International Journal of Computer Applications (0975 - 8887)
Volume 186 - No.74, March 2025

Detection and Control of Credit Card Fraud Attacks in Sliding Window with Exponential Forgetting

Alexander Stotsky
Data Science and AI
Department of Computer Science and Engineering
Chalmers University of Technology
Gothenburg SE - 412 96, Sweden

ABSTRACT

Credit card fraud causes significant financial losses and frequently occurs as fraud attack, defined as short-term sequence of fraudulent transactions associated with high transaction rates and amounts, business areas historically tied to fraud, unusual transaction times and locations and different types of errors.
Confidence interval method in the moving window with exponential forgetting is proposed in this paper which allows to capture recent changes in the shopping behaviour of the cardholder, detect fraudulent amounts and mitigate the attack. Fraud risk scoring method is used for estimation of the intensity of the fraudulent activity via monitoring of the transaction rates, merchant category codes, times and some other factors for detection of the start of the attack.
The development and verification are based on detailed analysis of the transaction patterns from the dataset, which represents an extensive collection of around 24.4 million credit card transactions from IBM financial database. Recommendations for further development of the detection techniques are also presented.

Keywords

credit card fraud attacks; time series analysis; detection & control in moving window with exponential forgetting; fraud risk scoring; monitoring of intensity of fraudulent activity; mitigation of fraud attack

1. CHALLENGES, PREVIOUS WORKS & NEW METHODS

Automation of the payment processes in e-commerce reduces economic costs but requires performance and robustness improvement of the real-time fraud detection systems especially in the event of a crisis or war, [1]. The achievement of the trade-off between fraud detection accuracy and minimization of false alarms (which is expensive and results in customer dissatisfaction) is the main challenge in real-time fraud detection systems.
Assuming that fraudulent transactions follow similar patterns the pattern recognition systems such as support vector machines, artificial neural networks, naive Bayesian networks, K - nearest neighbour methods, hidden markov models, fuzzy logic based systems and many other techniques can be applied for classification of transactions as fraudulent, see for example [2] -[4] and references therein. These methods do not take into account dynamics of fraud attacks (transient of fraudulent transactions) and dynamics of shopping behaviour of individual users and based on average quantities of the large number of transactions, which results in detection of almost every transaction as non-fraudulent due to overtraining and wrong correlations, [3].
Time series analysis where each discrete step is associated with the transaction number, which allows identification of the dynamic spending behavior of individual cardholder and flagging transactions that differ from that behavior as outliers shows better outcomes, see Figure 1.
The aim of fraudsters, scammers and cybercriminals is money and their behavior is associated with achievement of the highest benefits under time pressure. Maximization of the returns results in sudden increases in fraudulent activity, which can be defined as fraud attacks. Data analysis results in the following operational definition of the fraud attack.
Fraud attack is the sequence of fraudulent transactions (that can be conceptualized/framed as a dynamic system) which occur within the relatively short time interval (see Figure 2), usually initiated by automated software tools (often coming from the same IP address) associated with high transaction rates, typical merchant category codes (MCC) (business areas traditionally related to fraudulent activities, see Figure 3), unusual transaction times and locations (often with geographical mismatches, see Figure 1), technical glitches, insufficient balance, errors in CVV, PIN, zip codes and some other abnormal authorization errors.
Outlier detection in time series with adjustable deviation (which takes into account the factors described above) from the exponential moving average is the most efficient and implementable way for detection and control of fraud attack, see Figure 2 and Figure 4.
Notice that the values of purchases in time series could be large for maximization of the outcome or low aiming for disguising fraud. Therefore the beginning of attack is usually associated with the sequence of low-value transactions (which makes the detection difficult), but the attack includes always the sudden changes in spending behaviour associated with high transaction amounts, see Figure 2 and Figure 4. Detection of fraud attack can be more efficient than the detection of single fraudulent transaction due to availability of information about typical patterns/trends associated with elapsed time between transactions, merchant category codes, amounts, times, locations etc. Early detection of fraud attack that can be accomplished

using these patterns is essential for minimizing financial losses and maintaining trust of the customer.

Identification and description of the purchasing profile of individual cardholder is absolutely necessary for improvement of the detection accuracy. Moreover, accurate detection is impossible without accounting for variability in purchasing behaviour of individual cardholder. The detection system should be adaptive taking into account recent changes in purchasing behaviour, which significantly reduces false alarms. Notice that inclusion of the adaptive features into classical detection techniques described above could pose considerable difficulties.

Shopping Profile in Sliding Window. Recent changes in the shopping profile of the cardholder can be captured in a window of a certain size which is moving in time. In addition to instantaneous forgetting of the shopping behaviour with moving window the exponential forgetting (where the weights decrease exponentially for more remote transactions) is introduced inside of the window. In other words, combination of exponential forgetting and windowing techniques (where the forgetting factor and the window size are two adjustable parameters) opens new opportunities for performance improvement in time series analysis associated with better trade-off achievements in the case of rapid changes in the shopping profile, [5]. This results in significant improvement of the detection accuracy of fraud attacks.

