International Journal of AI, Big Data, Computational and Management Studies
Noble Scholar Research Group | Volume 7, Issue 2, 153-157, 2026
ISSN: 3050-9416 | https://doi.org/10.63282/3050-9416.IJAIBDCMS-V7I2P123

Original Article

Unsupervised Zero-Day Intrusion Detection in IoT Networks using Cycle-Consistent Adversarial Networks

Naresh Kalimuthu
Independent Researcher, USA.

Received On: 08/03/2026 ***Revised On: 10/04/2026*** ***Accepted On: 18/04/2026*** ***Published On: 24/04/2026***

Abstract: *The rapid growth of the Internet of Things (IoT) has created new vulnerabilities in global networks, as traditional signature-based intrusion detection systems (IDS) find it difficult to detect new "zero-day" threats. This study explores the use of Cycle-Consistent Adversarial Networks (CycleGAN) for unsupervised anomaly detection. By employing unpaired domain translation, CycleGAN models the statistical patterns of normal network traffic and detects intrusions through reconstruction-error analysis. The report addresses key challenges, including data imbalance, resource limitations in edge computing, and the emergence of polymorphic attacks. Experimental results show that CycleGAN-based frameworks achieve higher detection rates on benchmark datasets while keeping false positives low.*

Keywords: *Internet of Things (IoT), Zero-Day Detection, Cycle-Consistent Adversarial Networks (CycleGAN), Unsupervised Learning, Anomaly Detection, Network Security.*

1. Introduction

The modern digital environment is undergoing a fundamental shift driven by the widespread adoption of the Internet of Things (IoT) across industrial, commercial, and residential sectors. Experts estimate that the worldwide number of connected IoT devices could exceed 40 billion by 2034. This extraordinary connectivity has transformed fields such as smart manufacturing, healthcare, and urban planning, but it has also significantly expanded the cyber-attack surface. The main challenge of this era is balancing the social and economic benefits of IoT with the risks posed by billions of vulnerable endpoints.

Traditional network security measures, especially signature-based Intrusion Detection Systems (IDS), are becoming less effective in this evolving landscape. These systems rely on detecting predefined patterns, or "signatures," of known malicious activity. Although they excel at identifying established threats, they cannot detect "zero-day" attacks those that exploit unknown vulnerabilities or employ new tactics not yet recognized by security vendors. The time gap between the discovery of a zero-day vulnerability and the release of a security patch creates a critical "window of vulnerability," leaving networks exposed to significant breaches. In IoT networks, which comprise diverse devices with limited security features and infrequent updates, this window can remain open for long periods, making them especially attractive targets for advanced attackers.

To overcome the limitations of signature-based defenses, research has shifted towards anomaly-based intrusion detection, which focuses on detecting deviations from typical network behavior. Machine Learning (ML) and Deep Learning (DL) are now central tools, analyzing extensive, high-dimensional network traffic data to identify subtle signs of malicious activity. Nonetheless, many ML systems depend on supervised learning, requiring large, accurately labeled datasets of both safe and malicious traffic for training. In the case of zero-day threats, such labeled attack data is inherently unavailable, making supervised methods less effective for proactive security

Unsupervised learning offers a more practical option because it only needs unlabeled "normal" data to create a baseline of genuine network activity. Among unsupervised methods, Generative Adversarial Networks (GANs) have shown a remarkable ability to capture complex data patterns. A GAN comprises two neural networks in competition: a generator that tries to produce convincing synthetic data, and a discriminator that aims to distinguish real from fake samples. This rivalry helps the generator effectively learn the statistical nuances of the training data.

This report evaluates the effectiveness of Cycle-Consistent Adversarial Networks (CycleGAN) for unsupervised zero-day detection. CycleGAN, a specialized form of GAN, is designed for "unpaired domain translation," enabling it to learn mappings between two domains without needing direct sample-to-sample correspondence. In intrusion detection, this means the system can identify shifts and transformations in network traffic even without specific

attack examples. The cycle-consistency feature where a sample translated to another domain and then back should match the original offers a reliable way to measure reconstruction errors. When a zero-day anomaly packet is processed by a CycleGAN trained solely on normal traffic, the reconstruction fails, producing a high error score that signals potential threats to security teams. This research aims to combine these advanced adversarial techniques with the operational needs of IoT environments to create a comprehensive framework for securing future interconnected systems.

2. Research Topics

Developing a dependable CycleGAN-based intrusion detection system for IoT networks entails addressing several complex challenges. These stem from the distinctive characteristics of IoT ecosystems, the dynamic nature of cyber threats, and the mathematical challenges linked to adversarial techniques training.

