Differential Privacy-Aware Generative Adversarial Network-Assisted Resource Scheduling for Green Multi-Mode Power IoT

Sunxuan Zhang, Jiapeng Xue, Jiayi Liu, Zhenyu Zhou, *Senior Member, IEEE,* Xiaomei Chen, and Shahid Mumtaz, *Senior Member, IEEE*

***Abstract*—The low-carbon and efficient operation of smart parks requires high-precision and real-time energy management model training. Multi-mode power internet of things (PIoT) consisting of open radio access networks (O-RAN) and power line communications (PLC) can effectively improve the model training performance. However, the negative effects of network threats, such as model inversion attacks, cannot be neglected. To solve this problem, we propose a diFFerential pRivacy-aware gEnErative aDversarial netwOrk-assisted resource scheduling algorithM (FREEDOM). Firstly, we integrate a differential privacy mechanism with the energy management model training process and the related system model. Then, a joint resource scheduling optimization problem is constructed, the goal of which is to minimize the weighted sum of the global loss function, total energy consumption, and differential privacy cost under the long-term differential privacy constraint. The original problem is converted based on virtual queue theory and addressed by the FREEDOM. FREEDOM leverages a deep Q-learning network (DQN) to learn the resource scheduling strategy via differential privacy awareness. It improves optimization and convergence performances with the assistance of generative adversarial network (GAN). Simulation results show that FREEDOM can achieve excellent performances of global loss function, total energy consumption, differential privacy cost, and privacy preservation.**

***Index Terms*—Power internet of things (PIoT), green multi-mode network, energy management, model inversion attacks, resource scheduling, generative adversarial network (GAN), differential privacy awareness**

I. Introduction

Smart parks serve as the fundamental units of the new-generation power systems, integrating large-scale electrical equipment including distributed photovoltaics, load demands, and energy storage units [1]. These parks minimize carbon emissions and achieve energy demand-supply balance through advanced and real-time green energy management [2], [3]. The key to energy management is to unearth and configure the intricate relationships among data such as renewable energy production, load demands, and energy storage status based on federated learning (FL)-enabled model training, thereby learning the optimal energy management policy [4]. Power internet of things (PIoT) exemplifies a particular utilization of IoT technology in power systems, aimed at realizing intelligent acquisition, transmission, and control of data within power systems. By deploying PIoT devices within smart parks, it is possible to acquire real-time operational data from the power system, thereby facilitating the optimization of scheduling strategies. Therefore, the process begins by deploying large-scale PIoT devices on distributed electrical equipment to continuously collect data for model training. These data are then fed into local models, which are uploaded to a controller for global aggregation [5]. As a critical interface between devices and core networks, open radio access networks (O-RAN) provide real-time, efficient, and flexible communications for model interactions between PIoT devices and controllers in smart parks [6], [7]. It facilitates interaction across diverse devices and controllers through standardized interfaces, enabling data-driven optimization and automated control. On the other hand, power line communication (PLC) technology, as a cost-effective and convenient wired solution, compensates for the interference and attenuation caused by building obstructions of O-RAN, enhancing the coverage and scalability of communication networks in smart parks [8], [9]. Therefore, the integration of O-RAN with PLC forms a multi-mode PIoT to further improve model training performance.

In order to further enhance the precision of model training while ensuring the low-carbon operation of smart parks, network resources, e.g., training batch size, require to be scheduled in a flexible manner [10]. Unlike resource scheduling in other systems, the energy management of smart parks in multi-mode PIoT imposes stringent security demands. Due to the involvement of diverse devices and complex network components in O-RAN, malicious software and hardware could be introduced, elevating energy management risks. Furthermore, the open interfaces of O-RAN mean increased exposure to malicious attacks. Conversely, power lines are not designed specifically for data transmission, making the data conveyed through PLC susceptible to interception by attackers. The potential negative influence of malicious attackers on resource scheduling is a factor that warrants careful consideration. Although FL isolates global aggregation from raw data uploading to solve privacy disclosure, network threats such as model inversion attacks, can still recover private data

This work was supported by the Science and Technology Project of State Grid Corporation of China under Grant Number 52094021N010 (5400-202199534A-0-5-ZN). *(Corresponding author: Zhenyu Zhou.)*

Sunxuan Zhang, Jiapeng Xue, Jiayi Liu, Zhenyu Zhou, and Xiaomei Chen are with State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China. (E-mail: {sunxuan_zhang, jiapeng_xue, 120211020515, zhenyu_zhou, chxm}@ncepu.edu.cn). *Corresponding author: Zhenyu Zhou.*

Shahid Mumtaz is with Department of Applied Informatics, Silesian University of Technology Akademicka, Gliwice 16 44-100, Poland, and also with Departement of Engineering, Nottingham Trent University, Nottingham NG1 4BU, U.K. (E-mail: Shahid.mumtaz@ntu.ac.uk).

by inferring model parameters, seriously affecting the low-carbon operation of smart parks [11], [12]. For instance, attackers could infer the energy management patterns, peak power consumption periods, and long-term output regulations of distributed electrical equipment in smart parks. This inferred knowledge could enable targeted attacks on specific power grid service components, leading to security issues such as grid infrastructure damage and widespread power outages.

