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: *In the digital era, Kazakhstan's second-tier banks face compound threats such as cyberattack penetration, vulnerable authentication systems, and liquidity imbalances triggered by high-frequency trading. To address the above risks, this study constructs a hierarchical suspicious transaction detection mechanism based on recursive least squares and clustering, embeds distributed ledgers to ensure transaction immutability, and adopts machine learning dynamic thresholds to accommodate behavioral heterogeneity. With a focus on digital resilience, this study proposes the principles of network segmentation and micro-segmentation under a zero-trust architecture, plans a migration roadmap toward post-quantum cryptography, and presents a convergence paradigm of anti-money laundering rules and smart contracts in cross-border payment scenarios.*

Key words: *Financial security model; Suspicious transaction detection; Recursive least squares; Distributed ledger; Zero-trust architecture; Post-quantum cryptography; Smart contract*

Introduction

Kazakhstan's second-tier banks, amid the expansion of digital financial services, face multidimensional challenges in their security protection systems, including network penetration, authentication flaws, and algorithmic trading disturbances. Traditional static rules and post-event audits are unable to cope with the non-stationary fluctuations in transaction frequency and the continuous evolution of attack methods, thereby necessitating a new security model that integrates real-time detection, tamper-proof storage, and dynamic adaptive adjustment. This study aims to identify the evolutionary mechanisms of the unique threats faced by second-tier banks in the digital financial ecosystem and to propose technical protection architectures and resilience reconstruction strategies that are compatible with existing infrastructure. By introducing a hierarchical detection mechanism that combines recursive least squares and clustering, a verifiable embedding logic of distributed ledgers, and a dynamic threshold method driven by machine learning, this study can improve the efficiency of suspicious transaction identification while controlling the false positive rate. A future-oriented security model is constructed from three dimensions: zero-trust segmentation, post-quantum migration, and the integration of anti-money laundering rules with smart contracts, thereby providing theoretical support for Kazakhstan's second-tier banks to maintain 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 Targeting the Core Clearing System of Banks

Cyber attacks targeting the core clearing system of second-tier banks typically begin at vulnerable nodes on the external perimeter, such as public API interfaces or remote office access points. After obtaining low-level credentials through social engineering techniques, an attacker moves laterally along the internal network and gradually reaches the payment engine that processes real-time gross settlement. In this process, distributed denial-of-service traffic is often used as a cover to mask attempts to tamper with database transaction logs. Due to the hybrid architecture of legacy systems and modern cloud services in Kazakhstan's second-tier banks, blind spots exist in their network traffic monitoring, which allows malicious packets to bypass traditional firewalls and directly insert themselves into the message queues of SWIFT or local clearing channels.

Once the attacker enters the core area, he exploits operating system kernel vulnerabilities or unpatched message middleware to extract payment authorization tokens. Once these tokens are forged, the attacker can initiate interbank transfer instructions without secondary verification. The core clearing system typically separates transaction ordering from balance verification, and this design flaw gives the attacker a time window to cover up audit traces. Some second-tier banks, in order to improve processing speed, disable certain steps in transaction integrity verification, making it impossible to effectively distinguish malicious transactions from legitimate ones at the log level. The attacker finally disperses the funds to multiple shadow accounts through encrypted tunnels, thereby completing the illegal extraction of the liquidity pool.

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

Customer digital identity information faces multiple leakage channels during its storage and transmission in second-tier banks. Improper data export by internal employees, excessive authorization of third-party credit interfaces, and upload vulnerabilities in mobile application logs can all lead to the batch extraction of names, identification document numbers, and biometric templates. Once this information flows into the dark web, it provides the basic materials for account takeover attacks. More notably, many second-tier banks rely on SMS verification codes or static security questions, whose security highly depends on the exclusivity of mobile phone numbers, and this exclusivity almost fails in the face of SIM swap attacks[1].

The vulnerability of multi-factor authentication does not exist in isolation, but it is positively correlated with the degree of identity information leakage. When attackers have mastered the customer's full name, transaction habits, and common device fingerprints, they can bypass the knowledge-based authentication links and directly exploit leaked session tokens or password reset links. The biometric modules deployed by some second-tier banks fail to perform liveness detection on sensor inputs, allowing high-resolution photos or video replay attacks to deceive the system. Furthermore, the communication between the authentication server and the business server lacks end-to-end encryption, which enables man-in-the-middle attacks to tamper with subsequent authorization instructions after authentication has passed. This correlation indicates that the protection strength of identity information directly determines the actual upper limit of the multi-factor authentication system.

