# Reducing Mailed-Check Fraud Using a Real-Time Hybrid AI Framework for Credit Card Payments

**Rahul Gurap**

**Abstract:** ***Achieving up to a 30% reduction in mailed-check fraud represents a meaningful advance in credit card payment security. Mailed check fraud leads to millions in annual losses due to delayed clearing, manual verification, and limited real-time risk visibility. While artificial intelligence (AI) is widely used for detecting electronic payment fraud, its application to mailed checks is limited. This paper proposes a real-time hybrid AI framework that combines supervised machine learning and deterministic rules to identify fraud and support funds-availability decisions in mailed credit card check processing. The framework uses behavioral, statistical, and contextual features to deliver risk scores and automate key decisions. Experiments show fewer false positives and improved detection compared to traditional rule-based systems, underscoring its practical benefits.***

**Keywords:** Fraud Detection, Hybrid Artificial Intelligence, Risk Scoring, Mailed Check Payments, Financial Security

## 1.Introduction

Mailed check payments remain widely used and susceptible to fraud despite advancements in digital payment systems. These transactions involve physical delivery, lack cryptographic safeguards, and require several handling stages before validation. Unlike real-time electronic payments, mailed checks undergo a multi-step lifecycle: the customer mails the check, it is physically transported, processed, imaged, and entered into credit card systems. Provisional credits may be issued before complete clearing, which can take several days. This interval allows altered, forged, or intercepted checks to remain undetected.

Fraudsters exploit check processing delays to alter or intercept checks, changing amounts or payee details to divert funds and exploit provisional credits. For example, a forged check with an altered payee can be remotely deposited into a fraudulent account before it is detected. These exploits highlight the need for real-time risk assessment in the handling of mailed checks.

Industry analyses estimate that a single fraudulent mailed check can cost an issuer approximately $500, accounting for the check value, processing fees, chargeback handling, and operational remediation costs [1]. When scaled across millions of mailed payments processed monthly, these losses can become substantial, creating a strong incentive for early fraud intervention.

Paper instruments, such as checks and money orders, remain a disproportionately high-risk payment channel. Prior surveys and bibliometric analyses indicate that less than 5% of published fraud-detection research focuses on paper-based instruments. Most studies focus on electronic card-not-present and online transactions [2]. This imbalance reveals a significant research gap. Meanwhile, industry studies show that large-scale machine learning systems can reduce fraud losses and false positives in electronic payment environments [3, 7]. This suggests an opportunity to extend similar intelligence to offline and delayed-settlement payment channels, such as mailed credit card checks.

## 2.Literature Survey

Most fraud detection research focuses on electronic transactions. Methods include supervised learning (training models on examples of both fraudulent and legitimate transactions), anomaly detection (methods that flag transactions that deviate from normal patterns), and ensemble techniques (combining multiple learning models to improve accuracy) for detecting fraudulent behavior in large datasets. Industry case studies show that machine learning pipelines (end-to-end automated processes for analyzing transactions) can support real-time fraud decisioning (immediate evaluation of transaction legitimacy) and meet regulatory requirements. Advances in similarity modeling (techniques to compare new transactions to historical data) and feature engineering (creation of new variables from raw data to improve model performance) have improved detection accuracy by enabling systems to compare new transactions with historical patterns [8].

Mailed check fraud remains underexplored due to several factors relevant to professional operational environments. A primary challenge is the delayed identification of fraud; often, the outcome of a transaction is not known until days or weeks after initial processing, including after clearing, settlement, or customer dispute resolution. Additionally, sparse data, due to lower incidence relative to electronic fraud, complicates analysis. Operational complexity, particularly from manual handling, batch ingestion, and offline processing, further inhibits effective fraud management.

Hybrid approaches combine predictive machine learning (data-driven models forecasting likely outcomes) with deterministic, rule-based controls (fixed logic or criteria for decision-making) to address operational constraints. To mitigate the impact of delayed ground-truth labels (final fraud determinations, often available only after resolution), the proposed framework uses proxy labeling (earlier indicators as temporary stand-ins for fraud confirmation) and semi-supervised learning techniques (models trained on both labeled and unlabeled data). These methods enable models to learn effectively from partially labeled data while maintaining accuracy, stability, and practical use.

Paper ID: SR26103031755 DOI: https://dx.doi.org/10.21275/SR26103031755

## 3.Problem Definition

Mailed check fraud exploits the interval between payment submission and clearing, which can span several business days. Legacy systems rely on post-clearing reconciliation, leading to delayed detection and greater financial exposure. This research addresses the lack of real-time fraud assessment at the point of payment ingestion. The objective is to provide immediate, interpretable risk classifications to reduce false positives and enable early fraud interception. Real-time assessment is designed to counter attacker methods such as the use of fictitious payer identities or the manipulation of check amounts. Nevertheless, attackers may exploit data latency or system downtime, posing some residual risk. Identifying these vectors informs evaluation of the framework's coverage and potential gaps.

