International Journal of Intelligent Networks 5 (2024) 267–274

Contents lists available at ScienceDirect

International Journal of Intelligent Networks

journal homepage: www.keaipublishing.com/en/journals/international-journal-of-intelligent-networks

Generative artificial intelligence for distributed learning to enhance smart grid communication

Seyed Mahmoud Sajjadi Mohammadabadi [a,*], Mahmoudreza Entezami [b], Aidin Karimi Moghaddam [c], Mansour Orangian [d], Shayan Nejadshamsi [e]

[a] University of Nevada, Computer Science and Engineering Department, Reno, NV, USA
[b] Concordia University, Computer Science and Software Engineering Department, Montreal, Quebec, Canada
[c] University of Zanjan, Department of Mathematics and Computer Science, Zanjan, Iran
[d] Islamic Azad University, South Tehran Branch, Department of Engineering, Tehran, Iran
[e] Concordia University, Concordia Institute for Information Systems Engineering (CIISE), Montreal, Quebec, Canada

ARTICLE INFO

Keywords:
Energy forecasting
Generative AI
Smart grid
Communication efficiency
Distributed training
LSTM

ABSTRACT

Machine learning models are the backbone of smart grid optimization, but their effectiveness hinges on access to vast amounts of training data. However, smart grids face critical communication bottlenecks due to the ever-increasing volume of data from distributed sensors. This paper introduces a novel approach leveraging Generative Artificial Intelligence (GenAI), specifically a type of pre-trained Foundation Model (FM) architecture suitable for time series data due to its efficiency and privacy-preserving properties. These GenAI models are distributed to agents, or data holders, empowering them to fine-tune the foundation model with their local datasets. By fine-tuning the foundation model, the updated model can produce synthetic data that mirrors real-world grid conditions. The server aggregates fine-tuned model from all agents and then generates synthetic data which considers all data collected in the grid. This synthetic data can be used to train global machine learning models for specific tasks like anomaly detection and energy optimization. Then, the trained task models are distributed to agents in the grid to leverage them. The paper highlights the advantages of GenAI for smart grid communication, including reduced communication burden, enhanced privacy through anonymized data transmission, and improved efficiency and scalability. By enabling a distributed and intelligent communication architecture, GenAI introduces a novel way for a more secure, efficient, and sustainable energy future.

1. Introduction

The success of smart grids, a cornerstone of a clean and reliable energy future, hinges on robust communication infrastructure. The ever-increasing volume of data from sensors and devices in these modernized electrical networks threatens to impede this progress. This data deluge, fueled by the proliferation of advanced metering infrastructure (AMI) and home energy management systems (HEMS), strains existing communication infrastructure. As the Internet of Things (IoT) connects everyday devices to the grid, the continuous generation of massive energy datasets becomes a challenge. For example, a utility managing millions of customers could face an annual data intake exceeding 1 petabyte (PB). Traditional methods rely on transmitting raw data from grid edge devices (e.g., smart meters, sensors) to a central server for processing and analysis. This approach can leverage the computational power of centralized servers to train very powerful machine learning models. These models can then be used for critical tasks like anomaly detection, optimizing energy use, and forecasting demand. However, transmitting large volumes of raw data creates significant communication burdens and raises privacy concerns. The data may contain sensitive information about individual consumers' energy usage patterns.

Existing approaches to load modeling in smart grids fall into two main categories. **1) Model-based** approaches rely on complex mathematical equations derived from the physical characteristics of each load. This requires in-depth knowledge of every device within the grid, making them inflexible and difficult to scale for a large number of diverse loads. Additionally, these models struggle to account for user behavior, which can significantly impact energy consumption patterns. **2) Data-driven** approaches offer a more flexible alternative. They leverage real-world data collected from sensors and meters to learn the

* Corresponding author.
E-mail address: mahmoud.sajjadi@unr.edu (S.M. Sajjadi Mohammadabadi).

https://doi.org/10.1016/j.ijin.2024.05.007
Received 29 March 2024; Received in revised form 19 May 2024; Accepted 21 May 2024
Available online 25 May 2024

behavior of loads without relying on pre-defined assumptions. Techniques like machine learning and Generative Adversarial Networks (GANs) are employed in this approach. However, current data-driven methods have a key limitation: they primarily focus on understanding overall consumption patterns at the building or neighborhood level. This lack of granularity prevents them from capturing the behavior of individual loads, which is crucial for many advanced applications within smart grids.

GenAI offers a transformative solution by enabling a paradigm shift towards secure distributed learning at the grid edge. This approach empowers intelligent agents located at the grid edge to process and analyze data locally. GenAI techniques can be used to generate synthetic data that accurately reflects real-world grid conditions, but without revealing any sensitive information. Then, the synthetic data generated by the server after model aggregation from all grid edge devices can be used to train task-specific machine learning models on the central server, incorporating information from all clients across the grid. This enables a collaborative approach while preserving privacy.

This paper proposes a novel, hybrid approach to optimize smart grid communication by leveraging the power of Generative Artificial Intelligence (see Fig. 1). While traditional methods offer powerful centralized learning, they create communication bottlenecks. GenAI empowers intelligent agents located at the grid edge to process and analyze data locally. We explore how GenAI techniques, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), can be employed to generate synthetic data that accurately reflects real-world grid conditions. This significantly reduces the need for constant communication with a central server, alleviating communication burdens and enhancing overall grid efficiency.

