International Journal of Security and Its Applications
Vol. 10, No. 2 (2016), pp.247-254
http://dx.doi.org/10.14257/ijsia.2016.10.2.22

User Action-based Financial Fraud Detection Method by SVDD

Seong-Ho An[1], *Kihyo Nam[2], Mun-Kweon Jeong[2], Yong-Rak Choi[1]

[1]*Dept of Computer Engineering, Daejeon University, Yongun-dong, Dong-gu, Daejeon, Republic of Korea*
seongho.an83@gmail.com, yrchoi@dju.ac.kr
[2]*UMLogics Co., Ltd., 17, Techno 2-ro, Yuseong-gu, Daejeon, Republic of Korea*
{nkh, jmk}@umlogics.com

Abstract

This thesis proposes a method to detect sophisticated electronic financial frauds using SVDD. The financial industry detects electronic financial frauds using FDS, but its false positive rate is high enough to require additional authentications. It causes customers inconveniences and does not detect those sophisticated financial frauds. In order to resolve the aforementioned issues, this study proposes a method to detect such potential frauds by profiling and vectorizing user activities and device information by SVDD.

Keywords*: FDS, SVDD*

1. Introduction

As the use of smartphones and FinTech has become commonplace, the emphasis on electronic financial transaction has gained momentum. While growth on FinTech has made financial transaction much easier, frauds of electronic financial transaction have increased and become sophisticated. Certificates, passwords, and unique numbers on bank secret cards used for electronic finance can be stolen from smartphones or personal computers, and agencies that handle customers' personal information may fail to protect it. It has also made electronic financial frauds complicated and advanced. As the number of cases affected by such sophisticated frauds has steadily increased, existing measures like keyboard security, certificates, and additional passwords are not enough to deal with financial frauds. [1]

In order to handle increasing electronic financial frauds, Korean Financial Supervisory Service encourages the financial industry to set up FDS (Fraud Detection System) to protect customers and prevent any possible incidents of electronic financial transaction. Credit card companies set up FDS, and it has turned out to be effective. So Korean Financial Supervisory Service required the financial industry to set up FDS in 2014. [3]

Analysis ability to use inside knowledge is absolutely needed to effectively operate FDS, but Korean banks still have few operation experiences and have not accumulated enough data to detect electronic financial frauds. It leads to a high rate of false positive, more customer inconveniences, and more customer complaints. Customers tend to bear more rigorous customer verification including additional authentications or outbound calls due to those false positives.

This thesis designs and vectorizes user profiles of financial activities to detect sophisticated financial frauds and suggests a user action-based financial fraud detection method to predict and detect such frauds using SVDD.

* *Corresponding author*

ISSN: 1738-9976 IJSIA

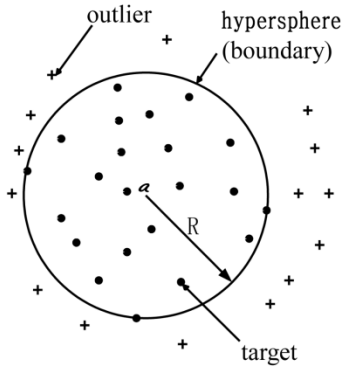
2. Data Profiling of Financial Transaction

2.1 Categorizing Detection of Existing Financial Frauds

Study on detection of financial frauds can be categorized into statistics analysis-based detection method and artificial intelligence detection method. Artificial intelligence detection method can be classified into data mining, pattern recognition, and machine learning. The detection method using data mining sorts or assembles data, and then automatically breaks it down into invalid transaction and valid transaction. [4] The method using pattern recognition and machine learning identifies frauds (i.e. invalid transactions) by figuring out characteristics of valid transactions and invalid ones based on transaction information. Leading methods are ones using artificial neural network and Bayesian modeling. [5]

SVDD, a method using data mining, is a one-class classification to identify invalid vectors in a specified vector group, first proposed by David [6]. This method maps data into higher dimension vector space through nonlinear transformation and produces the one-class classification of the same characteristics by producing boundaries of a hypersphere with the minimum radius that contains all of the mapped data. SVDD searches for a hypersphere and maps data in the hypersphere, so it is important to create an optimized hypersphere that contains most mapped data with the minimum radius. SVDD should create a hypersphere with its center a and radius $\mathcal{R}$. Suppose the researchers have a hypersphere with the set D, the hypersphere needs to be as small as possible and also needs to contain as many mapped data as possible. Once the hypersphere is formed through the mapping, SVDD rules data outside the hypersphere as invalid.

