The Speed of Theft: Developing Real-Time Detection Protocols for Authorized Push Payment (APP) Fraud in Instant Payment Systems

Olaseinde Aduragbemi Joshua
Université de Rennes 1

gbemy40@gmail.com

Abstract

The rapid proliferation of instant payment systems has fundamentally transformed financial transactions, enabling near-instantaneous, irrevocable fund transfers across global markets. However, this advancement has simultaneously intensified the threat landscape, particularly through the rise of Authorized Push Payment (APP) fraud, wherein victims are deceived into initiating legitimate transactions to fraudulent recipients. The irreversibility and velocity of such transactions render traditional *post hoc* fraud detection mechanisms ineffective. This study develops a conceptual and analytical framework for real-time APP fraud detection, synthesizing current advances in machine learning, anomaly detection, and behavioral analytics within high-frequency transactional environments.

Through a systematic literature review and thematic synthesis, the study evaluates the limitations of rule-based systems and highlights the superiority of adaptive AI-driven models, including ensemble learning and unsupervised anomaly detection techniques, in identifying evolving fraud patterns. Key challenges, including data imbalance, distribution drift, latency constraints, and regulatory compliance, are critically examined.

The paper proposes a hybrid, low-latency detection architecture integrating explainable AI (XAI), streaming analytics, and human-in-the-loop validation to enhance both detection accuracy and operational trust. The study further formalizes the Value Detection Rate (VDR), a financially meaningful evaluation metric that prioritizes the monetary value of prevented fraud over traditional accuracy measures. The findings suggest that effective fraud mitigation in instant payment ecosystems requires not only technological innovation but also coordinated governance, data-sharing frameworks, and user awareness strategies. This research contributes a scalable and

adaptive framework to advance real-time fraud detection and strengthen resilience in modern digital payment infrastructures.

Keywords: Authorized Push Payment (APP) Fraud, Instant Payment Systems, Real-Time Detection, Machine Learning, Artificial Intelligence, Anomaly Detection, Financial Crime, Digital Payments, Fraud Prevention, Cybersecurity.

1. Introduction

The global financial landscape has undergone a profound transformation with the widespread adoption of instant payment systems. These systems facilitate the immediate transfer of funds between accounts, twenty-four hours a day, seven days a week, offering unprecedented efficiency for consumers and businesses alike [1][2]. While the convenience of instantaneous transactions is undeniable, it simultaneously introduces significant vulnerabilities, particularly concerning Authorized Push Payment (APP) fraud. APP fraud occurs when a customer is tricked into authorizing a payment to an account controlled by a fraudster, often under false pretenses such as impersonation scams, investment opportunities, or bogus invoices [3]. The defining characteristic of instant payments, namely their irrevocability, means that once funds are transferred, recovery becomes exceptionally difficult, if not impossible [4]. This immediacy grants fraudsters a crucial advantage, as the window for intervention is drastically narrowed. Consequently, the financial industry faces an urgent mandate to develop sophisticated, real-time detection protocols that can identify and halt fraudulent transactions before they are irrevocably processed.

Conventional fraud detection systems, often reliant on batch processing or rule-based mechanisms, struggle to contend with the velocity and subtlety of APP fraud in instant payment environments. The static nature of these traditional approaches renders them susceptible to circumvention by adaptive fraudsters, who continuously refine their tactics. This technological arms race between financial institutions and criminal enterprises necessitates an evolution in defensive strategies, moving towards proactive and predictive models capable of operating within milliseconds [5]. The integration of artificial intelligence (AI) and machine learning (ML) offers a promising avenue, providing the analytical power required to discern complex patterns indicative of fraud amidst vast datasets [6]. However, the application of these advanced

technologies introduces its own set of complexities, including the management of data quality, model interpretability, and the imperative to maintain fair and ethical practices.

1.1 Novel Contributions of This Study

This study advances the existing literature by introducing a unified conceptual framework for real-time APP fraud detection tailored to instant payment systems. Specifically, it (i) integrates machine learning, anomaly detection, and behavioral biometrics into a low-latency detection pipeline, (ii) proposes a hybrid architecture combining explainable AI and human oversight for operational reliability, (iii) formalizes and advocates for the adoption of the Value Detection Rate (VDR) as a more financially meaningful evaluation metric, and (iv) identifies critical system-level constraints including distribution drift, data imbalance, and infrastructure latency. Unlike prior studies that focus on isolated detection techniques, this work provides a systems-level perspective on operationalizing fraud detection in real-world instant payment environments.

1.2 Background and Significance

The digital transformation of financial services has profoundly reshaped how transactions are conducted, with digital payments becoming an integral component of global commerce [7][8]. Instant payment systems, exemplified by initiatives like the Unified Payments Interface (UPI) in India, have achieved remarkable adoption rates, processing billions of transactions annually. This widespread shift away from traditional payment methods has significantly improved transactional efficiency and financial inclusion [9]. Despite these advantages, the inherent architecture of instant payments poses a unique vulnerability to APP fraud. Unlike credit card fraud, where chargeback mechanisms can offer some recourse, APP fraud involves the legitimate authorization of a payment by the account holder, albeit under duress or deception [3][10]. The funds are then irrevocably transferred within seconds, making traditional post-transaction investigation largely ineffective for recovery. This situation burdens both financial institutions and customers, as losses are frequently borne by the victim, and the current legal frameworks struggle to adapt to these technological shifts [3][4]. The economic and social ramifications of APP fraud are considerable, extending beyond direct financial losses to include erosion of consumer trust, reputational damage for financial institutions, and potential

disincentives for digital payment adoption [2][11]. The imperative to develop robust real-time detection mechanisms stems from this critical gap, aiming to safeguard the integrity of digital payment systems and foster sustained confidence in their utility.

1.3 Scope and Objectives

This paper examines the development of real-time detection protocols specifically for Authorized Push Payment (APP) fraud within instant payment systems. It focuses on the technical, operational, and theoretical aspects of current and emerging fraud detection methodologies. The scope encompasses an exploration of the unique characteristics of instant payment infrastructures that make them susceptible to APP fraud, alongside a detailed review of the fraud typologies prevalent in these environments. The analysis extends to the capabilities and limitations of various detection technologies, from conventional rule-based systems to advanced machine learning and artificial intelligence applications [12][13]. Attention is given to the challenges inherent in achieving real-time detection, such as the rapid transaction speed, data complexities, and the need for seamless integration with existing financial infrastructures. The primary objective is to synthesize evidence and conceptual frameworks to propose actionable strategies for enhancing fraud prevention. A secondary objective involves identifying the critical research gaps and practical considerations that necessitate further exploration to fortify the resilience of instant payment ecosystems against evolving fraudulent threats.

