An AI-Enabled Anti-Fraud Risk Early-Warning System and Empirical Evidence: A Big-Data and Multi-Model Framework

Zhuming Chen[1,*], Zhongyuan Li[2], Tian Qin[3]

[1] Shenhua AI Research Institute, Guangzhou Huashang College, Guangzhou, China 511300

[2] School of Accounting, Guangzhou Huashang College, Guangzhou, China 511300

[3] Department of Auditing, School of Accounting, Guangzhou Huashang College, Guangzhou, China 511300

* Email: lxy5882@gzhs.edu.cn (Corresponding Author)

Received 18 February 2025 · Revised 4 April 2025 · Accepted 19 May 2025 · Published 30 September 2025

Abstract

Telecom fraud has evolved into an industrialized and highly organized form of financial crime, exposing limitations in rule-based monitoring and siloed model deployment within commercial banks. This study proposes a lifecycle-oriented risk early-warning system that integrates three complementary modules: (i) a pre-event account-opening risk rating model driven by multi-source big-data profiling, (ii) an in-process transaction-level detector built on LightGBM for real-time identification of anomalous patterns, and (iii) a post-event linkage mining module that combines LPA-based community detection with knowledge-graph association analysis to uncover organized fraud networks and expand investigative leads. Using production data from Bank A, we demonstrate that the integrated framework improves both coverage and operational actionability, strengthens closed-loop risk governance across pre-event prevention, mid-event monitoring, and post-event tracing, and delivers measurable reductions in reported fraud risk relative to peer benchmarks. The proposed architecture offers a deployable blueprint for financial institutions seeking scalable, data-driven, and continuously updatable anti-fraud risk management.

Keywords: Telecom fraud; risk early warning; big-data analytics; LightGBM; community detection (LPA); knowledge graphs

An AI-Enabled Anti-Fraud Risk Early-Warning System and Empirical Evidence: A Big-Data and Multi-Model Framework

1. Introduction

In recent years, telecom fraud—where perpetrators exploit deception and digital channels to steal victims' funds through phone calls or online platforms—has become one of the fastest-growing property crimes in China. This crime exhibits new characteristics: increasingly diversified fraud scripts, more technologically sophisticated and professionalized operations, highly organized networked structures, and industrialized criminal activities, causing substantial losses to the public.

Current commercial bank account transaction identification models commonly suffer from limited functional coverage, miscalibrated thresholds, and slow iteration cycles, rendering traditional risk management systems inadequate for rapidly evolving fraud typologies. Establishing a systematic, accurate, and operationally actionable risk-monitoring framework has therefore become an urgent priority for financial institutions.

Existing research predominantly focuses on single-technology applications—either machine learning models or graph technologies—while integrated, multi-technology solutions remain underdeveloped. Moreover, data utilization during model development is often confined to single-perspective information, failing to leverage multi-dimensional behavioral, relational, and temporal signals to improve performance. In addition, many studies mainly address fraud detection during or after incidents, lacking a full-lifecycle monitoring framework spanning pre-incident screening, in-incident early warning, and post-incident relational mining.

To address these gaps, this study integrates big data analytics, machine learning, community (network) modeling, and knowledge graphs to establish a risk assessment framework for combating telecom fraud. Big data technology is employed to collect, process, and analyze massive datasets of banking transactions and customer information, forming the foundation for risk modeling. Machine learning algorithms such as LightGBM are utilized to develop near real-time transaction monitoring models that identify abnormal patterns and generate timely risk alerts. Community modeling techniques such as the LPA algorithm detect potential criminal groups by uncovering hidden structures through account-correlation analysis. Knowledge graph technology supports post-incident correlation analysis by constructing transactional relationship maps, enabling deep link analysis among fraudulent accounts to further expose criminal networks.

Using authentic data from Bank A as a case study, this paper identifies key risk characteristics and their influencing factors, and establishes an intelligent risk identification system covering pre-event, in-process, and post-event phases. Specifically, the system operates through three phases: the pre-event phase employs big data to build account-rating models; the in-process phase develops machine learning and suspicious-community detection models for fraud transactions; and the post-event phase utilizes

ISSN: 3068-3661 | © 2025 INATGI (Institute of Advanced Technology and Green Innovation)

knowledge graph technology to construct association models, ultimately forming a comprehensive risk identification framework integrating big data and AI technologies.

The primary contributions of this paper are threefold. First, multi-technology integration: by combining big data, machine learning, community modeling, and graph technology, we construct a coherent risk monitoring and identification system. Second, full-scenario coverage: from “pre-event” to “post-event” stages, the framework supports account-opening screening, transaction-time detection, community discovery, and relational tracing. Third, closed-loop risk governance: through in-depth mining of telecom-fraud accounts and fund-flow characteristics, multi-dimensional data are leveraged to improve model performance and support a closed-loop process of prevention, real-time monitoring and early warning, and post-incident association analysis (see Figure 1).

The remainder of this article is organized as follows: Section 2 reviews related literature; Section 3 analyzes the financial risk characteristics of telecom fraud; Section 4 presents a “pre-event” account-opening model based on big data; Section 5 develops an “in-process” machine learning model using LightGBM and a suspicious-community model employing LPA; Section 6 proposes a “post-event” association analysis model based on knowledge graphs; and Section 7 concludes with implications.

2. Literature Review

Big data and artificial intelligence technologies have been increasingly adopted in financial fraud prevention, including credit card fraud, anti-money laundering, insurance fraud, and other forms of financial misconduct (Al-Hashedi & Magalingam, 2021; Zhu et al., 2021). Recent research has also begun to move beyond the detection of isolated suspicious transactions toward the analysis of interconnected accounts, transaction relationships, and organized fraud networks (Usman et al., 2023). Existing fraud-detection approaches are commonly classified into supervised, unsupervised, semi-supervised, hybrid, and graph-based paradigms, with their suitability depending on data availability, fraud characteristics, and operational requirements (Hilal et al., 2022).

