International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

Harnessing GANs for Smarter Security: A Comprehensive Review of Generative Adversarial Networks in Autonomous Intrusion Detection Systems

Amarjeet Srivastava[1], Shiwangi Choudhary[2]
Dept. of Computer Science and Engineering,
Rameshwaram Institute of Technology & Management, (AKTU), Lucknow, India

Abstract— **In an era of rapidly evolving cyber threats, traditional intrusion detection systems (IDS) often fall short in adapting to novel attack patterns. Generative Adversarial Networks (GANs), a class of deep learning models known for their ability to generate synthetic yet realistic data, have emerged as a promising solution for enhancing the intelligence and adaptability of autonomous intrusion detection systems. This comprehensive review explores the integration of GANs in IDS frameworks, focusing on their potential to generate adversarial samples, detect zero-day attacks, and improve anomaly detection accuracy. The paper critically analyzes various GAN architectures—including standard GANs, conditional GANs, and CycleGANs—highlighting their applications in both training data augmentation and adversarial robustness. It also examines the challenges associated with GAN training instability, data privacy concerns, and real-time deployment in cybersecurity environments. Through an in-depth comparison of state-of-the-art research contributions, this review outlines the current trends, limitations, and future directions of GAN-driven autonomous intrusion detection systems. The goal is to provide researchers and practitioners with a consolidated foundation for further innovation in smarter, self-adaptive, and resilient cybersecurity solutions.**

Keywords— **Generative Adversarial Networks (GANs), Intrusion Detection Systems (IDS), Cybersecurity, Anomaly Detection, Adversarial Machine Learning, Deep Learning, Autonomous Security, Data Augmentation, Zero-Day Attacks, Network Security.**

I. INTRODUCTION

As digital infrastructures become increasingly complex and interconnected, ensuring robust cybersecurity has become a critical concern for organizations worldwide. Traditional Intrusion Detection Systems (IDS) primarily rely on static rule-based or signature-based mechanisms that are often inadequate against sophisticated, evolving cyber threats such as zero-day attacks and polymorphic malware [1][2]. To overcome these limitations, the research community has increasingly turned to Artificial Intelligence (AI) and, more specifically, Deep Learning (DL) techniques to build more adaptive and intelligent security systems [3].

Generative Adversarial Networks (GANs), introduced by Goodfellow et al. in 2014 [4], have shown remarkable success in various domains such as image synthesis, data augmentation, and anomaly detection due to their ability to learn and replicate complex data distributions. Comprising two neural networks—a generator and a discriminator—GANs operate in a minimax game setting where the generator attempts to produce data indistinguishable from real data, while the discriminator works to distinguish between real and synthetic samples. This adversarial training process results in highly realistic data generation capabilities, which can be harnessed for multiple applications in cybersecurity, particularly in intrusion detection [5].

In the context of IDS, GANs offer several advantages. They can generate high-quality synthetic attack data to augment scarce datasets, thereby improving model generalization and performance [6]. Furthermore, GANs have been leveraged to detect anomalies by learning the normal behavior of a system and flagging deviations that may indicate intrusions [7]. Emerging GAN variants such as Conditional GANs (cGANs) and Wasserstein GANs (WGANs) further enhance training stability and adaptability, making them suitable for real-time security applications [8][9].

Despite their potential, deploying GANs in IDS is not without challenges. Issues such as training instability, mode collapse, and the risk of generating adversarial attacks necessitate careful design and evaluation [10]. Additionally, real-time implementation and scalability in high-throughput network environments remain significant concerns [11].

This review presents a comprehensive analysis of the integration of GANs in autonomous intrusion detection systems. It categorizes and evaluates current research efforts, highlights the strengths and weaknesses of different approaches, and identifies key areas for future exploration. By systematically examining the state-of-the-art, this work aims to provide a solid foundation for developing smarter and more resilient security systems.

II. LITERATURE SURVEY

The integration of Generative Adversarial Networks (GANs) into Intrusion Detection Systems (IDS) has emerged as a promising direction in the cybersecurity landscape. Researchers have explored multiple GAN-based architectures to address issues such as data scarcity, class imbalance, anomaly detection, and the generation of adversarial examples, each contributing uniquely to the evolution of autonomous IDS.

