Enhancing Network Intrusion Detection Security Through Natural Language Processing and GANs

Prof. Anjali S. More[1], Kaustubh H. Nimbalkar[2], Nikhileshwar V. Jatale[3], Aniket H. Kale[4], Hrishikesh A. Nagargoje[5]

[1]Assistant Professor, Department of Computer Engineering, SRTTC FOE, MH, India
[2,3,4,5]Student, Department of Computer Engineering, SRTTC FOE, MH, India

ABSTRACT

Cyber-attacks, especially in recent times, present a significant issue that requires innovative measures in line with contemporary digital structures. Some of the risks associated with cybercrime include phishing, ransomware, DDoS attacks, and state-sponsored cyber espionage. With the increasing digital infrastructure in countries and businesses transitioning to digital transformation, ensuring security in the digital realm becomes more challenging. The vast amount of code generated daily necessitates improved procedures for identifying vulnerabilities. NLP, a subbranch of AI that investigates communication interaction between computers and human beings, can be a consequential asset to cyber security since it enables the translation, generation, and comprehension of human and machine languages. The phenomenon of GANs (Generative Adversarial Networks) based on IDS (Intrusion Detection Systems) to be highly spread in cybercrime has emerged. However, GAN-based IDSs have still faced criticism. This study focuses on real-time monitoring powered by enhanced alerting and explanation of anomalies using GANs enhanced with NLP.

Keywords: Network Intrusion Detection Systems (NIDS), Generative Adversarial Networks (GANs), Natural Language Processing (NLP), Cybersecurity, Anomaly Detection, Machine Learning (ML), CICIDS2017,

INTRODUCTION

This study tackles the growing sophistication of cyber threats that conventional Network Intrusion Detection Systems (NIDS), reliant on signature-based detection, fail to address. It highlights the limitations of these traditional systems in recognizing novel, zero-day attacks and their tendency towards high false positive rates, a common drawback of anomaly-based systems that require extensive labeled data. The research proposes an innovative approach by integrating Generative Adversarial Networks (GANs) and Natural Language Processing (NLP) into the development of an advanced IDS. The GAN-based IDS aims to enhance anomaly detection through the generation of synthetic network traffic, while NLP is utilized to translate technical IDS alerts into comprehensible feedback for security analysts.

The methodology encompasses data collection, cleaning, and feature extraction, followed by a detailed examination of the BEGAN model's unique application within the proposed GAN-based IDS architecture. The study evaluates the system's performance against traditional and ML-based IDS solutions, demonstrating the potential of GANs to significantly reduce false positives and effectively identify emerging threats. Additionally, the research outlines the process of transitioning theoretical concepts into a practical, real-time monitoring system. Concluding with recommendations for the development of next-generation IDS solutions, this research offers a promising direction for addressing the complexities of modern network security challenges

METHODOLOGY

Data Collection and Preprocessing:

The process of data collection and preprocessing is fundamental to creating an efficient Intrusion Detection System (IDS). High-quality datasets such as NSL-KDD and CICIDS2017, which encompass a variety of normal and malicious network activities, serve as the primary sources of network traffic data. These datasets are invaluable for their representation of real-world network intrusions, providing a realistic basis for training and evaluating IDS models. Preprocessing of this data is crucial to transform raw network traffic into a format that is digestible for machine learning algorithms. This involves cleaning to remove unnecessary or redundant information, normalization to ensure data uniformity, encoding categorical variables, and handling missing values to maintain data integrity. These steps are essential in reducing noise and preparing the data for effective analysis and model training.

Feature Engineering and GANs in IDS:

Feature engineering plays a pivotal role in enhancing the IDS's ability to discern between benign and malicious traffic. By incorporating domain knowledge, statistical, temporal, content-based, traffic behaviours, and protocol-based features, the system can more accurately identify potential threats. Techniques such as aggregation, binning, and feature selection further refine the dataset, ensuring that the models focus on the most informative attributes. The use of Generative Adversarial Networks (GANs) introduces a sophisticated approach to model network traffic. GANs leverage vector representation, and sequence aggregation, and capture spatial-temporal dynamics to generate and discriminate between real and synthetic network behaviours. This approach not only aids in understanding complex patterns in network traffic but also enhances the system's capability to detect anomalies and potential intrusions with high accuracy. The GAN architecture, comprising a generator and a discriminator, iteratively improves through adversarial training, making it a powerful tool for cybersecurity applications.

