International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Seeing the Unseen: A Hybrid Supervised-Unsupervised Learning Framework for Zero-Day Cyber Threat Detection in Digital Transactions

Noya Shaikh[1], Priya Budlani, Sonali Ingle, Vrishali Wable

[1]Ashoka Center for Business and Computer Studies, Nashik
Email: *noyashaikh786[at]gmail.com*

[2]Assistant Professor, Ashoka Center for Business and Computer Studies, Nashik
Email: *budlanipriya5[at]gmail.com*

[3]Assistant Professor, Ashoka Center for Business and Computer Studies, Nashik
Email: *sonali.ingle7512[at]gmail.com*

[4]Assistant Professor, Ashoka Center for Business and Computer Studies, Nashik
Email: *bykvrishali[at]gmail.com*

Abstract: ***The rapid growth of digital transactions has significantly increased exposure to cyber threats, particularly zero-day attacks that do not follow previously known patterns. Conventional security systems largely depend on signature-based or fully supervised learning techniques, making them ineffective against evolving and unseen attack behaviors. This paper proposes a hybrid machine learning framework that integrates unsupervised anomaly detection with supervised classification to enhance the detection of both known and unknown cyber threats in digital transaction environments. The unsupervised component identifies abnormal transaction patterns without relying on labeled data, while the supervised component classifies known malicious activities using historical attack information. By combining both learning paradigms, the proposed framework improves adaptability, reduces dependency on labeled datasets, and enhances early threat detection capabilities. The performance of the approach is conceptually evaluated using standard security metrics such as accuracy, precision, recall, and false alarm rate. This work presents a flexible and scalable direction for strengthening digital transaction security and offers a strong foundation for future implementation and real-world validation.***

Keywords: Zero-day attacks, Digital transactions, Anomaly detection, Hybrid learning, Machine learning security

1. Introduction

Digital transaction systems, including online payments, mobile banking, and e-commerce platforms, have become an essential part of modern financial ecosystems. The increasing reliance on these systems has led to a significant rise in transaction volumes and real-time financial interactions. While digital transactions offer speed, convenience, and accessibility, they also introduce serious security challenges. Cyber attackers continuously exploit vulnerabilities in transaction systems, leading to fraudulent activities, financial losses, and erosion of user trust.

Traditional transaction security mechanisms largely rely on rule-based systems or signature-based detection techniques that are designed to recognize previously known attack patterns. Although these methods are effective against familiar threats, they struggle to detect new and evolving attack behaviors. Modern cyber threats often emerge as zero-day attacks, where attackers intentionally modify their strategies to bypass existing security measures. As a result, such threats frequently remain undetected until substantial damage has already occurred.

Machine learning techniques have been widely adopted to enhance transaction security. Supervised learning models, such as decision trees, random forests, and neural networks, have demonstrated strong performance in detecting known fraudulent activities. However, these models depend heavily on labeled historical data, which is often limited, imbalanced, and costly to obtain in real-world financial systems. Since the majority of transactions are legitimate, supervised models may fail to generalize effectively to unseen attack patterns.

Unsupervised learning approaches, particularly anomaly detection techniques, offer an alternative solution by identifying unusual or abnormal transaction behavior without relying on labeled data. These methods are well suited for detecting unknown or emerging threats. However, purely unsupervised models may generate high false alarm rates, as not all anomalies represent malicious activities. This limitation reduces their reliability when deployed independently in practical transaction environments.

To overcome the limitations of individual learning approaches, this paper proposes a hybrid supervised–unsupervised learning framework for zero-day cyber threat detection in digital transactions. The proposed framework combines unsupervised anomaly detection to identify suspicious transaction behavior with supervised classification to recognize known malicious patterns. By integrating both techniques, the framework aims to improve detection accuracy, enhance adaptability to evolving attack strategies, and reduce dependence on labeled datasets.

The proposed approach provides a flexible and scalable direction for strengthening transaction security in dynamic cyber environments. It is particularly suitable for real-world

digital transaction systems where new attack patterns continuously emerge and labeled data is limited. This work lays the foundation for future implementation and evaluation of hybrid machine learning models in cyber security applications.

