Advances in Artificial Intelligence Research (AAIR)
ISSN 2757-7422
www.dergipark.com/aair/

Vol. 4 (No. 1), pp. 18-35, 2024
doi: 10.54569/aair.1442665
Published online: Aug 30, 2024

Generative Adversarial Networks in Anomaly Detection and Malware Detection: A Comprehensive Survey

Bishal KC [1,*], Shushant Sapkota [1], Ashish Adhikari [1]

[1] Kathmandu University, Department of Computer Science & Engineering, Nepal

Abstract

The swiftly changing panorama of machine learning has observed first-rate leaps within the field of Generative Adversarial Networks (GANs). In the beginning, the implantation of a deep neural network seemed quite difficult and poses challenges. However, with the rapid development of huge processing power, different machine learning models such as Convolutional Neural Networks, Recurrent Neural Networks, and GANs have emerged in the past few years. Following Ian Goodfellow's proposed GANs model in 2014, there has been a huge increase in the research focused on Generative Adversarial Networks. In the present context, not only GANs are used in feature extraction, but it proves itself worthy in the domain of anomaly and malware detection having firmly established in this field. Therefore, in our research paper, we conducted a comprehensive survey of prior and current research attempts in anomaly and malware detection using GANs. This research paper aims to provides detailed insights to the reader about what types of GANs are used for anomaly and malware detection with a general overview of the different types of GANs. These results are provided by analyzing both past and present GAN surveys performed, along with detailed information regarding the datasets used in these surveyed papers. Furthermore, this paper also explores the potential future use of GANs to overcome the advancing threats and malware.

Keywords: *Generative Adversarial Networks (GAN); Network Security; Deep Neural Network (DNN); Research Survey; Threat detection; Malware detection; Adversarial examples.*

1. Introduction

Malware also known as malicious software, is undesired programs designed to harm or exploit computer systems [1]. After the outbreak of COVID-19, the change in working environment from onsite office work to a work from home has rapidly increased cybercrimes. According to Statista [2], there were 5.4 billion malware attacks detected in 2021. The number of malware attacks reached 2.8 billion by mid-2022 over a short period of time. 560,000 new instances of malware are detected every day, contributing to the over 1 billion malware programs already been discovered [3]. From the insights of this data, we can infer that the number of cybercrimes continues to rise upward in 2024, raising concerns for the government and organizations. The cybercriminals are implementing Machine Learning models to automate and increase their capabilities which leads to more sophisticated and adaptive attacks which seems impossible to detect. With the current resources, it seems daunting to detect these types of attacks. As a result, Generative Adversarial Networks (GANs) emerge as a powerful tool to be used to see the unseen cybercriminals malicious behaviors. Generative Adversarial Networks, or GANs, serve as an architecture for training generative models, such as deep convolutional neural networks used for generating images [4]. These networks have emerged as a machine learning model proficient at creating new, previously unseen data samples realistically and synthesizing large datasets based on learned classes and features from an existing dataset [5]. GANs play a key role by assisting in the development of new datasets that replicates real-world cyber threats. These usage of GANs enables researchers and cybersecurity professionals to develop defense system to fight against a diversified possible attack, finally enhancing the security of networks against upcoming cyber threats.

This survey explores different research papers that provides insights regarding the effective detection of malware through the use of the Generative Adversarial Networks (GANs), highlighting the role of GANs in enhancing network security. Along with this, we focused for every individual and provide a clear understanding of the straightforward architecture and operational principles of GANs. Further, general overview of the various types of GANs are discussed, covering both widely accepted models and recent innovations proposed by researchers. In short, this paper will detail the various applications of GANs models in the malware research and network security, highlighting specific areas where GAN research significantly contributes.

2. Background

Reflecting on 2023 it stands out as the most successful year for cybercriminals. One of the prevalent threats

*Corresponding author
E-mail address: kcbisal1@gmail.com
Received: 27/Feb/2024; Accepted: 30/Aug/2024.

to organizations globally targeted 4,368 victims, marking an increase of over 55.5% compared to the previous year. The second and third quarter alone accounted for more victims (2,903) than totaled for the entire year 2022 [6]. These attacks are becoming more and more popular today. Malware is a software program that conducts virtually any behavior malicious attacker wants to perform [1]. These goals include interrupting system operations gained access to the system and network resources and gathering personal information without user consent. In the realm of cybersecurity Generative Adversarial Networks can act as a double-edged sword. MalGAN is a notable example of GANs producing malware. MalGAN was used to generate PE malware effectively bypassing a static PE malware detection engine [7]. This approach represented a direct and forceful attack on the engine as the system was trained to observe the outputs of the model while being aware of the exploratory inputs sent to the engine. On the other hand, generative Adversarial Networks (GANs) serve as a potent and innovative tool we can harness to continually improve detection systems and prevent ransomware attacks. Section 2.1 provides an overview of the fundamental working principles of Generative Adversarial Networks (GANs). In Section 2.2, previous work in the realm of malware detection and network security using GANs is discussed. The techniques for measuring the performance of GANs are outlined in Section 2.3, while Section 2.4 details the datasets utilized in the surveyed research paper. Section 3 introduces various modes or types of GANs employed in the study. Section 4 explores the diverse areas of application for GANs, and Section 5 outlines potential future uses of GANs.

2.1. Overview of GAN working

The Generative Adversarial Networks (GANs) include two neural networks i.e., Generator and discriminator as shown in **Figure 1**, which are in competition with each other, forming a Zero-sum game where one agent's win is at the expense of the other's loss. This approach, a generative model originally presented for the domain of unsupervised learning, has turned out to be useful not only in that domain, but also in semi-supervised learning, fully supervised learning, and reinforcement learning [8-10]. GANs use a training set to diverge from the training set, and are an outstanding tool for classifying different learning paradigms.

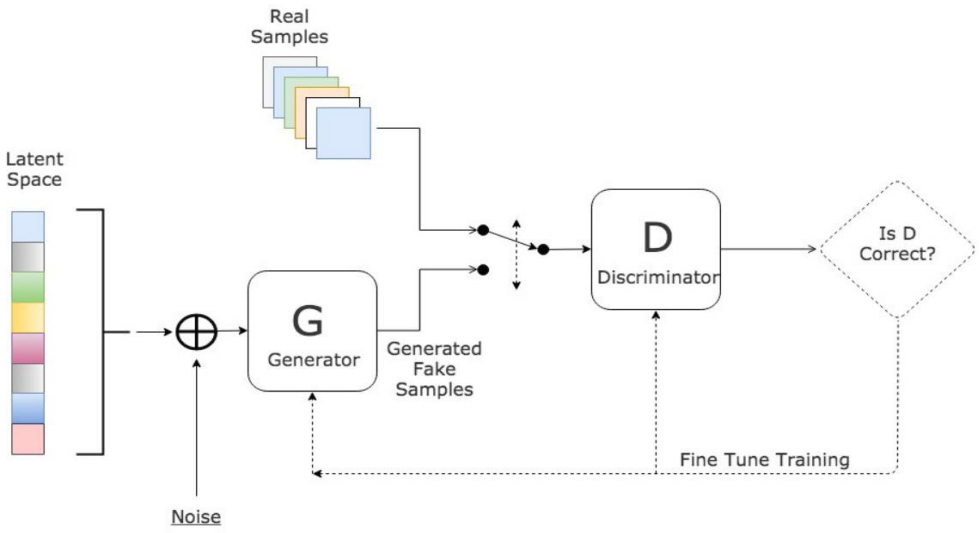

Figure 1. *Generative Adversarial Network, adopted from [11]*

2.1.1. Generator

The generator generates synthetic data that closely mimics the actual data from the training set [12]. It takes noisy signal as a stock input, and encodes it to data that ideally is impossible to distinguish from genuine ones. This process is that of knowledge acquisition through extraction of the underlying patterns and structures from the training data. The generator's goal during the training phase is to make the discriminator think that the fake samples are real.

2.1.2. Discriminator

The discriminator is the binary classifier that classifies data as real and fake [13]. It has a data set, which contains both the original samples and fake ones. Discriminator tries to precisely classify the origin of the input data. With time, the discriminator learns to discern more efficiently between the real and the generated samples

2.1.3. Feedback Loop

The interactions between a generator and discriminator of Generative Adversarial Networks (GANs)

generate a vital feedback loop. This cyclic interaction fosters a growth mindset in both strands of training. Generator improves its skill to produce realistic data to mislead the discriminator while the discriminator is getting better at this task [14]. The Generator trains its weights in response to Discriminator signal, thereby improving the generator's performance. Through this process, the complexity of both networks gets perfected using back-propagation as the mechanism [5]. On the other hand, this kind of adjustment takes place in a black box environment where researchers only control inputs and results while the underlying operations are simulated and calculated based on assumptions. Nonetheless, Arjovsky et al., 2017's approach also studies the mathematical foundations of GANs, building a framework for a deeper understanding of how adversarial training works [15].