2.1. Challenge 1: Severe Class Imbalance and the Scarcity of Zero-Day Data

In IoT security research, a major challenge is the vast difference between benign and malicious data. In typical networks, normal traffic vastly exceeds attack data, which is rare and often short-lived. This imbalance is even starker for zero-day attacks, with no prior data to learn from. Supervised machine learning models trained on such skewed datasets tend to favor the majority class normal traffic resulting in high accuracy for benign activity but causing many false negatives by missing rare attack types.

Traditional data balancing methods, like Synthetic Minority Over-sampling Technique (SMOTE) or random oversampling, often fall short for complex network flow data. These approaches usually interpolate linearly between existing minority samples, which doesn't capture the complex, non-linear relationships and temporal correlations characteristic of IoT protocols. Moreover, they cannot generate entirely new attack types; they merely replicate known ones. The challenge is to leverage CycleGAN's generative power not only to balance datasets but also to create realistic, unseen attack scenarios by converting normal traffic into anomalous patterns in an unsupervised way. This involves training the GAN to learn latent representations of "anomaly" concepts without explicit attack labels, requiring advanced architecture and loss functions.

2.2. Challenge 2: Computational Constraints and Real-Time Processing at the Edge

IoT networks are diverse, including everything from powerful gateways to highly resource-limited sensors and actuators. Most of these devices run on batteries and have limited processing power, memory, and storage, so they cannot support complex deep learning models. Nevertheless, the rapid pace of modern cyber-attacks demands real-time or near-real-time intrusion detection to stop lateral movement and data theft.

Adversarial networks like CycleGAN are costly to train and require substantial resources for inference. A typical CycleGAN setup includes two generators and two discriminators, each made up of multiple deep convolutional or residual layers. Running such models directly on IoT end-devices is often impractical both physically and economically. This highlights a key research challenge: how to develop a detection framework that uses CycleGAN's capabilities while meeting the strict latency and power constraints of IoT edge devices. The solution involves exploring hierarchical architectures, where heavy computation is shifted to gateways or cloud servers, and lightweight, optimized models perform local monitoring. Additionally, the system must sustain high throughput, processing thousands of packets per second without causing network delays.

2.3. Challenge 3: Adaptation to Dynamic Environments and Polymorphic Attack Evasion

IoT environments are naturally dynamic, with devices regularly connecting and disconnecting, firmware updates changing communication patterns, and user behavior varying seasonally or operationally. A static intrusion detection system, even one using advanced GANs, risks becoming outdated as the understanding of "normal" behavior evolves—a challenge called concept drift. If the IDS cannot adapt to these harmless changes, it will produce more false positives, resulting in "alert fatigue" and possibly causing operators to overlook real threats.

At the same time, cyber adversaries are increasingly using artificial intelligence to develop polymorphic and adversarial attacks aimed at bypassing current detection systems. Polymorphic malware changes its code or communication patterns with each version, while adversarial attacks add subtle, mathematically engineered tweaks to network traffic that evade traditional filters but can fool deep learning classifiers. Therefore, research should focus on developing a "robust" detection system that not only spots zero-day anomalies but also resists evasion techniques. This requires ongoing adversarial training, where GANs produce adversarial samples to strengthen the detector, creating a defensive loop that adapts to evolving threats.

3. Recommendations / Mitigation Strategies

Implementing a CycleGAN-based intrusion detection system (IDS) for IoT networks involves a systematic process that covers data preprocessing, model design, and deployment methods. These recommendations are based on the latest best practices in adversarial network research and IoT security.

3.1. Unsupervised Anomaly Detection via Cycle-Consistent Reconstruction

The main recommendation for zero-day detection involves implementing a reconstruction-based anomaly detection system. This setup trains a CycleGAN solely on normal network traffic, teaching it two directional mappings: one to encode data into a latent form and another to decode it back to the original. Since the model is optimized to reduce

cycle-consistency loss for legitimate traffic, it becomes very effective at reconstructing benign packets. When a zero-day attack happens, the network traffic exhibits statistical properties that deviate from the established normal distribution. Consequently, the CycleGAN cannot accurately reconstruct these unusual samples. The anomaly score is determined by the size of the reconstruction error, which is the difference between the original input and its reconstructed output.

This is the mathematical expression of the overall loss function used in this unsupervised detection process:

$$L_{Total} = L_{GAN}(G, D_Y) + L_{GAN}(F, D_x) + \lambda L_{cyc}(G, F)$$

Where:

- L_{GAN} represents the adversarial loss that ensures the generated samples are indistinguishable from the target distribution.
- L_{cyc} is the cycle-consistency loss
- λ is a hyperparameter that weights the importance of cycle consistency

By establishing a dynamic threshold for reconstruction error, the system can detect packets that deviate from the normal profile and flag them as potential zero-day threats.

