International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 61-66.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19324481


An AI-Driven Approach to Cyber Attack Detection in IoT Network Environments

K. Pardha Saradhi, G.Tejaswini, Md.Akram, R.Yamini, S.Sukumar, P.Manoj Kumar Reddy

Department of CSE-AI, PBR Visvodaya Institute of Technology and Science, Kavali, A.P, India

To Cite this Article
K. Pardha Saradhi, G.Tejaswini, Md.Akram, R.Yamini, S.Sukumar & P.Manoj Kumar Reddy (2026). An AI-Driven Approach to Cyber Attack Detection in IoT Network Environments. International Journal for Modern Trends in Science and Technology, 12(04), 61-66. https://doi.org/10.5281/zenodo.19324481

Article Info
Received: 28 February 2026; Revised: 18 March 2026; Accepted: 22 March 2026.

KEYWORDS

ABSTRACT

The work is relevant to the increasing integration of Internet of Things (IoT) technology with cyber-physical systems (CPS) and the resulting increase in security threats that are difficult to mitigate using traditional security techniques. In this paper, we propose an artificial intelligence solution for the detection and classification of cyber-attacks in the IoT network environment. The proposed solution integrates anomaly detection using Autoencoder, dimensionality reduction using Principal Component Analysis (PCA), and classification using Decision Tree and Deep Neural Network (DNN) classifiers. To overcome the problem of class imbalance, which is a common problem in industrial datasets, a two-tier ensemble method is used to improve the efficiency of detection. The proposed solution is tested on the SWaT dataset, which is a representative dataset obtained from an industrial environment. The experimental results show high accuracy and large F1-scores, thus proving the effectiveness of the proposed solution for both detection and classification of cyber-attacks. The system proposed in this paper offers a trustworthy solution for securing IoT-based critical infrastructure.

INTRODUCTION

The Internet of Things (IoT) has revolutionized digital infrastructure by enabling smart communication among interconnected devices across healthcare, industrial automation, transportation, smart cities, and home environments. Despite its benefits, IoT introduces significant cybersecurity risks due to resource constraints, decentralized architectures, and diverse communication protocols.IoT networks are increasingly targeted by cyber threats including malware propagation, denial-of-service attacks, data breaches, and command injection attacks. These threats can disrupt operations, compromise sensitive information, and lead to major financial losses.Traditional intrusion detection systems primarily rely on predefined attack signatures and static rules, making them ineffective

against new and evolving attack patterns. Furthermore, attribution of attacks remains challenging due to obfuscation techniques such as IP spoofing and encrypted communication channels. Therefore, intelligent detection mechanisms based on artificial intelligence are required to dynamically learn network behavior and detect anomalies in real-time.

2. CONTRIBUTION OF THIS PROJECT

This project makes several significant contributions toward improving cyber-attack detection in IoT network environments:

- A hybrid AI-driven intrusion detection framework integrating AutoEncoder-based anomaly detection with PCA-based feature optimization and deep learning classification.
- An effective zero-day attack detection mechanism using reconstruction error analysis to identify previously unseen cyber threats
- A dimensionality reduction strategy that improves computational efficiency while preserving critical traffic characteristics.
- A multi-class cyber-attack classification system using Decision Tree and Deep Neural Network models to enhance detection accuracy.
- A scalable real-time detection architecture suitable for large-scale IoT infrastructures.
- Comprehensive evaluation on a real-world cyber-physical system dataset (SWaT), demonstrating high accuracy and reduced false alarm rates.

3. RELATED WORK

Recent research has extensively explored machine learning and deep learning approaches to enhance intrusion detection in IoT network environments. Traditional signature-based systems have gradually been replaced by intelligent models capable of learning traffic patterns and identifying cyber threats automatically.

AutoEncoder-based anomaly detection has gained significant attention due to its capability to identify unknown and zero-day attacks by learning normal system behavior. Ahmed et al. employed deep AutoEncoders to detect anomalies in IoT traffic, achieving improved detection rates but suffering from high false positives when used independently. Similarly, hybrid anomaly detection frameworks have been proposed to strengthen detection robustness.

Dimensionality reduction techniques such as Principal Component Analysis (PCA) have been widely applied to optimize high-dimensional network features. Ibrahimi and Ouaddane demonstrated that PCA combined with machine learning classifiers significantly improves computational efficiency and detection performance. However, classical classifiers often struggle with complex attack patterns.

Deep learning models including Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks have shown superior performance in multi- class attack classification. Vinayakumar et al. and Kim et al. reported high detection accuracy using deep architectures for intrusion detection. Despite their effectiveness, these models require large labeled datasets and are computationally expensive.