This research paves the way for a more efficient and scalable smart grid communication architecture. By harnessing the power of GenAI, we can unlock the full potential of the smart grid, fostering a more resilient and sustainable energy future. Also this method addresses the challenge of data scarcity in data-driven load modeling by proposing a GenAI-based method for synthetic load profile generation. This method overcomes the limitations of model-based approaches, which require extensive knowledge of each load

2. Literature review

2.1. Machine learning and security considerations

The transformative potential of machine learning for smart grids has been extensively explored. Existing applications, from grid disturbance classification Wei et al. [1], and adaptive control with fuzzy systems Abdali and Monjezi [2], to forecasting Maleki [3], debugging fairness defects in deep neural networks Monjezi et al. [4] and various industry applications Darabi et al. [5]; Tavasoli et al. [6]; EskandariNasab et al. [7]; Wang et al. [8], demonstrate how AI can optimize energy use, seamlessly integrate renewable energy sources, manage storage effectively, and ultimately enhance grid resilience (as surveyed in Ali and Choi [9]). This aligns with the findings by Zhao et al. [10] who highlight AI's role in power load forecasting, energy use optimization, and fault detection, paving the way for a more efficient, reliable, and secure grid. As smart grid systems become increasingly complex, the ability to handle models of unknown systems becomes even more critical, as evidenced by research in multi-agent systems Jandaghi et al. [11]. Furthermore, advanced methods can address limitations of traditional approaches Borhani et al. [12], leading to better results.

Security considerations, however, are paramount. Sakib et al. [13] propose several deep learning models for short-term residential load forecasting. The study explores various models including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), and hybrid architectures like CNN-LSTM. Their findings demonstrate that the CNN-LSTM model achieves the best results in predicting weekly electricity consumption compared to other models. Machine learning fairness is being investigated through hyperparameter tuning, with tree-based algorithms showing promise Herrera et al. [14].

One crucial challenge in deploying real-world machine learning for smart grids is imbalanced data. This occurs when there's a significant difference in the number of examples between different classes. For instance, data representing normal grid conditions might be far more abundant than data signifying anomalies or faults. As discussed in Hasib et al. [15], there are three main approaches to address imbalanced data classification, data-level, algorithm-level, and ensemble methods. Imbalance can be addressed by modifying the data (oversampling, undersampling) or the learning algorithm (cost functions, class weights).

The interconnected nature of the smart grid and the two-way communication with consumers create a complex attack surface, as Grammatikakis et al. [16] emphasize. Their work on a collaborative intelligent intrusion response framework using federated learning demonstrates the critical need for securing these systems while maintaining data privacy between components. Kotsiopoulos et al. [17] review machine learning applications in various domains, highlighting their potential for smart grids. They discuss Industry 4.0 and data analytics, emphasizing how machine learning and deep learning can be used to analyze data and improve grid operations. Ahmad et al. [18] propose a framework for sustainable energy management that integrates smart grid panels with machine learning and IoT, showcasing the real-world application of these technologies for improved efficiency, reliability, and security.

Optimizing grid integration is another crucial aspect. Sulaiman et al. [19] address the challenges of integrating smart grid components into a large electrical power network by proposing a method using LSTM and recurrent Neural Networks (RNN). This exemplifies the ongoing

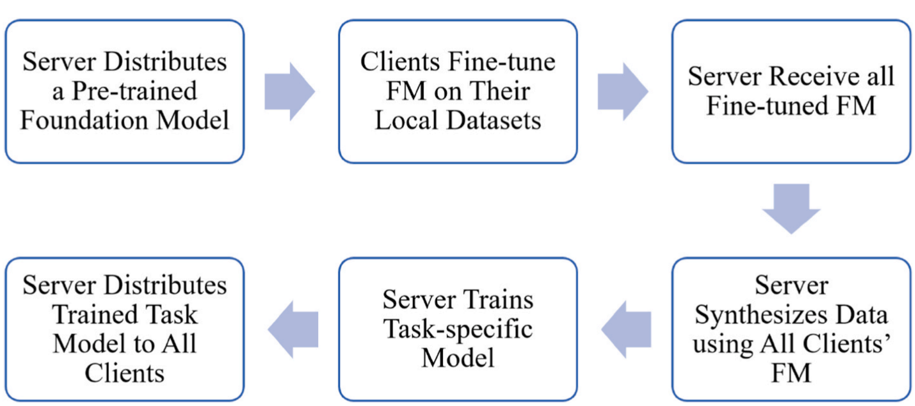

Fig. 1. Overview of data synthesis with GenAI in a distributed approach. Server distributes a pre-trained model. Clients fine-tune it locally and return it. The server then synthesizes data using these models to train a task-specific model, all while preserving privacy.

research efforts to optimize smart grid operations. Furthermore, machine learning empowers researchers to address specific grid challenges. For instance, Schieber et al. [20] delve deeper into energy-aware scheduling for batteryless devices, a critical aspect for optimizing energy usage. Similarly, Dolatabadi et al. [21] explore the use of PMUs to monitor voltage in microgrids and adjust power from DGs for optimal voltage profiles, highlighting machine learning's potential for real-time grid management. Smart grid development is influenced by various factors, including economics. For example, Razmi et al. [22] study how oil prices can indirectly impact smart grids by affecting consumption patterns. Torkaman et al. [23] focus on reconfiguring the power distribution network after faults to improve reliability, voltage profile, and power loss.

2.2. Communication efficiency and data privacy

The ever-growing volume of data in smart grids necessitates efficient communication protocols and robust data privacy mechanisms. The increasing sophistication of generative AI models, as exemplified by DeepFakes Zobaed et al. [24], necessitates robust data security measures to prevent the creation and injection of manipulated data into smart grids.

Tightiz and Yang [25] addressed this challenge by investigating various Internet of Things (IoT) protocols. Their work analyzes the strengths and weaknesses of protocols like IEC 61850, MQTT, CoAP, DDS, AMQP, and OPC UA, considering factors like latency, bandwidth, security, and scalability. This analysis provides valuable insights for selecting the most suitable communication protocol for specific smart grid applications. Data privacy is another critical concern in smart grids. Himthani and Prakash [26] proposed a novel approach using Generative Adversarial Networks (GANs) to encrypt smart grid data. Their technique hides this data within a cover image, offering an additional layer of security.