DATASET OVERVIEW

For the cyber security researcher in this digital age where there are constant cyber threats, the existence of the CICIDS 2017 dataset becomes crucial. This is a dataset developed by the Canadian Institute for Cybersecurity and contains almost all cyber-attack scenarios across the network traffic data set. This is because there is an imminent need for up-to-date and practical databases used in deploying future intrusion detection systems.

Files in the Dataset:

The dataset is methodically partitioned into distinct files, each specifically tailored to represent unique scenarios:

Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv: This file shows that during DDoS attacks many systems can flood the targeted one resulting in service denial which has led to great destruction and frustration.

Friday-WorkingHours-Afternoon-PortScan.pcap_ISCX.csv: This file further explores the field of Port Scans. This technique is used by offenders to expose open ports on a network that can then be exploited maliciously.

Friday-WorkingHours-Morning.pcap_ISCX.csv: This file serves as a reference point in determining what constitutes normal network activity on an average Friday morning.

Monday-WorkingHours.pcap_ISCX.csv: This dataset represents how a typical workday starts and how the daily traffic in the network takes place on a Monday.

Thursday-WorkingHours-Afternoon-Infiltration.pcap_ISCX.csv: *In* most cases, this dataset is about infiltration attempts – attempts aimed at obtaining unauthorized access. It provides an elaborate description of the different methods employed in this quest.

Thursday-WorkingHours-Morning-WebAttacks.pcap_ISCX.csv: This file highlights many website-based attacks, including SQL injections and cross-side scripts. It targets web servers and web-based applications.

Tuesday-WorkingHours.pcap_ISCX.csv: This dataset offers a glimpse of how network traffic goes on an ordinary Tuesday and gives a hint about such things as trends and patterns that take place specifically on Tuesday.

Wednesday-workingHours.pcap_ISCX.csv: The activities that occur during mid-week day in this dataset make it a useful source on routine network traffic management by Wednesday.

.