Differential privacy technology can prevent privacy leakage through adding suitable artificial noise to model training. In [13], Yan *et al.* designed a differential privacy-driven privacy-preservation asynchronous FL mechanism to improve the security and robustness of model aggregation. A coordinated differential privacy-enabled FL method was designed in [14] to increase the security level of client training. In [15], Yang *et al.* investigated a differential privacy policy based on the randomized response mechanism to effectively improve the security of IoT data collection. In [16], Iqbal *et al.* proposed an optimal differential privacy method to guarantee the privacy of personal data elements while permitting the retrieval of insightful information. However, adding substantial differential privacy noise may negatively affect the precision of model training. Incorporating less noise increases the risk of privacy leaks, which leads to extra compensation costs for privacy loss [17]. Therefore, the optimization of differential privacy scheduling is also crucial. The coordinated resource scheduling optimization for multi-mode PIoT under model inversion attacks faces several technical challenges, which are summarized in the following.

- *Differentiated Performance Metrics Guarantee:* Low-carbon energy management requires resource scheduling optimization, which calls for distinct performance metric demands. The optimization of training batch size scheduling can enhance the overall model accuracy to some extent, but it also increases the energy consumption generated during model training. Incorporating substantial differential privacy noise can ensure a degree of privacy preservation, albeit at the expense of achievable model accuracy. Conversely, minimal noise preserves model accuracy but may result in increased privacy breaches and higher privacy loss compensation.
- *The Coupling between Resource Scheduling Optimization and Differential Privacy Constraint:* Privacy loss is an accumulative process where minimal privacy breaches in the short term may result in excessive long-term privacy loss [18]. It is crucial to establish a long-term differential privacy constraint. However, the short-term resource scheduling strategy may not align with the long-term constraint. In other words, resource scheduling need to be optimized without considering future differential privacy information.
- *Poor and Slow Convergence of Resource Scheduling Optimization:* Traditional optimization methods such as stochastic programming and robust optimization are limited by accurate optimization modeling, making it challenging to guide resource scheduling optimization under uncertain information. Resource scheduling approaches based on traditional machine learning, while able to optimize resources with uncertain information, still face issues with slow convergence and unstable optimization performance, particularly when addressing resource scheduling problems with a large optimization space.

Deep reinforcement learning (DRL) is extensively used for optimizing resource scheduling with uncertain information. In [19], Kwon *et al.* designed a DRL-empowered coordinated cell sharing and resource scheduling mechanism for FL computation in the internet of underwater things under uncertain cell states and transmission gain, whose goal is to maximize the transmission rate. However, it neglects the joint guarantee of differentiated performance metrics. In [20], Zhao *et al.* proposed a resource allocation and device management method based on DRL with incompleted network information in industrial IoT, the goal of which is to jointly improve the delay, energy cost, and model precision. In [21], Zheng *et al.* investigated a resource scheduling optimization approach for the edge IoT based on DRL, which realizes the tradeoff between FL accuracy and energy cost. Although these works consider multi-objective optimization, the adverse impact on model training caused by malicious attacks is ignored. In [22], Okegbile *et al.* integrated DRL, differential privacy, and blockchain to design a resource scheduling mechanism for human digital twin model updating. The goal is to enhance precision and privacy while reducing communication overheads in the situation of unknown communication environments of base stations. However, this work does not consider the optimization of differential privacy budget, and cannot achieve fast convergence speed and stable optimization when facing a large optimization space.

Motivated by these challenges, we first establish the energy management model training and differential privacy model in multi-mode PIoT. Then, we formulate the coordinated resource scheduling optimization problem to minimize the weighted sum of the global loss function, total energy consumption of all devices, and differential privacy cost under a long-term differential privacy constraint. Next, by leveraging a virtual queue, the original problem is transformed into a sequence of optimization problems in short term. Finally, we design a diFferential pRivacy-aware gEnErative aDversarial netwOrk-assisted resource scheduling algorithM (FREEDOM) to address the converted problem. The main contributions are summarized below.