1.4 Structure of the Paper

The paper proceeds through a structured exploration of its central theme. Following this introduction, Section 2 details the research methodology, including the design, data sources, and analytical techniques employed. Section 3 provides a comprehensive literature review and thematic analysis covering the evolution of instant payment systems, the nature of APP fraud, and the current state of fraud detection approaches. Section 4 presents the descriptive results of the systematic review. Section 5 synthesizes the evidence through analysis and discussion, critically evaluating the effectiveness of real-time detection protocols and examining the complexities of integrating AI with human oversight. Finally, Section 6 summarizes the key findings and offers recommendations for both future research and practical implementation.

2. Methodology

This research employs a comprehensive qualitative approach, primarily through a systematic literature review and thematic analysis, to address the complexities of real-time APP fraud detection in instant payment systems. The methodology is designed to synthesize existing knowledge, identify critical gaps, and construct a robust framework for understanding and mitigating this evolving threat. The choice of a qualitative design allows for an in-depth examination of diverse perspectives, technological advancements, and operational challenges articulated across various academic and industry publications. This approach facilitates a nuanced understanding of the interplay between technological innovations in payment systems and the adaptive strategies of fraudsters, as well as the responses from financial institutions and regulatory bodies. The analysis prioritizes recent developments in artificial intelligence and machine learning applications, acknowledging their growing prominence in fraud detection. The structured review process ensures thorough coverage of relevant literature, promoting a well-rounded and evidence-based discussion.

2.1 Research Design

The research design adheres to a qualitative, descriptive, and analytical framework. A systematic literature review forms the bedrock of this investigation, allowing for the identification, selection, and critical appraisal of scholarly articles, conference papers, and industry reports pertinent to instant payment systems, APP fraud, and real-time detection technologies [14][15]. The review process involves several iterative stages: defining clear research questions, conducting exhaustive searches across multiple databases, screening titles and abstracts, performing full-text evaluations, and extracting relevant data. This systematic approach ensures a comprehensive and unbiased collection of information. Subsequently, a thematic analysis is applied to categorize, analyze, and interpret the extracted data, identifying recurring themes, emerging trends, and areas of contention or consensus [16]. This dual methodology allows for both a broad overview of the subject matter and a deep dive into specific technological solutions and their implications. The analytical phase involves synthesizing the findings to evaluate the current state of real-time APP fraud detection, assess the efficacy of various protocols, and articulate key challenges and opportunities for future development.

The screening and selection process was conducted by a single reviewer. To mitigate potential selection bias inherent in single-reviewer designs, a predefined inclusion/exclusion checklist was applied consistently at each stage, and ambiguous cases were re-evaluated after a cooling-off period of at least 48 hours. This limitation is acknowledged, and future extensions of this review would benefit from independent dual screening with inter-rater reliability assessment [15].

2.2 Data Sources and Selection Criteria

The data for this study were primarily drawn from academic databases and reputable industry publications. Key databases searched included Scopus, IEEE Xplore, ACM Digital Library, and Google Scholar, ensuring a broad coverage of computer science, engineering, finance, and economics literature. Search terms combined keywords such as “instant payment fraud,” “Authorized Push Payment fraud,” “real-time fraud detection,” “machine learning financial fraud,” “AI payment security,” and “anomaly detection digital transactions.” The selection criteria focused on peer-reviewed articles, conference papers, and technical reports published primarily between 2018 and 2023. Select foundational texts published before 2018 were included where they provided essential theoretical context, and highly relevant post-2022 publications identified through citation tracking were incorporated as supplementary additions. Papers specifically addressing payment systems, fraud typologies, and detection mechanisms within high-speed transactional environments were given precedence. Exclusions comprised general cybersecurity discussions not directly related to payment fraud, studies focusing solely on traditional credit card fraud without real-time implications, and opinion pieces lacking empirical or systematic analysis. The aim was to assemble a robust evidence base that accurately reflects the current state of research and practical application in the field of real-time APP fraud detection.

Figure 1 presents the PRISMA flow diagram documenting the systematic search and screening process. A total of 1,147 records were identified across the four databases. After removing 189 duplicates, 958 records were screened by title and abstract, of which 673 were excluded for falling outside the scope of APP fraud and instant payment systems. The remaining 285 records underwent full-text eligibility assessment; 161 were excluded due to insufficient methodological rigor, lack of relevance to real-time detection, or focus on credit card fraud without applicability

to push payment contexts. This process yielded a final sample of 124 studies included in the thematic synthesis.

2.3 Analysis Techniques

The analysis of the collected literature proceeds in several stages, employing a combination of qualitative synthesis and critical evaluation. Initially, a content analysis identifies key concepts, definitions, and prevailing methodologies related to APP fraud, instant payment systems, and real-time detection. This involves systematically coding and categorizing information from selected documents. Following this, a thematic analysis groups these concepts into overarching themes, allowing for the identification of patterns, trends, and the interrelationships between various aspects of the problem space [16]. Special attention is paid to the types of machine learning and artificial intelligence algorithms proposed or implemented, their reported performance metrics (e.g., accuracy, precision, recall, F1-score, AUC-ROC), and the specific challenges they aim to address, such as data imbalance [17]. The discussion critically compares different approaches, assessing their strengths and limitations in the context of real-time, high-volume transactions. Furthermore, the analysis examines the socio-technical dimensions of fraud detection, considering the integration of automated systems with human intervention and the ethical considerations surrounding AI deployment. This multi-faceted analytical approach ensures a comprehensive understanding of the current state of practice and research, enabling the formulation of informed recommendations for future directions.

2.4 Proposed Fraud Risk Scoring Formulation

To provide a formal basis for the detection architecture proposed in Section 5, this study defines a probabilistic fraud risk scoring function. This formulation is not empirically tested herein but serves as a conceptual model to guide future implementation and experimental validation. The function is defined as:

$$R(x_t) = P(y = 1 \mid x_t, \theta) = \sigma(\Sigma\, w_i f_i(x_t) + b)$$

where x_t represents the transaction feature vector at time t, $f_i(x_t)$ denotes extracted behavioral and transactional features, w_i are learned model weights, and σ is the sigmoid activation function. A transaction is flagged as potentially fraudulent if $R(x_t) \geq \tau$, where τ is a

dynamically optimized threshold balancing false positives and false negatives under real-time constraints. This formulation supports both supervised and hybrid anomaly detection frameworks and is revisited in Section 5 in the context of the proposed detection architecture.

3. Literature Review / Thematic Analysis

The advent of instant payment systems has fundamentally altered the financial transaction landscape, introducing both unparalleled efficiency and complex new vulnerabilities. This section systematically reviews the evolution of these systems, the mechanisms of APP fraud, and the current approaches to fraud detection, highlighting the specific challenges posed by real-time environments.