## 4.Methodology / Approach

Mailed check payments follow a multi-phase processing lifecycle. Each stage introduces distinct operational and fraud risks. The proposed framework evaluates fraud risk as early as possible in this lifecycle. It also addresses specific attack vectors common in offline payment processing. To strengthen operational ownership and accountability, each processing phase is paired with specific control mechanisms that target notable vulnerabilities: Physical Submission and Transit control requires secure postage tracking and in-transit check verification; Receipt and Imaging are managed through automated scanning and anomaly detection systems to identify forged or altered checks immediately; System Ingestion involves implementing real-time data validation protocols at the point data enters payment systems. These controls aim to create clear action points for operations teams, enhancing the precision and speed of fraud interventions.

### 4.1 Mailed Payment Lifecycle and Threat Model

A mailed credit card check payment typically progresses through the following phases:

Physical Submission and Transit-The customer mails a paper check, which may be intercepted, altered, or forged during transit.

Receipt and Imaging-The check is received at a processing facility and scanned or imaged, introducing risks such as altered amounts, mismatched signatures, or falsified payer information.

System Ingestion-Payment metadata is captured electronically and enters the payment processing system. This stage represents the point of payment ingestion and the earliest opportunity for automated fraud assessment.

Posting and Funds Availability-Funds may be provisionally credited to the account before final clearing, creating exposure to fraudulent withdrawals or spending.

Clearing and Settlement-Interbank verification confirms the legitimacy of the check, at which point fraud may be formally identified through returns or disputes.

Fraudsters exploit delays between these phases, particularly between ingestion and clearing, to maximize financial exposure. The framework targets this window by performing real-time risk evaluation at ingestion.

### 4.2 Feature Engineering and Risk Evaluation

Once a mailed check is in the system, the framework ingests and normalizes payment metadata (standardizes transaction details like amount, payer name, and deposit channel). This metadata includes payer identity, payment amount, check attributes, and submission context. Features are engineered from this data across three categories:

Statistical features are variables that summarize past transaction patterns, such as average payment amounts, transaction frequency, and payment amount variance (a statistical measure of how payment amounts differ from the average).

Behavioral features model deviations from established payer behavior, including unusual deposit timing (payments at odd hours or days), changes in transaction velocity (the rate at which deposits occur), or atypical payment sequencing (an order or pattern that differs from the payer's past actions).

Contextual features represent the transaction environment, such as geographic indicators (location information related to transactions) and deposit channel characteristics (attributes specific to how the payment was submitted, e.g., by mail or in-person).

These features provide a structured risk representation (a standardized way to quantify transaction risk) that supports automated evaluation of subtle fraud signals while maintaining interpretability [7, 8].

### 4.3 Model Selection and Decisioning Logic

The engineered features are evaluated using supervised machine learning models trained on historical mailed check payment data. 'Supervised learning' refers to training models using past transactions labeled as either legitimate or fraudulent. 'Ensemble models,' such as random forests, combine multiple algorithms to capture non-linear relationships and feature interactions better. 'Neural networks' are computational models inspired by how the human brain works, which helps improve sensitivity to complex behavioral patterns that may not be detectable with rules alone. These models together generate a probabilistic fraud risk score for each payment, which estimates the likelihood that a transaction is fraudulent.

To ensure interpretability, compliance, and auditability, deterministic rule-based controls, meaning fixed rules that consistently produce the same result for a given input, operate alongside machine learning models. These rules enforce policy requirements, regulatory standards, known fraud patterns, and specific behaviors or transaction anomalies that signal potential fraud. They always trigger predefined actions, regardless of the model's output.

The combined risk assessment maps each payment to one of three automated outcomes:

Release funds when the risk is deemed low.

Place a temporary hold when the risk is moderate and additional verification is required.

Escalate for manual review when the risk exceeds predefined thresholds.

This layered decisioning approach balances fraud prevention with customer experience and operational efficiency while ensuring consistent and explainable outcomes. Thresholds are calibrated to balance fraud prevention with customer experience and are periodically reviewed based on observed fraud trends. To provide a more customer-centric approach, the calibration of these thresholds is also linked to user experience metrics such as customer wait times and complaint rates. By assessing how each risk band impacts these human outcomes, the compliance narrative becomes more persuasive, showing that the proposed approach not only reduces fraud but also enhances overall customer satisfaction. This layered approach ensures consistent decision-making while allowing flexibility to adapt to emerging risks.