For illustration of the idea consider one step of the moving window of a size w which represents time evolution of the transaction amount:

$$\overbrace{y_{k-w+1}\ \lambda^{w-1}, \ldots\ldots,\ y_{k-1}\ \lambda,\ y_k}^{\text{window size} = w} \qquad \text{step k}$$

$$y_{k-w+2}\ \lambda^{w-1},\ \ldots\ldots,\qquad y_k\ \lambda,\ y_{k+1} \qquad \text{step k+1}$$

where recent shopping amount y_{k+1} enters the window and the discounted amount $y_{k-w+1}\ \lambda^{w-1}$ leaves the window, $k \geq w$. The static weighting sequence $1, \lambda, ..., \lambda^{w-1}$ discounts exponentially past shopping amounts in favour of more recent ones with the forgetting factor $0 < \lambda \leq 1$ inside of the sliding window. The window size and the forgetting factor are two adjustable parameters which allow to capture continuously changing shopping profile of a large number of cardholders. For example, the time series in the short window with rapid forgetting is able to capture fast deviations in shopping behaviour and to identify fraudulent attempts rapidly.

Notice that the detection accuracy in sliding window of individual cardholder is much higher compared to the methods based on average quantities of the large number of transactions.

The window size and the forgetting factor can be adjusted in real-time that changes the profile of the confidence interval (see Section 2) for avoidance of misdetection and improvement of the detection accuracy.

For example, consumer spending significantly increases on payday followed by brief surge in spending afterwards, which requires reduction of the window size and forgetting factor for accurate estimation of the shopping profile. The window size and forgetting factor can even be pre-calibrated provided that the payday of the cardholder is known.

Notice that the most pronounced spikes in spending around payday are associated with certain business areas such as restaurants, supermarkets, gas stations, beauty shops, etc. These areas are very well correlated with the frequencies of merchant category codes for legitimate transactions, see Figure 3(b). Notice also that fraudsters may use sharp rise in spending around payday to conceal the fraud attack. MCC signatures associated with fraudulent transactions, see Figure 3(a) can be used for detection of such attacks.

Detection of Fraud Attacks. Data analysis shows that fraudulent activities usually occur as fraud attacks. The shopping amounts are often not large in the beginning of the fraud attacks, which makes the detection difficult. However, detection of fraud attack can be more efficient than the detection of single fraudulent transaction due to availability of information about patterns of the attacks. Therefore the fraud detection system should be based on combination of different features which are able to correctly identify different parts of the fraud attack. The purchased amount should play the most important role in the detection of the attack. Taking into account that the amounts may be not large in the beginning of the attack the inter-transaction time gaps, merchant category codes, transaction times and locations, insufficient balance, technical glitches, errors in CVV, PIN, zip codes and other abnormal authorization errors should be used for identification of the start of the attack. The shopping amounts can then be controlled using confidence interval method minimizing losses.

2. CONFIDENCE INTERVAL FOR DETECTION & CONTROL

Outlier detection can be performed in moving window and associated with statistical hypothesis, where weighted average is compared to suspected outlier within the two sample t-test framework, [6]. Namely, the hypothesis that the mean value is equal to the outlier is taken as a null hypothesis, which is tested against the alternative hypothesis which indicates outlier in each step of sliding window. Such formulation gives the basis for fraud detection via the confidence interval method. This confidence interval has two end points in each step of sliding window, which determine the fraud detection thresholds. The endpoints are usually determined by the standard deviation, sample size (window size) and critical value taken from the Student distribution look-up table, under the assumption of normal distribution of the variable.

Due to the significant differences and volatility in the shopping behavior of individual cardholders, assumptions about the distributions within the window cannot be made, and a problem-oriented decision regarding the length of the confidence interval as the detection threshold is needed. In this particular problem the choice of the detection threshold is associated with the trade-off between fraud detection accuracy and misdetection. The amounts associated with current shopping behaviour should be located inside of the interval and all fraudulent amounts should exceed the threshold. It is clear that such trade-off is not achievable without real-time adjustments of the window size, the forgetting factor and the length of the interval, which is defined as follows:

$$\overline{y}_{w,\lambda,k}\ -\ c_k\ s_{w,\lambda,k} < a_k < \overline{y}_{w,\lambda,k}\ +\ c_k\ s_{w,\lambda,k} \qquad (1)$$

where a_k is the amount, $\overline{y}_{w,\lambda,k}$ is weighted mean value (exponential moving average) and $s_{w,\lambda,k}$ is weighted standard deviation, k is the step number. The amount located outside of the confidence interval (1) can be flagged as the fraudulent amount. The variable $c_k = c_k(\cdot)$ (which can be different for positive and negative amounts) determines the length of the interval and the window size and forgetting factor can be adjusted in real-time to avoid misdetection and to improve detection accuracy in the case of changes in shopping behaviour.

