AI-BASED INTRUSION DETECTION IN IOT SYSTEMS USING BERT

Patil Sakshi Chandrakant
Dept. Computer Science And Engineering (AI)
D.K.T.E. Society's Textile and Engineering Institute, Ichalkaranji, India
sakshipatil20102004@gmail.com

Musale Aarya Shantilal
Dept. Computer Science And Engineering (AI)
D.K.T.E. Society's Textile and Engineering Institute, Ichalkaranji, India
aaryamusale22@gmail.com

Patil Siddhi Ajit
Dept. Computer Science And Engineering (AI)
D.K.T.E. Society's Textile and Engineering Institute, Ichalkaranji, India
siddhip859@gmail.com

Patil Shivani Nagesh
Dept. Computer Science And Engineering (AI)
D.K.T.E. Society's Textile and Engineering Institute, Ichalkaranji, India
shivaninageshpatil@gmail.com

Abstract— The rapid growth of Internet of Things (IoT) devices has significantly increased the vulnerability of modern networks to cyber threats. Conventional Intrusion Detection Systems (IDS) are often inadequate in handling dynamic and complex attack patterns. This paper presents a comprehensive review of recent advancements in applying Natural Language Processing (NLP) techniques, particularly transformer-based Large Language Models such as BERT, GPT-2, and BART, for intrusion detection. By representing network traffic as sequential data similar to natural language, these models are capable of learning deep contextual relationships and temporal patterns. Various architectures, including lightweight models for IoT environments, predictive security frameworks, and meta-learning approaches for zero-day attack detection, are analyzed. The study highlights that LLM-based approaches outperform traditional machine learning and deep learning methods in terms of accuracy, adaptability, and robustness.

Keywords- Intrusion Detection System (IDS), BERT, BART, Large Language Models (LLMs), IoT Security, Transformers, Deep Learning.

I. *INTRODUCTION*

The rapid advancement of digital technologies has led to a significant increase in cybersecurity threats [3], with attackers continuously evolving their techniques to bypass conventional security mechanisms. Traditional Intrusion Detection Systems (IDS) primarily depend on predefined rules and manual feature engineering, which limits their ability to detect newly emerging and sophisticated attack patterns. As a result, these systems often struggle to operate effectively in dynamic and complex environments such as IoT networks.

Recent research has shifted towards utilizing the capabilities of pre-trained Large Language Models (LLMs) based on Transformer architectures [1]. These models, originally developed for Natural Language Processing (NLP), are highly effective in capturing contextual relationships within sequential data. This makes them well-suited for analyzing network traffic, which can be represented as structured sequences. In this review, we examine modern approaches that employ LLM-based models such as SecurityBERT, BARTPredict, and Siamese-based architectures to enhance intrusion detection performance and enable proactive cyber threat prediction.

I. LITERATURE REVIEW

The shift from traditional machine learning (ML) to deep learning (DL) and eventually to LLM-based IDS represents a significant evolution in cybersecurity.

BERT-Based Models: Bidirectional Encoder Representations from Transformers (BERT) has been extensively applied in various cybersecurity applications [1], including malware detection in Android systems, anomaly detection in log data, and security analysis in automotive networks. Its bidirectional architecture allows it to capture deeper contextual relationships within data sequences, making it more effective than traditional unidirectional models.

Proactive vs. Reactive Systems: Conventional IDS are primarily reactive in nature, meaning they detect threats only after they occur [2]. However, recent research focuses on proactive approaches that aim to predict potential cyberattacks before they happen. Generative models such as BART are being utilized to forecast malicious activities, enabling early threat mitigation.

IoT-Specific Challenges: IoT environments introduce unique challenges such as limited computational resources, high data variability, and the increasing use of encrypted traffic [1]. To address these issues, recent studies propose lightweight architectures and privacy-preserving techniques that allow efficient intrusion detection without requiring access to raw data payloads.

II. METHODS AND MODEL ARCHITECTURES

A number of advanced architectures have been proposed to integrate Large Language Models into intrusion detection systems, each focusing on improving detection accuracy and computational efficiency.

SecurityBERT **[1]**: SecurityBERT is specifically designed to operate in resource-constrained IoT environments where computational efficiency and privacy are critical concerns. It is based on a compact version of the BERT architecture, reducing the number of parameters while maintaining performance.

One of the key contributions of this model is its ability to transform raw network traffic data into structured representations using a privacy-preserving encoding technique. Instead of directly processing sensitive data, the model converts numerical features into a fixed-length encoded format. This approach ensures that confidential information is not exposed during processing while still allowing the model to learn meaningful patterns.

Additionally, SecurityBERT is optimized for deployment on edge devices due to its smaller size and faster inference time. This makes it highly suitable for real-time intrusion detection in IoT systems.

BARTPredict **[2]**: BARTPredict introduces a proactive approach to intrusion detection by combining prediction and evaluation mechanisms. Unlike traditional systems that detect attacks after they occur, this framework attempts to anticipate potential threats in advance.

In this model, a fine-tuned BART (Bidirectional and Auto-Regressive Transformer) model is used to predict future network traffic sequences based on historical data. These predicted sequences are then analyzed using a BERT-based classifier to determine whether they indicate malicious behavior.

This two-step process enables early identification of potential threats, allowing systems to take preventive measures before an actual attack occurs. As a result, BARTPredict enhances the overall security of IoT networks by shifting from reactive detection to proactive defense.

