Deep Learning-Based Anomaly Detection Models for Advanced Persistent Threats

OMOROGIE MICHAEL and NWABUDIKE UJU CYNTHIA
omorogiemichael2015@gmail.com
Department of Computer Science, Delta State Polytechnic Ogwashi-Uku, Delta State, Nigeria.

Abstract

Advanced Persistent Threats (APTs) represent sophisticated, long-term cyber-attacks that evade traditional security measures, posing significant risks to organizations worldwide. This paper explores deep learning-based anomaly detection models as a robust approach to identifying and mitigating APTs. We review key concepts in APTs and anomaly detection, delve into fundamental deep learning architectures such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, Auto encoders, Generative Adversarial Networks (GANs), and Transformers, and analyse their applications in detecting anomalous behaviours indicative of APTs. Drawing from recent literature, we examine hybrid models that integrate these techniques to enhance accuracy, reduce false positives, and handle complex network data. Case studies from benchmark datasets like UNSW-NB15 and NSL-KDD demonstrate high performance, with accuracies exceeding 98% in some implementations. Challenges such as data scarcity, model interpretability, and computational demands are discussed, alongside future directions including explainable AI and integration with threat intelligence. This work underscores the transformative potential of deep learning in cybersecurity, providing a comprehensive framework for researchers and practitioners to develop resilient detection systems.

Keywords: Deep Learning-Based, Anomaly Detection Models, Advanced Persistent Threats

Omorogie, M. & Nwabudike, U. C. (2026). Deep Learning-Based Anomaly Detection Models for Advanced Persistent Threats. Academy Journal of Information and Communication Technology (AJICT), 1(1), 1-5. https://ajict.mydspg.edu.ng

INTRODUCTION

In the evolving landscape of cybersecurity, Advanced Persistent Threats (APTs) have emerged as one of the most formidable challenges (Al-Rimy et al., 2020). APTs are characterized by their stealthy, targeted, and prolonged nature, often orchestrated by nation-state actors or well-resourced criminal groups to exfiltrate sensitive data or disrupt critical infrastructure. Unlike conventional attacks, APTs involve multiple stages, including reconnaissance, initial compromise, command and control, lateral movement, and data exfiltration, allowing adversaries to remain undetected for months or even years. Traditional signature-based detection systems, which rely on known patterns, are ineffective against APTs due to their use of zero-day exploits and adaptive tactics. Anomaly detection offers a promising alternative by identifying deviations from normal behaviour rather than matching predefined signatures (Sarker, 2021; Narayanan & Gupta, 2024). This approach is particularly suited to APTs, as it can flag subtle, anomalous activities such as unusual network traffic patterns or irregular system logs. Deep learning, a subset of machine learning, has revolutionized anomaly detection by enabling models to automatically learn complex features from vast datasets without extensive manual engineering. Techniques like neural networks can process high-dimensional data, capture temporal dependencies, and improve detection accuracy in dynamic environments. This paper provides a detailed examination of deep learning-based anomaly detection models for APTs. We begin with background on APTs and anomaly detection, followed by an overview of relevant deep learning fundamentals. A literature review synthesizes recent advancements, highlighting specific models and their performance. We then discuss experimental methodologies, results, challenges, and future directions. The goal is to offer a comprehensive resource that bridges theory and practice, emphasizing the need for innovative, sustainable cybersecurity solutions.

Background on Advanced Persistent Threats

APTs differ from other cyber threats in their persistence and sophistication. According to the NIST definition, APTs involve adversaries with significant resources who adapt to defenses and maintain access over extended periods (Walsh et al., 2024). Notable examples include the Solar Winds supply chain attack in 2020, (Walsh et al., 2024), which compromised thousands of organizations undetected for over a year, and Stuxnet, which targeted industrial control systems in 2009. These attacks exploit vulnerabilities across the cyber chain, from initial entry via phishing or malware to establishing command-and-control channels. The stealth of APTs stems from their "low-and-slow" approach, blending malicious activities with legitimate traffic to avoid triggering alerts (Eke et al., 2023). This makes rule-based or signature-based intrusion detection systems (IDS) inadequate, as they cannot detect novel or obfuscated threats (Ferrag et al., 2020). Anomaly detection, conversely, establishes baselines of normal behaviour such as typical network flow volumes, user access patterns, or log event sequences and flags outliers. There are two main types: statistical methods, which use probabilistic models like Gaussian distributions, and machine learning-based methods, which leverage algorithms to learn from data. In the context of APTs, anomaly detection must handle high-dimensional, time-series data from sources like network packets, system logs, and endpoint behaviour’s. Challenges include distinguishing benign anomalies (e.g., system updates) from malicious ones, managing noise in large-scale environments, and achieving low false positive rates to prevent alert fatigue. Deep learning addresses these by modeling non-linear relationships and temporal contexts, enabling more accurate and adaptive detection (Agarwal & Gupta, 2024; Kim et al., 2024).