1.3 Risk of Instantaneous Liquidity Imbalance Caused by High-Frequency Algorithmic Trading Behavior

High-frequency algorithmic trading in the digital platforms of second-tier banks typically manifests as a cycle of foreign exchange or securities order submissions and cancellations within milliseconds. To capture spread profits, such algorithms send a large number of tentative orders through the low-latency trading interfaces provided by the banks. When multiple algorithms operate simultaneously on the same liquidity pool, the order book depth is exhausted by spurious demand in an extremely short time, which forces large transactions of real customers to be executed at unfavorable prices. Some trading systems of Kazakhstan's second-tier banks do not impose order rate limits, thereby allowing algorithmic traders to create market panic through continuous order cancellations and trigger a chain reaction of automatic stop-loss instructions.

The direct consequence of instantaneous liquidity imbalance is the mismatch between the bank's own capital positions and the customers' assets under custody. The false trading signals generated by high-frequency algorithms mislead the bank's liquidity management model, which predicts funding needs based on historical volatility but fails to distinguish between algorithmic noise and real customer behavior. When extreme market fluctuations occur, algorithmic traders collectively exit the market, and the bank's clearing reserves may instantly fall below the minimum ratio required by regulations. Second-tier banks have to borrow funds from the interbank lending market at a high cost to fill the gap, and this process further exacerbates the liquidity strain within the system. A more insidious risk lies in the fact that some algorithms circumvent monitoring by splitting orders across multiple accounts, and the cumulative effect of such practices only becomes exposed during the settlement phase, at which point the bank cannot 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

When processing the transaction flow data of second-tier banks, the recursive least squares method can dynamically adjust the weights of historical observations through a forgetting factor. This characteristic makes it suitable for capturing the nonlinear drift of fund account behaviors, especially when the transaction frequency and amount distribution undergo non-stationary fluctuations due to changes in the external environment. The transaction records of Kazakhstan's second-tier banks often contain a large number of small frequent transfers and intermittent large-value remittances, and traditional statistical thresholds cannot cover both patterns simultaneously. After introducing recursive least squares, the system can update the regression coefficients in real time and generate a residual sequence for each transaction relative to the current behavioral pattern. The clustering algorithm then performs unsupervised grouping on the residual sequences and marks the data points that deviate from the center of the majority cluster as potentially suspicious objects. The two-layer mechanism operates in coordination, thereby avoiding the lagging response of a single classifier to new money laundering techniques.

At the specific deployment level, this mechanism maps the raw transaction logs of second-tier banks into three processing layers. The first layer retains structured fields such as customer identity, transaction counterparty, timestamp, and amount, and the recursive least squares algorithm fits the spending habit baseline of individual accounts at this layer. The second layer takes the regression residuals as input features and uses the APC-III clustering algorithm to form suspicious clusters based on the intra-group radius and distance threshold parameters. The third layer outputs a binary label, where normal transactions in the low-confidence interval are directly released, and abnormal transactions in the high-confidence interval are sent to a manual review queue. Empirical results show that this hierarchical detection mechanism identified approximately 30% of mixed suspicious samples on an offline dataset of a certain second-tier bank in Kazakhstan. The forgetting factor is set between 0.95 and 0.99, which allows the algorithm to strike a balance between tracking short-term fluctuations and maintaining long-term trend stability[3].

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

Distributed ledger technology provides second-tier banks with the capabilities of timestamp fixation for transaction records and multi-party consensus verification. Each transfer or account change operation is encapsulated as a block with a hash pointer and is linked in chronological order after the previous block. If a second-tier bank in Kazakhstan regularly anchors the log digest of its core clearing system to a permissioned blockchain, any back-end tampering with historical transactions will destroy the consistency of the hash chain. This embedding logic does not require all transaction data to be uploaded to the chain; instead, it performs hash operations on batches of transactions and then writes the root hash value into the distributed ledger. This approach reduces storage overhead while preserving auditability. When a regulator or an internal auditor needs to verify the authenticity of a specific transaction, the bank can provide the original data and the corresponding Merkle proof path, and the node network will recalculate the root hash and compare it with the record on the chain.

