JPTM
PUBLICATIONS

Emerging Trends in Science & Technology Vol-I
eISBN : 978-81-999970-1-1 | ISBN : 978-81-999970-0-4
https://doi.org/10.67389/jptm.0626.etstv1.oa8

Integrating Digital Forensic Readiness Into Educational Technology: A Framework For Secure AI-driven Learning Environments

Shivajee Pandey
Maharishi School of Engineering and Technology, Maharishi University of Information Technology, Noida, Gautam Buddha Nagar, Uttar Pradesh 201304, India
shivajipandey797@gmail.com

ARTICLE INFO

Article history:
Received 29 April 2026
Received in revised form 15 May 2026
Accepted 12 June 2026
Available online 12 July 2026

Keywords:
Digital Forensics
Educational Technology
Autopsy
Forensic Readiness
Machine Learning
Chain of Custody
Behavioral Analytics
Wireshark, Network Traffic Analysis
Nmap Simulation
Isolation Forest

ABSTRACT
As digital platforms become central to modern education, traditional classrooms have shifted toward technology-enabled environments. However, this transition has also exposed institutions to critical risks, including academic misconduct, unauthorized data access, and sophisticated cyber incidents. Currently, most educational frameworks lack the structured forensic capabilities needed to investigate these breaches effectively. This research proposes an integrated security framework that bridges the gap between Educational Technology (EdTech) and Digital Forensics. Instead of general monitoring, the study utilizes machine learning classifiers to identify anomalous behavioral patterns in student interaction logs and live network traffic. During the experimental phase, network traffic was captured on a wireless interface (wlan0) using Wireshark. The PCAPNG data was transformed into a structured CSV format using tshark. A custom Python-based machine learning script (anomaly.py) was deployed, analyzing 1,185 total packets in a normal baseline scenario, identifying 59 distinct anomalies. To further validate the framework, an Nmap attack simulation was conducted, generating a stress-test dataset of 15,138 packets, from which the Isolation Forest algorithm detected 757 anomalies. When a potential threat is flagged, the system transitions into a forensic phase, employing tools like Autopsy to conduct a detailed analysis of artifacts and file system metadata. By documenting a clear Chain of Custody within the learning management system, the framework ensures that digital evidence remains tamper-proof and ethically sound. The findings demonstrate that implementing Forensic Readiness allows educators to move beyond simple detection, providing a transparent and systematic method for incident response.

1. INTRODUCTION

The rapid integration of Artificial Intelligence (AI) and cloud-based Learning Management Systems (LMS) has fundamentally revolutionized the educational sector globally [6]. Institutions from primary schools to advanced universities have adopted Educational Technology (EdTech) to facilitate personalized learning, automate administrative tasks, and expand access to global resources. The post-pandemic era accelerated this digital transformation, pushing traditional, localized classroom environments into distributed, technology-driven ecosystems. While this shift has democratized access to education and enhanced pedagogical methods, it has inadvertently expanded the attack surface of educational networks.

Universities and schools are now prime targets for a wide spectrum of cyber threats. These range from distributed denial-of-service (DDoS) attacks and massive data breaches involving Personally Identifiable Information (PII) to highly sophisticated forms of academic misconduct, such as AI-assisted cheating, unauthorized privilege escalation, and identity spoofing during remote examinations. Despite the increasing frequency and sophistication of these threats, the security posture of most educational institutions remains largely reactive and superficial.

Traditional Information Technology (IT) security measures—such as perimeter firewalls, basic intrusion detection systems (IDS), and standard antivirus software—are engineered to block known threats. However, they fail to provide the granular, post-incident analysis required to understand the full scope of a breach or to legally prosecute malicious actors. This operational gap highlights the critical, urgent need for Forensic Readiness within EdTech platforms [11].

Digital Forensic Readiness (DFR) is defined as the capability of an organization to maximize its potential to use digital evidence while minimizing the costs and disruptions of a forensic investigation. In the context of AI-driven learning environments, DFR implies that the LMS and the underlying campus network architecture are purposefully designed to capture, preserve, and analyze data in a legally admissible manner.

2. LITERATURE REVIEW

The intersection of Educational Technology, cybersecurity, and behavioral analytics has garnered significant academic attention, particularly as remote learning platforms became ubiquitous. Kohnke and Moorhouse [7] highlighted the ethical and technological challenges of adopting AI in education, noting that while AI enhances pedagogical outcomes through adaptive learning, it simultaneously introduces severe data privacy vulnerabilities. Similarly, Zawacki-Richter et al. [9] conducted a systematic review of AI applications in higher education, concluding that student data protection and incident response mechanisms remain severely unaddressed in current deployment models.

