Journal of Theoretical and Applied Information Technology
28th February 2025. Vol.103. No.4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

GENERATIVE ADVERSARIAL NETWORKS FOR CYBER THREAT SIMULATION AND DEFENCE STRATEGIES

[1]DR. M. CHANDRA SEKHAR [2]VIJAYA ALUKAPELLY [3]TIRUGULLA NEELIMA [4]DR. RAJITHA KOTOJU, [5]DR. VADLAMANI VEERABHADRAM

[1]Professor, School of Computer Science Engineering, Presidency University, Itgalpura, Rajanukunte, Yelahanka, Bengaluru, Karnataka, Pin: 560119, India.
[2]Assistant professor, computer science and engineering, Geethanjali College of Engineering and Technology. Hyderabad, Telangana, India.
[3]Assistant Professor, computer science and engineering, TKR College of Engineering & Technology, Hyderabad, Telangana, India.
[4]Assistant professor, Department of Computer Science and Engineering, Mahatma Gandhi Institute Of Technology, (MGIT), Hyderabad 500075, India.
[5]Associate Professor, Department of CSE, CVR College of Engineering, Hyderabad Telangana India.
Emails:[1]chandragvp@gmail.com [2]vijayaphdsru@gmail.com [3]maalinimadhuri@gmail.com
[4]charuk.rajitha@gmail.com [5]drbhadram@gmail.com

ABSTRACT

The security of applications and networks is crucial and must be updated regularly. With ongoing technological innovations, adversaries continuously find new ways to compromise systems, highlighting the need to enhance cybersecurity. Traditional approaches like cryptography and firewalls have created safe coding systems and applications. However, due to the vast amounts of data flowing in today's cloud computing environment, it is essential to develop scalable methods for detecting intrusions. The emergence of artificial intelligence has made it possible to utilize deep learning models to detect cyber threats automatically. Literature suggests that there is a need for a generative adversarial network (GAN)-based deep learning framework to improve the quality of training, thereby enhancing the efficiency of intrusion detection. This paper proposes a GAN-based framework for automatically detecting cyber attacks. We use an improved CGAN model for the empirical study. We introduce an algorithm known as Learning-Based Cyber Attack Detection (LB-CAD), which leverages the enhanced CGAN model with the improved VGG16 model to optimize performance in defending against cyber attacks. Our empirical study, using a benchmark known as RT-IoT2022, revealed that the proposed method outperforms many existing approaches, achieving an accuracy of 97.62%. Therefore, the proposed framework can be integrated with existing applications to complement traditional security measures in a scalable manner.

Keywords : *Cybersecurity, Artificial Intelligence, Deep Learning, Cyber Defense Strategies, Generative Adversarial Network*

1. INTRODUCTION

Security plays a crucial role in protecting networks and applications in the real world. With the ever-increasing landscape of digital infrastructure worldwide, there is a pressing need for continuous improvement in security to safeguard the interests of all stakeholders. To enhance cybersecurity, it is essential to explore various approaches.

Cryptography and firewalls have traditionally been employed to protect information systems and networks from multiple threats. While these existing security measures are beneficial, it is time to advance security mechanisms in response to evolving attacks and adversaries' potential misuse of quantum computing. Additionally, the emergence of artificial intelligence offers promising opportunities for developing learning-based strategies to automatically detect security threats, providing an extra layer of protection for information systems and networks alongside traditional methods.

There are many existing methods found in the literature that exploit artificial intelligence to protect information systems. A network intrusion detection system has been proposed that leverages artificial intelligence to address data imbalances and enhance threat detection through the generation of synthetic data [1]. Deep learning methods for detecting cyberattacks in cyber-

physical systems have been explored, highlighting the associated challenges and identifying potential areas for future research [2]. A hybrid intrusion detection model has been introduced that combines machine learning and deep learning approaches to address security concerns and improve detection accuracy [3]. The significance of artificial intelligence and machine learning in detecting threats has been emphasized by analyzing current cybersecurity trends, challenges, and strategies [4]. Lastly, a review of artificial intelligence methods in cybersecurity discusses their advantages and disadvantages while suggesting directions for future research to strengthen defenses against malicious attacks [5]. From the literature, it was observed that cybersecurity is needed by exploiting generative adversarial network architectures to improve training and quality and enhance cyber defense strategies.

Our contributions in this paper are as follows: we propose a GAN-based framework for automatically detecting cyber attacks. Our empirical study utilizes an enhanced Conditional Generative Adversarial Network (CGAN) model. Additionally, we introduce an algorithm called Learning-Based Cyber Attack Detection (LB-CAD), which combines the enhanced CGAN model with an improved VGG16 model to optimize performance in defending against cyber attacks. Our empirical study, using a benchmark known as RT-IoT2022, demonstrates that the proposed method outperforms many existing approaches, achieving an accuracy of 97.62%. Thus, the proposed framework can be integrated with existing applications to complement traditional security measures in a scalable way.

The remainder of the paper is structured as follows: Section 2 reviews the literature on existing deep learning methods for enhancing cyber security. Section 3 presents the proposed methodology and algorithm for improving cyber security using a GAN-based approach. Section 4 discusses the results of our empirical study and compares the proposed system with various existing methods. Section 5 reviews the research conducted in this paper and addresses the study's limitations. Finally, Section 6 concludes our research and provides directions for future research opportunities.

2. RELATED WORK

Various methods exist based on deep learning to protect information systems. Park et al. [1] suggested a network intrusion detection system based on artificial intelligence that corrects data imbalance and improves threat detection by creating fake data. Gaba et al. [2] examined deep learning methods for cyberattack detection in cyber-physical systems, emphasizing difficulties and potential areas for further research. Sajid et al. [3] addressed security issues and increased detection accuracy by using a hybrid intrusion detection model that combines ML and DL approaches. Admass et al. [4] highlighted the importance of AI and machine learning in detecting threats in the future as it examines present cyber security trends, issues, and tactics. Okay et al. [5] examined artificial intelligence (AI) methods in cybersecurity, emphasizing their benefits and drawbacks and recommending further study enhancing defenses against hostile assaults.

Mehmood et al. [6] examined machine learning and quantum cybersecurity approaches, pointing out flaws and suggesting future advancements for solid security. Aurangzeb et al. [7] highlighted flaws and suggested quantum voting models for improved protection in evaluating intelligent grid security against deep black box assaults. Devi et al. [8] created the FOADL-EMAR method, which combines deep learning and optimization techniques to detect harmful behavior in smart cities better. Reliance on high-quality data is one of its limitations; nevertheless, more research in the future may offer better feature selection and more extensive testing. Abdi et al. [9] examined deep learning methods for proactive cyber protection in intelligent grids, pointing out obstacles and potential advancements. Among the limitations is the lack of thorough research on these tactics. Duraibi et al. [10] enhanced security through deep learning and feature selection in its proposed Improved Mayfly Optimization Algorithm for IoT intrusion detection. Reliance on specific algorithms is one of the limitations; further research may examine optimization strategies and broader applications.