- *Joint Guarantee of High Precision, Low Energy Consumption, and Privacy Preservation:* FREEDOM realizes the joint guarantee of model precision, energy consumption, and privacy preservation through the joint optimization of training batch size and privacy budget scheduling. The balance among the global loss function, total energy consumption, and differential privacy cost of PIoT devices is realized by dynamically adjusting the weights of differentiated performance metrics and learning the optimal resource scheduling strategies.
- *Differential Privacy Awareness:* To tackle the long-term differential privacy constraint, we transform it into the stability guarantee of a virtual queue. This is done by

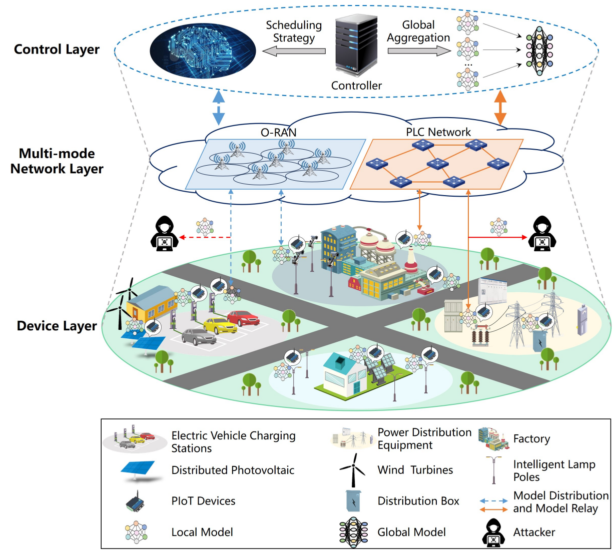

Fig. 1. The resource scheduling framework for green multi-mode PIoT.

redefining the optimization goal to include the virtual queue deficit fluctuation. Differential privacy awareness is realized by actively adjusting the resource scheduling strategy based on the dynamic changes of differential privacy queue deficit.

- *Intelligent Resource Scheduling with the Assistance of Generative Adversarial Network (GAN):* FREEDOM combines the deep Q-learning network (DQN) with GAN to enhance the optimization and convergence performance. The DQN-based generator provides a model-free mechanism to continuously adapt to the nonlinear relationship between the state and action space, learning the resource scheduling strategy despite uncertain information. The discriminator guides the training of the DQN-based generator by differentiating the expert policy from the generating strategy, supporting more accurate optimization and faster convergence speed.

The remainder of the structure of the paper is below. Section II introduces the system model. The problem formulation and transformation are introduced in Section III. FREEDOM is proposed in Section IV. Section V elaborates on the performance analysis and simulations. Section VI concludes this paper.

II. System Model

The proposed resource scheduling framework for green multi-mode PIoT is shown in Fig. 1. The goal is to train an energy management model based on FL. The considered framework includes a device layer, a multi-mode network layer, and a control layer [23]. In the device layer, PIoT devices are deployed on electrical equipment, e.g., distributed photovoltaic, wind turbines, and electric vehicle charging stations, to carry out local model training. Denote the number of devices as I, the set of which is $\mathcal{M} = \{m_1, \cdots, m_i, \cdots, m_I\}$. The multi-mode network layer contains O-RAN and PLC networks to provide local model uploading for devices. During this process, there is a risk of malicious attackers initiating model inversion attacks. They try to recover sensitive feature information of private data associated with energy management from the model. The control layer optimizes resource scheduling strategies and coordinates the global aggregation of the energy management model for devices through a controller [24], [25].

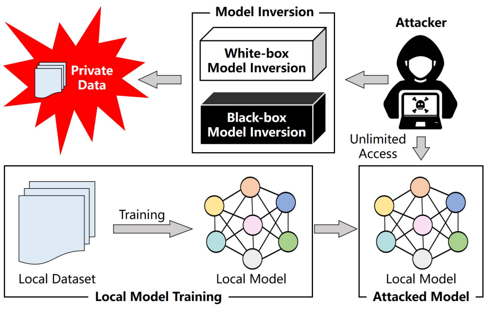

Fig. 2. The principle of model inversion attacks.

The total period of resource scheduling optimization is segmented into T iterations, collectively represented by the set $\mathcal{T} = \{1, \cdots, t, \cdots, T\}$. At the start of each iteration, the controller devises a resource scheduling strategy of training batch size and privacy budget for each device. Then, the controller utilizes O-RAN and PLC to distribute the resource scheduling strategy and the latest global model to each device. Subsequently, devices engage in local training of their models based on the scheduled batch size and preserve the privacy of model based on the scheduled privacy budget. Next, devices relay their local models back to the controller via multi-mode networks. Upon reception, the controller aggregates these local models to update the global model.

A. Model Inversion Attacks

The principle of model inversion attacks is shown in Fig. 2. Malicious attackers aim to recover the sensitive features of private energy management data from the uploaded models. Specifically, assume the attacked model is trained based on the data distribution $(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X}$ is the input dataset and $\mathcal{Y}$ is the corresponding target output dataset. The adversarial objective of malicious attackers is to obtain data feature $Z\left(\mathcal{X} | f(\mathcal{X}) = y\right)$, where y is the specific data label [26].