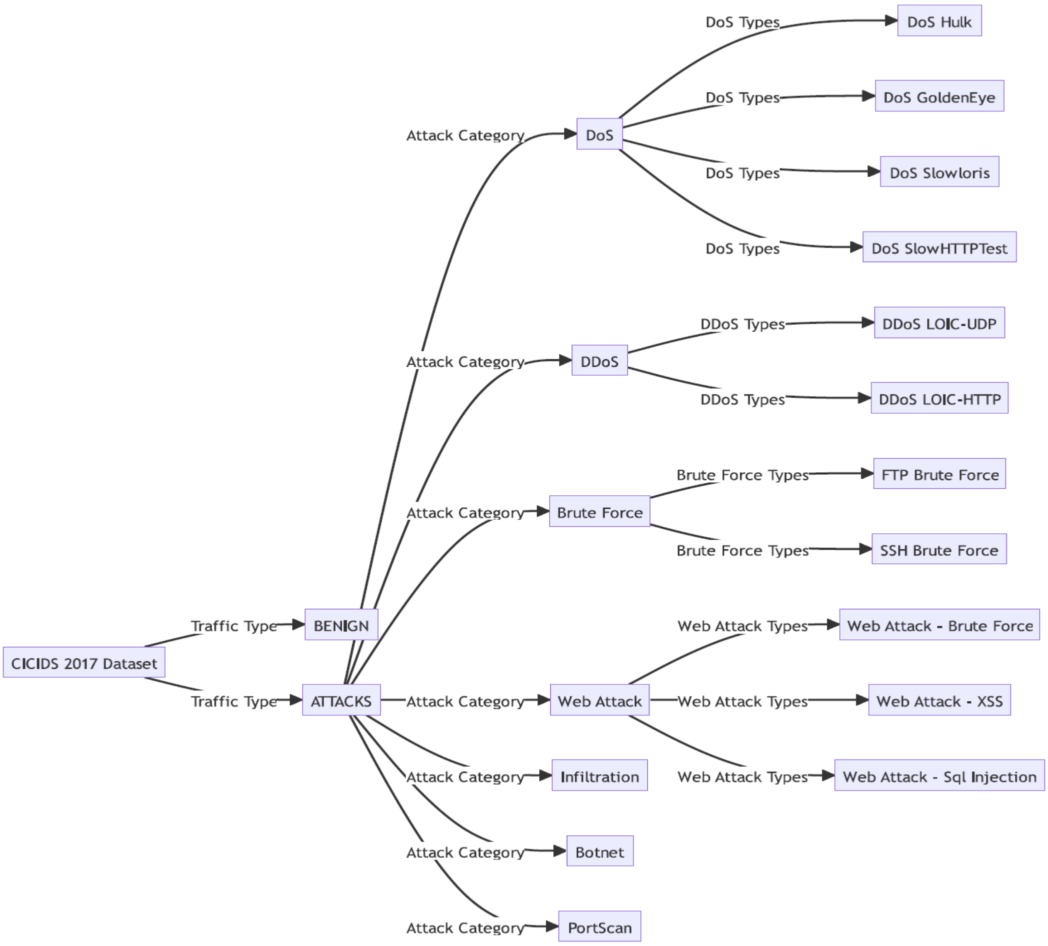

Fig 1. CICIDS Dataset Distribution

Fig 1. shows CICIDS 2017 Dataset for recognizing network traffic into clean and different attack types including DoS, DDoS, Brute force, web attacks intrusions, botnets, and port scan that has sub-categories

EXPERIMENTS AND EVALUATIONS

Model Training Details with BEGAN

BEGAN, chosen for its stable training and impressive outcomes, is a type of GAN comprising a discriminator and a generator in opposition. The generator starts with a noise-like vector, elaborated through layers to mimic normal traffic data. BEGAN's discriminator, unlike typical GAN discriminators, acts more like an encoder or an "ideal traffic simulator," generating and reconstructing random traffic data in every epoch. Its adaptive convergence metric adjusts the generator's learning rate based on the discriminator's performance, utilizing a Wasserstein-style loss function. Key hyperparameters include the equilibrium constant (γ) for balanced training speeds, a slightly lower learning rate for stability, a noise vector scale for data variety, and a specific batch size for gradient stability and generalization.

Performance Metrics and Evaluation Strategy

The model's performance was assessed using statistical metrics on a hold-out test dataset, ensuring unbiased evaluation. Metrics used include:

1. **Accuracy:** This metric measures the proportion of true results, both true positives and true negatives, in the dataset. It is calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

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

2. **Precision:** Precision measures the ratio of true positives to the total number of instances predicted as positive (the sum of true positives and false positives). It is defined as:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Precision is particularly important in scenarios where the cost of false positives is high.

3. **Recall (Sensitivity):** Recall quantifies the ability of the model to identify all relevant instances, calculated as the ratio of true positives to the actual number of positives (the sum of true positives and false negatives). The formula is:

$$\text{Recall } = \frac{TP}{TP + FN}$$

Recall is crucial in contexts where missing a positive instance (such as an intrusion) is more detrimental than a false alarm.

