International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V72I7P102

Volume 72 Issue 7, 12-16, July 2024
© 2024 Seventh Sense Research Group®

Original Article

Harnessing IoT Potential with Generative AI: Utilizations, Real-world Examples, and Boundaries

Tanvi Hungund[1], Shobhit Kukreti[2], Priyank Singh[3]

[1]California State University, Fullerton, United States.
[2]Carnegie Mellon University, United States.
[3]Rochester Institute of Technology, United States.

[1]Corresponding Author : Tsh110390@gmail.com

Received: 08 May 2024 Revised: 29 June 2024 Accepted: 13 July 2024 Published: 31 July 2024

Abstract - *The emergence of the Generative Pre-Trained Transformer (GPT) language model, commonly known as ChatGPT, has spotlighted the continuously evolving realm of Generative AI (GAI). With the current strides in Graphics Processing Units (GPUs), training and deploying deep generative models has become more accessible. Concurrently, advancements in edge computing have facilitated leveraging GAI's potential across various applications in the Internet of Things (IoT). This paper delves into the possibilities of amalgamating GAI with IoT technology to forge innovative solutions addressing shortcomings in diverse IoT domains. Specifically, it explores how GAI can mitigate challenges stemming from inadequate and incomplete data in IoT systems by generating synthetic data for training other deep models. Furthermore, it examines GAI's role in tailoring content produced by IoT devices and other collaborative applications. It also scrutinizes real-world implementations of this synergy. Finally, the article concludes by outlining the current limitations of GAI technology for IoT applications and proposing avenues for future improvement.*

Keywords - *Generative Pre-Trained Transformer (GPT), Graphics Processing Units (GPUs), Internet of Things (IoT), Generative AI (GAI), Data Privacy and Security.*

1. Introduction

The fusion of Generative Artificial Intelligence (GAI) with the Internet of Things (IoT) is revolutionizing the IoT landscape. GAI enables edge computing-based devices to analyze vast, disparate data sets and refine and augment them, benefiting other Deep Learning (DL) applications. This integration unleashes a potent synergy, opening avenues for diverse applications and opportunities. In 2022, the market for GAI in IoT was valued at USD 947.8 million, poised for substantial growth to an estimated USD 8,952.6 million by 2032, with an impressive projected Compound Annual Growth Rate (CAGR) of 25.9%. By harnessing GAI alongside IoT and edge computing, numerous possibilities emerge. For instance, IoT devices can generate synthetic data to complement real-time sensor data, addressing scarcity and incompleteness. This enhances the accuracy and resilience of models and algorithms built on such data. Moreover, GAI can predict and simulate sensor data, enabling proactive decision-making and predictive maintenance in IoT ecosystems. GAI enhances user experiences by generating personalized content and recommendations based on individual preferences and behaviours. For instance, IoT devices can leverage GAI in home settings to create personalized ambient lighting, curated music playlists, or bespoke artwork tailored to occupants' tastes and moods. In industrial environments, GAI can facilitate virtual prototyping and simulation of real-world scenarios, streamlining design optimization, testing, and validating IoT-enabled systems, thereby reducing costs and time-to-market for IoT deployments. GAI holds significant promise in augmenting various facets of IoT devices, from addressing data limitations to empowering predictive maintenance and enhancing user experiences. Although the full potential of GAI in IoT is yet to be realized, its integration promises innovative applications and expanded capabilities for IoT devices. As technology progresses and GAI techniques evolve, GAI is poised to play a pivotal role in enhancing IoT system functionality and user experiences. This article delves into these possibilities and explores real-world scenarios where GAI integration significantly impacts IoT device operations.