Lin et al. (2019) proposed IDSGAN, a framework that generates adversarial network traffic to test and strengthen the robustness of IDS models [1]. Their approach revealed vulnerabilities in machine learning-based IDS, suggesting the need for more resilient defensive mechanisms. Similarly, Ring et al. (2019) evaluated various GAN architectures for generating synthetic intrusion data and emphasized the importance of architecture selection for enhancing IDS datasets [2].

To tackle the problem of data imbalance, Shafiq et al. (2021) introduced a GAN-based augmentation approach that improved classifier performance on minority classes, particularly rare attack types [3]. This method proved beneficial for training deep neural networks in scenarios where labeled attack data is limited. A similar data augmentation approach was implemented by Li et al. (2020), who utilized Conditional GANs (cGANs) to generate class-specific samples, significantly enhancing detection accuracy [4].

Anomaly detection using GANs has also gained traction. Kim et al. (2019) proposed a GAN-based anomaly detector that models the distribution of normal network behavior and flags deviations as potential intrusions [5]. Their method was effective against unknown attack types and zero-day threats. Likewise, Zhou et al. (2021) developed a semi-supervised GAN that learned both labeled and unlabeled network traffic patterns for more comprehensive detection [6].

Several researchers have examined GAN stability and its impact on IDS performance. Arjovsky et al. (2017) introduced Wasserstein GANs (WGANs), which improved training stability by optimizing the Earth-Mover distance, thereby producing higher-quality samples for IDS training [7]. Mirsky et al. (2018) demonstrated how an ensemble of autoencoders (as in Kitsune) could benefit from synthetic data generated via GANs for enhanced online intrusion detection [8].

For real-time IDS applications, efficient GAN variants have been explored. Radford et al. (2016) introduced Deep Convolutional GANs (DCGANs), which utilize convolutional layers for faster and more accurate data generation in high-dimensional spaces [9]. These were adapted by subsequent researchers for streaming data analysis in cybersecurity environments.

Adversarial robustness has also been addressed in recent studies. Mao et al. (2022) used GANs not just for generating attacks but also for hardening models against adversarial examples by training them on synthetic malicious traffic [10]. This dual role—offensive and defensive—highlights GANs' versatility in the security domain.

Despite these advances, challenges remain. Training instability, mode collapse, and the risk of generating indistinguishable malicious traffic have been cited as critical issues [11]. Furthermore, issues of privacy and the ethical use of synthetic data in cybersecurity warrant ongoing investigation [12].

Collectively, these studies underscore the growing relevance of GANs in transforming IDS into smarter, more autonomous systems. Future work must focus on optimizing GAN architectures, addressing training challenges, and ensuring practical scalability in real-world networks.

TABLE 1: LITERATURE REVIEW TABLE BASED ON PREVIOUS YEAR RESEARCH PAPER KEY FINDINGS

S.No.	Author(s)	Year	Title	Methodology/ Model Used	Key Findings
1	Lin et al.	2019	IDSGAN: Generative Adversarial Networks for Attack Generation Against Intrusion Detection	GAN	Demonstrated GAN's ability to generate adversarial traffic that evades detection.
2	Shafiq et al.	2021	Data Augmentation Using GANs for Improved Detection of Network Intrusions	GAN-based data augmentation	Improved classifier accuracy on imbalanced datasets.
3	Li et al.	2020	Enhancing Intrusion Detection Using Conditional GANs	cGAN	Generated class-specific samples to boost IDS accuracy.
4	Kim et al.	2019	Anomaly Detection in CPS Using GANs	GAN	Detected unknown anomalies based on deviations from normal patterns.
5	Zhou et al.	2021	Semi-supervised IDS Based on GAN	Semi-supervised GAN	Effectively leveraged both labeled and unlabeled data.
6	Arjovsky et al.	2017	Wasserstein GAN	WGAN	Introduced training stability, useful for IDS data generation.
7	Radford et al.	2016	DCGAN for	DCGAN	Enabled high-