Model inversion attacks can be categorized into two kinds, namely, white box and black box. White-box attacks have unlimited access to the attacked model and its parameters, while black-box attacks require limited observations and interactions with data inputs and outputs to construct a similar model by inferring the attacked model behavior [27]. Obviously, defending against white-box attacks is more challenging. Therefore, this paper assumes that malicious attackers have unlimited access to the attacked model.

B. Local Model Training

Each device downloads the global model from the controller to update its local model, i.e., $\boldsymbol{\omega}_i(t-1) = \boldsymbol{\omega}_g\left(t-1\right)$, where $\boldsymbol{\omega}_i(t-1)$ is the local parameter of device m_i in the $(t-1)$-th iteration and $\boldsymbol{\omega}_g\left(t-1\right)$ is the global model parameter. Then, m_i adopts partial data samples from its dataset $\mathcal{D}_i$ to train the local model. Define $\alpha_i(t)$ as the batch size of m_i for local model training and $\alpha_i(t) \in \mathcal{A}_i(t) = \{1, 2, \cdots, |\mathcal{D}_i|\}$, where $\mathcal{A}_i(t)$ is the set of training batch size and $|\mathcal{D}_i|$ is the size of local dataset $\mathcal{D}_i$. The loss function is leveraged to measure the bias between the real output and the target output of a model. Define x_l and y_l as the input and target output of the l-th sample. The loss function of m_i is represented as

$$F_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t) = \frac{1}{\alpha_i(t)} \sum_{l=1}^{\alpha_i(t)} f(\boldsymbol{\omega}_i(t-1), x_l, y_l, t), \tag{1}$$

where $f(\boldsymbol{\omega}_i(t-1), x_l, y_l, t)$ is the single-sample loss function in the t-th iteration, representing the deviation of the real output from the target output. The local model parameter of m_i is updated by adopting the loss function based on the gradient descent approach, i.e.,

$$\boldsymbol{\omega}_i(t) = \boldsymbol{\omega}_i(t-1) - \lambda \nabla F_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t), \tag{2}$$

where λ is the updating step.

Define ζ_i as the CPU cycles needed for executing a single data sample and $\xi_i(t)$ as the CPU frequency options accessible for local model training. The energy consumption of local model training is calculated as

$$E_i^L(t) = e_i \zeta_i \alpha_i(t) \xi_i^2(t), \tag{3}$$

where e_i is the energy consumption coefficient.

C. Privacy Preservation Mechanism

To preserve the private data of the energy management model from model inversion attacks, the differential privacy framework is adopted. Malicious attackers are defended against by the differential privacy framework via adding a suitable level of randomness into the protected model [28].

Definition 1. *A randomized mechanism Ψ with the output range* Range(Ψ) *has the differential privacy if for any two datasets $\mathcal{D}_i$ and $\mathcal{D}'_i$ differing on at most one sample, and for any output datasets $\mathcal{O} \subseteq$* Range(Ψ)*, we have the following relationship*

$$\Pr\{\Psi(\mathcal{D}_i) \in \mathcal{O}\} \leq \exp(\epsilon_i(t)) \Pr\{\Psi(\mathcal{D}'_i) \in \mathcal{O}\}, \tag{4}$$

where $\Pr\{.\}$ *is the probability function.* $\epsilon_i(t)$ *is the privacy budget of* m_i.

In **Definition 1**, $\epsilon_i(t)$ is used to measure the privacy loss, representing the identifiable limit of all outputs between $\mathcal{D}_i$ and $\mathcal{D}'_i$ [29]. Since the distribution of differential privacy budget introduced in [18] meets a fixed gradient, we define $\epsilon_i(t) \in \mathcal{E}(t) = \{\epsilon_1, \cdots, \epsilon_k, \cdots, \epsilon_K\}$, where $\mathcal{E}(t)$ is the set of privacy budgets, ϵ_1 and ϵ_K are the minimum and maximum privacy budgets, and $\epsilon_k = \frac{(k-1)(\epsilon_K - \epsilon_1)}{K-1}$ is the k-th privacy budget. The randomized mechanism Ψ represents the strategy of introducing randomness into the model training process to preserve privacy. Thus, we adopt the Laplace mechanism on the local models of each device by adding Laplace noise to perturb the parameters. In addition to Laplace noise, there are several other types of noise commonly used for differential privacy, such as exponential and Gaussian mechanisms. Compared to other noises, Laplace noise allows for lower tail sensitivity and helps maintain the overall quality of the model [30]–[32]. The updating of the local model parameter of m_i is rewritten as