Within supervised learning, researchers have explored a range of feature-selection, sampling, cost-sensitive learning, and ensemble strategies to enhance fraud-detection accuracy. Rtayli and Enneya (2020), for example, combined random-forest-based feature selection with a support vector machine classifier to reduce redundant variables and improve credit-card risk identification. Vorobyev and Krivitskaya (2022) proposed an automated rule-induction approach that integrates distributed tree-based models with expert rules, thereby reducing false positives in bank anti-fraud systems. More recently, Esenogho et al. (2022) combined an LSTM-based AdaBoost ensemble with hybrid resampling for credit-card fraud detection, while Huang et al. (2024) addressed class imbalance using clustering-based undersampling and a hybrid neural network.

These studies indicate that effective anti-fraud systems benefit from combining different modeling stages rather than relying on a single classifier. Feature selection can reduce dimensionality, resampling can mitigate class imbalance, and supervised ensembles can combine model predictions for fraud classification (Rtayli & Enneya, 2020; Esenogho et al., 2022; Huang et al., 2024). More broadly, anomaly-detection research has shown that supervised, unsupervised, semi-supervised, and hybrid approaches each have distinct advantages depending on data availability, fraud type, and operational requirements (Hilal et al., 2022).

A persistent challenge in fraud modeling is class imbalance, because fraudulent observations usually account for only a small proportion of financial transactions. Sundarkumar and Ravi (2015) addressed this problem through a hybrid undersampling approach and reported improvements across multiple classifiers. Sundarkumar et al. (2015) further proposed a one-class support vector machine–based undersampling method for churn prediction and insurance fraud detection. Li et al. (2021) examined early default identification in online lending through a cost-sensitive multi-layer learning framework, demonstrating that imbalance handling and differentiated misclassification costs should be incorporated into model design rather than treated merely as preprocessing choices.

Recent studies have also combined optimization algorithms, ensemble learning, and interpretable machine learning. Ding et al. (2025) integrated particle swarm optimization with XGBoost and SHAP-based interpretation to improve automobile insurance fraud detection and identify the variables contributing to model predictions. Esenogho et al. (2022) used an LSTM-based AdaBoost classifier with SMOTE-ENN resampling, whereas Huang et al. (2024) combined transaction and identity features with clustering-based undersampling and a hybrid neural network. These studies illustrate how feature engineering, class balancing, and predictive models can support credit-card fraud detection.

Graph-based methods have gained increasing attention because many forms of financial fraud involve coordinated transactions, interconnected accounts, and organized networks rather than isolated individual behavior. Usman et al. (2023) employed a graph-based machine learning model for anti-money-laundering fraud control in the financial domain, showing that relational information can support the identification of suspicious transactions even when the available feature set is limited. Graph-based analysis therefore provides a useful complement to conventional transaction-level classification by capturing relationships and risk transmission among entities.

Comprehensive reviews further reinforce the importance of integrating multiple analytical technologies. Al-Hashedi and Magalingam (2021), reviewing 75 studies published between 2009 and 2019, documented the growing application of data mining, machine learning, big-data analytics, and artificial intelligence in financial fraud detection. Hilal et al. (2022) similarly summarized recent advances in anomaly-detection methods and emphasized the continuing challenges associated with class imbalance, evolving fraud

patterns, model interpretability, data quality, and real-time deployment. Zhu et al. (2021) further highlighted the increasing importance of intelligent financial fraud detection in the post-pandemic environment.

However, much of the existing literature focuses on individual fraud scenarios, specific transaction types, or isolated stages of fraud detection. Integrated end-to-end systems designed specifically for telecom-fraud governance in banking environments remain comparatively limited. Against this background, the present study conceptualizes telecom-fraud risk governance as a lifecycle problem and proposes an integrated "pre-event, in-process, and post-event" framework. The framework is intended to establish a closed-loop system comprising preventive screening, real-time monitoring and early warning, and post-incident correlation analysis, thereby improving both the precision and operational efficiency of telecom-fraud prevention in banking contexts.

3. Risk Characteristics Analysis of Telecom Fraud Funds

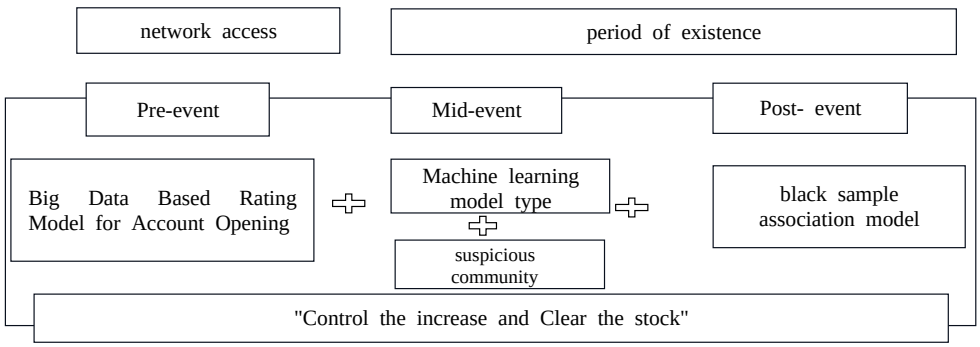
Telecom fraud schemes typically generate illicit proceeds that are routed through a multi-stage laundering chain using settlement accounts. The process can be summarized in two steps: (i) fraudulent acquisition of funds—commonly via bank transfers, wire transfers, or phishing-enabled credential theft—and (ii) rapid dispersion and concealment of the proceeds through successive transfers. Specifically, offenders often rely on personal accounts as relay nodes, performing layer-by-layer splitting and other structuring tactics to obscure the origin and traceability of the money flows.

