Generative AI-Empowered Secure Communications in Space-Air-Ground Integrated Networks: A Survey and Tutorial

Chenbo Hu, Ruichen Zhang, *Member, IEEE*, Bo Li, *Member, IEEE*, Xu Jiang, *Member, IEEE*,
Nan Zhao, *Senior Member, IEEE*, Marco Di Renzo, *Fellow, IEEE*, Dusit Niyato, *Fellow, IEEE*,
Arumugam Nallanathan, *Fellow, IEEE*, and George K. Karagiannidis, *Fellow, IEEE*

Abstract— **Space-air-ground integrated networks (SAGINs) face unprecedented security challenges due to their inherent characteristics, such as multidimensional heterogeneity and dynamic topologies. These characteristics fundamentally undermine conventional security methods and traditional artificial intelligence (AI)-driven solutions. Generative AI (GAI) is a transformative approach that can safeguard SAGIN security by synthesizing data, understanding semantics, and making autonomous decisions. This survey fills existing review gaps by examining GAI-empowered secure communications across SAGINs. First, we introduce secured SAGINs and highlight GAI's advantages over traditional AI for security defenses. Then, we explain how GAI mitigates failures of authenticity, breaches of confidentiality, tampering of integrity, and disruptions of availability across the physical, data link, and network layers of SAGINs. Three step-by-step tutorials discuss how to apply GAI to solve specific problems using concrete methods, emphasizing its generative paradigm beyond traditional AI. Finally, we outline open issues and future research directions, including lightweight deployment, adversarial robustness, and cross-domain governance, to provide major insights into GAI's role in shaping next-generation SAGIN security.**

Index Terms—**Space-air-ground integrated networks, generative AI, security threats, communication authenticity, communication confidentiality, communication integrity, communication availability.**

I. Introduction

A. Toward GAI-Enabled SAGIN Security

Over the past decade, the advancement of 5G and the prospects for 6G have led to significant progress in space-air-ground integrated networks (SAGINs), including the extensive deployment of low earth orbit (LEO) satellite constellations (e.g., Starlink and OneWeb), the development of high-altitude platform station technology, and the enhancement through software-defined networking (SDN) and network function virtualisation (NFV) technologies [1]. This multi-layer heterogeneous architecture aims to provide global three-dimensional seamless coverae by 2030, fulfilling the rigorous demands of future applications such as intelligent transportation, telemedicine, and smart cities, characterised by high reliability, ultra-low latency, and extensive connectivity. [2].

SAGIN communication faces numerous intricate security threats stemming from distinct problems, including multi-dimensional heterogeneity, dynamic topology, and resource limitations, in contrast to conventional single-domain networks (e.g., terrestrial and aerial networks) [3]. This significantly enhances the complexity of security measures implemented for such cross-domain systems. The ephemeral connectivity of cross-domain nodes makes static authentication susceptible to spoofing attacks. Furthermore, coordinated assaults across the physical, data link, and network layers are being escalated. Resource-limited nodes find it challenging to facilitate high-intensity real-time integrity verification. Consequently, the fundamental security needs of SAGIN communication—authenticity, confidentiality, integrity, and availability—are susceptible to risks of imbalance. Thus, guaranteeing secure communications in SAGINs is of paramount importance [4] [5] [6] [7].

However, security methods, as encryption and intrusion detection, frequently prove inadequate in the dynamic and heterogeneous SAGINs. Conventional artificial intelligence (AI) methodologies, such as deep learning (DL) and deep reinforcement learning (DRL), offer improved functionalities for anomaly detection, threat identification, and adaptive security strategies; however, they demonstrate considerable shortcomings in the context of SAGIN security. It necessitates substantial quantities of labelled samples for training; however, SAGIN attack samples are limited and demonstrate non-independent and identically distributed (non-IID) traits, resulting in a significant decline in detection accuracy. Architecturally, unimodal models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are inadequate in effectively capturing cross-domain attack features. Moreover, inadequate dynamic flexibility required regular retraining to address emerging threats, hence failing to satisfy real-time anti-jamming demands. These constraints hinder the efficacy of conventional AI in safeguarding SAGINs [10].

Generative AI (GAI), as a prominent subset of AI, presents

Chenbo Hu, Bo Li, and Xu Jiang are with the School of Information Science and Engineering, Harbin Institute of Technology (Weihai), Weihai 264209, China (e-mails: huchenbo@stu.hit.edu.cn; libo1983@hit.edu.cn; xjiang@hit.edu.cn).

Ruichen Zhang and Dusit Niyato are with the College of Computing and Data Science, Nanyang Technological University, Singapore (e-mails: ruichen.zhang@ntu.edu.sg; dniyato@ntu.edu.sg).

Nan Zhao is with the School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China (e-mail: zhaonan@dlut.edu.cn).

Marco Di Renzo is with Université Paris-Saclay, CNRS, Centrale Supélec, Laboratoire des Signaux et Systèmes, 91192 Gif-sur-Yvette, France, and also with King's College London, Centre for Telecommunications Research—Department of Engineering, WC2R 2LS London, U.K. (e-mail: marco.di-renzo@universite-paris-saclay.fr; marco.di_renzo@kcl.ac.uk).

Arumugam Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London and also with the Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, Korea (e-mail: a.nallanathan@qmul.ac.uk).

George K. Karagiannidis is with Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece (e-mail: geokarag@auth.gr).

TABLE I
Summary of Related Surveys, Where ●, ◐, and ○ Represent Comprehensive Review, Partial Review, and Not Review, Respectively

Ref.	Contribution	SAGIN	Security	GAI
[3]	A survey of security threats, attack methods, and defense schemes in SAGINs, addressing the unique vulnerabilities of heterogeneous SAGINs and explicitly covering traditional AI-enabled defenses.	●	●	○
[10]	A survey of AI applications across satellite systems, covering use cases, hardware implementation challenges, and future directions, with explicit focus on traditional AI with little coverage of GAI.	◐	◐	◐
[13]	A review of AI applications, including generative models for security enhancement, across diverse satellite communication challenges such as anti-jamming, interference management.	◐	◐	◐
[14]	A review of AI applications and challenges in 6G UAV-satellite networks, systematically classifying and comparing solutions for security risks including traditional AI and GAI.	◐	◐	◐
[15]	A survey of GAI applications for physical layer communication security, covering defense mechanisms like RF authentication, anomaly detection, and anti-jamming.	○	◐	●
[16]	A survey of physical layer security for low-altitude economy networks, highlighting the application of GAI for enhancing anti-eavesdropping strategies, anomaly detection, and optimizing defenses.	◐	●	◐
[17]	A survey of GAI models addressing security issues in physical layer communications, covering confidentiality, authentication, availability, resilience, and integrity.	○	●	●
[18]	A review of the fundamentals, applications, and challenges of GAI in mobile networking, with a specific focus on advancing security solutions (e.g., intrusion detection and jamming mitigation).	◐	◐	●
[19]	A survey on deploying AIGC services in mobile edge-cloud networks, covering architectures, generative models, and use cases, with a focus on low-latency, privacy, and resource efficiency.	◐	◐	●
[20]	A review of GAI applications within the IoEV, highlighting its critical role in addressing security threats like adversarial attacks, and cyber-physical anomaly detection.	◐	◐	●
[21]	A survey on deep learning in anomaly detection, highlighting the critical role of generative models for both data augmentation to and enhanced anomaly identification in security-critical domains like IoT.	○	●	●
[22]	A review of privacy and security concerns in GAI from, highlighting GAI's potential applications in solving security problems, such as anomaly detection and cybersecurity threat identification.	○	●	●
[23]	A survey on security and privacy challenges in generative data from AIGC, highlighting the applications of GAI in developing methods such as data synthesis and multimodal defense frameworks.	◐	●	●
[24]	A survey of architectures, applications, and security challenges of LLM-based edge intelligence, highlighting GAI's role in optimizing security solutions e.g., threat detection.	○	◐	●
Our paper	Focus on how GAI can address various security issues in SAGINs and discuss the development potential of GAI in SAGIN secure communications compared to traditional AI.	●	●	●

an innovative framework for addressing security limitations in SAGINs, using its robust skills in data production, semantic comprehension, and autonomous decision-making [11]. In contrast to conventional AI, which primarily analyses, interprets, and classifies data to address specific issues, GAI has superior capability in analysing and recording intricate multidimensional data distributions, facilitating the creation of highly correlated new data (e.g., photos, text, video) [12]. Specifically, although traditional AI depends on past data to identify known risks, GAI may generate variations of unknown attacks by analysing data distributions, thereby mitigating the lack of labelled data. Moreover, transformer models of GAI utilise multi-head attention techniques to correlate cross-domain traffic, hence identifying multi-domain attack patterns overlooked by unimodal CNNs. It utilises real-time dynamic defence via prompt engineering (PE) and in-context learning (ICL) to enable millisecond-level anti-jamming responses. In summary, GAI not only mitigates the intrinsic deficiencies of static defences and conventional AI but also establishes a proactive-adaptive defence system specifically designed for heterogeneous SAGINs through privacy-preserving synthesis, cross-domain attack simulation, high-fidelity reconstruction, and semantic strategy generation.

This paper offers thorough insights into how GAI can efficiently tackle the substantial security difficulties posed by SAGINs.

B. Related Surveys and Contributions

The security challenges and AI-driven methodologies for SAGINs and its sub-networks have garnered significant interest, as demonstrated in Table I. The research in [3] offered an extensive examination of security threats, attack methodologies, and defensive strategies, focussing on the distinct vulnerabilities in heterogeneous SAGINs and partially addressing conventional AI-based defence mechanisms. Since that time, specialised AI review functions for SAIGNs or their sub-networks have been developing. For instance, in [10], the utilisation of AI in satellite communications was thoroughly examined, concentrating mostly on conventional AI techniques for various security issues while providing minimal insight into GAI methodologies. The research in [13] specifically investigated AI applications, encompassing generative models for security improvement, in various satellite communication issues like beam hopping, anti-jamming, channel modelling, and space-air-ground integration. The authors of [14] thoroughly delineated the uses, problems, and future possibilities of AI in 6G UAV-satellite communication networks, carefully categorising and contrasting several AI solutions for security threats, encompassing classical AI and GAI.

Furthermore, certain surveys highlight the GAI for secure physical layer communication. The research in [15] examined GAI applications for the security of physical layer communication, encompassing defence strategies such as authentication, anomaly detection, and anti-jamming, while also tackling new dangers posed by GAI-driven adversarial attacks. [16] examined secure physical layer communication methods for low-altitude economic networking, emphasising the novel application of GAI to improve anti-eavesdropping tactics, anomaly detection, and the optimisation of security measures. Moreover, the authors in [17] presented an exhaustive survey of GAI models that tackle security concerns in physical layer communications, encompassing confidentiality, authentication, availability, resilience, and integrity.

Turning to mobile networks, [18] comprehensively reviewed the fundamentals, applications, and challenges of GAI in mobile and wireless networking, partially advancing wireless security solutions (e.g., intrusion detection, jamming mitigation, and generative steganography). [19] investigated the AI-generated content (AIGC) services in mobile edge-cloud networks, including architectures, generative models, implementation challenges, and real-world use cases, with a focus on low-latency, privacy, and resource efficiency. Targeting internet of electric vehicle networks (IoEVs), [20] categorized GAI applications across four layers (battery, electric vehicle, grid, and security), highlighting its critical role in addressing security challenges like adversarial attacks and cyber-physical anomaly detection. [21] reviewed deep learning advancements in anomaly detection, highlighting the critical role of generative models for both data augmentation to address class imbalance and enhanced anomaly identification in security-critical domains like IoT and cybersecurity.

Furthermore, [22] analyzed privacy and security concerns in GAI from five key perspectives (user, ethical, regulatory, technological, and institutional), while also highlighted GAI's potential applications in solving security problems, such as anomaly detection, cybersecurity threat identification, and enhancing autonomous system safety. [23] systematically reviewed security and privacy challenges in generative data from AIGC, analyzed through the lens of core information security properties (i.e., privacy, controllability, authenticity, and compliance), while emphasizing the transformative applications of GAI in developing innovative countermeasures such as synthetic data synthesis, multimodal defense frameworks, and adaptive security mechanisms. The study in [24] explored architectures, applications, and security challenges of large language model (LLM)-based edge intelligence, highlighting GAI's role in optimizing threat detection, vulnerability management, and automated security solutions for resource-constrained environments.

Unlike existing surveys and tutorials, which focus primarily on GAI applications in sub-networks of SAGINs or on a specific type of GAI-enabled security defense technique, our survey provides a detailed classification of attacks and GAI defenses across the entire SAGIN, considering security requirements and network architectures. Our work bridges an essential gap in current research by providing a comprehensive analysis of the status and potential of GAI applications in SAGINs, as well as a full comparison with traditional AI. Our work explains how GAI enhances SAGIN security compared to traditional AI, a topic that has been previously under-explored. The contributions of our survey are as follows:

- We present the architectures of SAGINs and the security challenges from a security perspective. Meanwhile, we outline the basic GAI models and summarize their suitability for solving security issues and advantages over traditional AI.
- We investigate four categories of security problems that may be encountered in SAGINs, namely authenticity failures, confidentiality breaches, integrity tampering, and availability disruptions, and outline the corresponding GAI-based approaches, categorized by physical layer, data link layer, and network layer.

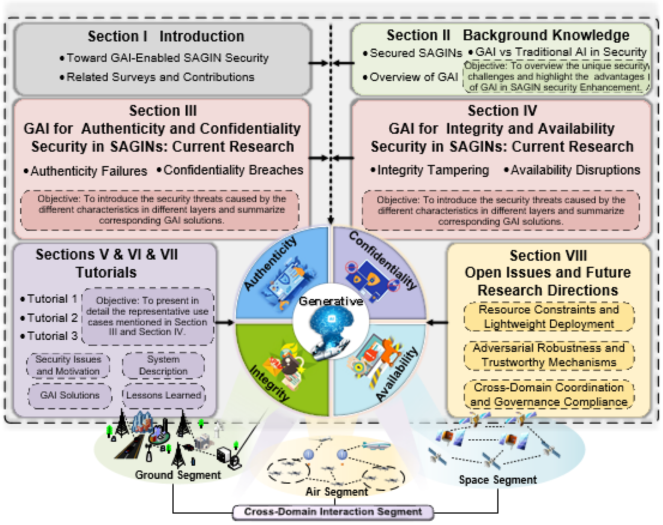
Section I Introduction
•Toward GAI-Enabled SAGIN Security
•Related Surveys and Contributions
Section II Background Knowledge
•Secured SAGINs
•GAI vs Traditional AI in Security
•Overview of GAI
Objective: To overview the unique security challenges and highlight the advantages of GAI in SAGIN security Enhancement.
Section III
GAI for Authenticity and Confidentiality Security in SAGINs: Current Research
•Authenticity Failures
•Confidentiality Breaches
Objective: To introduce the security threats caused by the different characteristics in different layers and summarize corresponding GAI solutions.
Section IV
GAI for Integrity and Availability Security in SAGINs: Current Research
•Integrity Tampering
•Availability Disruptions
Objective: To introduce the security threats caused by the different characteristics in different layers and summarize corresponding GAI solutions.
Sections V & VI & VII
Tutorials
•Tutorial 1
•Tutorial 2
•Tutorial 3
Objective: To present in detail the representative use cases mentioned in Section III and Section IV.
Security Issues and Motivation
System Description
GAI Solutions
Lessons Learned
Authenticity
Confidentiality
Generative
Integrity
Availability
Section VIII
Open Issues and Future Research Directions
Resource Constraints and Lightweight Deployment
Adversarial Robustness and Trustworthy Mechanisms
Cross-Domain Coordination and Governance Compliance
Ground Segment
Air Segment
Space Segment
Cross-Domain Interaction Segment

Fig. 1. Organization structure of this paper.

- We implement three tutorials that are employed to detail how GAI addresses the multi-class security issues in SAGINs, highlighting the advantages of GAI compared to traditional AI.
- We discuss the open issues and future research directions from the perspectives of resource constraints and lightweight deployment, adversarial robustness and trustworthy mechanisms, cross-domain coordination and governance compliance, respectively.

The remainder of this work is outlined in Fig. 1. From a perspective of security, we introdcue the architectures of SAGINs, overview the fundamental concepts of GAI, and compare the traditional AI with GAI in Section II. In Section III, a comprehensive exploration of GAI for authenticity failures and confidentiality breaches is presented. In addition, Section IV reviews the GAI solutions for integrity tampering and availability disruptions. In Sections V, VI, and VII, we conduct and analyze potential tutorials. Furthermore, we discuss open issues and discuss future research directions in Section VIII followed by Section IX as conclusion. For visibility, Table II lists the main acronyms quoted in this survey.

II. Background Knowledge

In this section, we first introduce the SAGINs from a security perspective, including components and security challenges. Subsequently, we provide an overview of GAI and compare its advantages over traditional AI in addressing SAGIN security issues.

A. Secured SAGINs

The distinctive characteristics of SAGINs, featured by self-organization, heterogeneity, and time-variability, introduce unprecedented challenges [6]. Among them, security vulnerabilities caused by architectural weaknesses, exposed nodes, and

TABLE II
LIST OF MAIN ABBREVIATIONS.

Abbreviation	Description	Abbreviation	Description
AAV	Autonomous Aerial Vehicle	IDS	Intrusion Detection System
ADS-B	Automatic Dependent Surveillance-Broadcast	IoT	Internet of Thing
AI	Artificial Intelligence	ISL	Inter-Satellite Link
AIGC	AI-Generated Content	LLM	Large Language Model
ARP	Address Resolution Protocol	LSTM	Long Short-Term Memory
BDS	Beidou Navigation Satellite	MAC	Media Access Control
CIR	Channel Impulse Response	MAD	Malicious Attack Detector
CNN	Convolutional Neural Network	MAVLink	Micro Air Vehicle Link
CSI	Channel State Information	MDP	Markov Decision Process
DDoS	Distributed Denial of Service	MEC	Multi-Access Edge Computing
DL	Deep Learning	MITM	Man-in-the-Middle
DNN	Deep Neural Network	ML	Machine Learning
DoS	Denial-of-Service	MTL	Multi-Task Learning
DP	Differential Privacy	RNN	Recurrent Neural Network
DRL	Deep Reinforcement Learning	RF	Radio Frequency
GAI	Generative AI	SAGIN	Space-Air-Ground Integrated Network
GAN	Generative Adversarial Network	SDR	Software-Defined Radio
GDM	Generative Diffusion Model	SVM	Support Vector Machine
GNSS	Global Navigation Satellite System	TBM	Transformer-Based Model
GPS	Global Positioning system	VAE	Variational Autoencoder

unsecured communication links pose critical threats to four fundamental communication security requirements: authenticity, confidentiality, integrity, and availability [3], [25].

- *Authenticity:* Authenticity ensures the legitimacy of network entities and the trustworthiness of data sources in SAGINs, and prevents unauthorized access through multi-layered authentication mechanisms [26]. Current implementations employ hardware-based identification (e.g., media access control (MAC) address), protocol-layer authentication (e.g., digital certificates), and biometrically-enhanced dynamic credentials [27].
- *Confidentiality:* Confidentiality ensures that data transmission is accessible exclusively to authorized users. Since the transmitted data contains information about the user's behavior, malicious attackers can indirectly infer sensitive information using unintentionally leaked available messages [7].
- *Integrity:* Integrity ensures data accuracy and reliability during transmission, storage, and processing. The attack against integrity is less intense but more sophisticated. Attack surfaces extend beyond simple original data modification to include protocol-level interference that disrupts information exchange processes from within network layers [28].
- *Availability:* Availability ensures authorized users to access the wireless network whenever and wherever they request it, even under adversarial conditions or system failures [29]. Attackers attempt to delay, block or even interrupt transmissions, thus rendering network resources unavailable.

The following discusses the security challenges and their impact on security requirements across four segments.

1) Secured Cross-Domain Interaction Segment: The cross-domain interaction mechanism consists of dynamic protocol stacks, multi-domain gateways, and intelligent control planes that coordinate the three heterogeneous segments: space, air, and ground [30], [31]. This framework serves as a vital facilitator for multi-domain coordination in SAGINs [1], [32], encountering two main security challenges: protocol heterogeneity [3] and vulnerabilities associated with delay-sensitive handovers [33]. The transient connectivity of dynamic nodes, such as satellites and autonomous aerial vehicles, makes static authentication mechanisms vulnerable to spoofing attacks. In these scenarios, adversaries can impersonate legitimate gateways and introduce false routing information, thereby compromising authenticity [34], [35]. Furthermore, confidentiality risks emerge from vulnerabilities in inter-satellite link (ISL) protocols, allowing eavesdroppers to capture cross-domain session keys through compromised key mapping tables at protocol gateways [36]. Divergent protocol implementations, such as the Consultative Committee for Space Data Systems (CCSDS), Micro Air Vehicle Link (MAVLink) for aerial platforms, and the 3rd Generation Partnership Project (3GPP) for ground networks, increase the vulnerability of data conversion processes to manipulation. For example, man-in-the-middle (MITM) attacks as described in [38] effectively modified AAV control commands during AAV-to-ground switching operations, thereby undermining data integrity. Additionally, threats to availability arise from the targeted jamming of cross-domain synchronisation signals, potentially disabling positioning systems in AAV-ground synchronisation [39].

2) Secured Space Segment: The space segment includes satellites, ground stations, and space-based Internet of Things (IoT) devices, utilising technologies such as LEO constellations. The broadened network exposure presents security risks, including susceptibility to assaults and data manipulation. Attackers can spoof global positioning system (GPS) signals by employing publicly accessible ephemeris data to distort satellite navigation. Although satellite laser communications provide narrow beams, high-altitude platforms can capture signals with sophisticated optical tracking, hence jeopardising confidentiality [42]. Furthermore, the constrained computational capabilities of on-board devices impede real-time high-strength integrity checks [43]. Moreover, adversaries can anticipate satellite trajectories and inundate satellite uplinks with high-power transmitters, significantly diminishing single-beam

TABLE III
SUMMARY OF SECURITY CHALLENGES IN SAGINS.

Segment	Authenticity	Confidentiality	Integrity	Availability
Cross-Domain Interaction	**• Spoofed Gateways:** Forge authentication to inject false routing [34], [35].	**• ISL Eavesdropping:** Intercept data in protocol conversion [36].	**• MITM in Handovers:** MITM tampers with control commands [38].	**• Cross-Layer Jamming:** Jamming attacks threaten cross-domain signals [39].
Space	**• GPS Spoofing:** Forge commands to manipulate signals [41].	**• Laser Interception:** High-altitude eavesdropping for narrow beams [42].	**• Real-Time Check Limits:** Limited resources restrict real-time validation [43].	**• High-Power Jamming:** Overwhelm satellite uplinks with high-power signals [44], [45].
Air	**• MAC Cloning:** Hijack networks via spoofed addresses [50].	**• mmWave Leakage:** Millimetre-wave links suffer from leakage [52], [53].	**• Path-Switching Injection:** Injecting data when switching paths in AAV networks [38].	**• DDoS Flooding:** Botnets overload aerial command channels [54], [55].
Ground	**• Rogue BS:** Fake BS harvests user data [61].	**• Edge Server Breaches:** Attacked MEC nodes leak sensitive data [62].	**• ML Traffic Inference:** ML deduces secrets from encrypted patterns [63].	**• Firmware Tampering:** Malware alters IoT device updates [64].

throughput [44], [45]. Sleep deprivation can be exploited by attackers to intentionally and persistently transmit erroneous control commands [45], [46] to deplete network resources.