$$\begin{aligned} \tilde{\boldsymbol{\omega}}_i(t) &= \boldsymbol{\omega}_i(t-1) \\ &\quad - \lambda \nabla \left[F_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t) + \text{Lap}(\frac{\delta_i^F(t)}{\epsilon_i(t)}) \right], \end{aligned} \tag{5}$$

where $\tilde{\boldsymbol{\omega}}_i(t)$ is the protected model parameter of m_i. Lap$(\delta_i^F(t)/\epsilon_i(t))$ is the Laplace noise drawn from a Laplace distribution with mean 0 and scale $\delta_i^F(t)/\epsilon_i(t)$. The probability density function of the added Laplace noise is $f(x) = \left(\epsilon_i(t)/2\delta_i^F(t)\right) \exp\left(\left(-\epsilon_i(t)/\delta_i^F(t)\right)|x|\right)$. $\delta_i^F(t)$ is the sensitivity of $F_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t)$.

Definition 2. *Define $F'_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t)$ as the loss function trained by $\mathcal{D}'_i$. For any two datasets $\mathcal{D}_i$ and $\mathcal{D}'_i$ differing on at most one sample, the sensitivity of $F_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t)$ satisfies*

$$\begin{aligned} &\delta_i^F(t) \\ &= \max_{\{\mathcal{D}_i, \mathcal{D}'_i\}} \|F_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t) - F'_i(\boldsymbol{\omega}_i(t-1), \alpha_i(t), t)\|. \end{aligned} \tag{6}$$

In **Definition 2**, the sensitivity value $\delta_i^F(t)$ relies on the loss function. Based on [33], $\delta_i^F(t)$ is obtained through the Stone-Weierstrass theorem when the loss function is known. Based on (5), we can obtain the theorem below.

Theorem 1. *For $\forall m_i \in \mathcal{M}$, the Laplace mechanism adopted in (5) meets the relationship defined in **Definition 1**.*

Proof. A similar proof is available in [30]. □

According to [34], the privacy budget can be regarded as the remuneration requested by each device for assuming the risk associated with privacy exposure. For simplicity, this compensation can be treated as a special cost, which is related to $\epsilon_i(t)$. The differential privacy cost of m_i is calculated as

$$C_i^P(t) = \mu_i^P \epsilon_i(t), \tag{7}$$

where μ_i^P is the unit differential privacy cost related to how much the device cares about its privacy compensation.

D. Local Model Uploading

After local model training and adding differential privacy, each device uploads its local model parameter through multi-mode networks to the controller. The uploading rate of m_i is given by

$$R_i(t) = B_i(t) \log_2 \left(1 + \frac{P_i(t) g_i(t)}{\sigma_0 + V_i(t)}\right), \tag{8}$$

where $B_i(t)$ is the uploading bandwidth of m_i. $P_i(t)$ is the uploading power of m_i. $g_i(t)$ is the channel gain. σ_0 is Gaussian white noise. $V_i(t)$ is electromagnetic interference power.

Define $|\tilde{\boldsymbol{\omega}}_i(t)|$ as the size of the local model parameter of m_i. The uploading delay and energy consumption are calculated as

$$\tau_i^U(t) = \frac{|\tilde{\boldsymbol{\omega}}_i(t)|}{R_i(t)}, \tag{9}$$

$$E_i^U(t) = P_i(t)\tau_i^U(t). \tag{10}$$

Define the total energy consumption of each device as the sum of energy consumption of local model training and energy consumption of model uploading, which is represented as

$$E_i(t) = E_i^L(t) + E_i^U(t). \tag{11}$$

E. Global Aggregation

The controller carries out global aggregation to update the global model after receiving the uploaded local model parameter from each device. The global aggregation is represented as

$$\boldsymbol{\omega}_g(t) = \sum_{i=1}^{I} \frac{\alpha_i(t)\tilde{\boldsymbol{\omega}}_i(t)}{\sum_{i=1}^{I}\alpha_i(t)}. \tag{12}$$

Define $F_g(\boldsymbol{\omega}_g(t), t)$ as the global loss function, which is utilized to measure the accuracy of the global model. $F_g(\boldsymbol{\omega}_g(t), t)$ is updated as

$$F_g(\boldsymbol{\omega}_g(t), t) = \sum_{i=1}^{I} \frac{\alpha_i(t)}{\sum_{i=1}^{I}\alpha_i(t)} F_i(\tilde{\boldsymbol{\omega}}_i(t), \alpha_i(t), t). \tag{13}$$

III. Problem Formulation and Transformation

In this section, we first elaborate on the long-term differential privacy constraint. Then, the joint optimization problem of resource scheduling is formulated. Finally, the problem transformation is introduced.