From a risk-monitoring perspective, the transaction network formed in this process can be operationally partitioned into three functional layers: upstream accounts (initial collection), midstream accounts (high-frequency splitting/relaying), and downstream accounts (final cash-out or conversion), as illustrated in Figure 2. This tiered structure implies that effective detection should focus not only on single-account anomalies, but also on cross-account propagation patterns and relay density along the chain.

4. Pre-event Account-Opening Model Based on Big Data

This section develops a risk rating model for personal account opening by integrating internal and external data (e.g., customer information, card information, and relevant watchlists) to assign accounts into high-, medium-, and low-risk tiers based on model scores, thereby enabling differentiated due diligence and transaction controls. Figure 3 illustrates the rating approach, and Table 1 presents the qualitative list filtering criteria.

Figure 1. Risk Identification Path of Telecom Fraud



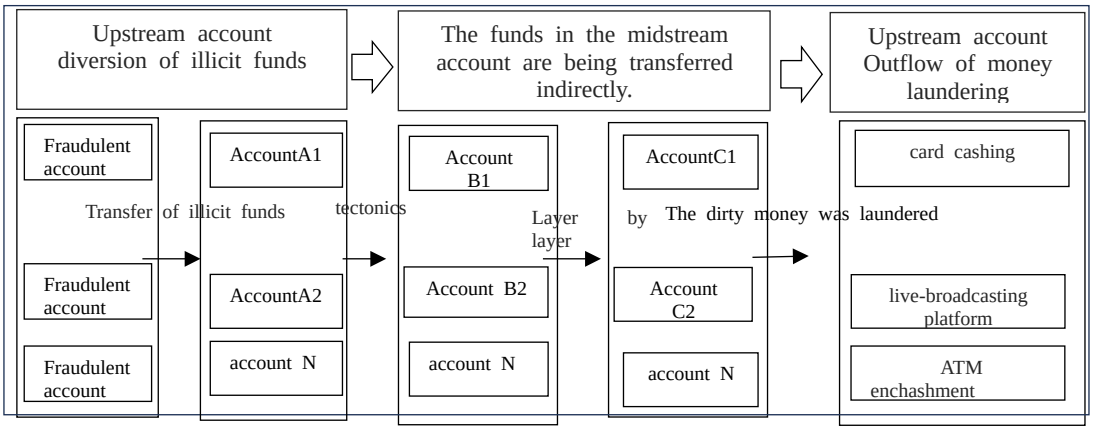


Figure 2. Flow of Funds from Telecom Fraud

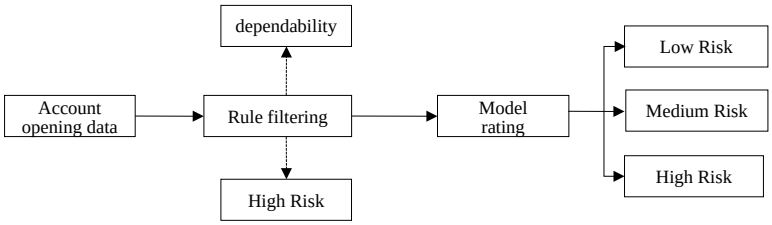


Figure 3. Rating Approach

Table 1. Qualitative List Filtering

Data source	Indicator type	Risk Rating indicator	Risk grade
External direct qualitative	Freezing and Locking date	Matching the List of Ministry of Public Security.	High risk
	Ministry of Public Security fraud data	National list of people involved in telecom fraud.	Black list
Internal data qualitative analysis	Customer tier	Diamond or Platinum cardholders	Dependability
	Internal VIP badge	VIP Customers	Dependability

4.1 Model Overview and Design

Building on prior work on risk-based evaluation and aligning with the Guidelines for Risk Assessment and Customer Classification Management of Financial Institutions in Anti-Money Laundering and Terrorist Financing, the model constructs a four-dimension scoring framework: (1) personal attributes, (2) financial operations, (3) geographic location, and (4) distribution channels.

Each sub-item is mapped to a five-level ordinal score (high → low, corresponding to 5, 4, 3, 2, 1), and the overall risk score aggregates weighted indicator scores within each dimension:

$$Score = \sum_{j=1}^{4} w_j \bullet \left(\sum_{i \in \Omega_j} V_i \bullet A_i \right) \quad (1)$$

where A_i is the indicator score, V_i is the indicator weight, and W_j is the dimension weight.