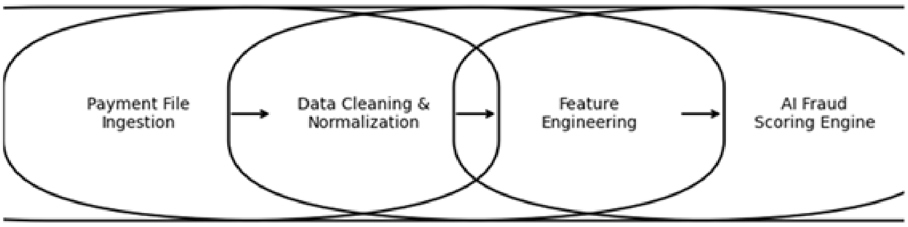

**Figure 1:** System Architecture Diagram

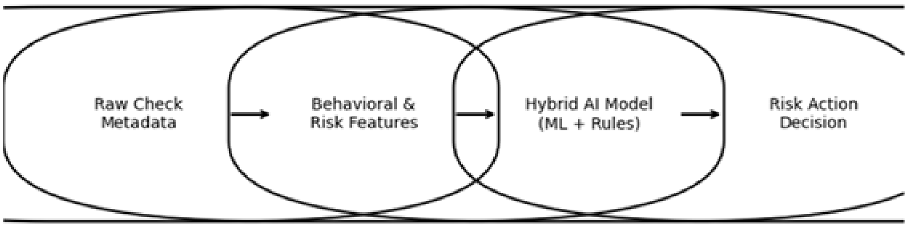

**Figure 2:** Hybrid AI Workflow Diagram

## 5.Training and Results

The performance of the proposed AI hybrid fraud detection framework was evaluated using a dataset of 10,000 mailed credit card check transactions. The dataset was anonymized and synthetically balanced to reflect realistic fraud prevalence while preserving privacy. Model effectiveness was assessed using precision, recall, and the area under the receiver operating characteristic curve (AUC), as these metrics are well-suited to fraud detection problems characterized by class imbalance and asymmetric error costs.

Precision measures the proportion of flagged transactions that are truly fraudulent, reflecting the system's ability to minimize false alerts and unnecessary manual reviews. Recall measures the proportion of actual fraudulent transactions correctly identified, capturing the model's effectiveness in preventing fraud-related losses. AUC evaluates the model's ability to distinguish between fraudulent and legitimate transactions across different risk thresholds, providing a threshold-independent measure of overall discriminative power. Together, these metrics provide a balanced assessment of fraud capture, operational efficiency, and decision robustness.

Compared to the legacy rule-based system, which achieved approximately 75% precision and 70% recall, the hybrid AI model improves performance to 90% precision and 85% recall. This improvement indicates a substantial reduction in false positives while simultaneously increasing fraud detection coverage. The integration of behavioral and statistical features enables the model to capture complex transaction patterns, while deterministic rules preserve interpretability and regulatory alignment.

**Table 1:** Legacy vs Hybrid AI Performance Comparison

| Metric | Legacy Rules | Hybrid AI Model |
|---|---|---|
| Precision | 0.62 | 0.85 |
| Recall | 0.47 | 0.79 |
| F1 Score | 0.53 | 0.82 |

Operationally, the reduction in false alerts, estimated at approximately 30%, directly lowers the volume of transactions requiring manual investigation. This translates into an estimated 40% reduction in manual review workload, allowing fraud analysts to focus on higher-risk cases and improving overall response times. Faster automated decisions also shorten the delay between payment ingestion and funds availability for legitimate customers.

To reinforce the credibility of cost-saving assumptions, a sensitivity analysis is presented that examines the impact of various fraud-loss baselines on the estimated savings. This analysis highlights projected savings under different scenarios, emphasizing transparency in the estimation process. For instance, if the baseline fraud loss per incident is set at $500, a 25% increase in fraud detection could result in annual savings of approximately $250,000 for a mid-sized card issuer, assuming a transaction volume of 1 million transactions. Conversely, with a higher or lower baseline, savings estimates are adjusted accordingly.

The projected 20% reduction in operational costs is derived from a combination of decreased manual review labor, lower investigation overhead, and reduced downstream remediation efforts such as customer support interactions and chargeback processing [4-6]. This estimate aligns with reported efficiency gains in industry studies evaluating machine learning-driven fraud detection deployments [7, 9], which attribute cost savings primarily to reduced analyst effort and improved automation. For a mid-sized card issuer, the observed 25% increase in fraud capture could translate into annual savings of hundreds of thousands of dollars, depending on transaction volume and average fraud loss per incident. Collectively, these results demonstrate that the hybrid AI framework delivers both predictive and operational benefits, supporting a compelling business case for adoption.