3.2. Data Augmentation and Cross-Domain Transformation

To address data scarcity and class imbalance issues, using CycleGAN as a data augmentation tool is recommended. This approach considers "Normal Traffic" and "Malicious Traffic" as separate domains. The generator learns to convert normal behavior data into synthetic intrusive data, effectively mimicking how legitimate traffic might look if compromised by an attack.

This transformation process provides several benefits:

- *Diversity of Training Data:* It produces a large number of synthetic attack samples that supplement the existing, likely small, dataset of known threats.
- *Generalization*: By adding variations to the synthetic data, the model learns to recognize not only specific attack signatures but also the overall statistical traits of malicious activity.
- *Adversarial Hardening:* Synthetic samples can be employed for adversarial training of secondary classifiers, like Random Forest or XGBoost, enhancing their resilience against subtle perturbations in polymorphic malware.

3.3. Hierarchical Deployment and Resource Management

To address the resource constraints of IoT devices, a decentralized hierarchical deployment of IDS components is essential. This report advocates a three-tier architecture.

Table 1: Layered Network Architecture for GAN-Based IoT Anomaly Detection and Resource Allocation

Network Layer	Component	Function	Resource Requirement
IoT Edge Device	Lightweight Filter	Performs initial screening of packets using simple rule-sets or decision trees.	Minimal (Low-power MCU)
IoT Gateway	CycleGAN Inference	Conducts real-time anomaly detection using the pre-trained CycleGAN model.	Moderate (Edge AI Platform)
Cloud/Fog Center	GAN Training & Analysis	Performs computationally intensive training and long-term behavioral pattern analysis.	High (GPU Clusters)

This hierarchy prioritizes training the adversarial network in the cloud, the most resource-intensive task, while delegating inference to gateways like NVIDIA Jetson Nano or similar hardware capable of real-time deep learning processing. This setup reduces the burden on individual sensors and ensures strong security across the network.

3.4. Optimized Training with WGAN-GP and Feature Engineering

To enhance the stability of the CycleGAN and avoid typical problems such as mode collapse where the generator yields limited data varieties the implementation should include Wasserstein GAN with Gradient Penalty (WGAN-GP). WGAN-GP employs the Wasserstein distance as its loss function, offering a more stable gradient for the generator and leading to higher-quality synthetic samples.

Additionally, the success of the IDS relies significantly on feature engineering. Instead of analyzing raw packet data, the system should derive high-level features that reflect the temporal and relational patterns of IoT traffic. Important suggested features include:

- Hurst Parameter: Used to assess the "long-term memory" of traffic patterns, aiding in differentiating genuine activity spikes from persistent malicious threats.
- Correlation Matrices: Converting tabular network flow data into image-like contour images through correlation matrices enables the CycleGAN to utilize robust 2D-CNN architectures for extracting features.
- Graph Embeddings: Modeling network flows as nodes in a graph enables capturing topological information, which simplifies identifying coordinated botnet activities such as Mirai.

4. Recommendations and Goals Achieved

The use of Cycle-Consistent Adversarial Networks and related generative methods has shown significant success in both theoretical models and practical pilot studies. These

developments confirm the effectiveness of adversarial learning in safeguarding IoT ecosystems.

4.1. Improvements in Detection Accuracy and Recall

A main objective of incorporating CycleGAN is to enhance detection rates for minority and zero-day attack classes. While standard machine learning models usually attain high overall accuracy, they tend to have low recall for rare attacks. Comparative studies indicate that GAN-based frameworks reliably close this gap.

Table 2: Performance Comparison of Advanced Machine Learning Models Across IoT and Network Intrusion Datasets

Dataset	Model Approach	Overall Accuracy	Minority Class Recall
KDD CUP 1999	CycleGAN (Augmentation)	99.81%	100.0%
BoT-IoT	WGAN-GP	99.99%	100.0%
CIC-DDoS2019	CycleGAN (Unsupervised)	97.79%	98.40%
NF-ToN-IoT-v2	XGBoost + Adv. Training	95.30%	94.50% (Adv. Data)
CIC IoT 2023	CZ-ResViT (Contour)	82.00% (Zero-Day)	81.00%

For example, in studies using the BoT-IoT dataset, traditional classifiers improved with SMOTE, reaching 99.99% accuracy overall but only identifying 77-80% of attack instances from the minority class. An optimized WGAN-GP architecture enhanced this detection rate to 100%. Similarly, in zero-day scenarios within the CIC IoT 2023 benchmark, using vision-based contour images from a residual framework achieved an 82% accuracy, markedly surpassing traditional deep learning models that had not seen the attack variants during training.