Randhawa et al. [11] presented Botshot, a GAN-based method that addresses the imbalance and shortage of datasets to improve botnet identification. Potential overfitting and dataset generalization are among the limitations. Subsequent investigations may explore more extensive dataset uses and enhance adversarial training techniques. Huang and Lei [12] created the Imbalanced Generative Adversarial Network (IGAN) to rectify the disparity in intrusion detection between classes. Potential overfitting is

one of the limitations; more research might improve robustness against unexplained abnormalities. Navidan et al. [13] examined GAN networking applications, suggesting a methodology for performance assessment, but acknowledge that further research is necessary. Rani et al. [14] created a unique IDS using a Deep Hierarchical Model, attaining a high degree of accuracy; nonetheless, further work is required for a more comprehensive assessment. Dhanya et al. [15] assessed several machine-learning models on the UNSW-NB15 dataset, obtaining high accuracy. However, further research should be done to investigate larger datasets and attack styles.

Jarrah et al. [16] identified fresh car cyberattacks using a revolutionary machine learning-based intrusion detection system (IDS). This system outperforms current approaches, but further study is required for broader use. Diaba et al. [17] provided a GSF-optimized TNN algorithm for SCADA systems that performs better than conventional approaches; nevertheless, more extensive attack scenarios should be the focus of future studies. Zoppi et al. [18] analyzed several ML algorithms for intrusion detection systems (IDSs), emphasizing the efficacy of unsupervised meta-learning against unidentified assaults; nonetheless, future research should improve flexibility to changing threats. Diaba et al. [19] provided a hybrid deep learning system that detects DDoS assaults in intelligent grids with 99.7% accuracy; nevertheless, future research needs to address more attack types. Aldhaheri et al. [20] examined IDS for IoT security based on deep learning, addressing existing issues and outlining potential future research avenues for advancement.

Presekal et al. [21] provided a hybrid GC-LSTM model that achieves over 96% accuracy for early-stage attack detection in power grids; nevertheless, scalability and adaption to changing threats should be the focus of future efforts. Shin et al. [22] provided an OSRDW approach that outperforms existing techniques in identifying unexpected cyberattacks; nevertheless, further research should be done to improve the approach's flexibility in various settings. Macas et al. [23] examined deep learning applications in cybersecurity, noting their drawbacks and difficulties, and recommended future lines of inquiry to enhance threat detection. Zhou et al. [24] evaluated adversarial assaults on deep learning and presented a complete analytical framework but stressed the necessity for uniform assessment methodologies. Skopik et al. [25] examined log analysis techniques using machine learning for intrusion detection, suggesting the AMiner pipeline but emphasizing the need for more reliable algorithms.

Mughaid et al. [26] provided a machine learning model that can identify phishing emails with high accuracy but stresses the need for constant adaption to changing phishing strategies. Gupta et al. [27] introduced MUSE, a deep learning system that accurately identifies fraudulent activity in healthcare networks while emphasizing the need for further scalability and adaptability. Aloqaily et al. [28] focused on the security issues that 5G and beyond will provide for cyber-physical systems, highlighting the necessity of flexible cybersecurity measures to counteract changing risks. Wang et al. [29] offer a layered deep learning approach that outperforms conventional IDS for cyberattack detection in SCADA systems. However, more research is needed to increase flexibility and real-time reaction. Garcia and Blandon [30] created Dique, a deep learning-based IDS/IPS with a 99.4% accuracy rate for identifying DoS assaults; nevertheless, more work may be done to improve defense mechanisms and flexibility.

Moti et al. [31] presented MalGan, a framework for IoT malware detection that combines CNN and GANs. It achieves better accuracy but still needs more efficiency improvements. Andresen et al. [32] used deep learning with GANs and CNNs to represent 2D traffic images and supplement data to improve the accuracy of network intrusion detection. Cai et al. [33] examined the privacy and security applications of GANs, stressing their benefits, drawbacks, and possible future research areas. Fard et al. [34] provided a GAN-based technique for identifying car cyberattacks and suggested improvements to improve stability and performance. Zhang et al. [35] indicated that CWGAN-CSSAE achieves excellent accuracy and F1 scores in intrusion detection despite unknown assaults and data imbalance.

Table 1: Summary Of Literature Findings

Reference	Approach	Technique	Algorithm	Dataset	Limitations / Future Scope
[3]	ML and DL	XGBoost and CNN	LSTM, GRU, and RNN algorithms	CIC IDS 2017, UNSW NB15, NSL KDD, and WSN DS dataset	Improve model training with GAN architectures.
[6]	ML and DL	Quantum techniques	Robust encryption algorithm	BraTS18 dataset	The model stability is yet to be established.
[8]	DL	OADL-EMAR technique	HHO algorithm	ISCX 2012 IDS datasets	Need to carry out scalability tests.
[10]	DL	IMFOHDL-ID technique	Mayfly Optimization Algorithm	TRA and TES datasets	Needs an adaptive approach and scalability in the methodology.
[11]	ML	GANs	C2ST Algorithm	CIC-2017 and CIC-2018 datasets	We need to make the IDS autonomous.
[16]	ML	ML techniques	DTW algorithm	CAN intrusion-dataset	It needs to be extended to VANET scenarios.
[21]	DL	GC-LSTM	TSC algorithm	OT network traffic dataset	Augmentation with GAN is to be explored.
[26]	ML	ML techniques	ML algorithm	Custom dataset	Feature selection methods are to be improved.
[34]	ML	Synthetic Minority Over-Sampling Technique (SMOTE)	ML algorithm	VX-Heaven dataset,	Evaluation is required with more threat scenarios.
[35]	ML	CWGAN	ML algorithm	KDDTestC, KDDTest-21, and UNSW-NB15 datasets,	To be evaluated with live network traffic.

Duy et al. [36] created a GAN-based DIGFuPAS system to produce adversarial traffic and increase IDS resilience against intrusions in SDN. Chen et al. [37] provided an unsupervised deep learning architecture for detecting fraud, enhancing system performance, and lowering false positives in AML systems. Terziyan et al. [38] suggested utilizing Generative Adversarial Networks (GANs) to overcome security flaws and provide a self-defense mechanism for AI in Industry 4.0. Yang et al. [39] provided an approach for assessing network security using adversarial deep learning, which improves attack categorization and evaluation flexibility but has difficulties with minority attack performance. Shi et al. [40] designed a spoofing assault with deep knowledge and GANs to produce identical signals, increasing attack success but posing difficulties with many antennas. Table 1 shows the literature findings. From the literature, it was understood that GAN-based architectures with deep learning could be

improved for better defense mechanisms towards cyber security.

3. MATERIALS AND METHODS

The proposed methodology for enhancing cybersecurity is based on deep learning and generative adversarial network (GAN) architecture. Given Georgia's improved performance in these architectures, even with limited training samples, it is evident that further enhancements can bolster cybersecurity defense strategies. This section outlines the proposed methodology, the framework, the underlying algorithm, details about the dataset, and the evaluation methodology.