2.2. Previous Work

Within the enormous area of application, the utilization of Generative Adversarial Networks (GANs) occupies numerous areas which are capable of classification, generalization, and feature extraction. The GAN models' adaptability allows them to be used in almost any field, and therefore, listing all the applications is a challenging task. Many surveys made attempts to quantify the far-reaching activities of GAN models in their respective domains. In the study we give a brief summary of the different types of GAN models examined in the related works, offering a concise overview of their varied applications.

In Z. Cai et al., 2021 work [16], the authors focus on cybersecurity by investigating application of deep learning methods in different areas. Our paper has the same scope, on the topic of employing the Generative Adversarial Networks (GANs) in the domain of malware research, and consists of both the generation and detection. In the paper of Z. Cai et al., 2021 [16], the machine learning research in cybersecurity is acquainted with the readers first. Their functionalities are explored through the enumeration of different adversarial networks. The author dives into the GAN's privacy and security details, which are discussed through various GAN models that have been designed to protect personal information. Besides, the article does comparative assessment of GAN-Based Mechanisms for Data Privacy Protection. The authors fully delve into the issue of model privacy and compares GAN-Based schemes for model privacy protection. This article focuses on the capabilities of various GANs in cybersecurity domains, illustrating the research results on the model robustness, malware detection, fraud prevention, vehicle security, industrial protocols, and more. Similarly, our research also focuses on how GAN models can be useful in detecting and safeguarding against malware and malicious activities. In further research, this paper recommends advancing GAN-based attack methods by lowering convergence rates, and resolving the mode collapse issue. Corresponding objective functions are defined, and probably approximately correct learning may be used for optimization of data volume. The GAN-based adversarial sample detection should be employed in the practical applications. Android malware detection by statistics, dynamic analysis, and white-box attacks are shown as prospects of development. It is also important to continually improve bioinformatic identification and industrial protection with the help of GANs technologies.

The paper Navidan et al., 2021 [17] reviewed GANs application in the cybersecurity and networking sector. The survey delves into various types of GAN models, categorizing their use cases into five main domains: mobile networks, network analysis, the Internet of Things, the physical layer, and cybersecurity. Remarkably, the authors built an interesting collection of the network related papers written using GANs, and each one was categorized on its merit. Additionally, the paper provides a set of qualitative assessment criteria common for this area, and compares the results of various GANs models on datasets obtained from the reviewed papers. Our research contributes by providing readers with knowledge on variety of GANs and their applications in computer and communication network areas.

Similarly, Dong-Ok Won et al., 2022 [18], address the threat of detecting zero-day malware by developing a system that is able to learn and detect related situations that have been generated to mimic real cases. Proposed PlausMal-GAN model implements generative adversarial networks (GANs) to generate plausible malware images that are visually appealing and unique, making use of a existing malware data. The discriminator, as a detector, gets trained on both genuine and false characteristics. It has a very strong and stable prediction capacity for same group of zero-day-malware images which provides accurate prediction results on average across a different range of representative GANs models. The authors assess both the standard GAN strategy (min-max), least-squares strategy, heuristic approach and a mix of those, as well as for the DCGAN, LSGAN, WGAN and E-GAN models that are included in the proposed architecture. This demonstrates the seeming success of the framework in identifying and anticipating many new resembling zero-day malware instances especially during the testing and updating of malware detection systems.

The primary objective of our survey is to offer readers a comprehensive understanding of the utilization of Generative Adversarial Networks (GANs). In other words, this research paper served as a bridge for those individuals who are interested in GANs and its applications in cybersecurity, particularly in malware and threat-related studies. Our overview provides a detailed examination that distinguishes itself from [5] and [19]. While aligning with these works, our paper explores various approaches, providing a deep analysis of the

current use of GAN models in computer malware research and network security along with the future usage.

Table 1. *Different topics discussed in the surveyed research paper.*

Topics	Z. Cai et al., 2021	Navidan et al., 2021	Dong-Ok Won et al., 2022
Malware Detection	✓	✓	✓
Fraud Detection	✓	✓	
Android Malware		✓	✓
Bioinformatic-Based Recognition	✓		
Android Security		✓	✓
Industry Protocol	✓		
Zero-Day Malware Detection			✓
Black-box API attacks		✓	
Adversarial Examples	✓	✓	✓
Malware Classification			✓
Model Privacy	✓		
Password attack		✓	
Vehicle Security	✓		
Botnet Detection	✓		
Network Intrusion Detection	✓	✓	
Data Privacy	✓		

Table 2. *Different GANS Models discussed in surveyed research paper.*

ANS Model	Z. Cai et al., 2021	Navidan et al., 2021	Dong-Ok Won et al., 2022
inilla GAN	✓	✓	✓
GAN	✓	✓	
CGAN	✓		✓
GAN			✓
GAN	✓	✓	✓
GAN-GP		✓	
GAN			✓
GAN		✓	✓
oGAN	✓	✓	
CGAN	✓		
GAN	✓	✓	✓
CGAN	✓		
oGAN	✓		
gGAN	✓		
cleGAN	✓		

We have cited three major papers that have already worked on Anomaly and Malware Detection. Table 1 presents the findings on various aspects of those surveyed papers that belong to the Anomaly and Malware Detection categories. This table provides a cross tabulation of the areas of concern. Likewise, Table 2 shows the different types of GANs described in those research papers and their uses in Anomaly and Malware Detection. This analysis in that paper demonstrates the different GAN models employed in the field and shows how each type handles the problems associated with the identification of anomalies and malware type. In this way, the presented insights can shape further research in this rapidly evolving field and offer a systematic overview of the current state of knowledge, as found in the existing literature.

2.3 Measuring Performance

The evaluation of the machine learning models is done through the application of a range of conventional metrics. Those statistics are universally acknowledged and are popularly used for evaluation purpose. Generally speaking, the metrics involved would be the Confusion Matrix, Classification Accuracy, Precision, Recall, F1 Score, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) as shown in **Figure 2**.

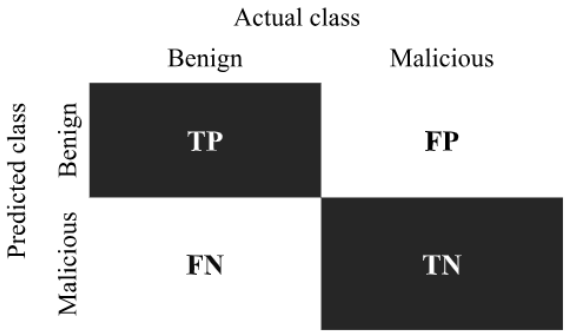

Figure 2. *Performance Metrics for Machine learning Classifications, adopted from* [20]

True Positive, True Positive refers to the number of benign samples being correctly identified as benign samples. True Negative, True Negative refers to the sum of malicious samples being correctly detected as malicious samples. False Positive, False Positive is calculated based on the number of malicious samples being incorrectly identified as benign samples. False Negative, False Negative refers to the number of benign samples that are incorrectly identified as malicious samples.

The classification accuracy of a model is simply measured by comparing the test samples that are correctly identified with total number of test samples. The accuracy of the model is given by:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision is calculated by comparing the number of correctly identified benign samples to the total predicted as benign. It measures the model's accuracy in classifying instances as benign and emphasizes its ability to minimize misclassifications.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

The Recall (True Positive Rate or Sensitivity), represents the ratio between the accurate detection of benign samples and the total number of actual benign samples.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1 score is the harmonic mean of precision and recall, it serves as a composite metric that captures the fundamental balance between precision and recall. The unified evaluation of the model's overall performance is provided by it which is given by,

$$F1\ Score = 2 * \frac{Precison * Recall}{Precision + Recall} \quad (4)$$

The Inception Score is a metric for human evaluation of the quality of image generative models that was developed by Salimans et al. and was published in 2016 [21]. This measure shows a particular good correlation with human evaluations of the generated images in the context of CIFAR-10 dataset. The Inception Score is a measure of a pre-trained Inception v3 Network on the ImageNet dataset, based on the network's outputs applied to generated images [22].

$$IS(G) = \exp(\mathbb{E}_{x \sim p_g} D_{KL}\,(p(y|x) \,||\, p(y))) \quad (5)$$

where $x \sim p_g$ refers that x is an image sampled from p_g, $D_{KL}\,(p \,||\, q))$ is the KL-divergence between the distributions p and q, $p(y|x)$ is the conditional class distribution, and p(y) = $\int_x p(y|x)p_g(x)$ is the marginal class distribution.

The Mode Score is the modified Inception Score [23] to overcome the sampling assessing issues by GAN. The Mode Score differs from the Inception Score in that it disregards the original probabilities. Since the generator results in few samples while the discriminator has more dominance, this score modifies the originating score [24]. Moreover, the Mode Score is an automated alternative to human annotators for quality evaluation of the DNA samples. Further details about Mode Score, including computational process, are

provided in the original article [25].