			Representation Learning		quality data generation for high-dimensional IDS datasets.
8	Mirsky et al.	2018	Kitsune: Online Network IDS	Autoencoders with GAN-augmented data	Improved online anomaly detection.
9	Mao et al.	2022	Adversarial Training with GAN-generated Attacks	GAN with adversarial training	Enhanced IDS robustness to adversarial traffic.
10	Luo & Nagarajan	2020	GAN-based Anomaly Detection for Cybersecurity	GAN	Showed potential for GANs in anomaly-based threat detection.
11	Hu et al.	2021	Privacy Risks in GAN-Generated Datasets	Survey	Raised concerns on privacy in synthetic data generation.
12	Roy et al.	2023	Survey on GANs for Network Security	Review	Provided taxonomy and trend analysis in GAN-based security.
13	Wang et al.	2021	GAN-based Intrusion Detection in IoT	GAN + CNN	Improved performance in IoT networks with limited data.
14	Huang et al.	2020	Synthetic Traffic Generation for IDS	GAN	Generated realistic traffic data for IDS model training.
15	Feng et al.	2022	Improving IDS via GAN-generated Samples	GAN	Balanced training data and enhanced detection rates.
16	Su et al.	2019	Defending IDS from GAN-based Attacks	Defensive GAN	Proposed adversarial-resistant model.
17	Nguyen et al.	2020	GAN for Intrusion Detection in Smart Grids	GAN	Provided early anomaly detection in energy networks.
18	Zhang et al.	2021	Conditional GAN for IoT Intrusion Detection	cGAN	Addressed attack classification in IoT environments.
19	Yao et al.	2021	GAN-enhanced Feature Selection for IDS	GAN + feature engineering	Boosted IDS performance through synthetic feature generation.
20	Reddy et al.	2020	Benchmarking GANs for IDS Dataset Generation	Comparative GAN architectures	Evaluated multiple GANs for IDS data quality.
21	Singh et al.	2022	Hybrid GAN for Cloud Security	Hybrid GAN	Increased detection in cloud-based IDS systems.
22	Goyal et al.	2020	GAN-based IDS for Smart City Applications	GAN	Used for smart city security against cyber threats.
23	Chen et al.	2019	TrafficGAN: Generative Model for IDS Simulation	TrafficGAN	Created realistic network behavior for testing IDS.
24	Wei et al.	2020	IDS Performance Enhancement with GAN-Generated	GAN	Reduced false positive rate and improved detection rate.

			Data		
25	Islam et al.	2023	Explainable GAN-IDS Model	XAI + GAN	Integrated explainability into GAN-based IDS decision-making.

III. INTRUSION DETECTION SYSTEMS

An Intrusion Detection System (IDS) is a cybersecurity mechanism designed to monitor and analyze network or system activities for signs of malicious behavior, policy violations, or security breaches. IDS plays a critical role in modern security infrastructures by detecting both known and unknown threats, alerting system administrators, and providing detailed logs for incident analysis.

Types of Intrusion Detection Systems

A. **Network-based IDS (NIDS):**
- Monitors traffic across entire networks.
- Positioned at key points like routers or gateways.
- Ideal for detecting widespread or external attacks.
- Example tools: Snort, Suricata.

B. **Host-based IDS (HIDS):**
- Installed on individual systems or servers.
- Monitors OS-level events like file changes, logs, and system calls.
- Better for detecting insider threats or local compromise.
- Example tools: OSSEC, Tripwire.

C. **Hybrid IDS:**
- Combines features of NIDS and HIDS.
- Offers a more comprehensive security coverage.
- Provides both network-level and host-level insights.

D. **Signature-based IDS:**
- Detects threats based on known patterns or signatures.
- Effective against known malware and attack vectors.
- Limitations: Fails to detect novel or zero-day attacks.

E. **Anomaly-based IDS:**
- Uses machine learning/statistical models to identify deviations from normal behavior.
- Effective against unknown and emerging threats.
- May produce higher false positives compared to signature-based systems.

F. **Core Components of an IDS**
- Data Collector: Gathers system/network traffic or audit logs.
- Detection Engine: Analyzes collected data to identify suspicious activity.
- Knowledge Base: Maintains rules, heuristics, or models used for threat detection.
- Response Mechanism: Generates alerts, logs incidents, or triggers automated responses.

G. **Key Functions and Benefits**
- Real-time Monitoring: IDS provides continuous surveillance of digital assets.
- Early Threat Detection: Identifies attacks before they escalate.
- Forensic Analysis: Logs events for post-incident investigation.
- Regulatory Compliance: Helps meet standards like PCI-DSS, HIPAA, and ISO 27001.
- Enhanced Security Posture: Strengthens network defense layers and reduces attack surfaces.