3.1 Proposed Framework

The proposed framework is based on deep learning and an enhanced cognitive genetic algorithm (CGA) in its architecture. The decision was to incorporate deep learning because these models can learn from data in-depth, providing more resilient cyber defense strategies. Furthermore, the framework utilizes an enhanced CGAN model, which has been shown to improve training quality. This enhancement is particularly beneficial for strengthening cyber defense strategies, even when a limited number of training samples are available.

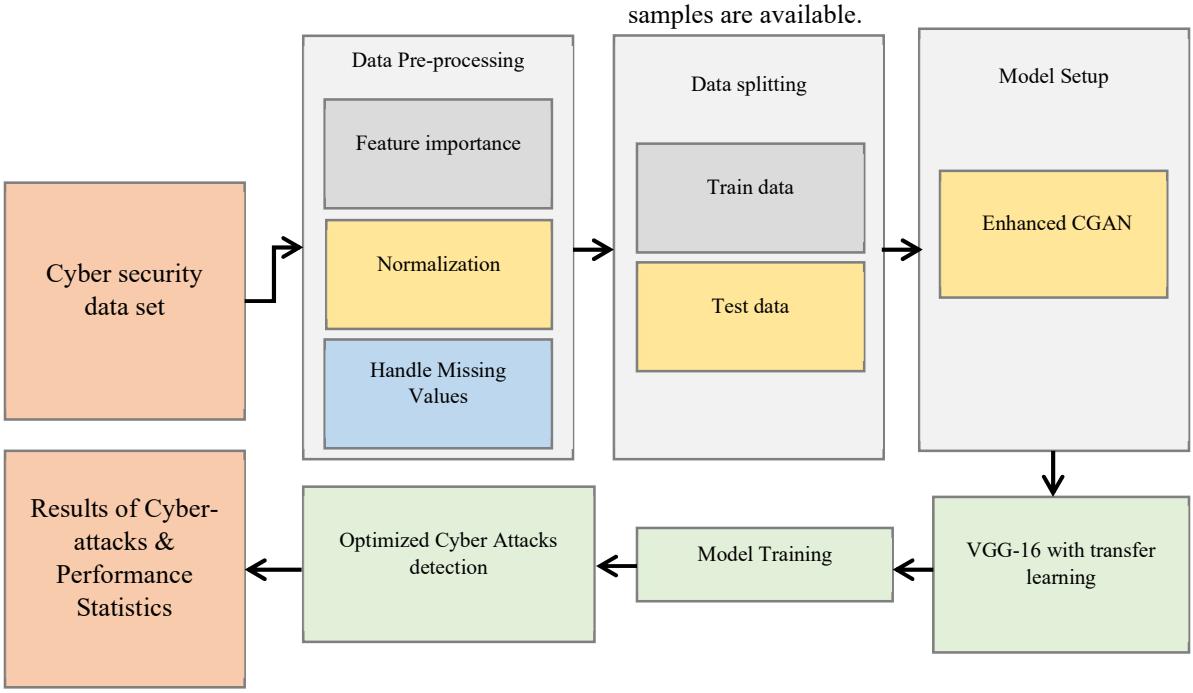

Figure 1: Proposed Enhanced CGAN-Based Framework For Enhancing Cybersecurity

The proposed framework, illustrated in Figure 1, is based on an enhanced CGAN-based architecture to improve cybersecurity. This framework is designed to accommodate any cybersecurity dataset, which undergoes preprocessing to compute feature importance, normalize values, and handle missing data—all enhancing data quality. After preprocessing, the data is prepared for the proposed deep learning architecture. Specifically, the dataset is divided into training and testing subsets to train the models effectively. The enhanced CGA model further improves data quality. Data augmentation is conducted to increase the diversity of data samples, after which the training begins. In this study, VGG16, a pre-trained deep learning model, has been enhanced using transfer learning techniques to detect cyberattacks effectively. This model incorporates possibilities for transfer learning and utilizes data provided by the GAN architecture to improve training quality. Once trained with sufficient data samples, the model is saved for future use. The saved model is a cyberattack detection system implemented in real-world applications. When new data arrives, the model is employed for intrusion detection, thereby enhancing cybersecurity. Finally, the model is evaluated to assess its effectiveness in detecting cyberattacks.

3.2 CGAN Architecture

Generative adversarial network architectures are found to be efficient in improving the diversity of samples in datasets. In this paper, we explore CGAN for this purpose only and try to improve it

further by generating more diversified samples that could improve the training process.