In the case of normal distribution of the transaction amounts the confidence interval contains 99.73% of the values within three standard deviations of the mean, $c = 3$ provided that $\lambda = 1$. The amounts outside the 3σ interval are inconsistent with the normal distribution associated with current shopping behaviour of the cardholder and can be identified as fraudulent. Control of fraud attack can be per-

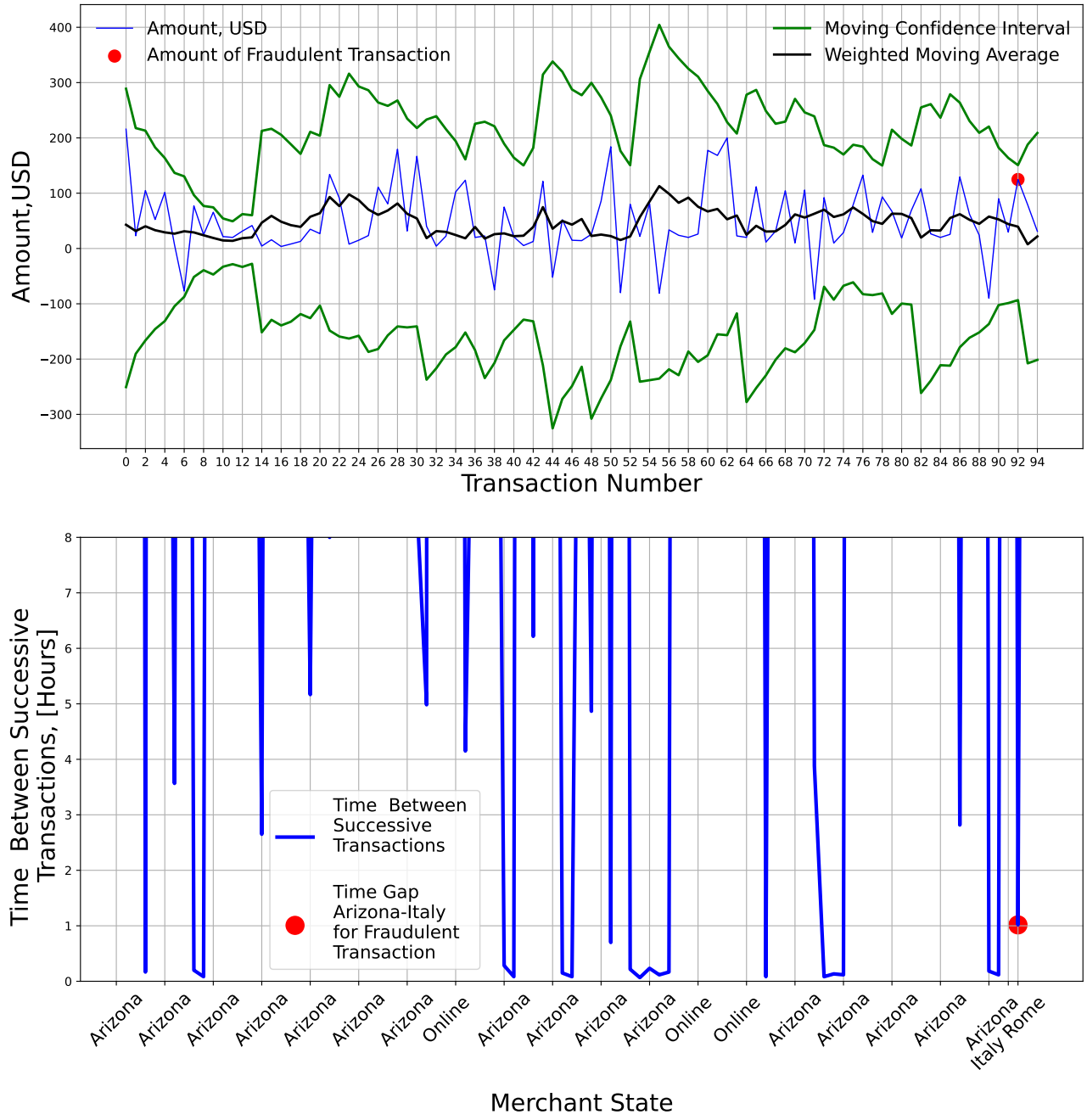


Fig. 1. Detection of single fraudulent transaction in the medium-sized window, $w = 8$ and rapid forgetting with $\lambda = 0.8$ is presented in the first plot of the Figure. The shopping amount is plotted with the blue line, exponential moving average is plotted with the black line, confidence interval is plotted with green lines and fraudulent amount is plotted with red round sign. Geographical information and inter-transaction time gaps are presented in the second plot of the Figure. The time gap between two successive transactions from Arizona and Italy (which does not align with usual locations) identifies the transaction from Rome as fraudulent transaction.

