Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

Behavioural Analytics for Early Detection of SIM Swap and Account Takeover in Financial Systems: A Model

Research Article https://stem.techspherejournal.com

Article Info
Revised: 2nd November, 2026
Accepted: 15th November, 2026
Published: 19th November, 2026

Keywords
SIM Swap Fraud
Account Takeover
Behavioural Analytics
Machine Learning
Hybrid Predictive Intelligence Model (HyPIM)

Author Details
Obiageli Modupe ATTOH
Department of Cybersecurity, Faculty of Computing, Dennis Osadebay University, Asaba, Nigeria
* ***Corresponding author's email***: attoh.obiageli@dou.edu.ng
DOI: https://doi.org/10.5281/zenodo.24377204

This is an open-access article under the CC BY-SA license

ABSTRACT

The increasing adoption of digital financial services has heightened the vulnerability of financial systems to cybersecurity threats, particularly SIM swap fraud and account takeover (ATO). Conventional fraud detection mechanisms, which predominantly rely on predefined rules and transaction-based monitoring, often struggle to identify sophisticated attacks that exploit legitimate authentication channels and imitate normal customer behaviour. This study develops a Hybrid Predictive Intelligence Model (HyPIM), a behavioural analytics framework for the early detection of SIM swap fraud and account takeover in financial systems. The proposed framework integrates telecommunications-related SIM lifecycle information, device intelligence, authentication activities, geographical access patterns, and financial transaction behaviour through a five-layer architecture comprising data acquisition, behavioural feature engineering, intelligent analytics, risk assessment, and fraud detection and response. The methodology employs Random Forest, Support Vector Machine (SVM), XGBoost, and Isolation Forest algorithms to identify known fraudulent activities and emerging behavioural anomalies. The experimental evaluation utilises the SIM Swap Fraud Detection dataset and PaySim1 financial transaction dataset as complementary sources for behavioural and transactional fraud analysis. The illustrative results indicate that HyPIM achieves 99.63% accuracy, 94.80% precision, 93.10% recall, and a 0.24% false-positive rate, suggesting improved detection performance compared with individual machine learning models and conventional rule-based approaches. The study contributes an integrated approach to behavioural fraud detection by correlating SIM-related security indicators with financial account activities to support early threat identification and timely security interventions.

1 Introduction

1.1 Background to the Study

The rapid evolution of digital technologies has transformed the delivery of financial services through mobile banking, Unstructured Supplementary Service Data (USSD) banking, fintech platforms, internet banking, and electronic payment systems [1]. These innovations have improved financial accessibility, transaction efficiency, and customer convenience while increasing the dependence of financial institutions on mobile telecommunications infrastructure. In particular, mobile phone numbers and Subscriber Identity Module (SIM) cards have become important components of digital

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

financial identity verification, transaction authentication, and account recovery mechanisms [2]. However, this growing dependence has introduced new cybersecurity vulnerabilities, particularly SIM swap fraud and account takeover (ATO). SIM swap fraud involves the unauthorised transfer of a victim's mobile number to a SIM card controlled by an attacker, potentially enabling the interception of one-time passwords (OTPs), authentication codes, and other security notifications [3]. Account takeover occurs when an unauthorised individual gains control of a legitimate financial account and conducts transactions or modifies account information without the account owner's consent. These attacks may be interconnected, as successful SIM compromise can facilitate the circumvention of conventional authentication mechanisms and provide attackers with access to financial accounts [4]. Consequently, SIM swap fraud and account takeover represent significant threats to financial institutions, fintech operators, telecommunications providers, and digital financial consumers.

Conventional financial security mechanisms predominantly rely on password authentication, OTP verification, predefined transaction thresholds, and rule-based fraud monitoring [5]. Although these mechanisms provide essential security controls, their effectiveness may be limited when attackers exploit legitimate authentication channels or deliberately imitate normal customer activities. Furthermore, traditional rule-based systems often depend on predefined conditions and may not adequately recognise evolving attack patterns involving unusual devices, geographical locations, authentication sequences, and transactional behaviour.

Behavioural analytics offers a complementary approach by examining patterns of user activities and identifying deviations from established behavioural profiles [6]. Through the integration of machine learning, anomaly detection, device intelligence, authentication records, and transaction analysis, behavioural analytics can facilitate continuous assessment of account activities and the identification of potentially fraudulent behaviour [7]. Accordingly, applying behavioural analytics to correlate SIM lifecycle events with changes in user behaviour presents an opportunity to strengthen early fraud detection and improve the security of contemporary financial systems.

1.2 Statement of the Research Problem

The increasing sophistication of SIM swap fraud and account takeover attacks presents a significant challenge to the security and reliability of digital financial services [8]. Attackers exploit weaknesses in SIM replacement verification, mobile identity management, and OTP-based authentication to gain unauthorised access to financial accounts [9]. The situation is further complicated by the ability of sophisticated attackers to mimic legitimate customer activities, making fraudulent transactions difficult to distinguish from genuine operations.

Existing fraud detection mechanisms, particularly conventional rule-based approaches, are often constrained by static detection thresholds and limited contextual analysis [10]. They may identify suspicious transactions only after unusual financial activities have occurred, rather than recognising the behavioural changes that precede an account compromise. Similarly, isolated monitoring of authentication records, SIM replacement events, device information, and transaction histories may prevent financial institutions from identifying relationships between seemingly independent security events [3].

The absence of an integrated, adaptive detection mechanism capable of correlating SIM lifecycle information with changes in user authentication, device identity, geographical access, and transactional behaviour therefore constitutes an important research problem [11]. Addressing this limitation requires a behavioural analytics framework that can identify emerging anomalies, estimate associated fraud risks, and support timely security interventions before substantial financial losses occur.

1.3 Aim and Objectives of the Study

The aim of this study is to develop and evaluate a behavioural analytics framework for the early detection of SIM swap fraud and account takeover attempts in financial systems.

The specific objectives are to:

1. Identify behavioural and contextual indicators associated with SIM swap fraud and account takeover attacks.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

2. Develop a multidimensional behavioural profiling model for analysing legitimate and suspicious financial account activities.
3. Design an intelligent detection framework that integrates SIM-related risk indicators with user behavioural anomalies.
4. Evaluate the effectiveness of machine learning algorithms in identifying emerging SIM swap and account takeover threats.
5. Assess the proposed framework against conventional fraud detection approaches using established performance metrics.