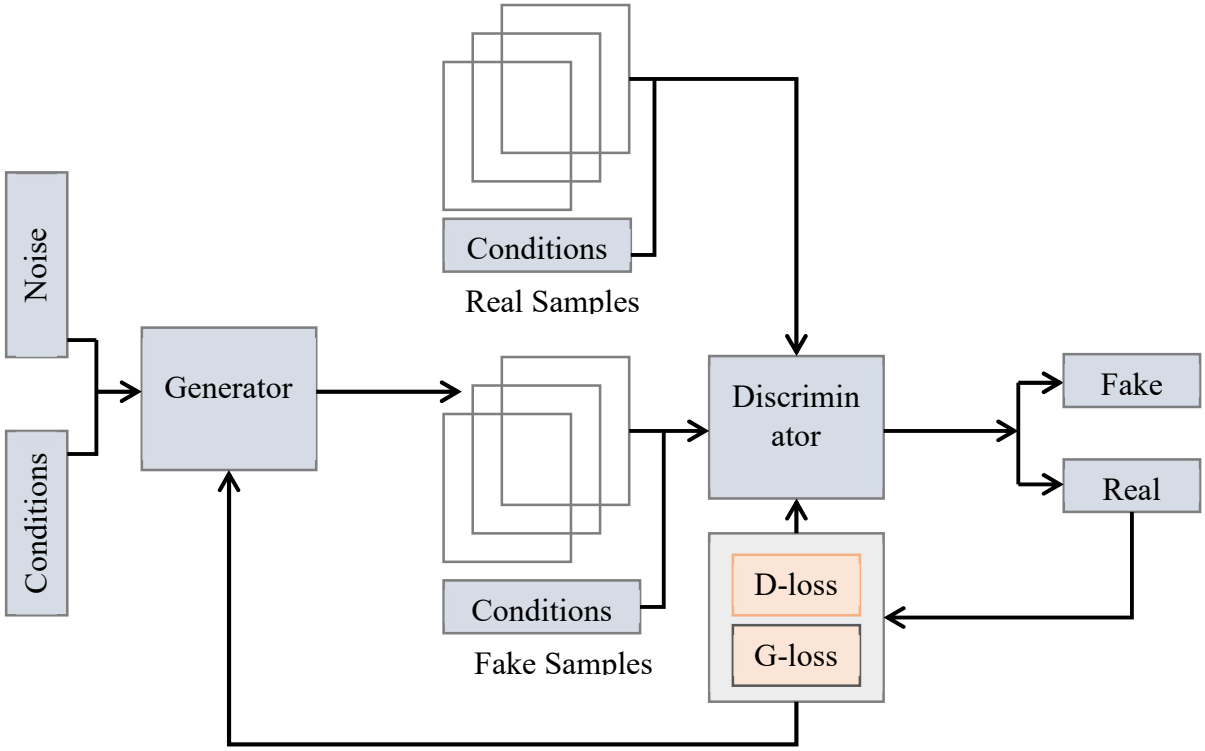

Figure 2: Conditional GAN Architecture Used In The Proposed System

The proposed deep learning framework is designed to enhance cybersecurity. This framework utilizes a conditional GAN (Generative Adversarial Network) architecture, as illustrated in Figure 2. The architecture was improved during the implementation to enhance its functionality and generate helpful training samples. It operates based on specific conditions that help improve the quality of the generated samples. The system consists of a generator and a discriminator. The generator inputs conditions and noise to produce various "fake" samples. These fake samples are then provided to the discriminator, who is responsible for classifying the samples as either fake or real and giving feedback. This process is akin to a non-cooperative game, and it involves continuous interaction between the generator and the discriminator to enhance data augmentation and ensure diversity in the dataset. The model periodically utilizes generator and discriminator losses to ensure adequate sample generation. Instead of directly employing deep learning models, this proposed system enhances training samples using the GAN architecture. The sophisticated training samples will enable deep learning models to acquire knowledge, ultimately acting as a robust cyber defense strategy. Once training is completed, the deep learning model, which utilizes data from the GAN architecture, will be capable of detecting cyberattacks with greater accuracy. In summary, the proposed deep learning framework with an enhanced CGAN model can significantly improve system performance by enhancing data augmentation and the overall quality of training for deep learning models. The deep learning model can incrementally acquire knowledge by utilizing transfer learning, thereby continuously enhancing the cyber defense strategy.

3.3 Proposed Algorithm

The proposed algorithm, Learning-Based Cyber Attack Detection (LB-CAD), is designed to be an adaptable defense strategy for detecting cyberattacks. It has provision for working on a benchmark dataset RT-IoT2022 to improve the diversity of samples using enhanced CGAN. Once the diversity of samples is increased, it has provision to train the VGG16 model with transfer learning to gain intelligence on cyberattack detection. The algorithm eventually exploits the trained model to detect intrusions as new test samples arrive.

Algorithm: Learning-Based Cyber Attack Detection (LB-CAD)
Input: RT-IoT2022 dataset D
Output: Cyberattack detection results R, performance statistics P

1. Begin
2. D'←DataPreprocessing(D)
3. (T1, T2)←DataPreparation(D')
4. Configure enhanced CGAN
5. Compile enhanced CGAN
6. (T1', T2')←RunEnhancedCGAN(T1, T2) //enhancing diversity of samples
7. Configure VGG16 model m
8. Compile m
9. m'←ModelTraining(T1', m)
10. Save m'
11. Load m'
12. R←AttackDetection(T2', m')
13. P←FindPerformance(ground truth, R)
14. Print R
15. Print P
16. End

Algorithm 1: Learning-Based Cyber Attack Detection (LB-CAD)

Algorithm 1 (LB-CAD) is designed to input a benchmark dataset RT-IoT2022 and produce intrusion detection results and performance of the model used in the study. The algorithm performs data preprocessing where it has provision to find feature meaningful and contributing features. This process helps in improving the quality of training in the proposed system. Apart from feature importance, the algorithm performs normalization and handles missing values as part of pre-processing. The data preparation part of the algorithm splits the dataset into training and test sets. Then, the algorithm exploits an enhanced CGAN model where the generator and discriminator networks play a non-cooperative game to generate more diversified samples that have the potential to leverage training quality. After the completion of an increasing diversity of training samples, the VGG16 model is used with transfer learning to learn from data and gain the knowledge required to develop cyber-attack defense strategies. The model works on the test samples for optimized cyberattack detection. Eventually, the algorithm provides cyberattack detection results and evaluates performance.

3.3 Dataset Details

The empirical study uses the RT-IoT2022 [41] dataset. The data is derived from real-time IoT infrastructure, which offers a wide range of possibilities with network traffic suitable for cyberattack detection and classification research.

3.4 Evaluation Methodology

Performance evaluation is carried out using the matrix derived from the confusion matrix, which is widely used in machine learning research to assess the ability of machine learning or deep learning models in classification tasks. Figure 3 shows a confusion matrix that helps compare the prediction results with ground truth.

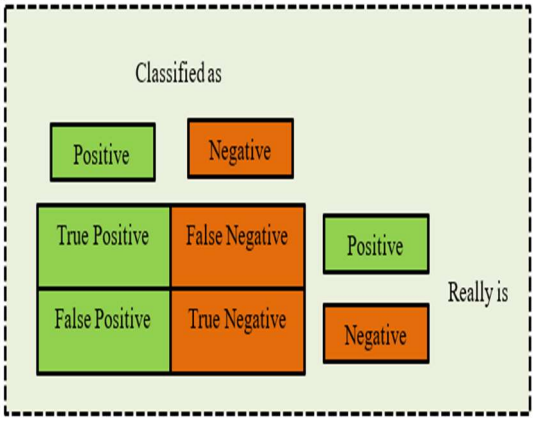

Figure 3: Confusion matrix

Based on the confusion matrix, the predicted labels of our method are compared with the ground truth to arrive at performance statistics. Eq. 1 to Eq. 4 express different metrics used in performance evaluation.

$$\text{Precision (p)} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall (r)} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{F1-score} = 2 * \frac{(p* r)}{(p+r)} \quad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The measures used for performance evaluation result in a value that lies between 0 and 1. These metrics are widely used in machine learning research.

4. EXPERIMENTAL RESULTS

This section presents the results of our empirical study, which was conducted using the RT-IoT2022 dataset. The proposed framework uses the Python programming language and machine learning libraries. The environment used for the empirical research is a PC with Windows 11 OS, and the system has a processor of 13th Gen Intel(R) Core(TM) i7-1355U, 1700 Mhz, 10 Core(s), and 12 Logical Processor(s). Before

Journal of Theoretical and Applied Information Technology
28th February 2025. Vol.103. No.4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

working with the proposed framework and underlying algorithm, exploratory data analysis is made.

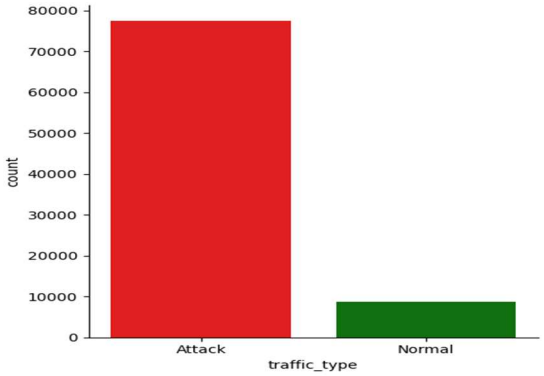

Figure 4: Data Distribution Dynamics In The Dataset

Figure 4 shows the data distribution dynamics in the data set. The horizontal axis shows two class labels, attack and normal, for which the number of samples available in the data set is provided in the vertical axis.

