The Financial Security Model of Kazakhstan's Second-Tier Banks in the Digital era

Yejiao Zhang*

Al-Farabi Kazakh National University,Almaty,050040, Kazakhstan
**Corresponding author: 793319550@qq.com*

Abstract: *With the progress of digitalisation in recent years, many new risks have appeared for second-tier banks in Kazakhstan, such as cybercrime, weak authentication, liquidity problems caused by high-frequency trading, etc. Given the problems mentioned above, a hierarchical suspicious transaction detection mechanism based on recursive least squares and clustering has been introduced; a distributed ledger will be employed to guarantee the immutability of transactions, and machine learning dynamic thresholds will be used to deal with the heterogeneity of behaviour. Given the present problems in digital resilience, this paper will present the concepts of network segmentation and micro-segmentation in a zero-trust model, propose a plan for migration to post-quantum cryptography, and put forward a convergence model for anti-money laundering rules and smart contracts in cross-border payments.*

Keywords: *Financial Security Model, Suspicious Transaction Detection, Recursive Least Squares, Distributed Ledger, Zero-Trust Architecture, Post-Quantum Cryptography, Smart Contract*

Introduction

With the development of digital finance in recent years, many problems have arisen in the security system of Kazakhstan's second-tier banks, such as network intrusion and weak authentication. With the increase in both the number and types of transactions recently, the old system of fixed rules and audit after the fact will no longer be practical; thus, a new security model that can detect problems in real time, record them permanently, and change dynamically is needed. This paper will investigate the reasons for the occurrence of some risks in second-tier banks during the era of digital finance and put forward corresponding technical protection structures and resilient-reconstruction plans that can be built on the current infrastructure. According to the ideas of recursive least squares and clustering, a hierarchical detection mechanism will be established to improve the efficiency of locating suspicious transactions in a distributed ledger, and a machine-learning-driven dynamic threshold method will be employed to reduce the false-positive rate. The three parts of the future-oriented security model are zero-trust segmentation, post-quantum migration and the combination of anti-money laundering rules with smart contracts; together they will provide theoretical support for second-tier banks in Kazakhstan to ensure the reliability of fund settlement and the integrity of transaction data.

1. Evolutionary Mechanisms of Security Threats to Second-Tier Banks in the Digital Financial Ecosystem

1.1 Penetration Paths of Cyber Attack Vectors on the Core Clearing System of Banks

Most of the cyberattacks on the core clearing system of second-tier banks come from weak links in the external part of the network, such as public API interfaces or remote office access points. Obtain the low-level credentials by means of social engineering, then conduct lateral movement in the internal network to finally reach the real-time gross settlement payment engine. At the same time, Distributed Denial of Service traffic is often employed to conceal the attempt to modify the database transaction log. Due to the hybrid nature of the legacy system and the modern cloud service at Kazakhstan's second-tier banks, there is a blind spot in network traffic monitoring, and thus malicious packets can bypass the traditional firewall and directly reach the message queue of SWIFT or the local clearing channel.

After the attacker enters the main area, they will use the vulnerability in the operating system kernel

or unpatched message middleware to obtain the payment authorisation token. The tokens above can be used to issue interbank transfer orders without further authorisation. Most clearing systems are not organised according to transactions, so they will not be affected by this design flaw, and thus the audit records cannot be deleted. Some second-tier banks have shortened some links in the transaction-integrity verification process for the sake of efficiency and can no longer distinguish between malicious and legitimate transactions at the log level. Finally, the attacker spreads the funds among several shadow accounts via encrypted tunnels and thus carries out an illegal withdrawal of liquidity from the pool.

1.2 Analysis of the Correlation between Leakage of Customer Digital Identity Information and Vulnerability of Multi-Factor Authentication

Customer digital identity information has many leakage paths in the storage and transmission of second-tier banks. Incorrect data export by internal staff, excessive authorisation for third-party credit interfaces, and upload vulnerabilities in mobile application logs can all result in the batch extraction of names, identification document numbers and biometric templates. This information is now on the dark web and can be used to carry out an account takeover attack. More seriously, many second-tier banks are still using SMS verification codes or fixed security questions; thus, the security depends on the uniqueness of mobile phone numbers, but this uniqueness can be compromised by SIM swap attacks[1].

The weakness of multi-factor authentication is not isolated, and it is positively correlated with the extent of leakage of identity information. Learn the full name of the customer, how they usually make transactions and which devices they use, so that you do not have to go through knowledge-based authentication and can directly use the stolen session token or password reset link. Some of the biometric modules in second-tier banks do not employ live detection of sensor data and can be fooled by high-resolution pictures or video playback. There is no end-to-end encryption for the communication between the authentication server and the business server, and after successful authentication, the authorisation information can be modified by a man-in-the-middle attack. Therefore, the scope of the protection for identity information will be the effective upper bound of the multi-factor authentication system.