Anomaly Detection Basics

Anomaly detection is a critical component of modern IDS, categorized into supervised, unsupervised, and semi-supervised paradigms (Sarker, 2021). Supervised methods require labelled data for training, which is scarce for APTs due to their rarity and variability. Unsupervised methods, such as clustering or isolation forests, identify anomalies without labels by assuming they are rare and distinct. Semi-supervised approaches train on normal data only, flagging deviations as anomalies. Key metrics for evaluating anomaly detection include accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). For APTs, recall is particularly important to minimize missed detections, while precision reduces false alarms. Datasets like NSL-KDD, UNSW-NB15, and CIC-IDS2017 are commonly used for benchmarking, simulating network intrusions including APT-like behaviours (Ferrag et al., 2020; Myneni et al., 2022). Traditional anomaly detection techniques, such as one-class SVM or k-nearest neighbours, struggle with scalability and feature extraction in complex data (Lee & Kim, 2025). Deep learning overcomes these limitations by using layered architectures to hierarchically learn representations, making it ideal for processing raw network flows or logs.

Deep Learning Fundamentals

Deep learning models excel in anomaly detection due to their ability to handle unstructured data and learn intricate patterns. Key architectures relevant to APT detection include:

Convolutional Neural Networks (CNNs): Designed for spatial data, CNNs use convolutional layers to extract features from network traffic images or matrices (Agarwal & Gupta, 2024). They are effective for identifying patterns in packet headers or flow statistics.

Recurrent Neural Networks (RNNs) and LSTMs: These capture temporal dependencies in sequential data like logs or traffic over time (Kim et al., 2024). LSTMs mitigate the vanishing gradient problem, making them suitable for detecting prolonged APT activities.

Auto Encoders: Unsupervised models that reconstruct input data, with anomalies identified by high reconstruction errors. Variants like variational auto encoders enhance robustness (Li & Liu, 2024).

Generative Adversarial Networks (GANs): Consist of a generator and discriminator; they generate synthetic anomalies for training, addressing data imbalance in APT datasets (Gupta & Sharma, 2025).

Transformers: Use self-attention mechanisms to model long-range dependencies in time-series data, outperforming RNNs in scalability and parallelization (Kim et al., 2024).

Hybrid models combine these, such as CNN-LSTM for spatial-temporal analysis or GAN-Transformer for augmented, context-aware detection.

Literature Review

Recent research has advanced deep learning applications for APT anomaly detection. One study proposes a fusion model integrating Isolation Forest for initial anomaly isolation, GAN for data augmentation, and Transformer for time-series modeling, achieving 96.75% accuracy on UNSW-NB15 with reduced false alarms (Kim et al., 2024). This approach handles high-dimensional network data effectively, improving stability in log analysis.

Another work introduces a hybrid CNN-LSTM model, leveraging CNN for spatial feature extraction from network logs and LSTM for temporal dependencies in event sequences (Alshamrani & Alqahtani, 2024). Tested on UNSW-NB15, it attained 98.5% accuracy, 97.8% precision, and 0.995 AUC, outperforming SVM and Random Forests by minimizing false positives. A comparative analysis of machine learning models on NSL-KDD showed a 6-layer deep neural network achieving 98.85% accuracy and 1.13% FPR, surpassing C5.0 decision trees (95.64%) and Bayesian networks (88.37%) (Abdulrazzaq & Al-Ani, 2023). The deep model's automatic feature extraction proved superior for APT classification.

A report on AI for APT detection emphasizes supervised and unsupervised learning, with deep learning techniques like CNNs and RNNs for complex pattern recognition (Walsh et al., 2024). It discusses frameworks like MLAPT for integrating low-level alerts into APT sequences, addressing challenges like novel attacks and low-and-slow movements.

An anomaly-based IDS using deep learning reviews models such as MLP (81.43% accuracy), CNN-BiLSTM (83.58%), and GRU-DAE (90.21%), proposing a sustainable system for real-time APT detection through data pre-processing, model training, and integration with NIDS/HIDS (Zhang et al., 2023). Other contributions include BiLSTM-GCN for explainable APT detection, genetic programming for feature evolution, and hybrid optimization with deep learning for scalable log analysis (Alsulami et al., 2024; Gupta & Sharma, 2025).