Figure 1. Basic Concept of SVDD



2.2 Profiling to Detect Financial Frauds

The researchers produced a hypersphere of its center a and radius $\mathcal{R}$ with a profiled data set of valid financial transactions through SVDD learning and identify invalid transactions by profiling subsequent transactions through SVDD. For this, they extracted feature values by profiling information from financial transactions by each user and create vectors for learning.

In previous researches, the researchers came up with customers' pattern information and profile information from electronic financial incidents in a Korean bank and set a detection rule based on the output. [7] They measured a starting coordinate, an ending coordinate, an inclination of fingers, and a scroll speed on the touchscreen of a smartphone and created a detection rule by applying them to a data mining algorithm. [8] In this thesis, we have the Table I that contains information regarding financial

transactions and setting information like device information based on the data used in the previous researches. The Table I also includes profiling information to handle potential invalid transactions. Data from the transaction information, device information, and activity information is used as feature values.

Table 1. Financial Information Profile by Type

Type	Variable	Remarks
Transaction information	Start_time	Start time of transfer
	Duration	Duration of transfer
	Bank_code	Bank code of transfer
	Bank_account	Whether the account has been used before
	Telephone_authentication	Whether the telephone authentication is successful per security rules
	Transfer_amount	Whether the amount is in the range of previous transfer amounts
Device information	Device_information	Unique information of the device
	OS	For PC or smartphone
	IP_address	Location information of the device
	MAC_address	Unique information of the device
	Country_code	Area code where the transaction took place
	Proxy_IP_address	Whether a proxy has been used
Activity information	Movement_speed	Usual user behavior
	Keyboard_typing_speed	Usual user behavior
	Failed_password_attempts	Determine invalidity regarding consecutive failed attempts
	Secret_card	Determine invalidity regarding consecutive failed attempts
	Certificate_used	User behavior
	IP_PW	User behavior

3. Detecting Financial Frauds using SVDD

3.1 Steps to Detect Financial Frauds

(Step 1) Creating a user profile vector

Compile n sets of profile information of legitimate financial transactions based on Table 1. Produce one vector X that includes n numbers converted from the n sets of profile information.

$$\mathbb{X} = \{x_1, x_2, x_3, \cdots, x_n\}$$

Repeat the step κ times and produce a vector group $\mathcal{D}$ that contains κ $\mathbb{X}$ vectors of legitimate financial transactions.

$$\mathcal{D} = \{\mathbb{X}_1, \mathbb{X}_2, \mathbb{X}_3, \cdots, \mathbb{X}_\kappa\}$$

(Step 2) Learning to define an area of legitimate financial transactions

Produce an optimized hypersphere by running the SVDD algorithm on the vector group $\mathcal{D}$ from the previous step. The results, center a and radius $\mathcal{R}$, are bases to detect financial frauds.