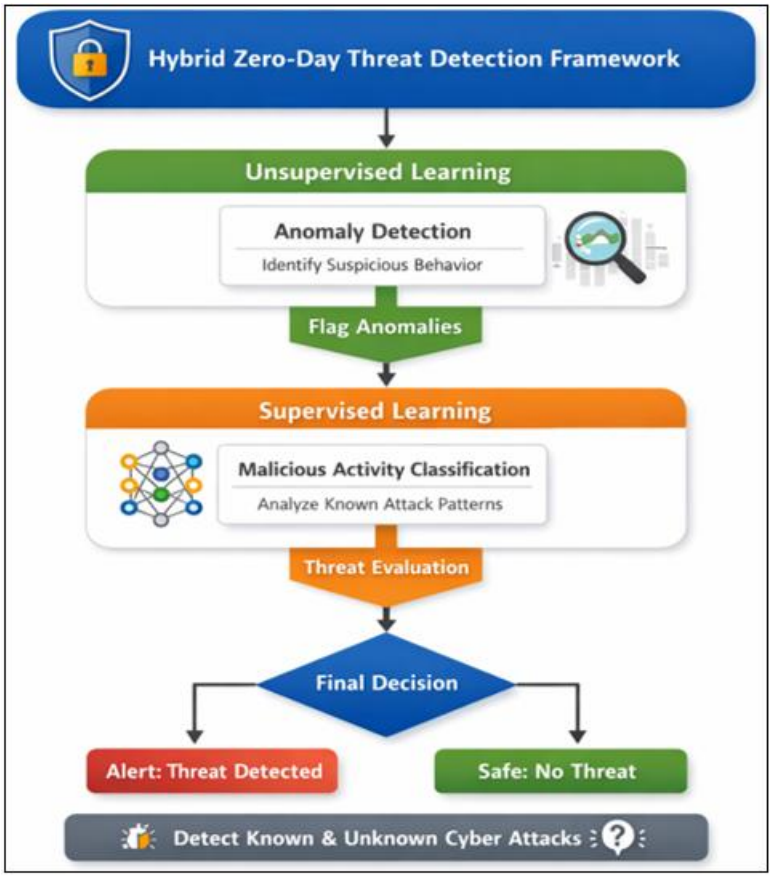
2. Methodology

This study proposes a conceptual hybrid supervised–unsupervised learning framework for zero-day cyber threat detection in digital transactions. The framework consists of two primary stages designed to work collaboratively.

The first stage employs unsupervised learning techniques to analyze transaction behavior and identify anomalies without relying on labeled data. This stage focuses on detecting unusual patterns that may indicate previously unseen or emerging threats.

The second stage applies supervised learning models trained on historical transaction data to classify known malicious activities. Transactions flagged as anomalous in the first stage are further evaluated to determine whether they represent genuine security threats.

The final decision is made by combining the outputs of both stages, enabling accurate detection of both known and unknown cyber-attacks.



3. Data Analysis

As this study focuses on a conceptual framework, direct implementation and dataset experimentation are not included. However, the proposed framework is designed to analyze transaction features such as transaction amount, frequency, location, device information, and behavioral patterns.

Performance evaluation is expected to be conducted using standard security metrics, including accuracy, precision, recall, and false alarm rate. These metrics provide insight into detection efficiency, reliability, and robustness.

4. Findings

The proposed hybrid framework is expected to provide improved detection of zero-day and unknown cyber threats compared to traditional single-model approaches. By integrating anomaly detection with supervised classification, the framework reduces dependency on labeled data and enhances adaptability to evolving attack behaviors.

Additionally, the combined approach is expected to lower false alarm rates while maintaining high detection accuracy, making it suitable for real-world digital transaction systems.

5. Conclusion

This paper presents a hybrid supervised–unsupervised learning framework for detecting zero-day cyber threats in digital transaction systems. By combining anomaly-based detection with supervised classification, the proposed approach addresses key limitations of traditional security models that rely solely on known attack patterns.

The framework offers improved adaptability, reduced dependence on labeled datasets, and enhanced resilience against evolving cyber threats. This work provides a strong foundation for future research, implementation, and experimental evaluation in the field of digital transaction security.

References

[1] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
[2] Fawcett, T., & Provost, F. (1997). Adaptive fraud detection. *Data Mining and Knowledge Discovery*, 1(3), 291–316.
[3] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), 1–58.
[4] Axelsson, S. (2000). The base-rate fallacy and the difficulty of intrusion detection. *ACM Transactions on Information and System Security*, 3(3), 186–205.
[5] Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud detection: A comparative study. *Decision Support Systems*, 50(3), 602–613.
[6] Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.
[7] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
[8] Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
[9] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
[10] Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *IEEE Symposium on Security and Privacy*, 305–316.
[11] Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques. *Computer Networks*, 51(12), 3448–3470.
[12] Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. *Military Communications and Information Systems Conference*.
[13] Lee, W., & Stolfo, S. J. (2000). A framework for constructing features and models for intrusion detection systems. *ACM Transactions on Information and System Security*, 3(4), 227–261.
[14] He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284.
[15] Javaid, A., Niyaz, Q., Sun, W., & Alam, M. (2016). A deep learning approach for network intrusion detection system. *IEEE International Conference on Bioinformatics and Biomedicine*, 21–26.
[16] Ring, M., Wunderlich, S., Scheuring, D., Landes, D., & Hotho, A. (2019). A survey of network-based intrusion detection datasets. *Computers & Security*, 86, 147–167.
[17] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.