3) Secured Air Segment: The air segment includes AAVs, airships, aerial base stations, and airborne sensors, distinguished by dynamic mobility and line-of-sight dependencies [47], [48]. Key security issues include AAV identity spoofing and the interception of wireless signals [49]. Malicious nodes can replicate AAV MAC addresses to infiltrate networks and seize control of swarms through deceptive link state announcements, compromising authenticity [50]. In [51], the radio frequency (RF) fingerprint authentication technique was introduced to extract hardware features of node communication modules to mitigate attacks. Millimetre-wave links in legitimate AAVs exhibit vulnerabilities regarding the leakage of unencrypted messages [52], [53]. The quantum key distribution technique [7] shows potential for mitigating decryption risks. Furthermore, node mobility enables MITM attacks that alter AAV commands, thereby undermining integrity [38]. The restricted computational capacity of on-board devices, akin to space nodes, hinders effective real-time integrity assessments [47]. Distributed denial of service (DDoS) attacks, particularly those involving numerous distributed bot-type attacks, inundate the air network, thereby limiting access for legitimate users and impairing availability [54], [55].

4) Secured Ground Segment: The ground segment is composed of multiple sub-networks, such as 5G/6G base stations [56], narrowband IoT networks [57], and worldwide interoperability for microwave access (WiMAX) [58]. Its distributed architectures and legacy protocols (e.g., message queuing telemetry transport (MQTT) [59] and constrained application protocol [60]) expose critical secure vulnerabilities. In terms of undermining authenticity, rogue base stations, masquerade as legitimate base stations to spoof the users and obtain their data [61]. Confidentiality leaks arise from compromised edge servers. Adversarial machine learning (ML) models can deduce sensitive industrial IoT data from encrypted multi-access edge computing (MEC) traffic patterns [62]. Integrity violations occur when malware modifies firmware updates in industrial IoT devices. For example, in [63], malware added malicious content to escalate an application's privileges, allowing hackers to exfiltrate sensitive data. Furthermore, availability disasters can stem from physical infrastructure damage, such as severed backbone fibre-optic cables [64].

In summary, SAGINs exhibit many new types of security challenges, and each segment confronts its own unique security threats, as summarized in Table III.

B. Overview of GAI

GAI is increasingly adopted across various fields of communication networks [65], and its distinctive features make it very suitable for SAGIN security enhancement. Secured GAI models be categorized as follows:

- *Variational Autoencoders (VAEs):* VAEs are generative frameworks that integrate DL with probabilistic graphical models, consisting of an encoder, a decoder, and a probabilistic latent space. The encoder compresses sensitive data into a probabilistic latent distribution, thereby enhancing confidentiality through non-linear dimensionality reduction. A joint source channel coding scheme for point-to-point wireless communications was proposed by [66], utilising vector quantised VAE (VQ-VAE) and achieving nearly 90% accuracy. This scheme applies to anti-eavesdropping within the space segment. Simultaneously, the decoder reconstructs data from latent samples, facilitating integrity verification via anomaly detection. The AC-VAE framework presented in [67] employed active learning alongside contrastive VAE-based models, resulting in F1 scores ranging from 0.68 to 0.96 using merely 3% of labelled data. This approach prevents dynamic routing table tampering in LEO satellite networks, thereby enhancing the efficiency of data integrity verification. VAEs, while adaptable, demonstrate blurred reconstructions and instability during training in high-dimensional contexts.
- *Generative Adversarial Networks (GANs):* GANs are unsupervised learning frameworks that model accurate data distributions through the adversarial training between a generator and a discriminator [68]. The discriminator detects anomalies by differentiating between legitimate data and adversarial samples. In [69], an FS-GAN algorithm was introduced, leveraging the principles of federated learning (FL), self-supervised learning, and GANs, resulting in an improvement of over 20% in average

classification performance. The enhancement of cross-domain protocol vulnerability detection is anticipated, along with a reduction in the risk of MITM hijacking during the conversion between MAVLink and hypertext transfer protocol (HTTP). The generator creates realistic attack scenarios to strengthen defences proactively. A mask guided adversarial training (MAGAT) structure was proposed by [70], integrating adversarial training with mask-guided operations to protect against malicious face editing. This architecture can be utilised to pre-train AAV biometric authentication systems, thereby mitigating identity authenticity failures resulting from MAC address forgery. Despite their advantages, GANs experience mode collapse and exhibit limited control over local feature generation, which constrains their ability to synthesise a diverse range of threats.

- *Generative Diffusion Models (GDMs):* GDMs are score-based generative models that learn data distributions through a bi-directional process of forward diffusion and reverse denoising [71], [72], [73]. The forward diffusion process progressively injects controlled Gaussian noise into real data until it becomes pure noise [75], whereas the reverse denoising network effectively recover corrupted raw data from adversarial environments iteratively by training neural networks [74]. As demonstrated in [76], a reverse denoising inspired DRL algorithm was proposed to balance content reconstruction fidelity and transmission efficiency in mobile AIGC networks. Complementary work in [77] presented an approximate message passing algorithm based on the reverse denoising process, which could achieve an improvement in signal reconstruction quality. These methods facilitate content reconstruction after data tampering in the space segment. While GDMs avoid mode collapse in GANs and exceed VAEs in generation fidelity, their reliance on 1000+ denoising steps limits real-time applications (e.g., real-time video encryption) [78].
- *Transformer-Based Models (TBMs):* TBMs are DL architectures based on a self-attentive mechanism that process sequential data through an encoder-decoder framework and perform well in natural language processing tasks [79]. The multi-head self-attention module captures global dependencies, overcoming the local constraints of CNNs and RNNs. The encoder captures the complex dependencies and analyzes input sequences for security detection, while the decoder generates target sequence using cross-attention. [80] developed TAEF, a transformer-autoencoder hybrid for hyperspectral anomaly detection, which could be adopted in identifying payload tampering that occurs in satellite-ground links and AAV-ground links. TBMs also advance physical layer security [81] and intrusion detection [82]. Despite these strengths, challenges such as high computational demands, strong data dependency, and integration complexity hinder widespread adoption.
- *LLMs:* LLMs are transformer-based neural networks trained on massive text corpora to understand and generate human-like text through contextual reasoning [83], [84]. While LLMs and TBMs are similar, LLMs uniquely rely on massive pre-trained language backbones and contextual reasoning to generate security policies (e.g., cross-domain rules) or parse natural-language threat descriptions. These capabilities are critical for addressing SAGIN's dynamic multi-domain collaboration and human-in-the-loop adaptation, which generic TBMs lack owing to task-specific rigidity. Specifically, LLMs consist of three core components: a pre-trained backbone (e.g., GPT-4 and LLaMA) for general language patterns, domain adaptation techniques (e.g., retrieval-augmented generation) to specialize in vertical fields, and tool interfaces enabling interaction with external systems (e.g., network analyzers). Security applications include natural-language-to-configuration translation, as implemented in ChatNet's encrypted SAGIN policy generation [83], and privacy-preserving sensor data transmission via collaborative mobile agent systems [85]. Despite these strengths, LLMs remain vulnerable to adversarial prompts and face challenges such as high computational costs and context window limitations [86], [87].

In summary, GAI models can be targeted to address authenticity failures, confidentiality breaches, integrity tampering, and availability disruptions in SAGINs through data generation, feature reconstruction or policy optimisation capabilities. The basic principles, applications in security, and features of these GAI models are concluded in Table IV.

C. GAI vs Traditional AI in Security

While traditional AI (e.g., discriminative AI) has demonstrated efficacy in addressing conventional security challenges. GAI, a transformative paradigm in AI's evolution, is revolutionizing the field through its unique features, such as data augmentation, scenario simulation, and cross-modal analysis [88], [89], [90]. The fundamental distinctions between these paradigms include the core paradigm, architecture, data dependency, and dynamic adaptation, which enable GAI as a superior solution for modern security enhancement.

- *Core Paradigm:* Traditional AI relies on discriminative models (classification/regression) to differentiate normal from abnormal patterns. However, its over-reliance on historical data restricts it to known threats, rendering it ineffective against dynamic unknown threats in complex SAGINs such as zero-day attack [91]. For instance, while support vector machines (SVMs) are effective for known threats [92], their accuracy drops to only 50% against unknown threats [93]. In contrast, GAI leverages generative models to synthesize diverse attack variants by learning underlying data distributions. FS-GAN in [69] could generate hybrid attack traffic without labeled data and improving detection accuracy more than 20%. This capability enables the generation of variant MITM attacks for cross-domain interactions, facilitating pre-training of SAGIN intrusion detection systems.
- *Architecture:* Conventional AI depends on established feature engineering and static architectures, such as CNNs and RNNs. This unimodal approach increases information

TABLE IV
SUMMARY OF BASIC PRINCIPLES, SECURITY APPLICATIONS, AND FEATURES OF TYPICAL GAI MODELS, WHERE GREEN TICK REPRESENTS STRENGTHS AND RED CROSS REPRESENTS WEAKNESSES.

Model	Principle	Security Applications	Pros & Cons
VAEs	Through an encoder and a decoder, the latent space can be learned, and new content is generated	• **VQ-VAE for Secure Coding:** A source channel coding scheme with above 90% accuracy [66]. • **AC-VAE for Anomaly Detection:** Utilize the active learning to improve VAE with less labels [67].	✓ Ideal latent-space anomaly detection ✗ Blurred reconstructions with weak protocol-level tamper detection
GANs	Model accurate data distributions through the adversarial training	• **FS-GAN for Anomaly Detection:** Distinguish legitimate data from adversarial samples [69]. • **MAGAT Against Malicious Face Editing:** Utilize MaGAT to defend against malicious face editing [70].	✓ Powerful attack traffic generation ✗ Easy mode collapse and dependence on empirical parameterization with poor performance
GDMs	Learn data distributions through forward diffusion and reverse denoising	• **DRL-GDM for Content Recovery:** Balance content recovery and transmission [76]. • **GDMs for Approximate Message Passing:** Utilize the reverse process to acheive the message passing [77].	✓ Robust model and flexible for real-time spoofing defense ✗ High energy consumption
TBMs	Self-attention-based process for sequential data	• **TAEF for Anomaly Detection:** Achieve detection on multiple real hyperspectral datasets [80]. • **TBMs for Intrusion Detection:** Detect sophisticated and dynamic threats [82].	✓ Well cross-modal protocol semantic analysis for cross-domain defense ✗ High computational load with limited real-time security in AAVs detection
LLMs	Generate human-like content through contextual reasoning	• **ChatNet for Eavesdropping Prevention:** Transform natural language into encrypted configurations [83]. • **LLMs for Privacy Preserving:** Present a split learning system with the privacy-preserving [85].	✓ Dynamic policy generation and natural-language threat parsing ✗ Context window limitations with cross-domain attack chain failure

leakage and diminishes cross-domain attack detection accuracy in SAGINs. A single CNN is insufficient to defend against diverse MITM attacks across multiple domains [94]. In contrast, GAI facilitates multimodal inputs and end-to-end generation through modular designs, such as progressive denoising in GDM and self-attention in TBM, enhancing the accuracy of threat detection via cross-modal correlation. The advantage was demonstrated through a cross-modal transformer (FmFormer) [95]. This framework enables the correlation of satellite telemetry, AAV sensor logs, and ground network traffic to detect multi-domain routing hijacking attacks in SAGINs, which single-modality CNNs are unable to identify.

- *Data Dependency:* Conventional AI depends on substantial labelled datasets, resulting in diminished performance in scenarios with limited data [96]. Centralised training raises privacy concerns, particularly in widely exposed SAGINs [97]. GAI addresses these limitations by employing unsupervised and few-shot learning, thereby mitigating label scarcity through the generation of synthetic data. As a result, it exhibits reduced response time in identifying unknown attacks and can effectively balance privacy with model performance through prospective simulation. A dilated convolutional transformer-based GAN (DCT-GAN) for time series anomaly detection was proposed in [98] to enhance model accuracy and generalisation, facilitating the precise detection of stealthy satellite-side attacks in data-scarce SAGINs.
- *Dynamic Adaptation:* Traditional AI relies on static models that require periodic retraining to adapt to evolving threats in SAGINs, resulting in delayed responses to emerging risks. Specifically, the underlying data distribution changes over time, which can render ML models trained on historical data obsolete [99] [100]. GAI, however, achieves real-time adaptability through PE and ICL, which enables milliseconds of real-time anti-jamming and on-demand dynamic optimization. [101] combined DL with generative model to utilize ICL for dynamically optimized defense against zero-day attacks, gaining a 15% performance improvement. Besides, ChatGPT was adopted for automated anomaly detection script generation via PE [102]. Hence, these methodologies enable real-time defense against evolving SAGIN threats without costly model retraining.

In summary, GAI has more powerful analytical capabilities and novel generative features than traditional AI, which makes it uniquely suited for tackling complex security challenges in SAGINs. The differences between them and examples of security enhancement via GAI are summarized in Table V.

The integration of GAI with SAGINs represents a significant advancement in communication technology [11]. The emergence of complex security threats, including cross-domain signal jamming and dynamic cyber attacks in SAGINs, highlights the potential application of GAI in threat modelling, real-time defence, and data privacy protection. In this context, GAI-enabled SAGIN security enhancement will serve as a crucial factor in the evolution of AI-SAGIN.

III. GAI FOR AUTHENTICITY AND CONFIDENTIALITY SECURITY IN SAGINS: CURRENT RESEARCH

Section III and IV review the role of GAI in mitigating SAGIN security threats through four security requirements as illustrated in Fig. 1. Meanwhile, this exploration incorporates the open systems interconnection (OSI) layered protocol layers (i.e., physical layer, data link layer, and network layer) under each requirement to bridge current research gaps. This is because different layers exhibit distinct vulnerabilities due to their underlying protocols as shown in Fig. 2. In this section,

TABLE V

Summary of Differences Between GAI and Traditional AI for Security Enhancement, Where Red Dot Represents the Drawbacks of Traditional AI, Green Dot Represents the Advantages of GAI, and Black Dot Represents the GAI Enhanced Security Examples.

Aspect	Traditional AI	Generative AI	SAGIN Security Enhancement via GAI
Core Paradigm	🔴 **Poor Dynamic Threat Detection:** Discriminative models are limited to classifying known threats.	🟢 **Superior Various Attack Defense:** Generative models create diverse samples to enhance robustness.	● **FS-GAN for Various Attack Generation:** FS-GAN generates MITM attacks for cross-domain interaction segment [69].
Architecture	🔴 **Poor Cross-Modal Detection:** Fixed architectures struggle to fuse multimodal data.	🟢 **Superior Cross-modal Detection:** Modular design enables cross-Modal correlation.	● **FmFormer for Cross-Modal Detection:** FmFormer detects the multi-domain routing hijacking attacks in SAGINs [95].
Data Dependency	🔴 **Over-reliance on Databases and Possible Privacy Issues:** Rely on amounts of labelled data with poor performance in few-shot scenarios.	🟢 **Powerful Data Generation and Prospective Privacy Protection:** Synthetic data generation addresses labeling gaps and preserves privacy.	● **DCT-GAN for Few-Shot Preservation:** DCT-GAN reduces detection latency and balances privacy via simulation, to detect stealthy attacks in data-scarce SAGINs [98].
Dynamic Adaptation	🔴 **Outdated Models Perform Poorly in New Threats Defense:** Static models fail to respond in real-time to dynamic threats.	🟢 **Superior On-Demand Optimization and Real-Time Anti-Jamming:** PE and ICL enable real-time strategy adjustments.	● **PE/ICL for Evolving Threats Defense:** ICL-enhanced generative model for SAGINs' zero-day defense [101] while PE-based anomaly detection through ChatGPT [102].

we first summarize the authenticity failures and confidentiality breaches that may be encountered in SAGINs, and overview the corresponding GAI-based security solutions in detail.

A. *GAI for Authenticity Failures*

1) Physical Layer: Failures in authenticity at the physical layer in SAGINs primarily arise from signal spoofing attacks, wherein attackers transmit either fabricated signals or relevant real signals to target devices. The openness of long-link signals, including satellite navigation signals like GPS [103] and the Beidou navigation satellite (BDS) system [104], as well as air signals such as automatic dependent surveillance-broadcast (ADS-B) [105], makes them susceptible to spoofing. Attackers utilise publicly accessible ephemeris data to spoof global navigation satellite system (GNSS) signals, leading to location inaccuracies and trajectory deviations in AAVs [41].

GANs improve security by facilitating a competitive interaction between the generator and the discriminator. This research aims to develop effective spearheads (spoofing jamming) and robust shields (spoofing defence) utilising GANs for anti-spoofing applications. In the air segment, the end-to-end FlightSense [105] integrated GAN and CNN methodologies to identify ADS-B spoofing and perform aircraft identification using raw I/Q signals, resulting in a detection accuracy of 98.87% and a classification accuracy of 99% on synthetic datasets. In the space segment, as shown in Fig. 3, [106] utilised GAN to attain high-precision signal detection in GNSS acquisition. Although these methods demonstrate high precision, they generally concentrate on individual attack types and exhibit limited scalability. To address these limitations, [107] developed a GAN-assisted contextual pattern-aware intrusion detection system that demonstrates high accuracy in identifying six types of attacks, including spoofing and replay, along with ultra-low latency in real-vehicle experiments. This model demonstrates significant applicability potential for SAGINs.

Hybrid GAI approaches are emerging beyond the single GAN. GenCoder [108] integrated a five-layer deep neural network (DNN) with VAE to effectively produce adaptive training data for unidentified attacks. Reference [109] integrated VAE with Wasserstein GAN (WGAN) for the detection of spoofing signals, utilising the accurate reconstruction capabilities of VAE alongside the feature learning strengths of WGAN. Additionally, the capabilities of TBMs to capture spatio-temporal features and identify anomalous patterns through a multi-head attention mechanism are incorporated with GANs. DroneDefGANt, as proposed by [110], represents a hybrid methodology that integrates GANs with transformer models to identify external threats, such as GPS spoofing, alongside internal failures, including actuator malfunctions.

Current research also explores GAI's integration with traditional authentication approaches such as RF fingerprinting (RFF), channel state information (CSI), and channel impulse response (CIR). For instance, [111] introduced GANSAT by combining GAN and satellite constellation RFF for GPS spoof detection and location estimation. This approach overcame conventional methods' environment-sensitivity limitations, but required location-specific training data with unverified cross-domain generalization. As a refinement, [112] employed GAN-CNN approach by analyzing the RFF of LoRa system, performing 92.4% authentication accuracy without channel compensation. In addition, [113] exploited conditional VAE (CVAE) to adaptively compress and extract discriminative RFF features to improve accuracy. Based on the CSI, [114] proposed a GDM-based secure sensing solution for ISAC system to extract the true CSI for sensing. [115] constructed a conditional GAN (CGAN) framework integrating long short-term memory (LSTM) and gated recurrent unit (GRU) networks to predict mobile channel responses. It achieved high accuracy but degraded without CSI training data. As a complement, [116] presented HVAE, combining an AE for CIR characteristics extraction and a VAE for the enhancement of CIR feature representation and authentication results.

2) Data Link Layer: The expiration of contextual authentication in dynamic SAGINs creates temporal gaps for attackers to forge temporary MAC identities. Typical threats include MAC address spoofing, address resolution protocol (ARP)

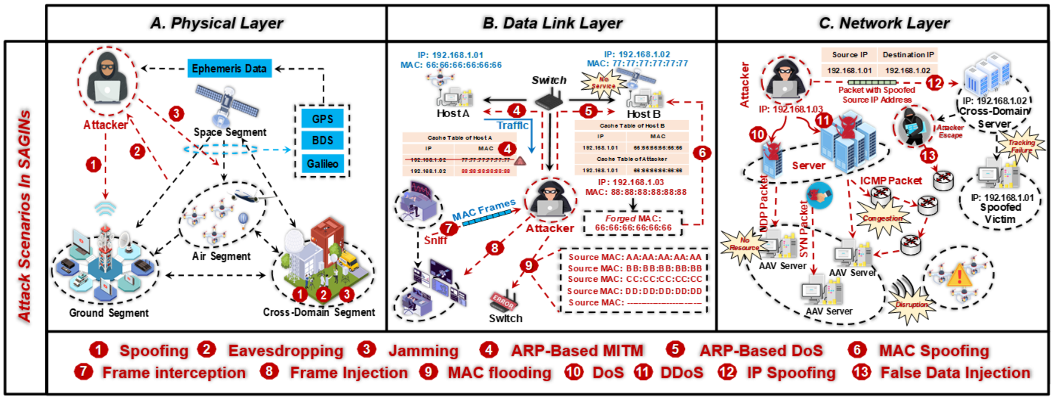

Fig. 2. Overview of typical attack scenarios at different layers in SAGINs. The 13 typical attack methods summarized above significantly cause authentication failures, confidentiality breaches, integrity tampering, and availability disruptions in SAGIN communications.

spoofing, and frame injection attack, facilitating DoS attack, MITM attack, and session hijacking attack [50].

MAC address serves as a unique physical credential for device access. Attackers can bypass the network authentication by cloning the legal MAC address [117] [118]. Although threshold-based sequence number analysis detects MAC spoofing, it yields false alarms due to frame loss. To overcome this drawback, an artificial neural network (ANN)-based detection method was proposed in [119] to analyze sequence number gaps and statistical distributions in MAC headers. The ANN framework could function as a pre-trained feature extraction layer for GAN discriminators or generators. Additionally, randomized dynamic defense [120] improve IoT security by adapting signal strength thresholds. Such methods are compatible with GAN integration for enhanced defense.

ARP maps logical address (i.e., IP address) to corresponding physical address (i.e., MAC address) without verifying the authenticity of responses to ARP requests [121]. This vulnerability enables attackers to execute ARP spoofing (i.e., ARP cache poisoning). It is an active attack method involving the injection of forged IP-MAC mappings into devices via unsolicited ARP messages. Attackers manipulate ARP cache tables without awaiting target-initiated queries, facilitating subsequent MITM and denial-of-service (DoS) attacks. The FS-GAN in Fig. 3 could classify ARP-spoofed frames adopting distributed GANs for sample generation and FL for training [69] .

In SAGINs, legitimate devices operate within open shared wireless media, allowing attackers to sniff MAC frames and forge spoofed frames for injection attacks such as false acknowledgment (ACK) and message frame injection [122]. These frames disrupt user decisions and exhaust resources, causing communication blockages. [123] designed an airspace anomaly detection method combining flight plan data with GAN-LSTM model for ADS-B attack. It could construct airspace image frames, utilizing GANs for temporal prediction and normalized cross-correlation for anomaly localization. Results demonstrated 92.3% average detection accuracy with 11.9% false positive rate and 6.2% false negative rate, effectively identifying stealthy attacks like frame injection while maintaining compatibility with ATC operational constraints.

3) Network Layer: Against threats like IP forgery and routing hijacks, GAI enhances end-to-end defense by synthesizing adversarial traffic for anomaly detection and generating dynamic authentication patterns to expose malicious actors.

In [124], the authors introduced a multi-architecture GAN framework aimed at mitigating data imbalance in network intrusion detection. The application of synthetic data to the CIC-IDS2017 benchmark resulted in a 6.18% enhancement in recall for Bot attack detection. However, real-time traffic in SAGINs is dynamic, and this method may have limitations. The Magteon-Turing L3TM framework in [125] introduced a GAI-driven method to address identity spoofing and traffic camouflage. The integration of Megatron-Turing NLG and Swarm OpenAI LLMs was achieved within a seven-layer architecture. 1) The data layer utilised GANs for the generation of synthetic hybrid traffic; 2) The traffic analysis layer implemented Langchain for semantic feature extraction, achieving 98.7% accuracy in detecting SYN Flood and IP spoofing; 3) The behavioural validation layer employed Llama models for protocol logic verification, identifying anomalous session patterns with a false positive rate of 1.5%. This study addresses the static rule limitations identified in [124] by employing dynamic LLM ensembles, which facilitate adaptive responses with sub-second latency.