The Fréchet Inception Distance (FID) [26] which is arguably the most widely used metric for testing the feature similarity between real and fake images has gained much attention. The assessment of GAN resembles the most with the Fréchet Inception Distance (FID) that is analyzed using the Inception V3 model, which is pre-trained on the ImageNet dataset. This FID score was named due to the activation function obtained through the Inception V3 model, providing a tool for classifying how dissimilar those distributions of authentic and fake images are.

Frechet Inception Distance (FID) technique for determining a "multivariate" normal distribution is as follows:

$$FID = || \mu_x - \mu_y ||^2 - Tr(\sum_x + \sum_y - 2\sqrt{\sum_x \sum_y} \quad (6)$$

where X and Y are the real and fake embeddings (activation from the Inception model) assumed to be two multivariate normal distributions. μ_x and μ_y are the magnitudes of the vector X and Y. Tr is the trace of the matrix and $\sum_x$ and $\sum_y$ are the covariance matrix of the vectors.

2.4. Datasets

To uncover various aspects of research, our surveyed paper used some in-depth datasets. CSI data set for wireless sensing and the Microsoft Malware Classification Challenge Dataset were the most important datasets which shaped the results. Here is a summary of the datasets employed in the survey article that we selected.

2.4.1 WIFI RSSI

Wi-Fi RSSI dataset [27] is designed for indoor user detection by signal strength strengthening its use by smart homes, security (finding criminals), and access point user counting. The dataset is used as a tool for building an optimized model for monitoring devices, which can be used to identify user location via Wi-Fi signal strength measurement. Signal strength data from different routers is used, and are classified as a problem of map making. The training of neural networks employs a fuzzy hybrid of Particle Swarm Optimization & Gravitational Search Algorithm (FPSOGSA) to achieve a higher accuracy. The dataset is aimed at applications that are in-house with an emphasis on the most efficient indoor user detection using Wi-Fi RSSI.

2.4.2 CSI

The CSI dataset [28] is an important layer for wireless sensing applications, considering the cases of activity recognition, people identification, and people counting with the support of Wi-Fi connection device. The gathered dataset, using a monitor router with Nexmon CSI, encompasses seven activities, ten clients, and more than 13.5 hours of channel readings. It establishes 242 Wi-Fi OFDM data sub-channels within this 80 MHz band to ensure a better unified foundation for the development of Wi-Fi-based wireless sensing solutions. The related work [29] highlights the possibilities of throughputs from commercial WiFi systems by utilization of Channel State Information (CSI). It concentrates on the active points CSI variations produced by body movement for activity recognition. The approach is represented by the process of feature extraction from CSI data streams and using machine learning techniques, which leads to the development of behavior recognition models.

2.4.3 Network traffic (KDD99)

The dataset known as the KDDCUP'99 dataset, or Network Traffic KDD99 dataset [30], features is especially important because it was created for the purpose of assessing the modern intrusion detection systems. It is a multifaceted exploit that consists of different attacks such as the connection details, traffic stats, system calls, and the user's authentication data. The data set permits the realization of multiclass classification, associating connections with safe and different attack types. Due to its numerous records, it can be used to develop models that would be helpful in security network. This is the reason why it is regarded as a critical data resource.

2.4.4 Deepsig RadioML 2016.10A

The "RADIOML 2016.10A" [31] historical dataset is a synthetic one which was created in GNU Radio programming. This library from Virtual Studio Technology introduced in 2016 contains 11 different types of modulation, 8 of which are digital while the remaining 3 are analog with different signal to noise ratios. These data were first reported at the 6th International Conference on GNU Radio. The 2016.05A data version is an updated and standardized version compared to the 2016.04C dataset, which is no longer available. The dataset may be designed for studies and experimentation in the domain of RF signaling processing and modulation characteristics.

2.4.5 Microsoft Malware Classification Challenge Dataset

The data of the Microsoft Malware Classification [32] has a large repository; its uncompressed size is almost 500 GB. The dataset is composed of nine well-known malware families that are covered by the entire dataset. This combination represents. For each malware file, the 20-bit hash is generated and then the integer is in turn converted into one of the family names. The raw data for every file is shown in a hexadecimal equivalent that represents the binary contents stripping off the headers such that only the analysis of the files is carried out without being biased. Concurrently with the raw data is given a text file, which has been enriched by different data like task names and strings which were obtained with the help of the IDA disassembler tool. The main issue of the participants was the classification of malware into the nine classes that had been already assigned.

2.4.6 Malimg Dataset

Malimg Dataset [35] includes 9,339 byteplot images of malware from 25 families. It is applied as a creative approach for malware visualization and classification based on the image processing [33]. Malware binaries are converted into gray images by utilizing visual resemblance across families.

3. Variant Models of GAN

The reviewed papers provide a quantitative study of different GANs and come with a more detailed analysis of their utilization. This part unravels the essence of GANs, where variety is emphasized and specific features are highlighted.

3.1. CGAN

The traditional GAN [14] do not have a control over the output. The drawback of the traditional Generative Adversarial Network (GAN) is resolved in the Conditioned Generative Adversarial Network (CGAN) [34], in which the model is built to a conditional form. By using the additional information "y" as we can see in figure 3, both the generator and discriminator are connected to more information which is also labeled or comes from other modalities. This extra data is added to both the discriminator and generator as a new input layer. The whole conditioning system gives the model a capability to produce outputs with the specific features or characteristics on the basis of the provided auxiliary information [34].

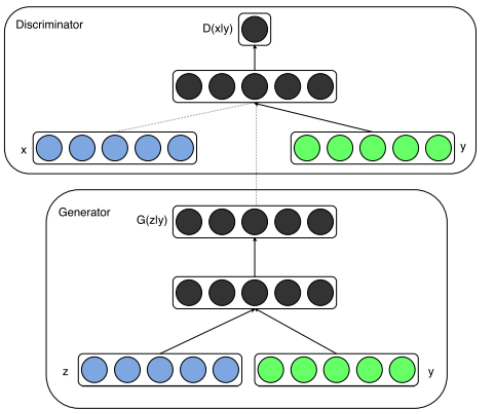

Figure 3. *The Architecture of Conditional GAN* [34]

3.2 DCGAN

DCGANs, which was proposed by Radford [13] and currently is the most popular and best performing GAN architecture that uses convolutional layers, batch normalization, Rectified Linear Unit (ReLU) activations and strided convolutions as represented in figure 4. Attentive modifications include employing convolutional layers instead of fully connected layers and comprising particular architectural rules for both generative and discriminative parts. The intermediate layers in the generator more often use the ReLU activations, while the output layer uses tanh or sigmoid to make sure the pixel values are within the norm for image data.

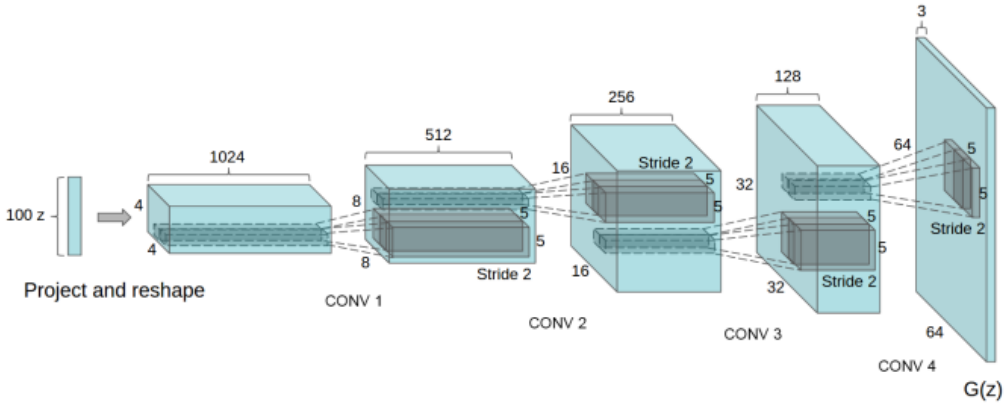

Figure 4. *The Architecture of DCGANs, adopted from* [13]

3.3 InfoGAN

InfoGAN [35], which is an unsupervised learning model, the objective is modified to maximize the mutual information between a subset of noise variables and observed data. This is done by giving the model extra data that contains aspects of the desired features and responding random noise which results to an artificial image creation [36]. In the figure 5, we can see that G(z,c) is generator networks function where z represents a random noise and c represents additional conditional information. This change indeed leads to the revelation of interpretable representations, the performance is profitably manifested on multiple image datasets as well, including MNIST, CelebA, and SVHN. Such method implies that an information cost can be very effective at teaching the generative model to represent patterns that naturally emerge in data, thereby encouraging meaningful and disentangled representations.