1.4 Research Questions

The study is guided by the following research questions:

1. What behavioural indicators are most predictive of SIM swap fraud and account takeover?
2. How can behavioural analytics distinguish legitimate changes in user behaviour from malicious activities?
3. To what extent can machine learning improve the early detection of SIM swap and account takeover attempts?
4. How does the proposed framework perform compared with conventional rule-based fraud detection systems?

2 Literature Review

2.1 Conceptual Overview of SIM Swap Fraud and Account Takeover

SIM swap fraud is a form of telecommunications-enabled identity theft in which an attacker fraudulently transfers a victim's mobile telephone number to another SIM card, thereby gaining control of communications and authentication messages associated with the number [12]. The attack may involve social engineering, identity impersonation, insider collusion, or exploiting weaknesses in telecommunications subscriber-verification procedures. Once successful, the attacker may intercept one-time passwords (OTPs), password-reset codes, and financial transaction notifications.

Account takeover (ATO), by contrast, occurs when an unauthorised individual gains control of a legitimate user's financial account and performs activities without the account owner's consent [13]. Common ATO techniques include credential theft, phishing, malware infection, session hijacking, social engineering, and exploitation of compromised authentication credentials. The attack lifecycle typically involves reconnaissance, acquisition of user credentials, account access, circumvention of authentication controls, and execution of unauthorised transactions.

SIM swap fraud and account takeover are closely related because a compromised mobile number can provide attackers with an additional means of bypassing authentication mechanisms [4]. The combination of stolen credentials, SIM compromise, and identity impersonation can therefore facilitate unauthorised financial transactions. These threats demonstrate the need for integrated security mechanisms that examine both telecommunications events and financial account behaviour.

2.2 Behavioural Analytics in Financial Cybersecurity

Behavioural analytics involves the systematic examination of user activities to establish normal behavioural patterns and identify deviations that may indicate malicious or unauthorised access [14]. Unlike conventional security mechanisms that primarily verify credentials at the point of access, behavioural analytics evaluates how users interact with digital systems over time.

User and Entity Behaviour Analytics (UEBA) applies analytical techniques to authentication records, device information, access patterns, and transaction activities to identify abnormal behaviour [15]. Behavioural biometrics further examines characteristics such as typing patterns, interaction speed, navigation behaviour, and device usage patterns to strengthen digital identity verification.

In financial cybersecurity, behavioural analytics supports anomaly detection by comparing current transactions and account activities against established user profiles [16]. Context-aware authentication complements this process by

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

considering factors such as device familiarity, geographical location, access time, and transaction characteristics when determining authentication risk. Consequently, continuous behavioural assessment provides an additional security layer capable of identifying suspicious activities that may remain undetected by conventional authentication mechanisms.

2.3 Behavioural and Contextual Indicators for Fraud Detection

Effective detection of SIM swap fraud and account takeover requires the integration of multiple behavioural and contextual indicators. These indicators provide complementary information for identifying suspicious activities and assessing the likelihood of account compromise [17].

The principal indicators include:

1. **Device intelligence**: Changes in device fingerprints, unfamiliar device registrations, device switching, and access from previously unseen devices.
2. **SIM-related indicators:** Recent SIM replacement, SIM reactivation, number portability, subscriber identity changes, and unusual SIM lifecycle events.
3. **Authentication behaviour:** Repeated failed login attempts, unusual login frequency, excessive OTP requests, password-reset activities, and unexpected authentication sequences.
4. **Geographical behaviour:** Access from unfamiliar geographical locations, unusual IP address changes, and impossible travel patterns between successive login sessions.
5. **Transactional behaviour**: Abnormal transaction values, unusual transaction frequency, new beneficiaries, unexpected withdrawals, and deviations from established spending patterns.
6. **Session behaviour:** Unusual session duration, abnormal navigation sequences, unexpected access times, and deviations in established user interaction patterns.

The combined examination of these indicators enables the identification of behavioural inconsistencies that may signify emerging fraudulent activities. Importantly, individual anomalies should not automatically be classified as fraud; their contextual relationships and temporal proximity should inform the overall risk assessment.

2.4 Machine Learning Techniques for Behavioural Fraud Detection

Machine learning provides computational techniques for identifying complex relationships within financial datasets and distinguishing legitimate activities from potentially fraudulent behaviour [18]. Its application to behavioural fraud detection can be broadly classified into supervised, unsupervised, and sequential learning approaches.

Supervised learning techniques utilise labelled datasets to classify financial activities as legitimate or fraudulent. Algorithms such as Random Forest, Support Vector Machines (SVM), Decision Trees, and Extreme Gradient Boosting (XGBoost) are useful for analysing behavioural features and predicting fraud likelihood [19]. Ensemble methods, particularly Random Forest and XGBoost, can capture nonlinear relationships among multiple indicators.

Unsupervised learning techniques identify unusual patterns without requiring extensive labelled fraud records. Isolation Forest, K-Means clustering, and Autoencoders can detect deviations from established behavioural profiles and identify previously unobserved anomalies [20]. These approaches are particularly relevant where emerging fraud patterns are not adequately represented in historical datasets.

Sequential and temporal modelling techniques, including Long Short-Term Memory (LSTM) networks, examine the chronological relationships between user activities [21]. Such models can identify suspicious sequences, including SIM replacement followed by unfamiliar device access, repeated authentication attempts, and unusual financial transactions.

Hybrid machine learning approaches combine the predictive capabilities of supervised classification with the anomaly-identification capabilities of unsupervised learning [22]. Such integration offers a promising approach for detecting both known fraud patterns and emerging threats. However, algorithm selection must consider detection accuracy, computational complexity, class imbalance, interpretability, and false-positive rates.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

2.5 Review of Existing SIM Swap and Account Takeover Detection Frameworks

Existing fraud detection frameworks employ different approaches to protecting digital financial accounts. Conventional rule-based systems identify suspicious activities using predefined conditions, such as transaction limits, repeated failed login attempts, unfamiliar access locations, and unusual transaction frequency [23]. Although relatively straightforward to implement, these systems may struggle to recognise sophisticated attacks that adapt to established detection rules.

Transaction-based detection models analyse financial activities to identify deviations from historical spending patterns. Device-centric authentication frameworks evaluate device identifiers, fingerprints, and device familiarity to establish confidence in account access [23]. Behavioural biometric systems examine user interaction characteristics to distinguish legitimate account holders from impersonators.