GAI is capable of predicting the time series of attacks in SAGINs via the generation of adversarial samples. A hybrid deep learning framework integrating remaining useful life GAN (RUL-GAN) for time-series prediction of DoS attacks was proposed by [126]. A domain-specific GAN mechanism was developed to synthesise distributed energy resource control messages (e.g., packet parameters, protocol features), revealing authenticity threats from adversarial message spoofing [127]. The framework incorporated conditional value-at-risk for quantifying tail risk and employed an ensemble learning-based bagging method to identify synthetic identities, resulting in an attack vector generation accuracy of 95.7%, a tail risk of 9.61% (with 95% stability), and an overall accuracy of 99%.

The GAN-DRL model presented in [128] incorporated an intelligent routing model to tackle the issue of detecting minor-class attacks resulting from imbalanced data in healthcare-consumer IoT, while simultaneously optimising network-layer routing. The approach utilised GAN for generating synthetic data to identify rare attacks, such as smurf attacks, and employed DRL to dynamically adjust routing policies. This resulted in a 12% improvement in throughput, a 20% reduction in latency, and a 30% increase in the probability of avoiding

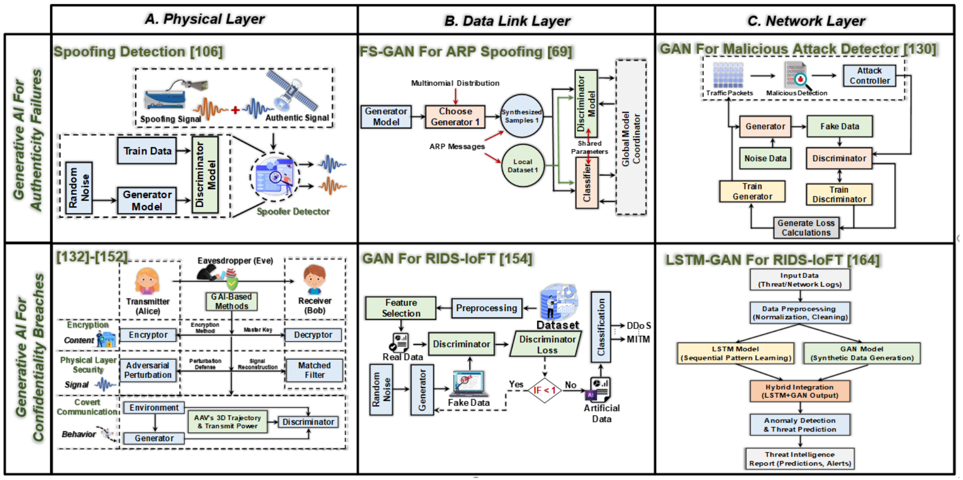

Fig. 3. Typical GAI-based approaches for authenticity failures and confidentiality breaches in SAGINs. For authenticity failures, [106] introduces a GAN-based GNSS spoofing framework; [69] illustrates a FS-GAN approach against ARP spoofing; [130] presents a GAN-based malicious attack detector. For confidentiality breaches, [132] - [152] designs multiple physical layer anti-eavesdropping methods by combining GAI with encryption, physical layer security, and covert communication; [154] proposes a GAN-based IDS for IoFT; [164] combines LSTM with GAN for IDS design.

malicious attacks within an SDN architecture. This approach demonstrates limited efficiency in generating high-dimensional traffic features. To tackle this problem, [129] proposed a lightweight GAN-DRL model compression technique aimed at enhancing the efficiency of high-dimensional traffic feature generation while mitigating the significant computational overhead in SDN-IoT environments. In addition, [130] proposed a distributed GAN training framework integrating a malicious attack detector (MAD) for multi-device collaborative novel attack detection as shown in Fig. 3.

4) Summary and Lessons Learned: As summarized in Table VI, GAI serves as a proactive tool for generating adversarial samples and a reactive shield for real-time anomaly detection, enabling adaptive defense in SAGINs. GANs excel at detecting spoofing through signal discrepancy analysis. Hybrid models integrating VAEs or transformers further enhance robustness against unknown attacks. Federated GANs and LLM-GAN hybrids demonstrate exceptional capabilities in mitigating MAC spoofing, IP forgery, and routing hijacks by leveraging synthetic traffic and semantic features to reduce false positives. Despite these advancements, challenges remain:

- *Generalization-Specialization Trade-off:* Models like CVAE-ZSL achieve high zero-shot detection accuracy but degrade in cross-domain scenarios.
- *Computational Overhead:* Computational overhead limits real-time high-dimensional traffic analysis.
- *Explainability Gaps:* Lack of explainable AI (XAI) integration undermines trust in hybrid model decisions.
- *Adversarial Robustness:* GAI-enhanced attack tools exploit static defenses, requiring co-designed countermeasures.

B. GAI for Confidentiality Breaches

1) Physical Layer: Eavesdropping attacks, referred to as stealth attacks, are markedly exacerbated by the intrinsic openness of wireless channels and the architectural heterogeneity present in SAGINs. Attackers leverage vulnerabilities in satellite link broadcasts, AAV relays, and terrestrial network density to intercept data through signal capture, protocol parsing, or cross-layer traffic correlation. Current countermeasures primarily depend on encryption methods for content protection, physical layer security for signal protection, and covert communication for behaviour concealment [131]. These approaches face challenges related to dynamic adaptability, computational overhead, and cross-domain coordination. Recently, GAI has improved these methods by incorporating dynamic learning and generative capabilities.

GAI-assisted encryption: Conventional encryption methods depend on a pre-shared secret key between communicating entities, which are susceptible to elevated error rates and key exposure. A WGAN-GP adversarial autoencoder was proposed to dynamically extract cross-layer channel features, thereby enhancing key capacity and reducing errors [132].

Additionally, the high spatio-temporal similarity of the CIR in high-density user scenarios enhances the likelihood of eavesdroppers successfully reconstructing keys. In this context, [133] proposed a GAN-assisted noise generation algorithm designed to inject adversarial noise into critical defence regions of dynamic LiFi networks. Moreover, frequent handovers in SAGIN elevate the risk of key leakage. An intelligent soft handover method for UAV-enabled cellular networks is presented, utilising GAN in conjunction with blockchain and physically unclonable function for lightweight message encryption [134].

In addition to single-attacker scenarios, [135] introduced a multiparty adversarial encryption model utilising GAN for secure multi-party communication. The introduction of an enhanced adversarial training framework featuring four types of attackers (including ciphertext-only, key-leaked, and chosen-plaintext attackers) enabled the model to attain information-theoretic security via synchronised neural network parameters and secret keys, effectively resisting multimodal threats in complex SAGINs.

Furthermore, current methods focus on encryption algorithms designed for “fixed eavesdroppers”. Eavesdroppers fail when they can flexibly adjust their methods based on the transmission policies of legitimate devices, a phenomenon we refer to as “evolved eavesdroppers” possessing adaptive learning capabilities. According to [136], eavesdroppers adjust neural

TABLE VI
SUMMARY OF GAI SOLUTIONS FOR SECURITY THREATS AFFECTING AUTHENTICITY IN SAGINS, WHERE MULTIPLE ATTACKS IN NETWORK LAYER REFER TO TRAFFIC INJECTION, SPOOFING, AND REPLAY ATTACKS.

Layer	Ref.	Segment	Security Threat	GAI Approach	Description
Physical Layer	[110]	Cross-Domain	GPS spoofing	TBM-GAN	DroneDefGANt for external attacks and internal faults
	[115]	Cross-Domain	Authentication breach	LSTM-CGAN	Physical-Layer authentication adopting CSI and CGAN-LSTM
	[106]	Space	GNSS spoofing	GAN	High-precision detection against wrong timing and positioning
	[109]	Space	Unknown spoofing	VAE-WGAN	A GNSS spoof detection model comprising VAE-WGAN
	[111]	Space	GPS spoofing	GAN	Combine GAN and RFF for detection and location estimation
	[105]	Air	ADS-B spoofing	CNN-based GAN	GAN-CNN for detecting ADS-B spoofing using raw I/Q signals
	[107]	Ground	Spoofing in IVN	GAN	A GAN-assisted contextual pattern-aware IDS
	[108]	Ground	Spoofing attack	DNN-based VAE	IDS to address the dynamic and evolving cyberthreatsy in IoV
	[112]	Ground	Spoofing for IoT	CNN-based GAN	Spoof detection and authentication using RFF-CWT-based GAN
	[113]	Ground	Spoofing for AIS	Conditional VAE	A RFFI-based CVAE detection framework
	[116]	Ground	Spoofing for IIoT	Hierarchical VAE	CIR-based authentication for static and mobile IIoT
Data Link Layer	[119]	Each Segment	MAC spoofing	ANN for GAN	An ANN-based spoofing detection method
	[69]	Each Segment	ARP spoofing	GAN	A federated self-supervised learning model for traffic analysis
	[123]	Air	ADS-B Spoofing	LSTM-GAN	An attack detection and marking model based on LSTM-GAN
	[120]	Ground	MAC spoofing	Adversarial for GAN	A randomized dynamic target defense framework for IoT system
Network Layer	[124]	Each Segment	Traffic injection	LSTM-GAN	A traffic augment solution for IDS system
	[125]	Each Segment	Traffic spoofing	LLM-GAN	A Magteon-Turing L3TM framework
	[126]	Ground	Multiple attacks	GAN	Time-series prediction of cyber attacks in vehicle networks
	[127]	Ground	Replay attack	GAN	GAN as an attack generation for training defense systems
	[128]	Ground	Minor-Class attack	DRL-based GAN	An anomaly detection system for optimum routing in H-CIoT
	[129]	Ground	Traffic injection	GAN	Lightweight GAN for traffic prediction and routing in SDN-IoT
	[130]	Ground	Multiple attacks	GAN	A GAN-based malicious attack detector for IoT systems

network parameters through a symmetric training process that emulates legitimate users, thereby increasing the risks to GAI-assisted SAGINs. The authors proposed a secure autoprecoder (ASAP) framework based on adversarial learning for MIMO wiretap channels featuring "evolved eavesdroppers." The two-stage adversarial training jointly optimises modulation and precoding, thereby enhancing the security-reliability trade-off.

Given the unique role of semantic communication in SAGINs [137], GAI-assisted encrypted semantic communication requires investigation. A GAN-based encrypted semantic communication system was proposed by [138]. An effective defence against eavesdropping attacks was achieved in a generic semantic communication scenario with shared knowledge through the design of a symmetric encryption structure and an adversarial training algorithm. The system improves the balance between privacy-preserving capabilities and versatility in SAGINs.

GAI-assisted physical layer security: GAI enhances physical layer security by leveraging channel randomness and active defense strategies against sophisticated eavesdropping. For instance, an adversarial training framework was proposed to harden modulation classifiers against evasion attacks [139]. By designing an iterative "Least Likely" white-box attack strategy, this active anti-eavesdropping method improved communication privacy by reducing the classification accuracy of eavesdroppers to random guessing levels (e.g., from 98% to 4%) in software-defined radio (SDR) tests.

In addition, beamforming significantly contributes to physical layer security. [140] proposed an actor-critic GDM (AC-GDM) scheme optimizing precoding and IRS phase shifts, leveraging GDM's denoising capability in IRS-assisted IoT. Another proactive defense method integrates perturbations into beamforming design [141]. As demonstrated in [142], an adversarial defense embedded waveform design (ADEWD) solution adopted GAN-generated amplitude-controlled perturbations superimposed on signals to disrupt eavesdropper identification while ensuring reliable communication. In addition, [143] extended this approach to cooperative beamforming in AAV swarm with a GDM-enabled twin delayed deep deterministic policy gradient (GDMTD3) method to tackle the non-convex, NP-hard dynamic problem and promote the secrecy rate.

Beyond secrecy rate, secrecy energy efficiency (SEE) is critical for energy-constrained SAGINs [144]. [145] designed a mixture of experts (MoE)-based GDM RL algorithm, improving SEE and resisting eavesdropping. Additionally, the authors in [146] further considered the energy constraints for HAP in SAGINs and proposed a GAN-empowered deep RL framework named Gen-DRL. Integrating bidirectional LSTM-enhanced GANs into the policy network, Gen-DRL dynamically predicted channel states and adapted to environment while capturing long-term temporal dependencies. It achieved 5.1%-12.43% higher SEE than benchmarks, exhibiting superior robustness against mobile user scenarios.

GAI-assisted covert communication: Encryption and physical layer security cannot prevent adversaries from detecting communication behaviors, risking data reliability and exposing source locations. In SAGINs, covert communication outperforms by ensuring "communication undetectability" (e.g., minimizing signal detection probability). [147] developed a GAN framework optimizing transmit power against eavesdroppers with uncertain thresholds in uplink FL. Extending this, the authors of [148] proposed a model-driven GAN (MD-GAN) framework, featuring a GAN-based joint trajectory and power optimization algorithm when only partial channel information is available. In addition, [149] introduced an AE-based GAN for water-to-air communication scenario based on optical wireless communication (OWC), which optimized a signal generator and a decoder to produce covert signals statistically indistinguishable.

For AAV-assisted jamming, [150] proposed a data-driven GAN (DD-GAN) framework to address the eavesdropping threats in downlink satellite-ground communication. Targeting AAV-assisted jamming scenarios with partial environmental knowledge, the method integrated genetic algorithm-generated samples to co-optimize AAV jamming power and 3D trajectory through adversarial training. For non-terrestrial networks (NTN) IoT, [151] developed a tripartite collaborative-adversarial network (TCAN) based on GANs. This framework integrated a generator (emulating legitimate transmitters), a cooperative classifier (receiver), and an adversarial discriminator (attacker) to dynamically balance covert signal design and detection. Beyond physical signals, text steganography is applied in covert communication. [152] proposed StegAbb, a GPT-3-based linguistic steganographic method for SDRs. Leveraging GPT-3's language generation capabilities, StegAbb transformed cryptographic abbreviations into natural-looking cover texts, optimizing the security-textuality balance.

2) Data Link Layer: Eavesdropping at this layer takes advantage of unencrypted protocol frame structures, such as MAC addresses, and inherent protocol vulnerabilities, facilitating data interception, tampering, and impersonation. Intrusion detection systems (IDSs) function as essential safeguards for SAGINs, integrating with GAI to facilitate adaptive threat detection, thereby ensuring data confidentiality and integrity.

In [153], the authors utilised WGAN to generate network traffic data, addressing data scarcity and class imbalance for multi-stage attack detection in distributed SAGINs. To mitigate the training instability and adversarial vulnerability of distributed GANs, [154] proposed a GAN-based robust intrusion detection system for the security of the Internet of Flying Things (IoFT), addressing the issue of limited attack diversity in public datasets. To enhance few-shot detection, [155] introduced a semi-supervised GAN for anomaly detection. The method addressed redundant feature issues and data imbalance in high-dimensional traffic data targeting SAGINs by employing positive-sample-only training.

Additionally, FL facilitates the improvement of model generalisation in generative artificial intelligence by enabling decentralised collaborative training while maintaining data privacy [156]. A framework for federated learning (FedDWM) that incorporates conditional GAN (CGAN) was proposed in [157] to tackle data privacy and class imbalance issues in IDS for SAGINs. The proposed algorithm utilises CGAN for synthesising attack samples to achieve balanced training and incorporates dynamic weight-momentum aggregation, resulting in improved model convergence and enhanced detection robustness. Evaluations of the CIC-IDS2017 and CSE-CIC-IDS2018 datasets revealed an accuracy of 95.74% and an F1-score of 94.29%. GAI-FL introduces a new framework for secure distributed content generation in SAGINs.

In addressing the multi-class IDS, [158] introduced an autoencoder-based multi-task learning (MTL) framework aimed at IoT networks. To address the limitations of single-task learning in the context of rare attacks and data imbalance, this method combined convolutional autoencoders for feature extraction with multi-task learning parameter-sharing mechanisms, further improved by stochastic weight averaging for model optimisation. The feature generation capability of the self-encoder serves as a foundation for botnet detection. For instance, [159] combined traditional machine learning with GANs to detect dynamic botnet traffic, achieving an accuracy exceeding 98% on the UNSW-NB15 dataset while minimising false positives.

To address the dynamic characteristics of SAGINs, [160] introduced a Wasserstein-distance-based composite GAN (WC-GAN) for dynamic access control in Internet of Vehicles (IoV) systems, effectively mitigating mode collapse and gradient vanishing issues present in conventional GANs. The WC-GAN produced synthetic behavioural data to enhance limited training samples and integrated it with real datasets to train a neural network for real-time risk assessment. The evaluations indicated that the hybrid-trained risk predictor achieved an accuracy of 87% and a false-negative rate of 5.2%.

3) Network Layer: Breaches at this layer pose a risk of exposing sensitive data, such as user identities and interaction patterns, which directly contradicts SAIGN's objectives of preserving privacy. GDM can be utilized to synthesize data to replace the real data for privacy protection and VAE can be adopted to generate synthetic samples of rare class for training against malicious data injection [161].

A GAN-based traffic feature hiding model (TFHM) was proposed in [162], utilising GRU-enhanced generators to maintain contextual dependencies in the context of encrypted traffic analysis. Building on this, [163] introduced a GAN-based chaotic logistic encryption technique for the protection of IoV trajectory data. This approach addresses the limitations of traditional chaotic encryption related to static key distribution and low plaintext-key correlation, while exhibiting resilience against noise, MITM, and differential attacks. Furthermore, the integration of GAN with LSTM facilitates dynamic defence and attack prediction. Fig. 3 demonstrates that [164] explored a hybrid LSTM-GAN model to mitigate the issues of high false positives and inadequate generalisation identified in [163]. This approach combines LSTM for capturing sequential dependencies with GANs for generating adversarial scenarios. Experimental results indicated a detection accuracy of 92.5%, a 15% decrease in false positives relative to standalone LSTM, and a 25% enhancement in threat prediction.

Additionally, [165] proposed a secure routing framework that integrates GAN and blockchain technologies for edge-assisted wireless sensor networks (WSNs). This framework facilitates the co-optimization of energy efficiency and confidentiality by employing dynamic bio-inspired clustering and reinforcement learning-based scheduling. Furthermore, [166] enhanced security by transitioning from reactive defence to proactive prediction. Utilising dynamic risk modelling and multi-policy collaboration with a Wasserstein distance-based

TABLE VII
SUMMARY OF GAI SOLUTIONS FOR SECURITY THREATS AFFECTING CONFIDENTIALITY IN SAGINS, WHERE MULTIPLE ATTACKS IN DATA LINK LAYER REFER TO DATA TAMPERING, PRIVACY LEAKAGE, AND IMPERSONATION, AND MULTIPLE ATTACKS IN NETWORK LAYER REFER TO PRIVACY LEAKAGE AND TRAFFIC EAVESDROPPING.

Layer	Type	Ref.	Segment	Security Threats	GAI Approach	Description
Physical Layer	**Encryption**	[132]	Each segment	Eavesdropping	WGAN-AE	A physical layer key generation method
		[135]	Each segment	Various attacks	GAN	A multiparty adversarial encryption model
		[136]	Each segment	Intelligent Eavesdropping	AE for VAE	A secure autoprecoder for MIMO wiretap channels
		[138]	Each segment	privacy leakage	GAN	An encrypted semantic system for privacy preserving
		[134]	Air	Eavesdropping	GAN	An intelligent soft handover for UAV-enabled framework
		[133]	Ground	Key leakage	GAN	Key defense in dynamic light-fidelity networks
	Physical Layer Security	[139]	Each segment	Eavesdropping	GAN	An enhanced and robust modulation classifier
		[146]	Satellite/Air	Eavesdropping	LSTM-GAN	A GAI-based DRL framework for SEE
		[143]	Air	Eavesdropping	GDM	Beamforming for multi-objective SEE optimization
		[140]	Ground	Eavesdropping	GDM	A actor-critic GDM-enhanced beamforming scheme
		[142]	Ground	Eavesdropping	GAN	An adversarial defense embedded waveform design
		[145]	Ground	Eavesdropping	MoE-GDM	A MoE-GDM-based resource allocation strategy
	Covert Communication	[147]	Ground	Eavesdropping	GAN	An optimization framework to counter attackers
		[148]	Air	Eavesdropping	GAN	A joint trajectory and power optimization algorithm
		[149]	Water/Air	Eavesdropping	AE-GAN	A covert signal generation scheme for OWC
		[150]	Satellite/Air	Eavesdropping	DD-GAN	Optimization for AAV's power and trajectory
		[151]	Air/Ground	Eavesdropping	GAN	A covert signal design and detection framework
		[152]	Each segment	Eavesdropping	LLM	A GPT-3-based linguistic steganographic method
Data Link Layer	**Intrusion Detection**	[155]	Each segment	Data tampering	GAN	A hybrid high-dimensional anomaly detection model
		[157]	Space/Ground	Privacy leakage	CGAN	CGAN-FL for data privacy and class imbalance
		[153]	Ground	Privacy leakage	WGAN	A traffic data generation scheme
		[154]	Ground	Privacy leakage	GAN	A GAN-based robust IDS for IoFT security
	Multi-class Intrusion Detection	[159]	Each segment	Impersonation	GAN	A hybrid approach for dynamic botnet traffic detection
		[158]	Ground	Multiple attacks	VAE	Autoencoder-based multi-task learning framework for IoT
		[160]	Ground	Multiple attacks	WC-GAN	A wasserstein-distance-based composite GAN in IoV
Network Layer	**Traffic Encryption**	[162]	Each Segment	Privacy leakage	GAN	Traffic feature hiding model for encrypted traffic analysis
		[165]	Each segment	Eavesdropping	GAN	A GAN-blockchain integrated secure routing framework
		[163]	Ground	Eavesdropping	GAN	A GAN-based chaotic logistic encryption method for IoV
	Intrusion Detection	[164]	Each segment	Multiple attacks	LSTM-GAN	LSTM-GAN for dynamic defense and attack prediction
		[166]	Ground	Multiple attacks	WC-GAN	A proactive prediction solution for IoV

combined GAN, communication efficiency was optimised while ensuring the confidentiality of IoV in SAGINs.

4) Summary and Lessons Learned: As summarized in Table VII, GAI enables dynamic encryption from channel characteristics, enhances physical layer security via noise-like signals, and optimizes covert communication parameters to evade detection. In addition, GAI synthesizes adversarial traffic for robust IDS, mitigates data scarcity in threat models, and obfuscates traffic patterns. The inherent adaptability of GAI models allows FL to learn and respond to the dynamic topologies, volatile channels, and diverse threat landscapes characteristic of SAGINs, offering a level of agility unattainable by static security solutions. Despite these significant promise, challenges remain:

- *Resource constraints:* High complexity and energy consumption hinder deployment on resource-limited edge/aerial/satellite nodes within SAGINs.
- *Fighting vulnerability:* GAI vulnerability to ”evolved eavesdroppers” enables training mimicry and evasion attacks.
- *Cross-domain integration:* Heterogeneous segment integration (satellite/aerial/terrestrial) demands standardized frameworks and cross-domain coordination.
- *High scalability:* Scalability to massive mobile networks requires real-time inference via algorithmic/system co-design.

IV. GAI FOR INTEGRITY AND AVAILABILITY SECURITY IN SAGINS: CURRENT RESEARCH

This section summarizes integrity tampering and availability disruptions that may be encountered in SAGINs, and overviews the corresponding GAI-based solutions in detail.