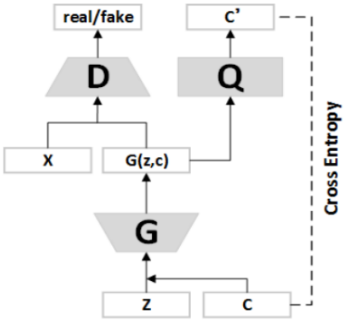

Figure 5. *The Architecture of InfoGAN, adopted from* [35]

3.4 CycleGAN

CycleGAN [38] is an images-to-images translation model that overcomes the challenge of unpaired training data. As we can see in the architecture of CycleGAN in figure 6, it extends Pix2Pix [39] by introducing a cycle consistency loss that helps the model to learn mappings between input and output domains without the need of one-to-one correspondence. Doing so supports numerous applications, for instance, changing SAR imagery into RGB or the other way around (source to target) by means of one model. The structure involves two generators and two discriminators that are trained at the same time, providing a variety of handling image translations with unpaired datasets.

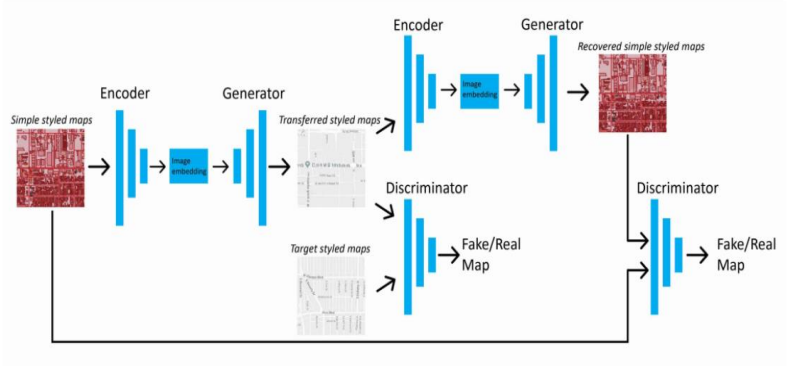

Figure 6. *The Architecture of CycleGAN, adopted from* [37]

3.5 ACGAN

The Auxiliary Classifier Generative Adversarial Network (AC-GAN) is one of the latest models that introduces the specific cost functions to the latent space structure of a traditional GAN in order to increase its capabilities for image resolution tasks [40]. Odena et al., 2016 [40] aimed at classifying the structure of the natural images that would include the down-sampling scheme to extract the basic features. According to Navidan et al., 2021 [17, 41], one of the main differences between ACGAN and CGAN is about setting class labels. While CGAN train generator to be classifier itself, ACGAN employ an additional decoder network to predict the labels. The model's efficacy however did not end there, as it also created malicious code for system interference, as highlighted in [41].

3.6 BiGAN

A crucial step forward in the development of GANs was the Bi-directional Generative Adversarial Network (BiGAN), which Donahue et al. [42] presented in 2017. The BiGAN structure incorporates three main components: a generator (G), a discriminator (D), and an extra encoder (E). They operate in a triangular relationship between each other. The generator does make the fake data, while the discriminator is responsible for the scrutinization of the real and the generated data. The last feature to be discussed is the encoder. It contributes to a bidirectional mapping between the data space and the latent space. This encoder performs not only real data to latent space mapping, but also synthetic data back to the same latent space fitting. Moreover, BiGAN [43] version has been studied further in the field of network intrusion detection. The objective here is to cut down on costs arising from extensive training while at the same time ensuring that the network data is sufficiently monitored by intrusion detectors.

3.7 TGAN

The model Temporal Generative Adversarial Nets (TGAN) [44] transforms the concept of video generation by introducing a dual-generator structure. Different with traditional ways, TGAN employs temporal generator which not only generates one latent variable at time but also changes it for every subsequent video frame. The purpose is to describe how the generator of the image finally synthesizes a complete video from these latent variables. The main goal of the inventive method is to make the grasping of momentary connections more accurate. The TGAN designers used the Wasserstein GAN model and introduced a new end-to-end training approach to ensure robustness during training. The time- and image-generating units show that the TGAN reaches a new level in semantic representations of unlabeled video content and generates lots of diversified and realistic video content. Moreover, Munoz et al., 2020 [45] presented a video generation model in this research which for the first time incorporated both natural spatiotemporal modeling and avoidance of computationally complex 3D architectures.

3.8 LSGAN

Classical GANs apply the sigmoid cross-entropy loss for training the discriminator, whereby it's possible to obtain vanishing gradients. To deal with this issue, LSGANs (the Least Squares Generative Adversarial Networks) is employed in the article [46] and it is applied least squares loss function for the discriminator. The adoption of LSGANs presents two notable advantages over regular GANs: the formation of better-quality images and higher stability in the algorithms learning process. This points out the effectiveness of LSGANs in unsupervised learning and expands their application conditions.

3.9 WGAN

The Wasserstein GAN (WGAN), whose conference paper was published in 2017 [47], introduces a unique approach to solve the problem encountered in the Generative Adversarial Networks (GANs). Unlike the common type of GANs based on the sigmoid cross-entropy loss function, which can be associated with vanishing gradients, WGAN instead relies on the Wasserstein distance, also known as Earth Mover (EM) distance. The main goal of the proposed WGAN architecture is to ensure stable training process, avoid mode collapses, give helpful learning curves for model optimizing (e.g. debugging and hyperparameter tuning). The distance and divergences between our two separate distributions: Pr, Pg ∈ Prog(X), such that Prob(X) is the "space of probability measures defined on X" [47]. We are computing the distance between the two distributions by the Total Variation (TV) distance which is defined as,

$$\delta(\mathbb{P}_r, \mathbb{P}_g) = \sup_{A \in \Sigma} |\mathbb{P}_r(A) - \mathbb{P}_g(A)| \tag{7}$$

The Kullback-Leibler (KL) and the Jensen-Shannon (JS) divergence are given as,

$$\mathrm{KL}\left(\mathbb{P}_r || \mathbb{P}_g\right) = \int \log\left(\frac{\mathrm{P_r(x)}}{\mathrm{P_g(x)}}\right) \mathrm{P_r(x)} d\mu(x) \tag{8}$$

where both $\mathbb{P}_r$ and $\mathbb{P}_g$ are assumed to be absolutely continuous, and therefore admit densities, with respect to the same measure μ defined on X^2The KL divergence is famously asymmetric and possibly infinite when there are points such that P_g (x) = 0 and P_r(x) > 0.

$$JS(\mathbb{P}_r,\mathbb{P}_g) = KL\ (\mathbb{P}_r||\mathbb{P}_m) + KL\ (\mathbb{P}_g||\mathbb{P}_m) \tag{9}$$

Where, $\mathbb{P}_m$ is the mixture ($\mathbb{P}_r$+ $\mathbb{P}_g$)/2. This divergence is symmetrical and always defined because we can choose μ = $\mathbb{P}_m$. The core of WGAN (Wasserstein Generative Adversarial Network) contains EM-distance – a distance generated by the Earth-Mover equation. In contrast to binary evaluations, the EM-distance determines the gradual distance between the produced results and the sought goals. This brings stability and performance to the WGANs.

$$W(\mathbb{P}_r, \mathbb{P}_g) = \inf_{\gamma \in \Pi(\mathbb{P}_r, \mathbb{P}_g)} \mathbb{E}_{(x,y)\sim\gamma}[||x - y||] \tag{10}$$

where Π ($\mathbb{P}_r$, $\mathbb{P}_g$) denotes the set of all joint distributions γ(x, y) whose marginals are respectively $\mathbb{P}_r$ and $\mathbb{P}_g$. Intuitively, γ(x, y) indicates how much "mass" must be transported from x to y in order to transform the distributions $\mathbb{P}_r$ into the distribution $\mathbb{P}_g$. The EM distance is the "cost" of the optimal transport plan.

3.10 WGAN-GP

The Wasserstein Generative Adversarial Network with Gradient Penalty (WGAN-GP) is an extension of the Wasserstein GAN (WGAN) that addresses potential issues related to weight clipping in the original WGAN. Proposed by Gulrajani et al. in 2017 [48], WGAN-GP introduces a gradient penalty term to the loss function as a regularization technique, eliminating the need for weight clipping. This modification improves the stability of the training process and mitigates potential mode collapse issues observed in standard GANs.