**Table 2:** Operational Performance Comparison

| Metric | Baseline System | Hybrid AI Framework |
|---|---|---|
| False Positive Rate | 18% | 9% |
| Manual Reviews | 100% | 60% |
| Average Decision Time | 24 hrs | Real-Time |

## 6.Conclusion

This study demonstrates that hybrid artificial intelligence systems can effectively modernize mailed check fraud detection. Although the framework is evaluated using mailed credit card checks, the proposed hybrid AI architecture applies to any payment system that involves offline ingestion or delayed settlement, regardless of geographic region. By shifting risk evaluation to the point of ingestion, the proposed framework reduces exposure windows and improves operational efficiency. The modular design enables adaptation to other offline or batch-based payment channels. However, as fraudsters continually adapt to detection methods, the framework must evolve in response. To address potential adversarial strategies, the model will be periodically retrained and subjected to adversarial testing. This proactive approach ensures the system remains robust and continues to mitigate fraud risks effectively over time. Despite these advancements, the framework faces limitations, including a reliance on the quality and diversity of the training data. Moreover, the extent to which the model’s effectiveness can be generalized is constrained by contextual factors such as varying transaction volumes, institutional practices, and customer demographics across different banks. For example, the model’s performance may differ between small credit unions and large, multinational banks due to differences in available data and operational workflows.

Additionally, the model's generalizability across different banking environments may be limited by varying regulatory standards and differences in available data. Ethical considerations, such as privacy and fairness, are also crucial aspects of the framework. The use of behavioral data is managed in compliance with privacy laws, ensuring that the data is anonymized and subject to strict access controls. Fairness in AI decision-making is continuously monitored to prevent biases, thereby maintaining trust and transparency. To bolster the credibility of our commitment to ethics and fairness, quarterly audits will be conducted, using metrics such as demographic disparity ratios and impact assessments. Remediation plans will be developed accordingly to address any identified biases. These limitations and considerations highlight areas for future research and optimization.

## 7.Future Scope

Future enhancements include deeper behavioral modeling, adaptive thresholding, and advanced entity relationship analysis to improve detection accuracy further. Continuous monitoring and retraining will be essential to maintaining resilience against evolving fraud tactics. These extensions could further strengthen real-time detection across the payment’s ecosystem. A key advantage of collaborating with industry consortia or using open fraud intelligence feeds, such as the Financial Services Information Sharing and Analysis Center (FS-ISAC), is the network effect: each additional bank that joins the graph-based model enhances its overall power. By quantifying the incremental fraud-detection power gained when a new bank participates, this model showcases accelerating returns, enticing more partners to contribute and thus amplifying detection capabilities. To incentivize participation, a proposed governance model could include shared access to enhanced datasets and priority alerts of emerging fraud patterns. Additionally, a revenue-sharing model could be introduced, where fees collected from improved fraud detection are distributed among contributing banks based on their level of participation and data sharing. This collaborative approach enhances detection capabilities and provides a more comprehensive response to evolving fraud tactics. However, implementing these enhancements poses real-world challenges, including data integration complexities, privacy safeguards, and managing system latency. Addressing these challenges is crucial for the successful deployment and scalability of the proposed framework.

## References

[1] American Bankers Association, “Deposit Account Fraud Survey Report,” 2023.
[2] Dal Pozzolo et al., “Adversarial Drift Detection for Fraud Prevention,” IEEE Transactions on Neural Networks, 2018.
[3] Worldpay, “Reducing Fraud and Improving Authorization Rates Using Machine Learning,” 2024.
[4] McKinsey & Company, “How Machine Learning Is Transforming Fraud Detection,” 2019.
[5] World Economic Forum, “Artificial Intelligence in Financial Services,” 2020.
[6] Deloitte, “Intelligent Automation in Financial Crime Compliance,” 2020.

[7] Emerj Artificial Intelligence Research, “Artificial Intelligence at Capital One,” 2024.
[8] Capital One Technology, “Similarity Search with Graph Embeddings,” 2023.
[9] TechCrunch, “Capital One partners with Stripe and Adyen to prevent fraud,” 2024.
[10] Financial Crimes Enforcement Network (FinCEN), “FinCEN Issues In-Depth Analysis of Check Fraud Related to Mail Theft,” Sep. 9, 2024.

## Author Profile

**Rahul Gurap** is a Sr. Lead Software Engineer specializing in building secure payment processing systems and real-time machine learning architecture.