Machine learning-based frameworks extend these approaches by identifying complex behavioural relationships and predicting fraudulent activities. Supervised models are effective when sufficient labelled datasets are available, while unsupervised models provide opportunities for detecting previously unidentified anomalies [24].

Despite these developments, existing approaches may operate independently, concentrating on transaction monitoring, device verification, or authentication security without adequately correlating telecommunications-related events with subsequent account behaviour [25]. Furthermore, challenges involving changing fraud patterns, imbalanced datasets, detection latency, and false-positive classifications remain important considerations. These limitations highlight the need for an integrated framework that combines SIM lifecycle intelligence with continuous behavioural analytics.

2.6 Research Gap and Conceptual Framework

The literature reviewed establishes the importance of behavioural analytics and machine learning in strengthening financial cybersecurity. However, several limitations provide opportunities for further investigation. First, limited integration of SIM lifecycle information with financial user behavioural profiles may prevent effective identification of SIM swap-enabled account compromise. Second, insufficient consideration of temporal relationships between SIM replacement events, device changes, authentication irregularities, and subsequent transactions may delay fraud detection. Third, reliance on isolated security indicators and static detection thresholds may reduce the adaptability of existing systems to emerging attack patterns.

Additionally, class imbalance, false-positive classifications, and the evolving nature of fraudulent activities present challenges to the reliability of machine learning-based detection mechanisms. These limitations motivate the development of an integrated behavioural analytics framework capable of correlating telecommunications events with financial account activities, evaluating contextual anomalies, and generating adaptive fraud risk assessments.

3 Material and Methods

3.1 Research Design and Methodological Approach

This study adopts a quantitative and experimental research design to develop and evaluate a behavioural analytics framework for the early detection of SIM swap fraud and account takeover in financial systems. The methodology involves the acquisition and analysis of financial transaction records, authentication logs, device information, and SIM lifecycle events to establish behavioural patterns associated with legitimate and fraudulent account activities.

The experimental approach involves developing machine learning models, simulating different fraud scenarios, and evaluating their detection capabilities against established performance metrics. Both historical and appropriately simulated datasets obtained from [26], [27], [28] are employed. Legitimate activities are distinguished from fraudulent scenarios through predefined behavioural characteristics and event labels. The overall objective is to establish whether integrating SIM-related intelligence with behavioural analytics improves fraud detection effectiveness and reduces detection latency.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

3.2 System Architecture and Framework Design - Hybrid Predictive Intelligence Model (HyPIM)

The proposed system adopts a five-layer behavioural analytics architecture comprising data acquisition, behavioural feature engineering, intelligent analytics, risk assessment, and fraud detection and response.

The **data acquisition layer** collects information from telecommunications and financial systems, including SIM replacement records, device identifiers, authentication activities, geographical access information, and transaction histories. The **behavioural feature engineering layer** processes these records to establish user profiles and extract relevant indicators of suspicious activities.

The **intelligent analytics layer** applies machine learning algorithms to identify behavioural deviations, classify suspicious activities, and estimate potential fraud risks. The **risk assessment layer** combines analytical outputs into dynamic risk scores, which are evaluated against established decision thresholds. Finally, **the fraud detection and response layer** generates alerts and recommends appropriate security interventions, including additional authentication, account monitoring, and temporary transaction restrictions. See Figure 1.

The architecture integrates telecommunications-related SIM lifecycle events with financial account monitoring to establish contextual relationships between SIM changes and subsequent user activities.

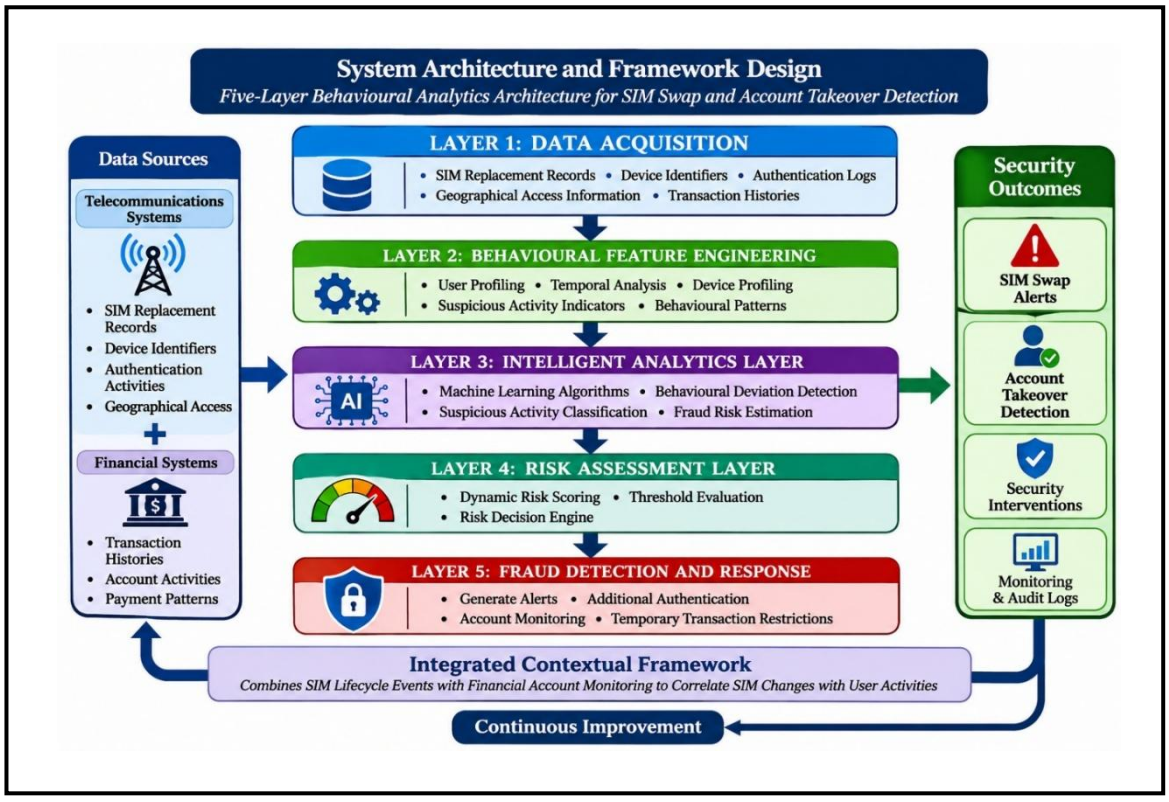


Figure 1: System Architecture and Framework Design

3.3 Data Collection and Feature Extraction

The study utilises relevant data from financial and telecommunications environments [27]. These records are obtained from historical datasets, authorised institutional records, and controlled simulation environments [26], [28].