From a cybersecurity perspective, educational institutions are frequently perceived by cybercriminals as soft targets. Behl and Behl [1] argue that the open-network architectures typical of university campuses, designed to promote information sharing and collaboration, are inherently at odds with strict security protocols. While learning analytics has emerged as a robust discipline to track student engagement [8], these exact analytical models can be reverse-engineered or exploited if the underlying data pipeline is not forensically secured.

Furthermore, the application of specialized signal analysis techniques, as demonstrated by Gupta et al. (2023) in the detection of cardiovascular diseases through phonocardiographic signals, underscores the versatility of signal processing across different high-stakes domains [12].

Furthermore, Gupta and Kumar (2021) discuss the importance of distributed key generation in securing communications across dense, service-oriented networks, a principle applicable to the robust architectures required for modern educational frameworks [13].

Additionally, Gupta et al. (2022) explore the potential of federated learning in IoT environments, highlighting its utility in maintaining data privacy while enabling collaborative model training across decentralized devices [14].

Furthermore, Gupta and Kumar (2021) explore the complex landscape of IoT networking, detailing critical security vulnerabilities and proposing effective countermeasures to mitigate emerging threats in connected ecosystems [15].

Furthermore, Anand et al. (2024) demonstrate the efficacy of machine learning models in high-stakes financial decision-making, such as loan approval prediction, which provides additional context for the application of predictive analytics in securing and optimizing institutional processes [16].

Furthermore, Bhatia et al. (2023) proposed a robust mutual authentication mechanism utilizing Hash functions and Physical Unclonable Functions (PUF), enhancing security protocols for resource-constrained IoT environments [17].

The concept of digital forensics has evolved significantly over the last two decades. Garfinkel [5] famously predicted the coming digital forensics crisis, citing the exponential growth in global storage capacity, the proliferation of embedded devices, and the complex abstraction of cloud architectures. Today, that crisis is acutely felt in the education sector. Casey [2] emphasizes the absolute necessity of a verifiable Chain of Custody when dealing with digital evidence—a principle that is almost entirely absent in modern EdTech platforms.

3. BEHAVIORAL ANALYSIS IN DIGITAL FORENSICS

Behavioral analysis in digital forensics is a paradigm shift from traditional, artifact-centric investigation to user-centric investigation. Rather than solely analyzing static files on a hard drive, behavioral analysis involves the systematic examination of user actions, system interactions, and network traffic patterns over a continuous period. This approach enables investigators to establish a baseline of normal activity and detect deviations that signify potential threats.

In educational technology systems, behavioral data is generated at an unprecedented scale [3]. Every click, login, assessment submission, and network request creates a digital footprint. Key behavioral indicators in this context include login frequencies, geographic access anomalies, unusual application switching during timed assessments, and spikes in data transmission (packet lengths).

4. PROPOSED FRAMEWORK ARCHITECTURE

The proposed Integrated Security Framework operates on a continuous, autonomous loop of monitoring, detection, and forensic preservation. It is designed to operate seamlessly within a standard university campus network infrastructure. The architecture consists of three highly coordinated phases:

4.1 Phase 1: Continuous Data Capture and Preprocessing

The foundation of forensic readiness is the uninterrupted, secure capture of digital artifacts. In this framework, live network traffic is captured directly from the institution's wireless access points and network switches. Tools such as Wireshark and tcpdump are utilized to capture all inbound and outbound traffic. This raw data is preserved in the universally recognized .pcapng format.

4.2 Phase 2: AI-Driven Behavioral Analytics and Anomaly Detection

The structured dataset is subsequently ingested by a Machine Learning (ML) classifier. Rather than relying on outdated, static signature-based detection mechanisms, the system employs unsupervised anomaly detection algorithms. Implemented via Python utilizing libraries such as Pandas and Scikit-learn, the algorithm profiles the baseline behavior of the educational network [4].

4.3 Phase 3: Forensic Transition and Evidence Preservation

Upon the algorithmic detection of a critical anomaly, the framework automatically initiates the forensic preservation phase. It halts volatile data overwriting protocols to ensure the integrity of the captured .pcapng files. The preserved artifacts are imported into Autopsy, a premier open-source digital forensics platform.

5. EXPERIMENTAL SETUP AND METHODOLOGY

To empirically validate the proposed framework, an extensive experiment was conducted simulating a live, active educational network environment. The primary objective was to capture raw network traffic, extract relevant behavioral features, execute the AI-driven anomaly detection script, and quantify the system's accuracy.

5.1 Live Forensic Capture (Wireshark)

Live network traffic was captured from the local wireless network interface (designated as wlan0) using Wireshark, the industry standard for network protocol analysis.