4.2. Reduction in False Positive Rates (FPR)

A key metric for operational security is the False Positive Rate (FPR), since a high FPR can cause operational fatigue and network downtime. CycleGAN's improved modeling of normal traffic distribution enables more accurate threshold settings.

Experimental results on the UGR'16 dataset show that a triple-discriminator GAN (TDCGAN) effectively minimized the false-alarm rate to nearly zero while discovering four new attack types. Another research found that IDS systems improved with GANs lowered false positives from 12.9% in traditional signature-based systems to 4.2%, nearly tripling precision. This improvement is especially important for IoT devices, where false alerts can deplete battery life and overload management systems.

4.3. Efficiency and Scalability at the Edge

Real-time detection without heavy computational load has been achieved through optimized architectures and hierarchical processing.

- Inference Latency: Models deployed on IoT gateways can process network flows with an average latency of 0.22 to 0.3 seconds. More optimized graph-based models, such as GAT-based IDS, achieved even lower latency around 3.1 ms per instance on an NVIDIA Jetson Nano thereby easily meeting the demands of high-speed network monitoring.
- Storage and Energy: New metrics like Accuracy-per-Joule (APJ) and F1-per-Joule (F1PJ) enable choosing architectures that maximize security while minimizing energy consumption. Optimized models, with memory footprints as tiny as 61.84 kB, are now suitable for deployment in the limited resources of smart agriculture and industrial automation.
- Scalability: By employing distributed preprocessing and chunked data loading, CycleGAN systems have effectively managed large-scale datasets like the 72-million-record BoT-IoT dataset without depleting system memory.

5. Conclusion

The integration of Cycle-Consistent Adversarial Networks (CycleGAN) marks a significant progress in identifying zero-day intrusions in IoT networks. Using an unsupervised approach, this method overcomes the weaknesses of signature-based detection and addresses data scarcity in supervised learning. The study shows that modeling normal behaviors through cycle-consistent reconstruction error offers a sensitive and reliable way to detect new threats. Additionally, CycleGAN-based synthetic data augmentation tackles class imbalance, boosting the recall of rare attack types without harming overall network performance. Moving toward hierarchical, edge-aware architectures allows these advanced deep learning models to operate in resource-limited environments, enabling real-time protection for vital infrastructures. Future work should focus on federated learning (FL) to facilitate collaborative model training while maintaining data privacy, and on developing attention-guided techniques to better identify subtle, low-probability anomalies in large volumes of traffic.

References

1. Surepalli, Sirisha & Sameera, Nerella. (Feb 2026). Unsupervised Intrusion Detection System for Zero-Day Attack Detection Using Machine Learning and Deep Learning. 10.1007/978-3-032-14197-2_39.
2. R. S. et al., "Distributed Preprocessing and Adversarial Training for Robust IoT IDS," arXiv:2507.19739v1, 2025. [Online]. Available: https://arxiv.org/html/2507.19739v1
3. Z. Dehghanian, S. Saravani, M. Amirmazlaghani, and M. Rahmati, "Anomaly Detection Using Complete Cycle Consistent Generative Adversarial Network," International Journal of Neural Systems, vol. 35, no. 02, 2550004, 2025. [Online]. Available: https://doi.org/10.1142/S0129065725500042