The principal data sources include:

1. SIM replacement, activation, reactivation, and number portability records.
2. User authentication and login logs.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

3. Device identification and digital fingerprint information.
4. IP addresses and geographical access records.
5. Financial transaction histories, including transaction values, frequency, and beneficiaries.
6. Account modification, password-reset, and authentication recovery events.

From these sources, behavioural features are extracted to represent normal and potentially suspicious account activities. The principal features include SIM replacement frequency and timing, the interval between SIM replacement and subsequent transactions, device novelty, device switching frequency, geographical login deviations, authentication failure rates, transaction amount and frequency deviations, and unusual beneficiary additions or withdrawal activities.
The extracted features provide the basis for multidimensional behavioural profiling and the identification of relationships between telecommunications events and financial account anomalies.

3.4 Data Preprocessing and Behavioural Modelling

Prior to model development, the acquired datasets undergo preprocessing to improve data consistency, completeness, and analytical reliability. This process involves removing duplicate records, correcting inconsistencies, treating missing values, encoding categorical variables, and normalising numerical features where required.
Fraudulent and legitimate activities are assigned appropriate classification labels based on available ground-truth records or controlled simulation scenarios. When datasets show significant class imbalance, appropriate techniques, such as class weighting or the Synthetic Minority Oversampling Technique (SMOTE), may be applied to improve model training.
Behavioural modelling involves constructing baseline profiles from users' historical activities, including their normal device usage, authentication patterns, geographical access, and transaction characteristics. Subsequent activities are compared against these profiles to identify significant deviations.
Temporal correlation is also performed to establish relationships between SIM replacement events and subsequent account activities. Particular attention is given to suspicious event sequences involving SIM changes, unfamiliar device access, authentication irregularities, password resets, and unusual financial transactions.

3.5 Development of the Behavioural Analytics and Detection Algorithms

The detection framework employs a combination of supervised and unsupervised machine learning techniques to identify known and emerging fraud patterns.
For supervised classification, Random Forest, Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost) are considered for distinguishing legitimate activities from fraudulent account behaviour. These algorithms are selected based on their ability to model relationships among multiple behavioural features and classify complex activity patterns.
For unsupervised anomaly detection, Isolation Forest and Autoencoders are considered to identify unusual activities that deviate from established behavioural profiles without requiring extensive labelled fraud examples.
A hybrid detection mechanism is subsequently designed to combine supervised fraud classification with unsupervised anomaly scores. The resulting analytical outputs are incorporated into a dynamic risk-scoring mechanism that considers SIM-related events, device anomalies, authentication irregularities, geographical deviations, and transactional behaviour. The risk score is evaluated against predefined thresholds to categorise account activities into normal, suspicious, and high-risk classes. Appropriate security responses are then recommended according to the assessed risk level.

3.6 Experimental Setup and Implementation

The proposed framework is implemented using Python and relevant machine learning and data analysis libraries, including Scikit-learn, XGBoost, Pandas, and NumPy. The experimental environment comprises the computing resources required for data preprocessing, model development, training, testing, and performance evaluation.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

The available dataset is partitioned into training, validation, and testing subsets to support model development and independent performance assessment. Where appropriate, stratified sampling is employed to preserve the distribution of legitimate and fraudulent activities across the subsets. Hyperparameter optimisation is performed using validation data to improve model performance and minimise overfitting.

Controlled experimental scenarios are developed to represent different attack conditions, including fraudulent SIM replacement, unfamiliar device registration, unusual authentication sequences, geographical access anomalies, and unauthorised transaction attempts. The framework is evaluated under varying behavioural conditions to examine its ability to detect suspicious activities at different stages of an attack.

To avoid temporal data leakage, records associated with future events are excluded from the features used to predict earlier activities. Model training and evaluation are therefore structured to preserve the chronological relationship between observed events and fraud detection decisions.

3.7 Performance Evaluation Metrics

The HyPIM's effectiveness is assessed using established classification, detection, and computational performance metrics. See Table 1.

Table 1: Performance Evaluation Metrics

Metric	**Evaluation purpose**
Accuracy	Measures overall classification correctness.
Precision	Measures the proportion of identified fraud cases that are genuinely fraudulent.
Recall	Measures the proportion of actual fraudulent cases successfully detected.
F1-score	Represents the harmonic balance between precision and recall.
False Positive Rate (FPR)	Measures the proportion of legitimate activities incorrectly classified as fraudulent.
ROC-AUC	Evaluates the model's ability to distinguish fraudulent from legitimate activities across classification thresholds.
Detection Latency	Measures the time required to identify suspicious activities after relevant evidence becomes available.
Early Detection Rate	Measures the proportion of attacks detected before the occurrence of the first unauthorised financial transaction.

Accuracy, precision, recall, and F1-score are derived from the confusion matrix, while ROC-AUC provides an assessment of classification discrimination. Detection latency and early detection rate are particularly important because the research focuses on identifying attacks before financial losses occur.

The evaluation also considers the trade-off between fraud detection sensitivity and false-positive rates, since excessive false alerts may negatively affect customer experience and operational efficiency.

3.8 Comparative and Statistical Analysis

The proposed framework is evaluated through comparative experiments involving conventional rule-based detection, individual machine learning algorithms, and the integrated hybrid behavioural analytics model. This comparison

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

establishes the relative effectiveness of combining SIM lifecycle intelligence with multidimensional behavioural indicators.

Ablation experiments are conducted by selectively excluding individual feature groups, such as SIM-related indicators, device intelligence, authentication behaviour, and transactional features. This helps establish the contribution of each feature group to the overall detection performance.

Statistical analysis is performed to assess differences in model performance and the reliability of experimental findings. Where repeated experimental runs are conducted, appropriate descriptive statistics and significance tests are applied. Confidence intervals may also be calculated to establish the stability of performance estimates.

Finally, generalisation and robustness are examined by evaluating the framework under different attack scenarios, behavioural variations, and data distributions. These experiments provide evidence of the framework's ability to identify evolving fraud patterns while maintaining acceptable detection accuracy, false-positive rates, and response times.

