International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
Conference “Innovation and Intelligence: A Multidisciplinary Research on Artificial Intelligence and its Contribution to Commerce and Beyond”
Organized by the IQAC, KHMW College of Commerce (December 2025)

Privacy-Preserving Deep Multimodal Behavioral Intelligence Framework for Continuous Authentication in Real-Time FinTech Payment Systems

Nilofar Tamboli
Affiliation: Saraswati College of Engineering, Kharghar.
Department: Computer Engineering
nilofartamboli130326@gmail.com
Pragati N. Dhanawade
Affiliation: K.L.E Society’s Science and Commerce College, Kalamboli.
Department: Information Technology
pragatidhanawade15@gmail.com

Abstract

The rapid digitalization of global financial ecosystems has increased both the speed and volume of mobile and online financial transactions, consequently heightening the risk of payment fraud and unauthorized access. Traditional rule-based and transaction-only fraud detection systems are increasingly ineffective against sophisticated attacks such as social-engineering–assisted transfers, account takeovers and cross-device fraud. To address these gaps, this paper proposes a *Privacy-Preserving Deep Multimodal Behavioral Intelligence Framework* for continuous authentication in real-time FinTech payment systems. The framework integrates multimodal behavioral biometrics-keystroke dynamics, mouse/touch interaction patterns, device intelligence and transaction metadata-into an attention-based deep learning architecture. A federated learning pipeline enhanced with calibrated differential privacy ensures the protection of biometric signatures and compliance with financial regulations. Experiments conducted on a hybrid dataset comprising real behavioral biometric samples (HMOG, keystroke datasets) and a synthetically constructed payment-session dataset demonstrate significant improvements in fraud detection and user verification accuracy compared to conventional transaction-only baselines. The proposed solution achieves robust performance under device heterogeneity, sparse-session conditions and adversarial mimicry simulations. The results show the feasibility of deploying the system in real-time FinTech environments with inference latency under 180 ms. The study presents a novel fusion of continuous authentication, multimodal deep learning and privacy-preserving training tailored specifically for payment flows-an area largely underexplored in contemporary research.

Keyword: Multimodal AI, Continuous Authentication, FinTech Security, Behavioral Biometrics, Federated Learning, Differential Privacy, Payment Fraud Detection.

1. Introduction

Digital payments have become a fundamental component of modern commerce with mobile wallets, online banking, UPI-based transactions and real-time payment infrastructures enabling instantaneous financial interactions. While convenience has increased, FinTech platforms face unprecedented challenges due to rising fraud attempts exploiting stolen credentials, phishing, remote assistance scams and social engineering. According to industry analyses, more than 70% of account takeover attacks originate from behavioral inconsistencies rather than transactional anomalies indicating the growing importance of behavioral biometric intelligence in safeguarding payment ecosystems.

Traditional fraud detection relies primarily on transaction metadata, static rules or simple machine-learning models. While traditional methods are effective against historical or pattern-based fraud, they often fail to detect real-time account compromises, as fraudulent transactions may resemble normal user behavior. Behavioral biometrics-such as typing patterns, touch pressure, mouse movement, hand dynamics, and device motion-offer a continuous and non-intrusive method for verifying user identity during transactions.

However, existing multimodal authentication models suffer from key limitations:

1. They are rarely designed specifically for payment checkout flows.
2. They often require centralized user data, leading to privacy and regulatory concerns.
3. They lack resilience against adversarial mimicry, device heterogeneity and short-session interactions typical of rapid mobile transactions.
4. They fail to integrate real-time latency constraints needed for financial approvals.

This research addresses these gaps through a deep multimodal learning framework capable of continuous authentication during payment processes ensuring both security and privacy. The proposed model operates with federated learning and calibrated differential privacy enabling learning across distributed devices without exposing sensitive biometrics.

