Novel Zero-Day Zero-Touch Intrusion Detection System Using Generative-AI and LLM

Md Mahinur Alam, Kanita Jerin Tanha, Md Raihan Subhan, and Taesoo Jun
Pervasive Intelligent Computing Lab, Department of IT Convergence Engineering,
Kumoh National Institute of Technology, Gumi 39177, South Korea
(mahinuralam213, kanitajerin17, raihan, taesoo.jun)@kumoh.ac.kr

***Abstract*—As communication technologies advance, the transmission of diverse data across distributed networks has become more common, increasing the vulnerability of network systems to security threats. Current Intrusion Detection Systems (IDS) struggle to detect zero-day attacks, as they depend heavily on traditional attack patterns and require manual updates and interventions, leading to operational overhead. To address these challenges, this paper introduces a novel zero-day zero-touch IDS that leverages Generative-AI to handle data imbalance and enhance the performance of existing systems. The proposed system utilizes Large Language Models (LLM) with context-aware capabilities to analyze network traffic in real-time, adapting to evolving threats and improving zero-day attack detection. We evaluated the system using benchmark datasets like CICIDS 2017 and BoT-IoT, and the results show that it outperforms existing AI-based IDS models in detecting zero-day attacks, achieving 98% accuracy and a 99% F1-score.**

***Index Terms*—BERT, Generative Adversarial Network (GAN), Intrusion detection, LLM, zero-day, zero-touch.**

With the advent of 5G technology, network systems have evolved into distributed architectures that process vast amounts of data from sensors, computing devices, and Internet of Things (IoT) devices. While this expansion improves connectivity, it also increases vulnerabilities by expanding the attack surface. This, combined with increasingly sophisticated cyber threats, has led to more frequent and severe cyberattacks across network infrastructures [1]. One such example is the cyberattack on SK Telecom on April 18, 2025, which compromised the personal data of over 23 million customers, nearly half of South Korea's population. This incident underscores the vulnerability of large-scale IoT networks and the need for robust Intrusion Detection Systems (IDS) capable of preventing zero-day attacks.

The growing importance of cybersecurity has driven extensive research efforts to counter network threats [2]. Although the majority of network traffic is benign, malicious activities, though rare, have the potential to cause significant service disruptions. Traditional IDS, especially those based on AI, struggle to detect zero-day attacks due to inherent data imbalance in network traffic. These systems often fail to learn the characteristics of rare yet critical threats, leaving networks vulnerable to evolving cyberattacks. Generative Adversarial Networks (GANs) have emerged as an effective solution to address this imbalance by generating synthetic data to improve detection performance [3]. Recent advances in Large Language Models (LLMs), such as BERT, have shown impressive capabilities in handling large, unlabeled datasets and tasks requiring nuanced understanding and generalization. This paper explores applying these advances to IoT security [4]. The contributions of this study are as follows:

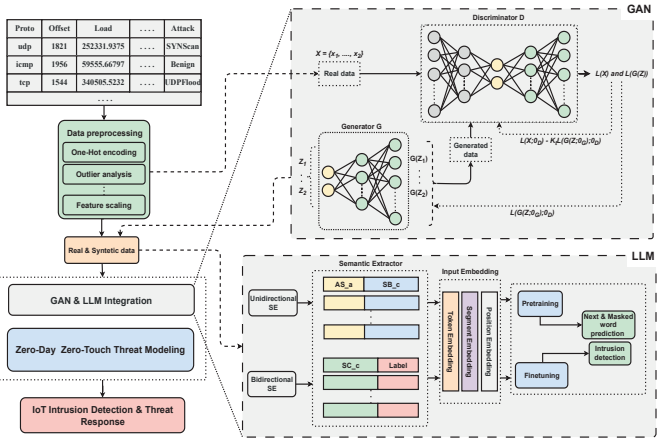

Fig. 1: Proposed zero-day zero-touch intrusion detection system.