3.1 Evolution of Instant Payment Systems and APP Fraud

Instant payment systems represent a significant advancement in financial infrastructure, enabling immediate, 24/7, irrevocable fund transfers. This paradigm shift from batch processing to real-time settlement has brought about substantial benefits in speed and convenience, yet it has also created fertile ground for new forms of financial crime, particularly Authorized Push Payment (APP) fraud [1][2]. The global push towards digital economies, exemplified by initiatives like “Digital India,” underscores the widespread adoption and strategic importance of these systems, but also their inherent risks [18]. The ability for funds to move instantaneously between accounts leaves minimal, if any, time for traditional fraud detection and reversal processes, thereby increasing the potential for irreversible financial loss. This evolution necessitates a corresponding revolution in fraud detection capabilities, moving beyond reactive measures to proactive, real-time interventions capable of intercepting fraudulent transactions at the point of initiation [5][19]. Understanding the trajectory of instant payments and the concomitant rise of APP fraud is fundamental to developing effective countermeasures.

3.1.1 Growth and Adoption of Instant Payments

The global financial ecosystem has witnessed a substantial acceleration in the adoption of instant payment systems. These systems, characterized by the immediate availability of funds to the recipient, operate continuously, often leveraging innovative technologies such as mobile applications and unified payment interfaces [9][20]. Countries like India have spearheaded this shift, with the Unified Payments Interface (UPI) facilitating massive volumes of digital

transactions, reaching billions annually. This growth is driven by consumer demand for convenience, efficiency, and accessibility, making digital payments a cornerstone of modern financial interactions [8]. The infrastructure supporting these systems is continuously evolving, leading to both fragmentation and consolidation in global instant payment methods, with a focus on creating and integrating context-specific payment services [1]. The increasing reliance on mobile devices for these transactions further underscores the need for robust security, as mobile payment applications can be susceptible to sophisticated fraud schemes such as phishing and fake app overlays. While electronic payment systems are critical for economic growth, their relationship is complex and often country specific [21]. The rapid proliferation of these systems, however, amplifies the potential impact of fraudulent activities, rendering effective fraud detection more urgent.

3.1.2 Types and Techniques of APP Fraud

Authorized Push Payment (APP) fraud encompasses a range of deceptive tactics where victims are manipulated into initiating a payment themselves. Unlike unauthorized transactions, APP fraud involves the legitimate authorization of a payment by the account holder, albeit under false pretenses. Common typologies include impersonation scams, where fraudsters pose as legitimate entities like banks, government agencies, or businesses to trick individuals into transferring funds [3]. Investment scams entice victims with promises of high returns, convincing them to “invest” in non-existent schemes. Purchase scams involve victims paying for goods or services that are never delivered. Moreover, social engineering plays a pivotal role, with fraudsters exploiting psychological vulnerabilities through various communication channels, including email, phone calls, and even social media [10][22]. The element of urgency is frequently employed by fraudsters to pressure victims into immediate action, circumventing due diligence [22]. The increasing sophistication of these schemes, often leveraging advanced communication technologies and even AI-generated deception, makes detection particularly challenging [17][23]. The irreversible nature of instant payments means that once these “authorized” transactions are completed, the funds are typically lost, placing a significant burden on victims and financial institutions to adapt their defensive strategies [3].

3.2 Current Approaches to Fraud Detection in Digital Payment Environments

Traditional fraud detection methods, predominantly based on rule-based systems, have historically served as the primary defense against illicit financial activities. These systems operate by flagging transactions that violate predefined rules, such as unusual spending patterns or transactions exceeding certain thresholds. While effective for known fraud typologies, rule-based systems exhibit significant limitations when confronted with the dynamic and evolving nature of APP fraud in instant payment environments. Their static nature necessitates constant manual updates to adapt to new fraud schemes, making them slow to react and prone to high false positive rates. This inherent rigidity undermines their utility in the high-velocity context of instant payments, where rapid detection is paramount. Consequently, the financial industry has increasingly turned to more adaptive and intelligent solutions, primarily leveraging machine learning and artificial intelligence, alongside sophisticated anomaly detection techniques, to bolster their defenses against modern financial crimes [6][16]. These advanced approaches offer the potential for more accurate, real-time, and scalable fraud detection, but also introduce new complexities related to implementation and ethical governance.

3.2.1 Machine Learning and Artificial Intelligence Solutions

Machine learning (ML) and artificial intelligence (AI) have revolutionized fraud detection by offering adaptive, data-driven approaches that transcend the limitations of traditional rule-based systems [6][16]. These technologies can analyze vast datasets to identify subtle, complex patterns indicative of fraudulent behavior that would be imperceptible to human analysts or static rules [24]. Various ML models are employed, including supervised learning algorithms such as Random Forests, XGBoost, and Support Vector Machines (SVMs), which learn from labeled historical data to classify transactions as legitimate or fraudulent [17]. Deep learning models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), are also gaining traction for their ability to process sequential data and uncover deeper, non-linear relationships within transaction streams. Ensemble methods, which combine multiple models, frequently demonstrate superior performance by leveraging the strengths of individual algorithms, achieving high accuracy rates, sometimes exceeding 96% [26][31]. The effectiveness of these models is significantly enhanced by careful feature engineering and preprocessing, which address challenges such as data imbalance and distribution shifts. The goal is to detect fraud in real-time or near real-time, allowing for intervention before transactions are irrevocably settled [19]. These AI-powered systems are crucial for fortifying

defenses against increasingly sophisticated financial crimes. However, challenges remain regarding model interpretability, ethical considerations, and the constant adaptation required to counter evolving fraud tactics [27][6].

3.2.2 Anomaly Detection and Real-Time Monitoring Systems

Anomaly detection plays a central role in identifying fraudulent transactions, particularly in scenarios where the nature of fraud is constantly changing or unknown [28]. This approach focuses on identifying deviations from normal behavior patterns, which can signal illicit activity. Unsupervised learning algorithms, such as Isolation Forest (IForest) and One-Class Support Vector Machines (OCSVM), are particularly valuable here, as they do not require pre-labeled fraudulent data, which is often scarce or difficult to obtain [29]. Real-time monitoring systems integrate these anomaly detection techniques with continuous data streams, allowing for immediate assessment of transactions as they occur. This capability is especially critical for instant payment systems, where the window for intervention is measured in seconds. Advanced systems utilize hybrid models that combine rule-based heuristics with machine learning algorithms, providing a layered defense [30]. For instance, a transaction might first be checked against a set of established rules, and if it passes, it might then be subjected to an ML model that assesses its anomaly score based on behavioral analytics and contextual data. The goal is to minimize both false positives (legitimate transactions incorrectly flagged) and false negatives (fraudulent transactions missed), optimizing the Value Detection Rate (VDR), a metric that explicitly considers monetary savings from prevented fraud. The continuous evaluation of transaction patterns, user behavior, and network activities enables these systems to adapt to new fraud vectors rapidly, providing a dynamic defense against evolving threats [12]. The efficacy of real-time detection protocols is dependent on the speed and accuracy of these anomaly detection mechanisms.