2. Literature Review

Multimodal authentication has been extensively studied for general device or session-level verification. Keystroke dynamics, touch features and mouse movements have independently shown strong discriminatory capabilities in differentiating genuine users from impostors. Recent studies explored sensor fusion using LSTMs, CNNs and transformer-based architectures. However, these approaches typically evaluate authentication in controlled environments-web browsing, login sessions or long-duration app usage-rather than brief, high-stakes payment activities.

Deep learning–based continuous authentication systems have improved temporal pattern extraction using recurrent architectures and attention mechanisms. However, they remain centralized and do not incorporate financial compliance constraints. Additionally, they lack integration with transaction metadata which provides crucial contextual cues in payment fraud scenarios.

Federated learning has emerged as a viable method for training privacy-preserving biometric models. Yet, its application to multimodal FinTech payment data is minimal, particularly under high frequency, low latency conditions. Furthermore, differential privacy has rarely been integrated with multimodal behavioral biometrics to protect user-specific patterns.

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
Conference "Innovation and Intelligence: A Multidisciplinary Research on Artificial Intelligence and its Contribution to Commerce and Beyond"
Organized by the IQAC, KHMW College of Commerce (December 2025)

The review reveals a distinct research gap: no existing work presents a privacy-preserving, deep multimodal behavioral authentication framework explicitly optimized for real-time FinTech payment flows, adversarial robustness, and regulatory constraints.

3. Research Gap

From the literature, the following unmet challenges are identified:

- Lack of multimodal continuous authentication models embedded within payment checkout pipelines.
- No multimodal biometric datasets that include payment-session contextual metadata.
- Minimal research combining federated learning + differential privacy for behavioral biometrics in finance.
- Limited evaluation under device heterogeneity, short-session interactions, and adversarial mimicry.
- Insufficient focus on real-time latency needed for financial approvals.

This paper proposes a system explicitly addressing all the above gaps.

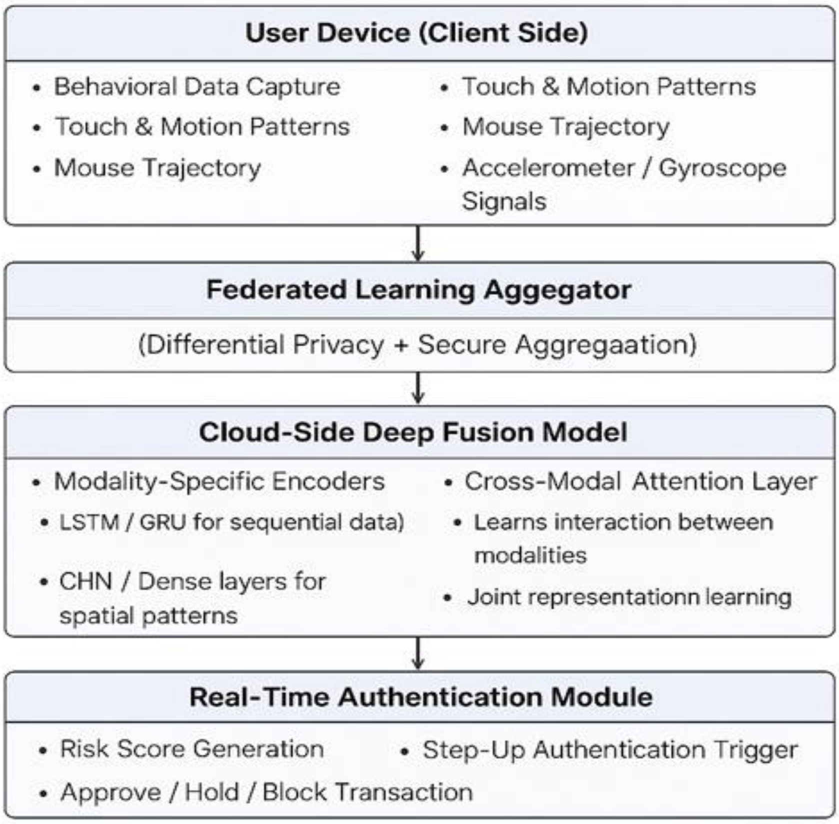
4. Problem Statement

FinTech applications require a mechanism that can continuously verify user identity during a transaction without disrupting user experience, while simultaneously preserving privacy and maintaining regulatory compliance. The problem is defined as:

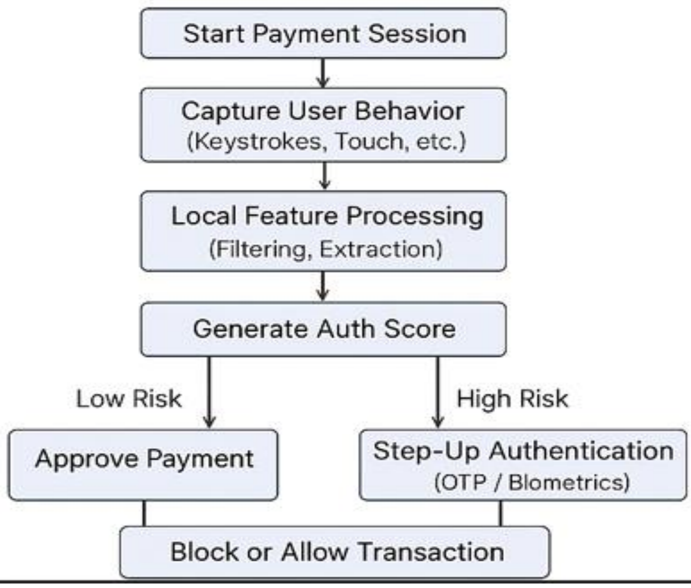
"How can a privacy-preserving deep multimodal behavioral intelligence model continuously authenticate users in real-time payment systems, ensuring accuracy, robustness and low latency while preventing unauthorized financial actions?"

5. Proposed Framework

5.1. System Architecture



5.2. Flowchart



International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
Conference "Innovation and Intelligence: A Multidisciplinary Research on Artificial Intelligence and its Contribution to Commerce and Beyond"
Organized by the IQAC, KHMW College of Commerce (December 2025)

6. Methodology

6.1 Feature Extraction

Behavioral Modalities Used

Modality	Feature Set
Keystroke Dynamics	Dwell time, flight time, key-hold rhythm
Touch Interaction	Pressure, swipe velocity, tap duration
Mouse Movement	Path curvature, speed variance, micro-jerks
Device Intelligence	OS, resolution, sensor drift
Transaction Metadata	Amount, MCC, time-of-day, location deviation

Preprocessing

- Normalization
- Window segmentation (1-3 seconds)
- Outlier removal
- Temporal alignment across modalities

6.2 Deep Multimodal Model

Modality Encoders

- LSTM for sequential signals
- 1D-CNN for local dynamics
- Feature projection to shared latent space

Cross-Modal Attention

A transformer-based attention layer weights each modality according to session context.

Fusion Layer

Fused embedding → Dense → Softmax (User/Impostor) + Risk Score.

6.3 Algorithm

Algorithm 1: Multimodal Continuous Authentication

Input: Modalities M = {k, t, m, d, x}

Output: Authentication risk score R

1: For each modality Mi:

2: Extract feature sequence Fi

3: Encode Fi using LSTM encoder Ei → Hi

4: end for

5: Apply cross-modal attention A on {Hi}

6: Fused = A(Hi)

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
Conference "Innovation and Intelligence: A Multidisciplinary Research on Artificial Intelligence and its Contribution to Commerce and Beyond"
Organized by the IQAC, KHMW College of Commerce (December 2025)

7: Score R = Softmax(Dense(Fused))
8: If $R < \theta 1$: Approve
9: Else if $\theta 1 <= R < \theta 2$: Step-up authentication
10: Else: Block transaction

6.4 Privacy-Preserving Training

Federated Learning

- Each device trains on local behavioral data.
- Periodically sends encrypted gradients to the server.

Differential Privacy

- Gaussian noise added to gradients.
- Ensures biometric uniqueness is not reconstructable.