2. Overview of Generative Artificial Intelligence (GAI)

Generative Artificial Intelligence (GAI) represents a dynamic frontier in AI, employing advanced techniques and models to generate fresh and original content. Unlike conventional AI methods that rely on analyzing existing data, GAI leverages extensive datasets to craft novel data based on learned patterns. Key to GAI's capabilities are generative models like Generative Adversarial Networks (GANs) and

Variational Autoencoders (VAEs). GANs, for instance, consist of a generator-discriminator duo working in concert. While the generator strives to produce data resembling the training data, the discriminator discerns between real and generated data. Through adversarial training, both networks refine their abilities, yielding outputs of increasing realism. GAI has achieved remarkable progress across diverse domains. In image generation, it produces intricate and lifelike images mirroring real-world objects, scenes, and even human faces. Similarly, in natural language processing, GAI facilitates the creation of coherent and contextually relevant text, aiding tasks such as language translation, text completion, and dialogue generation. Prominent large language models include GPT-3.5 and GPT-4 by OpenAI, Llama and Llama 2 by Meta, and Chinchilla AI by DeepMind. Cutting-edge image generation models include DALL-E and DALL-E 2 by OpenAI, Midjourney Art Generator, and Stable Diffusion by Stability AI. Video-generating models like Meta Make-A-Video, Google Imagen video, Synthesia Studio, and Krikey.ai have also delivered on their promises. For audio creation, popular models include MuseNet by OpenAI, Murf.ai AI voice generator, and Soundful AI music generator. Notably, GAI also extends to 3D visual object generation, with examples such as OpenAI's Point-E, Microsoft's Rodin Diffusion, and Google's DreamFusion.

2.1. Applications

2.1.1. Synthetic Sensor Data Generation:

GAI addresses a major challenge in IoT devices—lack of consistent training data—by generating synthetic sensor data mirroring real IoT sensor data. By learning from available data, generative models produce plausible values, enabling more comprehensive analysis and facilitating training and testing of IoT systems, particularly in scenarios with limited or inconsistent data availability.

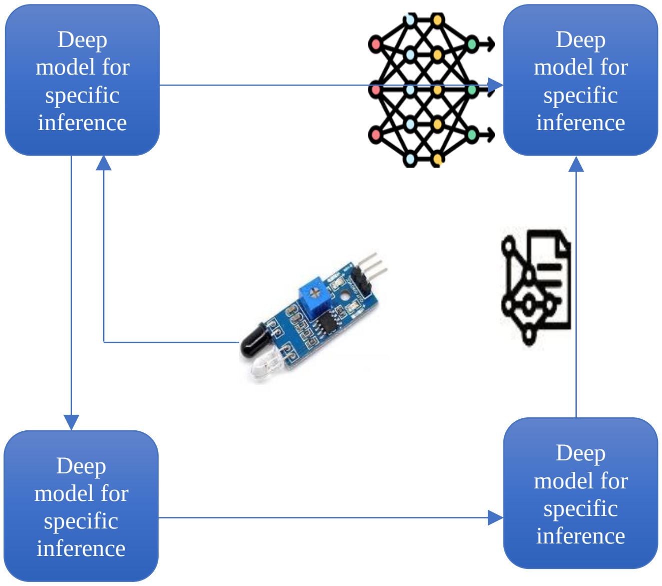
Deep model for specific inference
Deep model for specific inference
Deep model for specific inference
Deep model for specific inference

Fig. 1 Generating synthetic data

2.2. Personalized Device Response

GAI enhances user experiences by enabling IoT devices to offer personalized responses based on learned patterns of user preferences. By analyzing user behaviors, devices can make accurate predictions, leading to optimized performance, energy consumption, cost savings, and reduced maintenance downtime.

2.3. Autonomous Control

GAI revolutionizes autonomous IoT device control by generating logic and commands based on learned movement patterns, eliminating the need for individualized programming for each limb. This approach streamlines system training and enhances adaptability to varying environments.

2.4. Data Privacy and Security

GAI offers innovative solutions to protect data privacy by generating synthetic data for training and testing IoT systems without compromising user privacy. By analyzing normal behavior patterns, GAI aids in the early detection and mitigation of security threats, ensuring robust security measures in IoT networks and devices.