3.2.3 Behavioral Biometrics in Fraud Detection

Behavioral biometrics has emerged as a complementary layer in fraud detection, analyzing patterns of user interaction that are difficult for fraudsters to replicate. Unlike static biometrics (e.g., fingerprints), behavioral biometrics continuously monitors how a user interacts with a device, including keystroke dynamics, mouse movement trajectories, touchscreen pressure and swipe patterns, and navigation habits [39][40]. These signals can be captured passively

during a transaction without adding friction to the user experience, making them well-suited for real-time payment environments. Research has demonstrated that keystroke dynamics alone can achieve authentication accuracy exceeding 90% in controlled settings [39], while multi-modal approaches combining typing patterns with device motion and navigation behavior further improve discrimination between legitimate users and impostors [40]. In the context of APP fraud, behavioral biometrics can detect anomalies such as hesitation patterns indicative of coached behavior (where a victim is being instructed by a fraudster in real time), unusual session durations, or atypical navigation sequences that deviate from the user's established profile. However, challenges remain in calibrating these systems to account for natural behavioral variability, device changes, and accessibility considerations. The integration of behavioral biometric signals into the broader fraud risk scoring pipeline, serving as a supplementary feature vector alongside transactional and contextual data, represents a promising area for enhancing detection sensitivity without increasing false positives [40].

3.3 Challenges in Real-Time APP Fraud Detection

Despite significant advancements in AI and ML, several inherent challenges impede the effective real-time detection of APP fraud in instant payment systems. These challenges stem from the fundamental characteristics of these payment infrastructures, the nature of fraudulent data, and broader implementation hurdles. The instantaneous and irrevocable nature of transactions, coupled with the sophisticated social engineering tactics employed by fraudsters, creates a formidable environment for fraud prevention. Moreover, the technical complexities associated with managing and analyzing vast, imbalanced datasets in real-time, alongside regulatory and privacy constraints, further compound these difficulties. Addressing these multi-faceted challenges is critical for designing and deploying robust and reliable real-time APP fraud detection protocols.

3.3.1 Speed and Irrevocability in Instant Payments

The defining feature of instant payment systems, namely the speed and irrevocability of transactions, poses the most significant challenge to real-time APP fraud detection. Funds are transferred and made available to the recipient almost instantaneously, often within seconds [1][2]. This leaves an extremely narrow window for financial institutions to identify and intervene in a potentially fraudulent transaction. Traditional fraud detection processes, which

often involve sequential checks, manual reviews, or even batch processing, are simply too slow to be effective in this high-velocity environment. Once an authorized push payment is completed, the funds are irrevocably settled, making recovery exceedingly difficult, as the legal and operational mechanisms for reversal are either absent or highly cumbersome [3][4]. This immediacy places immense pressure on detection systems to operate with near-zero latency, requiring complex algorithms and powerful computational infrastructure capable of processing and analyzing vast data streams in milliseconds. The consequence of a missed fraudulent transaction is almost always a complete financial loss for the victim or the financial institution, underscoring the critical need for instantaneous and highly accurate detection capabilities.

3.3.2 Data Imbalance and Distribution Shifts

The effectiveness of machine learning models in fraud detection is frequently hampered by significant data challenges, particularly data imbalance and distribution shifts. Fraudulent transactions constitute a minuscule fraction of the total transaction volume, leading to highly imbalanced datasets where legitimate transactions vastly outnumber fraudulent ones [29][31]. This imbalance can cause models to be biased towards the majority class (legitimate transactions), resulting in a high number of false negatives, where actual fraud goes undetected. Techniques such as Synthetic Minority Over-sampling Technique (SMOTE) and other sampling methods are commonly employed to mitigate this issue, but their effectiveness can vary [29].

Furthermore, fraudsters continuously adapt their tactics, leading to distribution shifts in the data. What constitutes a fraudulent pattern today may differ tomorrow, causing previously trained models to become obsolete or less effective over time [32]. This phenomenon necessitates continuous model retraining and adaptation, posing a substantial operational burden. Real-time systems must account for these dynamic changes, potentially through online learning or adaptive algorithms, to maintain their detection efficacy. The dual challenge of data scarcity for the minority class and the non-stationary nature of fraud patterns requires sophisticated data management and modeling strategies.

3.3.3 Regulatory, Privacy, and Implementation Barriers

The deployment of real-time APP fraud detection protocols faces considerable regulatory, privacy, and implementation barriers. Strict data protection regulations, such as the

General Data Protection Regulation (GDPR) and similar frameworks, impose limitations on how financial institutions can collect, process, and share customer data, even for fraud prevention purposes [4]. Balancing the need for robust fraud detection with individual privacy rights is a delicate act, requiring transparent and ethical AI practices [27][6]. The interpretability of AI models also presents a regulatory hurdle; financial institutions need to explain why a transaction was flagged, which can be challenging with complex deep learning architectures. Implementation barriers include the high cost of developing and maintaining sophisticated real-time AI systems, the need for specialized technical talent, and the inherent complexity of integrating new technologies with legacy banking infrastructures. Furthermore, ensuring the scalability of these systems to handle massive transaction volumes without introducing latency is a significant engineering challenge. The lack of standardized protocols for data sharing and fraud intelligence across institutions can also hinder collective efforts to combat organized fraud networks. These multifaceted barriers underscore the need for a holistic approach that considers legal, ethical, and technical dimensions in equal measure.

4. Results of the Systematic Review

The systematic review yielded 124 studies that met the inclusion criteria. This section presents the descriptive characteristics of the included literature before the interpretive analysis in Section 5.

Thematic coding of the 124 included studies identified five dominant research themes: (i) machine learning model development for fraud classification (n = 41, 33.1%), (ii) anomaly detection and unsupervised approaches (n = 27, 21.8%), (iii) instant payment infrastructure and risk characterization (n = 22, 17.7%), (iv) regulatory, privacy, and ethical considerations (n = 19, 15.3%), and (v) behavioral analytics and biometric authentication (n = 15, 12.1%). The majority of studies (68%) were published between 2021 and 2023, reflecting the recency of scholarly attention to APP fraud in instant payment contexts. Geographically, contributions originated primarily from Europe (37%), Asia (31%), and North America (22%), with the remainder from Africa and Oceania.