A. Differential Privacy Constraint

The differential privacy mechanism measures the value of privacy loss through the privacy budget. However, in each iteration, the model interaction between the device and the controller results in a certain degree of privacy loss. The privacy loss accumulates with the number of iterations [14], [18]. A small privacy loss in the short term can accumulate into a significant privacy loss in the long term. Therefore, a long-term differential privacy constraint is defined as

$$\sum_{t=1}^{T} \epsilon_i(t) \le \epsilon_{i,\max}, \tag{14}$$

where $\epsilon_{i,\max}$ is the upper threshold of privacy budget.

B. Problem Formulation

Regarding the relationship among the global loss function, total energy consumption, and differential privacy cost, we give two remarks as follows.

Remark 1. *When the batch size $\alpha_i(t)$ is large adequately, the difference between the protected model parameter $\tilde{\boldsymbol{\omega}}_i(t)$ and the actual model parameter $\boldsymbol{\omega}_i(t)$ is infinitely close to 0, which ensures the convergence performance of $F_g(\boldsymbol{\omega}_g(T), T)$. However, the total energy consumption of each device will increase as well.*

Proof. According to (3) and (11), the increase of the total energy consumption is obvious, while a similar proof of the convergence performance guarantee of $F_g(\boldsymbol{\omega}_g(T), T)$ is available in [30]. □

Remark 2. *Increasing the privacy budget $\epsilon_i(t)$ will improve the convergence performance of $F_g(\boldsymbol{\omega}_g(T), T)$, ensuring the precision of the global model, but will reduce the level of privacy preservation and increase the differential privacy cost as well.*

Proof. According to (7), the increase of the differential privacy cost is obvious, while a similar proof of the relationship among $\epsilon_i(t)$, the convergence performance of $F_g(\boldsymbol{\omega}_g(T), T)$, and the level of privacy preservation is available in [35]. □

We construct a joint resource scheduling optimization problem for green multi-mode PIoT under the model inversion attacks. The goal is to minimize the weighted sum of the global loss function, total energy consumption, and differential privacy cost by jointly optimizing the scheduling of training batch size and privacy budget under the long-term constraint of differential privacy. The formulation of optimization problem is represented as

$$\begin{aligned}
\mathbf{P1}: &\min_{\{\alpha_i(t), \epsilon_i(t)\}} F_g(\boldsymbol{\omega}_g(T), T) + W_E \sum_{t=1}^{T}\sum_{i=1}^{I} E_i(t) \\
&+ W_C \sum_{t=1}^{T}\sum_{i=1}^{I} C_i^P(t), \\
\text{s.t. } & C_1 : \alpha_i(t) \in \mathcal{A}_i(t),\ \forall m_i \in \mathcal{M},\ \forall t \in \mathcal{T}, \\
& C_2 : \epsilon_i(t) \in \mathcal{E}(t),\ \forall m_i \in \mathcal{M},\ \forall t \in \mathcal{T}, \\
& C_3 : \sum_{t=1}^{T} \epsilon_i(t) \le \epsilon_{i,\max},\ \forall m_i \in \mathcal{M},
\end{aligned} \tag{15}$$

where W_E and W_C are the weights of total energy consumption and differential privacy cost. C_1 is the training batch size scheduling constraint. C_2 is the privacy budget scheduling constraint. C_3 is the long-term differential privacy constraint.

C. Problem Transformation

Since the resource scheduling strategy of each iteration is coupled with $F_g(\boldsymbol{\omega}_g(T), T)$ as well as the long-term constraint of differential privacy, **P1** is hard to be solved directly. Therefore, we first convert the long-term optimization goal to a series of short-term goals addressed in each iteration through the telescoping sum theory. $F_g(\boldsymbol{\omega}_g(T), T)$ is decoupled as

$$\begin{aligned}
&F_g(\boldsymbol{\omega}_g(T), T) \\
&= \frac{1}{T}\left[\sum_{t=1}^{T} F_g(\boldsymbol{\omega}_g(t), t) - \sum_{t=1}^{T} F_g(\boldsymbol{\omega}_g(t-1), t-1)\right],
\end{aligned} \tag{16}$$

where $F_g(\boldsymbol{\omega}_g(t-1), t-1)$ is known in the t-th iteration. Therefore, the optimization of $F_g(\boldsymbol{\omega}_g(T), T)$ can be converted to that of $F_g(\boldsymbol{\omega}_g(t), t)$ in the t-th iteration.