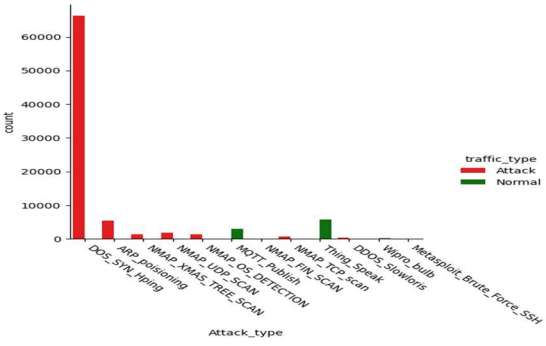

Figure 5: Different Types Of Attack Distribution Dynamics In The Dataset

Figure 5 illustrates different types of attack distribution dynamics in the data set. The data set contains multiple class labels about various kinds of attacks, which are reflected in the visualization. Different kinds of attacks are provided in original access, while the vertical axis shows the number of samples present in the data set for each attack category and also normal samples.

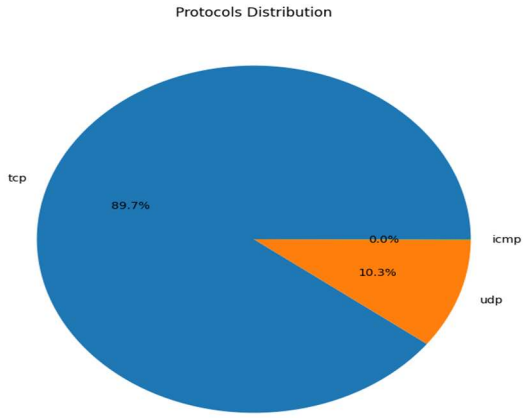

Figure 6: Data Distribution Dynamics In The Data Set In Terms Of Protocols

Figure 6 illustrates the data distribution dynamics in the data set for various protocols like TCP, UDP, and ICMP. Since these protocols are used in several types of networks, the dataset reflects the protocols and the percentage of samples available in the data set for each category of protocols. TCP protocol has the highest number of samples, with 89.70%.

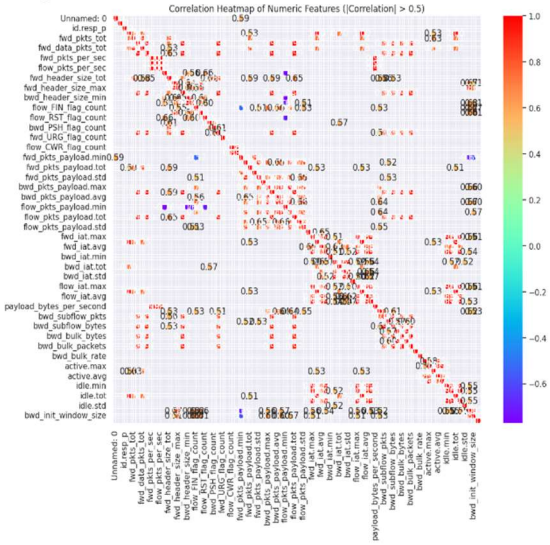

Figure 7: Correlation Heatmap Of Numerical Features Found In The Dataset

Figure 7 shows the correlation heat map of numerical features found in the data set. Each feature is correlated with itself and all other features to understand the correlation dynamics among the phases. The correlation value may be as low as 1 and as high as 1. The color scale provided in the correlation heat map visualizes the strength associated with correlations.

Journal of Theoretical and Applied Information Technology
28th February 2025. Vol.103. No.4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

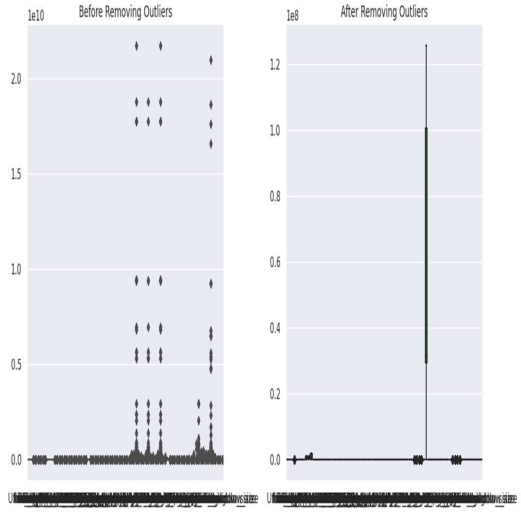

Figure 8: Showing The Visualization Reflecting Before And After Removing Outliers

Figure 8 shows the dynamics of numeric features before and after removing outliers. Since outliers can impact the accuracy of machine learning models, identifying and removing them is one of the preprocessing steps. Towards this end, the outliers have been removed, and the visualization shows before and after outlier removal (left and right, respectively).

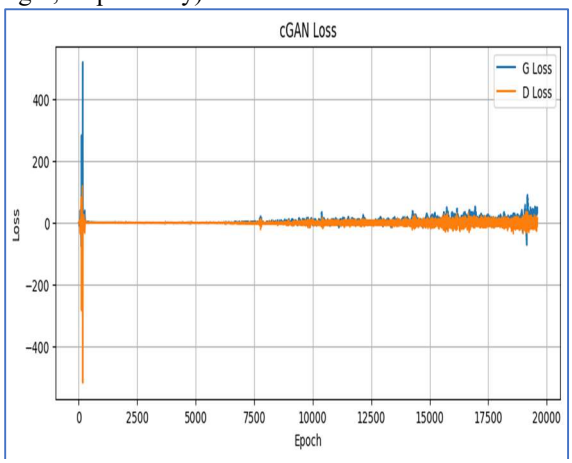

Figure 9: Loss Function Dynamics Associated With The CGAN

The loss function dynamics of CGAN are visualized in Figure 9. It provides the loss of generator and discriminator networks' dynamics against several epochs. The results of loss computation show that the CGAN exhibited relatively less loss value for both generator and discriminator, showing the model's efficiency.

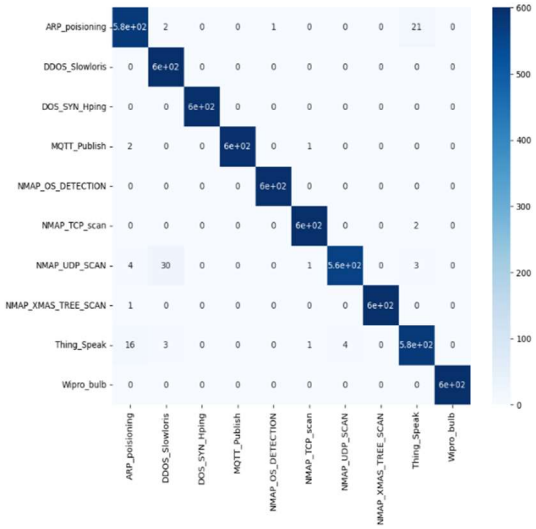

Figure 10: Confusion Matrix For The Classification Model

Figure 10 shows a confusion matrix for multiclass classification. The given data set has several classes reflecting various kinds of attacks. For each data class, the confusion matrix shows values from which performance metrics can be derived. The color scale at the right shows the strength of values associated with each class label.