2.4.1. Predictive Maintenance

By simulating real-world device behavior, GAI facilitates predictive maintenance through the creation of digital twins for IoT devices. These virtual replicas receive real-time data from sensors, enabling proactive identification of maintenance needs, resource optimization, and minimization of unexpected failures. Data anonymization is another area where GAI proves invaluable. By harnessing generative models, sensitive information can be transformed or obscured to render it unidentifiable while preserving the data's utility and integrity. This enables organizations to conduct data analysis and derive insights while adhering to privacy regulations and safeguarding individual confidentiality. Generative models can introduce noise or perturbations into the data, making it challenging to pinpoint specific individuals or extract sensitive details. Consequently, privacy is upheld even during aggregated data analysis, as depicted in Fig. 2.

Moreover, GAI enhances data security and mitigates potential vulnerabilities in edge computing-based IoT devices. In real time, abnormal patterns or malicious activities can be identified using generative anomaly and intrusion detection models. This facilitates prompt responses and the mitigation of security breaches, thereby safeguarding the integrity and privacy of the data collected by IoT devices.

In a study outlined by [7], an allowed method for sensor data obfuscation is proposed, employing a guided denoising diffusion model and a surrogate model for the intended inference. This obfuscation model, tested on a human activity recognition dataset, demonstrates a satisfactory balance between privacy and utility without necessitating knowledge of private attributes.

Table 1. illustrates the potential use cases for GAI in IoT devices, anticipated through future developments

Use Case	Description	Maturity
Code generation and debugging	LLMs can be used to create, complete, combine or debug code which aids developers in constructing innovative applications.	In use
Learning robotic movements	GAI can be used to automatically learn how a robot (humanoid or animal-like) can move more efficiently and safely.	First applications
Surveillance and security	GAI can understand a video's temporal and spatial elements based on previous information, and could predict the movement and actions of a potentially dangerous object or agent.	Years out
Social devices	LLMs can be fine-tuned on user data ultimately allowing the user to change settings of the machine by talking to it.	Years out

2.5. *Industrial IoT - A Real-World Scenario*

Industrial IoT (IIoT) encompasses networks of interconnected sensors, instruments, and other devices deployed to automate various industrial processes in factories, proving vital for manufacturing and energy management. GAI, in synergy with IoT devices, holds immense potential in real-world scenarios such as IIoT. For instance, GAI in IoT devices holds great promise in healthcare for tasks like enhancing medical image quality, facilitating medical training simulations, and improving patient monitoring through data augmentation.

Here, we explore the applications of GAI in IoT within the context of IIoT, painting a picture of a promising future. In IoT, crucial considerations include prioritizing safety, security, and low-latency communication for real-time operations. Traditional discriminative deep learning algorithms often struggle in these scenarios due to their reliance on large quantities of balanced labelled training data and their difficulty in handling uncertain input data.

To tackle these challenges, researchers have turned to deep generative models (DGMs), merging probabilistic modelling with the adaptability of deep learning, offering a more efficient approach for meeting the specific requirements of IIoT applications, as detailed in [14].

GAI models applied to various use cases in IIoT applications are highlighted in Table 2, providing a reassuring glimpse into their utility.

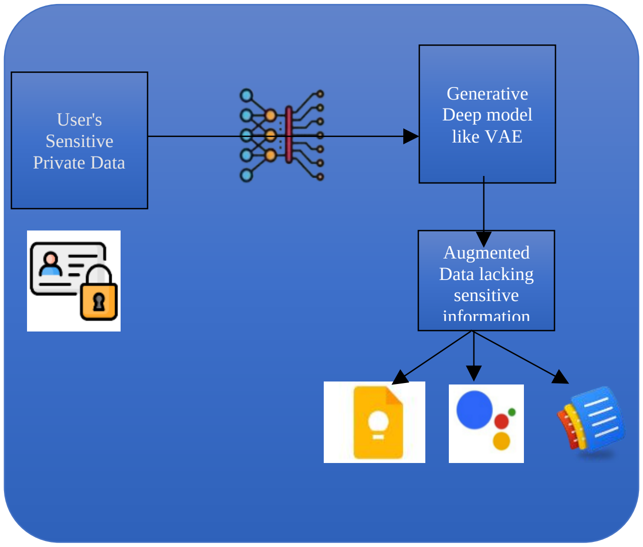
User's Sensitive Private Data
Generative Deep model like VAE
Augmented Data lacking sensitive information