For the coupling between the resource scheduling strategy and the long-term constraint, a virtual differential privacy deficit queue $H_i^{\epsilon}(t)$ corresponding to C_3 is introduced. The queue evolution is represented as

$$H_i^{\epsilon}(t+1) = \max\left\{H_i^{\epsilon}(t) + \epsilon_i(t) - \frac{\epsilon_{i,\max}}{T}, 0\right\}. \tag{17}$$

Based on the virtual queue theory [36], C_3 holds automatically when $H_i^{\epsilon}(t)$ is mean rate stable. To minimize the weighted sum while guaranteeing the long-term constraint of differential privacy to the maximum extent, $\mathbf{P1}$ can be transformed into

$$\begin{aligned}\mathbf{P2}: \min_{\{\alpha_i(t),\epsilon_i(t)\}} & \frac{1}{T}F_g(\boldsymbol{\omega}_g(t),t) + W_E\sum_{i=1}^{I}E_i(t) \\ & + W_C\sum_{i=1}^{I}C_i^P(t) + W_H\sum_{i=1}^{I}H_i^{\epsilon}(t)\epsilon_i(t), \\ \text{s.t. } & C_1, C_2, \\ & C_4: H_i^{\epsilon}(t) \text{ is mean rate stable,}\end{aligned} \tag{18}$$

where W_H is the corresponding weight of the differential privacy deficit queue.

IV. FREEDOM: Differential Privacy-Aware GAN-Assisted Resource Scheduling Algorithm

In this section, we propose FREEDOM to address $\mathbf{P2}$ and realize coordinated resource scheduling under model inversion attacks for green multi-mode PIoT.

A. MDP Model

$\mathbf{P2}$ is formulated as a Markov decision process (MDP). The specific introduction is below.

1) State Space: Define $\mathbf{H}^{\epsilon}(t) = \{H_1^{\epsilon}(t), \cdots, H_i^{\epsilon}(t), \cdots, H_I^{\epsilon}(t)\}$ as the set of differential privacy deficit queues and $\boldsymbol{\epsilon} = \{\epsilon_{1,\max}, \cdots, \epsilon_{i,\max}, \cdots, \epsilon_{I,\max}\}$ as the set of the upper thresholds of privacy budget. The state space is represented as $\mathbf{S}(t) = \mathbf{H}^{\epsilon}(t) \otimes \boldsymbol{\epsilon}$, where $\otimes$ is the Cartesian product.

2) Action Space: The action space of m_i is defined as $\mathbf{a}_i(t) = \mathcal{A}_i(t) \otimes \mathcal{E}(t)$.

3) Reward Function: The reward function is defined as the negative value of $\mathbf{P2}$'s optimization goal, i.e.,

$$\begin{aligned}\Omega(t) = & -\frac{1}{T}F_g(\boldsymbol{\omega}_g(t),t) - W_E\sum_{i=1}^{I}E_i(t) \\ & - W_C\sum_{i=1}^{I}C_i^P(t) - W_H\sum_{i=1}^{I}H_i^{\epsilon}(t)\epsilon_i(t).\end{aligned} \tag{19}$$

B. FREEDOM

DQN provides a model-free solution to the above MDP problem. However, traditional DQN-based approaches have a long process of decision-making exploration and optimization when facing high-dimensional state and action spaces, leading to slow convergence. Meanwhile, they do not consider the differential privacy awareness, resulting in poor resource scheduling performance. GAN continuously optimizes the relationship between a generator network and a discriminator network through adversarial learning, which is conducive to stable exploration and optimization as well as efficient convergence of DQN. Thus, we propose FREEDOM to sense the impact of differential privacy on resource scheduling, and improve the training and convergence performances with the assistance of GAN.

Algorithm 1 FREEDOM

```
Input: ${\mathcal{M}, \mathcal{N}, \mathcal{T}, \mathbf{S}(t), \boldsymbol{\epsilon}}$.
Output: ${\alpha_i(t), \epsilon_i(t)}$.
# Initialization:
    Initialize $\boldsymbol{\vartheta}_i^G(0)$, $\boldsymbol{\vartheta}_i^{tarG}(0)$, and $\boldsymbol{\vartheta}_i^D(0)$, $\forall m_i \in \mathcal{M}$. Set $\alpha_i(0) = 0$, $\epsilon_i(0) = 0$, $\Omega(0) = 0$, and $H_i^{\epsilon}(0) = 0$, $\forall m_i \in \mathcal{M}$.
For $t = 1, \cdots, T$ do
# Differential Privacy-aware DQN-based Generator Training:
    The controller draws an action of resource scheduling based on the $\varepsilon$-greedy approach.
    Each device executes $\alpha_i(t)$ and $\epsilon_i(t)$.
    The controller observes the global loss function, total energy consumption, differential privacy cost, and deficit queue performances, and calculates reward $\Omega(t)$ as (19).
    Update $H_i^{\epsilon}(t+1)$ as (17).
    Transfer $\mathbf{S}(t)$ to $\mathbf{S}(t+1)$, and update the experience replay pool $\mathcal{U}_i(t)$.
    Randomly sample $\tilde{\mathcal{U}_i}(t)$ and calculate $\eta_i^G(t)$ as (20).
    The controller updates $\boldsymbol{\vartheta}_i^G(t+1)$ as (22).
    Update $\boldsymbol{\vartheta}_i^{tarG}(t) = \boldsymbol{\vartheta}_i^G(t)$ every $T_0$ iterations.
# Discriminator Training:
    The controller calculates $\eta_i^D(t)$ as (23).
    The controller updates $\boldsymbol{\vartheta}_i^D(t+1)$ as (24).
end for
```