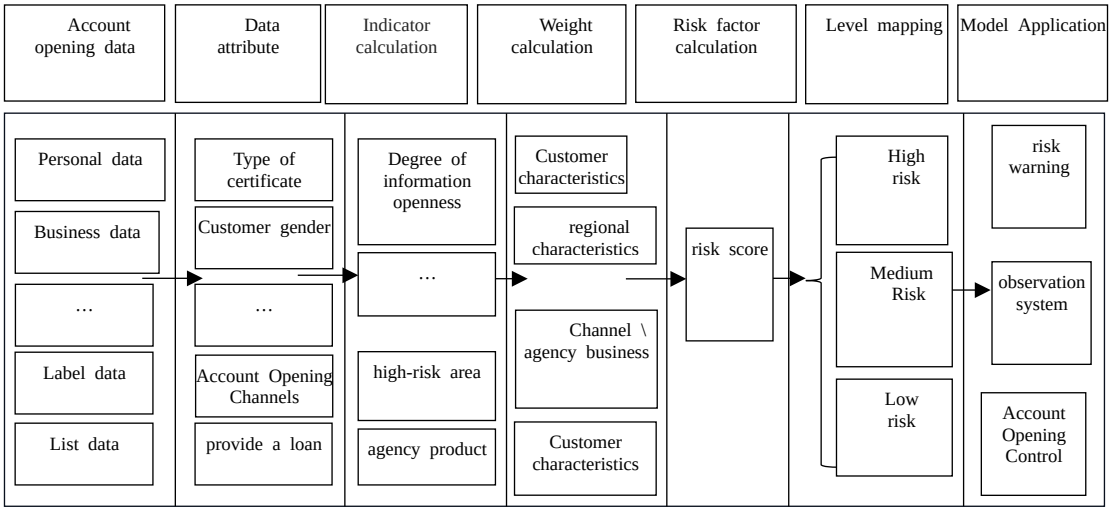


Figure 4. Scoring Path

4.2 Data Preparation

The study uses A Bank's new account-opening records from January–August 2023, split into a test set (2023-01-01 to 2023-05-31) and a validation set (2023-06-01 to 2023-08-31).

- Test set: 1.32 million full accounts, 235 black samples, 92,000 VIP accounts.
- Validation set: 370,000 full accounts, 96 black samples, 28,000 VIP accounts.

"Black samples" refer to accounts opened after 2023-01-01 that were later reported by public security authorities as involved in criminal activity; VIP data refer to platinum cardholders and above in the bank's hierarchy.

4.3 Weight Calculation and Model Performance Ranking

Dimension weights W_j are derived using the Analytic Hierarchy Process (AHP) via a four-order pairwise comparison matrix (Table 2) and its corresponding consistency-checked results (see Table 3). For implementation convenience, the final dimension weights are adjusted to 15%, 15%, 25%, and 45%. The final weights (15%, 15%, 25%, 45%) represent rounded approximations of the AHP-derived weights (14.11%, 14.11%, 26.31%, 45.47%), adjusted for operational simplicity while preserving the theoretically established priority hierarchy. The rounding maintains the consistency ratio within acceptable thresholds (CR < 0.1) and does not materially alter the model's discriminative power, as validated by the test set performance.

Table 2. Data on the Importance of the Four Dimensions

	Customer characteristics	**region**	**Agent and Channel**	**financing**
Customer characteristics	1	1	0.5	0.333
region	1	1	0.5	0.333
Agent and Channel	2	2	1	0.5
financing	3	3	2	1

Table 3. Results of the Hierarchical Analysis

	feature vector	**weighted value**	**principal eigenvalue**	**CI price**	**RI price**	**CR price**	**consistency check**
Customer characteristics	0.565	14.11%	4.01	0.003	0.89	0.004	pass through
region	0.565	14.11%					
Agent and Channel	1.052	26.31%					
financing	1.819	45.47%					

Indicator-level weights V_i are calibrated by comparing score distributions of black samples versus VIP accounts under each metric, and the discriminative power is grouped into four tiers (from no discrimination to excellent discrimination). The resulting indicator weights and scoring rules are summarized in Table 4:

Table 4. Indicator Weights and Scores

dimension	**A list of indicators**	**Sub-index directory**	**index weight**	**Indicator Details**	**value**
Customer characteristics	Customer Information Disclosure Level and Effectiveness	Can certificates be verified online	25%	yes	1
				deny	3
		The certificate has expired	25%	deny	1
				yes	3
		Gender and age	100%	17-55 years old and male	5
				17-55 years old and female	3
				Under 17, over 55 and male	2
				Under 17, over 55 and female	1
	channels for building or maintaining relationships with customers	Account opening or relationship change channel	25%	I handle it myself	1
				Batch account opening	3
				Agent account opening	5
	Related	The relevance	50%	The phone numbers are different	1

dimension	A list of indicators	Sub-index directory	index weight	Indicator Details	value
		between phone numbers and reserved addresses of different entities		The phone number or address is identical	3
				Phone number and address are the same	5
	circumstances	Is the legal representative of the enterprise or the income owner opening an account within the industry?	50%	yes	3
				deny	1
	Model warning situation	Model warning status (last six months)	75%	Unwarned count	1
				The model has triggered 0 times or issued 1 to 3 warnings in the past six months	3
				The model has triggered 3 times or issued more than 3 warnings in the past six months	5
regional and agency risk	Is it in a high-risk area	Is your hometown a high-risk area for phone scams?	75%	yes	5
				deny	1
		Open an account in another location	100%	yes	3
				deny	1
	agency transaction	Number of account opening requests processed	25%	0-3 times	1
				3-10 times	3
				More than 10 times	5
		Number of proxy transactions	25%	0 transactions	1
				1-5 transactions	2
				6-10 strokes	3
				11-15 transactions	4
				15 or more	5
financing	product	Status of electronic banking services activation	50%	Open online banking and mobile banking	5
				Open only one item	3
				No online banking or mobile banking services are available.	1
		Account status (last 3 months)	25%	Expenses exceed 3 times	3
				Expenses do not exceed 3 times	1
		Ladder status (last 6 months)	25%	no trade in the market	1
				one trade in which a stock is	3

dimension	A list of indicators	Sub-index directory	index weight	Indicator Details	value
				promoted or demoted	
				two or more swing trades	5
financing	product	Daily limit increase (last 6 months)	25%	no event day limit increase	1
				Daily limit increase for transactions with amounts up to 3,000 yuan	2
				Daily limit adjustment for transactions between 3,001 and 8,000 yuan	4
				Daily limit increase for transactions with amounts of 8,000 yuan or more	5
		temporary solution	75%	The reason for lifting control measures within the past three months is 3012.	5
				The restrictions were lifted within the past 3-6 months, with the reason being 3012.	4
				The reason for lifting restrictions in the past 6-12 months is 3012	3
				For cases with one year or longer of uncontrolled status, the reason is 3012	1
		Number of cards in the row	25%	One or fewer bank card records	5
				2-3 bank card records	4
				4-5 bank card records	2
				More than 5 bank card records	1
		loan documentation	75%	You have a loan and no overdue payments	1
				No loan history	3
				The loan is currently overdue	5
		Account Opening Channels	100%	counter ITM	1
				Mobile Business Development	3
				Personal Bank, Mobile Bank, WeChat Bank, External System	5

The model further transforms raw scores into a standardized score using min–max normalization, with the normalization base finalized at max = 11.5 and min = 3.5. Risk levels are then mapped as high-risk (65–100), medium-risk (35–65), and low-risk (0–35).