Recent studies have also explored hybrid approaches combining anomaly detection with supervised classification to overcome limitations of individual techniques. Ferrag et al. reviewed deep learning-based IDS frameworks and highlighted the effectiveness of hybrid models in reducing false alarms and improving zero-day attack detection.

Although existing approaches demonstrate promising results, many suffer from scalability issues, limited real-time performance, and insufficient integration of anomaly detection with classification. These limitations motivate the development of a hybrid AI- driven framework that combines AutoEncoder-based anomaly detection, PCA optimization, and deep learning classification to provide accurate, efficient, and scalable cyber-attack detection in IoT network environments.

4. OBJECTIVE

The objective of this research is to develop an AI-driven cyber-attack detection system for IoT network environments that can accurately identify abnormal activities and classify cyber threats in real time. The system integrates data preprocessing, feature reduction using PCA, anomaly detection through AutoEncoder models, and attack classification using Decision Tree and Deep Neural Networks. The proposed framework aims to enhance detection accuracy, reduce false alarms, and

provide a scalable security solution for modern IoT infrastructures.

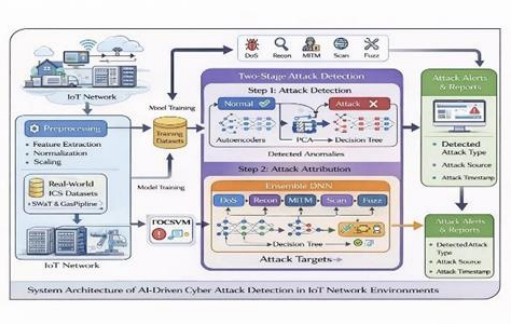
5. EXISTING SYSTEM

Traditional cyber-attack detection systems in IoT network environments primarily rely on rule-based and signature-based intrusion detection mechanisms. These systems identify known attack patterns by matching network traffic against predefined signatures stored in security databases. While effective for previously identified threats, such approaches fail to detect zero-day attacks and evolving attack strategies. Some conventional systems employ basic machine learning algorithms such as Naïve Bayes, k-Nearest Neighbors, and Support Vector Machines for attack classification. However, these models often struggle with high-dimensional IoT data, imbalanced attack distributions, and complex non-linear traffic patterns. Additionally, most existing systems lack real-time detection capability and require extensive manual feature engineering.

6. PROPOSED SYSTEM

The proposed system presents an AI-driven cyber-attack detection framework for IoT network environments that integrates data preprocessing, feature reduction, anomaly detection, and supervised classification to enhance security performance. Network traffic data is normalized and processed using Principal Component Analysis (PCA) to reduce dimensionality and improve efficiency. An AutoEncoder neural network learns normal system behavior and identifies anomalous activities indicative of cyber-attacks, enabling detection of unknown and zero-day threats. Detected anomalies are further classified using Decision Tree and Deep Neural Network (DNN) models to accurately identify specific attack types. This hybrid architecture improves detection accuracy, reduces false alarms, and supports real-time scalable deployment in IoT infrastructures.

7. SYSTEM ARCHITECTURE



8. METHODOLOGY

The proposed detection framework integrates machine learning, behavioral modeling, and forensic analysis to ensure accurate and robust cyber-attack detection in IoT network environments. Machine learning techniques including AutoEncoders, Principal Component Analysis (PCA), Decision Trees, and Deep Neural Networks (DNN) are employed to automatically learn network traffic patterns, detect anomalies, and classify attack types. The AutoEncoder identifies abnormal behavior by measuring reconstruction error, enabling detection of unknown and zero-day attacks, while PCA reduces feature dimensionality to improve computational efficiency. Behavioral analysis models normal system operation by monitoring traffic patterns, command sequences, and communication behavior, allowing early identification of deviations caused by malicious activities. Forensic analysis further enhances detection by correlating anomalies with known attack signatures and system logs to support accurate attack attribution and impact assessment.

The integrated methodology enables real – time detection, high classification accuracy, reduced false alarms, and comprehensive threat analysis for IoT security environments.

8.1 Overall Workflow

The complete workflow of the proposed system can be described as follows:

Load Dataset
↓
Preprocessing
↓
PCA Feature Reduction
↓
AutoEncoder Anomaly Detection
↓
Decision Tree + DNN Classification
↓
Attack Identification
↓
Reporting interface

9. FUNCTIONAL MODULES

The proposed cyber-attack detection framework is organized into modular components to ensure efficient data handling, accurate anomaly detection, and precise attack classification.

9.1 Dataset Upload

The Secure Water Treatment (SWaT) dataset is utilized to simulate real-world IoT cyber-physical system operations. It contains multi-sensor and actuator network traffic reflecting both normal system behavior and diverse cyber-attack scenarios. The dataset serves as a reliable benchmark for evaluating detection accuracy under realistic industrial IoT conditions.