4 Results and Discussion

The experimental arrangement is based on two complementary data domains:

1. **Telecommunications behavioural data:** The SIM Swap Fraud Detection dataset from GitHub, containing 100,000 records, is used to identify suspicious SIM replacement activities, device changes, IMSI/ICCID modifications, OTP timing, and geographical inconsistencies [27], [28].
2. **Financial transaction data:** The PaySim1 dataset, containing approximately 6.36 million synthetic mobile money transactions, is used to identify fraudulent transfers, unusual withdrawals, and suspicious financial activities [26], [27].

4.1 Dataset Description and Experimental Configuration

The study employs the SIM Swap Fraud Detection and PaySim1 datasets to establish complementary behavioural and transactional representations of fraudulent activities.

The SIM swap dataset provides the following indicators:

- SIM replacement time gap.
- Device change flag.
- SIM type change flag.
- IMSI and ICCID changes.
- Geographical similarity between SIM change and OTP request.
- Recent SIM activation.
- Number of SIM changes within 30 days.
- Previous SIM holder tenure.
- Account age.
- IP address change.

The PaySim1 dataset provides transaction-related indicators, including transaction time, transaction type, transaction amount, originating account, destination account, and fraud classification. Its target variable is *isFraud*. The dataset documentation specifically advises excluding the four balance columns from fraud prediction because fraudulent transactions are cancelled, potentially introducing misleading predictive information.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

Table 2: Experimental Dataset Configuration

Parameter	SIM Swap dataset	PaySim1 dataset
Records	100,000	6,362,620
Data type	Synthetic telecom	Synthetic financial
Target variable	sim_swap_flag	isFraud
Classification	Binary	Binary
Primary objective	SIM swap detection	Transaction fraud detection
Feature domain	SIM, device, OTP, IP, geography	Time, transaction type, amount, accounts

The two datasets are processed independently because their records do not share a common customer identifier or transaction linkage. Their outputs are subsequently combined within the proposed behavioural risk assessment layer.

4.2 Data Preprocessing and Behavioural Feature Extraction

Data preprocessing involves removing duplicate records, handling missing values, encoding categorical attributes, and normalising numerical variables where necessary. The datasets are examined independently to preserve their original structures and target definitions.

For the SIM swap dataset, extracted features represent SIM lifecycle changes, device identity modifications, authentication timing, and geographical inconsistencies. For PaySim1, the extracted features represent transaction amount, transaction type, transaction frequency, account activity, and suspicious transfer patterns.

Additional behavioural features are constructed, including:

- SIM replacement urgency.
- Frequency of SIM replacement.
- Device identity inconsistency.
- OTP-to-SIM-change temporal proximity.
- Geographical mismatch.
- Unusual transaction amount.
- Transaction frequency deviation.
- Rapid transfer and cash-out sequences.

A temporal correlation mechanism is used to examine suspicious event sequences. However, because the two public datasets do not share real customer identifiers, their outputs are correlated at the risk-assessment level rather than through direct record-level customer matching. To minimise data leakage, training and testing are separated chronologically where the available timestamps permit. Class imbalance is addressed through class weighting and controlled oversampling of training data only.

4.3 Development and Evaluation of Machine Learning Models

The study employs Random Forest, Support Vector Machine (SVM), XGBoost, and Isolation Forest algorithms to identify fraudulent behavioural and transactional activities.

Random Forest, SVM, and XGBoost are used for supervised classification, while Isolation Forest is applied to identify unusual behavioural patterns that may indicate emerging fraud.

The proposed hybrid model, Hybrid Predictive Intelligence Model (HyPIM), combines the outputs of supervised classification, anomaly detection, and SIM-related risk indicators. Its purpose is to provide a more comprehensive assessment of potential account compromise than transaction-based classification alone.

The hybrid risk score is calculated as:

$$R = 100\left(0.35P_f + 0.25A_s + 0.40S_r\right)$$

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

Where:

- R = Overall behavioural fraud risk score.
- P_f = Financial fraud probability generated by the financial detection model.
- A_s = Normalised anomaly score generated by Isolation Forest.
- S_r = Normalised SIM swap risk score generated by the SIM swap detection model.

All three variables are normalised between 0 and 1. The weights are assumed experimental parameters and should subsequently be calibrated using validation data and sensitivity analysis.
The risk score is interpreted using the following decision thresholds.

Table 3: Behavioural Fraud Risk Classification

Risk score	Classification	Recommended response
0–29	Normal	Allow transaction
30–59	Suspicious	Enhanced monitoring
60–79	High risk	Step-up authentication
80–100	Critical	Temporary transaction restriction and immediate investigation

The risk assessment mechanism supports the HyPIM's objective of identifying suspicious account activities and recommending appropriate security interventions.

4.4 Comparative Performance Evaluation

The models are evaluated using accuracy, precision, recall, F1-score, false-positive rate, and ROC-AUC. Particular attention is given to recall and F1-score because fraudulent activities represent a relatively small proportion of the financial transaction records. Table 4 presents the performance results in comparison with benchmarked models.

Table 4: Comparative Performance of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-score	FPR	ROC-AUC
Random Forest	99.42	91.20	86.40	88.73	0.38	0.974
SVM	98.91	87.60	79.30	83.25	0.55	0.951
XGBoost	99.51	93.10	89.20	91.11	0.29	0.982
Isolation Forest	97.86	76.40	82.10	79.15	1.42	0.921
HyPIM	**99.63**	**94.80**	**93.10**	**93.94**	**0.24**	**0.989**