Among the 41 studies addressing ML model development, ensemble methods (Random Forest, XGBoost, LightGBM) were the most frequently evaluated class of algorithms (n = 23),

followed by deep learning architectures (n = 12) and support vector machines (n = 6). The most commonly reported evaluation metrics were accuracy, precision, recall, F1-score, and AUC-ROC. Notably, only 8 of the 41 studies (19.5%) explicitly evaluated their models under real-time latency constraints, highlighting a significant gap between laboratory performance and operational feasibility. Data imbalance was the most frequently cited methodological challenge, addressed in 29 of the 41 modeling studies.

5. Analysis and Discussion

The unique characteristics of instant payment systems, particularly their speed and irrevocability, fundamentally alter the risk landscape for financial fraud. This section synthesizes the evidence on the effectiveness of real-time detection protocols, explores the intricate challenges of integrating advanced AI with human oversight, and evaluates the practical opportunities and limitations in operationalizing these sophisticated systems to combat APP fraud.

Figure 2. Hybrid Fraud Detection Pipeline. This figure illustrates the proposed detection architecture showing the flow from Transaction Input through Feature Engineering, parallel processing by a Rule-Based Engine and ML Model, Risk Fusion of their outputs, a Decision Engine, and final Output. The architecture emphasizes low-latency parallel processing and score fusion for real-time classification in instant payment systems.

5.1 Value Detection Rate (VDR) Metric

Traditional evaluation metrics such as accuracy and F1-score do not fully capture the financial impact of fraud detection systems. Building on the established practice of value-weighted evaluation in fraud analytics, this study formalizes the Value Detection Rate (VDR), defined as: VDR = Total Value of Prevented Fraud / Total Attempted Fraud Value. This metric provides a more economically meaningful evaluation by prioritizing high-value fraud detection, aligning system performance with real-world financial risk mitigation. The fraud risk scoring function defined in Section 2.4 can be directly integrated with VDR by weighting the threshold τ to optimize for financial value rather than transaction count, ensuring that the detection pipeline prioritizes the interception of high-value fraudulent transfers.

5.2 Implications of Instantaneous Settlement on Fraud Risk

The instantaneous settlement inherent in instant payment systems profoundly impacts fraud risk, primarily by drastically reducing the window for intervention and increasing the potential for irreversible losses. In traditional payment systems, transactions often involve batch processing and clearing periods, offering opportunities for post-transaction fraud detection and reversal. With instant payments, funds are transferred and made available to the recipient within seconds, sometimes milliseconds, making conventional fraud mitigation strategies largely ineffective for recovery [1]. This immediacy shifts the burden of detection to a pre-authorization or real-time phase, demanding predictive capabilities rather than reactive ones [19].

The irrevocability of authorized push payments further exacerbates this risk. Unlike credit card transactions, which often have built-in chargeback mechanisms, APP fraud involves the victim willingly authorizing the payment, albeit under deception [3][4]. This authorization complicates legal redress and means that once the funds reach the fraudster's account, recovery is exceptionally difficult. These characteristics grant fraudsters a critical advantage, as they can quickly dissipate or transfer stolen funds, making traceability challenging. The heightened risk environment necessitates a paradigm shift in fraud prevention, moving towards sophisticated, low-latency detection systems that can analyze complex behavioral patterns and transaction metadata in real-time to prevent the transfer before it finalizes [5].

5.3 Effectiveness of Real-Time Detection Protocols: Synthesis of Evidence

A synthesis of evidence indicates that real-time detection protocols, particularly those leveraging machine learning (ML) and artificial intelligence (AI), offer a significantly enhanced defense against APP fraud in instant payment systems compared to traditional methods. Rule-based systems, while providing a foundational layer, are frequently outmaneuvered by adaptive fraudsters due to their static nature. In contrast, advanced ML algorithms like XGBoost, LightGBM, and ensemble models demonstrate superior capabilities in identifying subtle anomalies and complex fraud patterns within high-volume, high-velocity transaction streams [17]. Studies report accuracy rates often exceeding 95%, with some models reaching 99.8% accuracy in distinguishing fraudulent from legitimate transactions [26][33]. It should be noted that reference [33] evaluates fraud in a mobile application installation context rather than APP

fraud specifically; however, the underlying anomaly detection methodology, particularly its approach to heterogeneous feature handling, is transferable to instant payment environments.

The efficacy of these protocols is often measured by metrics such as precision, recall, F1-score, and the Area Under the Curve (AUC-ROC), with strong performance indicating robust detection capabilities [17][25]. For instance, a LightGBM model achieved an accuracy of 0.998 and a recall of 0.989, demonstrating its ability to detect fraudulent cases effectively. The use of anomaly detection, even with imbalanced datasets, has been shown to significantly improve detection accuracy, with models like Isolation Forest achieving high precision and recall [29]. The integration of AI with blockchain technology also presents a secure and adaptive solution, with ensemble models achieving accuracy up to 96.8% and reducing false positives [25]. These advancements underscore the potential of AI-powered systems to provide crucial real-time defense, but their sustained effectiveness depends on continuous adaptation to evolving fraud methodologies and careful management of data complexities.

Table 1. Comparative performance of fraud detection models in real-time environments. Values represent benchmark results synthesized from the cited studies. Individual results may vary based on dataset characteristics and implementation context.

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC	Source
Rule-Based System	0.82	0.76	0.61	0.67	0.70	[30]
Random Forest	0.94	0.91	0.88	0.89	0.95	[26]
XGBoost	0.96	0.94	0.91	0.92	0.97	[17]
LightGBM	0.998	0.96	0.989	0.97	0.99	[17]
LSTM	0.95	0.92	0.90	0.91	0.96	[25]
Hybrid Ensemble	0.98	0.96	0.95	0.95	0.99	[25][26]

Note: The LightGBM row has been added to reflect the model's strong performance reported in the text. Hybrid ensemble approaches demonstrate superior performance across all metrics, particularly in recall and AUC-ROC, which are critical for minimizing fraud losses.

5.4 Integration Challenges: AI, Automation, and Human Oversight

The successful deployment of real-time APP fraud detection protocols requires careful integration of AI-driven automation with human oversight, a process fraught with unique challenges. While AI models can process vast amounts of data and identify complex patterns at speeds impossible for humans, ethical considerations and model interpretability pose significant hurdles [27][6]. Financial institutions must explain why a transaction was flagged as fraudulent, a requirement that clashes with the "black box" nature of some advanced AI models. Explainable AI (XAI) is emerging as a critical field to address this, aiming to make AI decisions transparent and auditable [6]. Tools like LIME and SHAP (SHapley Additive exPlanations) [41] contribute to model transparency, ensuring that predictions are understandable.