4.4 Model Evaluation and Validation

Validation runs on the test set confirm the model's segmentation effectiveness and stability. The reported results show that the model achieves 53.37% coverage of black samples within the high-risk tier while maintaining 80.74% coverage of VIP accounts within the low-risk tier, supporting practical deployment for front-end screening and tiered controls.

Performance gains from the training to validation period are attributed to the inclusion of new metrics (e.g., cross-regional account deactivation, quota adjustments, and model alert frequency), which become more salient in later periods and improve discrimination.

5. In-process Machine-Learning Modeling Based on LightGBM and Suspicious Community Detection Using LPA

5.1 Machine learning model based on LightGBM

5.1.1 Data Preparation

The machine-learning model takes A Bank accounts as the prediction target, where accounts later confirmed in public-security notifications are treated as black samples. The modeling adopts a "day + account" granularity. The training set spans September 2022 to July 2023, and the test set covers August 2023. The training set contains 806,383 white samples and 139,002 black samples (total 945,385), while the test set contains 92,581 white samples and 2,397 black samples (total 94,978).

5.1.2 Construction of the Risk Indicator System

The personal telecom-fraud model organizes indicators into five dimensions:

(1) Basic statistical features (e.g., total transaction amount/count; credit/debit amount and count);

(2) Special amount patterns (e.g., rounded-amount transactions, 9*-ending** funds, ratios and counts of special debit/credit amounts);

(3) Zero-balance indicators (e.g., low-balance behavior such as decimal-amount frequency, balance < 10 yuan, and days with balance < 10 yuan);

(4) Abnormal remarks (e.g., note patterns such as "member numbers + letters," "digital currency + letters," "margin financing deposits," and "game top-ups");

(5) Transaction timing patterns (e.g., nighttime frequency and share, and the share of transactions < 10 yuan).

5.1.3 Feature selection

To mitigate overfitting from high-dimensional inputs, a wrapper-based feature-selection strategy is used. After removing low-quality features, the study applies single-feature AUC screening: because an AUC of 0.5 implies no better-than-random discrimination, features with single-feature AUC < 0.51 are excluded. The final model retains 200 features, with the overall procedure summarized in Figure 5.

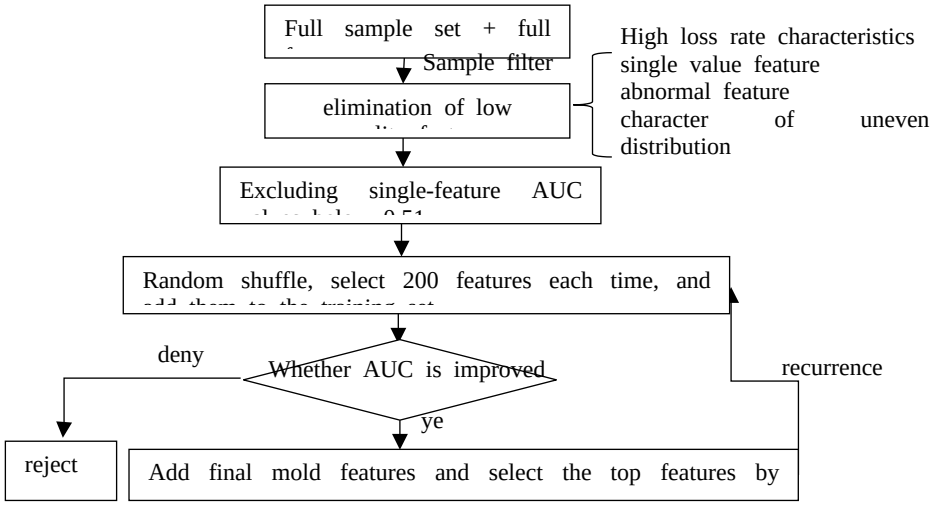


Figure 5. Machine Learning Model Feature Selection Process

5.1.4 Model Training and Evaluation

The modeling uses 945,385 training samples, split 8:2 into a training set (756,308) and a validation set (189,077) for validation. Model evaluation emphasizes validity (see Table 5) and stability (see Table 6).

In the validity assessment, the key focus is whether the model can capture more fraudulent accounts while keeping false positives at an acceptable level. For stability, the PSI is used; PSI values within the 0.1 range indicate strong stability between training and validation distributions.