4. K. Nitrat, N. Suetrong and N. Promsuk, "Zero-Day Attack Detection in IoT Networks Using a Residual Vision Transformer-Based Approach With Zero-Shot Learning," in IEEE Open Journal of the Communications Society, vol. 6, pp. 7405-7423, 2025, doi: 10.1109/OJCOMS.2025.3604826.
5. Ioannou I, Vassiliou V. Generative Adversarial Networks for Energy-Aware IoT Intrusion Detection: Comprehensive Benchmark Analysis of GAN Architectures with Accuracy-per-Joule Evaluation. Sensors (Basel). 2026 Jan 23;26(3):757. doi: 10.3390/s26030757. PMID: 41682273; PMCID: PMC12899382.
6. Fang, M., Wang, Y., Yang, L., Wu, H., Yin, Z., Liu, X., Xie, Z., & Kong, Z. (2024). Reinventing Web Security: An Enhanced Cycle-Consistent Generative Adversarial Network Approach to Intrusion Detection. *Electronics*, *13*(9), 1711. https://doi.org/10.3390/electronics13091711
7. Alshehri, Mohammed & Saidani, Oumaima & Al Malwi, Wajdan & Asiri, Fatima & Latif, Shahid & Khattak, Aizaz & Ahmad, Jawad. (2025). A Hybrid Wasserstein GAN and Autoencoder Model for Robust Intrusion Detection in IoT. Computer Modeling in Engineering & Sciences. 143. 3899-3920. 10.32604/cmes.2025.064874.
8. F. S. Atedjio, J. -P. Lienou, F. F. Nelson, S. S. Shetty and C. A. Kamhoua, "CycleGAN-Gradient Penalty for Enhancing Android Adversarial Malware Detection in Gray Box Setting," in IEEE Access, vol. 12, pp. 162685-162696, 2024, doi: 10.1109/ACCESS.2024.3486734.
9. Sridharan, S., Patil, S., Shobha, T., Pai, P. (2025). Hybrid machine learning–based intrusion detection for zero-day attack prevention in digital education networks. International Journal of Safety and Security Engineering, Vol. 15, No. 8, pp.1703-1713. https://doi.org/10.18280/ijsse.150815
10. Hashim, Khalid & Mohd, Yusnani & Shahbudin, Shahrani. (2025). Mitigating Zero-Day Vulnerabilities in IIoT Systems: Challenges and Advances in AI-Powered Intrusion Detection Systems. Mesopotamian Journal of CyberSecurity. 5. 1184-1198. 10.58496//MJCS/2025/063.
11. Li, E., Shang, Z., Gungor, O., & Rosing, T. (2025). SAFE: Self-Supervised Anomaly Detection Framework for Intrusion Detection. ArXiv. https://arxiv.org/abs/2502.07119
12. Zahra, Fatima & Bostanci, Yavuz & Soyturk, Mujdat. (2024). Unsupervised Machine Learning for Anomaly Detection in Wi-Fi Based IoT Networks. 10.1109/ICCSPA61559.2024.10794232.
13. Dehghanian, Z., Saravani, S., Amirmazlaghani, M., & Rahmati, M. (2023). Spot The Odd One Out: Regularized Complete Cycle Consistent Anomaly Detector GAN. *ArXiv*. https://arxiv.org/abs/2304.07769
14. Jiang, YaPing & Zhang, ZhengHe & Ge, YangTao. (2024). CycleGAN-based intrusion detection data augmentation model. 350. 10.1117/12.3031413.
15. Ioannou I, Vassiliou V. Generative Adversarial Networks for Energy-Aware IoT Intrusion Detection: Comprehensive Benchmark Analysis of GAN Architectures with Accuracy-per-Joule Evaluation. Sensors (Basel). 2026 Jan 23;26(3):757. doi: 10.3390/s26030757. PMID: 41682273; PMCID: PMC12899382.
16. Jamoos, M.; Mora, A.M.; AlKhanafseh, M.; Surakhi, O. A New Data-Balancing Approach Based on Generative Adversarial Network for Network Intrusion Detection System. Electronics 2023, 12, 2851, doi:10.3390/electronics12132851.
17. Allagi, S., Pawan, T., & Leong, W. Y. (2025). Enhanced Intrusion Detection Using Conditional-Tabular-Generative-Adversarial-Network-Augmented Data and a Convolutional Neural Network: A Robust Approach to Addressing Imbalanced Cybersecurity Datasets. *Mathematics*, *13*(12), 1923. https://doi.org/10.3390/math13121923
18. Gao, Ziyuan. (2025). Anomaly Detection for Enhancing IoT Device Security Using Machine Learning: A Comparative Study of Four Lightweight Models Based on the IoT-23 Dataset. ITM Web of Conferences. 80. 01027. 10.1051/itmconf/20258001027.
19. Wakili, A., & Bakkali, S. (Mar 2026). ZeroDefense: An adaptive hybrid fusion-based intrusion detection system for zero-day threat detection in IoT networks. Journal of Electronic Science and Technology, 24, Article 100345. https://doi.org/10.1016/j.jnlest.2026.100345
20. Lenort, M., Szygula, J., Marek, D., Marszalek, K., & Domanski, A. (2025). Comparative analysis of generator architectures in CycleGAN for image style transfer. 2025 IEEE International Conference on Big Data (BigData), 3979–3987. https://doi.org/10.1109/BigData66926.2025.11401638
21. D. P. Kavadi et al., "Design of an Integrated Model Combining CycleGAN, PPO, and Vision Transformer for Adaptive Scene Rendering in the Metaverse," in IEEE Access, vol. 13, pp. 21117-21138, 2025, doi: 10.1109/ACCESS.2025.3532327.