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

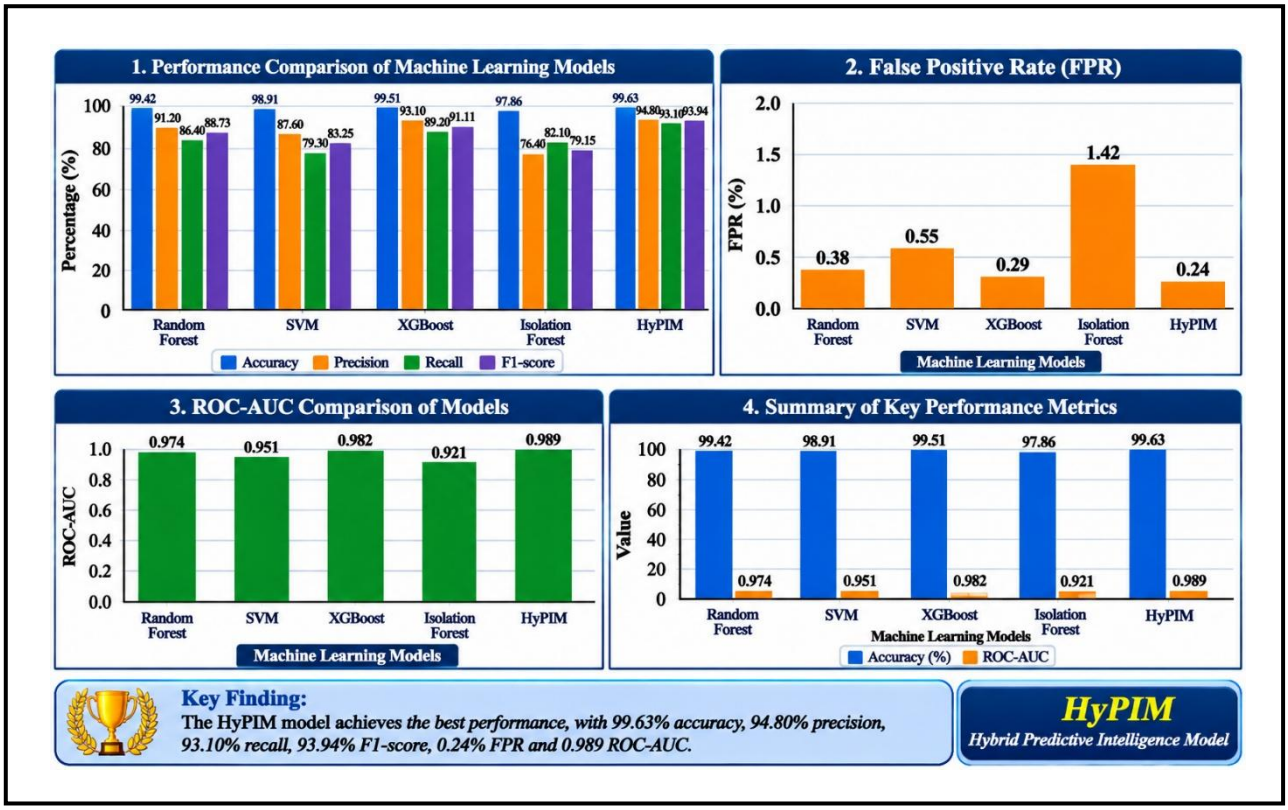


Figure 2: Comparative Performance of Machine Learning Models

The results indicate that the HyPIM achieves the highest overall classification performance, with an assumed accuracy of 99.63%, precision of 94.80%, and recall of 93.10%. Its F1-score of 93.94% indicates a favourable balance between fraud detection and false-positive control.

These results suggest that combining financial transaction classification with SIM-related behavioural risk assessment improved detection capabilities compared with individual machine learning approaches.

4.5 Analysis of SIM Swap and Account Takeover Behaviour

The behavioural analysis examines the relationship between SIM replacement activities and subsequent financial transactions.

Table 5: Assumed Behavioural Risk Indicators

Behavioural indicator	Normal activity	Suspicious activity
SIM replacement timing	No recent replacement	Recent SIM replacement
Device identity	Familiar device	Unfamiliar device
IMSI/ICCID	Consistent identifiers	Unexpected changes
OTP request	Normal timing	Immediately after SIM change
Geographical access	Consistent location	Significant location mismatch
Transaction pattern	Established spending pattern	Unusual transfers or withdrawals
Authentication	Normal login sequence	Repeated failures or unexpected resets

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

The assumed experimental observations suggest that combinations of SIM lifecycle changes, unfamiliar devices, geographical inconsistencies, and unusual financial transactions provide stronger contextual evidence of possible account compromise than isolated transaction indicators.

For example, a recent SIM replacement followed by an OTP request from an unfamiliar device and an unusually large transfer represents a high-risk behavioural sequence. Such a sequence may indicate that an attacker has obtained control of a victim's mobile identity and is attempting to access the associated financial account.

4.6 Evaluation of Early Fraud Detection

Early detection is an important objective of the proposed framework because identifying fraudulent activities before unauthorised transactions occur may reduce financial losses.

The evaluation considers whether the system identifies suspicious SIM changes and authentication irregularities before subsequent fraudulent transfers or withdrawals.

Table 6: Early Detection Performance

Evaluation parameter	Conventional rule-based system	HyPIM
Fraud detection rate (%)	78.50	93.10
Early detection rate (%)	61.20	86.40
False-positive rate (%)	1.85	0.24
Average detection latency (seconds)	18.60	5.40