1.3 Risk of Sudden Liquidity Shortage caused by High-Frequency Algorithmic Trading Behavior

High-frequency algorithmic trading on the digital platforms of second-tier banks usually takes the form of a cycle of foreign exchange or securities order submissions and cancellations in milliseconds. To share in the profits, these algorithms will send a large number of tentative orders via the low-latency trading interface of the bank. When many algorithms work together in the liquidity pool, the artificial demand will quickly reduce the depth of the order book, and large orders from real customers will have to be executed at a worse price. Some of the trading systems at Kazakhstan's second-tier banks do not have an order-rate limit, and thus algorithmic traders can cause a market panic by repeatedly cancelling orders and triggering a chain reaction of automatic stop-loss orders.

There will be a liquidity shortage at the bank, and it will be unable to pay out funds from customer deposits on time. The false trading signals produced by high-frequency trading algorithms are misleading to the liquidity management model of the bank, and although this model is based on historical volatility, it cannot distinguish between algorithmic noise and actual changes in customer behaviour. If there is a significant drop in prices, the algorithmic traders will sell off rapidly, and at that time, the bank's clearing reserve may fall below the legal minimum. Second-tier banks need to borrow from the interbank lending market at a high cost to cover the deficit, and this will further worsen the liquidity problems of the whole system. Another serious problem is that some algorithms will go unnoticed; they may distribute the order among several accounts, and the total effect of this behaviour will only be known at the settlement stage, at which time the bank will be unable to reverse the already-executed transactions[2].

2. Technical Adaptation of the Security Protection Architecture for Kazakhstan's Second-Tier Banks

2.1 Hierarchical Suspicious Transaction Detection Mechanism Based on Recursive Least Squares and Clustering Algorithms

Recursive Least Squares is used to update the weights of the historical data in a time-series manner by adding a forgetting factor. Therefore, it can be used to handle the non-linear drift of fund account behaviour and changes in the distribution of transaction frequency and amount caused by fluctuations in the external environment. Most of the transaction data for second-tier banks in Kazakhstan are small and frequent payments, along with some large transfers; thus, the old statistical threshold cannot cover both kinds of patterns. Recursively least squares is employed to iteratively update the regression coefficients and obtain a sequence of residuals according to the current state of each transaction. Then, the clustering algorithm will group the residual sequences without any prior knowledge and mark the data points that are far from the center of the main cluster as potentially suspicious objects. Together, the two levels will be able to keep up with the new forms of money laundering.

At the level of particular deployment, the above mechanism will map the raw transaction logs of the second-tier bank to the three processing stages. The first layer stores the structured field data of customer information, transaction counterparty, timestamp and amount, etc., and recursive least squares is used to fit the spending habit baseline for individual accounts at this level. The second layer takes the regression residuals as its input features and then the APC-III clustering algorithm is used to form suspicious clusters based on the parameters of the intra-group radius and distance threshold. The output of the third layer is a binary label; that is to say, normal transactions in the low-confidence interval are released immediately, and abnormal transactions in the high-confidence interval are added to the manual review queue. Based on the above data, the hierarchical detection method can find about 30% of the total suspicious samples in the offline dataset of a second-tier bank in Kazakhstan. Set the forgetting factor to be between 0.95 and 0.99 so that both short-term fluctuations and long-term trends can be considered by the algorithm[3].

2.2 The Embedding Logic of Distributed Ledger Technology for Transaction Tamper-Proof Verification

Distributed Ledger Technology can be employed to append timestamps and multi-party consensus verification information to the transaction data of second-tier banks. Every transfer or account modification operation is a block that has a hash pointer and is chronologically linked to the previous block. If the second-tier bank in Kazakhstan frequently commits the log digest of its core clearing system to a permissioned blockchain, any modification of the historical transactions in the back-end will result in an inconsistency of the hash chain. The embedding logic does not need to store all the transaction data on the chain; instead, it will calculate the hash of a group of transactions and save only the root hash in the distributed ledger. It is not very big, but still reasonable. When the regulator or internal auditor needs to verify whether a particular transaction is genuine, the bank can provide the original data and the corresponding Merkle proof path, and then the node network will recompute the root hash and compare it with the stored information in the chain.

At the time of distributing the ledger, the interface latency from the old database of the second-tier bank to the blockchain network needs to be addressed. The response time of the traditional relational database is a few milliseconds, so the consensus algorithm has added some confirmation delay. Given that second-tier banks in Kazakhstan have a large number of transactions, sidechain or state channel technology can be used to carry out many off-chain transactions, and only batched hash values need to be published on the main chain. A Smart Contract module will be employed to automate the verification logic; that is to say, after generating the batch file, the contract will determine whether the digital signature of the file originated from the hardware security module of an authorised administrator. If the signature is valid and the timestamp has not exceeded the upper limit, then store the root hash and metadata of the batch in the contract. A bad actor tries to post a fake group but fails the verification and is not allowed. Embedding logic can be used to solve the cross-bank problem; that is, when several second-tier banks are added to the same consortium blockchain, no single party can arbitrarily alter an approved batch of transactions, and thus the foundation of trust for clearing and settlement will be strengthened.