Federated learning presents another promising approach for balancing communication efficiency and data privacy. This technique allows different clients or entities, like utility companies, to collaboratively train models without directly sharing sensitive data. Mohammadabadi et al. [27] introduced a federated learning method for event classification in power grids using Phasor Measurement Unit (PMU) data. This approach not only protects privacy but also reduces the need for extensive data transmission, leading to faster training and improved communication efficiency. While federated learning offers privacy benefits, it can experience slow training due to device variations. To address this, Mohammadabadi et al. [28,29] further propose a dynamic tiering system in their research. This system offloads computations to a central server for faster training while still maintaining data privacy on participating devices. Their work highlights the potential for further optimization using GenAI within federated learning frameworks, paving the way for even more efficient and secure smart grid management.

2.3. Generative AI for smart grid applications

Machine learning offers a powerful toolkit for optimizing smart grids, but challenges arise when dealing with vast amounts of unlabeled data or limited access to high-quality data. Active learning addresses this by strategically selecting informative data points for labeling, reducing human effort, where Support Vector Machines (SVMs) can be used within Active Learning framework to efficiently identify informative data points in large datasets for labeling Jahan et al. [30].

GenAI, particularly Generative Adversarial Networks (GANs), are emerging as powerful tools for smart grid applications. Large language models can outperform humans on complex tasks like the GMAT exam Ashrafimoghari et al. [31]. Li et al. [32] propose a method using GANs to generate high-frequency data from low-frequency measurements. This high-frequency data can be instrumental for various grid management tasks, such as improving fault detection and optimizing energy use. However, GANs can be susceptible to a phenomenon called mode collapse, where the model gets stuck generating a limited set of outputs instead of diverse, realistic data Yeom et al. [33]; Rombach et al. [34]. This can hinder the generalizability and effectiveness of GAN-based applications in smart grids.

To address this challenge, researchers are exploring alternative generative AI models to enhance communication in power systems Santos et al. [35]; Sajjadi Mohammadabadi [36]. Wang and Zhang [37] propose using stable diffusion models, which excel at generating diverse and high-quality data. Unlike GANs, stable diffusion models achieve this by iteratively adding noise to real data and then learning to remove it, essentially reversing the noise addition process. This technique allows the model to capture the underlying data distribution more effectively, producing more unique and representative samples that are better suited for real-world applications in smart grids.

Although recent advancements in AI have improved efficiency in smart grids, a critical gap remains in fully harnessing GenAI for smart grids, particularly within distributed learning frameworks. This paper addresses this gap by proposing a novel method that leverages pre-trained GenAI models to generate synthetic data on the server-side. This approach facilitates privacy-preserving, collaborative training of machine learning models in a federated learning setting. Moreover, the utilization of tailored datasets generated through our method has the potential to enhance model performance in smart grid applications.

3. Generative AI-powered distributed learning

Smart grids are intelligent electricity that efficiently deliver power through interconnected components, as illustrated in Fig. 2. There are four main elements: distributed energy resources (DERs), distribution/transmission network, customer network, and power system operators. DERs (e.g., rooftop solar panels), generate electricity, while transmission and distribution lines carry it to customers. These customers can also be producers with their own solar panels and major consumers with electric vehicles (EVs). Power system operators manage the entire process (see Fig. 3).

Secure and stable data communication is vital in smart grids. It flows from three key sources: DERs, the distribution network, and customers, all reaching system operators. This data enables advanced analytics and machine learning to enhance grid reliability and optimize power generation planning. However, the high volume generated by numerous sensors and devices, like high-frequency PMUs (e.g., 120 Hz), creates a significant communication burden. On top of the communication challenges, data privacy of customers is a paramount concern in smart grids. This data contains information about energy consumption patterns, which can be exploited by malicious actors in different ways. This could enable targeted burglaries where thieves use insights from energy data to identify times when homeowners are likely to be away.

Considering the challenges of data communication and privacy in smart grids, we propose a novel approach for enhancing data privacy while facilitating data exchange for machine learning applications. Unlike conventional centralized training methods, which can compromise privacy and incur high communication overhead, our approach leverages pre-trained GenAI models. These models are first fine-tune at each client (i.e., data collection point) before being transferred to a central server (i.e., smart grid operator). At the server, the fine-tuned models from each data collection point are used to generate synthetic data. This synthetic data is then used to train a global machine learning model. This novel method achieves two critical goals; first, it effectively mitigates privacy risks by ensuring clients never share their raw consumption data. Since the models themselves don't contain the raw information, only the knowledge learned from it, privacy concerns are significantly reduced. Second, this approach dramatically reduces communication burden. Instead of transmitting vast amounts of raw data from each client, only the significantly smaller, fine-tuned models are transferred. This minimizes bandwidth usage and improves the

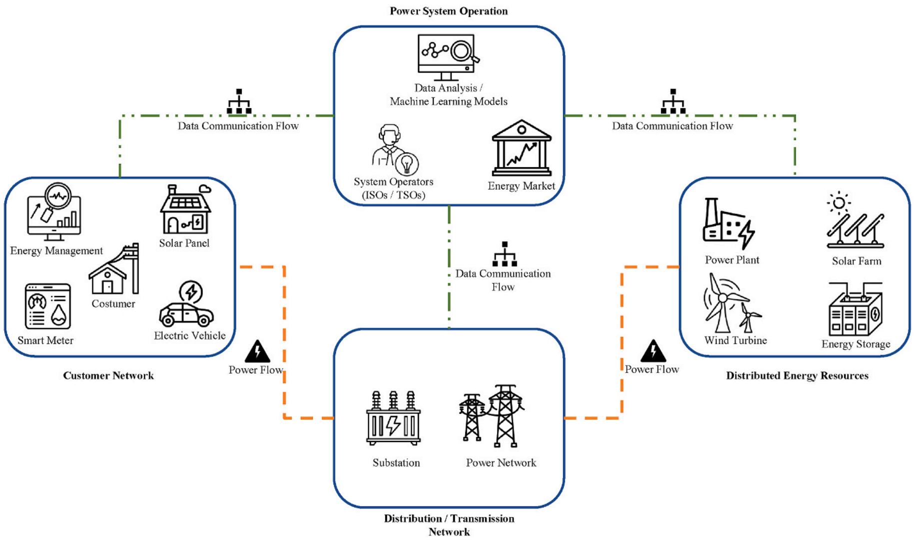

Fig. 2. Overview of Components and Communication in a Smart Grid. Data flow from DERs, grids, and customers empowers analytics for a reliable and optimized grid, but managing data volume and privacy presents Challenges.