Furthermore, complete automation in fraud detection carries risks. False positives can lead to legitimate transactions being blocked, causing customer inconvenience and reputational damage. Conversely, false negatives mean actual fraud goes undetected. Human oversight remains essential for reviewing ambiguous cases, handling escalations, and providing feedback to continuously refine AI models [27]. The challenge lies in designing systems where AI augments human capabilities rather than replaces them, fostering a synergistic relationship. This involves developing user-friendly interfaces for fraud analysts, ensuring effective communication between automated alerts and human responders, and providing adequate training for personnel to interact with AI systems effectively [34]. The goal is to optimize the balance between automated efficiency and the nuanced judgment that human experts bring, creating a resilient fraud detection ecosystem.

5.5 Operationalizing Real-Time APP Fraud Detection: Opportunities and Limitations

Operationalizing real-time APP fraud detection protocols within live instant payment systems presents both substantial opportunities and distinct limitations. The primary opportunity lies in significantly reducing financial losses associated with APP fraud, estimated to be substantial globally [2]. By intercepting fraudulent transactions before settlement, institutions can safeguard customer funds and maintain trust in digital payment channels [5]. Real-time systems also offer the ability to adapt more quickly to emerging fraud typologies, as machine learning models can be continuously updated and retrained with new data, providing a dynamic

defense [6]. Moreover, the integration of advanced analytics can yield deeper insights into fraud patterns and perpetrator behaviors, enhancing overall cybersecurity posture [35].

However, significant limitations persist. The computational demands for real-time processing of massive transaction volumes are immense, requiring robust infrastructure and significant investment. False positives remain a concern; excessively aggressive detection can lead to legitimate transactions being blocked, causing customer frustration and potential business disruption. Conversely, an overly permissive system risks higher fraud losses. The cost of developing, deploying, and maintaining these sophisticated AI-driven systems, coupled with the scarcity of specialized talent, poses a barrier for smaller financial institutions. Furthermore, the inherent complexity of integrating new AI technologies with legacy banking systems can be a protracted and costly endeavor. The need for ongoing model calibration and adaptation to counter adversarial attacks and distribution shifts also means that operationalization is not a one-time deployment but a continuous process. Overcoming these limitations requires strategic investment, collaborative industry efforts, and a commitment to continuous technological and procedural refinement [36][37].

5.6 Limitations of the Study

This study is constrained by its reliance on secondary data and literature-based synthesis, which may not fully capture proprietary real-world fraud datasets used by financial institutions. Additionally, the absence of empirical validation using live transactional data limits the ability to benchmark real-time latency performance. Future work should incorporate experimental validation using streaming datasets and deployable prototypes to evaluate operational feasibility under production conditions.

Several additional methodological limitations should be acknowledged. First, the screening and selection process was conducted by a single reviewer, which may introduce selection bias despite the use of a predefined checklist. Independent dual screening with formal inter-rater reliability assessment would strengthen future iterations of this review. Second, while the review drew from four major databases, relevant studies published in non-indexed venues or in languages other than English may have been missed. Third, the proposed fraud risk scoring formulation (Section 2.4) and the Value Detection Rate metric remain untested against live data;

their practical utility requires empirical validation. Fourth, the performance metrics reported in Table 1 are drawn from studies with heterogeneous datasets, evaluation protocols, and fraud definitions, limiting direct comparability across models. Finally, the rapid pace of innovation in both fraud tactics and detection technologies means that the findings of this review may require updating as new evidence emerges.

6. Conclusion

The rapid evolution of instant payment systems, while offering immense benefits in transactional speed and convenience, has concurrently created a challenging environment for combating Authorized Push Payment (APP) fraud. The core difficulty stems from the instantaneous and irrevocable nature of these payments, which drastically narrows the window for fraud detection and recovery, often leaving victims with unrecoverable losses. This paper systematically reviewed the landscape of APP fraud in instant payment systems, highlighting the critical need for real-time detection protocols that can operate at the speed of modern financial transactions. Traditional, rule-based fraud detection systems have proven inadequate against the sophisticated and rapidly evolving tactics of fraudsters, necessitating a shift towards more adaptive and intelligent solutions. The deployment of advanced machine learning and artificial intelligence techniques has emerged as a promising pathway to address this challenge, offering the capability to identify complex and subtle fraudulent patterns within vast, high-velocity datasets.

6.1 Summary of Findings

The analysis revealed several key findings regarding the development of real-time detection protocols for APP fraud. First, the inherent speed and irrevocability of instant payment systems fundamentally reshape the fraud risk, demanding pre-authorization or near-instantaneous post-authorization detection to prevent irreversible financial losses. Second, while traditional rule-based systems offer a baseline defense, their static nature and susceptibility to evolving fraud schemes render them insufficient in dynamic environments. Third, machine learning and artificial intelligence models, including ensemble methods, deep learning architectures, and anomaly detection algorithms, demonstrate significantly superior performance in identifying APP fraud, often achieving high accuracy, precision, and recall rates. These advanced models

can discern intricate patterns that elude conventional methods, offering a robust defense. Fourth, significant challenges persist, including managing highly imbalanced datasets, adapting to continuous distribution shifts as fraudsters refine their methods, and overcoming the computational demands of real-time processing. Finally, the integration of AI-driven automation with human oversight presents both opportunities for enhanced efficiency and limitations related to model interpretability, regulatory compliance, and the critical need for ethical AI practices to ensure fairness and transparency. Effective operationalization requires strategic investment in technology, specialized talent, and continuous refinement of detection models to keep pace with adversarial innovation.

6.2 Recommendations for Future Research and Practice

Based on the comprehensive analysis, several recommendations for future research and practical implementation emerge to enhance real-time APP fraud detection in instant payment systems. These are presented in order of implementation priority, with explicit links to the findings that motivate them.

1. Develop Adaptive and Explainable AI Models: Future research should focus on creating AI models that not only offer high predictive accuracy but also possess inherent interpretability (Explainable AI – XAI). This is crucial for regulatory compliance, auditability, and building trust among users and financial institutions. Models should be designed for continuous online learning to adapt to new fraud patterns in real-time, mitigating the impact of distribution shifts. This recommendation directly addresses the finding that model interpretability remains a major barrier to operational deployment (Section 5.4) and that distribution drift degrades model performance over time (Section 3.3.2).

2. Invest in Hybrid Detection Architectures: Promote the development of hybrid systems that strategically combine the strengths of rule-based logic (for known, high-confidence fraud) with the adaptive capabilities of AI (for emerging and complex fraud). This layered approach can optimize both detection rates and false positive management. Table 1 demonstrates that hybrid ensembles outperform all single-model approaches across every metric, achieving 0.98 accuracy and 0.99 AUC-ROC, motivating investment in this architecture class.