Table 5. Performance of Machine Learning Models on Test Sets

top	thresholds	Precision	recall	F1_score
100	0.99	0.97	0.08	0.14
200	0.98	0.97	0.14	0.25
500	0.95	0.94	0.28	0.43
1000	0.91	0.65	0.49	0.48
1500	0.87	0.45	0.53	0.44
2000	0.84	0.34	0.56	0.39

Table 6. The Machine Learning Model's Training and Validation Sets (PSI)

probability distribution	**training set**		**test set**	
probability value	Precision	psi	Precision	psi
(0.0, 0.1]	0.3301	0.0000	0.3328	0.0000
(0.1, 0.2]	0.1725	0.0000	0.1747	0.0000
(0.2, 0.3]	0.1145	0.0003	0.1205	0.0003
(0.3, 0.4]	0.0797	0.0004	0.0855	0.0004
(0.4, 0.5]	0.0595	0.0005	0.0650	0.0005

probability distribution	training set		test set	
(0.5, 0.6]	0.0499	0.0000	0.0509	0.0000
(0.6, 0.7]	0.0454	0.0000	0.0441	0.0000
(0.7, 0.8]	0.0466	0.0008	0.0409	0.0008
(0.8, 0.9]	0.0549	0.0016	0.0459	0.0016
(0.9, 1.0]	0.0469	0.0012	0.0396	0.0012

To reduce future tuning costs while maintaining long-run effectiveness, the system is designed for dynamic updating: it retrains on the latest black samples after model updates and reorders the 200 input features to adapt to emerging fraud behaviors and transfer tactics. Empirically, from October 2023 to February 2024, the model maintained over 47% coverage, supporting rapid and automated detection of telecom-fraud risks from transaction and account data.

5.2 Suspicious community detection using the LPA algorithm

5.2.1 Model Principle and Rationale

The Label Propagation Algorithm (LPA) is a graph clustering method (Raghavan et al., 2007). It proceeds as follows: (1) initialize each node with a unique label; (2) iteratively update each node's label to match the most frequent label among its neighbors (with random tie-breaking); and (3) stop when nodes reach a majority-label consensus, at which point nodes sharing a label form a community.

Because LPA scales well (near-linear time complexity) and does not require pre-specifying the number of communities, it is suitable for large, complex transaction networks. After community mining, the framework allows designing monitoring metrics to trigger early warnings for high-risk groups. The pipeline comprises four steps: node–edge design, data processing, graph construction, and group mining (Figure 6).

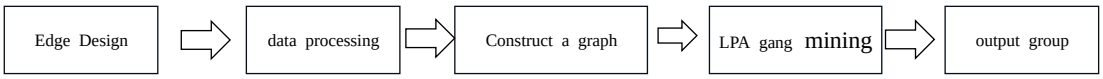


Figure 6. The Overall Concept of the Group Model

5.2.2 Node–edge design

In the transaction relationship graph, nodes represent accounts/cards, and edges represent transaction relationships. Edge direction can encode fund flow (e.g., from the involved account to the counterparty). For visualization, card numbers are used as node labels; edges carry behavioral attributes such as cumulative transaction count, cumulative transaction amount, and active transaction days.

5.2.3 Data Processing

With node–edge definitions in place, the model traces transaction counterparts of involved accounts through direct and second-degree counterparties. If account A transacts with B, C, and D, then {B, C, D} are A's first-degree counterparties; if B further transacts with {F, G}, then {F, G} become A's second-degree related counterparties (Figure 7).

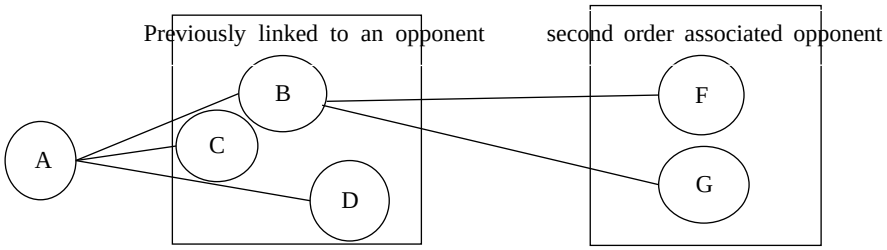


Figure 7. Association Diagram of the Suspicious Community

5.2.4 LPA Algorithm for Mining and Outputting Risk Groups

Using the full set of black samples, point–edge processing on second-degree counterparties produced 691,511 entries. Defining a gang as having at least three members yielded 5,232 gangs. After applying the risk strategy—each gang must include ≥2 public-security-reported black samples, have ≥5 members, contain ≥3 bank-issued cards, and include ≥1 nationally reported black sample—34 suspicious gangs remained. This produced 1,680 machine-learning pre-warning entries and 1,354 newly added accounts. After municipal and branch verification, 379 risk accounts were confirmed, giving a supplementary confirmation rate of 27.99%.

6. Post-event association analysis model based on knowledge graph

Telecom fraud is typically orchestrated by organized syndicates, whose accounts and actors are connected through transaction links and, in many cases, shared physical or digital devices used during account operations. To support post-event tracing and deeper exposure of hidden networks, we leverage the account information disclosed by public security authorities to build a graph-based correlation model that integrates both static and dynamic signals (see Figure 8). Static correlation focuses on identity and account attributes, while dynamic correlation captures transaction behavior and fund-flow patterns for relationship discovery and risk confirmation.

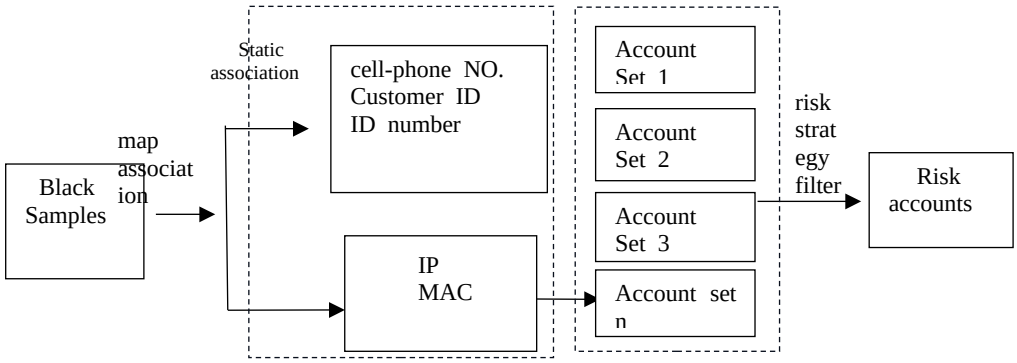


Figure 8. Conceptual Diagram of the Association Model

6.1 Account and Non-account Association Model

We first construct an account–non-account relationship knowledge graph by extracting key identifiers—such as ID number, mobile number, and customer ID—and performing correlation analysis between accounts and these non-account entities. The resulting model contains six feature families and nine indicators (see Table 7).