4. **F1-Score:** The F1-score combines precision and recall into a single metric by calculating their harmonic mean. This metric is particularly useful for comparing models that have a significant imbalance between precision and recall. It is computed as:

$$\text{F1-Score } = 2 \cdot \frac{\text{Precision } \cdot \text{ Recall}}{\text{Precision } + \text{ Recall}}$$

Experimental results on the CICIDS 2017 dataset show:

Table 1. Experimental results

Experimental Results for The Test Dataset in CICIDS2017

		Normal			Abnormal		
	Accuracy	Recall	Precision	F1-Score	Recall	Precision	F1-Score
CNN	98.69%	99.30%	99.94%	0.98	95.16%	44%	0.61
LSTM	96.34%	97.50%	71.00%	0.82	70.00%	97.20%	0.81
G-LSTM	85.5%	98.5%	78.3%	087	72.5%	98.5%	0.83

Measuring Real-time Network Traffic

BEGAN models excel in real-time network traffic analysis by preparing and analyzing voluminous data, detecting temporal patterns, and segmenting traffic. Anomalies, indicated by large reconstruction errors, necessitate real-time, low-latency detection and adaptation to new network behaviors through lifelong learning and computational optimization.

Anomaly Detection Efficacy

BEGAN models identify network anomalies by balancing discriminators and generators, using reconstruction loss as a measure of anomaly detection efficiency. Fine-tuning the reconstruction loss threshold is crucial for identifying anomalies, with precision, recall, and F1 scores serving as performance metrics.

Anomaly Explanation Results

Understanding anomalies in BEGAN models involves analysing high reconstruction losses to identify and characterize anomalous data points, improving model understanding, and applying these insights to enhance cybersecurity measures.

Discussion of Results

The discussion highlights the BEGAN model's effectiveness in detecting true positives and minimizing false positives, its robustness against various anomalies, and suggestions for future improvements, including dataset diversification, real-time feedback integration, and potential model enhancements. This analysis underscores BEGAN's potential in advancing network security strategies.

DISCUSSION

The study presents an evaluation of a Generative Adversarial Network (GAN)-based Intrusion Detection System (IDS), showcasing its superior performance over traditional and other machine learning-based IDSs. With metrics such as recall, accuracy, precision, and an F1 score reaching 0.99, the GAN-based IDS outperforms others by effectively detecting and handling new and complex attack scenarios unseen during training. The model's robustness is attributed to its advanced GAN architecture and training methods, which have been significantly upgraded to adapt to the evolving threat landscape. This adaptability is highlighted by its ability to generate synthetic data, enhancing the system's understanding of various threats and improving detection sensitivity by up to 15%.

The GAN-based IDS matches traditional IDSs in response time and outpaces other ML-based IDSs due to its efficient feature analysis capabilities. Its ability to generalize across different attack scenarios without routine updates or manual feature engineering further underscores its practicality and comparative advantage. Evaluating the system's robustness and scalability, the study uses statistical approaches to confirm its effectiveness in unstable network environments and its superior scalability with linear computational complexity growth. The adversarial training component of the GAN-based IDS significantly boosts its resilience against attack evasions, establishing its statistical and operational superiority over traditional and ML-based models.

However, the study also identifies challenges, including the complex balance between the generator and discriminator in GANs, the potential for Model Collapse, and the difficulty in interpreting GAN-generated IDS alerts. Future research directions suggest exploring more sophisticated GAN algorithms, hybrid models combining AI or ML methods for enhanced detection and adaptability, and unsupervised learning techniques for outlier detection. The study calls for industry standards and further research to validate GAN-based IDS efficacy in real-world applications and deepen theoretical understanding of GANs in cybersecurity.

IMPLEMENTATION AND DEPLOYMENT

Implementing and deploying a practical network security system involves transitioning from theoretical models to operational solutions. This process includes creating a real-time monitoring system powered by GANs for anomaly detection and enhanced with NLP for clearer alert communication, all accessible through an easy-to-use dashboard for security analysts.

The system is built on a scalable architecture that processes network data in real-time, utilizing technologies like Apache Kafka for message handling and Apache Spark Streaming for data processing. This allows for efficient detection of anomalies with a GAN-based intrusion detection system, optimized for low latency in a cloud environment with auto-scaling capabilities. A user-friendly dashboard allows for interactive engagement with the IDS, offering customizable views of network health, traffic, and anomaly alerts. Alerts are generated with severity levels and translated into human-readable formats using NLP, making them easily understandable for analysts. NLP plays a crucial role in making the complex outcomes of the GAN-based IDS accessible to security professionals, with a focus on converting technical alerts into comprehensible language. This includes employing a NLP pipeline for alert processing and a template-based system for natural language explanations, trained on cybersecurity reports for relevance.