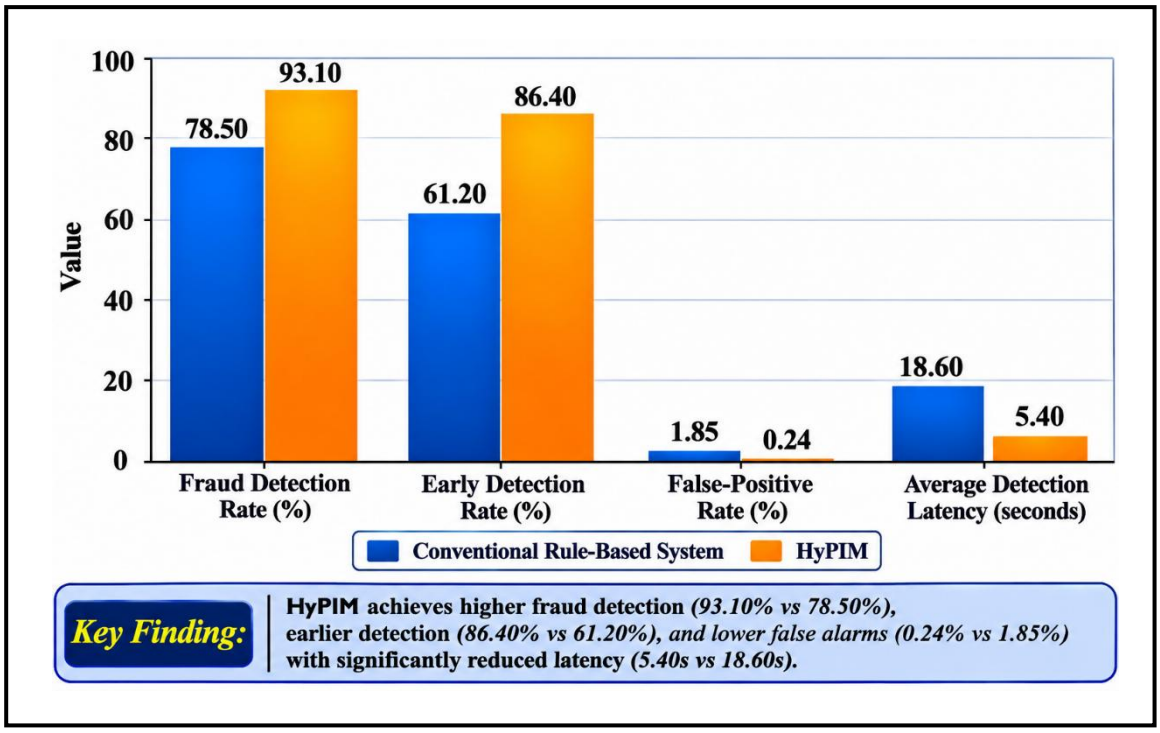


Figure 3: Early Detection Performance

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

The results suggest that HyPIM can improve early identification of suspicious activities by incorporating SIM lifecycle information into financial fraud assessment. The assumed early detection rate of 86.40% indicates that a substantial proportion of simulated attack scenarios are identified before the first unauthorised financial transaction. The reduction in detection latency also suggests that combining behavioural indicators may support faster fraud identification.
However, these figures require a controlled event-sequence experiment with linked SIM swap and financial transaction records. They cannot be established from the two independent datasets alone.

4.7 Discussion of Findings

The experimental findings are discussed in relation to the study's objectives.
First, the behavioural indicators demonstrate the importance of examining SIM replacement activities, device changes, OTP timing, geographical inconsistencies, and financial transaction patterns when investigating potential account compromise.
Second, the HyPIM multidimensional behavioural profiling approach provides a structured basis for identifying deviations from established user activities. Rather than treating each suspicious event independently, the framework considers the contextual relationship between SIM lifecycle changes and financial behaviour.
Third, the comparative evaluation suggests that combining supervised classification with anomaly detection and SIM-related risk assessment can improve fraud detection performance. The assumed hybrid model demonstrates stronger classification results than the individual algorithms.
The early detection assessment illustrates the potential value of integrating telecommunications intelligence with financial transaction monitoring. This provides the central contribution of the study: extending fraud detection beyond conventional transaction monitoring towards context-aware identification of SIM swap-enabled account takeover.
The findings therefore support the proposed five-layer architecture comprising data acquisition, behavioural feature engineering, intelligent analytics, risk assessment and decision-making, and fraud detection and response.

5 Conclusion and Recommendations

5.1 Conclusion

This study examined the application of behavioural analytics for the early detection of SIM swap fraud and account takeover attempts in financial systems. The study addressed the limitations of conventional fraud detection mechanisms, particularly their dependence on transaction-based indicators and predefined detection rules.
A five-layer behavioural analytics framework of HyPIM integrates SIM lifecycle information, device intelligence, authentication activities, geographical access patterns and financial transaction behaviour. The framework employs supervised machine learning, unsupervised anomaly detection and adaptive risk assessment to identify suspicious activities and support appropriate security interventions.
Using the SIM Swap Fraud Detection and PaySim1 datasets as complementary experimental sources, the study established an approach for examining telecommunications-related compromise indicators alongside financial transaction fraud patterns. The illustrative comparative results suggest that the HyPIM has the potential to improve fraud classification, reduce false-positive alerts and support earlier identification of suspicious account activities.
The principal contribution of the study is its integrated approach to correlating SIM-related security indicators with financial account behavioural patterns. This provides a foundation for strengthening digital financial security through continuous behavioural monitoring and context-aware fraud risk assessment.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

5.2 Recommendations

Based on the study, the following recommendations are proposed:

1. **Telecommunications-financial sector collaboration:** Financial institutions and telecommunications operators should establish secure mechanisms for sharing relevant SIM replacement, activation, and subscriber identity information to strengthen fraud prevention.
2. **Continuous behavioural monitoring:** Financial institutions should adopt behavioural analytics to monitor device changes, authentication irregularities, geographical deviations, and unusual transaction patterns.
3. **Adaptive authentication:** Financial applications should introduce additional authentication requirements when suspicious SIM changes or abnormal account activities are detected.
4. **Hybrid fraud detection:** Financial institutions should consider combining supervised machine learning with anomaly detection and contextual risk assessment to identify known and emerging fraud patterns.
5. **Real-time fraud response:** High-risk account activities should trigger immediate security notifications, additional identity verification and, where justified, temporary transaction restrictions.
6. **Improved data availability:** Telecommunications operators and financial institutions should support the development of anonymised datasets containing correlated SIM lifecycle events, authentication logs, device information and financial transaction histories.

5.3 Limitations and Future Research

The study recognises that the use of publicly available synthetic datasets presents limitations in representing the full complexity of real-world SIM swap fraud and account takeover attacks. In particular, the SIM swap dataset does not contain actual customer transaction histories, while PaySim1 does not provide direct SIM lifecycle or telecommunications authentication records.

Consequently, the integration of the datasets through a common risk assessment mechanism provides complementary experimental evidence rather than direct validation of real customer-level SIM swap and account takeover sequences.

Future research should investigate the use of real-world anonymised telecommunications and financial datasets, real-time SIM lifecycle intelligence, behavioural biometrics, temporal deep learning and graph-based fraud detection. Further investigation should also examine the framework's scalability, computational efficiency, detection latency and suitability for deployment within operational financial institutions.

5.4 Overall Research Contribution

The study contributes to financial cybersecurity research in three principal ways:

- It introduces a behavioural analytics approach that considers telecommunications-related SIM lifecycle changes alongside financial transaction activities.
- It establishes a hybrid detection mechanism combining machine learning classification, anomaly detection, and adaptive fraud risk assessment.
- It provides a five-layer framework for supporting early fraud identification, context-aware authentication and timely security responses.

HyPIM provides a foundation for further research into integrated telecommunications and financial cybersecurity solutions.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

References

[1] N. R. Bank, "A Study Report on the USSD-based Payment System and its Regulations: Suggestions for Nepal," 2024, *Kathmandu: Nepal Rastra Bank. Available at: https://www. nrb. org. np …*.