The WGAN-GP loss function includes the Wasserstein distance term and the gradient penalty term:

$$L = \mathbb{E}_{\tilde{x}\sim\mathbb{P}_g}[D(\tilde{x})] - \mathbb{E}_{x\sim\mathbb{P}_r}[D(x)] + \lambda \mathbb{E}_{\tilde{x}\sim\mathbb{P}_{\hat{x}}}[(||\nabla_{\hat{x}} D(\hat{x})||_2 - 1)^2] \tag{11}$$

Where, $\mathbb{E}_{\tilde{x}\sim\mathbb{P}_g}[D(\tilde{x})] - \mathbb{E}_{x\sim\mathbb{P}_r}[D(x)]$ is the original critic loss, λ is the gradient penalty coefficient and $\lambda \mathbb{E}_{\tilde{x}\sim\mathbb{P}_{\hat{x}}}[(||\nabla_{\hat{x}} D(\hat{x})||_2 - 1)^2]$ is the gradient penalty.

3.11 E-GAN

Evolutionary Generative Adversarial Networks (EGAN) [49] departs from standard GANs in its incorporation of an evolutionary algorithm. In EGAN, a grouping of producers {G} continues to evolve in the discriminator-defined ecosystem. For every generator in the population, a spot in the solution space of the generative network will be assigned. The evolutionary process involves three stages at each step: Another method, called variation, creates various children from the parents after they randomly mutate; while evaluating them using a fitness function that takes the discriminator as an argument and selecting those who are better. Discriminative network (D) changes continuously, as it is responsible to tell apart the real and generated samples, and it provides an adaptive loss for driving the evolution of the generator population.

$$LD = -\mathbb{E}_{x\sim p_{data}}[log\ D(x)] - \mathbb{E}_{y\sim p_g}[log(1 - D(y))] \tag{12}$$

The fitness function gives feedback about the quality of samples that are generated in the current discriminator environment and this guides the evolution of the population towards having leading solutions.

3.12 BEGAN

The BEGAN model introduced by Berthelot et al. in 2017 [50], is generated by the aim of achieving equilibrium between generator and discriminator dynamics. This model applies a autoencoder-based architecture, including a notion of diversity in the loss function. The equilibrium is kept by ensuring the diversity ratio is maintain, the model can achieve stable and high-quality image. The BEGAN Loss function is given as,

$$LD = L(x) - k_t.\ L(G(z_D)) \qquad \text{for } \theta_D \tag{13}$$

$LG = L(G(z_G))$ for θ_G

$k_{t+1} = k_t + \gamma_k (\gamma L(x) - L(G(z_G)))$ for each training step t

3.13 ProGAN

The ProGAN is a breakthrough in the field of generative adversarial network architectures, achieved by Karras et al. [51] in 2017. Such technique is created by rising the resolution step by step in training corresponding to building more layers of both generator and discriminator networks. The sequential development makes the model to first of all handle major structures in the images distribution and then goes to details. Simultaneous growth of generator and discriminator networks, with carefully tuned addition of new layers, makes the process stable and robust. The ProGAN distinguishes from other models as it is able to generate high-quality images of high-resolution by using modern loss functions like WGAN-GP and LSGAN. This method agrees with the concept of getting easier questions at first while in training, what maintains the model's success.

3.14 MsgGAN

MsgGAN, the Multi-Scale Gradient Adversarial Generative Network, offers a solution providing these GANs with stability to quickly adapt to any new data sets. Per Karnewar & Wang, 2020 [52], MSGGAN resolves the problem of noisy gradients in 2018 by allowing for a multi-scale transfer of gradients from the discriminator to the generator. In this way, it creates an even higher resolution image synthesis that can be used as an effective substitute to standard progressive growing techniques. The MSGGAN protocol performance is stable under different datasets, resolutions, domains, loss functions, and network architecture - exceeding or matching the best GANs in the majority of cases.

4. Areas of Use

The emergence of Generative Adversarial Networks (GANs) has demonstrated that these tools are entirely interdisciplinary with applications in a variety of fields, including cybersecurity. These powerful models are not limited to a singular task; they find a multitude of applications through cybersecurity categories. Here GANs brought in to improve different areas of vulnerability in the system and provide new solutions and insights in this field. Here we demonstrate different areas where GANs have a potential and, by that, explain why these algorithms have such a wide use and why they are so important to the progress of the whole field.

4.1. Anomaly Detection

Anomaly detection based on GANs has been gaining in popularity because attacks are growing ever more complicated and new detection methods are needed. Chandola et al., 2009 [53] systematically covered the field of anomaly detection through an in-depth analysis of challenges, anomaly types, and various detection methods. Specifically, Schlegl et al., 2019 [54] proposed AnoGAN, which is based on the deep convolutional generative adversarial network (DCGAN) [13] and trains a generator and discriminator on the normal data through unsupervised learning. Efficient GAN-Based Anomaly Detection (EGBAD) was proposed by Zenati et al., 2018 [55] which makes use of the BiGAN [42] architecture allowing learning of an encoder mapping input samples to the latent representations during the adversarial training. Akcay et al., 2018 [56] demonstrated GANomaly approach, and Luo et al., 2022 [57] worked on convergence issues using exponential weighted moving average methods and Convolutional Autoencoders. Furthermore, Xia et al., 2022 [58] have summarized the issues, predicted the trends and given the research directions for the future. In the addressing of industrial limitations, Goetz & Humm, 2023 [59] presented an unsupervised, decentralized, real-time anomaly detection concept for cyber-physical production systems using multiple 1D convolutional Autoencoders. Similarly, Lim et al., 2024 [60] carried out a systematic review on GANs and their effectiveness, with suggestions for future research directions. The repeated enhancement of anomaly detection methods that overcomes the limitations while adding more capabilities is the key feature of GANs in anomaly detection.

4.2 Cyber Intrusion and Malware Detection

The application of the Generative Adversarial Networks (GANs) in cybersecurity also extends to different applications such as defending against malicious traffic and improving the capability of Intrusion Detection Systems (IDSs). Reflecting the vulnerability of IDS to adversarial examples [61], Usama et al., 2019 suggested a Generative Adversarial Network-based attack that used the traffic features, making the IDS unable to detect malicious traffic. To this end, a GAN-based defense was applied to enhance the resilience of IDSs against security breaches. In the last method, an innovative method called IDSGAN is proposed, based on WGAN [62], which generates adversarial attacks with the aim of bypassing the detecting system. It does that based on the generated generator, discriminator, and black-box IDSs following an unknown architecture of IDSs in real

environments. This shows the GANs capability in fitting any environment like in-vehicle networks, where the need for high intrusion detection accuracy is a priority for driver's and passengers' safety. Seo et al., 2018 [63] suggested a GIDS (GAN-based IDS) for detecting unknown attacks, and this system was shown to be an excellent real-time intrusion detector with superior performance reported.

Salem et al., 2018 [64] employed GANs as a weapon to solve the skew problem in HIDSs by converting normal data to anomalies that were so successfully dealt with, the HIDSs' performance improved significantly. Likewise, GAN for Android malware identification by Shahpasand et al., 2019 [65] is used for black-box attacks which the adversaries know nothing about the inside of the network. In this case, GANs are the driving force of such attack campaigns in that they provide the attackers with adversarial variations that enable them to manipulate the malware code and consequently avoid detection. In addition, GANs have been instrumental in the study of malware across varied operating systems including Linux and Windows among others. Kargaard et al, 2017 [66] applied GANs in malware detection by transforming malware binaries into image for GAN training. Kim et al. [67, 68] introduced an approach called the transferred deep convolutional generative adversarial network (tDCGAN) for zero-day malware detection. This example of GANs demonstrates the adaptability of GANs to deal with different types of malwares on different operating systems. These applications show a wide range of GANs applications in cybersecurity: from traffic perturbation and intrusion detection up to malware or pattern-lock system security.

4.3. Security Attacks

Generative Adversarial Networks (GANs) are widely implemented in the security domains for the purpose of stimulating and manipulating data. One of the key roles of GANs in cybersecurity is the generation of real but synthetic samples which help in the study of vulnerabilities, penetration testing and the development of defense mechanisms. Imitating the patterns and behaviors, GANs have an ability to strengthen the systems against the possible APTs (Advanced Persistent Threats). The studying by Chowdhary et al., 2023 [69] presents a novel autonomous web application security testing framework based on the Generative Adversarial Networks (GANs) in order to detect vulnerabilities, even the corresponding emerging threats, such as Cross-Site Scripting (XSS) [70] and SQL Injection [71]. The framework uses semantic tokenization for extracting key features for XSS attacks and conditioned sequence GAN for producing attack payloads. Moreover, vulnerable passwords such as those used by Hashcat [72] and JtR [73] can be cracked rapidly, and therefore, represent a great threat to IoT networks. Recent efforts to better the password-cracking efficiency have considered techniques like the Markov model [74] and probabilistic context-free grammar (PCFG) [75]. More significantly, PassGAN has introduced the incorporation of Generative Adversarial Networks (GANs) as a novel approach. The Nam et al., 2020 [76] propose a more advanced model which is constructed on top of the improvements done in PassGAN and PCFG [77]. The Cyber-Physical Production System (CPPS) [77] is vulnerable to cross-domain attacks from the complex interactions between the cyber and physical environments. To deal with this weakness, Chhetri et al., 2019 [78] conceived GAN-Sec that is based on CGAN (conditional generative adversarial network). Through the implementation of GAN-Sec, it is possible to determine whether the key security parameters such as confidentiality, availability and integrity, are appropriately adhered to. Rigaki et al., 2018 [79] propose using GANs that generate network traffic and consequently allow malware to mimic other network patterns by GAN-obtained parameters inclusion in the source code of malware. In conclusion, GANs play an important role in cybersecurity by imitating data patterns, conducting penetration testing, scanning for vulnerabilities, and exposing the security weaknesses in password. This adds to the effectiveness of cybersecurity.