formed by adjusting the length of the interval via variable c_k.
The model (1) is self-learning (due to the sliding window and forgetting) which constantly updates the window and gives more weight to recent data. Calculation of the weighted mean and standard deviation in each step accurately describes current shopping profile of the cardholder. The window size w should be sufficiently small in order to capture fast changes in shopping behaviour that limits application of different types of models. The accuracy of autoregres-

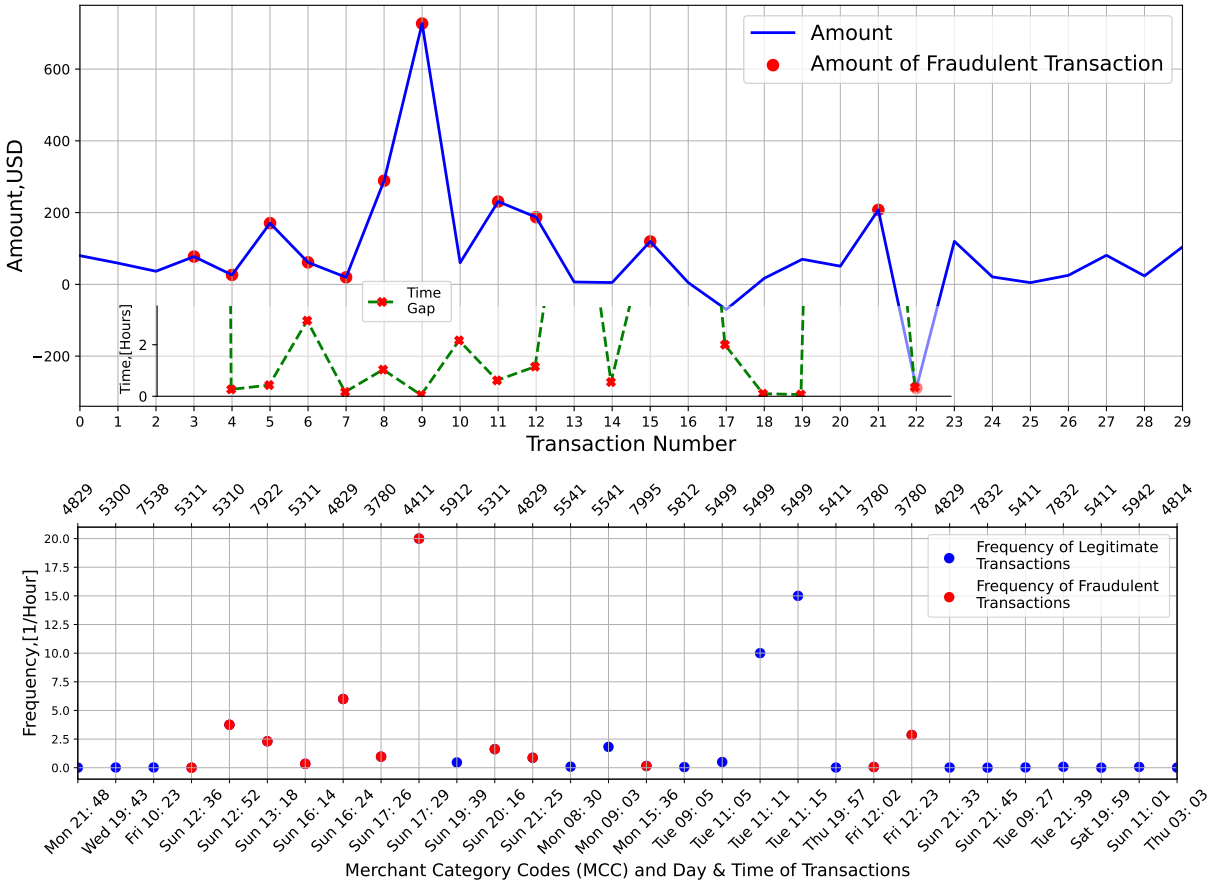


Fig. 2. The fraud attack (defined as the sequence of fraudulent transactions) is plotted in the first plot. Shopping amount in US dollars is plotted with blue line. Fraudulent transactions are indicated with red round signs. Inter-transaction time gaps (in hours) for fraudulent transactions are plotted with cross signs of a red color in the subplot. Fraud attack is characterized by high transaction rate (calculated as inverse of elapsed time) identified by small time gap between transactions. Fraud attack was performed as the sequence of online transactions with well established signature of merchant category codes (MCC), see the second plot and Figure 3, technical glitch and CVV error on Sunday (at unusual time). Small time between successive transactions, MCC, technical glitch and unusual transaction time flagged the transactions 3 and 4 as fraudulent and associated with the start of the attack. Early detection of the start of the attack is essential for mitigation of the financial losses.