A. GAI for Integrity Tampering

1) Physical Layer: Attacks such as jamming, spoofing, and tampering can all affect data integrity to some degree [167], [26]. Data anomaly detection and data reconstruction are the primary means of verifying and ensuring data integrity [17].

Anomaly detection: It identifies deviations from normal patterns in datasets. GAI improves this capability by offering self-supervised anomalous data samples for training, facilitating high-dimensional anomaly detection and accurate localisation in complex and unfamiliar scenarios. A discriminative autoencoding framework (Dis-AE) was introduced by [168], which synergistically integrates GAN and autoencoders for semi-supervised anomaly detection. Dis-AE encounters implementation challenges in SAGINs owing to the complexity of

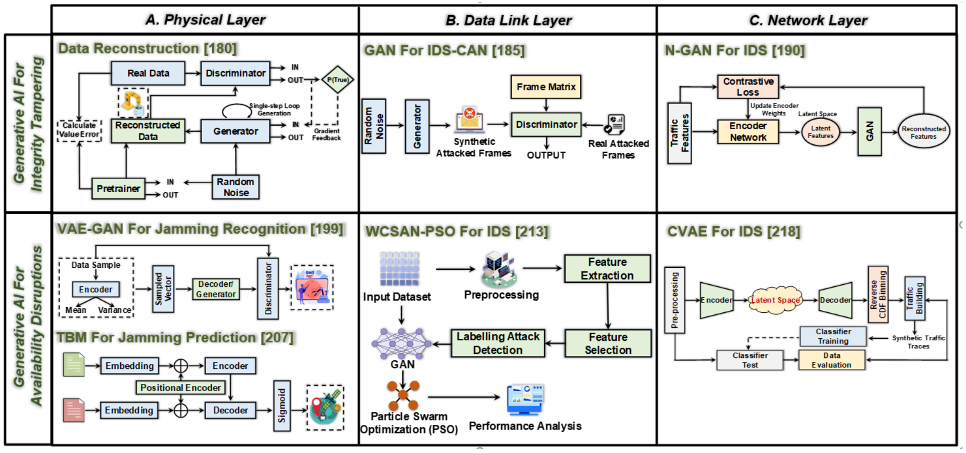

Fig. 4. Typical GAI-based approaches for integrity tampering and availability disruptions in SAGINs. For integrity tampering, [180] demonstrates a data reconstruction method based on MTS-GAN; [185] introduces a GAN-based IDS for CAN; [190] illustrates an encoder-based GAN for IDS. For availability disruptions, [199] and [207] present VAE-GAN-based and TBM-based approaches for jamming recognition and jamming prediction, respectively; [213] integrates PSO with GAN for IDS; [218] introduces how to apply VAE in IDS.

high-speed spatio-temporal data. To tackle these challenges, [169] introduced an anomaly detection model that combines improved GAN (IGAN) with LSTM networks for ADS-B-based air transportation systems.

In addition to high-precision detection, [170] facilitates accurate localisation of anomalous data. The authors introduced a cross-correlation graph-based encoder-decoder GAN for the purposes of anomaly detection and localisation. Expanding upon [170], [171] addressed the issue of pattern sensitivity in scenarios involving sample imbalance. A hybrid model combining K-means, VAE, and Support Vector Data Description (SVDD) was developed to tackle anomaly detection in multidimensional spacecraft telemetry characterised by imbalanced and unlabelled samples. In the context of contaminated datasets, frameworks based on normalising flows have been effective in reducing training biases [172]. Additionally, a convolutional variational autoencoder-GAN (CVAE-GAN) has facilitated zero-shot anomaly detection in AAV systems in the absence of contaminated data [173].

To enhance defence against multimodal threat coverage in AAV swarms, [174] proposed a hybrid anomaly detection system that integrates GANs and FL. The framework addresses in-flight anomalies, such as GPS jamming and spoofing, as well as network attacks including blackhole, grayhole, and flooding. It integrates an unsupervised stacked autoencoder with federated learning for flight anomaly detection and employs a supervised LightGBM classifier augmented by generative adversarial networks for data balancing. To address data heterogeneity, [175] utilised deep federated learning (DFL) and VAEs as local models in satellite communication systems, focussing on mitigating privacy leakage and fairness issues inherent in centralised methods.

Current methodologies primarily focus on recognising the information contained within the IP/TCP header, rather than directly analysing the payload. The multimodal process of LLM interprets data content as a specific ”language” of the device, facilitating deeper semantic feature extraction from network behavioural characteristics to content semantics. A content-based IoT detection framework utilising ChatGPT 4.0 Turbo embeddings and LSTM temporal modelling attained an accuracy of 99.75% through the extraction of semantic features from network behaviour [176].

Data reconstruction: Data reconstruction in SAGINs faces challenges including heterogeneity and resource constraints. GAI models exhibit the ability to reconstruct high-dimensional data, including channel states and images, from low-dimensional observations. For instance, [177] presented an adversarial autoencoder (AAE) framework designed for simultaneous anomaly detection and signal reconstruction in sub-Nyquist sampled spectrum monitoring scenarios relevant to 6G and satellite systems. This architecture combines an autoencoder with adversarial discriminators to enable simultaneous signal recovery. The AAE exhibits a delay in detection and has restricted capabilities for attack localisation. In response to these limitations, [178] introduced a CVAE-based deep learning approach for the detection of cyber-physical attacks in distribution systems, attaining a detection accuracy of 98% on the BATADAL dataset.

To improve high-precision data reconstruction, [179] established a framework based on a denoising diffusion probabilistic model (DDPM) for the compression of satellite monitoring data. The proposed method mitigates the limitations of existing dictionary-dependent sparse representation techniques by utilising deep generative modelling to effectively capture the intricate distributions of vibration signals. Similarly, as shown in Fig. 4, [180] proposed a high-precision reconstruction method utilising multi-component time series GAN (MTS-GAN) to address electromagnetic data loss caused by equipment failures in SAGINs.

Ensuring the integrity of image data is essential in SAGINs. A multi-modality semantic-aware framework for vehicular networks was introduced by [68], utilising GAI-based image generation and reconstruction. A single-input multi-task reconstruction framework utilising an efficient pyramidal GAN for remote sensing images was proposed by [181]. This method is hindered by over-smoothed textures resulting from pixel-wise reconstruction loss. A GAN-based super-resolution method (SRGAN) for AAV image enhancement was designed to address the critical challenge of low-resolution artefacts resulting from hardware constraints and motion blur

[182]. SRGAN primarily emphasises pixel-wise feature mapping, thereby overlooking the spatial-spectral interdependencies present in multisensor data. In response to this challenge, [183] proposed TransGAN-CFR, a novel framework that combines transformer and GAN for cross-modal reconstruction. The method utilised window-based multi-head self-attention and depthwise-convolution-enhanced feedforward networks to achieve accurate multispectral texture recovery.

2) Data Link Layer: Integrity at the data link layer requires that data frames remain unchanged and consistent throughout transmission, protecting against unauthorised alteration or corruption. Threats predominantly leverage protocol frame structures, such as ethernet frame tampering, and vulnerabilities, including ARP/DHCP spoofing and VLAN hopping attacks, to forge, inject, or truncate frames.

The authors in [184] introduced a GAN-enhanced hybrid machine learning framework aimed at the real-time detection of ARP spoofing attacks. The approach combined dynamic ARP cache validation with CNN classifiers, utilising GAN-generated synthetic attack traffic to enhance model generalisation to previously unobserved attack patterns. It relies on static feature engineering and lacks the capability to dynamically generate adversarial attack samples.

The enhanced GAN-based IDS for automotive CAN networks was proposed in [185], as illustrated in Fig. 4. An enhanced discriminator was developed to detect data tampering through the verification of signal value ranges. Experiments indicated enhanced efficacy in identifying DoS, injection, masquerade, and data tampering attacks, attaining 99.9% precision and recall.

To address the dynamic and heterogeneous characteristics of SAGINs, [186] proposed a GAN-enhanced anomaly detection framework that generates synthetic attack patterns, such as data tampering and frame spoofing. The lightweight classifiers were developed to detect unauthorised alterations in network frames. This framework is limited as it does not incorporate proactive remediation mechanisms and functions exclusively at the threat identification stage. In response, [187] proposed a hybrid security evaluation framework for future IoV, which addresses integrity verification challenges in complex attack scenarios through the multidimensional integration of vehicle dynamics modelling. The authors in [188] integrated a generative neural network for data imputation with blockchain technology to provide tamper-proof storage, facilitating the prediction of missing attributes and ensuring dataset completeness. Evaluations demonstrated a 17%-23% improvement in imputation accuracy compared to traditional SVM method, while blockchain consensus mechanisms reduced tampering risks by 34%.

3) Network Layer: The integrity of the network layer in SAGINs is crucial for ensuring reliable routing and packet delivery; however, it is susceptible to threats such as MITM attacks and malicious routing injection, particularly during dynamic topology changes. GAI tackles these challenges by employing adversarial traffic generation for routing verification hardening, spatiotemporal anomaly reconstruction, and topology-aware integrity self-healing mechanisms, thereby establishing a proactive defence against multi-domain coordinated attacks.

In the context of anomaly detection, [189] proposed a GAN-based framework that utilised latent interactions to identify hardware trojans (HTs) affecting routing components and to classify attack types. Evaluations indicated a 91Additionally, [190] introduced N-GAN to address the high false alarm issue by integrating partial attack semantics via weak supervision, in contrast to the unsupervised GAN-based IDS presented in [189]. The study in [191] introduced a GAN-based method for generating synthetic attack traffic to mitigate issues of data scarcity and class imbalance in intrusion detection systems (IDS). The authors utilised vanilla GAN, WGAN, and conditional tabular GAN (CTGAN) to generate high-fidelity Botnet attack samples from the CIC-IDS2017 dataset, resulting in an increase in the Botnet detection F1-score from 0.60 to 0.90, while ensuring robustness across various attack classes. Furthermore, [130] examined a GAN-based dynamic MAD, in which the generator produced varied attack patterns while the discriminator worked to improve detection robustness collaboratively.

Stealthy cyber attacks deliberately avoid detection by imitating legitimate network behaviours and conforming to protocol specifications while carrying out malicious payloads. In [192], the authors developed a hybrid adversarial framework that integrates GAN and constrained optimisation to generate protocol-compliant attack samples in industrial control systems. The approach focused on injection, function code, and reconnaissance attacks, implementing ICS-specific constraints such as immutable packet headers while modifying negotiable features. This resulted in over 80Extending beyond single GANs, [193] designed a hybrid GAN-driven framework for secure cluster-based routing in ad-hoc networks, jointly addressing energy efficiency and malicious node resilience. Additionally, LLM can collaborate effectively with GANs. A GAN-based intelligent fuzzer (DCGAN-MNFuzz) was introduced by [194], which generated protocol-aware mutated payloads through adversarial training. This was combined with an LLM-powered dynamic risk assessment engine, validated in a smart airport IoT testbed.

In addition to IDS design, GAI enables protocol and packet encryption. In this context, [195] proposed a GAN-based framework for generating balanced datasets to mitigate imbalanced attack detection in drone video analytics, specifically focussing on replay, packet injection, and physical capture threats. The integration of MAVLink's continuous authentication methods, such as digital signatures, along with lightweight encryption, improved data integrity against network-layer tampering while preserving operational efficiency.

4) Summary and Lessons Learned: As summarized in Table VIII, GAI enables proactive defense through two synergistic capabilities: high-dimensional anomaly detection and robust data reconstruction. Meanwhile, generative models transcend traditional threshold-based methods by enabling semantic-aware validation—interpreting payloads as domain-specific "languages" and verifying contextual coherence beyond syntactic checks. Despite these transformative advantages, critical challenges remain:

- *Resource-needs conflict:* Computational intensity of GAI

TABLE VIII
SUMMARY OF GAI SOLUTIONS FOR SECURITY THREATS AFFECTING INTEGRITY IN SAGINS. MULTIPLE ATTACKS IN PHYSICAL LAYER REFER TO JAMMING, SPOOFING, AND TAMPERING; MULTIPLE ATTACKS IN DATA LINK LAYER REFER TO ARP/DHCP SPOOFING AND VLAN HOPPING ATTACK; MULTIPLE ATTACKS IN NETWORK LAYER REFER TO MITM ATTACKS AND MALICIOUS ROUTING INJECTION.

Layer	Type	Ref.	Segment	Security Threats	GAI Approach	Description
Physical Layer	**Anomaly Detection**	[168]	Each segment	Data tampering	AE-GAN	A semi-supervised anomaly detection model
		[170]	Each segment	Data tampering	GAN	A cross-correlation graph-based encoder-decoder
		[171]	Space	Data tampering	VAE	A hybrid high-precision detection model
		[175]	Space/Ground	Data tampering	VAE	A heterogeneous multidimensional detection model
		[169]	Air	Data tampering	LSTM-GAN	A detection model for ADSB-based system
		[173]	Air	Multiple attacks	VAE-GAN	A CVAE-GAN for zero-shot learning for AAV
		[174]	Air	Multiple attacks	GAN	Anomaly detection for anomalies and attacks
		[176]	Ground	Multiple attacks	LLM	Anomaly detection using ChatGPT 4.0 Turbo
	Data Reconstruction	[180]	Space/Air/Ground	Multiple attacks	GAN	A high-precision reconstruction method
		[181]	Satellite/Ground	Multiple attacks	GAN	A multitask image reconstruction framework
		[177]	Space	Multiple attacks	VAE	Joint anomaly detection and signal reconstruction
		[179]	Space	Multiple attacks	GDM	High-precision reconstruction for satellite data
		[183]	Space	Multiple attacks	TBM-GAN	TransGAN-CFR for cross-modal reconstruction
		[182]	Air	Multiple attacks	GAN	Super-resolution for AAV image enhancement
		[178]	Ground	Multiple attacks	VAE	A detection approach for cyber-physical attack
		[68]	Ground	Multiple attacks	GAI	A multi-modality semantic-aware framework
Data Link Layer	**Intrusion Detection**	[184]	Each segment	ARP spoofing	GAN	A hybrid ML framework for real-time detection
		[185]	Each segment	Multiple attacks	GAN	A IDS for automotive CAN networks
	Prediction	[188]	Each segment	Frame tampering	GAN	Data imputation for secure blockchain
	Anomaly Detection	[186]	Ground	Frame spoofing	GAN	The lightweight frame classifier
		[187]	Ground	Frame tampering	GAN	A security evaluation framework for future IoVs
Network Layer	**Intrusion Detection**	[189]	Each segment	Packet dropping	GAN	GAN-based detection for hardware trojans
		[190]	Each segment	Packet dropping	GAN	GAN for partial semantic attack
		[191]	Each segment	Multiple attacks	GAN	A synthetic attack traffic generation approach
		[192]	Each segment	Stealthy attack	GAN	A framework for protocol-compliant attack samples
		[195]	Air	Packet injection	GAN	GAN for authentication and lightweight encryption
		[194]	Air/Ground	Multiple attacks	LLM-GAN	A GAN-based intelligent fuzzer
		[130]	Ground	Multiple attacks	GAN	A GAN-driven dynamic malicious attack detector
		[193]	Ground	Stealthy attacks	GAN	GAI for secure cluster-based routing

models (e.g., transformer-GAN hybrids) conflicts with satellite/avionic latency/energy constraints.

- *Over-reliance on dataset:* Training-data dependence creates vulnerability to adversarial poisoning that may induce model biases.
- *Explainability deficit:* Explainability deficits in deep generative models hinder breach root-cause analysis.
- *Cross-Domain gaps:* Layer-specific defenses (e.g., physical-layer recovery, network-layer verification) lack holistic orchestration across SAGINs.

B. GAI for Availability Disruptions

1) Physical Layer: Jamming attacks compromise communication availability by transmitting high-power signals that degrade legitimate signal-to-noise ratio (SNR). GAI-based research addresses jamming recognition and mitigation.

Jamming recognition: Protection mechanisms in SAGINs initiate dynamic spectrum reconfiguration and topology adaptation for swift defence. GANs utilise zero-sum game dynamics for the recognition of multi-attack signals, akin to spoofing detection [106]. Leveraging this advantage, [196] introduced a robust IDS utilising GAN and adversarial sample regularisation, effectively addressing spoofing and jamming attacks encountered by AAVs that depend on GPS navigation. Nonetheless, a single GAN would face challenges in managing multi-dimensional data characterised by complex relationships, such as high-resolution images. To address this issue, [197] developed a conditional tabular GAN (CTGAN) that synthesises data rows from discrete columns, effectively resolving unbalanced distributions. Furthermore, TBM may be employed to improve the efficacy of GAN. The DroneDefGANt, as discussed in [110], utilised the multi-head attention mechanism of transformers to address discriminator gradient challenges in GANs. Furthermore, autoencoders enhance GANs in situations with limited labelled data, such as AAV swarm [174].

In contrast to GANs, VAEs generate interpretable features for jamming recognition by utilising latent space recognition. In [198], the authors introduced a VAE-based unsupervised framework for detecting anomalous interference in MIMO-OFDM ISAC through reconstruction probabilities. The advantages of VAEs are particularly evident in few-shot learning. The study [199] explored the latent space of small sample datasets through the use of VAE, subsequently sampling and decoding this latent space to facilitate dataset expansion prior to GAN discrimination. Centralised VAEs encounter difficulties in distributed SAGINs because they are susceptible to jamming. A federated augmented aggregate training algorithm was proposed by [200], which integrates spectral function feature extraction through CVAE to address the issue of unknown interference detection in distributed SAGINs. Additionally, in response to intelligent jammers that replicate legitimate signals, [201] introduced a VAE that integrates a dynamic adaptive spectro-temporal resilient filter (DASTRF), further augmented by a vision transformer (ViT) and LSTM. The proposed FL-ViT-LSTM-VAE enhances the detection of signal-replicating smart jammers through the optimisation of time-frequency distribution feature separation, thereby ensuring the safe operation of AAVs.

Jamming recognition models are primarily characterised by TBMs [202] and GDMs [203]. For instance, [202] introduced a distributed radar multi-interference identification method utilising a transformer network and adaptive beamforming to mitigate the decline in identification performance caused by the overlap of multiple interference sources in the time-frequency domain. Additionally, to address the issues of inefficient intrusion detection and privacy leakage in SAGIN with heterogeneous data, the authors in [203] developed a STINIDF framework that collaboratively trains the conditional diffusion model (DP-CDM) via federated learning (FL) and produces global traffic data by integrating the differential privacy (DP) mechanism.

Jamming mitigation: Jamming mitigation involves both active signal suppression and passive data reconstruction methods. Active mitigation relies on the awareness of jammers and the surrounding environment. According to the available literature, [204] introduced a collaborative reinforcement learning algorithm that employs a mixture Gaussian distribution model, integrating GAN localisation with time difference of arrival techniques. In situations characterised by unknown or incomplete information, [205] developed an intelligent spectrum access algorithm that integrates GAN and DRL to mitigate interference, even in cases with up to 90% missing data. Furthermore, [206] introduced a vision transformer-based adaptive blind beamforming method (ViT-BF) aimed at jammer suppression. A proactive jamming prediction method that integrates a pseudo-random algorithm with a transformer module to forecast jammer behaviour was proposed by [207].

Passive data reconstruction emphasises recovery on the receiver's side. A conditional diffusion model (CDM)-based anti-jamming algorithm was implemented to ensure accurate reception of satellite navigation signals in complex SAGINs [208]. The algorithm incrementally introduced noise and learnt the noise distribution via the forward diffusion process, while systematically denoising the noise-embedded QPSK signal as a condition in the reverse diffusion process. Furthermore, [209] proposed GAN-inspired anti-jamming techniques for semantic communication, ensuring decoding consistency during attacks. Additionally, few-shot sample scenarios in IoT networks were examined in [210], highlighting that various jamming styles were rare and challenging to mitigate effectively. The authors proposed a meta-learning and multi-task approach for source separation to address unknown interference.

2) Data Link Layer: Availability threats, such as MAC flooding and ARP-based DoS, encompass adversarial actions that interfere with protocol operations or deplete resources, consequently denying legitimate access to network services. GAI can improve the resilience of the data link layer protocols in SAGINs against jamming or resource-exhaustion attacks through the simulation of intricate adversarial scenarios and the development of adaptive defence strategies.

In [211], the authors introduced a GAN-enhanced IDS combined with a LSTM-based MAC protocol to mitigate availability threats in underwater wireless sensor networks (UWSNs). The framework utilises GAN for real-time assessment of acoustic channel quality, including noise patterns and malicious signal interference, alongside LSTM-MAC for adaptive medium access control. This approach dynamically optimises contention access periods and implements reactive jamming to counter DoS attacks, resulting in a detection accuracy of 98.9% with 5% malicious nodes present. Based on this, [212] examined cross-layer attacks that encompass the data link and network layers, including flooding attacks that utilise IPv6 router advertisements, resulting in link layer congestion. The authors proposed a DL-based approach for detecting router advertisement flooding DDoS attacks in IPv6 networks. Integrating feature ranking algorithms with a RNN addresses the vulnerability of the neighbour discovery protocol (NDP), specifically mitigating RA flooding attacks.

The aforementioned studies depend heavily on static feature libraries and are unable to manage adversarial perturbation traffic in real time. To address this issue, [213] proposed an adversarial attack detection framework (WCSAN-PSO) utilising an optimised weighted conditional stepwise adversarial network alongside particle swarm optimisation, as illustrated in Fig. 4. The integration of GAN-generated adversarial samples with feature selection techniques enhanced the robustness of IDS. The framework demonstrated an accuracy of 99.36% for normal traffic and 98.55% for malicious traffic when assessed using the CIC-IDS2017 dataset. Utilising GAI to dynamically generate attack patterns and optimise model parameters enhanced the generalisation of IDS against evolving threats, consequently improving service availability.

3) Network Layer: The availability of the network layer in SAGINs is compromised by dynamic topology-based DDoS attacks, resource depletion in high-latency links, and routing flooding attacks. GAI facilitates federated anomaly detection and supports privacy-preserving adversarial training.

VAE demonstrates effective performance in traffic anomaly detection within SAIGINs. For instance, [214] presented a two-phase DDoS mitigation framework powered by a VAE, which incorporated cyclic queuing to optimise the balance between persistent resource consumption and detection efficiency for edge devices in dynamic SAGIN environments. Furthermore, [215] introduced an unsupervised intrusion detection system utilising deep autoencoders to capture spatiotemporal patterns of normal network flows within industrial control systems. This method attained a detection accuracy of 98.8% for DDoS attacks, accompanied by a false alarm rate of only 1.13%. However, the method presented in [215] is limited by inadequate privacy protection when dealing with non-independent and identically distributed data (non-IID) and exhibits weak cross-device generalisation. In response to this issue, [216] examined a FL-VAE scheme that integrates variational autoencoders with momentum-accelerated FL, with a specific focus on DDoS mitigation in SAGINs. The frame-

TABLE IX
SUMMARY OF GAI SOLUTIONS FOR SECURITY THREATS AFFECTING AVAILABILITY IN SAGINS, WHERE MULTIPLE ATTACKS IN NETWORK LAYER REFER TO DDOS ATTACK AND ROUTING FLOODING.