Using these account–non-account association rules, we conducted a zero-based scan verification across all accounts as of September 20, 2023. The results (see Table 8) show that the risk confirmation rate for account–non-account associations—especially those involving ID-number linkage and agent-assisted account opening—reached 52.45%. For accounts flagged by ID-related risk signals, the bank can deploy preemptive controls (e.g., restricting daily transaction frequency, total transaction amount, and channels) before manual review. With timely verification and transaction restriction, high-risk accounts can be managed earlier, thereby curbing telecom-fraud risk more effectively.

Table 7. Non-Account Relationship Association Diagram Rules

order number	rule of graph	Rule indicator
feature_01	The same ID number as the fraudulent account	Other accounts under the name of the black sample customer
		The ID number matches that of a fraudulent account from another bank.
feature_02	The phone number is the same as the account involved in the scam.	The account phone number matches the number linked to the scam account
		The number of linked accounts with the same phone number in the bank is <=5
feature_03	Open accounts in the same batch as fraudulent accounts	The batch size of accounts opened in bulk matches the black sample
feature_04	And the accounts involved in fraud, along with the agents...	The agent who handles the sample of the proxy account opening
		The account opening agent is the same as the

order number	rule of graph	Rule indicator
		agent handling the fraudulent account.
feature_05	Same phone number	Accounts with phone numbers matching those in the public security report
feature_06	The phone number matches the ID number	Accounts with the same phone number and ID number as those reported by the police

Table 8. Performance of the Results Regarding Non-Account Relationship Associations

rule of graph	Linked accounts	amount of risk recognition	risk confirmation rate
Linking the ID number of the fraudulent account	10240	5371	52.45%
Linking phone numbers to accounts involved in fraud	24175	3684	15.24%
Association of agents with accounts involved in fraud	47	19	40.43%
Batch account creation and linking for fraudulent accounts	8	3	37.50%
Report mobile number association	1752	539	30.76%
Link the phone number to the account holder's ID number	8580	952	11.10%

6.2 Account and Account Relationship Association Model

The account-to-account relationship model targets the deeper structure of fraudulent capital chains, aiming to interrupt fund-flow pathways by identifying and prioritizing suspicious accounts. Based on account-chain relationships, the system generates a risk list database, then applies risk-strategy filtering to output an actionable warning list.

Suspected List Database Generation. The database can be generated through (i) direct transaction relationships, or (ii) mining black samples and their second-order transactional relationships. For direct transactions, the system primarily defines accounts using their transaction-based linkage rules (see Table 9).

Linking an account to its counterparties (see Figure 9), a one-way arrow denotes a one-directional fund link over the past 7 days, while a two-way arrow indicates mutual transfers. Brackets denote different account-set groupings used for screening and policy execution.

Experimental evidence indicates that, from over 1.2 million accounts on the anti-fraud blacklist, the model established a risk-list database of 190,000 entries through account-relationship analysis. After applying risk-policy filters, 13,861 accounts were alerted; following bank verification, 8,733 were confirmed as high risk, yielding a 63.00% confirmation rate.

Table 9. Logic for Generating the List Database

Association rule name	order number	metric
direct correlation	rule_01	Number of transactions in fraudulent personal accounts across the industry in the past 30 days
	rule_02	The number of counterparties in the bank's blacklisted sample from the past 30 days
indirect linkage	ind_01	The total number of transactions between the account and its personal counterparties over the past 7 days
	ind_02	The total transaction amount between the account and its personal counterparties over the past 7 days
	ind_03	Total credit transactions between the account and its personal counterparties over the past 7 days
	ind_04	Cumulative credit transactions between the account and its personal counterparties over the past 7 days
	ind_05	The cumulative number of debit transactions between the account and its personal counterparties over the past 7 days
	ind_06	The cumulative debit transactions between the account and its personal counterparties over the past 7 days

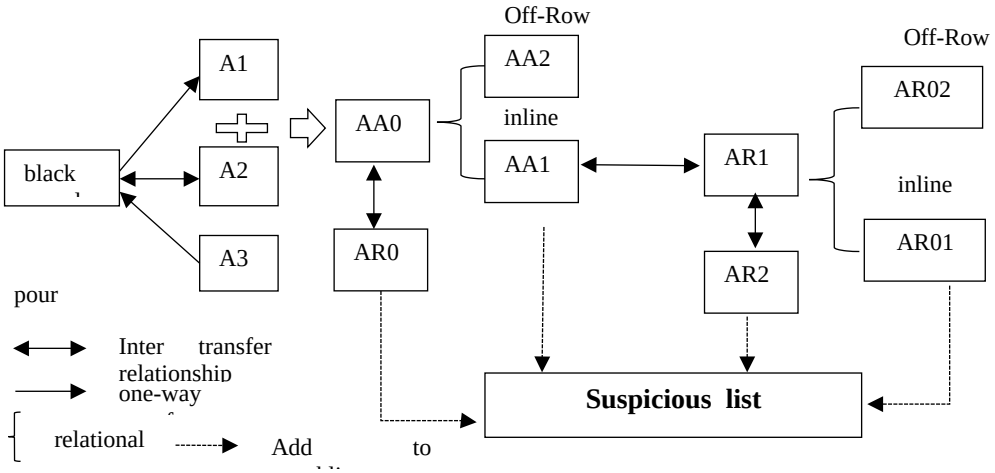


Figure 9. Suspicious Account Identification Workflow Based on Black Sample Account Diffusion

A1: Downstream account set of the black sample

A2: Upstream and downstream account set of the black sample

A3: Upstream account set of the black sample

AA0: The account set composed of A1 and A2.