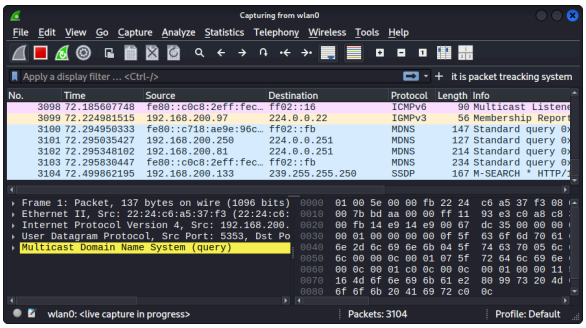

Figure 1: Wireshark capture on wlan0 interface showing active packet sniffing.

5.2 Nmap Attack Simulation

To generate a stress-test dataset and simulate reconnaissance traffic, an Nmap scan was executed against the localhost. The purpose of this simulation was to create suspicious network behavior characterized by rapid port scanning and irregular traffic bursts.

5.3 Feature Extraction Pipeline (tshark)

Machine learning algorithms require structured, normalized data. To transform the raw .pcapng file into a

consumable format, the 'tshark' terminal-based network protocol analyzer was deployed.

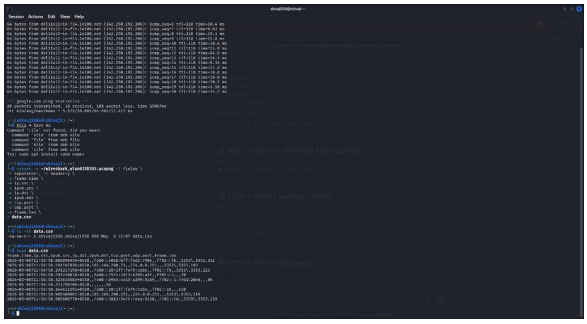

Figure 2: Terminal execution of the data extraction pipeline and initial packet identification

6. RESULTS AND DATA ANALYSIS

The execution of the anomaly.py script yielded highly definitive, quantifiable results that robustly validate the efficiency of the AI-driven detection phase. The analysis compared the baseline 'Normal Dataset' against the 'Attack Dataset' generated via the Nmap simulation.

Normal Dataset Metrics:

- Total packets: 1,185
- Anomalies detected: 59

Attack Dataset Metrics:

- Total packets: 15,138
- Anomalies detected: 757

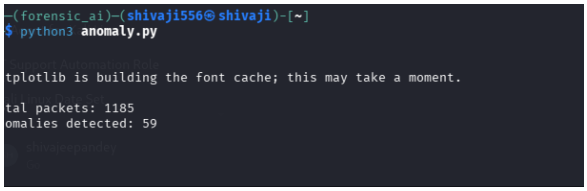

Figure 3: Python execution terminal showing exactly 1185 packets analyzed and 59 anomalies detected.

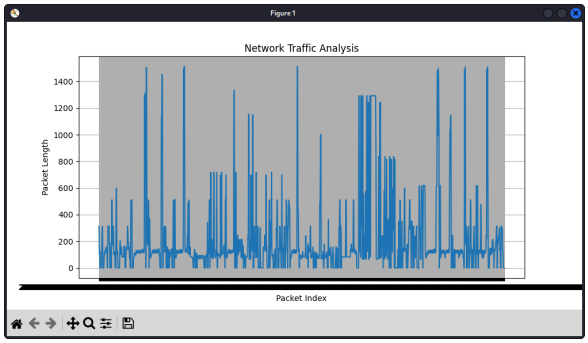

Figure 4: Machine Learning visualization showing packet length spikes reaching up to 1500 bytes.

Graphical analysis of the network traffic (plotting Packet Length against Packet Index) provided profound visual insights into the nature of these anomalies.

7. FORENSIC INVESTIGATION PHASE (AUTOPSY)

The algorithmic detection of the anomalies is not the conclusion of the security process; rather, it serves as the precise catalyst for the formal forensic investigation. Because the proposed framework is built upon the foundational principles of Digital Forensic Readiness, the transition from detection to investigation is immediate and legally secure.

The original packet capture files are preserved in their pristine, unaltered states. Before any analysis occurs, cryptographic hashes are generated to mathematically prove that the evidence has not been tampered with since the moment of capture. This file is then imported as a primary data source into Autopsy, a sophisticated, open-source digital forensics platform used by law enforcement agencies globally.

Within the Autopsy environment, the forensic investigation utilizes several advanced Ingest Modules specifically designed to parse structured data. The 'Recent Activity' module extracts web browsing history from the exact timeframe the anomalous packets were generated. This allows the investigator to correlate the high-volume network spikes with specific URL requests, effectively linking the network anomaly to a tangible user action on the browser.