H. **Challenges in IDS**
- High False Positive Rate: Especially in anomaly-based systems.
- Resource Intensive: Requires computational and human resources for effective monitoring.
- Encrypted Traffic Analysis: Detection is harder in environments with end-to-end encryption.
- Evasion Techniques: Skilled attackers may use stealth tactics to bypass detection.

I. **Advancements and Trends**
- Machine Learning & Deep Learning: Used to improve anomaly detection capabilities.
- Generative Adversarial Networks (GANs): Employed to generate synthetic data or adversarial attacks to improve IDS robustness.
- Edge and Cloud-based IDS: Offering scalable and distributed detection.
- Behavioral IDS: Focuses on user behavior analytics (UBA) to detect insider threats.

Intrusion Detection Systems remain a foundational element in cybersecurity defense. With evolving attack surfaces and increasingly sophisticated threat actors, modern IDS solutions are being enhanced using artificial intelligence, automation, and adaptive analytics. The integration of GANs and deep learning models is pushing the boundaries of traditional IDS, enabling smarter and more autonomous threat detection systems.

IV. CONCLUSION

Intrusion Detection Systems (IDS) are indispensable components in the cybersecurity framework, playing a pivotal role in identifying, mitigating, and preventing malicious activities across networks and host systems. As cyber threats grow in complexity and frequency, traditional IDS approaches often struggle with high false positives, limited adaptability to novel attacks, and difficulties in handling imbalanced datasets.

The integration of Generative Adversarial Networks (GANs) into IDS represents a transformative shift toward smarter, more autonomous security solutions. GANs can generate realistic synthetic data to address data scarcity, simulate adversarial attacks to enhance system resilience, and support anomaly detection by learning complex data distributions. These

capabilities significantly enhance the performance, adaptability, and robustness of IDS, making them more effective in detecting both known and unknown intrusions.

However, deploying GANs in IDS also brings challenges such as model instability, training complexity, and potential misuse for generating adversarial traffic. Therefore, continued research, optimization of GAN architectures, and the incorporation of explainability and interpretability are essential for ensuring trust and reliability in GAN-based IDS.

In conclusion, harnessing GANs for intrusion detection holds immense promise for the future of cybersecurity, offering innovative paths toward building intelligent, adaptive, and proactive defense mechanisms capable of safeguarding increasingly complex digital ecosystems.