Fig. 2 Anonymization via augmentation

3. Primary Applications of DGMs in IIoT Systems Include

3.1. *Anomaly Detection*

Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) have demonstrated efficacy in anomaly detection. They learn the underlying distribution of the given data and determine if a sample belongs to that distribution by comparing it to the nearest sample in the generated distribution. Given the substantial volume of multidimensional data produced in the context of 6G IIoT, researchers in [8] devised an autoregressive exogenous model (ARX) to remove noise from the data, thereby enhancing anomaly detection. In [9], another group of researchers employed Conditional Generative Adversarial Networks (CGANs) to identify security anomalies.

3.2. *Trust Boundary Protection*

Within IIoT, trust-boundary techniques are deployed to partition networks, effectively segregating IIoT processes and data storage based on user access privileges. Trust-boundary protection relies on intrusion detection as a fundamental aspect of regulating access levels. The DIGFuPAS system [10] is dedicated to enhancing the effectiveness of Intrusion Detection Systems (IDS) against adversarial attacks. It utilizes a Wasserstein Generative Adversarial Network (WGAN) to analyze network traffic flow meticulously. Conversely, ARIES [11] is a multilayered IDS that merges unsupervised GANs with supervised decision trees and Support Vector Machines (SVMs). ARIES' initial layer conducts supervised classification of attacks, such as denial of service and brute force, while subsequent layers identify abnormalities in packet and operating data.

3.3. *Network Traffic Prediction*

Predicting end-to-end network traffic is paramount for various security and management functions in IIoT.

Moreover, Quality of Service (QoS) heavily relies on factors like packet size distributions and interarrival time. To address limited data samples in channel fading models, Liu et al. [12] utilized GANs to dynamically learn wireless channel distributions within factory environments. Meanwhile, the versatility of GANs has been showcased in traffic classification scenarios with less than 20% labelled traffic flows [15].

3.4. Platform Monitoring

An RNN-based VAE is deployed for motor fault detection using motor vibration time-domain signals [13]. Authors employ a deep variational information bottleneck approach to extract latent variables related to quality, minimizing mutual information between latent variables and observations while maximizing process quality. GANs also assist in overcoming challenges related to limited fault data availability and datasets in manufacturing IIoTs, particularly in the context of predictive maintenance.

4. Computational and Energy Requirements

Integrating GAI with edge computing-based devices necessitates consideration of computational resources and energy efficiency. IoT devices often need more computational capabilities, memory, and power. The computationally intensive nature and large size of GAI models pose challenges for deployment on resource-constrained IoT devices. Optimization techniques and hardware enhancements are necessary for real-time GAI on IoT devices while maintaining energy efficiency.

4.1. Time Efficiency

Real-time inference and latency are critical factors in many IoT applications requiring timely decision-making based on data analysis. However, generating content using complex GAI models can introduce significant latency, impacting real-time or near-real-time processing requirements.

4.1.1. Data Reliability

Data quality and distribution play vital roles in the effectiveness of GAI models. The success of these models heavily relies on the quality and diversity of the training data. Data in IoT environments can be noisy, incomplete, or biased due to environmental conditions or sensor limitations. Addressing data biases and ensuring high-quality data collection practices become critical to training robust and reliable GAI models for IoT applications.

4.2. Risk of Misuse

Ethical and legal implications arise from the potential misuse of GAI in IoT contexts, such as deepfakes, which raise concerns about impersonation or unauthorized content creation. Adherence to ethical guidelines and legal frameworks is essential to prevent malicious activities or harm.

4.3. Inexplainable Complexities

Interpretability and explainability pose significant challenges in GAI. The inner workings of GAI models are often considered black boxes, making it difficult to understand the decision-making process or explain the generated content, particularly in sensitive IoT applications requiring transparency and accountability.

4.4. Generalization

Adaptability and generalization are vital in dynamic IoT environments. GAI models trained on static datasets may struggle to generalize to new scenarios as IoT deployments are subject to changes in conditions or device configurations.