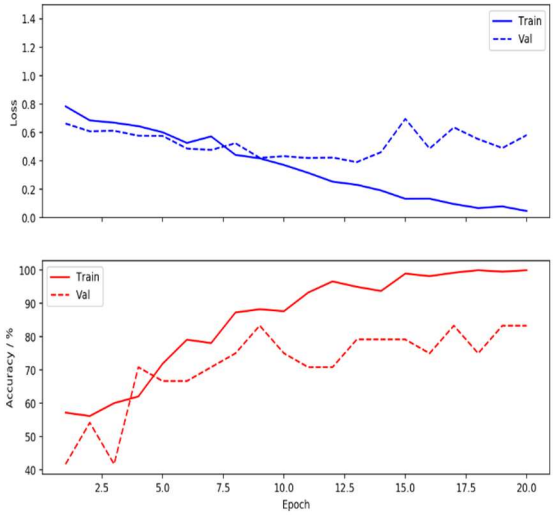

Figure 11: Loss And The Accuracy Graphs For The VGG-16 Model

Feature 11 shows the loss and accuracy dynamics of the VG16 model with transfer learning used for multi-class classification. The loss and accuracy dynamics are provided against several epochs. The model was executed for 20 epochs, and for each epoch, the last value is gradually decreased while the accuracy value is slowly increased until convergence.

Table 2: Performance Comparison Among Deep Learning Models Used For Cyberattack Detection

Model	Precision	Recall	F1-Score	Accuracy
BaseLine CNN	85.21	85.32	85.26	85.51
BaseLine LSTM	87.23	87.8	87.51	87.54
ResNet50	90.37	90.21	90.28	90.8
VGG-16	92.15	92.63	92.38	92.54
CGAN with VGG-16 (Proposed)	97.36	97.03	97.19	97.62

Table 2 compares the performance of deep learning models used for cyber attack detection, including the proposed and existing models.

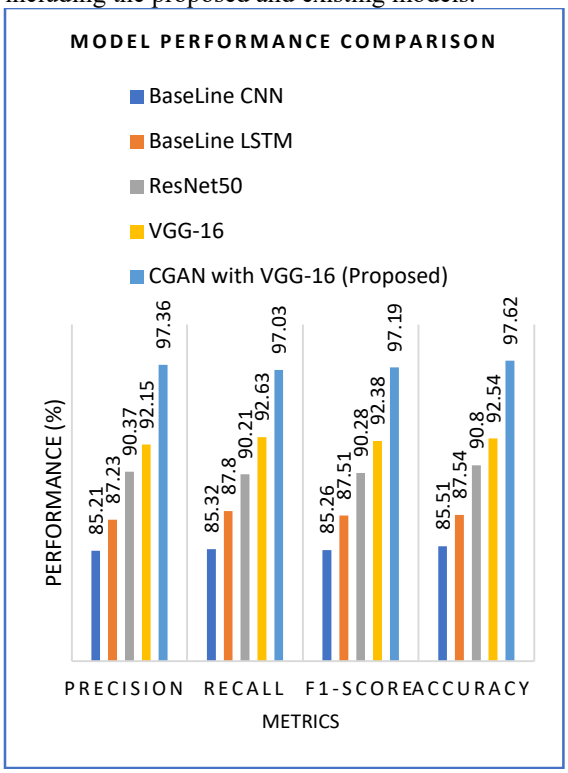

Figure 12: Performance Comparison Of Deep Learning Models In Cyber Attack Detection

Figure 12 compares the performance of various deep learning models used in cyberattack detection. The experimental results demonstrate that each model exhibits different performance levels due to its underlying architecture and operational methodology. In terms of precision, the baseline CNN achieved 85.21%, while the baseline LSTM attained 87.23%. The ResNet50 model achieved 90.37%, the VGG16 model reached 92.15%, and the proposed model achieved the highest precision at 97.36%. Regarding recall, the baseline CNN model showed a recall rate of 85.32%. The baseline LSTM model achieved 87.80%, the ResNet50 model reached 90.21%, the VGG16 model attained 92.63%, and the proposed model achieved the highest recall rate with 97.03%. For the F1 score metric, the baseline CNN model scored 85.26%, the baseline LSTM model reached 87.51%, the ResNet50 model achieved 90.28%, and the proposed model achieved the highest F1 score of 97.19%. Finally, the baseline CNN model recorded an accuracy of 85.51%, the baseline LSTM model achieved 87.54%, the ResNet50 model reached 90.80%, and the VGG16 model attained 92.54%. In comparison, the proposed model achieved the highest accuracy at 97.62%.

5. DISCUSSION

With the emergence of artificial intelligence, it has become clear that solutions are being developed for various problems across different domains. Recently, AI has played a significant role in a variety of applications. Since security is crucial for protecting applications and information systems, leveraging artificial intelligence to create a scalable, real-time intrusion detection system is essential. This system should be able to identify threats during runtime and intelligently implement defense strategies. As new technologies and innovations proliferate, security vulnerabilities are also increasing. Adversaries exploit the latest technologies to breach security systems, highlighting the need to continuously improve security standards, especially in light of emerging quantum computing technologies. In this context, we propose a gain-based architecture that utilizes a non-cooperative game theory approach to enhance the training samples. This improvement can improve training quality for deep learning models, ultimately enhancing defense strategies. Using transfer learning to leverage its performance, we have improved the pre-trained deep learning model, VGG16. The enhancement of training samples and the transfer learning applied to VGG16 could significantly improve the system's performance in detecting attacks. However, the proposed system has certain limitations, as discussed in section 5.1.

5.1 Limitations

The proposed system discussed in this paper demonstrates improved performance compared to existing systems, providing effective defense strategies against various attacks. However, it does have certain limitations. The dataset used for the empirical study consists of a finite number of samples, indicating a need for more diverse samples to generalize the findings. Additionally, it is worth noting that the proposed system employs neural networks as components of both the generator and administrator networks, which could be further enhanced by incorporating more advanced deep learning models.