5. Challenges and Future Improvements

Generative Adversarial Network (GAN) has an unparalleled impact on the aspects of contemporary human existence today. In this section, we are discussing about challenges and the improvements of GANs that we can do in the future:

5.1. Transfer learning and generalization capabilities

Generative models, although effective, can encounter difficulty in adapting acquired information to new and diverse conditions which restrict their practical marketability in real-life situations [80]. Scientists are on the lookout for various approaches like transfer learning [81], meta-learning [82], and domain adaptation that could be used to make the generalization capabilities of generative models better. This is very critical for diffusion to different sectors. Contrary to GAN-based anomaly detection models, the main disadvantage is that they face difficulties to adapt to diversified network structures, consequently, leading to their subpar performance across different datasets [81]. The main problem is generalizing the acquired knowledge outside the particular area the system was trained in. For this issue the answer is to let transfer learning skills developed, which should allow the model to acquire knowledge from domain one and apply it successfully in various settings without considerable retraining. In the next experiment, we will have to improve the transfer learning

methods mainly by preventing the negative transfer and exploring unsupervised methods. This would increase the model's robustness, freeing it from dependence on labelled data in such a case. However, the anomaly detection in GAN-assisted anomaly setting remains difficult due to inefficient GAN architectures, vulnerability to adversarial attacks and lower interpretation capability. Research and development work in this field may demonstrate more effective and productive techniques of network anomaly detection.

5.2. Hyperparameter optimization

A little acknowledged subject in discrepancy detection within `GANs is the role of hyperparameter choice on the model performance, reported by Soenen et al., 2021 [83]. Hyperparameter optimization involves fine tuning the different areas of the GAN and tries to tune down learning rates and network architectures, activation functions, and loss functions to come up with an optimal balance for improvement of anomaly detection effectiveness. Future work on improving the GAN-based methods could focus on the new model's architectures and the training strategies, as proposed in the work by Xia et al., 2022 [84]. Moreover, the progression of GAN improvement could be achieved by implementing visualization techniques for attention maps obtained by residual channel attention module [85]. Efforts to strengthen GANs may include a review of aids and techniques and devising methods of correcting the missing values [86]. These fields of study are having a positive impact on the GANs performance in the area of anomaly detection and the cybersecurity sector is taking special interest in understanding the issues of model performance, comprehensibility, and reliability.

5.3. Legal and Regulatory Challenges: Navigating the Legal Landscape

Developing AI techniques at a remarkable speed brings some cyber security issues, especially the intellectual property and liability problems associated with anomaly detection. The generation of content that might be similar to copyrighted material poses issues of ownership and attribution which become the domain of the legal framework in relation to generated content. These consequences are especially important for the cybersecurity community which will face such problems as in the cases of misinformation or malicious use where the role of liability is still vague. Resolving these legal and regulatory issues calls for combined participation of policy makers, legal experts, AI community to suggest rules for ownership and liability. In network security and anomaly detection, the field of integrative generative AI holds huge potential for transformation. Nevertheless, legal and ethical issues should be taken into account as well as the solution of the technical difficulties so that the deployment will be performed in a safe way and to promote innovative solutions in the fight against cyber threats. Through creating a cooperation between academia, industry, and regulatory institutions, an interdisciplinary approach could be achieved by incorporating technological developments, ethical rules, and legal regulations to improve generative AI technologies in the field of cybersecurity.

6. Conclusion

In conclusion, this research work aims to present an extended review of Generative Adversarial Networks (GANs) and their usefulness in malware analysis and computer security. In light of these questions, the present paper provides a good starting point for understanding the potential of GANs in anomaly and malware detection by analysing the current state of GAN fundamentals, finding the deficits in the existing research, and discussing the evaluation metrics. The discussion sheds adequate light to several GAN models – Deep Convolutional GANs (DCGANs), Conditional GANs (CGANs), and Wasserstein GANs (WGANs) while enumerating their distinguishing features and areas of application. GANs hold great potential to improve the working of IDSs, malware identification, and anomaly detection along with producing realistic datasets and emulating attacks. Contrary to the traditional systems, the GANs play a dual role both as a shield against cyber-attacks and as enabler of an attack.

To guide future research, the following recommendations are made: The future work should focus on improvements in various training methods and better tuning of the model for enhancing the developments of GAN. Researchers need to concentrate on collecting better quality sets of data and focusing on issues like stability in training process and adversarial robustness. Further, the proposed idea of incorporating transfer learning and generalization of GANs in various tasks and datasets will enhance the performance of GANs even further. Hyperparameter tuning is also required to enhance the model's accuracy, and, therefore, the ability to generate more accurate outcomes. Further research works have more useful applications of GANs in advanced neural cryptography and better attack representations that are beyond the current cybersecurity measures. Thus, establishing comprehensive knowledge about GANs and their multifaceted usages, researchers can make a huge leap in furthering the developments in network security. This detailed examination also seeks to encourage and guide more research initiatives on the constantly growing cybersecurity domain, specifically in the field of anomaly and malware detection, thus assuring GANs' steady and significant contribution to combating current and future cyber threats.

Declaration of interest: The authors declare that there is no conflict of interest.

Appendix A: List of Abbreviations

Abbreviation	Description
ACGAN	Auxiliary Classifier Generative Adversarial Network
AI	Artificial Intelligence
BEGAN	Boundary Equilibrium Generative Adversarial Network
BiGAN	Bi-directional Generative Adversarial Network
CGAN	Conditioned Generative Adversarial Network
CIFAR-10	Canadian Institute for Advanced Research - 10
CNN	Convolution Neural Network
CPPS	Cyber-Physical Production System
CSI	Channel State Information
CycleGAN	Cycle-Consistent Generative Adversarial Network
DCGAN	Deep Convolutional Generative Adversarial Network
DP-SGD	Differentially-Private Stochastic Gradient Descent
E-GAN	Evolutionary Generative Adversarial Network
EM	Earth Mover
FID	Fréchet Inception Distance
FPSOGSA	Fuzzy hybrid of Particle Swarm Optimization & Gravitational Search Algorithm
GAN	Generative Adversarial Network
GNU	GNU's Not Unix
HIDS	Host-based Intrusion Detection System
IDS	Intrusion Detection System
InfoGAN	Information Maximizing Generative Adversarial Network
JtR	John the Ripper
KDD	Knowledge Discovery in Databases
KL	Kullback Leibler
LSGAN	Least Squares Generative Adversarial Network
MalGAN	Malicious Generative Adversarial Network
ML	Machine Learning
MsgGAN	Multi-Scale Gradient Generative Adversarial Network
OFDM	Orthogonal Frequency Division Multiplexing
PassGAN	Password Generative Adversarial Network
PATE-GAN	Private Aggregation of Teacher Ensembles Generative Adversarial Network
PCFG	Probabilistic Context Free Grammar
PII	Personally Identifiable Information
Pix2Pix	Image-to-Image Translation with Conditional Adversarial Networks
PlausMal-GAN	Plausible Malware Generative Adversarial Network
PPAPNet	Privacy-Preserving Adversarial Protector Network
ProGAN	Progressive Generative Adversarial Network
RadioML	Radio Machine Learning
RSSI	Received Signal Strength Indicator
tanh	hyperbolic tangent
tDCGAN	Transferred Deep-Convolutional Generative Adversarial Network
TGAN	Temporal Generative Adversarial Nets
WGAN	Wasserstein Generative Adversarial Network
WGANGP	Wasserstein Generative Adversarial Network with Gradient Penalty
XSS	Cross-Site Scripting

References

[1] U. Bayer, A. Moser, C. Kruegel and E. Kirda, "Dynamic Analysis of Malicious Code," vol. 2, pp. 66-67, 2006; doi: 10.1007/s11416-006-0012-2.

[2] A. Petrosyan, "Annual number of malware attacks worldwide from 2015 to 2022," 2023. [Online]. Available: https://www.statista.com/statistics/873097/malware-attacks-per-year-worldwide/. [Accessed 2 2 2024].

[3] N. J. Palatty, "30+ Malware Statistics You Need To Know In 2024," Astra, 2023. [Online]. Available: https://www.getastra.com/blog/security-audit/malware-statistics/. [Accessed 2 2 2024].