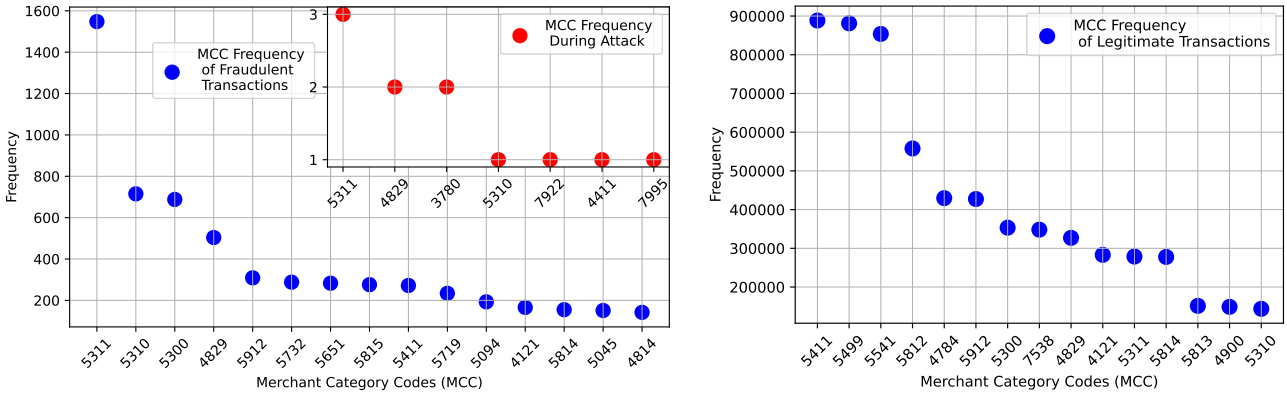
sive moving average (ARMA) models, neural network models and some others is not sufficient for reliable detection of the outliers due to insufficient amount of data available inside of short windows, [7]. Confidence interval method based on the exponential moving average in short sliding window with prioritization of recent transactions is the only method for robust detection of the outliers in the case of rapid changes in shopping behavior.

3. DETECTION OF SINGLE FRAUDULENT TRANSACTION

Detection of single fraudulent transaction in the window of the size $w = 8$ with forgetting factor $\lambda = 0.8$ is presented in the first plot of Figure 1, where the shopping amount is plotted with the blue line, weighted moving average is plotted with the black line, confidence interval is plotted with green lines and fraudulent amount is plotted with red round sign. The plot shows that current shopping behavior of the cardholder can be very well described by the confidence interval method, where the amounts of legitimate transactions are located inside of the interval. Fraudulent transaction is independent event, which is not associated with current shopping behaviour and has a strong chance (due to independency) of falling outside the confidence interval, see the first plot of Figure 1.

Several recent transactions are accounted only in calculation of the confidence interval which allows to describe rapidly changing shopping behaviour of the cardholder. Notice that all types of shopping behaviour can be described using different window sizes and forgetting factors.

Detection of fraudulent amount is difficult using the first plot of Figure 1 only since the amount is located almost on the upper limit. The amounts of the transactions located close to the confidence limits can be identified as suspicious and verified for fraudulency. Geographical information (geographical mismatches) and inter-transaction time gaps presented in the second plot of Figure 1 can be used for verification of suspicious transactions. A one-hour time gap between two successive transactions from Arizona and Italy is too small and



(a) (b)

Fig. 3. Merchant category codes (MCC) are associated with different types of businesses. Fraudulent activities can be associated with certain merchant category codes since certain business areas are more prone to fraud. The frequencies of fraudulent transactions calculated from historical data as a function of MCC are plotted in Figure 3(a) with blue round sings. The subplot shows MCC frequencies plotted with red color during fraud attack illustrated in Figure 2 and Figure 4. Figure 3(a) shows very strong correlation of the frequency of fraudulent merchant category codes from historical data and the codes obtained during the attack. The code 5311 indicates 'Department Stores', the code 5310 is associated with 'Discount Stores', the code 5300 is related 'Wholesale Clubs' and the code 4829 is assigned for 'Money Orders - Wire Transfer'.
Figure 3(b) shows the frequencies of MCC for legitimate transactions, where the code 5411 is associated with 'Grocery Stores or Supermarkets', the code 5499 is related to 'Misc Food Stores - Specialty Markets & Convenience' and the code 5541 is assigned for 'Gas/Service Stations with/without Ancillary Services'. The Figure shows that fraudulent transactions are very well separated from legitimate transactions in terms of merchant category codes.

the latter transaction is identified as fraudulent, see the second plot of Figure 1.