During the process of embedding the distributed ledger, the interface latency between the second-tier bank's existing database and the blockchain network must be addressed. Traditional relational databases aim for millisecond-level responses, whereas the consensus mechanism introduces additional confirmation waiting time. Given the transaction throughput characteristics of Kazakhstan's second-tier banks, sidechain or state channel technologies can be adopted to process high-frequency transactions off-chain, and only the batched hash values are submitted to the main chain. A smart contract module is designed to automatically execute the verification logic: whenever a batch file is generated, the contract checks whether the digital signature of the file comes from the authorized administrator's hardware security module. If the signature is valid and the timestamp does not exceed the preset window, the contract records the root hash and metadata of the batch. When a malicious node attempts to submit a forged batch, the signature verification step directly rejects it. This embedding logic also supports cross-bank scenarios: when multiple second-tier banks share the same consortium blockchain, no single party can unilaterally modify a confirmed transaction batch, thereby enhancing

the trust foundation of clearing and reconciliation.

2.3 Machine Learning Behavioral Anomaly Scoring and Dynamic Threshold Adjustment Method

The machine learning model generates a behavioral anomaly score for each operation by performing feature engineering on the account's historical transaction sequences. The features used include the periodic deviation of the transfer time, the change in the concentration of counterparty accounts, and the multiple of the standard deviation between the transaction amount and its historical mean. The customer base of Kazakhstan's second-tier banks includes individual merchants and small enterprises, and their fund flow patterns differ significantly, so a single global threshold cannot adapt to such heterogeneity. By adopting an unsupervised model based on isolation forests or autoencoders, the system maps each transaction to an anomaly score between 0 and 1 without relying on labeled data. High-scoring transactions represent candidates that highly deviate from the account's own regular pattern or the overall population distribution[4].

The dynamic threshold adjustment method solves the problem of an increased false positive rate of fixed cutoff points during periods of transaction volume fluctuation. This adjustment mechanism takes the sequence of anomaly scores within a sliding time window as its input and calculates the moving average and standard deviation of the window's quantiles. When the median within the window shifts significantly relative to the previous window, the system automatically raises or lowers the decision threshold, thereby maintaining the expected proportion of suspicious transactions within a preset range. For example, during a surge in transaction volume around holidays, the dynamic threshold rises correspondingly to avoid marking a large number of normal high-frequency transfers as anomalies. The model also introduces a feedback loop, which reinjects samples of false negatives and false positives confirmed by human review into the training set, and updates the feature weights and the splitting rules of the isolation trees. When deploying this method, Kazakhstan's second-tier banks need to pay attention to the handling strategy for cold-start accounts, that is, in the early stage when historical transaction data are lacking, they should use a population statistical threshold as a temporary substitute, and then switch to the personalized dynamic adjustment mode after accumulating sufficient behavioral samples.

3. Reconstruction Strategy of the Security Model for Second-Tier Banks Aiming at Digital Resilience

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

The zero-trust architecture assumes that no trusted areas exist within the internal network of a second-tier bank, and any access request, regardless of its source, must undergo independent authentication and authorization. Based on this premise, the internal network of a second-tier bank in Kazakhstan is divided into multiple logical micro-segments, and each micro-segment corresponds to a specific business function domain, such as the payment processing domain, the customer data domain, and the clearing and reconciliation domain. The communication between the micro-segments is implemented through a software-defined perimeter, and the default policy is to deny all traffic. Only when the requesting party simultaneously meets the three conditions of device health status, user identity strength, and contextual behavioral score does the system temporarily open the designated ports and protocols. This segmentation method effectively limits the lateral movement radius of an attacker after a single point of breach. Even if an application server is compromised, the attacker cannot easily reach the core accounting database.

The specific deployment of micro-segmentation must take into account the existing network topology and traffic characteristics of second-tier banks. Second-tier banks in Kazakhstan typically operate hardware devices from multiple coexisting generations, and some legacy switches do not support virtual network tagging. The solution is to adopt an agent-based isolation scheme, which installs a lightweight firewall module inside the workload to control inbound and outbound traffic at the host granularity. For cross-segment data exchange scenarios, such as calling transaction flows from the customer information micro-segment to the anti-money laundering micro-segment, the system sets up a dedicated gateway for protocol conversion and content filtering. The management plane centrally issues security policies, each micro-segment regularly reports access logs, and anomalies in the logs trigger dynamic adjustments to the isolation rules. This deployment principle also requires that the

bank's internal operation and maintenance roles follow the principle of least privilege: a database administrator can only access the storage node for which he is responsible and cannot query unauthorized tables across zones.