[4] J. Brownlee, "How to develop an auxiliary classifier GAN (AC-GAN) from scratch with Keras," 2021. [Online]. Available: https://machinelearningmastery.com/how-to-develop-an-auxiliary-classifier-gan-ac-gan-from-s. [Accessed 2 2 2024].

[5] A. Dunmore, J. Jang-Jaccard, F. Sabrina and J. Kwak, "Generative Adversarial Networks for Malware Detection: a Survey," *ArXiv*, vol. abs/2302.08558, 2023; doi: 10.48550/arXiv.2302.08558.

[6] S. Gihon, "Ransomware Trends Q4 2023 Report," 2024. [Online]. Available: https://cyberint.com/blog/research/ransomware-trends-and-statistics-2023-report/. [Accessed 2 2 2024].

[7] W. Hu and Y. Tan, "Generating Adversarial Malware Examples for Black-Box Attacks Based on GAN," *ArXiv*, vol. abs/1702.05983, 2017.

[8] T. Salimans, I. J. Goodfellow, W. Zaremba, V. Cheung, A. Radford and X. Chen, "Improved Techniques for Training GANs," *ArXiv*, vol. abs/1606.03498, 2016.

[9] P. Isola, J.-Y. Zhu, T. Zhou and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5967-5976; doi: 10.1109/CVPR.2017.632.

[10] J. Ho and S. Ermon, "Generative Adversarial Imitation Learning," in *Neural Information Processing Systems*, 2016.

[11] A. Gharakhanian, "Generative Adversarial Networks – Hot Topic in Machine Learning," 2017. [Online]. Available: https://www.kdnuggets.com/2017/01/generative-adversarial-networks-hot-topic-machine-learning.html. [Accessed 2 2 2024].

[12] J. He, Y. Nie and Z. Mao, "Analysis of Image Generation by different Generator in GANs," *Journal of Physics: Conference Series*, vol. 1903, 2021.

[13] A. Radford, L. Metz and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," *CoRR*, vol. abs/1511.06434, 2015.

[14] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville and Y. Bengio, "Generative Adversarial Nets," in *Neural Information Processing Systems*, 2014.

[15] M. Arjovsky and L. Bottou, "Towards Principled Methods for Training Generative Adversarial Networks," *ArXiv*, vol. abs/1701.04862, 2017.

[16] Z. Cai, Z. Xiong, H. Xu, P. Wang, W. Li and Y.-L. Pan, "Generative Adversarial Networks," *ACM Computing Surveys (CSUR)*, vol. 54, pp. 1-38, 2021.

[17] H. Navidan, P. F. Moshiri, M. Nabati, R. Shahbazian, S. A. Ghorashi, V. Shah-Mansouri and D. Windridge, "Generative Adversarial Networks (GANs) in Networking: A Comprehensive Survey & Evaluation," *ArXiv*, vol. abs/2105.04184, 2021.

[18] D.-O. Won, Y.-N. Jang and S.-W. Lee, "PlausMal-GAN: Plausible Malware Training Based on Generative Adversarial Networks for Analogous Zero-Day Malware Detection," *IEEE Transactions on Emerging Topics in Computing*, vol. 11, pp. 82-94, 2023; doi: 10.1109/TETC.2022.3170544.

[19] I. K. Dutta, B. Ghosh, A. H. Carlson, M. W. Totaro and M. A. Bayoumi, "Generative Adversarial Networks in Security: A Survey," *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0399-0405, 2020.

[20] M. K. Prabakaran, P. M. Sundaram and A. D. Chandrasekar, "An enhanced deep learning-based phishing detection mechanism to effectively identify malicious URLs using variational autoencoders," *IET Information Security*, vol. 17, no. 3, pp. 315-551, 2023; doi: 10.1049/ise2.12106.

[21] T. Salimans, I. J. Goodfellow, W. Zaremba, V. Cheung, A. Radford and X. Chen, "Improved Techniques for Training GANs," *ArXiv*, vol. abs/1606.03498, 2016.

[22] S. T. Barratt and S. Rishi, "A Note on the Inception Score," *ArXiv*, vol. abs/1801.01973, 2018.

[23] A. Borji, "Pros and Cons of GAN Evaluation Measures," *ArXiv*, vol. abs/1802.03446, 2018.

[24] A. Dunmore, J. Jang-Jaccard, F. Sabrina and J. Kwak, "A Comprehensive Survey of Generative Adversarial Networks (GANs) in Cybersecurity Intrusion Detection," *IEEE Access*, vol. 11, pp. 76071-76094, 2023; doi: 10.1109/ACCESS.2023.3296707.

[25] T. Che, Y. Li, A. P. Jacob, Y. Bengio and W. Li, "Mode Regularized Generative Adversarial Networks," *ArXiv*, vol. abs/1612.02136, 2017.

[26] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818-2826; doi: 10.1109/CVPR.2016.308.

[27] J. G. Rohra, B. Perumal, S. J. Narayanan, P. Thakur and R. B. Bhatt, "User Localization in an Indoor Environment Using Fuzzy Hybrid of Particle Swarm Optimization & Gravitational Search Algorithm with Neural Networks," in *International Conference on Soft Computing for Problem Solving*, 2016.

[28] F. Meneghello, N. D. Fabbro, D. Garlisi, I. Tinnirello and M. Rossi, "A CSI Dataset for Wireless Human Sensing on 80 MHz Wi-Fi Channels," *IEEE Communications Magazine*, vol. 61, pp. 146-152, 2023.

[29] S. Yousefi, N. Hirokazu, S. Dayal, S. Ermon and S. Valaee, "A Survey on Behavior Recognition Using WiFi Channel State Information," *IEEE Communications Magazine*, vol. 55, pp. 98-104, 2017.

[30] "KDD Cup 1999 Data," 1999. [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. [Accessed 2 2 2024].

[31] "Historical Dataset: RADIOML 2016.10A.," 2016. [Online]. Available: https://www.deepsig.ai/datasets. [Accessed 2 2 2024].

[32] R. Ronen, M. Radu, C. Feuerstein, E. Yom-To and M. Ahmadi, "Microsoft Malware Classification Challenge," *ArXiv*, vol. abs/1802.10135, 2018.

[33] L. Nataraj, S. Karthikeyan, G. Jacobe and B. S. Manjunath, "Malware images: visualization and automatic classification," in *Visualization for Computer Security*, 2011.

[34] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," *ArXiv*, vol. abs/1411.1784, 2014.

[35] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever and P. Abbeel, "InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets," in *Neural Information Processing Systems*, 2016.

[36] X. Li, L. Chen, L. Wang, P. Wu and W. Tong, "SCGAN: Disentangled Representation Learning by Adding Similarity Constraint on Generative Adversarial Nets," *IEEE Access*, vol. 7, pp. 147928-147938, 2019.

[37] "How CycleGAN Works?," [Online]. Available: https://developers.arcgis.com/python/guide/how-cyclegan-works/. [Accessed 2 2 2024].

[38] J.-Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2242-2251, 2017.

[39] P. Isola, J.-Y. Zhu, T. Zhou and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967-5976, 2016; doi: 10.1109/CVPR.2017.632.

[40] A. Odena, C. Olah and J. Shlens, "Conditional Image Synthesis with Auxiliary Classifier GANs," in *International Conference on Machine Learning*, 2016.

[41] R. Nagaraju and M. Stamp, "Auxiliary-Classifier GAN for Malware Analysis," *ArXiv*, vol. abs/2107.01620, 2021.

[42] J. Donahue, P. Krähenbühl and T. Darrell, "Adversarial Feature Learning," *ArXiv*, vol. abs/1605.09782, 2016.

[43] W. Xu, J. Jang-Jaccard, T. Liu, F. Sabrina and J. Kwak, "Improved bidirectional GAN-based approach for network intrusion detection using one-class classifier," *MDPI (Basel, Switzerland)*, 2022; doi: 10.3390/computers11060085.

[44] M. Saito, E. Matsumoto and S. Saito, "Temporal Generative Adversarial Nets with Singular Value Clipping," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2849-2858; doi: 10.1109/ICCV.2017.308.

[45] A. Munoz, M. Zolfaghari, M. Argus and T. Brox, "Multi-Variate Temporal GAN for Large Scale Video Generation," *ArXiv*, vol. abs/2004.01823, 2020.

[46] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang and S. P. Smolley, "Least Squares Generative Adversarial Networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2813-2821; doi: 10.1109/ICCV.2017.304.

[47] M. Arjovsky, S. Chintala and L. Bottou, "Wasserstein GAN," *ArXiv*, vol. abs/1701.07875.

[48] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. C. Courville, "Improved Training of Wasserstein GANs," in *Neural Information Processing Systems*, 2017.