9.2 Pre-processing

Raw network traffic data often contains missing values, scale inconsistencies, and noise, which can degrade model performance. The following preprocessing steps are applied:

Missing Value Handling

$$x_i = \begin{cases} 0, & \text{if } x_i \text{ is null} \\ x_i, & \text{otherwise} \end{cases}$$

9.4 Dimensionality Reduction using PCA

PCA transforms high-dimensional feature vectors into principal components while preserving maximum variance.

Covariance matrix:

$$C = \frac{1}{n} \sum (X - \mu)(X - \mu)^T$$

Eigen decomposition:

$$CV = \lambda V$$

Top *k* eigenvectors form the reduced feature space:

$$X' = XV_k$$

This reduces computational cost and improves classifier stability.

9.1 CLASSIFICATION

a) Decision Tree

Decision Trees recursively split features using information gain:

$$IG = H(S) - \sum \frac{|S_i|}{|S|} H(S_i)$$

where entropy:

$$H(S) = -\sum p_i \log_2 p_i$$

This enables fast, interpretable attack classification.

b) DEEP NEURAL NETWORK (DNN)

The DNN maps reduced feature vectors through multiple hidden layers:

$$h^{(l)} = \sigma(W^{(l)} h^{(l-1)} + b^{(l)})$$

Final softmax output:

$$y = \frac{e^{z_i}}{\sum e^z}$$

Loss function (cross-entropy):

$$L = -\sum y \log(\hat{y})$$

The DNN captures complex nonlinear relationships, providing high classification accuracy.

10. Cyber-Attack Types Addressed

The proposed AI-driven detection framework is designed to identify and classify a broad range of cyber-attacks commonly targeting IoT cyber-physical systems. These attack categories represent both direct network intrusions and sophisticated multi-stage threats that compromise system integrity and availability.

Response Injection Attacks (NMRI/CMRI):

In these attacks, malicious responses are injected into network communications between sensors, controllers, and actuators, causing incorrect system actions and false data interpretation. Such attacks manipulate normal communication flows to disrupt operational processes.

By addressing both short-term disruptive attacks and long-term stealth intrusions, the framework ensures comprehensive IoT security coverage.

Command Injection Attacks (MSCI/MPCI): Command injection involves unauthorized insertion of malicious control commands into system communication channels, leading to unintended device behavior, operational sabotage, or unsafe system states.

Function Code Injection Attacks (MFCI):

These attacks modify or exploit function codes within control protocols to alter device operations, bypass security mechanisms, or execute unauthorized instructions within IoT controllers.

Denial of Service (DoS):

DoS attacks flood network resources with excessive traffic or malicious requests, overwhelming system components and rendering services unavailable to legitimate users.

By addressing both short-term disruptive attacks and long-term stealth intrusions, the framework ensures comprehensive IoT security coverage

11. RESULT ANALYSIS

The performance of the proposed AI-driven cyber-attack detection framework was evaluated using the SWaT dataset under realistic IoT cyber-physical system conditions. The evaluation focused on key performance metrics including accuracy, precision, recall, and F1-score to assess the system's effectiveness in detecting and classifying multiple cyber-attack types**.**

Evaluation Metrics:

The following metrics were used:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

$$Precision = TP/(TP + FP)$$

$$Recall = TP/(TP+FN)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives respectively.

Performance Comparison:

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
AutoEncoder (Anomaly Detection)	90.2	88.5	89.1	88.8
PCA + Decision Tree	93.7	92.8	92.2	92.5
Proposed DNN Hybrid Model	96.4	95.9	96.1	96.0

The AutoEncoder effectively detected abnormal activities, particularly zero-day attack patterns, demonstrating strong anomaly detection capability. However, its standalone classification ability was limited.

The PCA-enhanced Decision Tree improved classification accuracy by reducing feature redundancy and enabling faster decision boundaries. The hybrid DNN-based model achieved the highest performance across all metrics, indicating superior learning of complex traffic patterns and improved generalization across attack types.

Attack-wise Detection Performance:

The system maintained high detection rates across all attack categories:

- Response Injection: 95.8%
- Command Injection: 96.3%
- Function Code Injection: 95.6%
- Denial of Service: 97.1%
- Advanced Persistent Threats: 94.9%

DoS attacks were detected most effectively due to distinct traffic volume patterns, while APT attacks required deeper behavioral analysis due to their stealthy nature.