6. CONCLUSION AND FUTURE WORK

We propose a GAN-based framework for the automatic detection of cyber attacks. Our study utilizes an enhanced Conditional Generative Adversarial Network (CGAN) model and introduces an algorithm called Learning-Based Cyber Attack Detection (LB-CAD). This algorithm leverages the enhanced CGAN model with an improved VGG16 model to optimize performance in defending against cyber attacks. Our empirical study, conducted using the RT-IoT2022 benchmark, demonstrated that our proposed method outperforms many existing approaches, achieving an impressive accuracy of 97.62%. This framework can be integrated with existing applications to complement traditional security measures in a scalable way. Furthermore, we have identified several opportunities for enhancing the proposed framework based on our empirical findings. For instance, the generator and discriminator networks could be replaced with more advanced deep-learning models to generate more diverse samples. Additionally, we see potential in utilizing a variety of deep learning models for detection and classification purposes in our future endeavors.

REFERENCES

[1] Cheolhee Park, Jonghoon Lee, Youngsoo Kim, Jong-Geun Park, Hyunjin Kim, and Dowon Hong. (2023). An Enhanced AI-Based Network Intrusion Detection System Using Generative Adversarial Networks. *IEEE*. 10(3), pp.2330 - 2345. http://DOI:10.1109/JIOT.2022.3211346

[2] SHIVANI GABA, ISHAN BUDHIRAJA, VIMAL KUMAR, SHESHIKALA MARTHA, JEBREEL KHURMI, AKANSHA SINGH, KRISHNA KANT SINGH, S. S. ASKAR, AND MOHAMED ABOUHAWWASH. (2024). A systematic analysis of enhancing cyber security using deep learning for cyber-physical systems. *IEEE*. 12, pp.6017 - 6035. http://DOI:10.1109/ACCESS.2023.3349022

[3] Muhammad Sajid, Kaleem Razzaq Malik, Ahmad Almogren, Tauqeer Safdar Malik, Ali Haider Khan, Jawad Tanveer, and Ateeq Ur Rehman. (2024). Enhancing intrusion detection: a hybrid machine and deep learning approach. *Springer*. 13(123), p.1-24. https://doi.org/10.1186/s13677-024-00685-x

[4] Wasyihun Sema Admass, Yirga Yayeh Munaye, and Abebe Abeshu Diro. (2024). Cyber security: State of the art, challenges and future directions. *Elsevier*. 2, pp.1-9. https://doi.org/10.1016/j.csa.2023.100031

[5] MERVE OZKAN-OKAY, ERDAL AKIN, ÖMER ASLAN, SELAHATTIN KOSUNALP, TEODOR ILIEV, IVAYLO STOYANOV, AND IVAN BELOEV. (2024). A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions. *IEEE*. 12, pp.12229 - 12256. http://DOI:10.1109/ACCESS.2024.3355547

[6] ABID MEHMOOD, ARSLAN SHAFIQUE, MOATSUM ALAWIDA, AND ABDUL NASIR KHAN. (2024). Advances and Vulnerabilities in Modern Cryptographic Techniques: A Comprehensive Survey on Cybersecurity in the Domain of Machine/Deep Learning and Quantum Techniques. *IEEE*. 12, pp.27530 - 27555. http://DOI:10.1109/ACCESS.2024.3367232

[7] Muhammad Aurangzeb, Yifei Wang, Sheeraz Iqbal, Ausnain Naveed, Zeeshan Ahmed, Mohammed Alenezi, and Mokhtar Shouran. (2024). Enhancing cybersecurity in smart grids: Deep black box adversarial attacks and quantum voting ensemble models for blockchain privacy-preserving storage. *Elsevier*. 11, pp.2493-2515. https://doi.org/10.1016/j.egyr.2024.02.010

[8] Roopa Devi E. M, Naif Almakayeel, E. Laxmi Lydia. (2024). Improved sand cat swarm optimization with deep learning based enhanced malicious activity recognition for cybersecurity. *Elsevier*. 98, pp.187-198. https://doi.org/10.1016/j.aej.2024.04.053

[9] Nima Abdi, Abdullatif Albaseer, and Mohamed Abdallah. (2024). The Role of

Deep Learning in Advancing Proactive Cybersecurity Measures for Smart Grid Networks: A Survey. *IEEE*. 11(9), pp.16398 - 16421. http://DOI:10.1109/JIOT.2024.3354045

[10] SALAHALDEEN DURAIBI AND ABDULLAH MUJAWIB ALASHJAEE. (2024). Enhancing Cyberattack Detection Using Dimensionality Reduction with Hybrid Deep Learning on Internet of Things Environment. *IEEE*. 12, pp.84752 - 84762. http://DOI:10.1109/ACCESS.2024.3411612

[11] Randhawa, R. H., Aslam, N., Alauthman, M., Rafiq, H., & Comeau, F. (2021). Security Hardening of Botnet Detectors Using Generative Adversarial Networks. IEEE Access, 9, 78276–78292. doi:10.1109/access.2021.3083421

[12] Huang, S., & Lei, K. (2020). IGAN-IDS: An Imbalanced Generative Adversarial Network towards Intrusion Detection System in Ad-hoc Networks. Ad Hoc Networks, 102177. doi:10.1016/j.adhoc.2020.102177

[13] Navidan, H., Moshiri, P. F., Nabati, M., Shahbazian, R., Ghorashi, S. A., Shah-Mansouri, V., & Windridge, D. (2021). Generative Adversarial Networks (GANs) in networking: A comprehensive survey & evaluation. Computer Networks, 194, 108149. doi:10.1016/j.comnet.2021.108149

[14] S. V. JANSI RANI, IACOVOS I. IOANNOU, PRABAGARANE NAGARADJANE, CHRISTOPHOROS CHRISTOPHOROU, VASOS VASSILIOU, HARSHITAA YARRAMSETTI, SAI SHRIDHAR, L. MUKUND BALAJI, AND ANDREAS PITSILLIDES. (2023). A Novel Deep Hierarchical Machine Learning Approach for Identification of Known and Unknown Multiple Security Attacks in a D2D Communications Network. *IEEE*. 11, pp.95161 - 95194. http://DOI:10.1109/ACCESS.2023.3308036

[15] Dhanya K. A., Sulakshan Vajipayajula, Kartik Srinivasan, Anjali Tibrewal,. (2023). Detection of network attacks using machine learning and deep learning models. *Elsevier*. 218, pp.57-66. https://doi.org/10.1016/j.procs.2022.12.401

[16] OMAR Y. AL-JARRAH, KARIM EL HALOUI, MEHRDAD DIANATI, AND CARSTEN MAPLE. (2023). A novel detection approach of unknown cyber-attacks for intra-vehicle networks using recurrence plots and neural network. *IEEE*. 4, pp.271 - 280. http://DOI:10.1109/OJVT.2023.3237802

[17] Sayawu Yakubu Diaba, Theophilus Anafo, Lord Anertei Tetteh, Michael Alewo Oyibo, Andrew Adewale Alola, Miadreza Shafie-khah, and Mohammed Elmusrati. (2023). SCADA securing system using deep learning to prevent cyber infiltration. *Elsevier*. 165, pp.321-332. https://doi.org/10.1016/j.neunet.2023.05.047