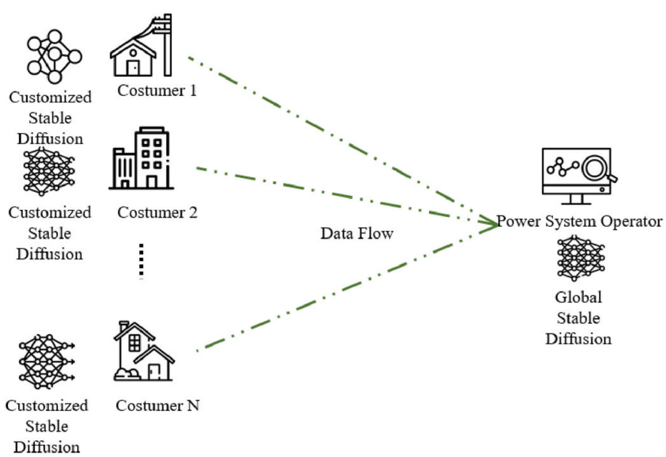

Fig. 3. Overview of Distributed Learning using Generative Models: Leveraging Synthetic Data for Privacy-Preserving Training.

overall efficiency of the data exchange process.

Following the federated learning approach with Generative AI, we can define the optimization problem that the server aims to solve. This involves minimizing an objective function, represented by the following equation:

$$\min_{\theta}\mathscr{L}(\mathbf{x},y;\theta)+\lambda\Omega(\mathbf{x},y;\theta) \tag{1}$$

where $\boldsymbol{x}$ represents the input, y represents the target values, $\mathscr{L}(\cdot)$ denotes the loss function, and $\Omega(\cdot)$ denotes the regularization term. The parameter λ controls the trade-off between fitting the training data and preventing overfitting. This formulation captures the server's objective in the smart grid; iteratively refining model parameters to minimize a combined function. This function balances the model's fit to the synthetic data (captured by the loss term) with its overall complexity (penalized by the regularization term). This approach prevents overfitting and improves the model's generalization to unseen data.

To solve Equation (1), the server requires access to all client data. However, in our proposed method, to preserve privacy, clients transmit fine-tuned generative models (instead of raw data) to the server. The server then utilizes these models to generate synthetic data for the training process.

3.1. Synthetic data generation for smart grids

Our proposed method addresses privacy concerns and communication overhead associated with data exchange in smart grids by leveraging GenAI models for synthetic data generation. By fine-tuning pre-trained GenAI models, we can generate synthetic data that closely resembles real-world data distributions, eliminating the need for directly exchanging raw data. Several GenAI models are well-suited for this task, offering flexibility and scalability in generating synthetic data tailored to specific smart grid applications. These GenAI models can categorized in: (i) Generative Adversarial Networks (GANs): GANs consist of a generator and a discriminator network trained simultaneously. The generator aims to create realistic data samples, while the discriminator distinguishes between real and synthetic data. (ii) Variational Autoencoders (VAEs): VAEs are probabilistic generative models that learn a latent representation of the data distribution. By sampling from this latent space, VAEs can generate synthetic data points. (iii) Foundation Models: These are large, pre-trained models trained on massive amounts of unlabeled data. They can be fine-tuned for downstream tasks, including synthetic data generation in the smart grid domain. Foundation models offer the potential to leverage knowledge from diverse data sources, potentially leading to richer and more informative synthetic data compared to other GenAI approaches.

This approach offers an efficient solution for federated learning in smart grids by addressing both privacy and efficiency concerns. The synthetic data we generate preserves the statistical properties of real data, allowing for effective model training without compromising individual privacy. Furthermore, this approach significantly reduces communication overhead. Instead of uploading large volumes of raw data, often measured in gigabytes per year from high-frequency sensors, consumers only need to transmit fine-tuned model to the server. This reduction in data size by several orders of magnitude benefits power system devices and resource-constrained IoT devices on the edge of the network.

Given an input data point x_t where t indicates timestamp, the GenAI model aims to create a synthetic data point, denoted by $\widetilde{x}_t$. This synthetic data point will resemble the original data point while preserving privacy. We aim to generate a dataset of such synthetic samples, $\{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_k\}$, where k is the desired number of synthetic samples. These synthetic samples will then be used instead of raw data exchange for effective model training. We denote the generator model as *Gen* with a set of parameters ω, aims to approximate the underlying data distribution p

(x). This can express as:

$$\widetilde{x}_t = Gen(t_{1:t-1};\omega) \quad \text{s.t.} \quad \widetilde{x}_t \approx x_t \tag{2}$$

Leveraging this formulation, each agent can fine-tune its local GenAI model (denoted by Gen_i with its specific parameters ω_i) based on its own data. This allows for the generation of synthetic data points that are tailored to the specific characteristics of each agent's data, even if the data distributions differ across agents (i.e., data heterogeneity).

3.2. Synthetic diffusion data

Our proposed method leverages GenAI models, specifically diffusion models, for high-quality synthetic data generation to solve Eq. (2). Diffusion models are a type of deep generative model trained to learn the underlying structure of real data distributions Sohl-Dickstein et al. [38]. This allows them to generate diverse and realistic synthetic samples compared to other methods like Generative Adversarial Networks (GANs) which can suffer from training instability. Additionally, diffusion models boast a fixed learning procedure and high-dimensional latent variables, leading to more accurate and more stable data generation quality compared to AutoEncoder or flow-based methods Croitoru et al. [39].

Once the GenAI models are trained locally, they are transmitted to the server. The server then leverages these models to generate synthetic data. This process utilizes models from all clients, resulting in diverse synthetic data that captures the statistical properties of the original information without revealing individual client details.