4. DETECTION OF THE START OF ATTACK: FRAUD RISK SCORING METHOD

Fraud risk scoring is well established technique widely used by financial institutions for detecting a range of fraudulent activities. The early detection of the start of the fraud attack is crucial for minimizing financial losses. The start of the attack is usually characterized by relatively small amounts and the amount monitoring methods cannot be used for detection, unfortunately.
Therefore the detection of the start of the fraud attack should be based on estimation of the intensity of fraudulent activity. *The intensity* refers to the degree of activity (associated with the frequency of fraudulent transactions) in the areas which are regularly linked to fraud.
Fraud risk scoring is applied here for detection of the start of the fraud attack via estimation of the intensity of the fraudulent activity. First, a number of factors associated with the strength of the attack is identified for quantification. Each factor in the list is quantified using risk score. The intensity is estimated via total fraud risk score defined as the sum of individual scores of different factors. The start of the attack is detected when the cumulative score goes beyond the predefined threshold.
The following factors (listed in order of priority, starting with the most important one) can be used for quantification of the fraudulent activity and detection of the start of the attack:

(1) high transaction rates, see Figure 2
(2) merchant category codes which are often associated with fraudulent activities, see Figure 2 and Figure 3
(3) transactions from different locations with small inter-transaction time gap (transactions with geographical inconsistencies), transactions from unusual locations and transactions from other countries (where strong customer authentication is not required), see Figure 1
(4) frequent transactions which occur at unusual times, see Figure 2
(5) transaction errors, like CVV, PIN, zip codes (applicable in the USA only) and other abnormal authorization errors, technical glitches, network connection errors and insufficient balance
(6) some other factors which can also be identified as suspicious.

Start of the attack which involves the sequence of several fraudulent transactions is identified via monitoring of inter-transaction time gap (which is the main variable in detection of the start of attack), see Figure 2. The scores are calculated for current and previous transactions provided that small inter-transaction time gap appears in the time series. If the total score of two transactions is not high enough for detection of the start of the attack the next transaction is added in the estimation, provided that inter-transaction time gap is small. If the time gap for the next transaction is large enough (larger than 8 hours) the detection is canceled. The end of attack can be detected using the same method. This method accounts for dynamics of transaction rates and equivalent to monitoring of inter-transaction time gaps in sliding window, whose size is adjusted for accumulation of additional information for more reliable detection.
Notice that the transaction amount is not included in the list since the start of attack is usually characterized by relatively small amounts. Larger amounts can be identified using confidence interval method,