[18] Tommaso Zoppi, Andrea Ceccarelli, Tommaso Puccetti, and Andrea Bondavalli. (2023). Which algorithm can detect unknown attacks? Comparison of supervised, unsupervised and meta-learning algorithms for intrusion detection. *Elsevier*. 127, pp.1-12. https://doi.org/10.1016/j.cose.2023.103107

[19] Sayawu Yakubu Diaba, and Mohammed Elmusrati. (2023). Proposed algorithm for smart grid DDoS detection based on deep learning. *Elsevier*. 159, pp.175-184. https://doi.org/10.1016/j.neunet.2022.12.011

[20] Alyazia Aldhaheri, Fatima Alwahedi, Mohamed Amine Ferrag, and Ammar Battah. (2023). Deep learning for cyber threat detection in IoT networks: A review. *Elsevier*. 4, pp.110-128. https://doi.org/10.1016/j.iotcps.2023.09.003

[21] Alfan Presekal, Alexandru Ştefanov, Vetrivel S. Rajkumar, and Peter Palensky. (2023). Attack graph model for cyber-physical power systems using hybrid deep learning. *IEEE*. 14(5), pp.4007 - 4020. http://DOI:10.1109/TSG.2023.3237011

[22] GUN-YOON SHIN, DONG-WOOK KIM, AND MYUNG-MOOK HAN. (2023). Open Set Recognition with Dissimilarity Weight for Unknown Attack Detection. *IEEE*. 11, pp.102381 - 102390. http://DOI:10.1109/ACCESS.2023.3277871

[23] Mayra Macas, Chunming Wu, and Walter Fuertes. (2022). A survey on deep learning for cybersecurity: Progress, challenges, and opportunities. *Elsevier*. 212, pp.1-40. https://doi.org/10.1016/j.comnet.2022.109032

[24] SHUAI ZHOU, CHI LIU, DAYONG YE, TIANQING ZHU, WANLEI ZHOU, PHILIP S. YU, U. (2022). Adversarial attacks and defenses in deep learning: From a perspective of cybersecurity. *ACM*. 55(8), pp.1-39. https://doi.org/10.1145/3547330

[25] Florian Skopik, Markus Wurzenberger, and Max Landauer. (2022). Detecting unknown

cyber security attacks through system behavior analysis. *Sprigner*., p.103–119. https://doi.org/10.1007/978-3-031-04036-8_5

[26] Ala Mughaid, Shadi AlZu'bi, Adnan Hnaif, Salah Taamneh, Asma Alnajjar, and Esraa Abu Elsoud. (2022). An intelligent cyber security phishing detection system using deep learning techniques. *Springer*. 25, p.3819–3828. https://doi.org/10.1007/s10586-022-03604-4

[27] Lav Gupta, Tara Salman, Ali Ghubaish, Devrim Unal, Abdulla Khalid Al-Ali, and Raj Jain. (2022). Cybersecurity of multi-cloud healthcare systems: A hierarchical deep learning approach. *Elsevier*. 118, pp.1-23. https://doi.org/10.1016/j.asoc.2022.108439

[28] Moayad Aloqaily, Salil Kanhere, Paolo Bellavista, and Michele Nogueira. (2022). Special issue on cybersecurity management in the era of AI. *Springer*. 30(39), pp.1-7. https://doi.org/10.1007/s10922-022-09659-3

[29] Wu Wang, Fouzi Harrou, Benamar Bouyeddou, Sidi-Mohammed Senouci, and Ying Sun. (2022). A stacked deep learning approach to cyber-attacks detection in industrial systems: application to power system and gas pipeline systems. *Springer*. 25, p.561–578. https://doi.org/10.1007/s10586-021-03426-w

[30] JUAN FERNANDO CAÑOLA GARCIA, AND GABRIEL ENRIQUE TABORDA BLANDON. (2022). A deep learning-based intrusion detection and preventation system for detecting and preventing denial-of-service attacks. *IEEE*. 10, pp.83043 - 83060. http://DOI:10.1109/ACCESS.2022.3196642

[31] Moti, Z., Hashemi, S., Karimipour, H., Dehghantanha, A., Jahromi, A. N., Abdi, L., & Alavi, F. (2021). Generative adversarial network to detect unseen Internet of Things malware. Ad Hoc Networks, 122, 102591. doi:10.1016/j.adhoc.2021.102591

[32] Andresini, G., Appice, A., De Rose, L., & Malerba, D. (2021). GAN augmentation to deal with imbalance in imaging-based intrusion detection. Future Generation Computer Systems, 123, 108–127. doi:10.1016/j.future.2021.04.017

[33] Cai, Z., Xiong, Z., Xu, H., Wang, P., Li, W., & Pan, Y. (2021). Generative Adversarial Networks. ACM Computing Surveys, 54(6), 1–38. doi:10.1145/3459992

[34] Kavousi-Fard, A., Dabbaghjamanesh, M., Jin, T., Su, W., & Roustaei, M. (2020). An Evolutionary Deep Learning-Based Anomaly Detection Model for Securing Vehicles. IEEE Transactions on Intelligent Transportation Systems, 1–9. doi:10.1109/tits.2020.3015143

[35] Zhang, G., Wang, X., Li, R., Song, Y., He, J., & Lai, J. (2020). Network Intrusion Detection Based on Conditional Wasserstein Generative Adversarial Network and Cost-Sensitive Stacked Autoencoder. IEEE Access, 8, 190431–190447. doi:10.1109/access.2020.3031892

[36] Duy, P. T., Tien, L. K., Khoa, N. H., Hien, D. T. T., Nguyen, A. G.-T., & Pham, V.-H. (2021). DIGFuPAS: Deceive IDS with GAN and function-preserving on adversarial samples in SDN-enabled networks. Computers & Security, 109, 102367. doi:10.1016/j.cose.2021.102367

[37] Chen, Z., Soliman, W. M., Nazir, A., & Shorfuzzaman, M. (2021). Variational Autoencoders and Wasserstein Generative Adversarial Networks for Improving the Anti-Money Laundering Process. IEEE Access, 9, 83762–83785. doi:10.1109/access.2021.3086359

[38] Terziyan, V., Gryshko, S., & Golovianko, M. (2021). Taxonomy of generative adversarial networks for digital immunity of Industry 4.0 systems. Procedia Computer Science, 180, 676–685. doi:10.1016/j.procs.2021.01.290

[39] Yang, H., Zeng, R., Xu, G., & Zhang, L. (2021). A network security situation assessment method based on adversarial deep learning. Applied Soft Computing, 102, 107096. doi:10.1016/j.asoc.2021.107096

[40] Shi, Y., Davaslioglu, K., & Sagduyu, Y. E. (2020). Generative Adversarial Network in the Air: Deep Adversarial Learning for Wireless Signal Spoofing. IEEE Transactions on Cognitive Communications and Networking, 1–1. doi:10.1109/tccn.2020.3010330

[41] RT-IoT2022 dataset. Retrieved from https://www.kaggle.com/datasets/joebeachcapital/real-time-internet-of-things-rt-iot2022