profile from each customer's transaction history, calculate the degree of abnormality when the customer performs a new transaction, and detect the high degree of abnormality as fraudulent activity. To calculate the anomaly when a customer makes a new transaction using their proposed method, the amount of transaction history for each customer must be significant. As shown before, in the transaction data we analyzed, 10% of the customers have less than 10 transaction histories, making it difficult to calculate the anomaly level. Sonal et al. compared hidden Markov models, deep learning, and neural networks as methods for detecting fraud in bank transaction logs. They validated them on a data set consisting of 700 samples. The objective of all of these studies was to detect fraudulent activity targeting online banking. The following limitations exist with respect to the data used to evaluate the studies. The data set used by Carminati et al. to detect unauthorized fund transfers records only transactions as expected, with no unauthorized transactions present. They conduct their evaluations by adding artificially created data on fraudulent transactions to a data set that only records transactions as expected. Thus, it is unclear whether unauthorized electronic fund transfers follow a realistic pattern. Sonal et al. use a dataset for their evaluation, but there is a problem in that they do not provide details on what kind of data it is.

8. Summary

This study developed a method to detect fraudulent activity in online banking with high accuracy by applying supervised machine learning. We evaluated the proposed method on a large set of online banking transaction data provided by a major Japanese bank for the period March 2019 to May 2020, and showed that it can detect fraudulent activity with extremely high accuracy. For example, FPR=0.000 and FNR=0.005 can be achieved for a security policy that minimizes false positives. We confirm that the proposed method does not suffer from overlearning. We compared the proposed method with a detection method using an IF-THEN rule created by an operator, which is actually in operation in a bank. The results show that the IF-THEN rule can achieve FPR as low as that of the proposed method, but the FNR is as low as 0.568 and misses many fraudulent activities. Thus, we have successfully demonstrated that the proposed method is practical for the detection of fraudulent activity in real online banking services. Evaluation of the method when the trend of fraudulent activity changes significantly over a long period of operation is a future issue.

References

[1] Wells-Fargo-Bank: First in Online Banking, available from ⟨https://www.wellsfargohistory.com/first-in-online-banking/⟩ (accessed 2022-11-21).

[2] Wei, W., Li, J., Cao, L., Ou, Y. and Chen, J.: Effective detection of sophisticated online banking fraud on extremely imbalanced data, *Springer US World wide web* (*Bussum*), pp.449–475 (2012).

[3] Financial-Services-Agency: Anti-money laundering and counter-terrorist financing measures P.129, available from ⟨https://www.fatf-gafi.org/media/fatf/documents/reports/mer4/Mutual-Evaluation-

Report-Japan-2021.pdf⟩ (accessed 2022-11-21).

[4] Sahin, Y., Bulkan, S. and Duman, E.: A cost-sensitive decision tree approach for fraud detection, *Lecture Notes in Artificial Intelligence*, Vol.6870, pp.5916–5923 (2013).

[5] National-Police-Agency: Securing Safety in Cyberspace, available from ⟨https://www.npa.go.jp/hakusyo/h29/english/p28-29_WHITE_PAPER_2017_E_28.pdf⟩ (accessed 2022-11-21).

[6] Wang, C. and Zhu, H.: Representing Fine-Grained Co-Occurrences for Behavior-Based Fraud Detection in Online Payment Services, *IEEE Trans. Dependable and Secure Computing*, Vol.19, No.1, pp.301–315 (2022).

[7] Shi, E., Niu, Y., Jakobsson, M. and Chow, R.: Implicit Authentication through Learning User Behavior, *ISC 2010, LNCS*, Vol.6531, pp.99–113 (2010).

[8] Malekian, D. and Hashemi, M.R.: An adaptive profile based fraud detection framework for handling concept drift, *10th International ISC Conference on Information Security and Cryptology, ISC*, pp.1–6 (2013).

[9] MUFG-Bank: Personal banking services for residents of Japan, available from ⟨https://www.bk.mufg.jp/global/productsandservices/p_banking.html⟩ (accessed 2022-11-21).

[10] National-Police-Agency: Threats in Cyberspace in 2019, available from ⟨https://www.npa.go.jp/publications/statistics/cybersecurity/data/R01_cyber_jousei_eng.pdf⟩ (accessed 2022-11-21).

[11] National-Police-Agency: Crime Situation in 2020, available from ⟨https://www.npa.go.jp/english/crime_situation_in_2020_en.pdf⟩ (accessed 2022-11-21).

[12] MINT: Govt shares data on online banking fraud and how many cases solved, available from ⟨https://www.livemint.com/news/india/govt-shares-data-on-online-banking-fraud-and-how-many-cases-solved-11660007363092.html⟩ (accessed 2022-11-23).

[13] UKFINANCE: FRAUD - THE FACTS 2021, available from ⟨https://www.ukfinance.org.uk/policy-and-guidance/reports-publications/fraud-facts-2021⟩ (accessed 2022-11-23).

[14] Yunao, Z., Mori, K., Sakurai, Y., Tsubone, M., Iijima, R., Asomura, I., Sakamoto, T., Shimaoka, M. and Mori, T.: Can End-Users Detect a Phishing Site? An Interactive User Behaviour Study Through the Eye Tracking and Semi-Structured Interviews, *IPSJ SIG Technical Report, Information Processing Society of Japan*, pp.1–8 (2020).

[15] Khonji, M., Iraqi, Y. and Jones, A.: Phishing Detection: A Literature Survey, *IEEE Communications Surveys and Tutorials*, pp.2091–212 (2013).