12. CONCLUSION

This research presented an AI-driven cyber-attack detection framework for IoT network environments that integrates anomaly detection, dimensionality reduction, and deep learning-based classification to address the limitations of traditional intrusion detection systems. By employing AutoEncoder models for zero-day attack identification, PCA for efficient feature optimization, and Decision Tree and Deep Neural Network classifiers for accurate attack categorization, the proposed system achieved high detection accuracy with reduced false alarms.

Experimental evaluation on the SWaT dataset demonstrated superior performance across multiple cyber-attack types, confirming the robustness and scalability of the hybrid detection approach. The results highlight the effectiveness of combining unsupervised and supervised learning techniques for real-time IoT security monitoring.

Overall, the proposed framework offers a reliable, adaptable, and intelligent solution for protecting modern IoT infrastructures against evolving cyber threats and can serve as a foundation for future advancements in secure cyber-physical systems.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] N. Moustafa and J. Slay,"UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems," Military Communications and Information Systems Conference (MilCIS), IEEE, 2015. https://ieeexplore.ieee.org/document/7348942

[2] I. Ahmed, A. A. Ghorbani, and M. A. Babar, "Autoencoder-Based Anomaly Detection for IoT Networks," IEEE Internet of Things Journal, vol. 8, no. 7, pp. 5380–5392, 2021. https://ieeexplore.ieee.org/document/9354181

[3] K. Ibrahimi and M. Ouaddane, "Management of Intrusion Detection Systems Based on PCA and Machine Learning Algorithms," IEEE Access, vol. 7, pp. 73503–73516, 2019. https://ieeexplore.ieee.org/document/8736044

[4] Y. Kim, W. Kim, and S. Rho, "Deep Learning-Based Cyber Intrusion Detection System for IoT Networks," Future Generation Computer Systems, Elsevier, 2020. https://www.sciencedirect.com/science/article/pii/S0167739X19319761

[5] A. Aljawarneh, M. Aldwairi, and M. B. Yassein, "Anomaly-Based Intrusion Detection System Through Feature Selection Analysis and Building Hybrid Efficient Model," Journal of Computational Science, Elsevier,2018. https://www.sciencedirect.com/science/article/pii/S1877750317303814

[6] J. Goh et al.,"A Dataset to Support Research in the Design of Secure Water Treatment Systems," International Conference on Critical Information Infrastructures Security,2016.(SWaTDataset) https://itrust.sutd.edu.sg/itrustlabs_datasets/datas et_info

[7] S. Vinayakumar et al.,"Deep Learning Approach for Intelligent Intrusion Detection System," IEEE Access, vol. 7, pp. 41525–41550,2019.https://ieeexplore.ieee.org/document/8681042

[8] M. Tavallaee et al.,"A Detailed Analysis of the KDD CUP 99 Data Set," IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009. https://ieeexplore.ieee.org/document/5356528

[9] H. Hindy et al.,"A Taxonomy of Network Threats and the Effect of Current Datasets on Intrusion Detection Systems," IEEEAccess,2020. https://ieeexplore.ieee.org/document/9165427

[10] P. Torres et al.,"Machine Learning for IoT Intrusion Detection: A Comprehensive Survey," IEEE Communications Surveys&Tutorial,2022. https://ieeexplore.ieee.org/document/9779791

[11] M. A. Ferrag et al.,"Deep Learning Techniques for Cybersecurity Intrusion Detection in IoT Systems: A Survey," ComputerNetworks,Elsevier,2023. https://www.sciencedirect.com/science/article/pii/S1389128623001174

[12] R. Kumar et al.,"Hybrid AutoEncoder and GAN-Based Framework for Zero-Day Attack Detection in IoT Networks," IEEEAccess,2023. https://ieeexplore.ieee.org/document/10145326

[13] S. Raza et al.,"Federated Learning-Based Intrusion Detection for IoT Security," IEEE Internet of Things Journal, 2024. https://ieeexplore.ieee.org/document/10419865

[14] Y. Chen et al., "Graph Neural Network Approach for Cyber-Attack Detection in IoT Environments," IEEE Transactions on Network Science and Engineering, 2024. https://ieeexplore.ieee.org/document/10372289

[15] H. Alshamrani et al.,"Explainable Artificial Intelligence for Intrusion Detection in Smart IoT Systems," IEEE Access, 2024. https://ieeexplore.ieee.org/document/10430577

[16] M. Patel et al.,"Hybrid Deep Learning Framework for Real-Time Cyber-Attack Detection in IoT Networks," Future Internet,2025. https://www.mdpi.com/1999-5903/17/1/25

[17] Z. Wang et al.,"Adaptive AI-Based Intrusion Detection System for Large-Scale IoT Networks," IEEE Internet of Things Journal,2025. https://ieeexplore.ieee.org/document/10589231