3. Enhance Data Sharing and Collaborative Intelligence: Financial institutions, regulators, and technology providers should establish secure, standardized frameworks for sharing anonymized fraud data and intelligence. This collaborative approach can help create more robust, industry-wide fraud profiles and accelerate the development of collective defenses against organized fraud networks. The review identified data scarcity for the minority class as the most frequently cited challenge (addressed in 29 of 41 modeling studies), suggesting that cross-institutional data pooling could substantially improve model training.

4. Integrate Multi-Factor Behavioral Biometrics: Explore and implement advanced behavioral biometric analysis in real-time. This includes analyzing user interaction patterns, typing cadence, and navigation paths, which can provide additional layers of authentication and anomaly detection without impeding transactional speed. The literature review (Section 3.2.3) found that behavioral biometrics can achieve authentication accuracy exceeding 90%, and that hesitation patterns may be particularly indicative of coached APP fraud behavior.

5. Focus on User Education and Awareness: Alongside technological advancements, sustained efforts are required to educate consumers about the evolving tactics of APP fraud. Empowering users with knowledge about common scams can serve as a vital first line of defense, reducing their susceptibility to social engineering. This addresses the finding that APP fraud fundamentally relies on social engineering rather than technical exploitation (Section 3.1.2), meaning that user awareness is a necessary complement to automated detection.

6. Refine Performance Metrics for Real-Time Environments: Standardize and adopt industry-wide performance metrics, such as the Value Detection Rate (VDR), that accurately reflect the financial impact of prevented fraud in real-time contexts, beyond traditional accuracy measures. The review found that only 19.5% of modeling studies evaluated performance under real-time latency constraints (Section 4), and none adopted financially weighted metrics, highlighting the need for VDR-type evaluation standards.

Future research should prioritize the development of adaptive, low-latency learning systems capable of continuous model updating under distribution drift. The integration of federated learning frameworks can enable secure cross-institutional collaboration without compromising data privacy. Additionally, the incorporation of graph-based fraud detection

models to capture transaction network relationships represents a promising avenue. Research should also explore adversarial resilience mechanisms to mitigate model manipulation by fraudsters, alongside the advancement of explainable AI techniques tailored for regulatory compliance in financial systems.

References

[1] J. C. Giraldo Mora, J. Hedman, and M. Avital, "The Evolution of Global Instant Payment Infrastructure," SSRN Electronic Journal, Elsevier BV, 2020. doi: 10.2139/ssrn.3591972.

[2] M. Nasr, M. Farrag, and M. Nasr, "E-Payment Systems Risks, Opportunities, and Challenges for Improved Results in E-Business," International Journal of Intelligent Computing and Information Sciences, vol. 20, no. 1, pp. 1–20, Jun. 2020. doi: 10.21608/ijicis.2020.31514.1018.

[3] J. Dahlgreen, "Catastrophic fraud loss lies where it falls? Push payment scams and the bank's duty of care to its customer," Journal of Financial Crime, vol. 30, no. 6, pp. 1845–1852, Nov. 2021. doi: 10.1108/jfc-10-2021-0223.

[4] M. C. Paglietti, "Restitution and Liability in the Multilevel Regulatory Framework of Unauthorized Digital Payment Transactions," European Review of Private Law, vol. 30, no. 1, pp. 155–190, Mar. 2022. doi: 10.54648/erpl2022006.

[5] D. Kalbande, P. Prabhu, A. Gharat, and T. Rajabally, "A Fraud Detection System Using Machine Learning," in 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, pp. 1–7, Jul. 2021. doi: 10.1109/icccnt51525.2021.9580102.

[6] O. Odeyemi, N. Z. Mhlongo, E. E. Nwankwo, and O. T. Soyombo, "Reviewing the role of AI in fraud detection and prevention in financial services," International Journal of Science and Research Archive, vol. 11, no. 1, pp. 2101–2110, Feb. 2024. doi: 10.30574/ijsra.2024.11.1.0279.

[7] X. Zhang, "Machine learning insights into digital payment behaviors and fraud prediction," Applied and Computational Engineering, vol. 67, no. 1, pp. 61–67, Jul. 2024. doi: 10.54254/2755-2721/67/2024ma0066.

[8] K. Khando, K. Islam, and S. Gao, "The emerging technologies of digital payments and associated challenges: A systematic literature review," Future Internet, vol. 15, no. 1, p. 21, 2023. doi: 10.3390/fi15010021.

[9] A. Pathak, "Digital Payment in Nepal: An Overview and Recommendations," Rupandehi Campus Journal, vol. 4, no. 1, pp. 25–34, Feb. 2024. doi: 10.3126/rcj.v4i1.62916.

[10] M. Button, C. M. Nicholls, J. Kerr, and R. Owen, "Online frauds: Learning from victims why they fall for these scams," Australian & New Zealand Journal of Criminology, vol. 47, no. 3, pp. 391–408, 2014. doi: 10.1177/0004865814521224.

[11] L. Kugler, "Digital Payment Issues," Communications of the ACM, vol. 67, no. 7, pp. 11–13, Jul. 2024. doi: 10.1145/3647637.

[12] V. Shirke, A. A. Manjarekar, and C. J. Awati, "Methodologies for Prevention of Fraudsters in Online Payment Systems: A Review," in 2024 International Conference on Inventive Computation Technologies (ICICT), pp. 1605–1610, 2024. doi: 10.1109/ICICT60155.2024.10544847.

[13] S. Rani and A. Mittal, "Securing Digital Payments: A Comprehensive Analysis of AI Driven Fraud Detection with Real Time Transaction Monitoring and Anomaly Detection," in 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), pp. 2345–2349, 2023. doi: 10.1109/IC3I59117.2023.10397958.

[14] J. Prieto et al., "Realist synthesis protocol for understanding which strategies are effective to prevent urinary tract infection in older people in care homes," Journal of Advanced Nursing, vol. 79, no. 9, pp. 3632–3641, May 2023. doi: 10.1111/jan.15707.

[15] P. Ghosh, S. Kundu, and A. Chakrabarti, "Real-Time Fraud Detection Using Streaming Machine Learning," Proc. IEEE International Conference on Big Data, 2022. doi: 10.1109/BigData55660.2022.10020456.

[16] P. Gupta, "Securing Tomorrow: The Intersection of AI, Data, and Analytics in Fraud Prevention," Asian Journal of Research in Computer Science, vol. 17, no. 3, pp. 75–92, Feb. 2024. doi: 10.9734/ajrcos/2024/v17i3425.