[49] C. Wang, C. Xu, X. Yao and D. Tao, "Evolutionary Generative Adversarial Networks," *IEEE Transactions on Evolutionary Computation*, vol. 23, pp. 921-934, 2019; doi: 10.1109/TEVC.2019.2895748.

[50] D. Berthelot, T. Schumm and L. Metz, "BEGAN: Boundary Equilibrium Generative Adversarial Networks," *ArXiv*, vol. abs/1703.10717, 2017.

[51] T. Karras, T. Aila, S. Laine and J. Lehtinen, "Progressive Growing of GANs for Improved Quality, Stability, and Variation," *ArXiv*, vol. abs/1710.10196, 2017.

[52] A. Karnewar, O. Wang and R. S. Iyengar, "MSG-GAN: Multi-Scale Gradient GAN for Stable Image Synthesis," *ArXiv*, vol. abs/1903.06048, 2019.

[53] V. Chandola, A. Banerjee and K. Vipin, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1-15:58, 2009.

[54] T. Schlegl, P. Seeböck, S. Waldstein, G. Lang and U. Schmidt-Erfurth, "f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," *Medical Image Analysis*, vol. 54, pp. 30-44, 2019.

[55] C.-S. Houssam Zenati and Foo, B. Lecouat, G. Manek and V. R. Chandrasekhar, "Efficient GAN-Based Anomaly Detection," *ArXiv*, vol. abs/1802.06222, 2018.

[56] D. Akçay and B. D. Akçay, "Effect of media content and media use habits on aggressive behaviors in the adolescents," *The European Research Journal,* vol. 5, no. 3, 2019; doi: 10.18621/eurj.395892.

[57] X. Luo, Y. Jiang, E. Wang and X. Men, "Anomaly detection by using a combination of generative adversarial networks and convolutional autoencoders," *EURASIP Journal on Advances in Signal Processing,* vol. 2022, pp. 1-13, 2022.

[58] X. Xia, X. Pan, N. Li, X. He, L. Ma, X. Zhang and N. Ding, "GAN-based anomaly detection: A review," *Neurocomputing 493,* 2022; dio: 10.1016/j.neucom.2021.12.093.

[59] C. Goetz and B. Humm, "Decentralized Real-Time Anomaly Detection in Cyber-Physical Production Systems under Industry Constraints," *Artificial Intelligence Enhanced Health Monitoring and Diagnostics),* vol. 23, no. 9, p. 4207, 2023; doi: 10.3390/s23094207.

[60] W. Lim, S. K. C. Y. Sheng and B. T. T. C. C. L. Lau, "Future of generative adversarial networks (GAN) for anomaly detection in network security: A review," *Computers & Security,* 2024.

[61] M. Usama, M. Asim, S. Latif, J. Qadir and Ala-Al-Fuqaha, "Generative Adversarial Networks For Launching and Thwarting Adversarial Attacks on Network Intrusion Detection Systems," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC),* 2019, pp. 78-83; doi: 10.1109/IWCMC.2019.8766353.

[62] L. Zilong, Y.-y. Shi and X. Zhi, "IDSGAN: Generative Adversarial Networks for Attack Generation against Intrusion Detection," *ArXiv,* vol. abs/1809.02077, 2018.

[63] E. Seo, H. M. Song and H. K. Kim, "GIDS: GAN based Intrusion Detection System for In-Vehicle Network," in *2018 16th Annual Conference on Privacy, Security and Trust (PST),* 2018, pp. 1-6; doi: 10.1109/PST.2018.8514157.

[64] M. A. Salem, S. Taheri and J.-S. Yuan, "Anomaly Generation Using Generative Adversarial Networks in Host-Based Intrusion Detection," *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON),* pp. 683-687, 2018.

[65] M. Shahpasand, L. Hamey, D. Vatsalan and M. Xue, "Adversarial Attacks on Mobile Malware Detection," in *2019 IEEE 1st International Workshop on Artificial Intelligence for Mobile (AI4Mobile)*, 2019, pp. 17-20; doi: 10.1109/AI4Mobile.2019.8672711.

[66] J. Kargaard, T. Drange, A.-L. Kor, H. Twafik and E. Butterfield, "Defending IT systems against intelligent malware," in *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 2018, pp. 411-417; doi: 10.1109/DESSERT.2018.8409169.

[67] J.-Y. Kim, S.-J. Bu and C. Sung-Bae, "Malware Detection Using Deep Transferred Generative Adversarial Networks," in *International Conference on Neural Information Processing*, 2017.

[68] J.-Y. Kim, S.-J. Bu and S.-B. Cho, "Zero-day malware detection using transferred generative adversarial networks based on deep autoencoders," *Inf. Sci.,* Vols. 460-461, pp. 83-102, 2018.

[69] A. Chowdhary, K. Jha and M. Zhao, "Generative Adversarial Network (GAN)-Based Autonomous Penetration Testing for Web Applications," *Sensors (Basel, Switzerland),* vol. 23, 2023.

[70] M. Singh, P. Singh and P. Kumar, "An Analytical Study on Cross-Site Scripting," in *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*, 2020, pp. 1-6.

[71] R. Shobana and M. Suriakala, "A Thorough Study On Sql Injection Attack-Detection And Prevention Techniques And Research Issues," *Journal of Information and Computational Science,* vol. 10, no. 5, 2020.

[72] C. Binnie, "Password Cracking with Hashcat," in *Linux Server Security: Hack and Defend*, 2016; doi: 10.1002/9781119283096.ch9.

[73] K. Marchetti and P. Bodily, "John the Ripper: An Examination and Analysis of the Popular Hash Cracking Algorithm," in *2022 Intermountain Engineering, Technology and Computing (IETC)*, 2022, pp. 1-6. doi: 10.1109/IETC54973.2022.9796671; doi: 10.1109/IETC54973.2022.9796671

[74] M. Kuperberg, "Markov Models," in *Dependability Metrics*, 2005, pp. 48-55. doi: 10.1007/978-3-540-68947-8_8.

[75] Z. Chi and S. Geman, "Estimation of Probabilistic Context-Free Grammars," *Computational Linguistics,* vol. 24, no. 2, pp. 298-305, 1998.

[76] S. Nam, S. Jeon, H. Kim and J. Moon, "Recurrent GANs Password Cracker For IoT Password Security Enhancement †," *Sensors (Basel, Switzerland),* vol. 20, 2022.

[77] L. Monostori, "Cyber-physical production systems: roots from manufacturing science and technology," *at - Automatisierungstechnik,* vol. 63, pp. 766-776, 2015.

[78] S. R. Chhetri, A. B. Lopez, J. Wan and M. A. Al Faruque, "GAN-Sec: Generative Adversarial Network Modeling for the Security Analysis of Cyber-Physical Production Systems," in *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2018, pp. 770-775; doi: 10.23919/DATE.2019.8715283.

[79] M. Rigaki and S. Garcia, "Bringing a GAN to a Knife-Fight: Adapting Malware Communication to Avoid Detection," in *2018 IEEE Security and Privacy Workshops (SPW)*, 2018, pp. 70-75; doi: 10.1109/SPW.2018.00019.

[80] D. Saxena and J. Cao, "Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions," *ACM Computing Surveys,* vol. 54, no. 3, pp. 1-42, 2021; doi: 10.1145/3446374

[81] J. Li, Y. Liu and L. Qijie, "Generative Adversarial Network and Transfer Learning Based Fault Detection for Rotating Machinery with Imbalance Data Condition," *Measurement Science and Technology*, vol. 33, no. 4, 2022; doi: 10.1088/1361-6501/ac3945.

[82] A. Yang, C. Lu, J. Li, X. Huang, T. Ji, X. Li and Y. Sheng, "Application of meta-learning in cyberspace security: a survey," *Digital Communications and Networks*, vol. 9, no. 1, pp. 67-78, 2023; doi: 10.1016/j.dcan.2022.03.007.

[83] J. Soenen, K. Leuven, E. V. Wolputte, L. Perini, V. Vercruyssen, W. Meert, J. Davis and H. Blockeel, "The Effect of Hyperparameter Tuning on the Comparative Evaluation of Unsupervised Anomaly Detection Methods," 2021.

[84] X. Xuan, X. Pan, N. Li, X. He, L. Ma, X. Zhang and N. Ding, "GAN-based anomaly detection: A review," *Neurocomputing*, vol. 493, pp. 497-535, 2022; dio: 10.1016/j.neucom.2021.12.093.

[85] Z. Dehghanian, S. Saravani, M. Amirmazlaghani and M. Rahmati, "Spot The Odd One Out: Regularized Complete Cycle Consistent Anomaly Detector GAN," 2023.

[86] J. Fu, W. Lina, J. Ke, K. Yang and R. Yu, "GANAD:A GAN-based method for network anomaly detection," 2023; doi: 10.21203/rs.3.rs-2081269/v1.