Deep Learning-Based Models for APT Anomaly Detection

Hybrid CNN-LSTM Model

The CNN-LSTM hybrid is a prominent model for APT detection (Alshamrani & Alqahtani, 2024). Data pre-processing involves cleaning, normalization, and one-hot encoding on datasets like UNSW-NB15. The architecture features three CNN layers with ReLU activation and max pooling for spatial features, followed by two LSTM layers (50 units each) for temporal analysis, ending with a softmax output. Training uses back propagation, grid search for hyper parameters, and 5-fold cross-validation. This model captures both static patterns in traffic and dynamic sequences in logs, making it effective for multi-stage APTs.

Isolation Forest-GAN-Transformer Fusion

This model addresses data scarcity and time-series complexity (Gupta & Sharma, 2025). Isolation Forest isolates anomalies via random trees, GAN generates synthetic data to balance classes, and Transformer uses self-attention for global context. On datasets like CIC-IDS2017, it achieves high F1-scores (92.84%) and efficiency (low flops and inference time). It's scalable for large logs, reducing false alarms in network anomaly detection.

Multi-Layer Deep Neural Network

A 6-layer DNN with four hidden layers (50 neurons each) and Maxout activation excels in feature extraction without manual intervention. On NSL-KDD, it classifies normal vs. anomaly with 98.85% accuracy, low FPR, and high specificity, demonstrating deep learning's superiority over traditional ML for APT patterns.

Other Models

BiLSTM-GCN combines bidirectional LSTMs for sequence modeling with graph convolutions for relational data, providing explainability. Genetic programming evolves features for enhanced detection, while hybrid optimization with DL improves scalability for persistent threats.

Case Studies and Applications

In practical applications, these models have been tested on real-world scenarios. For instance, the CNN-LSTM model detected APT-like intrusions in simulated enterprise networks, identifying lateral movement through anomalous user behaviours. The fusion model applied to log analysis in cloud environments reduced detection time for stealthy exfiltration. Academic frameworks like HOLMES correlate alerts for APT sequence detection, integrating DL with graph analysis. Commercial tools incorporate similar techniques for endpoint detection and response (EDR), enhancing threat hunting in sectors like finance and government.

Challenges and Future Directions

Despite advancements, challenges persist. Data scarcity for APTs leads to imbalanced datasets, addressed partially by GANs but requiring better synthetic generation. Model interpretability is crucial for trust; explainable AI (XAI) techniques like SHAP can help. Computational demands limit deployment in resource-constrained environments, necessitating efficient architectures. Adversarial attacks on DL models, where attackers poison data, pose risks, mitigated by robust training. Future directions include integrating DL with semantic reasoning for hybrid frameworks, exploring multi-agent reinforcement learning for adaptive defenses, and focusing on sustainability by optimizing energy use in training. Benchmarking on diverse datasets and real-time implementation will drive progress.

Conclusion

Deep learning-based anomaly detection models offer a powerful paradigm for combating APTs, with hybrid approaches achieving superior performance in accuracy and efficiency. By leveraging architectures like CNN-LSTM and GAN-Transformer, these models address the stealth and persistence of APTs, outperforming traditional methods. While challenges remain, ongoing research in XAI and adaptive learning promises further enhancements. Organizations should adopt layered defenses incorporating these technologies to safeguard against evolving threats.

REFERENCES

Abdulrazzaq, S., & Al-Ani, M. S. (2023). Advanced Persistent Threats Detection based on Deep LearningApproach.ResearchGate.https://www.researchgate.net/publication/371033431_Advanced_Persistent_Threats_Detection_based_on_Deep_Learning_Approach

Agarwal, A., & Gupta, R. (2024). Deep Learning Techniques for Anomaly Detection in Cybersecurity,ResearchGate.https://www.researchgate.net/publication/389089829_Deep_Learning_Techniques_for_Anomaly_Detection_in_Cybersecurity

Al-Rimy, B. A. S., Maarof, M. A., & Shaid, S. Z. M. (2020). Early detection of the advanced persistent threat attack using performance metrics. https://arxiv.org/pdf/2009.10524