[16] cybersecurity-infrastructure-security agency: Report Phishing Sites, available from ⟨https://www.cisa.gov/uscert/report-phishing⟩ (accessed 2022-11-21).

[17] Indian-Computer-Emergency-Response-Team: Phishing websites hosted on NGROK platform, targeting Indian banking customers, available from ⟨https://www.cert-in.org.in/s2cMainServlet?pageid=PUBADV01&CACODE=CICA-2021-2948⟩ (accessed 2022-11-21).

[18] Business-Standard: A new phishing attack lurking to scam banking customers: Advisory, available from ⟨https://www.business-standard.com/article/finance/a-new-phishing-attack-lurking-to-scam-banking-customers-advisory-121081100780_1.html⟩ (accessed 2022-11-21).

[19] Asomura, I. and Takeda, Y.: Sandbox: Proposal of bootable system snapshot for physical machine (2017).

[20] Rawat, R., Rimal, Y.N., William, P., Dahima, S., Gupta, S. and Sankaran, K.S.: Malware Threat Affecting Financial Organization Analysis Using Machine Learning Approach, *International Journal of Information Technology and Web Engineering*, Vol.17, No.1 (2022).

[21] National-Police-Agency: Crime Situation in 2017, available from ⟨https://www.npa.go.jp/publications/statistics/cybersecurity/data/H29_cyber_jousei_eng.pdf⟩ (accessed 2022-11-21).

[22] imperva: Vishing Attack, available from ⟨https://www.imperva.com/learn/application-security/vishing-attack/⟩ (accessed 2022-11-24).

[23] Lakshmanan, R.: Hackers Using Vishing to Trick Victims into Installing Android Banking Malware, available from ⟨https://thehackernews.com/2022/10/hackers-using-vishing-tactics-to-trick.html?m=1⟩ (accessed 2022-11-24).

[24] LaCour, J.: Vishing Volume Increases 554% in 2021, available from ⟨https://www.phishlabs.com/blog/vishing-volume-increases-554-in-2021/⟩ (accessed 2022-11-24).

[25] Financial-Services-Agency: Anti-Money Laundering, Counter Financing of Terrorism, and Counter-Proliferation Financing P.10, P.11, P.55, available from ⟨https://www.fsa.go.jp/en/news/2022/20221007/20221007.pdf⟩ (accessed 2022-12-07).

[26] Chandola, V., Banerjee, A. and Kumar, V.: Anomaly detection: A survey, *ACM Computing Surveys*, pp.1–58 (2009).

[27] Breiman, L.: Random Forests, *Machine Learning*, Vol.45, pp.5–32 (2001).

[28] Chen, T., Guestrin, C. and Claims, A.I.: XGBoost: A Scalable Tree Boosting System, *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.785–794 (2016).

[29] Jijo, B.T. and Abdulazeez, A.M.: Classification Based on Decision Tree Algorithm for Machine Learning, *The Journal of Applied Science and Technology Trends*, pp.20–28 (2021).

[30] Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S. and Jiang, C.: Random forest for credit card fraud detection, arXiv:1305.1707 (2013).

[31] Devi, D., Biswas, S.K. and Purkayastha, B.: A Cost-sensitive weighted Random Forest Technique for Credit Card Fraud Detection, *2019 10th International Conference on Computing, Communication and Networking Technologies* (*ICCCNT*) (2019).

[32] Longadge, R., Dongre, S. and Malik, L.: Class Imbalance Problem in Data Mining Review, *International Journal of Computer Science and Network* (*IJCSN*) (2013).

[33] Abdallah, A., Maarof, M.A. and Zainal, A.: Fraud detection system: A survey, *Journal of Network and Computer Applications*, pp.90–113 (2016).

[34] Pozzoloa, A., CaelenbYann, O., Borgnea, A., Waterschootb, S. and Bontempia, G.: Learned lessons in credit card fraud detection from a practitioner perspective, *Expert Systems with Applications*, pp.4915–4928 (2014).

[35] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique, *Journal of Artificial Intelligence Research*, pp.4915–4928 (2002).

[36] Carminati, M., Caron, R., Maggi, F., Epifani, I. and Zanero, S.: BankSealer: A decision support system for online banking fraud analysis and investigation, *Elsevier Computers & Security* (2015).

Ichiro Asomura is currently a manager at SKY Perfect JSAT Corporation. He received his B.E. degree from Dokkyo University in 1996 and M.E. degree from Advanced Institute of Industrial Technology in 2016. He is working toward his Ph.D. degree at Waseda University since 2018. His research interests include Financial Crime Prevention and machine learning.

Ryo Iijima received his Ph.D. in engineering from Waseda University in 2022. He is currently a researcher at the National Institute of Advanced Industrial Science and Technology. His research interests include hardware security, sensor security, and signal processing. He received the Best Paper Award at the Computer Security Symposium in 2018. He is a member of the IEEE, AAAS, and IPSJ.

Tatsuya Mori is currently a professor at Waseda University, Tokyo, Japan. He received B.E. and M.E. degrees in applied physics, and Ph.D. degree in information science from the Waseda University, in 1997, 1999 and 2005, respectively. He joined NTT lab in 1999 and moved to Waseda University in 2013. From Mar. 2007 to Mar. 2008, he was a visiting researcher at the University of Wisconsin-Madison. He has engaged in the research of network measurement, security, and privacy. He has received many best paper awards including NDSS 2020 and EuroUSEC 2021. Dr. Mori is a member of ACM, IEEE, IEICE, and IPSJ.