SiamXBERT **[4]**: SiamXBERT is designed to address one of the most critical challenges in cybersecurity: detecting unknown or zero-day attacks. Traditional models struggle with such attacks because they rely heavily on previously seen patterns.

To overcome this limitation, SiamXBERT utilizes a Siamese network combined with meta-learning techniques. It processes both flow-level and packet-level data, enabling it to capture multiple perspectives of network traffic.

The model learns similarity relationships between different traffic patterns, which allows it to identify anomalies even when encountering new types of attacks. Furthermore, it requires only a small amount of labeled data for training, making it highly efficient in real-world scenarios where labeled datasets are limited.

BERT-based Network for Intrusion Detection System **[3]**: This approach treats network traffic data as a form of language, enabling the application of Natural Language Processing techniques to intrusion detection. Instead of analyzing raw numerical data directly, network features such as IP addresses, port numbers, and protocol types are converted into structured sequences.

A pre-trained BERT model is then used to extract high-level features from these sequences. Due to its bidirectional architecture, BERT is capable of

understanding complex relationships and dependencies within the data.

These extracted features are further processed using sequence-based models, such as recurrent neural networks, to classify network activity as normal or malicious. This combination of NLP and deep learning techniques enhances the model's ability to detect complex attack patterns.

A Deep Learning based IDS using Transformers **[5]**: This model combines multiple transformer architectures, including BERT and GPT-2, to improve the overall performance of intrusion detection systems. The integration of these models allows the system to handle diverse and complex network scenarios more effectively.

One of the key objectives of this approach is to reduce computational complexity while maintaining high accuracy. This is achieved by selecting only the most relevant features from the dataset, thereby minimizing unnecessary processing.

The model is typically evaluated using benchmark datasets such as CICIDS2017 and is designed for binary classification tasks, where it distinguishes between normal and malicious network traffic. This approach ensures efficient detection while keeping computational costs manageable.

COMPARATIVE ANALYSIS

Experimental evaluations across various datasets demonstrate the superiority of LLM-based approaches.

Model	Dataset	Primary Metric	Result
SecurityBERT	Edge-IIoTset	Accuracy	98.2%
BARTPredict	CICIoT2023	Accuracy	98.0%
SiamXBERT	CICIoT2023	Unknown F1-Score	78.8% improvement over baselines
BERT-GRU	Public Datasets	Detection Performance	Outperforms ML/DL methods

Table 1: Experimental Evaluation

Key Findings:

Data Efficiency: SiamXBERT **[4]** demonstrates extreme data efficiency, outperforming baselines while being trained on only 100 samples per class compared to thousands for traditional models.

Detection vs. Prediction: While BERT **[2]** models excel at identifying active attacks, BART-based frameworks offer the added benefit of preemptive mitigation.

Computational Efficiency: SecurityBERT **[1]** achieves an inference time of less than 0.15 seconds on a standard CPU, proving that LLMs can be made practical for real-time edge deployment.

III. DISCUSSION

The integration of Large Language Models into intrusion detection systems represents a significant shift in how cybersecurity problems are addressed. Unlike traditional approaches that rely on predefined rules or statistical methods, LLM-based models analyze network traffic as structured sequential data, enabling a deeper understanding of complex patterns and behaviors [1].

One of the major advantages of these models is their ability to perform effectively even in environments where data is encrypted. By focusing on traffic patterns rather than payload content, models such as SiamXBERT [4] demonstrate that high detection accuracy can be achieved without compromising privacy.

Additionally, the use of transformer architectures improves the ability to capture long-range dependencies and contextual relationships within network data. This enhances the system's capability to detect sophisticated and previously unseen attack patterns. The incorporation of meta-learning techniques further strengthens adaptability, allowing models to generalize across different datasets and scenarios [2].

However, despite these advantages, certain challenges remain. These include high computational requirements for large-scale models, potential performance variations across different datasets, and the need for standardized methods to represent network traffic as input sequences. Addressing these challenges is essential for improving real-world deployment of LLM-based intrusion detection systems.

These observations highlight that LLM-based intrusion detection systems are not only more intelligent but also more adaptable to evolving cybersecurity threats, making them a promising direction for future research.

IV. CONCLUSION

In conclusion, the adoption of large language models in intrusion detection systems represents a significant advancement in cybersecurity, particularly for IoT environments. These models leverage their ability to understand contextual and sequential patterns, resulting in improved detection accuracy and efficiency. Techniques such as privacy-preserving encoding, meta-learning, and predictive modeling further enhance their applicability in real-world scenarios. Despite existing challenges like cross- dataset generalization and standardization issues, LLM-based IDS solutions demonstrate strong potential for future research and deployment. Continued improvements in model optimization and real-time adaptability will further strengthen their role in next-generation security systems.

V. REFERENCES

[1] Ferrag et al., "Revolutionizing Cyber Threat Detection with Large Language Models: A privacy-preserving BERT-based Lightweight Model for IoT/IIoT Devices," 2024.

[2] Diaf et al., "BARTPredict: Empowering IoT Security with LLM-Driven Cyber Threat Prediction," 2025.

[3] Yang and Peng, "BERT-based network for intrusion detection system," 2025.

[4] Ali et al., "Unknown Attack Detection in IoT Networks using Large Language Models: A Robust, Data-efficient Approach," 2026.

[5] Athul K and Anita John, “A Deep Learning based Intrusion Detection System using Transformers.”