Diffusion models, distinguished from other latent variable models by deriving the posterior from a Markov chain, consist of two key processes: the diffusion process and the reverse process. The diffusion process gradually adds Gaussian noise to the original data, as formulated below:

$$p(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0) := \prod_{t=1}^{T} p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}), \tag{3}$$

Here, $p(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0)$ represents the diffusion process, where $\boldsymbol{x}_0$ is the original data sampled from the real data distribution $p(\boldsymbol{x}_0)$. The set $\{\boldsymbol{x}_t\}_{t=1}^{T}$ comprises latent variables, noisy data, sharing the same dimensionality as $\boldsymbol{x}_0$. t and T refer to the timestep and the total number of diffusion timesteps, respectively. Since the diffusion process follows a Markov chain structure, with $p(\boldsymbol{x}_t)$ depending only on $p(\boldsymbol{x}_{t-1})$, the joint distribution of noisy data $p(\boldsymbol{x}_{1:T})$ can be expressed as the product of successive diffusion steps. At each diffusion step, denoted by $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$, the process takes $\boldsymbol{x}_{t-1}$ as input and produces $\boldsymbol{x}_t$ by adding some noise. This operation is defined as follows:

$$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) := \mathcal{N}\left(\boldsymbol{x}_t; \sqrt{1-\eta_t}\,\boldsymbol{x}_{t-1}, \eta_t \boldsymbol{I}\right), \tag{4}$$

Here, $\eta_t \in (0, 1)$ represents the noise schedule. Ultimately, as T tends to infinity, $\boldsymbol{x}_T$ will conform to an isotropic Gaussian distribution.

Following Wang and Zhang [37], and by defining $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ as Gaussian noise, $\overline{\gamma}_t = \prod_{m=1}^{t} \gamma_m$, and $\gamma_t = 1 - \eta_t$, we can derive a closed form for the above process:

$$\begin{aligned} p(\boldsymbol{x}_t|\boldsymbol{x}_0) \quad &:= \mathcal{N}\left(\boldsymbol{x}_t; \sqrt{\overline{\gamma}_t}\,\boldsymbol{x}_0, (1-\overline{\gamma}_t)\boldsymbol{I}\right), \\ \boldsymbol{x}_t \quad &:= \sqrt{\overline{\gamma}_t}\boldsymbol{x}_0 + \sqrt{(1-\overline{\gamma}_t)}\boldsymbol{\epsilon}, \end{aligned} \tag{5}$$

The reverse process aims to denoise $\boldsymbol{x}_t$ to recover $\boldsymbol{x}0$, enabling the recreation of data samples from Gaussian noise. Due to the difficulty of directly and explicitly formulating the reverse step as in Eq. (4), we approximate it using a neural network model. Similar to the diffusion step, the reverse process is represented as $D(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t;\omega)$, where $D(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t;\omega)$ denotes the reverse step implemented by the neural network. This step takes $\boldsymbol{x}_t$ and t as inputs, estimating mean and variance of the Gaussian distribution.

Given that the reverse process also conforms to a Markov chain, its formulation can be derived from the reverse step as follows:

$$D(\boldsymbol{x}_{0:T};\omega) := D(\boldsymbol{x}_T) \prod_{t=1}^{T} D(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t;\omega), \tag{6}$$

Accurate noise estimation at each step is essential for effective denoising at each timestep. The objective of denoising process of the diffusion models can be formulated as:

$$\min(\mathcal{L}(D(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t;\omega), p(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t,\boldsymbol{x}_0))), 1 \le t \le T, \tag{7}$$

where $p(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t,\boldsymbol{x}_0)$ represents the posterior conditional probability of the diffusion process, reflecting the ground truth of added noises. The objective is to minimize $\mathcal{L}$, typically Kullback-Leibler divergence in diffusion models.

3.3. Dataset aggregation and task model training

After receiving the fine-tuned GenAI models Gen_i from all available clients, the server leverages them to create synthetic data for training the task model. We denote the set of synthetic datasets generated by these N models as $\mathcal{D} = \mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_N$. The server aggregates these datasets to form a combined dataset, denoted by $\mathcal{D}_{\text{combined}}$. To minimize the objective function defined in Equation (1), the server iteratively utilizes $\mathcal{D}_{\text{combined}}$ and solves the function for K iterations.

For solving this optimization problem, gradient descent is a commonly used algorithm in machine learning models. The update rule for parameter θ in the gradient descent algorithm can be expressed as:

$$\theta_{k+1} = \theta_k - \alpha \nabla(\mathcal{L}(\mathbf{x}, y;\theta_k) + \lambda\Omega(\mathbf{x}, y;\theta_k)) \tag{8}$$

where α is the step size, and ∇ denotes the gradient operator. This update rule iteratively adjusts the parameters θ in the direction opposite to the gradient of the objective function, aims to convergence towards the optimal solution. Adjusting the step size α appropriately is crucial for ensuring convergence and efficiency of the optimization process.

Training the model on the aggregated dataset $\mathcal{D}_{\text{combined}}$ allows for effective learning while preserving the privacy of individual clients' data. Additionally, incorporating data augmentation techniques further enhances the model's ability to generalize and adapt to diverse scenarios.

4. Experiments and discussion

The convergence of federated learning (FL) with GenAI models opens exciting possibilities for enhancing smart grid functionalities. By leveraging synthetic data generated by GenAI models, FL can overcome privacy concerns and data scarcity limitations, paving the way for advancements in various smart grid applications. This includes real-time anomaly detection in sensor data for proactive maintenance, optimizing demand response programs to reduce peak loads, and facilitating the seamless integration of renewable energy sources. In this paper, we will focus on the critical application of load forecasting. We aim to explore how GenAI-powered synthetic data can empower FL models to predict future energy demand with greater accuracy and efficiency, ultimately contributing to a more reliable and sustainable smart grid.