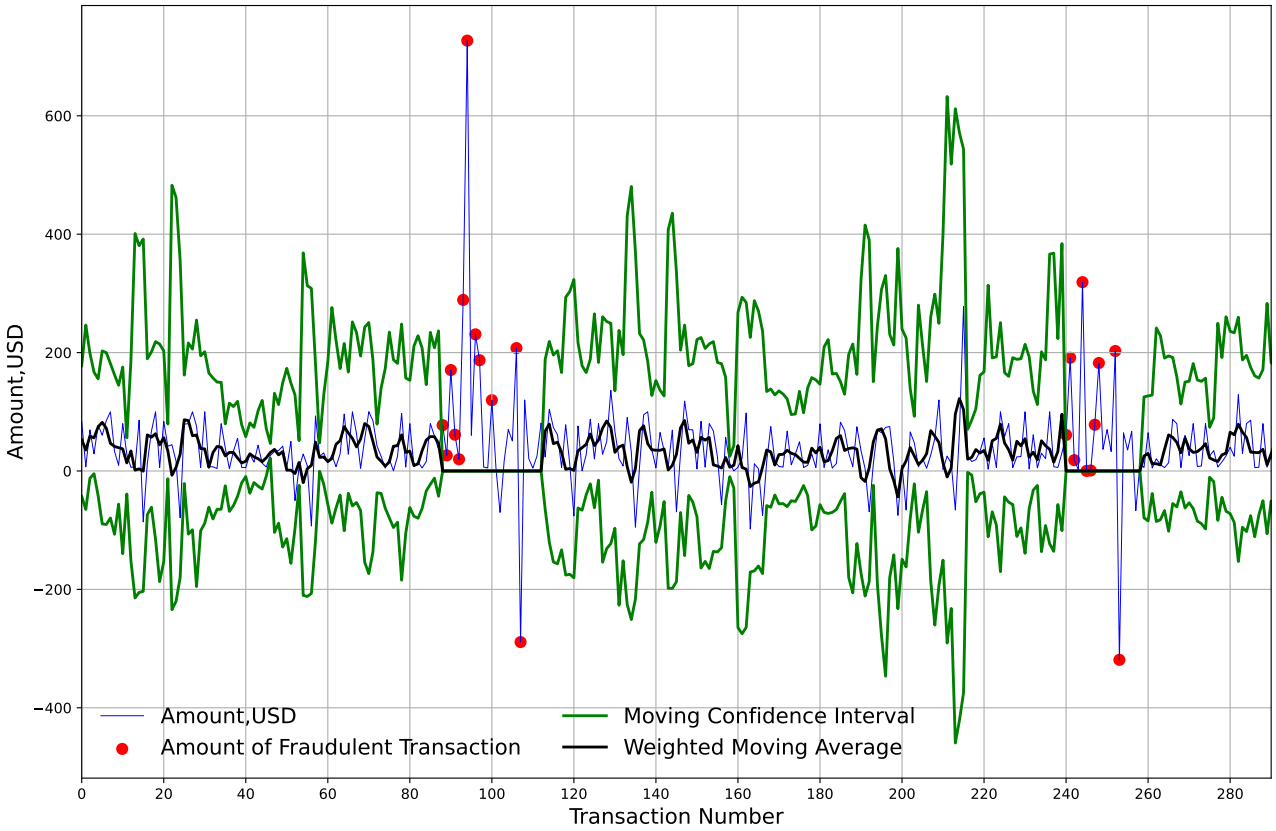


Fig. 4. Detection and control of the fraud attack is presented in this Figure. The confidence interval defined in (1) is plotted with green lines for $w = 3$ and $\lambda = 0.9$. The transaction amount is plotted with the blue line, weighted moving average is plotted with the black line and fraudulent amounts (similar to Figure 2) are plotted with red round signs. Fraudulent transaction is detected if shopping amount exceeds the threshold. The first two transactions in the attack were flagged as fraudulent due small inter-transaction time gap, technical glitch and unusual time. The confidence interval was assigned to zero in each step of moving window and several online fraudulent transactions with high rates and amounts, which took place at unusual time with technical glitches and CVV errors have been blocked and the fraud attack was prevented. Similar procedure was applied to the second fraud attack shown in the Figure. The ability of the detection system to recover after the attacks is simply associated with resetting of the length of the confidence interval.

quantified with the scores and added to the cumulative score which estimates the intensity of the attack. Notice that the confidence interval method in moving window allows flexible quantification of large transaction amounts with respect to the confidence limits, taking into account current shopping profile of the cardholder, see Figure 1. For example, sudden increase in spending around payday results in adjustments of the window size and forgetting factor, which imply changes in the confidence profile and the scores, see Section 1.

Figure 2 shows spending amount in US dollars plotted with the blue line. Fraudulent transactions are indicated with red round signs. Inter-transaction time gaps (in hours) for fraudulent transactions are plotted with cross signs of a red color in the subplot. Fraud attack is characterized by high transaction rate (sudden increase in activity) identified by small time gaps between transactions. Fraud attack was performed as sequence of online transactions with technical glitch and CVV error on Sunday (at unusual time).

Small time between successive transactions, merchant category codes 5311 and 5310 (business areas which are traditionally associated with fraudulent activities), see Figure 3, technical glitch and unusual transaction time resulted in high total risk which flagged the transactions 3 and 4 as fraudulent. Notice that the amounts of the first two fraudulent transactions can not be identified as unusual amounts, see Figure 4. Nevertheless, these fraudulent transactions indicate the start of the fraud attack.

Notice that legal transactions performed by the cardholder can be present inside of the fraud attack, see for example swipe transactions 10,13,14 and $16 - 20$ in Figure 2. These transactions can be identified as non-fraudulent transactions inside of the fraud attack using merchant category codes which are not associated with fraudulent activities. For example, the transactions $16 - 20$ are associated with food stores and supermarkets and can be identified as non-fraudulent transactions. It means that the cardholder was driving the car (see the code 5541 in Figure 3 of transaction 14) and purchased the food during the fraud attack.

Notice also that the first fraudulent transaction in the attack is impossible to identify using Bayes theorem for example, where the probability of fraud is the product of probabilities of the components, including probability of the amount being fraudulent (which is very

low since the amount is low) that, in turn, implies a low resulting probability of fraud. In addition, Bayesian methods do not usually account for dynamic behaviour, inter-transaction time gaps, MCC, technical glitches, CVV errors, which are the main variables in detection. The efficiency of detection methods which do not account the most significant variables can be low.