The proposed system includes (1) a Zero-Day Zero-Touch Intrusion Detection System, which autonomously detects unseen attacks using Generative AI and LLMs, enabling real-time response with limited labeled data. (2) GAN-based Synthetic Traffic Generation to address rare zero-day attacks and improve detection beyond traditional NIDS. (3) BERT-ID Hybrid Model, leveraging Byte Sentences (BS) and a Semantic Extractor (SE) for preprocessing, with pre-training tasks like Masked Byte Word Model (MBWM) and Next Byte Sentence Prediction (NBSP) to improve detection accuracy.

I. Proposed Zero-Day Zero-Touch IDS

The proposed intrusion detection system combines Generative-AI and LLM to detect zero-day attacks in IoT networks. It preprocesses network traffic through one-hot encoding, outlier analysis, and feature scaling. A GAN generates synthetic attack data while BERT analyzes contextual semantics of network traffic, enabling autonomous detection of unseen threats, Fig. 1 demonstrates the overall system.

In the GAN system, the generator minimizes $\mathcal{L}_G = \mathbb{E}_{Z\sim p_Z(z)}\left[\log(1 - D(G(Z)))\right]$ while the

TABLE I: Comparison of intrusion detection performance of the proposed system with existing methods.

Dataset	CIC-IDS 2017								BoT-IoT							
Model	Without GAN				With GAN				Without GAN				Without GAN			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
BLM-Chain [4]	85%	83%	84%	85%	88%	88%	87%	86%	84%	82%	84%	84%	87%	87%	86%	87%
ByteSGAN [5]	89%	89%	91%	91%	92%	93%	93%	92%	88%	87%	89%	90%	91%	92%	92%	92%
VAE [2]	87%	86%	86%	87%	89%	88%	89%	89%	86%	85%	85%	86%	90%	89%	88%	90%
Proposed	**92%**	**92%**	**91%**	**91%**	**98%**	**99%**	**98%**	**99%**	**93%**	**92%**	**92%**	**92%**	**97%**	**98%**	**98%**	**98%**

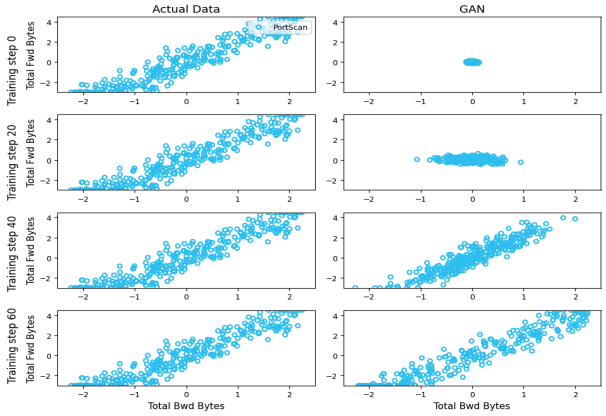

Fig. 2: Synthetic network traffic data generation of PortScan class.

discriminator maximizes $\mathcal{L}_D = \mathbb{E}_{X\sim p_{\text{real}}(X)}[\log D(X)] + \mathbb{E}_{Z\sim p_Z(z)}[\log(1-D(G(Z)))]$. This adversarial process creates a minimax game where the generator constantly improves at creating realistic attack patterns while the discriminator becomes more adept at distinguishing real from synthetic data. Through iterative training, the GAN converges to produce high-fidelity synthetic attack vectors that closely resemble potential zero-day threats, effectively expanding the training dataset beyond known attack signatures.