2.3 Method of Machine Learning Behavioral Anomaly Scoring and Dynamic Threshold Adjustment

According to the transaction record of the user, this model will determine the extent of the current operation abnormality. The attributes are the periodic deviation of the transfer time, changes in the proportion of counterparty accounts, and the number of standard deviations from the mean for a transaction amount. The customer base of Kazakhstan's second-tier banks is mainly individual merchants and small businesses, and their fund flow patterns are quite different; therefore, a single global threshold cannot be used. Isolation Forest or an Autoencoder unsupervised model is used to assign an anomaly score of 0 to 1 to each transaction in the absence of labelled data. High-scoring transactions are those that are far from both the normal behaviour of the account and the distribution among all users[4].

To reduce the number of false alarms in a busy period with many transactions, a variable threshold can be employed instead of a fixed one. The adjustment mechanism takes the sequence of anomaly scores in a sliding time window as input and computes the moving average and standard deviation of quantiles for that window, etc. If the median of the window is significantly different from that in the previous window, automatically adjust the decision threshold up or down to keep a certain proportion of suspicious transactions. For example, there are a large number of transactions during the holiday season, so the dynamic threshold has been increased to avoid flagging many normal high-frequency transfers as abnormal. The model also has a feedback loop; it adds samples of false negatives and false positives that have been confirmed by humans to the training set and then updates the feature weights and splitting rules of the isolation trees. Based on the above method, the second-tier banks in Kazakhstan need to address the problem of cold-start accounts; that is to say, at the beginning when there is little historical transaction data, a population statistical threshold should be used as a temporary substitute, and then switch to a personalised dynamic adjustment mode after collecting a sufficient number of behavioural samples.

3. Reconstruction Strategy for the Security Model of Second-Tier Banks in the Digital Age

3.1 Deployment Principles of Internal Network Segmentation and Micro-segmentation under the Zero-Trust Architecture

A Zero Trust model does not assume that there is a trusted area in the network of a second-tier bank, so all access requests must be independently verified and authorised according to their origin. Based on the above, the internal network of a second-tier bank in Kazakhstan has been divided into several logical micro-segments, and each micro-segment is connected to a particular part of the business, such as the payment processing area, the customer data section, and the clearing and reconciliation zone. A Software-Defined Perimeter will be used to control communication in the micro-segment, and the default policy for this will be to deny all traffic. Only when all three conditions for the healthy state of the device, the strength of the user's identity and the contextual behaviour score are met will the system temporarily open the corresponding port and protocol. Therefore, the range of a side movement by an attacker after a single point of entry will be small. Even if the application server has been hacked, one cannot directly access the main accounting database.

According to the present network topology and traffic conditions of the second-tier banks, set up micro-segmentation accordingly. Most of the second-tier banks in Kazakhstan have old hardware, and some of the old switches do not support virtual network tagging. Therefore, an agent-based isolation scheme will be employed to put a small firewall module in the workload and handle all the incoming and outgoing traffic at the host level. For cross-segment data exchange, such as invoking the transaction flow in the customer information micro-segment from the anti-money laundering micro-segment, a special gateway is employed to carry out protocol conversion and content filtering. The management plane centrally distributes the security policy, each micro-segment regularly reports the access log, and if there is any anomaly in the log, the isolation rule will be changed. Based on the deployment rules, the internal operation and maintenance responsibilities of the bank should be in line with the principle of least privilege; for example, a database administrator should only be able to access the storage node they are responsible for and not query unauthorised tables in other parts of the system.

3.2 The Migration Path of Cryptographic Protocols to Post-Quantum Cryptography in Light of the Quantum Computing Threat

The first danger that quantum computers will pose to the current cryptographic system at second-tier banks is public-key cryptography, that is, RSA and elliptic-curve cryptography. A computer with a large number of qubits can solve the discrete logarithm problem and the problem of large integer factorization in polynomial time, thus breaking the algorithms of digital signatures and key exchange. Second-tier banks in Kazakhstan have started to apply the above cryptographic system to cross-border payments, message authentication and certificate management, but the current cryptographic protocols are vulnerable to attack by future quantum computers. There will be no change to the algorithms, but all the certificates, communication protocols and hardware security modules of the system need to be updated instead[5].