5. CONTROL OF FRAUD ATTACK IN THE CASE OF UNCERTAINTY

When the start of the attack is detected the control variable c_k and the mean value are assigned to zero to control the attack, see Figure 4. A number of online fraudulent transactions which took place at unusual times with high rates and amounts, well established MCC pattern, technical glitches and CVV errors has been blocked and the fraud attack was prevented. Notice that the average time of online attack is known and determines recovery after the attack, which is implemented as resetting of the length of confidence interval, see Figure 4.
Fraudsters continuously enhance their strategies and invent new forms of attacks in emerging business areas such as quasi cash (cryptocurrency), code 6051 and others. Existing detection algorithms may not be able to identify new patterns and attack mitigation strategies can be applied in the case of uncertainty, where the total score is not high, but the attack takes place. Additional confirmation of the transaction in the form of single-use authentication code sent via sms can be requested in such cases. Data enrichment methods can also be applied in the case of uncertainty and additional relevant information such as details of the transaction, IP address, and other supporting information can be requested. Finally, the transactions with certain merchant category codes which are likely associated with fraudulent activities can be blocked or limited to small amounts, like 50 USD, which results in significant savings.
In the case of misdetection the transactions with merchant category codes which are rarely associated with fraudulent activities can be processed that increases customer satisfaction.
Finally, the algorithm developers should monitor and proactively model new types of attacks which will likely appear in the future, outpacing fraudsters and protecting interests of the customers. To this end the simulation tools for modeling and detection of the attacks should be developed.

6. CONCLUSION & OUTLOOK

New concept of fraud attack, which allows accurate detection and control the sequences of fraudulent transactions, reduces financial losses and increases the trust of customers was introduced in this paper. It was shown that the beginning of attack is usually associated with a number of small transactions and the attack includes always the sudden changes in spending behaviour associated with high transaction amounts. The detection of such attacks is a complicated problem, which was divided in two parts.
The first part is associated with monitoring of the transaction amounts using confidence interval method in the moving window with exponential forgetting which allows to capture recent changes in the shopping behaviour of the cardholder. This part is able to capture fraudulent transactions of relatively large amounts as deviations from current shopping profile.
The second part, which is assigned for detection of the beginning of the attack employs fraud risk scoring method for estimation of the intensity of fraudulent activity. High transaction rates, business areas traditionally associated with fraudulent activities, unusual transaction times and locations, different types of errors and other factors are used for detection of the start of the attack.
The attack is controlled (after detection of the start) by adjusting the length of the confidence interval. A number of control strategies can be used for mitigation of the attack: 1) all the transactions located outside of the interval can be blocked or limited to small amounts, 2) the transactions with merchant category codes which are likely associated with fraudulent activities can be blocked, 3) the transactions with merchant category codes which usually are not associated with fraudulent activities can be executed that increases customer satisfaction, and 4) many other actions.
Further development of advanced real-time algorithms for detection of the fraud attacks based on data analysis and simulation tools is required since cybercriminals continuously improve their strategies and create new methods. Developed detection algorithms should be robust against disturbances in data communication systems and electricity grids for safe payments in crisis situations and war, [1].

Disclosure Statement

This research was not supported by any organization. The dataset contains around 24.4 million credit card transactions generated by multi-agent virtual world simulation performed by IBM. The details of the generation method can be found in [8]. The simulations were performed in Python.

7. REFERENCES

[1] Riksbank Payments report 2024. Riksbank, Sweden, 2024. https://www.riksbank.se/

[2] Panigrahi S., Kundu A., Sural S. & Majumdar A., Credit card fraud detection: a fusion approach using Dempster-Shafer theory and Bayesian learning. Information Fusion, 2009, 10: 354-363. https://www.sciencedirect.com/science/article/pii/S1566253509000141

[3] Banirostam H., Banirostam T., Pedram M. & Rahmani A. A model to detect the fraud of electronic payment card transactions based on stream processing in big data. Journal of Signal Processing Systems, 2023, 95: 1469 - 1484. https://link.springer.com/article/10.1007/s11265-023-01903-6

[4] Hafez I., Hafez A., Saleh A., Abd El-Mageed A. & Abohany A. A systematic review of AI-enhanced techniques in credit card fraud detection. Journal of Big Data,2025,12:6. https://doi.org/10.1186/s40537-024-01048-8

[5] Stotsky A., Kaczmarz Projection Algorithms with Rank Two Gain Update. J Sign Process Syst, 2024, 96:327–332. https://doi.org/10.1007/s11265-024-01915-w

[6] Stotsky A. Automotive Engines: Control, Estimation, Statistical Detection. Springer-Verlag, Berlin-Heidelberg, 2009. https://link.springer.com/book/10.1007/978-3-642-00164-2

[7] Moschini, G., Houssou, R., Bovay J., Robert-Nicoud S. Anomaly and fraud detection in credit card transactions using the ARIMA model. Eng. Proc., 2021, 5(1): 56. https://www.mdpi.com/2673-4591/5/1/56

[8] Altman E. Synthesizing credit card transactions. arXiv:1910.03033v1 [cs.DB],2019. https://doi.org/10.48550/arXiv.1910.03033