AA1: The off-balance-sheet account set in AA0.

AR0: The upstream and downstream account set with AA0

AR1: The upstream and downstream account set with AA1

AR01: The intra-bank account set within AR1

AR2: A set of accounts that are upstream and downstream to AR1.

AR02: A set of accounts outside the bank in AR1.

Acknowledgement

The authors sincerely appreciate the three anonymous reviewers for their targeted and constructive comments, which have greatly improved the rigor of the system architecture design, the clarity of the empirical analysis framework, and the practical application value of the research conclusions. We also extend our gratitude to Guangzhou Huashang College and Shenhua AI Research Institute for providing the research environment, computing support and institutional conditions necessary for this work. Special thanks go to the partner commercial bank for granting access to the production-level transaction dataset that supports the empirical verification of this study. We also recognize the editorial team of the journal for their efficient and professional work throughout the submission and review process.

Author Contributions

Conceptualization, Z.M.C.; Methodology, Z.M.C.; Formal Analysis, Z.M.C.; Investigation, Z.M.C. and Z.Y.L. and T.Q.; Data Curation, Z.M.C. and Z.Y.L. and T.Q.; Software, Z.M.C. and Z.Y.L. and T.Q.; Validation, Z.M.C. and Z.Y.L. and T.Q.; Resources, Z.Y.L. and T.Q.; Visualization, Z.M.C. and Z.Y.L. and T.Q.; Writing – Original Draft, Z.M.C.; Writing – Review & Editing, Z.M.C. and Z.Y.L. and T.Q.; Supervision, Z.M.C.; Project Administration, Z.M.C. We have read and approved the final manuscript and take responsibility for the work.

Funding

This research did not receive any financial support from external funding programs.

Conflicts of Interest

The authors declare that there are no financial or professional conflicts of interest that could affect the objectivity of the research design, data analysis, or conclusion presentation.

Ethical Approval and Consent

This study was conducted using desensitized, aggregated banking transaction data. It does not involve human subjects, identifiable personal privacy information, or animal experiments, and thus does not require ethical review or informed consent procedures.

Data Availability Statement

The core indicator systems, model parameters and performance metrics supporting the findings of this study are all presented in the tables of the main text. The raw transaction dataset used in the research belongs to the partner commercial bank and cannot be publicly disclosed due to commercial confidentiality agreements. The analysis code for the model framework can be requested from the corresponding author on a reasonable basis, subject to compliance with data use specifications.

Declaration of Generative AI Use

Generative artificial intelligence tools were used to assist with language polishing of the manuscript text and the initial drafting of schematic diagrams, including the risk identification path diagram, fund flow schematic and model architecture diagrams. All AI-generated content has been carefully reviewed and verified by the authors, who take full responsibility for the final content of the manuscript.

References

Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review, 40, 100402. doi: 10.1016/j.cosrev.2021.100402

Ding, N., Ruan, X., Wang, H., & Liu, Y. (2025). Automobile insurance fraud detection based on PSO-XGBoost model and interpretable machine learning method. Insurance: Mathematics and Economics, 120, 51–60. doi: 10.1016/j.insmatheco.2024.11.006.

Esenogho, E., Mienye, I. D., Swart, T. G., Aruleba, K., & Obaido, G. (2022). A neural network ensemble with feature engineering for improved credit card fraud detection. IEEE Access, 10, 16400–16407. doi: 10.1109/ACCESS.2022.3148298.

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. Expert Systems with Applications, 193, 116429. doi: 10.1016/j.eswa.2021.116429.

Huang, H., Liu, B., Xue, X., Cao, J., & Chen, X. (2024). Imbalanced credit card fraud detection data: A solution based on hybrid neural network and clustering-based undersampling technique. Applied Soft Computing, 154, 111368. doi: 10.1016/j.asoc.2024.111368.

Li, Z., Zhang, J., Yao, X., & Kou, G. (2021). How to identify early defaults in online lending: A cost-sensitive multi-layer learning framework. Knowledge-Based Systems, 221, 106963. doi: 10.1016/j.knosys.2021.106963.

Rtayli, N., & Enneya, N. (2020). Selection features and support vector machine for credit card risk identification. Procedia Manufacturing, 46, 941–948. doi: 10.1016/j.promfg.2020.05.012.

Sundarkumar, G. G., & Ravi, V. (2015). A novel hybrid undersampling method for mining unbalanced datasets in banking and insurance. Engineering Applications of Artificial Intelligence, 37, 368–377. doi: 10.1016/j.engappai.2014.09.019.

Sundarkumar, G. G., Ravi, V., & Siddeshwar, V. (2015). One-class support vector machine based undersampling: Application to churn prediction and insurance fraud detection. In 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) (pp. 1–7). IEEE. doi: 10.1109/ICCIC.2015.7435726.

Usman, A., Naveed, N., & Munawar, S. (2023). Intelligent anti-money laundering fraud control using graph-based machine learning model for the financial domain. Journal of Cases on Information Technology, 25(1), 1–20. doi: 10.4018/JCIT.316665.

Vorobyev, I., & Krivitskaya, A. (2022). Reducing false positives in bank anti-fraud systems based on rule induction in distributed tree-based models. Computers & Security, 120, 102786. doi: 10.1016/j.cose.2022.102786.

Zhu, X., Ao, X., Qin, Z., Chang, Y., Liu, Y., He, Q., & Li, J. (2021). Intelligent financial fraud detection practices in post-pandemic era. The Innovation, 2(4), 100176. doi: 10.1016/j.xinn.2021.100176.