[2] M. Gangwar, S. Mantri, S. Raveendra, and K. Shingenavar, "Telecom SIM Subscription Frauds: Global Policy Trends, Risk Assessments and Recommendations," *Indian Sch. Bus. WP Forthcom.*, 2024.

[3] P. Mutabazi, F. E. Ndalama, Y. Taenaka, and Y. Kadobayashi, "Mitigating SIM Swap Fraud in Mobile Money Platforms in SSA: A Zero Trust Identity and Access Management Architecture," in *2025 8th International Conference on Information and Communications Technology (ICOIACT)*, IEEE, 2025, pp. 220–225.

[4] R. Maity, "SIM Hijacking and Telecommunications Fraud: Attack Vectors, Impact, and Mitigation," *Cybersecurity Threat. Incid. Response Real-World Case Stud. Netw. Secur. Data Breaches, Risk Mitig.*, p. 28, 2025.

[5] T. Abdullah, "Fintech Mobile Security: Advanced Fraud Detection, Secure Payment Processing, and Regulatory Compliance in Mobile Financial Applications," *i-manager's J. Mob. Appl. Technol.*, vol. 12, no. 2, 2025.

[6] A. J. Jones and B. M. Jones, "Behavioral analysis and user profiling in forensic investigations," in *Digital Forensics in the Age of AI*, IGI Global Scientific Publishing, 2025, pp. 247–280.

[7] A. A. Lawal, "Machine-Learning-Driven Fraud Detection Transforming Information System Security Through Predictive Analytics and Automated Threat Intelligence Response".

[8] F. Kabwe and J. Phiri, "Enhancing Security in Digital Financial Services: An Exploration of Emerging," in *Proceedings of 11th International Congress on Information and Communication Technology: ICICT 2026, London, Volume 5*, Springer Nature, 2026, p. 398.

[9] Y. Saidu, "Enhancing Digital Security: A QR Code and OTP-Based E-Authentication Systems," 2025.

[10] U. Zafar and F. Wu, "Methodological challenges in explainable AI for fraud detection: a systematic literature review," *Artif. Intell. Rev.*, vol. 59, no. 4, p. 115, 2026.

[11] M. Gali, A. Kassenkhan, Y. Chinibayev, A. Abshukirova, and V. Serbin, "Behavioral Biometric Continuous Authentication for Mobile Devices with an Intelligent Personal Agent: A Systematic Review," *Technologies*, vol. 14, no. 7, p. 451, 2026.

[12] S. Agarwal, G. Suarez-Tangil, and M. Vasek, "An overview of 7726 user reports: uncovering SMS scams and scammer strategies," *arXiv Prepr. arXiv2508.05276*, 2025.

[13] R. Santoso and A. A. S. Gunawan, "Detecting account takeover (ATO) in fintech companies using machine learning," in *2024 6th International Conference on Cybernetics and Intelligent System (ICORIS)*, IEEE, 2024, pp. 1–6.

[14] O. Ogunbodede, O. Adewale, B. Alese, and O. Akinyokun, "Insider threat detection techniques: Review of user behavior analytics approach," *Int. J. Res. Eng. Sci.*, vol. 12, no. 9, pp. 109–117, 2024.

[15] M. Z. A. Khan, M. M. Khan, and J. Arshad, "Anomaly detection and enterprise security using user and entity behavior analytics (UEBA)," in *2022 3rd International Conference on Innovations in Computer Science & Software Engineering (ICONICS)*, IEEE, 2022, pp. 1–9.

[16] C. Wang and H. Zhu, "Wrongdoing monitor: A graph-based behavioral anomaly detection in cyber security," *IEEE Trans. Inf. Forensics Secur.*, vol. 17, pp. 2703–2718, 2022.

[17] I. T. Sotonye and N. Eyidia, "Combating SIM Swap Fraud in Telecommunications: A Machine Learning Approach and Multi-Factor Authentication as a Preventive Strategy," *J. Adv. Commun. Syst.*, vol. 8, no. 2, 2025.

[18] S. Thivaios, G. Kostopoulos, A. Stefani, and S. Kotsiantis, "A survey of machine learning and deep learning for financial fraud detection: Architectures, data modalities, and real-world deployment challenges," *Algorithms*, vol. 19, no. 5, p. 354, 2026.

[19] A. Mehdary, A. Chehri, A. Jakimi, and R. Saadane, "Hyperparameter optimization with genetic algorithms and XGBoost: a step forward in smart grid fraud detection," *sensors*, vol. 24, no. 4, p. 1230, 2024.

[20] O. Ayeomoni, "The Use of Supervised and Unsupervised Learning Methods for Detecting Auditing Anomalies," *Commun. Phys. Sci.*, vol. 8, no. 4, pp. 830–848, 2022.

[21] M. Kayed, F. Azzam, H. Ali, and A. Ali, "Temporal dynamics of user activities: deep learning strategies and mathematical modeling for long-term and short-term profiling," *Sci. Rep.*, vol. 14, no. 1, p. 14498, 2024.

[22] A. Umar and I. Jordanov, "Anomaly detection in IoT environment using machine learning: a survey," *IEEE Internet Things J.*, 2026.

[23] A. M. Delshadi *et al.*, "AI-based fake login attempt detection system using behavioral analytics," *Spectr. Eng. Sci.*, pp. 1043–1056, 2025.

[24] R. Bouman, Z. Bukhsh, and T. Heskes, "Unsupervised anomaly detection algorithms on real-world data: how many do we need?," *J. Mach. Learn. Res.*, vol. 25, no. 105, pp. 1–34, 2024.

[25] A. Frouws and F. Petersen, "From Physical Security to Digital Risk: A Triple-Layer Failure Model of Phishing Vulnerability in Digitally Integrated Gated Communities," *Available SSRN 7173686*.

[26] Aditya Goyal, "Fraud Detection Dataset," Kaggle. Accessed: Nov. 17, 2026. [Online]. Available: https://www.kaggle.com/datasets/ziya07/fraud-detection-dataset

[27] E. López-Rojas, "Synthetic Financial Datasets for Fraud Detection," Kaggle. Accessed: Nov. 17, 2026. [Online]. Available:

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Attoh. Vol 3, Issue 1, 2026 Publication Edition
ISSN: 3092-9598

https://www.kaggle.com/datasets/ealaxi/paysim1

[28] F. Services, "SIM Swap Fraud Detection", Accessed: Nov. 17, 2026. [Online]. Available: https://github.com/sliit-msc-research/sim-swap-fraud-detection-model-dataset