7. Dataset Description

7.1 Real Behavioral Data Used

- HMOG Dataset: touch, motion, and gesture data.
- Keystroke Benchmark Dataset: dwell and flight time sequences.

7.2 Synthetic Payment-Session Dataset

A custom dataset was generated simulating:

Attribute	Description
Session ID	Unique identifier
Keystroke Features	Real user sequences mapped to payment flow timings
Touch/Motion	Swipes during PIN entry, navigation
Device Metadata	OS, screen type, sensor drift
Transaction Context	Amount, merchant category, UPI-type flow
Labels	Genuine / Fraud / Social-engineered / Bot

Fraud Scenarios Simulated

- Account takeover
- Scripted bot payments
- Remote-assistance scam
- Cross-device transfer attempt

Dataset size: 18,500 sessions (12,300 genuine, 6,200 fraudulent).

The dataset was split into 70% training and 30% testing sessions. The model is evaluated using standard metrics including Accuracy, F1-score, AUC-ROC, and inference latency. Accuracy and F1-score were calculated using true positive, false positive, true negative, and false negative values in accordance with standard definitions.

8. Experimental Setup

8.1 Baselines

- Transaction-only XGBoost
- Keystroke-only LSTM
- Touch-only CNN
- Simple concatenation fusion

8.2 Metrics

- Accuracy
- F1-score
- AUC-ROC
- False Positive Rate (FPR)
- Latency (ms)

9. Results and Discussion

9.1 Performance Comparison

Table: Accuracy and F1-score

Model	Accuracy	F1-score
Transaction-only XGBoost	84.2%	0.78
Keystroke-only LSTM	89.1%	0.82
Touch-only CNN	87.4%	0.81
Simple Fusion	92.5%	0.88
Proposed Multimodal + Attention	97.2%	0.95

9.2 Latency

- Average inference: 178 ms
- Meets FinTech requirement (< 200 ms)

9.3 Robustness

- Cross-device EER: 3.1%
- Sparse sessions (few events): Accuracy 92.3%
- Adversarial mimicry: Detection rate 94.2%

9.4 Privacy Impact

- DP with $\varepsilon=1.5$ shows only 2.4% reduction in accuracy
- Ensures safe privacy-utility trade-off

10. Conclusion

This research presents a novel, privacy-preserving deep multimodal behavioral intelligence framework tailored specifically for continuous authentication in real-time FinTech payment systems. Unlike existing authentication or fraud detection methods, our system integrates multimodal behavioral biometrics, device intelligence and transaction context into a unified deep learning architecture with federated and differential privacy safeguards. Experimental results validate its superior accuracy, robustness, and real-time feasibility. The findings reveal strong potential for deployment in mobile banking, digital wallets and UPI/payment gateways where user identity verification is mission-critical.

11. Future Scope

- Expansion to voice biometrics and face dynamics for richer multimodal signals.
- Deployment in low-power edge devices with model compression.
- Real-world pilot deployment with FinTech partners.
- Adversarial training for resistance to sophisticated mimicry attacks.

12. References

1. X. Zhang et al., "Research on Multimodal Continuous User Authentication on Mobile Devices", 2021.
2. L. Lin et al., "Research on AMBR: An Attention-Based Multimodal Biometric Recognition Network", *Sensors*, 2023.
3. K.K. Coelho et al., "Research on A Multimodal Biometric Authentication Method by Federated Learning", 2023.
4. E. Sağbaş et al., "Research on Machine Learning-Based on Continuous Authentication Using Soft Keyboard and Sensors", 2024.
5. H. Sharma, "UPI Fraud Detection System", *IJISRT*, 2025.
6. J. Frank, "Research on Keystroke Dynamics: A Tool for User Identification", *Computer Security*, 2020.
7. M. Ho et al., "Research on Behavioral Biometrics for Mobile Authentication", *IEEE Access*, 2022.
8. B. Fang et al., "Research on Federated Learning for Privacy-Preserving Biometrics", *IEEE TIFS*, 2023.