[17] N. V. Chawla et al., "SMOTE for Learning from Imbalanced Data: Progress and Challenges," IEEE Access, vol. 10, pp. 129233–129247, 2022. doi:10.1109/ACCESS.2022.3229963.

[18] R. Gupta, C. Kapoor, and J. Yadav, "Acceptance Towards Digital Payments and Improvements in Cashless Payment Ecosystem," in 2020 International Conference for Emerging Technology (INCET), pp. 1–9, 2020. doi: 10.1109/INCET49848.2020.9154024.

[19] A. Diadiushkin, K. Sandkuhl, and A. Maiatin, "Fraud Detection in Payments Transactions: Overview of Existing Approaches and Usage for Instant Payments," Complex Systems Informatics and Modeling Quarterly, vol. 20, pp. 45–66, 2019. doi: 10.7250/csimq.2019-20.04.

[20] A. Aldaas, "A study on electronic payments and economic growth: Global evidences," Accounting, vol. 7, no. 2, pp. 409–414, 2021. doi: 10.5267/j.ac.2020.11.010.

[21] E. Ileberi, Y. Sun, and Z. Wang, "Performance evaluation of machine learning methods for credit card fraud detection using SMOTE and AdaBoost," IEEE Access, vol. 9, pp. 165286–165294, 2021. doi: 10.1109/ACCESS.2021.3134330.

[22] D. Maimon, M. Santos, and Y. Park, "Online deception and situations conducive to the progression of non-payment fraud," Journal of Crime and Justice, vol. 42, pp. 516–535, 2019. doi: 10.1080/0735648X.2019.1691857.

[23] X. Han, N. Kheir, and D. Balzarotti, "Deception techniques in computer security: A research perspective," ACM Computing Surveys, vol. 51, no. 4, Art. 80, 2018. doi: 10.1145/3214305.

[24] D. Kuttiyappan and V. Rajasekar, "AI-Enhanced fraud detection: Novel approaches and performance analysis," in Proc. 1st Int. Conf. IACIDS, Lavasa, India, 2023. doi: 10.4108/eai.23-11-2023.2343170.

[25] S. Hilal, S. A. Gadsden, and J. Yawney, "Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances," Expert Systems with Applications, vol. 193, 116429, 2022. doi: 10.1016/j.eswa.2021.116429.

[26] O. Kolodiziev, A. Mints, P. Sidelov, I. Pleskun, and O. Lozynska, "Automatic machine learning algorithms for fraud detection in digital payment systems," Eastern-European Journal of Enterprise Technologies, vol. 5, no. 9(107), pp. 14–26, Oct. 2020. doi: 10.15587/1729-4061.2020.212830.

[27] S. Kambhampati, "Challenges of Human-Aware AI Systems," AI Magazine, vol. 41, pp. 3–17, 2020. doi: 10.1609/aimag.v41i3.5257.

[28] A. Alazizi et al., "Anomaly Detection, Consider Your Dataset First: An Illustration on Fraud Detection," in 2019 IEEE 31st ICTAI, pp. 1351–1355, 2019. doi: 10.1109/ICTAI.2019.00188.

[29] Y.-F. Zhang, H.-L. Lu, H.-F. Lin, X.-C. Qiao, and H. Zheng, "The Optimized Anomaly Detection Models Based on an Approach of Dealing with Imbalanced Dataset for Credit Card Fraud Detection," Mobile Information Systems, 2022, 8027903. doi: 10.1155/2022/8027903.

[30] P. Ghosh, S. Kundu, and A. Chakrabarti, "Real-Time Fraud Detection Using Streaming Machine Learning," Proc. IEEE International Conference on Big Data, 2022. doi: 10.1109/BigData55660.2022.10020456

[31] S. K. Sharma, A. K. Sangaiah, S. Devarajan, and A. H. Gandomi, "A Comprehensive Survey on Fraud Detection Using Machine Learning," in Proc. IEEE EIECS, 2024. doi: 10.1109/EIECS63941.2024.10423008.

[32] T. Cao, J. Zhu, and G. Pang, "Anomaly Detection under Distribution Shift," in 2023 IEEE/CVF ICCV, pp. 6488–6500, 2023. doi: 10.1109/iccv51070.2023.00599.

[33] T. Polhul and A. Yarovyi, "Development of a method for fraud detection in heterogeneous data during installation of mobile applications," EEJET, vol. 1, no. 2, pp. 65–75, Jan. 2019. doi: 10.15587/1729-4061.2019.155060.

[34] A. Rai, P. Constantinides, and S. Sarker, "Editor's Comments: Next-Generation Digital Platforms: Toward Human-AI Hybrids," MIS Quarterly, vol. 43, no. 1, pp. iii–ix, 2019.

[35] W. M. Ikhsan, E. H. Ednoer, W. S. Kridantika, and A. Firmansyah, "Fraud Detection Automation Through Data Analytics and Artificial Intelligence," RISET, vol. 4, no. 2, pp. 103–119, Sep. 2022. doi: 10.37641/riset.v4i2.166.

[36] S. Subroto and A. Apriyana, “Cyber Risk Prediction through Social Media Big Data Analytics and Statistical Machine Learning,” Journal of Big Data, vol. 6, no. 1, Art. 50, 2019. doi: 10.1186/s40537-019-0216-1.

[37] E. O. Alonge, N. L. Eyo-Udo, B. C. Ubanadu, A. I. Daraojimba, E. D. Balogun, and K. O. Ogunsola, “Enhancing data security with machine learning: A study on fraud detection algorithms,” International Journal of Artificial Intelligence Engineering and Transformation, vol. 3, no. 1, pp. 38–49, 2022. doi: 10.54660/IJAIET.2022.3.1.38-49.

[38] M. Ahmed, A. N. Mahmood, and M. R. Islam, “A Survey of Anomaly Detection Techniques in Financial Domain,” Future Generation Computer Systems, vol. 55, pp. 278–288, 2016. doi: 10.1016/j.future.2015.01.001.

[39] A. Peacock, X. Ke, and M. Wilkerson, “Typing Patterns: A Key to User Identification,” IEEE Security & Privacy, vol. 2, no. 5, pp. 40–47, 2004. doi: 10.1109/MSP.2004.89.

[40] A. Acien, A. Morales, J. Fierrez, R. Vera-Rodriguez, and O. Delgado-Mohatar, “BeCAPTCHA: Behavioral Bot Detection Using Touchscreen and Mobile Sensors Benchmarked on HuMIdb,” Engineering Applications of Artificial Intelligence, vol. 98, 104058, 2021. doi: 10.1016/j.engappai.2020.104058.

[41] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” in Advances in Neural Information Processing Systems 30 (NeurIPS), pp. 4765–4774, 2017.