4.1. Experiment setting

Data Preparation. Our experiments leverage the Pecan Street dataset Parson et al. [40], a well-established benchmark for residential energy consumption data. This dataset provides high-resolution (e.g., 1-min interval) measurements from a large number of households, offering a realistic representation of real-world load profiles. For synthetic data generation, we employ a U-Net model Ronneberger et al. [41] pre-trained on a massive dataset. U-Nets are a type of convolutional neural network architecture specifically designed for image segmentation tasks. However, recent research has demonstrated their effectiveness in time series forecasting tasks as well Yin et al. [42]; Madhusudhanan et al. [43]. While other foundation models like GPT-3 can be explored in future work, U-Nets offer a good balance between performance and computational efficiency for this application.

For the forecasting task, we utilize a LSTM network Graves and Graves [44] as the core machine learning model running on the server side. LSTMs are well-suited for tasks involving sequential data like load forecasting, as they can effectively capture temporal dependencies within the data.

Performance Metrics. To effectively assess the quality of the proposed method, particularly for time series data, we evaluate the performance of the methods in terms of Root Mean Squared Error (RMSE), measuring the disparity between individual synthetic and real samples. Accuracy is also considered, with predictions deemed accurate if the error is less than 5 %. Additionally, privacy preservation and communication efficiency are evaluated using various benchmarks.

Benchmark. To demonstrate the effectiveness of our proposed method and compare its performance against leading approaches, we select the following benchmarks based on current state-of-the-art methods.

1. **Individual Training:** Where we compare the average performance of clients. Each client trains its model independently without collaborating with others. We then average the performance across all clients.
2. **Centralized Training:** In this scenario, all clients send their data to the server, and the model is trained centrally.
3. **Federated Learning:** In a Federated Learning setting, clients iteratively train the model with the assistance of the server. In each iteration, they send their updated model to the server for aggregation.
4. **GAN:** A GAN-based framework proposed in El Kababji and Srikantha [45] for load pattern synthesis.
5. **Federated-WDCGAN:** A Wasserstein deep convolutional conditional GAN proposed in Chen et al. [46] to generate energy consumption data.
6. **Conditional DM:** This benchmark represents a Conditional Diffusion Model introduced by Wang and Zhang [37], that progressively learns to denoise a latent representation of the target data distribution.

4.2. Next-day energy demand prediction

In this experiment, we evaluate the effectiveness of our proposed method for next-day energy demand prediction at the individual household level. The server-side LSTM model is trained for 300 epochs using the synthetic data generated from the data of 10 homes. Fig. 4 illustrates the performance of the LSTM model with one hidden layer in predicting the next-day energy demand for two sample households. The plot compares the predicted demand (in kWh) with the actual observed consumption. As evident from the figure, the proposed method achieves good accuracy in predicting the next-day energy demand for these particular households.

4.3. Comparison to baselines

Table 1 compares the accuracy of the proposed method with six existing approaches, also evaluating their Root Mean Squared Error (RMSE). The table shows that our method achieves better performance on RMSE and accuracy compared to these benchmarks. This demonstrates that synthetic data generated using our method effectively captures the underlying patterns in real-world energy consumption data while preserving privacy. Speccifically our method offers a significant privacy advantage compared to the centralized approach. Unlike the centralized approach, which requires sharing raw client data, our method protects privacy by keeping all data on the client side. This allows for accurate predictions without compromising sensitive information.

4.4. Communication efficiency

Our proposed method offers significant advantages in terms of communication efficiency and privacy protection compared to traditional approaches. By leveraging synthetic data generated locally at each household, the communication burden is reduced by 58 % compared to a scenario where all raw data is uploaded to the server for centralized training. This translates to a substantial decrease in network traffic and associated costs, especially for resource-constrained devices on the edge of the smart grid.

Furthermore, our approach enhances privacy by eliminating the need for households to share their raw energy consumption data. Instead, only anonymized synthetic data is transmitted to the server, significantly reducing the risk of sensitive information leakage. This prevents the server or any potential adversaries from gleaning insights into individual consumption patterns that could be used for malicious purposes, such as inferring occupancy or daily routines.

Table 1
Comparison of performance and privacy preservation in different benchmarks for predicted versus observed next-day energy demand.

Method	Accuracy (%)	RMSE	Privacy
Individual Training	83	2.4529	Yes
Centralized Training	94	0.8931	No
Federated Learning	92	1.2360	Partial
GAN	90	1.5584	Yes
Federated-WDCGAN	91	1.1694	Yes
Conditional DM	93	0.9806	Yes
Proposed Method	**94**	**0.9026**	**Yes**

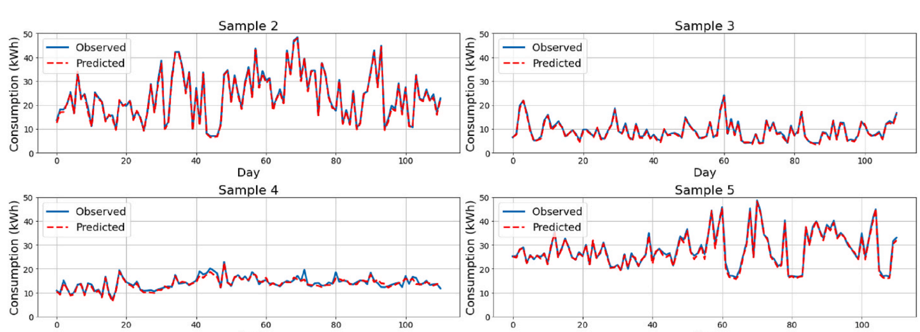

Fig. 4. Comparison of Predicted and Observed Next-Day Energy Demand for Sample Households using the Proposed Method.