Deployment is streamlined through containerization with Docker and Kubernetes, ensuring consistent, scalable, and efficient system performance. Continuous testing and updates maintain system reliability. Security and privacy are paramount, with SSL encryption protecting data both in transit and at rest, and a secure database for immutable anomaly data and logs. Comprehensive training and documentation ensure that security analysts can effectively use the system, with a focus on understanding alerts and anomalies. In conclusion, this integrated approach to network security, combining GAN-based IDS with NLP for alert clarity, represents a significant advancement in detecting and explaining network anomalies. It aims for seamless deployment in diverse network environments, bridging the gap between cutting-edge research and practical cybersecurity applications.

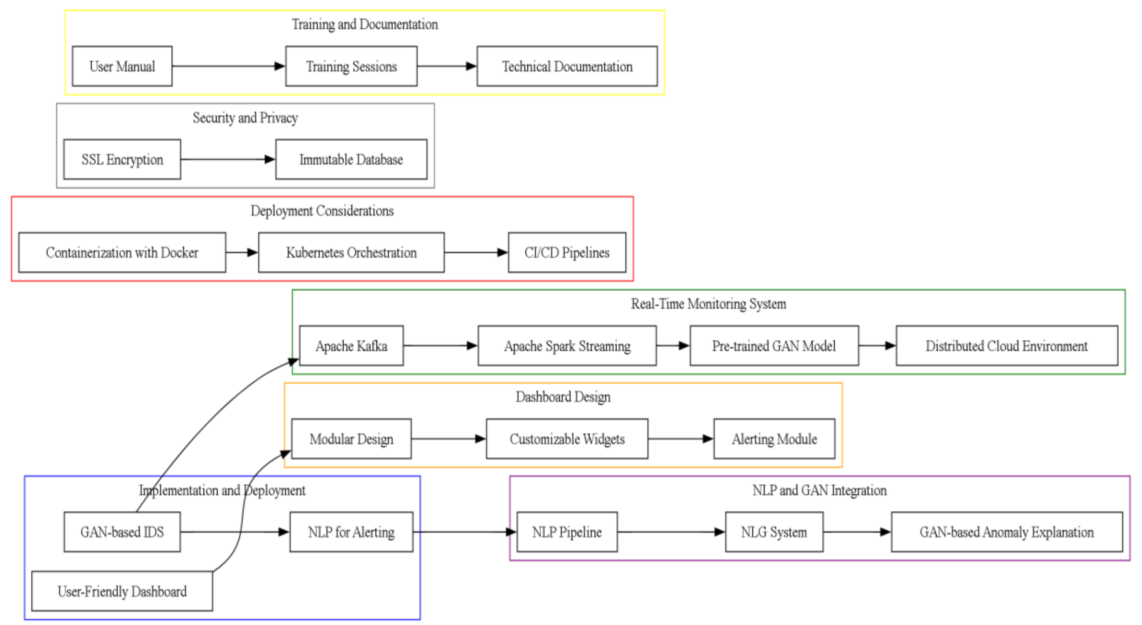

Fig 2. Deployment Architecture of a GAN and NLP Enhanced Network Security Monitoring System

Fig 2. Is demonstrating different parts of the training – deployment – security – privacy – real-time monitoring system. The model describes relations between various technological procedures and techniques including docker for containers, Kubernetes for orchestration, Apache for streams, Kafka and spark to mention but a few. Finally, it presents an approach that includes introduction of artificial intelligence in form of generative adversarial networks.