Layer	Type	Ref.	Segment	Security Threats	GAI Approach	Description
Physical Layer	**Jamming Recognition**	[196]	Space/Air	GPS jamming	GAN	A robust IDS based on GAN and sample regularization
		[197]	Air	Jamming attack	GAN	Conditional tabular GAN to handle multi-dimensional data
		[110]	Air	Jamming attack	TBM-GAN	DroneDefGANt smooths gradient during backpropagation
		[174]	Air	Jamming attack	AE-GAN	AE-GAN solves insufficient labelled data in AAV swarm
		[198]	Ground	Jamming attack	VAE	Unsupervised detection for MIMO-OFDM ISAC system
		[199]	Each segment	Jamming attack	VAE-GAN	Jamming recognition based on AC-VAEGAN
		[200]	Air/Ground	Unknown jamming	VAE	A federated augmented aggregate training algorithm
		[201]	Air/Ground	Intelligent jamming	VAE	A dynamic adaptive spectro-temporal resilient filter
		[202]	Ground	Jamming attack	TBM	A distributed radar multi-interference identification method
		[203]	Space/Ground	Privacy leakage	GDM	STINIDF through FL and generated traffic data
	Jamming Mitigation	[204]	Air	Jamming attack	GAN	A collaborative RL algorithm based on GAN
		[205]	Ground	Jamming attack	GAN	An GAN-DRL-based spectrum access algorithm
		[206]	Space	Jamming attack	TBM	A vision transformer based blind beamforming method
		[207]	Air	Jamming attack	TBM	Jamming prediction integrates pseudo-random with TBM
		[208]	Space	Signal jamming	GDM	A GDM-based anti-jamming algorithm
		[209]	Each segment	Jamming attack	GAN	A anti-jamming scheme for semantic communication
		[210]	Ground	Jamming attack	TBM	A sample less source separation algorithm
Data Link Layer	**Intrusion Detection**	[211]	Underwater	MAC flooding	LSTM-GAN	A GAN-enhanced IDS integrated with MAC protocol
		[212]	Cross-Domain	Flooding DDoS	GAN	Detection for routing flooding in IPv6 networks
		[213]	Each segemnt	MAC flooding	GAN	An adversarial attack detection framework
Network Layer	**Intrusion Detection**	[214]	Each segment	DDoS attack	VAE	A VAE-powered two-phase DDoS mitigation framework
		[215]	Each segment	DDoS attack	VAE	An unsupervised IDS using deep autoencoders
		[216]	Each segment	DDoS attack	VAE	A FL-VAE scheme combining VAE with FL
		[217]	Each segment	Multiple attacks	VAE	VAE is combined with neural architectures
		[218]	Each segment	Multiple attacks	VAE	A privacy-preserving traffic generation method
		[219]	Each segment	Multiple attacks	GAN	A hybrid IDS integrating CGAN and gradient boosting
		[220]	Each segment	DDoS attack	LLM	A GAI-driven framework using pre-trained transformers
		[221]	Each segment	DDoS attack	LLM	A GAI-powered real-time DDoS detection framework

work tackled feature heterogeneity among distributed nodes via dynamic client sampling and adaptive model retraining, resulting in a 99.97% attack recognition precision on non-IID traffic and a 32% reduction in cross-node data exchange.

VAE can be integrated with neural architectures to provide defence mechanisms against attacks in SAGINs. The study [217] utilised the anomaly detection capability of VAEs by employing probabilistic reconstruction of multi-dimensional traffic patterns. Additionally, [218] introduced a privacy-preserving traffic generation method utilising CVAE, which tackles security availability issues in SAGINs arising from dynamic traffic patterns and device heterogeneity. The authors in [219] proposed a hybrid IDS that integrates conditional GAN (CGAN) and extreme gradient boosting (XGBoost) for small sample datasets. This method decreased the physical layer deployment requirements in wireless sensor networks by utilising adversarial data generation. The approach attained 99.9% detection accuracy and a 1.83% reduction in false alarms on the NSL-KDD and CICIDS2017 datasets by integrating GAI with lightweight classifiers, thereby improving network-layer anti-jamming capabilities, especially for resource-constrained nodes in SAGINs.

In the context of DDoS vulnerabilities, [220] presented PLLM-CS, a framework driven by GAI that utilises pre-trained transformers. The model restructured network data into context-aware token sequences, effectively capturing spatiotemporal patterns in cyber threats via self-attention mechanisms, thereby overcoming the limitations of traditional IDSs in resource-constrained satellite environments. The centralised architecture presented in [220] is at odds with the distributed characteristics of SAGINs, facing challenges related to dynamic network traffic and the complexity of heterogeneous cross-domain protocols. Therefore, [221] developed Llama2-Defender, leveraging LLMs for contextual reasoning across heterogeneous nodes to overcome generalization limitations in aerial-terrestrial networks.

4) Summary and Lessons Learned: As summarized in Table IX, GAI enables robust jamming recognition by distinguishing stealthy adversarial patterns from legitimate transmissions. Meanwhile, diffusion models facilitate high-fidelity signal reconstruction from severely corrupted observations, while adversarial learning frameworks generate adaptive beamforming and spectrum access policies that dynamically circumvent interference. Crucially, GAI's synthetic data generation capability overcomes the critical limitation of attack sample scarcity, enabling the training of detectors for novel or evolving threats without requiring extensive real-world datasets. Despite these advances, critical challenges remain:

- *Fidelity vs. deployability trade-offs:* While model distillation and FL offer partial mitigation, fundamental trade-offs persist between generative fidelity and edge deployability.

- *Intrinsic vulnerability:* Maliciously crafted inputs could poison generative training data or deceive detectors through perturbation attacks, potentially turning defensive systems into availability liabilities.
- *Cross-Layer coordination:* While GAI excels at layer-specific countermeasures (e.g., physical-layer beamforming and network-layer DDoS mitigation), holistic frameworks for orchestrating generative defenses across SAGIN's protocol stack require development.

V. Tutorial 1: GDMTD3 for Multi-Objective Aerial Collaborative Secure Communication Optimization

A. Motivation

The allocation of resources for secure communication in SAGINs presents a critical and complex challenge under various attacks. This issue pertains to the dynamic allocation of constrained resources, such as transmit power, computing power, and routing, to optimise security metrics, including secrecy rate and SEE, while concurrently reducing operational costs, such as energy consumption and latency. The primary challenge is optimising problems characterised by high-dimensionality, non-convexity, and NP-hardness within the dynamic adversarial conditions and strict resource limitations of SAGINs. For instance, [143] presented a multi-objective problem aimed at optimising the excitation current weights and three-dimensional trajectories of UAVs to achieve a balance between secrecy rate and energy consumption in a dynamic high-dimensional space. Traditional AI methods, such as DRL approaches, encounter significant challenges in accurately modelling the complex probability distributions associated with high-dimensional continuous action spaces. These methods often exhibit high policy variance, unstable convergence, and suboptimal Pareto frontiers in the context of secure resource allocation [78] [222].

GDMs offer a significant advantage compared to traditional DRL in addressing complex resource allocation issues. DRL utilises deterministic or simplified stochastic policy networks, whereas GDMs employ a progressive denoising mechanism to accurately capture and sample from complex, high-dimensional state-action distributions. This facilitates the production of varied and high-quality candidate solutions for DRL's policy networks, efficiently examining the complex trade-offs between opposing objectives (e.g., secrecy rate versus energy). The integration of GDMs into the DRL framework significantly enhances the representational capacity of the policy, addressing the limitations of traditional DRL in navigating complex solution spaces and achieving robust convergence towards near-Pareto-optimal strategies under uncertainty. The GDM-driven DRL framework creates a new and scalable model for secure and efficient SAGINs.

B. System Description

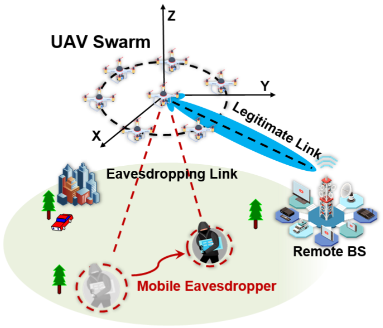

Fig. 5. A UAV swarm-enabled surveillance network transmits sensitive data to a RBS, persistently threatened by a mobile eavesdropper attempting interception through time-varying wiretap channels [143].

Fig. 5 illustrates that the UAV swarm-enabled secure surveillance network encounters ongoing threats from mobile eavesdroppers seeking to intercept communications via time-varying wiretap channels across discrete time slots. Collaborative beamforming (CB) using UAV virtual antenna arrays (UVAAs) improves directional security but creates a significant energy-security trade-off. The continuous repositioning of UAVs to achieve optimal beam patterns significantly increases energy consumption in the pursuit of maximising secrecy rates. Within this dynamic threat environment, a multi-objective optimisation problem for aerial secure communication and energy efficiency (ASCEE-MOP) is formulated to maximise the secrecy rate and minimise energy consumption as follows:

$$\max_{\boldsymbol{I},\boldsymbol{q}} \quad \left(\sum_{n=1}^{N} R_{SE}[n], -\sum_{n=1}^{N} E[n]\right) \tag{1a}$$

$$\text{s.t.} \quad 0 \leq I_k^U[n] \leq 1, \forall k \in \{1, ..., K\}, \tag{1b}$$

$$X_{\min} \leq x_k^U[n] \leq X_{\max},\ \forall k \in \{1, ..., K\}, \tag{1c}$$

$$Y_{\min} \leq y_k^U[n] \leq Y_{\max},\ \forall k \in \{1, ..., K\}, \tag{1d}$$

$$Z_{\min} \leq z_k^U[n] \leq Z_{\max},\ \forall k \in \{1, ..., K\}, \tag{1e}$$

$$0 \leq v_k^U[n] \leq V_{\max}, \forall k \in \{1, ..., K\}, \tag{1f}$$

$$||q_{k_1}[n], q_{k_2}[n]|| \geq D_{\min}^U, \forall k_1, k_2 \in \{1, ..., K\}. \tag{1g}$$

where $R_{SE}[n]$ and $E[n]$ are the achievable secrecy rate and the flight energy consumption, respectively. K and N are the number of UAVs and time slots, respectively. $\boldsymbol{I} = \{I_k^U[n]\}_{k\in K, n\in N}$ and $\boldsymbol{q} = \{q_k^U[n]\}_{k\in K, n\in N}$ are the excitation current weight matrix and the position matrix of UAVs at all time slots, respectively. $q_k^U[n] = (x_k^U[n], y_k^U[n], z_k^U[n])$ is the the coordinate of UVAA center. (1b) expresses the range constraint of the excitation current weight. Moreover, Constraints (1c), (1d), and (1e) restrict the flight area of the UAV. Constraint (1f) is the speed constrain of the UAV, and Constraint (1g) is imposed to guarantee the minimum distance between two UAVs.

This problem is non-convex, NP-hard, and dynamic due to eavesdropper mobility. Existing approaches for such optimization problems typically decompose them into separable convex subproblems solved iteratively. However, solution accuracy critically depends on the decomposition strategy. Moreover, dynamic factors, such as mobile eavesdroppers and time-varying channels, impose prohibitive computational overhead for real-time SAGIN operations. DRL provides an efficient

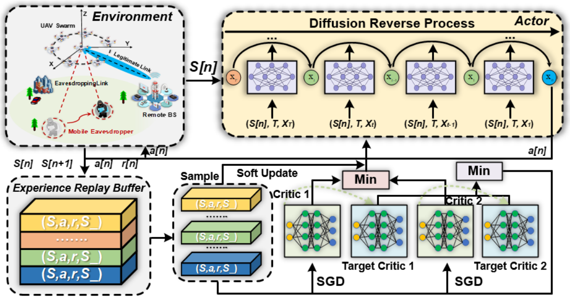

Fig. 6. The proposed GDMTD3 framework, which integrates a GDM into actor network to capture complex state features and generate optimal actions from environmental states [143].

framework for adaptive sequential decision-making through policy gradient frameworks that encode environmental dynamics into a Markov decision process (MDP). This paradigm efficiently navigates the solution space of ASCEE-MOP without explicit problem decomposition. Furthermore, GDM can be adopted in policy network to generate high-quality solutions for excitation current weights and UAV localisation, mapping the state of the environment directly to the optimal action.

C. GAI-Based Solutions

Generally, the GDM-driven DRL approach significantly improves the robustness and convergence of the policies by fusing the distributional modelling capability of diffusion models with the decision optimization framework of DRL. It is divided into three main steps when solving complex single-objective and multi-objective optimization problems with high-dimensional state-action spaces. Here, we take the GDMTD3 in [143] as an example to illustrate.

- *Step 1:* Modelling the dynamic optimisation problems as MDP tuples is fundamental, where the MDP tuple is $< \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma >$. The environment state at slot n is $\boldsymbol{s}[n]$. The action set is $\boldsymbol{a}[n] = (\boldsymbol{I}[n], \boldsymbol{q}[n])$. The reward function is defined as $r[n]$. $\mathcal{P}$ is the transition probability of the state, and γ is the discount factor. This formulation transforms ASCEE-MOP into a tractable DRL problem where UAV swarms interacts with the environment to maximize cumulative discounted rewards.
- *Step 2:* We integrate the GDM with the actor network of TD3, where the multi-layer perceptrons (MLPs)-based actor network struggles with the non-linear state space induced by mobile eavesdroppers and multi-objective trade-offs. GDMs address this limitation by modeling high-dimensional distributions for more balanced and optimized decisions in uncertain and dynamic environments. As shown in Fig. 6, GDMTD3 integrates GDM with the actor network for enhanced state-action distribution capture and a more diverse set of potential actions. The core is GDM-based action sampling. This process is similar to the Gen-DRL in [146], but this tutorial analyzes the more complex multi-objective optimisation rather than the single-objective optimisation in [146]. Furthermore, while [140], [145], and [165] consider resource optimisation in secure communication, their schemes address lower dimensions andthe algorithms may fail in dynamic and complex SAGINs. Using GDM-driven DRL to implement secure security resource allocation would support them being more applicable to real SAGINs to some extent. Specifically, in GDMTD3, based on the current state $\boldsymbol{s}[n]$, action $\boldsymbol{a}[\boldsymbol{n}]$, and initialized Gaussian distribution $\boldsymbol{x_T} \sim \boldsymbol{\mathcal{N}(0, I)}$, a denoising distribution $\boldsymbol{\varepsilon_{\theta_d}(x_t, t, s[n])}$ is deduced and then the mean of the reverse process $\boldsymbol{\kappa_{\theta_d}}$ can be computed as follows:

$$\boldsymbol{\kappa_{\theta_d}(x_t, t, s[n]) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t \cdot \varepsilon_{\theta_d}}{\sqrt{1 - \overline{\alpha}_t}} \right)}, \tag{2}$$

  where $\boldsymbol{\beta_t}$ is the variance function of variance preserving stochastic differential equations and $\boldsymbol{\alpha_t = 1 - \beta_t}$. $\boldsymbol{\overline{\alpha}_t = \prod_{k=1}^{t} \alpha_k}$. Then, a reparameterization trick that facilitates differentiable sampling is employed to compute the distribution $\boldsymbol{x_{t-1}}$ as follows:

$$\boldsymbol{x_{t-1} = \kappa_{\theta_d}(x_t, t, s) + (\widetilde{\beta}_t/2)^2 \odot \varepsilon}, \tag{3}$$

  where $\boldsymbol{\widetilde{\beta}_t}$ is a predetermined variance factor and $\odot$ is the operator of Hadamard product. Finally, we can obtain the generative distribution $\boldsymbol{p_{\theta_d}(x_0)}$ as follows:

$$\boldsymbol{p_{\theta_d}(x_0) = p(x_T) \prod_{t=1}^{T} p_{\theta_d}(x_{t-1}|x_t)}, \tag{4}$$

  where $\boldsymbol{p(x_T)}$ represents a standard normal distribution. Once the generative distribution $\boldsymbol{p_{\theta_d}(x_0)}$ is successfully trained, we can sample the action $\boldsymbol{x_0}$ from (4).
- *Step 3:* Now, we can conduct the training and execution process. The RBS governs the training process via the actor-critic framework of GDMTD3. During this step, environment interactions of the UAV swarm are continuously recorded and cached in a replay buffer. Upon completion of the training cycle, the actor network is deployed across the UAV swarm, enabling autonomous real-time adaptation to dynamic threats for sustained secure communication during operational execution.

D. Numerical Results

As shown in Fig. 6 of [143], the proposed GDMTD3 achieves a 20% higher average reward per episode compared to benchmarks (TD3, PPO, DDPG, SAC, and transformer-based TD3 methods). This performance gain stems from its generative diffusion mechanism, which enhances exploration-exploitation trade-offs in high-dimensional state-action spaces, thereby maximizing cumulative rewards.

Fig. 7 demonstrates the superior performance of GDMTD3 in terms of average secrecy rate and average energy consumption compared to baselines. This gain stems from dynamic joint optimisation of UAV excitation currents and spatial configurations, effectively countering eavesdropping threats. It exhibits 18% lower average energy consumption than conventional strategies, which is attributed to the joint optimisation of beamforming parameters and trajectory planning. Collectively, these results confirm that the GDM-DRL framework uniquely balances two competing objectives: maximizing secrecy rates while minimizing energy consumption. We can further conclude that the proposed GDM-DRL approach still performs

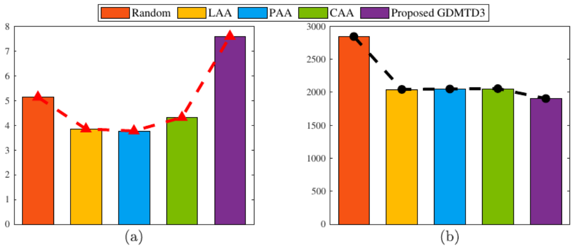

Fig. 7. Comparison results of the proposed GDM-enabled DRL approach and other four deployment policies. (a) Average secrecy rate per step [bps/Hz]. (b) Average flight energy consumption per step [J] [143].

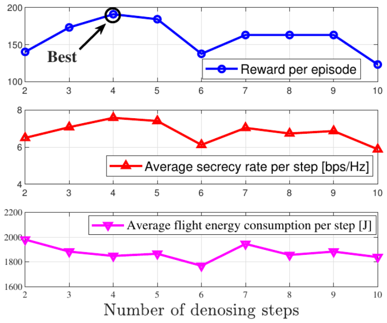

Fig. 8. Comparison of curves of GDMTD3 with different denoising steps [143]. When the number of denosing steps is 4, the reward reaches its maximum value and the average security rate is the highest and the energy consumption is the lowest. As the number of denosing steps is raised or lowered, the performance starts to decrease.

well when the security resource optimisation is more complex, e.g., cross-layer or cross-domain resource optimisation, as considered in [39], [132], and [212]. Traditional DRL, on the other hand, fails further.

In addition, the denoising step count critically modulates noise suppression efficacy and overfitting vulnerability in GDMTD3. As validated in Fig. 8, performance peaks at 4 steps for ASCEE-MOP. Beyond this threshold, marginal utility attenuation occurs due to noise-pattern overfitting, inducing oscillatory artifacts in action generation. This establishes 4-step diffusion as the optimal robustness-efficiency trade-off.

E. Lessons Learned

GDM-driven DRL robustly models high-dimensional state-action spaces to resist adversarial perturbations and data noise. Their integration replaces conventional policy networks with iterative diffusion processes, generating high-quality action sequences via multi-step refinement and overcoming traditional DRL's local convergence in complex spaces. With this mechanism, the challenge of [140], [145], and [165] not being able to cope with high-dimensional dynamic security optimization problems is addressed. It also shows vast potential in higher dimensional cross-layer and cross-domain optimization problems in [39], [132], and [212].

VI. Tutorial 2: GDM for Privacy-Preserving Mobile Crowdsensing

A. Motivation

The collection and sharing of extensive sensing data in SAGINs, including satellite telemetry, AAV trajectories, and user behaviours, poses significant risks to privacy. Attackers can acquire raw data through eavesdropping on wireless links and compromising edge servers or IoT devices, resulting in the exposure of user identity, location, and behavioural information. Traditional privacy-preserving methods exhibit significant limitations: static encryption techniques struggle to accommodate the dynamic topology of SAGIN, resulting in increased latency [131]; DP necessitates a trade-off between data utility and anonymity, thereby reducing the effectiveness of data analytics [203]; and FL utilising discriminative AI safeguards local data but suffers from slow convergence and accuracy degradation in heterogeneous device environments [156]. None of these approaches can effectively protect privacy while simultaneously addressing the constraints of dynamism, utility preservation, and real-time requirements.

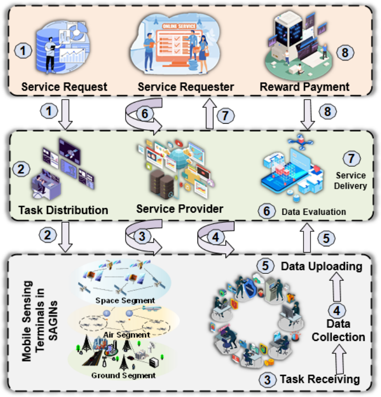

Fig. 9. The privacy preservation framework, where GAI-enabled SPPMCS can alleviate the privacy preservation concerns on sensing data, and MST's identification and location information [161].

GAI, particularly GDM, overcomes the aforementioned limitations via paradigm reconstruction. Sensitive privacy attributes are removed from the original data using a forward diffusion process, followed by the generation of synthetic data that retains equivalent statistical features while lacking privacy information, based on backward denoising [161]. In contrast to DRL, GDM mitigates issues related to high variance and sub-optimal policies in complex privacy-utility trade-offs in reinforcement learning [78]. In comparison to GAN and VAE, GDM offers a more stable training process and superior generation quality, effectively addressing the pattern collapse associated with GAN and the ambiguous reconstruction issues of VAE [71] [73].

The typical application involves secure and privacy-preserving mobile crowdsensing (SPPMCS) as discussed in [161]. This system employs GDM to substitute raw sensing data, such as vehicle images, with synthetic samples that do not compromise privacy. This approach effectively eliminates potential privacy leakage pathways and quantitatively enhances the equilibrium between privacy and task accuracy via the privacy-preserving utility index (PPUI). Consequently, it equips the SAGIN with a dynamic privacy mitigation mechanism unattainable by conventional AI methods.

B. System Description

As illustrated in Fig. 9, SPPMCS comprises three core entities: a service requester (SR) for task generation and specification, a service provider (SP), and distributed mobile sensing terminals (MSTs) performing SP-coordinated data acquisition. Specifically, this system operates in three phases: *i) Task Generation and Allocation:* The SR generates sensing tasks (specifying data type, volume, and quality) and distributes them to SPs. SPs then authenticate MSTs using broadcast encryption and assign decryption keys. However, dynamic task adjustments (e.g., for real-time data needs) incur significant key-management overhead and delays under traditional encryption, degrading system efficiency. *ii) Data Collection and Submission:* MSTs submit collected data to SPs. While conventional discriminative AI techniques (e.g., federated learning) preserve local data privacy, they suffer from slow convergence and reduced accuracy due to edge device heterogeneity (varying computational and communication capabilities). *iii) Result Evaluation and Reward Payment:* SPs evaluate data quality and reward MSTs. Some RL and blockchain-based approaches, face high computational costs and slow convergence, hindering real-time verification of reward transactions.