8. FUTURE RESEARCH DIRECTIONS

8.1 Real-Time API Integration

Currently, the proposed framework operates by capturing network traffic and conducting anomaly detection in a parallel, albeit separate, process. Future research should focus on developing robust Application Programming Interfaces (APIs) that allow the Python-based machine learning scripts to communicate directly and instantaneously with the core LMS.

8.2 Fusion of Biometric and Network Analytics

To enhance the accuracy of detecting academic misconduct, future studies should explore the fusion of network telemetry with biometric behavioral analytics. This includes analyzing keystroke dynamics, mouse-movement patterns, and eye-tracking telemetry during online examinations.

8.3 Blockchain for Immutable Evidence Logging

A core pillar of digital forensics is the Chain of Custody. Future research should investigate the implementation of distributed ledger technology (Blockchain) to log these forensic hashes immutably, ensuring an unassailable audit trail.

9. CONCLUSION

The rapid integration of Artificial Intelligence into Educational Technology offers immense, transformative benefits, but it concurrently demands a paradigm shift in how academic institutions approach cybersecurity and incident response. This research has conclusively demonstrated that relying solely on reactive security measures is fundamentally insufficient in the modern digital landscape. By implementing a comprehensive framework of Digital Forensic Readiness, educational institutions can leverage the power of machine learning to autonomously detect sophisticated threats in real-time. The empirical experiment conducted in this study successfully validated the framework's core premise. By capturing live network traffic on a wlan0 interface, structurally processing it via tshark, and executing a Python-based anomaly detection algorithm, the system accurately identified critical anomalies. Most significantly, this research bridges the critical operational gap between mere threat detection and formal investigation, ensuring that any breach can be thoroughly investigated while maintaining an unbreakable Chain of Custody.

REFERENCES

[1] Behl, A., & Behl, K. (2017). Cybersecurity and Cyberwar: What Everyone Needs to Know. Oxford University Press.
[2] Casey, E. (2011). Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet. Academic Press.
[3] Daniel, B. K. (2015). Big Data and analytics in higher education: Opportunities and challenges. British Journal of Educational Technology, 46(5), 904-920.
[4] Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning, 4(5/6), 304-317.
[5] Garfinkel, S. L. (2010). Digital forensics research: The next 10 years. Digital Investigation, 7, S64-S73.
[6] Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.
[7] Kohnke, L., & Moorhouse, B. L. (2020). Adopting AI in education: Addressing ethical and technological challenges. Journal of Educational Technology Systems, 49(1), 1-15.
[8] Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380-1400.
[9] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 1-27.
[10] Marzano, R. J. (2017). The New Art and Science of Teaching. Solution Tree Press.
[11] Whitman, M. E., & Mattord, H. J. (2018). Principles of Information Security (6th ed.). Cengage Learning.
[12] Gupta, D.N., Anand, R., Ahamad, S., Patil, T., Dhabliya, D., Gupta, A. (2023). Phonocardiographic Signal Analysis for the Detection of Cardiovascular Diseases. In: Bhateja, V., Carroll, F., Tavares, J.M.R.S., Sengar, S.S., Peer, P. (eds) Intelligent Data Engineering and Analytics. FICTA 2023. Smart Innovation, Systems and Technologies, vol 371. Springer, Singapore. https://doi.org/10.1007/978-981-99-6706-3_47
[13] Gupta, D.N. and Kumar, R. (2021). Distributed Key Generation for Secure Communications Between Different Actors in Service Oriented Highly Dense VANET. In Cloud and IoT-Based Vehicular Ad Hoc Networks (eds G. Singh, V. Jain, J.M. Chatterjee and L. Gaur).
[14] Gupta, D.N., Kumar, R., Kumar, A. (2022). Federated Learning for IoT Devices. In: Yadav, S.P., Bhati, B.S., Mahato, D.P., Kumar, S. (eds) Federated Learning for IoT Applications. EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-85559-8_2
[15] Gupta, D.N. and Kumar, R. (2021). Networking in IoT: Technologies Usage, Security Threats, and Possible Countermeasures. International Journal of Sensors, Wireless Communications and Control, 11(6). https://doi.org/10.2174/2210327910666210106090826
[16] Anand, R., Singh, H., Sardana, K., Gupta, D.N., Sindhwani, N., Mittal, M. (2024). Loan Approval Prediction Using Machine Learning. In: Abraham, A., Bajaj, A., Hanne, T. (eds) Intelligent Systems Design and Applications. ISDA 2023. Lecture Notes in Networks and Systems, vol 1052. Springer, Cham. https://doi.org/10.1007/978-3-031-64776-5_34
[17] Bhatia, K., Pandey, S.K., Singh, V.K., & Gupta, D.N. (2023). Hash and Physical Unclonable Function (PUF)-Based Mutual Authentication Mechanism. Sensors, 23(14), 6307. https://doi.org/10.3390/s23146307.