(Step 3) Detecting financial frauds

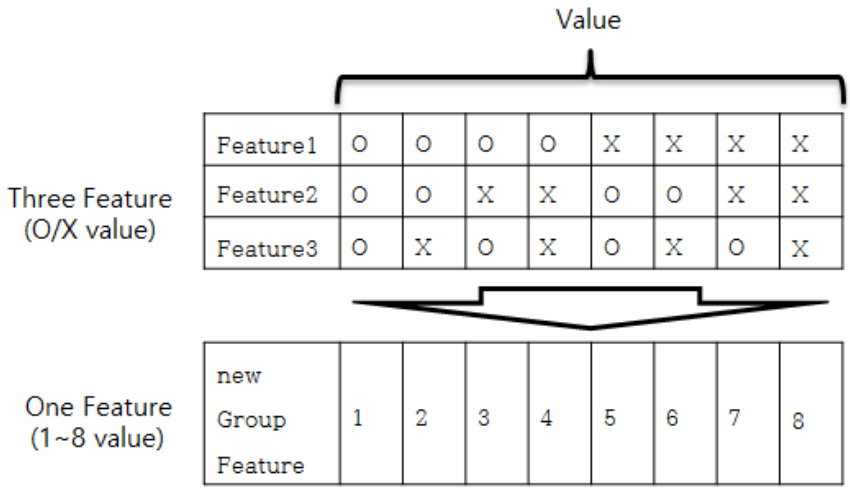
The researchers can determine financial frauds by running the SVDD algorithm the new vector z with the existing vector group $\mathcal{D}$. If the vector z satisfies the SVDD formula below, they can rule that the vector is of legitimate financial transactions. Otherwise it is of financial frauds.

$$\| z - a \|^2 = K(z,z) - 2\sum_i a_i K(z, d_i) + \sum_{i,j} a_i a_j K(d_i, d_j) \leq \mathcal{R}^2$$

3.2 Optimization of Feature

SVDD is highly accurate, but the higher the number of Feature are available, the longer SVDD runs to completion. Since financial transactions take place in real time, a high number of Feature makes it difficult to adopt the SVDD-based system. To resolve this issue, the researchers can handle possible diminished performance by creating new Group Features using the "Feature Scaling" method shown in Fig 2. Fig 2 depicts three Features with 8 O/X values each join to become one Feature with a scale value of 1~8.

Figure 2. Feature Scaling Method



4. Evaluation of SVDD-based FDS

4.1. Evaluation Criteria

4.1.1 Evaluation method

In order to evaluate research materials in this thesis, the researchers created a hypersphere with center a and radius $\mathcal{R}$ by running the SVDD algorithm on each user's profile data of legitimate financial transactions. Then they measured a ratio using n vectors of legitimate ones and m vectors of frauds.

4.1.2 Evaluation data

For evaluation purpose, the researchers used 500 vectors of legitimate financial transactions to be "trained" and 100 other vectors to evaluate.

The researchers created 500 vectors of legitimate financial transactions based on Table I. They created a hypersphere with center a and radius $\mathcal{R}$ by running the SVDD algorithm on a vector group $\mathcal{D}$ of those 500 vectors.

The researchers created 100 vectors for evaluation using the same method.

Jae Hoon Park's thesis mentions signs of financial frauds as shown in Table II[7]. The table is a result of analysis of patterns of actual 500 financial frauds in the bank A since 2013.

Table 2. Pattern of FDS Construction

Classification	Item	Analysis
Transaction period of attacked users	Transaction Time	Midnight transaction (0 am ~ 4am)
	Initial / Final Transaction	Deviation from normal transaction period
Mediums	New medium	Access with new medium
	Number of mediums	Using 2 or more mediums for attack
	Local	Access from outside of usual local
Daily Transaction to other banks	Daily Transaction frequency	Exceeding daily transaction frequency limit
	Daily Transaction Amount	Exceeding daily transaction Amount limit
Remittance bank	Initial remittance bank	Transfer to unprecedented bank(more than 300,000 KRW)
Attacked Saving Account	Withdrawal account balance	Withdrawal minimum balance of Savings Account

The researchers created a vector $\mathbb{X}$ of financial frauds using Table II and normal data $\mathcal{D}$. They created 100 vectors of financial frauds by adding alterations of 10%, 20%, and 30%, respectively, to applicable feature values.

The researchers repeated the procedure 5 times and produced profile data for 5 users.

4.2. Evaluation Result

In this paper, the researchers used LIBSVM to empirically evaluate detection of fraud detection through the SVDD algorithm after they obtained feature values through user profiling and vectorize them.

The researchers created a hypersphere with center a and radius $\mathcal{R}$ using 100 normal vectors of legitimate transactions and 100 fraud vectors. They deduced the results in Table III by running the SVDD algorithm with the hypersphere.