C. GAI-Based Solutions

GDM-driven privacy preservation mechanisms contribute to robust security defenses against threats like malicious data injection, unauthorized access, and spectrum manipulation, while simultaneously enhancing protection for both data content and terminal identification/location in SAGINs. Here, we introduce the GDM-driven SPPMCS scheme to compensate for the lack of dynamic adaptability of static encryption [138], poor accuracy of differential privacy data [102], and slow convergence of FL-based traditional AI [156] [157]. Main steps in Fig. 10 are as follows:

- *Step 1:* In sensing task publishment, SRs outsource application-specific data requirements to SPs, leveraging SP resources to overcome collection and analysis limitations.
- *Step 2:* In sensing terminal recruitment, the SP recruits qualified MSTs based on SR specifications utilizing a utility-driven reward model where compensation scales with data quality and objective fulfillment.
- *Step 3:* In sensing data collection, MSTs collect data per SR specifications. While high-capacity MSTs fulfill multi-type tasks independently, others form coalitions to achieve full coverage. This cooperation enhances task completeness but necessitates proportional reward redistribution.
- *Step 4:* We conduct GDM to finish synthetic data generation. Specifically, the GDM process involves two stages: (1) a forward diffusion phase that incrementally adds noise to original training images to learn latent features; (2) a reverse diffusion (denoising) phase that reconstructs realistic synthetic images from noise. By replacing a portion of real-world data with these synthetic samples during data submission, MSTs mitigate privacy exposure without degrading downstream task performance (e.g., vehicle detection). GDM eliminates the lack of dynamic adaptability inherent in encryption in [138]. Synthetic data requires no cryptographic keys or complex key-distribution protocols, enabling seamless adaptation to new data types/privacy requirements. This GDM mechanism can also be adopted to enhance DP with poor data accuracy in [102], by generating structurally consistent, high-fidelity synthetic data. In addition, GDM avoids the slow convergence of FL-based privacy in [157]. It operates locally or at the SP without iterative parameter exchanges, reducing latency. Synthetic data submission incurs minimal overhead compared to FL's multi-round coordination.

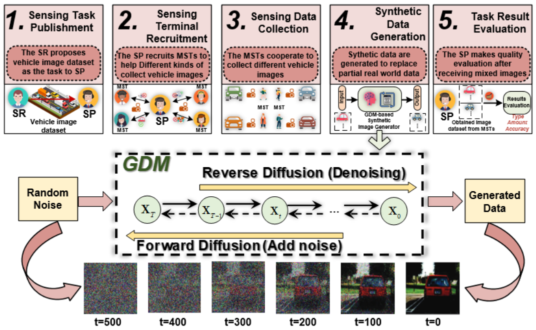
1. Sensing Task Publishment
The SR proposes vehicle image dataset as the task to SP
SR
SP
Vehicle image dataset
2. Sensing Terminal Recruitment
The SP recruits MSTs to help Different kinds of collect vehicle images
MST
MST
SP
MST
MST
3. Sensing Data Collection
The MSTs cooperate to collect different vehicle images
MST
MST
4. Synthetic Data Generation
Sythetic data are generated to replace partial real world data
Input
GDM-based Synthetic Image Generator
Output
5. Task Result Evaluation
The SP makes quality evaluation after receiving mixed images
SP
Obtained image dataset from MSTs
Results Evaluation
Type
Amount
Accuracy
GDM
Reverse Diffusion (Denoising)
Random Noise
X_T
X_{T-1}
X_t
...
X_0
Generated Data
Forward Diffusion (Add noise)
t=500
t=400
t=300
t=200
t=100
t=0

Fig. 10. A GDM-enabled SPPMCS framework for sensing data content protection, where GDM is adopted to synthesize vehicle images. The GDM process extracts features from original images by adding noise during the forward diffusion stage. A subsequent denoising process in the reverse diffusion stage then generates the target synthetic image. By replacing real-world images with these privacy-free synthetic counterparts, the framework effectively alleviates sensing data privacy concerns [161].

- *Step 5:* Hybrid data submissions from MSTs undergo SP-led quality attestation—measuring critical dimensions, which directly determines reward distribution via mechanism-defined payment rules anchored in quality-effectiveness proof.

D. Numerical Results

Since GDM can generate new synthetic data to replace the original data to be analyzed and processed, it can lower data attacks and privacy leakage risks for the original data. Simulations are based on the YOLOv3 model. The GDM parameters are configured with 500 iterations, a mini-batch size of 32, a learning rate of 0.005, and a squared gradient decay factor of 0.9999.

Fig. 11 illustrates the performance of YOLOv3 when trained with varying proportions of GDM-generated synthetic data. Increased integration of synthetic data correlates with a reduction in YOLOv3 detection accuracy. While synthetic data effectively mitigates privacy leakage from raw data, its inclusion inevitably compromises downstream task performance. To address this issue, Fig. 12 introduces the privacy protection utility index (PPUI), quantifying the balance between privacy preservation and data utility. PPUI is calculated as a weighted sum of the synthetic data proportion in the training dataset and the average accuracy of the downstream task model, both normalized to the range [0, 1]. The index demonstrates an inverse relationship with synthetic data proportion: higher proportions enhance privacy protection but degrade model accuracy, thereby reducing PPUI. The strategy maximizing PPUI achieves an optimal privacy-accuracy equilibrium.

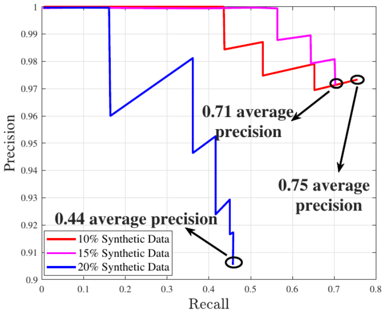

Fig. 11. Precision recall performance for YOLOv3 detection model with different percentage of synthetic data for training dataset [161].

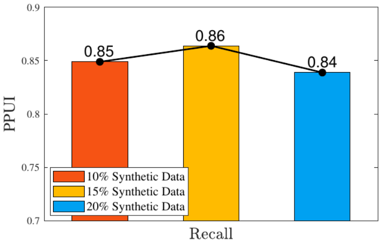

Fig. 12. Privacy-preserving utility index performance with different percentage of synthetic data for training dataset [161].

E. Lessons Learned

GDM-driven privacy-preserving mechanisms that generate high-fidelity synthetic data for high-altitude platform remote sensing data, UAV swarm cooperative path data, and SAGIN resource scheduling data. It generates statistical attributes that are consistent with, but free of, privacy properties. It can effectively deal with the privacy preservation challenges encountered in encryption [138], differential privacy [102], and FL-based traditional AI [156] [157].

VII. Tutorial 3: GAI for Multi-Modality Semantic-Aware Communications in SAGINs

A. Motivation

Real-time secure transmission and defense requirements face great challenges in dynamic and vulnerable SAGIN environments. For example, high-resolution image transmission delays cannot meet millisecond security decision-making needs [68]. The static models employed in traditional AI require periodic retraining to adapt to dynamic threats in SAGINs, resulting in delayed responses to emerging risks [99] [100]. In addition, data generation and reconstruction based on a single textual prompt can deviate significantly from real-world scenarios (e.g., CNN-based image recognition) [94].

The GAI approaches enhanced by multi-modal data generation and semantic compression can effectively break through the above limitations to achieve real-time adaptive security communication and defense [176] [209]. Semantic compression-based VAEs constribute to probabilistic compression of SAGIN data, saving SAGIN resource. In cross-modal alignment, the text-image semantic space can be unified, eliminating semantic ambiguities caused by dynamic environment changes. GDM balances generation quality and latency by adjusting the number of denoising steps, which significantly improves environment adaptability. A representative GAI-based multi-modality semantic-aware framework is proposed in [68] to promote the service quality of GAI for SAGINs. In this case, text and image data are exploited to create multimodal content to provide more reliable and secure communications for vehicle networks, mainly including GAI-assisted skeleton-semantic co-generation and DRL-assisted co-optimization.

B. System Description

We take the vehicle-to-vehicle (V2V) networks as an example to demonstrate the advantages of the GAI-based multi-modality semantic-aware framework. The Fig. 1 in [68] illustrates the GAI-enabled V2V networks, where multi-camera systems are employed to capture real-time road imagery for safety applications, such as collision alerts during accidents. Other major applications include navigation and route optimization, insurance and risk assessment, traffic simulation and prediction, and driving data generation. However, conventional generative models relying solely on textual descriptors exhibit high uncertainty, often generating inaccurate visual reconstructions (e.g., misrepresenting accident scenes or contextual details like license plates). Therefore, while the application of GAI in V2V shows great potential, it also faces critical challenges, including real-time data processing and decision making and adapting to dynamic and unpredictable environments. These unreliability challenges necessitate a robust generative architecture capable of preserving semantic fidelity to ensure safety-critical decision-making in V2V networks.

C. GAI-Based Solutions

We conduct a multimodal semantic-aware GAI framework for V2V networks, leveraging complementary data modalities (text and images) to enhance situational awareness. Multimodal fusion mitigates ambiguities inherent in unimodal data, yielding superior environmental perception. As depicted in Fig. 13, semantic skeletons derived from textual and visual inputs enable generative models to produce highly accurate reconstructions of road scenes. In addition, By fusing textual semantics and structural skeletons, this framework establishes tamper-evident environmental fingerprints, ensuring trustworthy perception against malicious attacks in [211] and [213].

- *Step 1: Semantic information extraction:* Textual semantics are extracted from road images via a secure encoder, detecting objects (vehicles, pedestrians) and attributes (position, trajectory).
- *Step 2: Image skeleton Extraction:* A streamlined structural skeleton is derived by identifying salient edges and contours within the road scene. This compact representation preserves critical topological features, enabling efficient downstream processing (e.g., object recognition, lane detection) while discarding superfluous details.
- *Step 3: Wireless transmission:* Semantic text and structural skeletons are fused into a lightweight data package for V2V broadcast. This approach reduces bandwidth

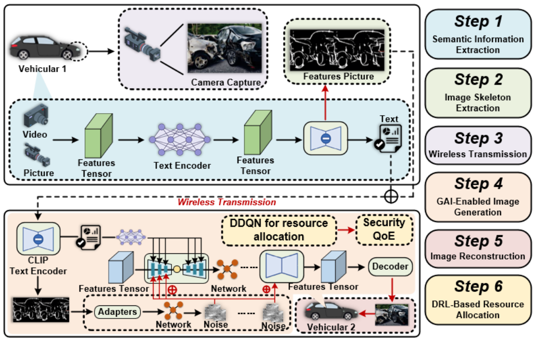

Fig. 13. Proposed GAI-enabled multi-modality semantic-aware communication framework for V2V networks [68].

consumption by $>90\%$ compared to raw image transmission, ensuring reliable delivery of essential environmental data under constrained network resources.

- *Step 4: GAI-enabled image generation:* A generative model synthesizes road scenes by integrating the structural skeleton (spatial foundation) with semantic descriptors (contextual attributes). This dual-input approach ensures high-fidelity reconstructions across diverse scenarios (e.g., weather variants, traffic densities). Model efficiency is tunable via diffusion-step optimization. In addition, this enhanced GAI can be utilized to generate adversarial scenarios to proactively identify hazards [213] and contribute the traffic real-time generation in [124]. Cross-modal validation detects synthesized deepfakes targeting SAGIN perception.
- *Step 5: Image reconstruction:* Onboard intelligence compares GAI-reconstructed scenes with real-time feeds. Potential hazards (e.g., obstacles, accidents) are identified through real-time scene comparison, triggering multimodal (auditory/visual) alerts to prompt driver intervention. The proposed framework can alleviate the problems of insufficient accuracy and real-time performance in [177], [179], and [182].
- *Step 6: DRL-based resource allocation:* For the proposed GAI framework, we propose a double deep Q-network (DDQN)-based approach to optimize the security quality of experience (QoE) in V2V networks within the constraints of the transmission power budget and the probability of successful transmission for each vehicle. More specific process can be found in [68].

D. Numerical Results

We evaluate the performance of the proposed GAI-DDQN algorithm for multi-modality semantic-aware framework. In Fig. 4 (a) of [68], the proposed DDQN strategy shows superior convergence dynamics versus benchmarks (e.g., DQN-based, greedy-based, and random-based GAI approaches). In addition, as shown in Fig. 14, we analyze the variation of QoE with the size of image payload. The QoE metric, evaluated per timestep in unconstrained environments, quantifies generative image fidelity critical for security-sensitive perception. The DDQN-based approach consistently outperforms benchmarks, demonstrating its efficacy in adversarial SAGIN conditions. Crucially, increased image payload enhances system QoE by conveying richer environmental signatures, enabling more robust anomaly detection against spoofing or data tampering attacks in [189], [214], and [217]. This payload-QoE synergy directly strengthens spatial-temporal integrity verification across ground-air-space links.

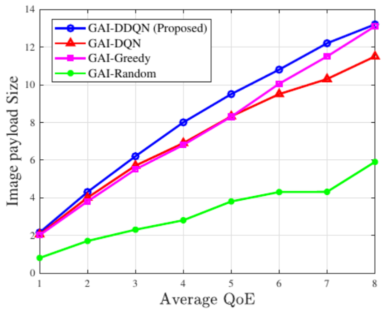

Fig. 14. Average QoE of different methods versus image payload size [68].

E. Lessons Learned

The transmission delays impeding real-time security decisions in SAGINs, and traditional AI fails to address due to their latency and single-modality inaccuracies. The proposed GAI-based multi-modality semantic-aware framework overcomes these limitations through cross-modal fusion and semantic compression, enabling tamper-evident environmental fingerprints and adaptive diffusion-step optimization for real-time threat response. It can effectively enhance cross-domain spoofing detection in [115], data tampering resilience in [189], and integrity verification of SAIGN links in [187]. Ultimately, the co-design of generative semantics and DRL resource allocation establishes a new paradigm for security-assured dynamic perception in safety-critical SAGIN operations.

VIII. Open Issues and Future Research Directions

This section explores the open issues and potential research directions for the application of GAI in SAGIN security.

A. Resource Constraints and Lightweight Deployment Issues

The high computational overhead inherent in current GAI models is in fundamental conflict with the limited resources (i.e., computing power and energy) of the SAGIN nodes. Existing compression techniques (e.g., knowledge distillation) often struggle to balance real-time security reasoning and threat detection accuracy with dynamic topologies. Advancing theoretical efficiency boundaries for edge-deployed models remains imperative.

1) Neural Architecture Search (NAS)-Driven GAI Model Compression: NAS leverages ML methods to automate the neural network architecture design in a data-driven manner [223], eliminating extensive manual effort required to explore vast network structure spaces for optimal task-specific efficiency. Within SAGINs, NAS can automatically generate lightweight GAI models with high compression ratio and low latency (e.g., adversarial sample generator and anomalous traffic synthesiser) to run real-time threat detection directly on satellite/IoT devices, reducing cloud dependency. Novel NAS frameworks should automate compression of GAI models for SAGIN-edge devices, optimizing layer pruning and quantization while preserving anomaly detection accuracy.

2) MoE-Driven Lightweight GAI Model: The MoE-driven GAI addresses SAGIN deployment challenges through modular task-specialization [84]. Its dynamic routing mechanism activates only relevant lightweight expert modules (e.g., compact models for generating defensive data), significantly reducing computation load on satellites. The modular design facilitates cross-platform adaptation and independent updates, enabling collaborative threat analysis without raw data transfer. This architecture delivers complex security capabilities to resource-constrained nodes through elastic deployment.

3) Edge-Cloud Collaborative Generative Frameworks: These frameworks mitigate resource constraints through layered task offloading and semantic communication [19]. Edge devices execute lightweight submodules (e.g., compressed diffusion models) for localized tasks, while dynamically offloading compute-intensive steps (e.g., multi-step denoising) to the cloud. Transmitting latent features instead of raw data significantly reduces bandwidth consumption. Federated knowledge distillation facilitates continuous compression of cloud-trained large models into micro-experts deployed at the edge, supporting offline generation of security decoys.

B. Adversarial Robustness and Trustworthy Mechanisms Issues

Current GAI-driven SAGIN security defenses are vulnerable to dynamically evolving adversarial attacks. Static defence models cannot counter adaptive attackers such as "evolutionary eavesdroppers", which continuously analyze defenses and adjust attack strategies in real-time, utilizing the static nature of defenses and the flaws of black-box decision-making.

1) Explainable Generative Frameworks: Explainable generative frameworks are expected to address the opaque decision-making of black-box models in SAGIN security [88]. These frameworks employ multi-modal explanation techniques to trace adversarial samples through cross-domain feature attribution. In addition, domain knowledge constraints, such as communication protocols and physical-layer features, are embedded into generative models to restrict output spaces and suppress adversarial perturbations. Robustness verification tools mathematically certify explanation integrity against attacks, ensuring auditability for critical missions. Integrated into detection pipelines, they provide interpretable alerts through XAI interfaces.

2) Digital Twin-GAI Fusion for Proactive Defense: Digital twin (DT)-GAI synergy establishes a dynamic cyber-physical testing environment for SAGIN resilience [224]. Using GANs, the DT synthesizes high-fidelity attack scenarios by leveraging historical attack data and threat intelligence. The DT to iteratively optimizes defense policies through millions of simulated attack cycles, employing techniques like OpenAI's "red teaming" to stress-test robustness. Crucially, concept drift adaptation mechanisms inspired by "digital cousins" that generalize physical environments—enable extrapolation to unseen attack variants, enhancing cross-domain generalization.

3) Continual Learning-Driven GAI for Sustainable Adaptive Defense: This framework mitigates catastrophic forgetting in SAGINs by synergizing adversarial robustness with trustworthy mechanisms through continual learning [225]. Federated continual learning enables edge devices to locally train adversarial detectors on emerging cross-domain threats. Stable knowledge retention leverages Hessian-based regularization to freeze critical weights from prior tasks, mathematically certifying that new adaptations do not degrade defenses against known threats, thus ensuring robustness continuity. This end-to-end approach bridges adaptive defense and trustworthy operation for SAGINs' dynamic threat landscape.

C. Cross-Domain Coordination and Governance Compliance Issues

In SAGINs, GAI-enabled cross-domain defense faces "governance-efficiency paradox": multi-source heterogeneous data integration challenges and governance-compliance barriers. Cross-domain attacks (e.g., "LEO relay with UAV poisoning") require fused information, but inter-domain data sovereignty conflicts and regulatory fragmentation create "governance silos" among satellite/aerial/ground nodes, hindering security policy synchronization.

1) Semantics-Driven Policy Generation: Multilevel generative models establish an end-to-end semantics-to-policy mapping framework to resolve defense policy fragmentation challenges in cross-domain SAGIN [76]. LLMs such as GPT-4, parse threat descriptions and extract semantic feature vectors, and GANs map semantic vectors to specific defense rules. In addition, a predefined security policy ontology library can also be constructed to map heterogeneous protocol semantics—including satellite authentication rules, airborne encryption policies, and ground access control policies—into a unified framework. Through FL, GAI aggregates localized policy features from each domain and synthesizes cross-domain coordinated policies based on semantic similarity. This approach unifies heterogeneous protocol semantics, enabling panoramic attack chain cognition.

2) Neuro-Symbolic Quantum-Driven GAI Architectures: This direction fuses the dynamic learning capability of neural networks with the interpretable rule engine of symbolic systems , combining quantum computing to accelerate the cross-domain verification process and crack the governance compliance and real-time contradiction [226]. It can be achieved through three key capabilities: (1) A neuro-symbolic layer encoding governance rules (e.g., spectrum allocation) for interpretable compliance while learning multi-domain threats; (2) A GAI module simulates cross-domain conflicts (such as SAGIN resource contention) to draft real-time protocols and optimize scheduling; (3) A quantum engine accelerates multi-domain policy verification, slashing audit time, and concurrently secures communications via quantum encryption. This integration enables dynamic cross-domain policy alignment and real-time collaborative governance.

IX. Conclusion

This review presents an overview of GAI-enabled secure communications for SAGINs, emphasizing the enhanced effectiveness of GAI compared to traditional AI in protecting SAGINs. This analysis thoroughly examines the architecture of SAGINs and the specific security challenges faced. The efficacy of different GAI models in addressing security issues has been examined. This survey provides a comprehensive analysis of GAI-based methods addressing authenticity failures, confidentiality breaches, integrity tampering, and availability disruptions in SAGIN communications. The three tutorials conducted provide a detailed exploration, highlighting the superior efficiency of GAI in addressing security threats in SAGINs relative to traditional AI. We have identified open issues and associated research directions, offering insights into the future of GAI-enabled SAGIN security.

References

[1] H. Gao, R. Cao, W. Xu, C. Yuan, and H. -H. Chen, "Space-Air-Ground Integrated Networks with Task-Driven Connected Intelligence," *IEEE Wireless Commun.*, vol. 32, no. 2, pp. 254-261, Apr. 2025.

[2] O. Kodheli et al., "Satellite Communications in the New Space Era: A Survey and Future Challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 70-109, 1st Quart., 2021.

[3] H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, "A Survey on Space-Air-Ground-Sea Integrated Network Security in 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 53–87, 1st Quart., 2021.

[4] P. Ray, "A review on 6G for space-air-ground integrated network: Key enablers, open challenges, and future direction," *J. King Saud. Univ-com.*, vol. 34, no. 9, pp. 6949–6976, 2022.

[5] J. Liu, L. Bai, C. Jiang, and W. Zhang, *Space-Air-Ground Integrated Network Security*. Springer, Berlin, Germany, 2023.

[6] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-Air-Ground Integrated Network: A Survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, 4th Quart., 2018.

[7] M. Xu et al., "Quantum-Secured Space-Air-Ground Integrated Networks: Concept, Framework, and Case Study," *IEEE Wireless Commun.*, vol. 30, no. 6, pp. 136-143, Dec. 2023.

[8] N. Kato, Z. M. Fadlullah, F. Tang, B. Mao, S. Tani, A. Okamura, and J. Liu, "Optimizing Space-Air-Ground Integrated Networks by Artificial Intelligence," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 140–147, 2019.

[9] S. Mahboob and L. Liu, "Revolutionizing future connectivity: A contemporary survey on AI-empowered satellite-based non-terrestrial networks in 6G," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 1279-1321, 2nd Quart., 2024.

[10] G. Fontanesi et al., "Artificial intelligence for satellite communication: A survey," *IEEE Commun. Surveys Tuts.*, 2025.

[11] R. Zhang et al., "Generative AI for space-air-ground integrated networks," *IEEE Wireless Commun.*, vol. 31, no. 6, pp. 10-20, Dec. 2024.

[12] R. Zhang et al., "Embodied AI-enhanced vehicular networks: An integrated vision language models and reinforcement learning method," *IEEE Trans. Mobile Comput.*, 2025.

[13] F. Fourati and M.-S. Alouini, "Artificial intelligence for satellite communication: A review," *Intell. converg. netw.*, vol. 2, no. 3, pp. 213–243, 2021.

[14] S. Hashima, A. Gendia, K. Hatano, O. Muta, M. S. Nada, and E. M. Mohamed, "Next-Gen uav-satellite communications: AI innovations and future prospects," *IEEE Open J. Veh. Technol.*, 2025.

[15] N. Van Huynh et al., "Generative AI for physical layer communications: A survey," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 3, pp. 706-728, Jun. 2024.

[16] L. Cai et al., "Secure physical layer communications for low-altitude economy networking: A survey," 2025, *arXiv:2504.09153v1*.

[17] C. Zhao et al., "Generative AI for secure physical layer communications: A survey," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 3-26, Feb. 2025.

[18] T. -H. Vu, S. Kumar Jagatheesaperumal, M. -D. Nguyen, N. Van Huynh, S. Kim, and Q. -V. Pham, "Applications of generative AI (GAI) for mobile and wireless networking: A survey," *IEEE Internet Things J.*, vol. 12, no. 2, pp. 1266-1290, 15 Jan. 2025.

[19] M. Xu et al., "Unleashing the power of edge-cloud generative AI in mobile networks: A survey of AIGC services," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 1127-1170, 2nd Quart., 2024.

[20] H. Zhang et al., "The role of generative artificial intelligence in internet of electric vehicles," *IEEE Internet Things J.*, vol. 12, no. 6, pp. 6208-6232, Mar. 2025.

[21] H. Huang, P. Wang, J. Pei, J. Wang, S. Alexanian, and D. Niyato, "Deep learning advancements in anomaly detection: A comprehensive survey," 2025, *arXiv:2503.13195v1*.