REFERENCES

[1] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, “Building an intrusion detection system using a filter-based feature selection algorithm,” IEEE Transactions on Computers, vol. 65, no. 10, pp. 2986–2998, 2016.
[2] K. Kendall, “A database of computer attacks for the evaluation of intrusion detection systems,” MIT Thesis, 1999.
[3] R. Sommer and V. Paxson, “Outside the closed world: On using machine learning for network intrusion detection,” in IEEE Symposium on Security and Privacy, 2010, pp. 305–316.
[4] I. Goodfellow et al., “Generative adversarial nets,” in Advances in Neural Information Processing Systems (NeurIPS), 2014, pp. 2672–2680.
[5] M. Lin, C. Dou, and Y. Zhou, “IDSGAN: Generative adversarial networks for attack generation against intrusion detection,” in IEEE Access, vol. 7, pp. 160288–160298, 2019.
[6] A. Shafiq, M. A. Shah, A. Wahid, et al., “Data augmentation using GANs for improved detection of network intrusions,” Future Generation Computer Systems, vol. 118, pp. 87–99, 2021.
[7] Y. Luo and Y. Nagarajan, “Anomaly detection for cyber security using GANs and deep learning,” in Proceedings of the 2020 International Conference on Artificial Intelligence, 2020, pp. 156–161.
[8] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein GAN,” arXiv preprint, arXiv:1701.07875, 2017.
[9] M. Mirsky et al., “Kitsune: An ensemble of autoencoders for online network intrusion detection,” in NDSS Symposium, 2018.
[10] T. Salimans et al., “Improved techniques for training GANs,” in Advances in Neural Information Processing Systems (NeurIPS), 2016.
[11] S. Roy, D. Kundur, and R. Shandilya, “A survey of recent trends in using GANs for network security,” ACM Computing Surveys (CSUR), vol. 55, no. 2, pp. 1–36, 2023.
[12] L. Lin, Y. Shi, and Z. Ye, "IDSGAN: Generative adversarial networks for attack generation against intrusion detection," in Proc. IEEE ICC, 2019, pp. 1–6.
[13] M. Shafiq, F. A. Khan, and M. A. Khan, "A novel approach to enhance the performance of network intrusion detection systems using generative adversarial networks," IEEE Access, vol. 9, pp. 4661–4671, 2021.
[14] W. Li et al., "Improving the performance of intrusion detection system using conditional generative adversarial network," Computer Networks, vol. 171, p. 107138, 2020.
[15] G. Kim, S. Lee, and S. Kim, "Anomaly detection in cyber-physical systems using GAN-based autoencoder," Sensors, vol. 19, no. 19, p. 4375, 2019.
[16] Y. Zhou, X. Zhang, and S. Zhang, "Semi-supervised anomaly detection for network traffic based on generative adversarial networks," Neurocomputing, vol. 439, pp. 1–9, 2021.
[17] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," arXiv preprint arXiv:1701.07875, 2017.
[18] Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, 2016.
[19] Y. Mirsky et al., "Kitsune: An ensemble of autoencoders for online network intrusion detection," in Proc. NDSS, 2018.
[20] C. Mao et al., "Adversarial training with GAN-generated attacks to improve IDS robustness," IEEE Access, vol. 10, pp. 82313–82325, 2022.
[21] Z. Luo and V. Nagarajan, "Anomaly detection using generative adversarial networks for cyber security," in Proc. IEEE C3S2E, 2020, pp. 51–58.
[22] H. Hu et al., "Membership inference and attribute inference attacks against generative models," in Proc. IEEE S&P, 2021, pp. 669–686.
[23] A. Roy, S. Sharma, and R. Ghosh, "A survey on applications of generative adversarial networks in cyber security," ACM Comput. Surv., vol. 55, no. 1, pp. 1–36, 2023.
[24] Y. Wang et al., "Intrusion detection in IoT using deep learning with GAN-based data augmentation," Wireless Networks, vol. 27, pp. 2125–2136, 2021.
[25] L. Huang et al., "Generating synthetic network traffic using GANs for intrusion detection training," in Proc. IEEE ICNC, 2020, pp. 167–171.
[26] Feng et al., "Enhancing intrusion detection systems with GAN-generated synthetic samples," Future Generation Computer Systems, vol. 127, pp. 112–121, 2022.
[27] J. Su et al., "Defending against GAN-based adversarial attacks in IDS," Computers & Security, vol. 87, p. 101568, 2019.
[28] N. Nguyen and L. Nguyen, "Smart grid intrusion detection using deep learning and GANs," in Proc. IEEE ISI, 2020, pp. 1–6.
[29] B. Zhang et al., "cGAN-based intrusion detection system for IoT networks," IEEE IoT Journal, vol. 8, no. 18, pp. 14234–14245, 2021.
[30] H. Yao, Y. Wang, and L. Zhang, "Feature augmentation for network intrusion detection using GANs," Journal of Network and Computer Applications, vol. 180, p. 103004, 2021.
[31] T. Reddy et al., "Comparative study of GAN variants for intrusion detection dataset generation," in Proc. ACM SAC, 2020, pp. 905–910.

[32] A. Singh et al., "Hybrid GAN-based intrusion detection system for cloud computing environments," IEEE Trans. Cloud Comput., Early Access, 2022.
[33] D. Goyal and M. Arora, "Using GANs for smart city cybersecurity intrusion detection," in Proc. IEEE SMARTCOMP, 2020, pp. 177–182.
[34] Q. Chen, W. Wang, and J. Xu, "TrafficGAN: Generative adversarial network based traffic simulation for IDS," IEEE Access, vol. 7, pp. 46098–46107, 2019.
[35] X. Wei et al., "Improving IDS performance using synthetic samples generated by GANs," Information Sciences, vol. 540, pp. 101–116, 2020.
[36] M. Islam et al., "Explainable AI for GAN-based IDS," Expert Systems with Applications, vol. 215, p. 119269, 2023.