3.2 Migration Path from Cryptographic Protocols to Post-Quantum Algorithms from the Perspective of Quantum Computing Threats

The threat of quantum computing to the existing cryptographic protocols of second-tier banks mainly affects public-key cryptosystems, including RSA and elliptic curve cryptography. A computer with a sufficient number of quantum bits can solve discrete logarithm and large integer factorization problems in polynomial time, thereby breaking the algorithms used for digital signatures and key exchange. Second-tier banks in Kazakhstan rely heavily on the above cryptographic systems in cross-border payments, message authentication, and certificate management, and their cryptographic protocols face potential risks from future quantum attacks. The migration to post-quantum algorithms is not a simple algorithm replacement, but rather involves full-stack adaptation of certificate systems, communication protocols, and hardware security modules[5].

The migration path can begin with a hybrid encryption mode, that is, using both traditional elliptic curve key encapsulation and a lattice-based post-quantum encapsulation mechanism in the existing TLS session. After the two outputs are combined through an XOR operation, the system generates the session key, thereby ensuring that even if one algorithm is broken, the other algorithm still provides protection. Second-tier banks in Kazakhstan can prioritize the deployment of the hybrid scheme between the core clearing system and the SWIFT access gateway, and then gradually introduce post-quantum signature algorithms into software update package verification and log hash chain anchoring steps. Hardware accelerators need to be replaced or upgraded because the key sizes and computational overhead of lattice-based algorithms far exceed those of existing schemes. During the migration process, the bank sets up a dual-certificate period, where each network element holds both a traditional certificate and a post-quantum certificate, and a load balancer determines the negotiation strategy. After international standards organizations publish official post-quantum cipher suites, the bank will then disable support for traditional algorithms.

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

Cross-border digital payments involve fund transfers across multiple jurisdictions, and second-tier banks in Kazakhstan play the role of correspondent banks or remittance intermediaries in these transactions. Traditional anti-money laundering rules rely on manual review and post-event reporting, and they face efficiency bottlenecks when dealing with high-frequency, small-value cross-border payment flows. The programmability of smart contracts allows anti-money laundering rules to be directly encoded as predefined conditions, and when a transaction meets a specific trigger pattern, the smart contract automatically executes actions such as freezing, delaying, or rejecting the payment. For example, when a receiving account receives funds from remitters in multiple different countries within a short period of time and the cumulative amount exceeds a certain limit, the smart contract can temporarily withhold the transfer and invoke the on-chain risk assessment module to recalculate the anomaly score of that account.

The integration paradigm requires Kazakhstan's second-tier banks to build a consortium distributed ledger, where all parties involved in cross-border payments (including the remitting bank, intermediary banks, and the receiving bank) jointly maintain the same set of anti-money laundering rules. The rules are described in a domain-specific language, including parameters such as transaction amount thresholds, geographical risk coefficients, and account lifecycles. The smart contract automatically extracts field values during each cross-chain message transmission and compares them against the blacklist address database and behavioral pattern database stored on the chain. Transactions that hit a rule are not rejected immediately; instead, they enter a multi-signature confirmation dispute phase, where fast-track reviews are conducted by on-duty nodes assigned by each bank in turn. The execution trace of the entire process is recorded on the ledger, forming an immutable chain of audit evidence. This paradigm reduces subjective bias in manually filing suspicious transaction reports and enhances the transparency and traceability of cross-border fund flows.

Conclusion

In response to the financial security challenges faced by Kazakhstan's second-tier banks in the digital era, this study systematically discusses the three aspects of threat evolution mechanisms, protection architecture adaptation, and security model reconstruction. The hierarchical detection mechanism based on recursive least squares and clustering algorithms identifies approximately 30 percent of mixed suspicious samples in offline data tests, while the distributed ledger embedding and machine learning dynamic threshold methods respectively solve the problems of transaction tamper-proof verification and behavioral heterogeneity adaptation. The three strategies of zero-trust micro-segmentation deployment, post-quantum cryptography migration, and the integration of anti-money laundering rules with smart contracts jointly enhance the digital resilience of the bank system.

Future directions can be explored in the following areas. First, this study can combine recursive least squares with deep temporal networks to further improve the detection sensitivity for complex money laundering network paths. Second, this study can explore online Bayesian methods without forgetting factor parameters to reduce the dependence of recursive estimation on initial value settings. Third, this study can introduce zero-knowledge proof protocols into the consortium blockchain to achieve compliance verification of cross-border payments while protecting customer transaction privacy. In addition, in view of the unique hardware generation gaps and network topology constraints of Kazakhstan's second-tier banks, subsequent work can design lightweight edge computing nodes to offload part of the anomaly scoring and isolation decisions to branch-level gateways, thereby reducing the computational load on the core 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.