BERT processes network traffic as Byte Sentences, using masked byte prediction with loss function $\mathcal{L}_{\text{MBWM}} = -\sum_i \log P(\hat{X}_i|X_{\text{masked}})$ and next byte prediction with $\mathcal{L}_{\text{NBSP}} = -\sum_i \log P(\text{Next Byte}|\text{Previous Byte})$. The MBWM task randomly masks portions of network packets, forcing BERT to reconstruct the original data based on surrounding context, thus learning the inherent structure of legitimate traffic versus malicious patterns. The NBSP task enhances BERT's ability to understand sequential dependencies in network flows, critical for detecting anomalous packet sequences indicative of attacks. Fine-tuning uses $\mathcal{L}_{\text{fine-tuning}} = -\sum_i Y_{\text{true}} \log(Y_{\text{pred}})$, which optimizes the model's classification.

II. Experimental Results and Discussion

The CICIDS 2017 and BoT-IoT datasets, enhanced with 30% synthetic data, were used for evaluation. The GAN generates realistic attack data, improving over time, as shown in Fig. 2, simulating rare attack patterns to enhance detection in imbalanced datasets.

The proposed system outperforms BLM-Chain, ByteSGAN, and VAE, achieving 98% accuracy, 99% precision, and 99% F1-score on both datasets (Table I). In comparison, ByteSGAN achieved 93% accuracy. The system also excels in classifying attack types (e.g., Brute Force, DDoS, Infiltration) with higher precision and recall, as seen in Fig. 3, showing its superior generalization to evolving threats.

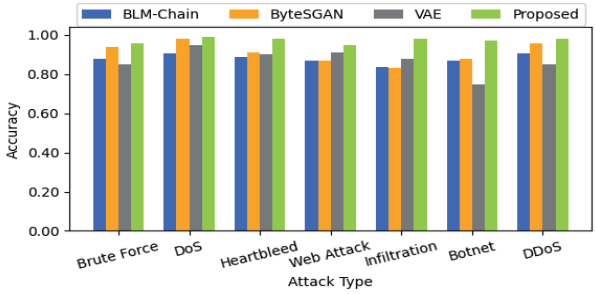

Fig. 3: Classification generalization of CIC-IDS 2017 dataset.

III. Conclusion

This paper introduced a Zero-Day Zero-Touch Intrusion Detection System combining Generative AI for synthetic data generation and LLMs for contextual analysis to detect evolving cyber threats in IoT networks. The system outperformed existing systems with 98% accuracy, 99% precision, recall, and F1-score, demonstrating the effectiveness of GAN-based data generation and LLM-based analysis in overcoming data imbalance and enhancing real-time intrusion detection.

Acknowledgment

This research was funded by the Innovative Human Resource Development for Local Intellectualization Program (IITP-2025-RS-2020-II201612, 34%) through IITP under MSIT, the Basic Science Research Program (2018R1A6A1A03024003, 33%) through NRF, and the Information Technology Research Center (ITRC) Program (IITP-2025-RS-2024-00438430, 33%) funded by MSIT through IITP.

References

[1] O. H. Abdulganiyu, T. Ait Tchakoucht, and Y. K. Saheed, "A systematic literature review for network intrusion detection system (ids)," *International journal of information security*, vol. 22, no. 5, pp. 1125–1162, 2023.

[2] C. Liu, R. Antypenko, I. Sushko, and O. Zakharchenko, "Intrusion detection system after data augmentation schemes based on the vae and cvae," *IEEE Transactions on Reliability*, vol. 71, no. 2, pp. 1000–1010, 2022.

[3] A. Majeed and S. O. Hwang, "Moving conditional gan close to data: Synthetic tabular data generation and its experimental evaluation," *IEEE Transactions on Big Data*, 2024.

[4] M. Golam, M. M. Alam, D.-S. Kim, and J.-M. Lee, "Blm-chain: Ai-driven blockchain for uav threat resistance in iobt," in *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2024, pp. 1609–1613.

[5] P. Wang, Z. Wang, F. Ye, and X. Chen, "Bytesgan: A semi-supervised generative adversarial network for encrypted traffic classification in sdn edge gateway," *Computer Networks*, vol. 200, p. 108535, 2021.