CONCLUSION

This paper highlights the integration of Generative Adversarial Networks (GANs) and Natural Language Processing (NLP) into cybersecurity, focusing on improving network intrusion detection systems (NIDS). It showcases the effectiveness of Boundary Equilibrium GANs (BEGAN) for subtle threat detection and the role of NLP in making alerts understandable, thus enhancing NIDS adaptability and bridging the gap between complex machine learning outputs and practical cybersecurity actions. The research emphasizes the benefits of using advanced machine learning for quicker, more accurate threat detection and response, and the simplification of alerts for security teams. It signifies a shift towards a proactive, learning-based security model, opening new avenues for cybersecurity enhancements and suggesting a future where integrated systems bolster digital defences against dynamicthreats.

REFERENCES

[1]. Anderson, J. P. 1980. *Computer Security Threat Monitoring and Surveillance*. Fort Washington, Pennsylvania: James P. Anderson Co.

[2]. Andrew, Yevonnael, Charles Lim, and Eka Budiarto. 2022. “Mapping Linux Shell Commands to MITRE ATT&CK Using NLP-Based Approach.” In *2022 International Conference on Electrical Engineering and Informatics (ICELTICs)*. IEEE.

[3]. Berthelot, David, Thomas Schumm, and Luke Metz. 2017. “BEGAN: Boundary Equilibrium Generative Adversarial Networks.” *arXiv [Cs.LG]*. http://arxiv.org/abs/1703.10717.

[4]. Dock, Moby. 2022. “Docker: Accelerated Container Application Development.” Docker. May 10, 2022. https://www.docker.com/.

[5]. Eskin, Eleazar. 2002. *A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data*. Boston, MA: Springer.

[6]. Goodfellow, Ian J. 2014. “Generative Adversarial Nets.” *Advances in Neural Information Processing Systems* 27: 2672–80.

[7]. “IDS 2017.” n.d. Unb.Ca. Accessed November 8, 2023. https://www.unb.ca/cic/datasets/ids-2017.html.

[8]. Lunt, Teresa F. 1992. “A Real-Time Intrusion Detection Expert System (IDES).” *Computer Science Laboratory, SRI International*.

[9]. Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. “Efficient Estimation of Word Representations in Vector Space.” *arXiv [Cs.CL]*. http://arxiv.org/abs/1301.3781.

[10]. Mimno, David, Matt Hoffman, and David Blei. 2012. “Sparse Stochastic Inference for Latent Dirichlet Allocation.” *arXiv [Cs.LG]*. http://arxiv.org/abs/1206.6425.

[11]. “Production-Grade Container Orchestration.” n.d. Kubernetes. Accessed November 8, 2023. https://kubernetes.io/.

[12]. Rayavarapu, Swarajya, Tammineni Madhuri, Gottapu Shanmukha Prasanthi, and Gottapu Santosh Saibhushana Rao. n.d. “Generative Adversarial Networks for Anomaly Detection in Cyber Security: A Review.” In *Proceedings of the Fourth International Conference on Electronics and Sustainable Communication Systems (ICESC-2023)*. IEEE.

[13]. Roesch, Martin. 1999. “Snort - Lightweight Intrusion Detection for Networks.” In *Proceedings of the 13th USENIX Conference on System Administration*.

[14]. Sax, Matthias J. 2022. “Apache Kafka.” In *Encyclopedia of Big Data Technologies*, 1–8. Cham: Springer International Publishing.

[15]. Singh, Pramod. 2019. “Spark Structured Streaming.” In *Learn PySpark*, 49–65. Berkeley, CA: Apress.

[16]. Subbureddiar, Ramamoorthy, Srinivas Mukkamala, Madhukumar Shankarpani, and Andrew H. Sung. 2007. “Mining Audit Data for Intrusion Detection Systems Using Support Vector Machines and Neural Networks.” *International Journal on Information Sciences and Computing* 1 (1): 47–57. https://doi.org/10.18000/ijisac.50010.

[17]. Zenati, Houssam, Chuan Sheng Foo, Bruno Lecouat, Gaurav Manek, and Vijay Ramaseshan Chandrasekhar. 2018. “Efficient GAN-Based Anomaly Detection.” *arXiv [Cs.LG]*. http://arxiv.org/abs/1802.06222.