[22] A. Golda et al., "Privacy and security concerns in generative AI: A comprehensive survey," *IEEE Access*, vol. 12, pp. 48126-48144, 2024.

[23] T. Wang et al., "Security and privacy on generative data in AIGC: A survey," *ACM Comput. Surv.*, vol. 57, no. 4, pp. 1-34, 2024.

[24] O. Friha, M. Amine Ferrag, B. Kantarci, B. Cakmak, A. Ozgun, and N. Ghoualmi-Zine, "LLM-Based edge intelligence: A comprehensive survey on architectures, applications, security and trustworthiness," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 5799-5856, 2024.

[25] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727-1765, Sept. 2016.

[26] M. Shakiba-Herfeh, A. Chorti, and H. V. Poor, "Physical layer security: Authentication, integrity, and confidentiality," *Physical Layer Security*, Cham, Switzerland: Springer, 2021, pp. 129–150.

[27] S. Yao, J. Guan, Y. Wu, K. Xu, and M. Xu, "Toward secure and lightweight access authentication in sagins," *IEEE Wireless Commun.*, vol. 27, no. 6, pp. 75-81, Dec. 2020.

[28] M. Sherman, S. Shao, X. Sun, and J. Zheng, "Counter uav swarms: Challenges, considerations, and future directions in uav warfare," *IEEE Wireless Commun.*, vol. 32, no. 1, pp. 190-196, Feb. 2025.

[29] Y. Gao, T. Zhou, W. Zheng, H. Yang, and T. Zhang, "High-Availability authentication and key agreement for internet of things-based devices in industry 5.0," *IEEE Trans. Ind. Inform.*, vol. 20, no. 12, pp. 13571-13579, Dec. 2024.

[30] H. Zhao, K. Ren, T. Yue, C. Zhang, and S. Yuan, "TransFG: A cross-view geo-localization of satellite and uavs imagery pipeline using transformer-based feature aggregation and gradient guidance," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1-12, Jan. 2024.

[31] Bariah Lina et al., "RIS-Assisted space-air-ground integrated networks: New horizons for flexible access and connectivity," *IEEE Netw.*, vol. 37, no. 3, pp. 118-125, Jun. 2023.

[32] P. Zhang, C. Wang, N. Kumar, and L. Liu, "Space-Air-Ground integrated multi-Domain network resource orchestration based on virtual network architecture: A DRL method," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2798-2808, Mar. 2022.

[33] D. Wang, T. He, Y. Lou, L. Pang, Y. He, and H. -H. Chen, "Double-Edge computation dffloading for secure integrated space–air–aqua networks," *IEEE Internet Things J.*, vol. 10, no. 17, pp. 15581-15593, Sept. 2023.

[34] J. Zhou, M. Hu, C. Ma, Z. Liu, W. Jian, and C. Zhou, "Identification of guidance model for GNSS spoofing in rotary-wing uavs," in *Proc. 11th Int. Forum Elect. Eng. Autom. (IFEEA)*, Shenzhen, China, 2024, pp. 1152-1158.

[35] G. Rigoni, N. Scremin, and M. Conti, "Towards a self-rescuing system for uavs under GNSS attack," in *Proc. 20th Int. Conf. Wirel. and Mobile Comput., Netw. and Commun. (WiMob)*, Paris, France, 2024, pp. 339-346.

[36] L. Huang, P. Liu, X. Chen, C. Jiang, L. Kuang, and J. Lu, "A consolidated game framework for cooperative defense against cross-domain cyber attacks in satellite-enabled internet of things," *IEEE Internet Things J.*, 2025.

[37] M. Chen, K. Zhang, D. Zhang, L. Yuan, and Y. Zhai, "Design of Network Security Data Exchange Management System Based on Blockchain," in *Proc. Int. Conf. Telecommun. Electron. Inform. (ICTEI)*, Lisbon, Portugal, 2023, pp. 506-510.

[38] M. Li, H. Cui, C. Shan, X. Du, and M. Guizani, "IH-SESD: Modeling information hiding with super resolution enhancement and significant region detection for uav networks," *IEEE Internet Things J.*, 2024.

[39] Z. Wang, R. Liu, Q. Liu, L. Han, and J. S. Thompson, "Feasibility study of uav-assisted anti-jamming positioning," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7718-7733, Aug. 2021.

[40] S. Salim, N. Moustafa, and M. Reisslein, "Cybersecurity of satellite communications systems: A comprehensive survey of the space, ground, and links segments," *IEEE Commun. Surveys Tuts.*, vol. 27, no. 1, pp. 372-425, Feb. 2025.

[41] D. She, W. Wang, Z. Yin, J. Wang, and H. Shan, "GPS spoofing attack recognition for uavs with limited samples," *IEEE Internet Things J.*, vol. 12, no. 1, pp. 250-261, Jan. 2025.

[42] B. Li, Z. Fei, C. Zhou, and Y. Zhang, "Physical-Layer security in space information networks: A survey," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 33-52, Jan. 2020.

[43] Y. Gu, Y. Ma, X. Wang, and A. Shah, "Security energy efficiency optimization and analysis of aerial-irs-assisted uav-mec system," *IEEE Access*, vol. 12, pp. 118953-118967, Aug. 2024.

[44] V. S. R. Kantheti, C. -H. Lin, S. -C. Lin, and L. C. Chu, "Anti-Jamming resilient leo satellite swarms," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Boston, MA, USA, 2023, pp. 77-82.

[45] A. Talgat, R. Wang, M. A. Kishk, and M. -S. Alouini, "Enhancing physical-layer security in leo satellite-enabled iot network communications," *IEEE Internet Things J.*, vol. 11, no. 20, pp. 33967-33979, Oct. 2024.

[46] J. M. Willis, R. F. Mills, L. O. Mailloux, and S. R. Graham, "Considerations for secure and resilient satellite architectures," in *Proc. Int. Conf. Cyber Conflict U.S. (CyCon U.S.)*, Washington, DC, USA, 2017, pp. 16-22.

[47] O. Ceviz, S. Sen, and P. Sadioglu, "A survey of security in uavs and fanets: Issues, threats, analysis of attacks, and solutions," *IEEE Commun. Surveys Tuts.*, 2024.

[48] J. Jung, S. Ahn, S. Kwon, S. -I. Park, and J. Kang, "Optimal aav 3D trajectory design and resource allocation for secure mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 74, no. 3, pp. 5281-5286, Mar. 2025.

[49] L. Gupta, R. Jain, and G. Vaszkun, "Survey of Important Issues in UAV Communication Networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123-1152, 2nd Quart. 2016.

[50] F. Du, J. Ge, W. Wang, Y. Zou, S. -Y. Chang, and W. Fan, "Exploiting the vulnerabilities in mavlink protocol for uav hijacking," in *Proc. 17th Int. Conf. Secur. Inf. Netw. (SIN)*, Sydney, Australia, 2024, pp. 1-8.

[51] D. Lin and W. Wu, "Optimization of a secure uav-based iot: RF-Fingerprint authentication and resource allocation," *IEEE Internet Things J.*, vol. 10, no. 21, pp. 19208-19217, Nov. 2023.

[52] Y. Zhu, G. Zheng, and M. Fitch, "Secrecy rate analysis of uav-enabled mmwave networks using matérn hardcore point processes," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, pp. 1397-1409, Jul. 2018.

[53] Z. Tong, J. Wang, X. Hou, C. Jiang, and J. Liu, "UAV-Assisted covert federated learning over mmWave massive mIMO," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11785-11798, Sept. 2024.

[54] Z. Zhang, Y. Huang, W. Huang, W. Pan, Y. Liao, and S. Zhou, "An auto-upgradable end-to-end preauthenticated secure communication protocol for uav-aided perception intelligent syste," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 30187-30203, Sept. 2024.

[55] S. Dahmane, M. B. Yagoubi, C. A. Kerrache, P. Lorenz, N. Lagraa, and A. Lakas, "Toward a secure edge-enabled and artificially intelligent internet of flying things using blockchain," *IEEE Internet Things J. Mag.*, vol. 5, no. 2, pp. 90-95, Jun. 2022.

[56] F. M. Awaysheh, M. Alazab, S. Garg, D. Niyato, and C. Verikoukis, "Big data resource management & networks: Taxonomy, survey, and future directions," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 4, pp. 2098-2130, 4th Quart., 2021.

[57] K. Kousias et al., "A large-scale dataset of 4g, nb-iot, and 5g non-standalone network measurements," *IEEE Commun. Mag.*, vol. 62, no. 5, pp. 44-49, May 2024.

[58] S. Achraou, D. El Khamlichi, A. Zakriti, M. El Bakkali, and S. Ben Haddi, "Innovative design of a compact diplexer using heterogeneous filters for 5g and wimax," in *Proc. Int. Microw. Antemm. Symp. (IMAS)*, Marrakech, Morocco, 2024, pp. 1-4.

[59] V. Voloshyn, M. S. Khan, G. Srivastava, and D. M, "Analysis of nist lightweight cryptographic algorithms performance in iot security environments based on MQTT," in *Proc. IEEE Int. Conf. Wirel. Commun. New. (WCNC)*, Dubai, United Arab Emirates, 2024, pp. 1-6.

[60] C. -S. Park, "Security architecture for secure multicast coap applications," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3441-3452, Apr. 2020.

[61] O. A. Saad, A. A. Etman, M. A. Abdel-Malek, and M. Azab, "A proactive crowdsourcing framework for fake base station detection and avoidance," in *Proc. IEEE Annu. Comput. Commun. Workshop Conf. (CCWC)*, Las Vegas, NV, USA, 2025, pp. 00814-00820.

[62] C. Wang, Z. Yuan, P. Zhou, Z. Xu, R. Li, and D. O. Wu, "The security and privacy of mobile-edge computing: An artificial intelligence perspective," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 22008-22032, Dec. 2023.

[63] R. G, V. P, and A. S, "Evading machine-learning-based android malware detector for iot devices," *IEEE Syst. J.*, vol. 17, no. 2, pp. 2745-2755, Jun. 2023.

[64] K. -I. Kitayama et al., "Security in photonic networks: threats and security enhancement," *J. Lightwave Technol.*, vol. 29, no. 21, pp. 3210-3222, Nov. 2011.

[65] J. Wen et al., "Generative AI for low-carbon artificial intelligence of things with large language models," *IEEE Internet Things J. Mag.*, vol. 8, no. 1, pp. 82-91, Jan. 2025.

[66] M. Nemati, J. Park, and J. Choi, "VQ-VAE Empowered Wireless Communication for Joint Source-Channel Coding and Beyond," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, 2023, pp. 3155-3160.

[67] Z. Li et al., "Situation-Aware multivariate time series anomaly detection through active learning and contrast vae-based models in large distributed systems," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2746-2765, Sept. 2022.

[68] R. Zhang et al., "Generative AI-Enabled vehicular networks: Fundamentals, framework, and case study," *IEEE Netw.*, vol. 38, no. 4, pp. 259-267, Jul. 2024.

[69] Y. Xiao et al., "Distributed traffic synthesis and classification in edge networks: A dederated self-supervised learning approach," *IEEE Trans. Mobile Comput.*, vol. 23, no. 2, pp. 1815-1829, Feb. 2024.

[70] S. Luo and F. Huang, "MaGAT: Mask-Guided adversarial training for defending face editing gan models from proactive defense," *IEEE Signal Process. Lett.*, vol. 31, pp. 969-973, 2024.

[71] H. Cao et al., "A survey on generative diffusion models," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 7, pp. 2814-2830, Jul. 2024.

[72] F. -A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 10850-10869, Sept. 2023.

[73] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 6840–6851.

[74] H. Du et al., "Exploring collaborative distributed diffusion-based aigc in wireless networks," *IEEE Netw.*, vol. 38, no. 3, 2024.

[75] J. He et al., "Securing federated diffusion model with dynamic quantization for generative ai services in multiple-access artificial intelligence of things," *IEEE Internet Things J.*, vol. 11, no. 17, pp. 28064-28077, Sept. 2024.

[76] J. Liu et al., "Optimizing resource allocation for multi-modal semantic communication in mobile aigc networks: A diffusion-based game approach," *IEEE Trans. Cogn. Commun. Netw.*, 2025.

[77] P. Dai, K. Yue, R. Jin, T. M. Wu, and K. Xiong, "Enhancing approximate message passing via diffusion models towards on-device intelligence," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Denver, CO, USA, 2024, pp. 890-895.

[78] H. Du et al., "Enhancing deep reinforcement learning: A tutorial on generative diffusion models in network optimization," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 4, pp. 2611-2646, 4th Quart., 2024.

[79] X. Li et al., "Transformer-Based visual segmentation: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 12, pp. 10138-10163, Dec. 2024.

[80] Z. Wu and B. Wang, "Transformer-Based autoencoder framework for nonlinear hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1-15, 2024.

[81] W. Xie, J. Zou, J. Xiao, M. Li, and X. Peng, "Quan-Transformer based channel feedback for ris-aided wireless communication Systems," *IEEE Commun. Lett.*, vol. 26, no. 11, pp. 2631-2635, Nov. 2022.

[82] U. C. Akuthota and L. Bhargava, "Transformer-Based intrusion detection for iot networks," *IEEE Internet Things J.*, vol. 12, no. 5, pp. 6062-6067, Mar. 2025.

[83] R. Zhang et al., "Interactive AI with retrieval-augmented generation for next generation networking," *IEEE Netw.*, vol. 38, no. 6, pp. 414-424, Nov. 2024.

[84] R. Zhang *et al.* "Generative AI agents with large language model for satellite networks via a mixture of experts transmission," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 12, pp. 3581-3596, Dec. 2024.

[85] M. Xu et al., "When large language model agents meet 6g networks: Perception, grounding, and alignment," *IEEE Wireless Commun.*, vol. 31, no. 6, pp. 63-71, Dec. 2024.

[86] R. Zhang et al., "Toward democratized generative AI in next-generation mobile edge networks," *IEEE Netw.*, 2025.

[87] Y. Qu, M. Ding, N. Sun, K. Thilakarathna, T. Zhu, and D. Niyato, "The frontier of data erasure: A survey on machine unlearning for large language models," *Computer*, vol. 58, no. 1, pp. 45-57, Jan. 2025.

[88] T. Senevirathna, V. H. La, S. Marchal, B. Siniarski, M. Liyanage, and S. Wang, "A Ssurvey on XAI for 5g and beyond security: Technical aspects, challenges and research directions," *IEEE Commun. Surveys Tuts.*, 2025.

[89] O. Kuznetsov, P. Sernani, L. Romeo, E. Frontoni, and A. Mancini, "On the integration of artificial intelligence and blockchain technology: A perspective about security," *IEEE Access*, vol. 12, pp. 3881-3897, 2024.

[90] H. Du et al., "The age of generative AI and AI-generated everything," *IEEE Netw.*, vol. 38, no. 6, pp. 501-512, Nov. 2024.

[91] Y. Lian et al., "Deep reinforcement learning-based moving target defense for multicast in software-defined satellite networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Denver, CO, USA, 2024, pp. 4786-4791.

[92] K. Rathor, K. V, K. Sunanda, R. K, A. Shobana, and M. Anusuya, "Enhancing network security against apts through svm-based network traffic analysis: Identifying anomalies in communication flows," in *Proc. Int. Conf. Comput. Data Sci. (ICCDS)*, Chennai, India, 2024, pp. 1-6.

[93] C. Hutabarat and Y. Asnar, "Development of machine learning module in intrusion detection system for unknown threat detection," in *Proc. IEEE Int. Conf. Data Softw. Eng. (ICoDSE)*, Gorontalo, Indonesia, 2024, pp. 114-119.

[94] J. Luo, "Man-in-the-middle intrusion detection based on CNN-LSTM model," in *Proc. IEEE Int. Conf. Electron. Technol., Commun. Inf. (ICETCI)*, Changchun, China, 2023, pp. 1-7.

[95] G. Wu, Y. Zhang, L. Deng, J. Zhang, and T. Chai, "Cross-Modal learning for anomaly detection in complex industrial process: Methodology and benchmark," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 35, no. 3, pp. 2632-2645, Mar. 2025.

[96] N. Amjad, H. Afzal, M. F. Amjad, and F. A. Khan, "A multi-classifier framework for open source malware forensics," in *Proc. IEEE 27th Int. Conf. Enabling Technol., Infrastruct. Collaborative Enterprises (WET-ICE)*, Jun. 2018, pp. 106–111.

[97] Y. Ahmed, A. T. Asyhari, and M. A. Rahman, "A cyber kill chain approach for detecting advanced persistent threats," *Comput., Mater. Continua*, vol. 67, no. 2, pp. 2497–2513, 2021.

[98] Y. Li, X. Peng, J. Zhang, Z. Li, and M. Wen, "DCT-GAN: Dilated convolutional transformer-based GAN for time series anomaly detection," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 3632-3644, Apr. 2023.

[99] R. Zhang, K. Xiong, Y. Lu, D. W. K. Ng, P. Fan, and K. B. Letaief, "SWIPT-Enabled cell-free massive mimo-noma networks: A machine learning-based approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 6701-6718, Jul. 2024.

[100] B. F. Balogun, K. Tripathi, S. Tiwari, J. S. S. Mohan, and A. K. Tyagi, "A blockchain-based deep learning approach for cyber security in next generation medical cyber-physical systems," *J. Auto. Intell.*, Mar. 2024.

[101] P. P. Kundu, T. Truong-Huu, L. Chen, L. Zhou, and S. G. Teo, "Detection and classification of botnet traffic using deep learning with model explanation," *IEEE Trans. Dependable Secure Comput.*, early access, Jun. 2022.

[102] M. Gupta, C. Akiri, K. Aryal, E. Parker, and L. Praharaj, "From ChatGPT to ThreatGPT: Impact of generative AI in cybersecurity and privacy," *IEEE Access*, vol. 11, pp. 80218–80245, 2023.

[103] Z. Wu, C. Liang, and Y. Zhang, "Blockchain-Based authentication of GNSS civil navigation message," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 4, pp. 4380-4392, Aug. 2023.

[104] Z. Wu, R. Liu, and H. Cao, "ECDSA-Based message authentication scheme for BeiDou-II navigation satellite system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 4, pp. 1666-1682, Aug. 2019.

[105] N. S. Joseph, C. Banerjee, E. Pasiliao, and T. Mukherjee, "FlightSense: A spoofer detection and aircraft identification system using raw ADS-B data," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Atlanta, GA, USA, 2020, pp. 3885-3894.

[106] J. Li, X. Zhu, M. Ouyang, W. Li, Z. Chen, and Q. Fu, "GNSS spoofing jamming detection based on generative adversarial network," *IEEE Sens. J.*, vol. 21, no. 20, pp. 22823–22832, Oct. 2021.

[107] J. Qin, Y. Xun, Z. Deng, and J. Liu, "GPIDS: GAN assisted contextual pattern-aware intrusion detection system for IVN," *IEEE Trans. Veh. Technol.*, vol. 73, no. 9, pp. 12682-12693, Sept. 2024.

[108] M. Smolin, "GenCoder: A generative AI-based adaptive intra-vehicle intrusion detection system," *IEEE Access*, vol. 12, pp. 150651-150663, 2024.

[109] A. Iqbal, M. N. Aman, and B. Sikdar, "A deep learning based induced GNSS spoof detection framework," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 2, pp. 457-478, 2024.

[110] H. El alami and D. B. Rawat, "DroneDefGANt: A generative AI-based approach for detecting uas attacks and faults," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Denver, CO, USA, 2024, pp. 1933-1938.

[111] D. Roy, T. Mukherjee, A. Riden, J. Paquet, E. Pasiliao, and E. Blasch, "GANSAT: A GAN and satellite constellation fingerprint-based framework for GPS spoof-detection and location estimation in GPS deprived environment," *IEEE Access*, vol. 10, pp. 45485–45507, 2022.

[112] D. Huang and A. Al-Hourani, "Physical layer spoof detection and authentication for iot devices using deep learning methods," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 2, pp. 841-854, 2024.

[113] Q. Jiang and J. Sha, "Physical layer authentication via conditional variational auto-encoder for ais," *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 2, pp. 513-521, Jun. 2025.

[114] J. Wang et al., "Generative AI based secure wireless sensing for ISAC networks," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 5195-5210, 2025.

[115] K. S. Germain and F. Kragh, "Mobile physical-layer authentication using channel state information and conditional recurrent neural networks," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Helsinki, Finland, 2021, pp. 1-6.

[116] R. Meng et al., "Physical-Layer authentication based on hierarchical variational autoencoder for industrial internet of things," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2528-2544, Feb. 2023.

[117] H. Alipour, Y. B. Al-Nashif, P. Satam, and S. Hariri, "Wireless anomaly detection based on IEEE 802.11 behavior analysis," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 10, pp. 2158-2170, Oct. 2015.

[118] E. Letsoalo and S. Ojo, "Survey of media access control address spoofing attacks detection and prevention techniques in wireless networks," in *Proc. IST-Africa Week Conf.*, Durban, South Africa, 2016, pp. 1-10.

[119] C Benzaïd et al., "Intelligent detection of MAC spoofing attack in 802.11 network," in *Proc. ACM Int. Conf.*, 2016.

[120] P Madani, N Vlajic, and I Maljevic, "Randomized moving target approach for mac-layer spoofing detection and prevention in IoT systems," *Digital Threats: Research and Practice*, vol. 3, no. 4, Dec. 2022.

[121] D. J. Tian, K. R. B. Butler, J. I. Choi, P. McDaniel, and P. Krishnaswamy, "Securing ARP/NDP from the ground up," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 9, pp. 2131-2143, Sept. 2017.

[122] W. Kim, S. Kim, and H. Lim, "Malicious data frame injection attack without seizing association in IEEE 802.11 wireless lANs," *IEEE Access*, vol. 9, pp. 16649-16660, 2021.

[123] M. Yue, H. Zheng, H. Cui, and Z. Wu, "GAN-LSTM-Based ADS-B attack detection in the context of air traffic control," *IEEE Internet Things J.*, vol. 10, no. 14, pp. 12651-12665, Jul. 2023.

[124] MG Constantin et al., "Exploring generative adversarial networks for augmenting network intrusion detection tasks," *ACM Trans. Multimedia Comput.*, vol. 21, no. 1, pp. 1-19, Jul. 2024.

[125] M. Nadeem and C. Hongsong. "Advancing social network security with magteon-turing L3TM: A multi-layered defense system against cyber threats," *Comput. Netw.*, vol. 267, pp. 1389-1286, 2025.

[126] H. Tanyıldız et al., "Detection of cyber attacks in electric vehicle charging systems using a remaining useful life generative adversarial network," *Sci. Rep.*, vol. 15, no. 1, 2025.

[127] M. S. Munir, S. Proddatoori, M. Muralidhara, W. Saad, Z. Han, and S. Shetty, "A zero trust framework for realization and defense against generative AI attacks in power grid," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Denver, CO, USA, 2024, pp. 2482-2488.

[128] Zabeehullah, N. A. Khan, I. U. Din, A. Almogren, A. Altameem, and M. Guizani, "Secure and efficient AI-SDN-based routing for healthcare-consumer internet of things," *IEEE Trans. Consum. Electron.*, 2025.

[129] Zabeehullah, Q. M. u. Haq, F. Arif, N. A. Khan, M. S. Anwar and W. Alhalabi, "A secure AI framework for intelligent traffic prediction and routing in SDN based consumer internet of things," *IEEE Trans. Consum. Electron.*, 2025.

[130] S. Bethu, "Malicious attack detection in IoT by generative adversarial networks," *SN Comput. Sci.*, vol. 6, no. 4, 2025.