The first stage of migration can be a hybrid encryption mode, that is to say, both the old elliptic curve key encapsulation and a lattice-based post-quantum encapsulation method are employed in the current TLS session. The two outputs are combined with an XOR operation to get the session key, so if only one of the algorithms is known, the other can still provide some protection. Second-tier banks in Kazakhstan can first concentrate on building the hybrid scheme for the core clearing system and the SWIFT access gateway, and then add post-quantum signature algorithms at the time of verifying the software update package and anchoring the log hash chain. The hardware accelerators need to be replaced or upgraded because the key sizes and computational overhead of lattice-based algorithms are too large for the current schemes. A dual-certificate period will be used during migration to save both the old and the new certificates on all network equipment, and then a load balancer will select one. After the international standards organisation publishes the new post-quantum cipher list, we will stop using the old ones.

3.3 Integration Paradigm of Anti-Money Laundering Rules and Smart Contracts in Cross-border Digital Payment Scenarios

Cross-border digital payments are cross-border fund transfers, and the second-tier banks in Kazakhstan act as correspondent banks or remittance institutions for such transactions. The old anti-money laundering rules are based on manual checks and after-the-fact reporting, so they are not very effective in dealing with many small-amount, frequent cross-border payments. Programmability of smart contracts can be used to set a particular condition for the rules of anti-money laundering; that is, when a transaction meets a certain trigger pattern, a smart contract will automatically take some action, such as blocking, delaying or refusing payment. For example, if a receiving account has been receiving funds from remitters in many different countries over a short period, and the total amount exceeds a certain limit, the smart contract can temporarily block the transfer and have the on-chain risk assessment module recalculate the anomaly score of that account.

By the second tier, a model of integrated services will be built on a consortium distributed ledger in Kazakhstan, and all parties to the cross-border payment transactions (such as remittance banks, intermediary banks and receiving banks) will need to comply with the same anti-money laundering regulations. The Rule Language is a specific type of language, and its parameters include the transaction amount threshold, geographical risk coefficient, account life cycle, etc. Automatically extract the field values of the cross-chain message at the time of transmission, compare them with the blacklist address database and behavioral pattern database stored on the chain, and so on. If it meets the above conditions, it will not be immediately rejected; instead, it will be put in a multi-signature confirmation dispute period and subject to fast-track review by the on-duty nodes appointed by the respective banks in sequence. The execution trace of the whole process is stored in the ledger as an immutable chain of audit records. Therefore, there will be less subjectivity in the manual submission of suspicious transaction reports, and cross-border fund flows will be more transparent and traceable.

Conclusion

To solve the problem of financial security for second-tier banks in Kazakhstan in the digital era, the three links of the threat evolution mechanism, the adjustment of the protection architecture, and the reconstruction of the security model will be systematically studied in this paper. According to the hierarchical detection mechanism of recursive least squares and clustering algorithms, about 30% of the mixed suspicious samples can be found in the offline data test, and problems with transaction

tamper-proof verification and adaptation to behavioural heterogeneity can be solved by embedding in distributed ledgers and using machine learning dynamic threshold methods. The three ways to increase the digital resilience of the bank system are to introduce zero-trust micro-segmentation, switch to post-quantum cryptography and combine anti-money laundering rules with smart contracts.

The following are some possible directions for the future. First, we will present the combination of Recursive Least Squares and deep temporal networks in this paper to improve the detection accuracy of complex money-laundering network paths. Second, one can use online Bayesian methods to track the factor parameters and thus reduce the dependence of recursive estimation on the starting value. Third, it is possible to add zero-knowledge proof protocols on the consortium blockchain to verify the compliance of cross-border payments and protect the privacy of customer transaction data. Given that the hardware and network topology of the various generations of second-tier banks in Kazakhstan are different, some lightweight edge computing nodes can be added to offload some of the anomaly scoring and isolation decisions at the branch-level gateway and reduce the computational load on the central system.

References

[1] Zhang Chao. Digital Empowerment of Supply Chain Finance: A Four-Dimensional Reshaping of Enterprises, Banks, Platforms, and Security. Cybersecurity and Informatization, no. 04 (2026): 1-2.

[2] Yuan Guofang and Ning Xueping. The Formation Mechanism and Prevention of Systemic Financial Risks. Gansu Social Sciences, no. 04 (2024): 185-194.

[3] Yu Han. How to Walk the Path of Digital Transformation for Commercial Banks. China Bank Insurance News, December 12, 2023, p. 008, Fintech.

[4] Ke Sijie. Digital Transformation of Rural Commercial Banks under the 'Open Bank' Model. China Foreign Investment, no. 16 (2022): 105-107.

[5] Zhang Qian and An Zhen. Thoughts and Suggestions on the Development of Digital Inclusive Finance in Small and Medium-Sized Banks. Information and Communications Technology and Policy, no. 03 (2022): 82-85.