5. Conclusion

Emerging communication challenges threaten to impede the growth of smart grids. The ever-increasing volume of data transmitted from sensors and devices strains existing infrastructure, leading to bottlenecks and inefficiencies. Generative Artificial Intelligence presents a compelling solution by enabling distributed learning at the grid edge. This approach empowers intelligent agents, such as smart meters and distributed generators, to process and analyze data locally. By leveraging pre-trained GenAI models, these agents can generate synthetic data that accurately reflects real-world grid conditions. This synthetic data can then be used to train local machine learning models for critical tasks like anomaly detection, energy use optimization, and demand forecasting. GenAI adoption within smart grids unlocks a multitude of benefits. It fosters a significant reduction in communication burden by alleviating the need for constant communication with a central server. This not only frees up valuable bandwidth for essential grid operations but also enhances overall system efficiency. Furthermore, GenAI safeguards privacy by eliminating the transmission of raw energy consumption data. Instead, anonymized synthetic data, preserving the statistical properties of real data without compromising individual details, is transmitted. This approach mitigates the risk of cyberattacks and protects consumer privacy. In conclusion, GenAI has the potential to revolutionize smart grid communication, paving the way for a more secure, efficient, and scalable energy infrastructure that underpins a sustainable future.

CRediT authorship contribution statement

Seyed Mahmoud Sajjadi Mohammadabadi: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mahmoudreza Entezami:** Writing – review & editing, Resources, Funding acquisition. **Aidin Karimi Moghaddam:** Writing – review & editing, Resources, Funding acquisition. **Mansour Orangian:** Writing – review & editing, Resources, Funding acquisition. **Shayan Nejadshamsi:** Writing – review & editing, Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Wei, A. Chammam, J. Feng, A. Alshammari, K. Tehranian, N. Innab, W. Deebani, M. Shutaywi, Power system monitoring for electrical disturbances in wide network using machine learning, Sustain. Comput.: Inform. Syst. 42 (2024) 100959.
[2] A. Abdali, V. Monjezi, Machine learning-based flexible link robot control, Comput. Method. Eng. Sci. 1 (2023) 12–17.
[3] M. Maleki, Forecasting Tech Sector Market Downturns Based on Macroeconomic Indicators, 2024 arXiv preprint arXiv:2404.10208.
[4] V. Monjezi, A. Trivedi, G. Tan, S. Tizpaz-Niari, Information-theoretic testing and debugging of fairness defects in deep neural networks, in: 2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE), IEEE, 2023, pp. 1571–1582.
[5] N. Darabi, S. Tayebati, S. Ravi, T. Tulabandhula, A.R. Trivedi, et al., Starnet: Sensor Trustworthiness and Anomaly Recognition via Approximated Likelihood Regret for Robust Edge Autonomy, 2023 arXiv preprint arXiv:2309.11006.
[6] M. Tavasoli, E. Lee, Y. Mousavi, H. Barahouei Pasandi, A. Fekih, Wipe: A novel web-based intelligent packaging evaluation via machine learning and association mining, IEEE Access (2024).
[7] M. EskandariNasab, Z. Raeisi, R.A. Lashaki, H. Najafi, A gru–cnn model for auditory attention detection using microstate and recurrence quantification analysis, Sci. Rep. 14 (2024) 8861.
[8] S. Wang, X. Huang, P. Liu, M. Zhang, F. Biljecki, T. Hu, X. Fu, L. Liu, X. Liu, R. Wang, et al., Mapping the landscape and roadmap of geospatial artificial intelligence (geoai) in quantitative human geography: an extensive systematic review, Int. J. Appl. Earth Obs. Geoinf. 128 (2024) 103734.
[9] S.S. Ali, B.J. Choi, State-of-the-art artificial intelligence techniques for distributed smart grids: a review, Electronics 9 (2020) 1030.
[10] X. Zhao, Y. Guo, X. Guo, H. Li, Artificial intelligence applications and prospects for the smart grid, in: 2023 Panda Forum on Power and Energy (PandaFPE), IEEE, 2023, pp. 1844–1848.
[11] E. Jandaghi, D.L. Stein, A. Hoburg, M. Zhou, C. Yuan, Composite Distributed Learning and Synchronization of Nonlinear Multi-Agent Systems with Complete Uncertain Dynamics, 2024 arXiv preprint arXiv:2403.00987.
[12] M. Borhani, M.R. Kalbasi, M. Psenicka, B. Falahatkar, A.D. Crawford, Identification and isolation of sprmatogonial stem cells from beluga (huso huso) testicular tissue, Aquatic Physiol. Biotechnol. 10 (2022) 1–24, https://doi.org/10.22124/japb.2021.20253.1430.
[13] S. Sakib, K.M. Hasib, I.K. Tasawar, A.K. Tanzeem, M.F. Arefin, S. Islam, M.S. Alam, A data-driven hybrid optimization based deep network model for short-term residential load forecasting, in: 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), IEEE, 2021, pp. 187–193.
[14] S.R. Herrera, V. Monjezi, V. Kreinovich, A. Trivedi, S. Tizpaz-Niari, Predicting Fairness of Ml Software Configuration, 2024 arXiv preprint arXiv:2404.19100.
[15] K.M. Hasib, M.S. Iqbal, F.M. Shah, J.A. Mahmud, M.H. Popel, M.I.H. Showrov, S. Ahmed, O. Rahman, A Survey of Methods for Managing the Classification and Solution of Data Imbalance Problem, 2020 arXiv preprint arXiv:2012.11870.
[16] K.P. Grammatikakis, I. Koufos, N. Kolokotronis, A collaborative intelligent intrusion response framework for smart electrical power and energy systems, in: Proceedings of the 17th International Conference on Availability, Reliability and Security, 2022, pp. 1–10.
[17] T. Kotsiopoulos, P. Sarigiannidis, D. Ioannidis, D. Tzovaras, Machine learning and deep learning in smart manufacturing: the smart grid paradigm, Comput. Sci. Rev. 40 (2021) 100341.
[18] A.Y.B. Ahmad, P. William, D. Uike, A. Murgai, K. Bajaj, A. Deepak, A. Shrivastava, Framework for sustainable energy management using smart grid panels integrated with machine learning and iot based approach, Int. J. Intell. Syst. Appl. Eng. 12 (2024) 581–590.
[19] A. Sulaiman, B. Nagu, G. Kaur, P. Karuppaiah, H. Alshahrani, M.S.A. Reshan, S. AlYami, A. Shaikh, Artificial intelligence-based secured power grid protocol for smart city, Sensors 23 (2023) 8016.
[20] B. Schieber, B. Samineni, S. Vahidi, Interweaving real-time jobs with energy harvesting to maximize throughput, in: International Conference and Workshops on Algorithms and Computation, Springer, 2023, pp. 305–316.
[21] S.H.H. Dolatabadi, A. Soleimani, A. Ebtia, M. Shafie-khah, T.H. Bhuiyan, Enhancing voltage profile in islanded microgrids through hierarchical control strategies, Elec. Power Syst. Res. 231 (2024) 110270.
[22] S.F. Razmi, L. Torki, M.J. Razmi, E.M. Dowlatabadi, The indirect effects of oil price on consumption through assets, Int. J. Energy Econ. Pol. 12 (2022) 236–242.
[23] H. Torkaman, M. Shadaei, N. Deyhimi, H. Torkaman, Multi-objective method to reconfigure radial power distribution system with dgs considering transformer tap-changing, in: 2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), IEEE, 2021, pp. 1–4.
[24] S. Zobaed, F. Rabby, I. Hossain, E. Hossain, S. Hasan, A. Karim, K. Md Hasib, Deepfakes: detecting forged and synthetic media content using machine learning. Artificial Intelligence in Cyber Security: Impact and Implications: Security Challenges, Technical and Ethical Issues, Forensic Investigative Challenges, 2021, pp. 177–201.
[25] L. Tightiz, H. Yang, A comprehensive review on iot protocols' features in smart grid communication, Energies 13 (2020) 2762.
[26] V. Himthani, V. Prakash, Generative adversarial network-based deep learning technique for smart grid data security, in: Artificial Intelligence and Machine Learning in Smart City Planning, Elsevier, 2023, pp. 303–315.
[27] S.M.S. Mohammadabadi, Y. Liu, A. Canafe, L. Yang, Towards distributed learning of pmu data: a federated learning based event classification approach, in: 2023 IEEE Power & Energy Society General Meeting (PESGM), IEEE, 2023, pp. 1–5.
[28] S.M.S. Mohammadabadi, S. Zawad, F. Yan, L. Yang, Speed up Federated Learning in Heterogeneous Environment: A Dynamic Tiering Approach, 2023 arXiv preprint arXiv:2312.05642.
[29] S.M.S. Mohammadabadi, L. Yang, F. Yan, J. Zhang, Communication-efficient Training Workload Balancing for Decentralized Multi-Agent Learning, 2024 arXiv preprint arXiv:2405.00839.
[30] S. Jahan, M.R. Islam, K.M. Hasib, U. Naseem, M.S. Islam, Active learning with an adaptive classifier for inaccessible big data analysis, in: 2021 International Joint Conference on Neural Networks (IJCNN), IEEE, 2021, pp. 1–7.
[31] V. Ashrafimoghari, N. Gürkan, J.W. Suchow, Evaluating Large Language Models on the Gmat: Implications for the Future of Business Education, 2024 arXiv preprint arXiv:2401.02985.
[32] F. Li, D. Lin, T. Yu, Improved generative adversarial network-based super resolution reconstruction for low-frequency measurement of smart grid, IEEE Access 8 (2020) 85257–85270.
[33] T. Yeom, C. Gu, M. Lee, Dudgan: improving class-conditional gans via dual-diffusion, IEEE Access (2024).
[34] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolution image synthesis with latent diffusion models, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10684–10695.
[35] S. Santos, T. Breaux, T. Norton, S. Haghighi, S. Ghanavati, Requirements Satisfiability with In-Context Learning, 2024 arXiv preprint arXiv:2404.12576.