Table 3. Detection Accuracy of the Test Results

	Legitimate transaction Detection accuracy (%)	Fraud Detection accuracy (%) [10% alteration]	Fraud Detection accuracy (%) [20% alteration]	Fraud Detection accuracy (%) [30% alteration]
1st run	98	88	92	95
2nd run	98	92	94	98
3rd run	97	85	92	96
4th run	99	85	93	98
5th run	96	87	93	96
Average	97.6	87.4	92.8	96.6

Detection rate of legitimate transactions through learning is 97.6%, and detection rates for frauds are 87.4%, 92.8%, and 96.6%, respectively, depending on alteration percentage.

5. Conclusion

This thesis has described the user action-based financial fraud detection method by SVDD to detect potentially sophisticated financial frauds. The researchers have gone over other existing financial fraud detection methods and proposed a detection method to make patterns of user activities using SVDD derived from data mining. To detect users' financial frauds, the method performs profiling of users' devices and activities and vectorize them. Then it trains them using SVDD to detect financial frauds.

The analysis of research results is to create an optimized hypersphere based on the vector group $\mathcal{D}$ of profile of legitimate transactions and SVDD. Then the researchers measured a ratio using n vectors of legitimate ones and m vectors of frauds. For evaluation purpose, they used 500 vectors of legitimate financial transactions to be "trained" and 100 other vectors to evaluate and ran LIBSVM to execute the evaluation. Detection rate of legitimate transactions through learning is 97.6%, and detection rates for frauds are 87.4%, 92.8%, and 96.6%, respectively, depending on alteration percentage.

In this thesis, the researchers used a limited number of Features, so learning data may not be enough. If they can apply user information, device information, and activity information from actual financial transactions to create Features, we can use those Features to create an optimized hypersphere. Such a hypersphere can help us detect financial frauds more accurately.

Acknowledgements

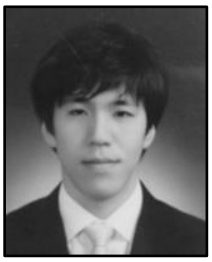
This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No. R-20150521-001431, Fraud Detection System development by analysis of user action pattern in various web environment)

References

[1] Financial supervisory service, "Electronic Fraud Prevention Services.", **(2013)**.
[2] Financial Security Researchers, "Fraud Detection System Technical Guide", **(2014)**.

[3] S. H. Jeong, "A Survey of Fraud Detection Research based on Transaction Analysis and Data Mining Technique", Jounal of The Korea Institute of Information Security & Cryptology, vol.25, no.6, **(2015)**.
[4] P.K Chan, "Distributed data mining in credit card fraud detection", Intelligent Systems and their Applications, IEEE, vol. 14, no. 6, **(2002)**.
[5] C. Phua, "A Comprehensive Survey of Data Mining-based Fraud Detection Research", Intelligent Computation Technology and Automation (ICICTA), **(2010)**.
[6] T. David and D. Robert, "Support vector data description", Machine Learning, vol. 54, no. 1, **(2004)**, pp. 45-66.
[7] J. H. Park, "Effective Normalization Method for Fraud Detection Using a Decision Tree", Jounal of the Korea Institute of Information Security & Cryptology, vol.25, no.1, **(2015)**.
[8] H. Y. Min, "Outlier Detection Method for Mobile Banking with User Input Pattern and E-finance Transaction Patter", Journal of Korean Society for Internet Information, vol.15, no. 1, **(2014)**, pp. 157-170,.

Authors



Seong-Ho An,
Master of Science in Computer Engineering
Daejeon University
Republic of Korea

Kihyo Nam
Ph.D of Industrial Engineering(Korea University)
CISA (Certified Information Systems Auditor)
CISSP (Certified Information Systems Security Professional)
Adjunct Professor (Konkuk University)
Republic of Korea

Mun-Kweon Jeong
Master of Information Industrial Engineering
(Chungbuk University)
Director-General of UMLogics Co., Ltd.
Republic of Korea

Yong-Rak Choi
Professor of Computer Engineering (Daejeon University)
Ph.D of Computer Science (Chung-Ang University)
Republic of Korea