5. Conclusion

This article underscores the potential of combining Generative AI (GAI) with Internet of Things (IoT) technology. Leveraging GAI, IoT systems can overcome challenges posed by data insufficiency and generate synthetic data for training machine learning models.

Additionally, GAI can personalize content generated by IoT devices, facilitate proactive decision-making, ensure data privacy, and enhance predictive maintenance in IoT systems. The synergy between GAI and IoT holds promise for revolutionizing technology interaction, fostering innovation, and driving growth. Future research in this domain should address challenges and explore new applications, further enhancing the synergy between GAI and IoT.

References

[1] Moustafa Alzantot, Supriyo Chakraborty, and Mani Srivastava, "Sensegen: A Deep Learning Architecture for Synthetic Sensor Data Generation," *2017 IEEE International Conference on Pervasive Computing and Communications Workshops*, Kona, HI, USA, pp. 188-193, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[2] Mauajama Firdaus et al., "I Enjoy Writing and Playing, Do You?: "A Personalized and Emotion Grounded Dialogue Agent Using Generative Adversarial Network," *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 2127-2138, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Minrui Xu et al., "Generative AI-Empowered Simulation for Autonomous Driving in Vehicular Mixed Reality Metaverses," *IEEE Journal of Selected Topics in Signal Processing*, vol. 17, no. 5, pp. 1064-1079, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Hui Wang et al., "Privacy-Preserving Federated Generative Adversarial Network for IoT," *2021 International Conference on Networking and Network Applications*, Lijiang City, China, pp. 75-80, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Mohamed Amine Ferrag, Merouane Debbah, and Muna Al-Hawawreh, “Generative AI for Cyber Threat-Hunting in 6G-Enabled IoT Networks,” *23rd International Symposium on Cluster, Cloud and Internet Computing Workshops*, Bangalore, India, pp. 16-25, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Sourajit Behera, and Rajiv Misra, “Generative Adversarial Networks Based Remaining Useful Life Estimation for IIoT,” *Computers & Electrical Engineering*, vol. 92, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Xin Yang, and Omid Ardakanian, “Privacy through Diffusion: A White-Listing Approach to Sensor Data Anonymization,” *CPSIoTSec '23: Proceedings of the 5th Workshop on CPS&IoT Security and Privacy*, Copenhagen Denmark, pp. 101-107, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Guangjie Han et al., “Anomaly Detection Based on Multidimensional Data Processing for Protecting Vital Devices in 6G-Enabled Massive IIoT,” *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5219-5229, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Viacheslav Belenko et al., “Evaluation of GAN Applicability for Intrusion Detection in Self-Organizing Networks of Cyber Physical Systems,” *2018 International Russian Automation Conference (RusAutoCon)*, Sochi, Russia, pp. 1–7, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Phan The Duy et al., “DIGFuPAS: Deceive IDs with GAN and Function-Preserving on Adversarial Samples in SDN-Enabled Networks,” *Computers & Security*, vol. 109, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Panagiotis Radoglou Grammatikis et al., “Aries: A Novel Multivariate Intrusion Detection System for Smart Grid,” *Sensors*, vol. 20, no. 18, pp. 1-20, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Chen-Feng Liu, and Mehdi Bennis, “Data-Driven Predictive Scheduling in Ultra-Reliable Low-Latency Industrial IoT: A Generative Adversarial Network Approach,” *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications*, Atlanta, GA, USA, pp. 1-5, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Yang Huang, Chiun-Hsun Chen, and Chi-Jui Huang, “Motor Fault Detection and Feature Extraction Using RNN-Based Variational Autoencoder,” *IEEE Access*, vol. 7, pp. 139086–139096, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Suparna De et al., “Deep Generative Models in the Industrial Internet of Things: A Survey,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 5728–5737, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Laisen Nie et al., “Network Traffic Prediction in Industrial Internet of Things Backbone Networks: A Multitask Learning Mechanism,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 7123–7132, 2021. [CrossRef] [Google Scholar] [Publisher Link]