[36] S.M. Sajjadi Mohammadabadi, Advancing Communication Efficiency in Electric Vehicle Systems: A Survey of Generative Ai and Distributed Machine Learning Strategies, 2024, https://doi.org/10.2139/ssrn.4791891. SSRN 4791891.
[37] Z. Wang, H. Zhang, Customized load profiles synthesis for electricity customers based on conditional diffusion models, IEEE Trans. Smart Grid (2024).
[38] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, S. Ganguli, Deep unsupervised learning using nonequilibrium thermodynamics, in: International Conference on Machine Learning, PMLR, 2015, pp. 2256–2265.
[39] F.A. Croitoru, V. Hondru, R.T. Ionescu, M. Shah, Diffusion models in vision: a survey, IEEE Trans. Pattern Anal. Mach. Intell. (2023).
[40] O. Parson, G. Fisher, A. Hersey, N. Batra, J. Kelly, A. Singh, W. Knottenbelt, A. Rogers, Dataport and nilmtk: a building data set designed for non-intrusive load monitoring, in: 2015 Ieee Global Conference on Signal and Information Processing (Globalsip), IEEE, 2015, pp. 210–214.
[41] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.
[42] L. Yin, L. Wang, T. Li, S. Lu, J. Tian, Z. Yin, X. Li, W. Zheng, U-net-lstm: time series-enhanced lake boundary prediction model, Land 12 (2023) 1859.
[43] K. Madhusudhanan, J. Burchert, N. Duong-Trung, S. Born, L. Schmidt-Thieme, Yformer: U-Net Inspired Transformer Architecture for Far Horizon Time Series Forecasting, 2021 arXiv preprint arXiv:2110.08255.
[44] A. Graves, A. Graves, Long short-term memory, Superv. Sequen. Labell. Recurrent Neural Ntworks (2012) 37–45.
[45] S. El Kababji, P. Srikantha, A data-driven approach for generating synthetic load patterns and usage habits, IEEE Trans. Smart Grid 11 (2020) 4984–4995.
[46] Z. Chen, J. Li, L. Cheng, X. Liu, Federated-wdcgan: a federated smart meter data sharing framework for privacy preservation, Appl. Energy 334 (2023) 120711.