[131] X. Jiang et al., "Covert communication in UAV-assisted air-ground networks ," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 190-197, Aug. 2021.

[132] J. Han, Y. Zhou, G. Liu, T. Liu, and X. Zeng, "A novel physical layer key generation method based on WGAN-GP adversarial autoencoder," in *Proc. 4th Int. Conf. Commun. Inf. Syst. Comput. Eng. (CISCE)*, Shenzhen, China, 2022, pp. 1-6.

[133] E. Mahalal, M. Ismail, Z. -Y. Wu, M. M. Fouda, and Z. Md Fadlullah, "GAN-Assisted secret key generation against eavesdropping in dynamic indoor LiFi networks," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Washington, DC, USA, 2024, pp. 1-5.

[134] R. Karmakar, G. Kaddoum, and O. Akhrif, "IntSHU: A security-enabled intelligent soft handover approach for UAV-aided 5G and beyond," *IEEE Trans. Cogn. Commun. Netw.*, 2025.

[135] I. Meraouche, S. Dutta, S. K. Mohanty, I. Agudo, and K. Sakurai, "Learning multi-party adversarial encryption and its application to secret sharing," *IEEE Access*, vol. 10, pp. 121329-121339, 2022.

[136] X. Wang, Z. Zheng, and Z. Fei, "ASAP: adversarial learning based secure autoprecoder design for MIMO wiretap channels," *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1915-1919, Sept. 2022.

[137] P. Jiang, C. -K. Wen, X. Li, S. Jin, and G. Y. Li, "Semantic satellite communications based on generative foundation model," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 7, pp. 2431-2445, Jul. 2025.

[138] X. Luo, Z. Chen, M. Tao, and F. Yang, "Encrypted semantic communication using adversarial training for privacy preserving," *IEEE Commun. Lett.*, vol. 27, no. 6, pp. 1486-1490, Jun. 2023.

[139] K. W. McClintick, J. Harer, B. Flowers, W. C. Headley, and A. M. Wyglinski, "Countering physical eavesdropper evasion with adversarial training," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1820-1833, 2022.

[140] J. Zhang, Z. Liu, X. Feng, H. Yang, and S. Liang, "Enhanced secure beamforming for irs-assisted iot communication using a generative-diffusion-model-enabled optimization approach," *IEEE Internet Things J.*, vol. 12, no. 10, pp. 13398-13414, May, 2025.

[141] R. Zhang, K. Xiong, Y. Lu, B. Gao, P. Fan, and K. B. Letaief, “Joint coordinated beamforming and power splitting ratio optimization in mu-miso swpt-enabled HetNets: A multi-agent DDQN-based approach,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 2, pp. 677-693, Feb. 2022.

[142] P. Qi, Y. Meng, S. Zheng, X. Zhou, N. Cheng, and Z. Li, “Adversarial defense embedded waveform design for reliable communication in the physical layer,” *IEEE Internet Things J.*, vol. 11, no. 10, pp. 18136-18153, May, 2024.

[143] C. Zhang et al., “Multi-Objective aerial collaborative secure communication optimization via generative diffusion model-enabled deep reinforcement learning,” *IEEE Trans. Mobile Comput.*, vol. 24, no. 4, pp. 3041-3058, Apr. 2025.

[144] Y. Guo et al., “Secrecy Energy Efficiency Maximization in IRS-Assisted VLC MISO Networks with RSMA: A DS-PPO Approach,” *IEEE Trans. Wireless Commun.*, 2025.

[145] C. Zhang et al., “Enhancing physical layer communication security through generative AI with mixture of experts ,” *IEEE Wireless Commun.*, vol. 32, no. 3, pp. 176-184, Jun. 2025.

[146] A. Kakati, G. Li, E. Moustapha Diallo, L. Chiru Kawala, N. Hussain, and A. B. M. Adam, “Toward proactive, secure and efficient space-air-ground communications: Generative AI-based DRL framework,” *IEEE Open J. Commun. Soc.*, vol. 6, pp. 1284-1298, 2025.

[147] Y. Feng, Y. Jiang, and Y. Wang, “GAN-Based covert communications against an adversary with uncertain detection threshold in federated learning networks,” in *Proc. Int. Conf. Netw. Netw. Appl. (NaNA)*, Qingdao, China, 2023, pp. 613-618.

[148] Z. Li, X. Liao, J. Shi, L. Li, and P. Xiao, “MD-GAN-Based uav trajectory and power optimization for cognitive covert communications,” *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10187-10199, Jun. 2022.

[149] Q. Hu, N. Huang, Y. Chen, and C. Gong, “Covert water-to-air optical wireless communication based on an adversarial autoencoder,” *IEEE Trans. Cogn. Commun. Netw.*, 2025.

[150] S. Jia, L. Xiaomeng, L. Xiaomin, T. Zhuangzhuang, and H. Junfan, “Covert leo satellite communication aided by generative adversarial network based cooperative UAV jamming,” *China Commun.*, vol. 21, no. 9, pp. 27-39, Sept. 2024.

[151] J. An, B. Kang, Q. Ouyang, J. Pan, and N. Ye, “Covert communications meet 6G NTN: A comprehensive enabler for safety-critical IoT,” *IEEE Netw.*, vol. 38, no. 4, pp. 17-24, Jul. 2024.

[152] M. V. Namitha, G. R. Manjula, and M. C. Belavagi, “StegAbb: A cover-generating text steganographic tool using GPT-3 language modeling for covert communication across SDRs,” *IEEE Access*, vol. 12, pp. 82057-82067, 2024.

[153] M. Grekov and A. Sychugov, “Distributed detection of anomalies in the network flow using generative adversarial networks,” in *Proc. Int. Russ. Autom. Conf. (RusAutoCon)*, Sochi, Russian Federation, 2022, pp. 332-336.

[154] T. Gaber, T. Ali, M. Nicho, and M. Torky, “Robust attacks detection model for internet of flying things based on generative adversarial network (GAN) and adversarial training,” *IEEE Internet Things J.*, vol. 12, no. 13, pp. 23961-23974, Jul. 2025.

[155] W. Yang and C. Zeng, “A hybrid anomaly detection model based on GANomaly in cloud environment,” in *Proc. IEEE 5th Int. Conf. Big Data Artif. Intell. (BDAI)*, Fuzhou, China, 2022, pp. 51-56.

[156] Y. Liu, J. Yin, W. Zhang, C. An, Y. Xia, and H. Zhang, “Integration of federated learning and AI-generated content: A survey of overview, opportunities, challenges, and solutions,” *IEEE Commun. Surveys Tuts.*, 2024.

[157] W. Jiang, H. Han, Y. Zhang, J. Mu, and A. Shankar, “Intrusion detection with federated learning and conditional generative adversarial network in satellite-terrestrial integrated networks,” *Mobile Netw Appl.*, 2024.

[158] H. Dong and I. Kotenko, “An autoencoder-based multi-task learning for intrusion detection in IoT networks,” in *Proc. IEEE Ural-Sib. Conf. Biomed. Eng., Radioelectron. Inf. Technol. (USBEREIT)*, Yekaterinburg, Russian Federation, 2023, pp. 1-4.

[159] H. Sonune and N. Kulkarni, “Leveraging AI for enhanced botnet detection-A review of machine learning approach for cybersecurity,” in *Proc. IEEE Int. Conf. Blockch. Distrib. Syst. Secur. (ICBDS)*, Pune, India, 2024, pp. 1-8.

[160] Y. Liu et al., “An access control mechanism based on risk prediction for the IoV,” in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Antwerp, Belgium, 2020, pp. 1-5.

[161] Y. Yang et al., “Generative AI for secure and privacy-preserving mobile crowdsensing,” *IEEE Wireless Commun.*, vol. 31, no. 6, pp. 29-38, Dec. 2024.

[162] Y. Huang, Y. Chen, Y. Zhang, J. Zou, Z. Tan, and N. Hu, “TFHM: A traffic feature hiding scheme based on generative adversarial networks,” in *Proc. IEEE 7th Int. Conf. Data Sci. Cyberspace (DSC)*, Guilin, China, 2022, pp. 175-182.

[163] X. Lu and W. Song, “Improved trajectory data encryption method for internet of vehicles using GAN-based chaotic logistic algorithm,” *Alex. Eng. J.*, vol. 114, 2025.

[164] K. Danladi Sankara, G. Saritha, S. Anitha Elavarasi, S. Saradha, and D. Arul Kumar, “A hybrid LSTM-GAN model for predictive cyber threat intelligence and anomaly detection,” in *Proc. Int. Conf. Integr. Intell. Commun. Syst. (ICIICS)*, Kalaburagi, India, 2024, pp. 1-6.

[165] K. H. V. Prasad and S. Periyasamy, “Secure-Energy efficient bio-inspired clustering and deep learning-based routing using blockchain for edge assisted WSN environment,” *IEEE Access*, vol. 11, pp. 145421-145440, 2023.

[166] Y. He, M. Kong, C. Du, D. Yao, and M. Yu, “Communication security analysis of intelligent transportation system using 5G internet of things from the perspective of big data ,” *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2199-2207, Feb. 2023.

[167] N. Kumar and A. Chaudhary, “Surveying cybersecurity vulnerabilities and countermeasures for enhancing UAV security,” *Comput. Netw.*, vol. 252, 2024.

[168] S. Mao, J. Guo, and Z. Li, “Discriminative autoencoding framework for simple and efficient anomaly detection,” *IEEE Access*, vol. 7, pp. 140618-140630, 2019.

[169] X. Guo, C. Zhu, J. Yang, and Y. Xiao, “An anomaly detection model for ADS-B systems based on improved GAN and LSTM networks,” in *Proc. IEEE 21st Int. Conf. Commun. Technol. (ICCT)*, Tianjin, China, 2021, pp. 802-809.

[170] H. Liang, L. Song, J. Du, X. Li and L. Guo, “Consistent anomaly detection and localization of multivariate time series via cross-correlation graph-based encoder–decoder GAN,” *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1-10, 2022, Art no. 3504210.

[171] G. Yan, Q. Wang, Z. Liu, C. Su, and R. Gao, “Anomaly detection in multidimensional spacecraft telemetry via an integrated K-means, VAE and SVDD model,” in *Proc. 6th Int. Conf. Electron. Eng. Inform. (EEI)*, Chongqing, China, 2024, pp. 445-449.

[172] M. Kim, J. Yu, J. Kim, T. -H. Oh, and J. K. Choi, “An iterative method for unsupervised robust anomaly detection under data contamination,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 10, pp. 13327-13339, Oct. 2024.

[173] C. Li et al., “A zero-shot fault detection method for UAV sensors based on a novel CVAE-GAN model,” *IEEE Sensors J.*, vol. 24, no. 14, pp. 23239-23254, Jul. 2024.

[174] L. M. Da Silva, I. G. Ferrão, C. Dezan, D. Espes, and K. R. L. J. C. Branco, “Anomaly-Based intrusion detection system for in-flight and network security in uav swarm,” in *Proc. 21st Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Warsaw, Poland, 2023, pp. 812-819.

[175] S. Salim, N. Moustafa, and A. Almorjan, “Responsible deep-federated-learning-based threat detection for satellite communications,” *IEEE Internet Things J.*, vol. 12, no. 5, pp. 4807-4819, Mar. 2025.

[176] T. Wang, Z. Zhao, and K. Wu, “Exploiting LLM embeddings for content-based IoT anomaly detection,” in *Proc. IEEE Pac. Rim Conf. Commun. Comput. Signal Process. (PACRIM)*, Victoria, BC, Canada, 2024, pp. 1-6.

[177] H. Zhang, Z. Song, J. Yang, and Y. Gao, “Adversarial autoencoder empowered joint anomaly detection and signal reconstruction from sub-nyquist samples,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 3, pp. 618-628, Jun. 2023.

[178] H. H. Addeen, Y. Xiao, and T. Li, “A CVAE-based anomaly detection algorithm for cyber physical attacks for water distribution systems,” *IEEE Access*, vol. 12, pp. 48321-48334, 2024.

[179] Z. Gu, G. Tang, and J. Ma, “Compressive-Sensing reconstruction for satellite monitor data using a deep generative model,” *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1-15, 2024, Art no. 2535015.

[180] L. Guo, Y. Liu, Y. Li, and K. Yang, “High-Precision reconstruction method based on MTS-GAN for electromagnetic environment data in SAGIoT MTS-GAN,” *EURASIP J. Adv. Signal Process.*, vol. 2023. no. 1, 2023.

[181] M. Shao, C. Wang, W. Zuo, and D. Meng, “Efficient pyramidal GAN for versatile missing data reconstruction in remote sensing images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1-14, 2022, Art no. 5626014.

[182] A. Rouhbakhshmeghrazi, B. Li, W. Iqbal, and G. Alizadeh, “Super-Resolution reconstruction of UAV images with GANs: Achievements and challenges,” in *Proc. 3rd Int. Conf. Cyber-Phys. Soc. Intell.*, Doha, Qatar, 2024, pp. 1-6.

[183] C. Li, X. Liu, and S. Li, “Transformer meets GAN: Cloud-Free multispectral image reconstruction via multisensor data fusion in satellite images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1-13, 2023.

[184] S. Majumder, M.K. Deb Barma, and A. Saha, "ARP spoofing detection using machine learning classifiers: an experimental study," *Knowl. Inf. Syst.*, vol. 67, vol. 1, pp. 727-766, 2025.

[185] G. Xie, L. T. Yang, Y. Yang, H. Luo, R. Li, and M. Alazab, "Threat analysis for automotive CAN networks: A GAN model-based intrusion detection technique," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4467-4477, Jul. 2021.

[186] N. Kumar, K. Berwal, R. Verma, and S. K. Vishvakarma, "Ai-Driven strategies for securing internet of Battlefield things (IoBT)," in *Proc. IEEE Int. Conf. ICT Bus. Ind. Gov. (ICTBIG)*, Indore, India, 2024, pp. 1-6.

[187] N. Sun, W. Wang, K. Liu, D. Li, and J. Lü, "Hybrid framework for security evaluation in internet of vehicles," *Comput. Secur.*, vol. 153, 2025.

[188] S. Li, W. Liu, Y. Wu, and J. Zhao, "Generative architecture for data imputation in secure blockchain-enabled spatiotemporal data management," *J. Web Eng.*, vol. 23, no. 1, pp. 111-163, Jan. 2024.

[189] K. Wang, H. Zheng, Y. Li, J. Li, and A. Louri, "AGAPE: Anomaly detection with generative adversarial network for improved performance, energy, and security in manycore systems," in *Proc. Design Autom. Test Eur. Conf. Exhib. (DATE)*, Antwerp, Belgium, 2022, pp. 849-854.

[190] A.S. Iliyasu and H. Deng, "N-GAN: A novel anomaly-based network intrusion detection with generative adversarial networks," textitInt. j. inf. tecnol., vol. 14, pp. 3365-3375, 2022.

[191] X. Zhao, K. W. Fok, and V. L.L. Thing, "Enhancing network intrusion detection performance using generative adversarial networks," *Comput. Secur.*, vol. 145, 2024.

[192] J. Chen, X. Gao, R. Deng, Y. He, C. Fang, and P. Cheng, "Generating adversarial examples against machine learning based intrusion detector in industrial control systems," *IEEE Trans. Depend. Sec. Comput.*, vol. 19, no. 3, pp. 1810-1825, May-June 2022.

[193] R. Prabha, G.A. Senthil, and S. Suganthi, "Cluster head based secure routing using optimized dual-discriminator conditional generative adversarial network in wireless ad-hoc networks," *Peer Peer Netw. Appl.*, vol. 16, no. 6, pp. 2747-2760, 2023.

[194] M. Tanvir Masud, N. Koroniotis, M. Keshk, B. Turnbull, S. Kasra Kermanshahi, and N. Moustafa, "Generative fuzzer-driven vulnerability detection in the internet of things networks," *Appl. Soft Comput.*, vol. 174, 2025.

[195] A. E. Morel, Z. Murry, K. Kostage, C. Qu, and P. Calyam, "Enhancing drone video analytics security management using an AERPAW testbed," in *Proc. IEEE Int. Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Vancouver, BC, Canada, 2024, pp. 1-6.

[196] M. Asif, M. A. Rahman, K. Akkaya, H. Shahriar, and A. Cuzzocrea, "Adversarial data-augmented resilient intrusion detection system for unmanned aerial vehicles," in *Proc. IEEE Int. Conf. Big Data*, Sorrento, Italy, 2023, pp. 5428-5437.

[197] B. S. Sarikaya and Ş. Bahtiyar, "Generative Adversarial Networks for Synthetic Jamming Attacks on UAVs," in *Proc. 9th Int. Conf. Comput. Sci. Eng. (UBMK)*, Antalya, Turkiye, 2024, pp. 760-765.

[198] L. Arcangeloni, E. Testi, and A. Giorgetti, "Jamming detection in MIMO-OFDM ISAC systems using variational autoencoders," in *Proc. IEEE Int. Symp. Syst. Eng. (ISSE)*, Perugia, Italy, 2024, pp. 1-7.

[199] Y. Tang, Z. Zhao, X. Ye, S. Zheng, and L. Wang, "Jamming recognition based on AC-VAEGAN," in *Proc. 15th IEEE Int. Conf. Signal Process. (ICSP)*, Beijing, China, 2020, pp. 312-315.

[200] A. Meftah, T. N. Do, G. Kaddoum, and C. Talhi, "Federated learning-enabled jamming detection for stochastic terrestrial and non-terrestrial networks," *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 1, pp. 271-290, Mar. 2025.

[201] H. Bouzabia, A. Meftah, and G. Kaddoum, "Federated learning-enabled smart jammer detection in terrestrial and non-terrestrial heterogeneous joint sensing and communication networks," *IEEE Commun. Lett.*, vol. 28, no. 9, pp. 2026-2030, Sept. 2024.

[202] R. Li, Z. Hu, D. Tian, P. He, Z. Liang, and Q. Liu, "A distributed radar multi-jamming recognition method based on transformer network," in *Proc. IEEE Int. Conf. Signal Inf. Data Process (ICSIDP)*, Zhuhai, China, 2024, pp. 1-6.

[203] J. He, X. Li, X. Zhang, W. Niu, and F. Li, "A synthetic data-assisted satellite terrestrial integrated network intrusion detection framework," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 1739-1754, 2025.

[204] Y. Zhu et al., "Mixture gaussian distribution-based collaborative reinforcement learning for 3D UAV localization optimization against jamming attacks," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Cape Town, South Africa, 2024, pp. 529-534.

[205] H. Han et al., "Better late than never: GAN-Enhanced dynamic anti-jamming spectrum access with incomplete sensing information," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1800-1804, Aug. 2021.

[206] İ. Aras, I. Erer, and E. Akdemir, "Vision transformer based adaptive beamforming for GNSS bands," in *Proc. 32nd Telecommun. Forum (TELFOR)*, Belgrade, Serbia, 2024, pp. 1-4.

[207] I. Elleuch, A. Pourranjbar, and G. Kaddoum, "Leveraging transformer models for anti-jamming in heavily attacked UAV environments," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 5337-5347, 2024.

[208] Z. Zhang, Y. Wen, J. Zhang, D. Zhang, and J. Ren, "Conditional diffusion model for anti-interference in train satellite navigation receiver signals," in *Proc. 8th Int. Conf. Commun. Inf. Syst. (ICCIS)*, Shenzhen, China, 2024, pp. 102-106.

[209] R. Tang, D. Gao, M. Yang, T. Guo, H. Wu, and G. Shi, "GAN-Inspired intelligent jamming and anti-jamming strategy for semantic communication systems," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Rome, Italy, 2023, pp. 1623-1628.

[210] P. Guo et al., "Few-Shot source separation for IoT anti-jamming via multi-task learning and meta-learning," *IEEE Internet Things J.*, vol. 12, no. 11, pp. 16761-16776, Jun. 2025.

[211] S. Rajasoundaran, et al. "Secure and optimized intrusion detection scheme using LSTM-MAC principles for underwater wireless sensor networks," *Wirel. Netw.*, vol. 30, no. 1, pp. 209-231, 2024.

[212] A. H. Hasan, M. Anbar, and T.A. Alamiedy, "Deep learning approach for detecting router advertisement flooding-based DDoS attacks," *J. Amb. Intel. Hum. Comp.*, vol. 14, no. 6, pp. 7281-7295, 2023.

[213] K. Barik, S. Misra, and L. Fernandez-Sanz, "Adversarial attack detection framework based on optimized weighted conditional stepwise adversarial network," *Int. J. Inf. Secur.*, vol. 23, no. 3, pp. 2353-2376, 2024.

[214] R. Yaegashi, E. Takeshita, and Y. Nakayama, "Two-Stage DDoS mitigation with variational auto-Encoder and cyclic queuing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seoul, Korea, 2022, pp. 5421-5426.

[215] I. Ortega-Fernandez, et al., "Network intrusion detection system for DDoS attacks in ICS using deep autoencoders," *Wirel. Netw.*, vol. 30, no. 6, pp. 5059-5075, 2024.

[216] B. Nugraha, K. Kota, and T. Bauschert, "Leveraging federated learning and variational autoencoders for an enhanced anomaly detection system," in *Proc. IEEE 10th Int. Conf. Netw. Softwarization (NetSoft)*, Saint Louis, MO, USA, 2024, pp. 166-174.

[217] J. Singh, J. Ali, P. Sharma, V. Kumar, M. Sharma, and R. Singh, "Mitigating distributed denial of service (DDoS) attacks in cloud networks using neural networks," in *Proc. 2nd Int. Conf. Adv. Comput. Commun. Technol. (ICACCTech)*, Sonipat, India, 2024, pp. 161-166.

[218] G. Aceto et al., "Synthetic and privacy-preserving traffic trace generation using generative AI models for training network intrusion detection systems," *J. Netw. Comput. Appl.*, vol. 229, 2024.

[219] T. Sood, "Intrusion detection system in wireless sensor network using conditional generative adversarial network," *Wirel. Pers. Commun.*, vol. 126, no. 1, pp. 911-931, 2022.

[220] M. Hassanin et al., "PLLM-CS: Pre-trained large language model (LLM) for cyber threat detection in satellite networks," *Ad Hoc Netw.*, 2025.

[221] M. Mahmoodi and S. M. Jameii, "Utilizing large language models for DDoS attack detection," in *Proc. OPJU Int. Technol. Conf. Smart Comput. Innov. Adv. Ind. 4.0 (OTCON)*, Raigarh, India, 2024, pp. 1-6.

[222] R. Zhang, K. Xiong, Y. Lu, P. Fan, D. W. K. Ng, and K. B. Letaief, "Energy Efficiency Maximization in RIS-Assisted SWIPT Networks With RSMA: A PPO-Based Approach," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 5, pp. 1413-1430, May 2023.

[223] E. Thomas, M. Jan Hendrik, and H. Frank. "Neural architecture search: A survey," *J. Mach. Learn. Res.*, vol. 20, no. 55, pp. 1–21, 2019.

[224] A. Kharbouch, F. H. Aghdam, N. Gholipoor, and M. Rasti, "Digital-Twin-6G empowered future smart grid applications," *IEEE Wireless Commun.*, vol. 32, no. 3, pp. 90-97, Jun. 2025.

[225] H. Guo et al., "A comprehensive survey on continual learning in generative models," 2025, *arXiv:2506.13045v3*.

[226] S. K. Jagatheesaperumal, S. Ali, A. Alotaibi, K. Muhammad, V. H. C. De Albuquerque, and M. Guizani, "Generative AI-Enhanced neuro-symbolic quantum architectures for secure communications and networking," *IEEE Netw.*, 2025.