Alshamrani, A., & Alqahtani, A. S. (2024). A Deep Learning Approach to Detecting Advanced Persistent Threats in Cybersecurity. International Journal of Advanced Engineering and Management.https://ijaem.net/issue_dcp/A%20Deep%20Learning%20Approach%20to%20Detecting%20Advanced%20Persistent%20Threats%20in%20Cybersecurity.pdf

Alsulami, A. A., Alqahtani, A. S., & Alshehri, M. D. (2024). Explainable deep learning approach for advanced persistent threats detection. Artificial Intelligence Review, 57(10), Article 289. https://doi.org/10.1007/s10462-024-10890-4

Chandrasekaran, S., & Gupta, R. (2025). A computational framework for IoT security integrating deep learning and semantic reasoning. Scientific Reports, 14, Article 93898. https://doi.org/10.1038/s41598-025-93898-2

Eke, H. N., Petrovski, A., & Ahriz, H. (2023). Advanced Persistent Threat Detection Through Multi-Layered,Machine,Learning.Preprints.org. https://www.preprints.org/frontend/manuscript/5b192e325327f008e712e754c35e7e25/download_pub

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications, 50, Article 102419. (Referenced in broader literature; specific link not provided in search.)

Ghafir, I., Saleem, J., Hammoudeh, M., Faour, H., Prenosil, V., Jaf, S., Jabbar, S., & Baker, T. (2018). Security threats to critical infrastructure: The human factor. The Journal of Supercomputing, 74(10), 4986-5002. (Broader context; specific APT focus in related works.)

Gupta, S., & Sharma, A. (2025). Advanced Persistent Threat Detection Using Optimized and Hybrid Deep Learning Models. Security and Privacy, 8(1), e70011. https://doi.org/10.1002/spy2.70011

Kim, J., Shin, N., Jo, S. Y., & Kim, S. H. (2024). Deep Learning-based Anomaly Detection and Log Analysis for Computer Networks. https://arxiv.org/pdf/2407.05639

Lee, S., & Kim, J. (2025). Machine Learning-Based Intrusion Detection Systems: Capabilities, Methodologies, and Open Research Challenges. https://www.techrxiv.org/users/747440/articles/1256948-machine-learning-based-intrusion-detection-systems-capabilities-methodologies-and-open-research-challenges

Li, Y., & Liu, Q. (2024). Deep Learning-Based Network Security Threat Detection and Defense. International Journal of Intelligent Systems, 15(11), 64. https://thesai.org/Downloads/Volume15No11/Paper_64-Deep_Learning_Based_Network_Security_Threat_Detection.pdf

Myneni, S., Chowdhary, A., Sabur, A., Alshamrani, S. S., Agrawal, A., Srivastava, G., & Huang, D. (2022). DAPT 2020 - Constructing a Benchmark Dataset for Advanced Persistent Threats. In Proceedings of the 2020 IEEE International Conference on Big Data (pp. 1-10). (Referenced in report.)

Narayanan, A., & Gupta, R. (2024). Advanced Data Anomaly Detection with Machine Learning. Acceldata Blog. https://www.acceldata.io/blog/advanced-data-anomaly-detection-with-machine-learning-a-step-by-step-guide

Nguyen, T. T., & Armitage, G. (2008). A survey of techniques for internet traffic classification using machine learning. IEEE Communications Surveys & Tutorials, 10(4), 56-76. (Broader ML context.)

Raman, M. G., Somu, N., Kirthivasan, K., Liscano, R., & Sriram, V. S. (2017). An efficient intrusion detection system based on hypergraph - Genetic algorithm for parameter optimization and feature selection in support vector machine. Knowledge-Based Systems, 134, 1-12. (Referenced in related work.)

Sarker, I. H. (2021). Machine Learning-Based Anomaly Detection Model for Cybersecurity. International Information and Engineering Technology Association. https://www.iieta.org/journals/isi/paper/10.18280/isi.290628

Walsh, M., Worrell, C., & Scanlon, T. (2024). Toward the Use of Artificial Intelligence (AI) for Advanced Persistent Threat Detection (CMU/SEI-2024-TR-001). Software Engineering Institute, Carnegie Mellon University. https://www.sei.cmu.edu/documents/6061/Toward-the-Use-of-Artificial-Intelligence-AI-for-Advanced-Persistent-Threat-Detection_.pdf

Zhang, H., Li, J., & Li, X. (2023). Anomaly-Based Intrusion Detection System To Detect Advanced Persistent Threats. E3S Web of Conferences, 399, Article 01106. https://doi.org/10.1051/e3sconf/202339901106