ISBN: 978-81-988536-2-2

Chapter 8: Fraud Detection and Cyber Risk Prevention in Digital Commerce: A Machine Learning-Centric Framework

Mrs. PMG. Jegadhambal

[1]Assistant Professor, Department of Computer Science Engineering, Vels Institute of Science Technology and Advanced Studies

Digital transactions have become ubiquitous, with e-commerce, fintech, and digital banking platforms processing billions of payments daily. However, this surge in digital activity has also led to an alarming rise in **cyber fraud incidents**, ranging from synthetic identity fraud to bot-driven checkout attacks. The **2023 Cybersecurity Report** by Verizon indicated that over 74% of breaches involved human and behavioral manipulation, while over 35% were financially motivated, specifically targeting payment systems (Nair et al., 2024)**.** In this high-stakes ecosystem, **fraud detection systems (FDS)** are evolving from static rule engines to dynamic, AI-driven platforms capable of real-time evaluation of transactional trustworthiness.

8.1. Anomaly Detection

Anomaly detection refers to the identification of patterns in data that do not conform to expected behavior. In the context of financial fraud, anomalies often manifest as:

- Unusual transaction amounts
- Abnormal frequency of transactions
- Geolocation discrepancies
- Device switching mid-session (Elia et al.,2024)

8.2. Types of Machine Learning Models in Anomaly Detection

A. Supervised Models

These models are trained using labeled datasets where instances of fraud and legitimate transactions are known. Algorithms include:

- **Logistic Regression**
- **Random Forests**
- **Support Vector Machines (SVM)**
- **XGBoost (**Agyemang 2024)

Limitations: Requires large labeled datasets, suffers from class imbalance.

B. Unsupervised Models

Trained on normal data to learn the distribution; anomalies are identified as outliers.

- **Isolation Forest**
- **One-Class SVM**
- **Autoencoders**
- **K-Means Clustering**

C. Semi-Supervised and Hybrid Models

Combines both labeled and unlabeled data. Particularly effective in minimizing false positives.

8.3 Deep Learning for Complex Fraud Patterns

- **Recurrent Neural Networks (RNNs)**: Detect temporal patterns in sequential transactions.
- **Variational Autoencoders (VAEs)**: Capture subtle deviations in high-dimensional transaction vectors.
- **Graph Neural Networks (GNNs)**: Model user-merchant relationships to detect collusive fraud (Gandhar et al., 2024)

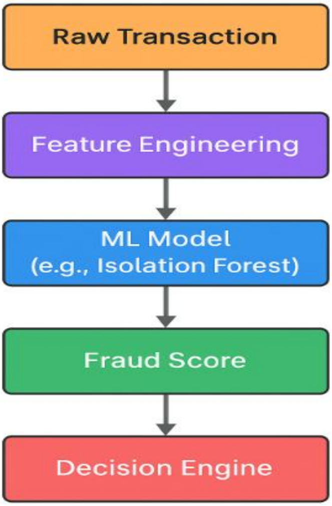

Figure 8.1: ML Workflow for Transactional Anomaly Detection

8.4. Device Fingerprinting in Fraud Detection

i) Concept of Device Fingerprinting

Device fingerprinting is a sophisticated tracking technique used to uniquely identify and authenticate a device based on its specific configuration and environment. Unlike traditional cookies, device fingerprints do not rely on local storage but instead collect a wide array of passive data points from a user's device and browser during interaction with a website or application (Selvasundaram et al.,2025).

This technique is widely applied in fraud detection, cyber risk prevention, user verification, and behavioral analytics due to its robustness and resistance to deletion or spoofing.

Device fingerprinting creates a unique signature for each user device using parameters such as:

Table 8.1: Parameters in Device fingerprinting

Parameter	**Description**
OS Version	Captures the exact operating system and version (e.g., Windows 11, iOS 17)
Screen Resolution	Captures width × height (e.g., 1920x1080), color depth, and scaling factor
Browser Plugins	Lists installed browser extensions/plugins and their versions
Installed Fonts	Enumerates local fonts available to the browser, differing across systems
Time Zone	Detects local time zone and any deviations from UTC

Unlike cookies, fingerprints persist across sessions and cannot be easily deleted.

ii) Techniques in Device Fingerprinting

- **Canvas Fingerprinting**: Uses rendering of HTML5 canvas elements.

- **WebGL & Audio Fingerprints**: Exploit subtle differences in rendering or processing.
- **Hardware-Specific Fingerprints**: Gyroscope behavior or battery usage patterns (Durey, 2021)

Use Case in Fraud Detection

Device fingerprints can help identify:

- Account sharing
- Device spoofing
- Bot-driven attacks
- Suspicious access patterns (e.g., high-risk geolocation + unknown device)

Table 8.2: Comparison of Fingerprinting Methods

Method	**Accuracy**	**Evasiveness Resistance**	**Common Use**
Canvas Fingerprinting	High	Medium	Web Apps
Audio Fingerprinting	Medium	High	Mobile Apps
IP + User-Agent	Low	Low	Legacy FDS

8.5. Security and Anti-Fraud Use Cases

1. **Multi-factor Risk Scoring**

   Device fingerprinting is often integrated into **fraud scoring systems**. For example, if a known user logs in from a device with a different fingerprint, the system may:

   - Trigger an alert
   - Initiate additional authentication
   - Block the session

2. **Bot and Emulator Detection**

   Bots often operate in controlled, non-diverse environments. Fingerprints that appear *too uniform* or lack common entropy features (e.g., missing plugins or standardized fonts) can be flagged.

3. **Account Takeover Prevention (ATO)**

   By maintaining a historical device fingerprint per account, any deviation from known devices can be interpreted as a possible account takeover attempt.