The framework of FREEDOM is shown in Fig. 3. The controller acts as an agent to execute FREEDOM, which maintains a generator based on differential privacy-aware DQN and a discriminator for each device. The generator G_i consists of a main network with parameter $\boldsymbol{\vartheta}_i^G(t)$ and a target network with parameter $\boldsymbol{\vartheta}_i^{tarG}(t)$. The discriminator D_i maintains a discriminator network with $\boldsymbol{\vartheta}_i^D(t)$. The principle of differential privacy awareness is to incorporate the differential privacy queue deficit into the calculation of reward function and state-action value. FREEDOM can continuously fit the complex mapping between states and actions based on the dynamic changes of differential privacy queue deficit. Hence, the controller can actively learn the resource scheduling strategy to improve the global loss function, energy consumption, and differential privacy cost. The implementation procedures of FREEDOM are summarized in Algorithm 1.

1) Initialization: Initialize generator network parameters $\boldsymbol{\vartheta}_i^G(0)$ and $\boldsymbol{\vartheta}_i^{tarG}(0)$, and discriminator network parameter $\boldsymbol{\vartheta}_i^D(0)$. Set $\alpha_i(0) = 0$, $\epsilon_i(0) = 0$, $\Omega(0) = 0$, and $H_i^{\epsilon}(0) = 0$.

2) Differential Privacy-aware DQN-based Generator Training: The training purpose of the generator is to improve its ability to generate the resource scheduling strategy that confuses discriminator D_i in a best-effort way. The training

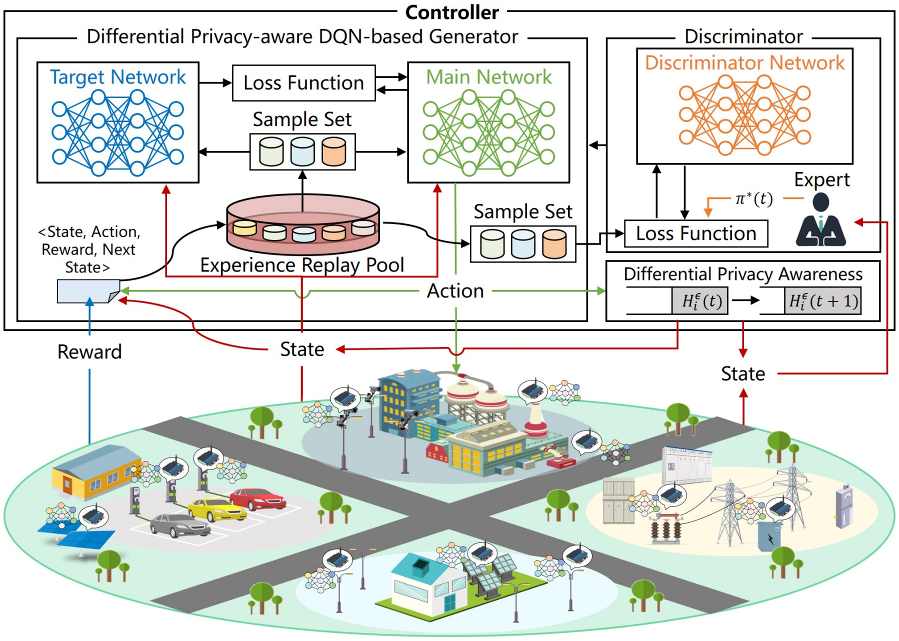

Fig. 3. The framework of FREEDOM.

process of the generator is introduced as follows.

Action Drawing and Reward Calculation: In each iteration, the controller utilizes the ε-greedy approach to draw the action of resource scheduling. Each device performs the drawn action. Then, the controller observes multiple performances such as global loss function, total energy consumption, differential privacy cost, and differential privacy queue deficit, and calculates $\Omega(t)$ based on (19).

Differential Privacy Deficit Queue Updating and State Transition: Then, the controller updates $H_i^\epsilon(t+1)$ as (17), transfers the state $\mathbf{S}(t)$ to the next state $\mathbf{S}(t+1)$, generates experience information $u_i(t) = \{\mathbf{S}(t), \mathbf{a}_i(t), \Omega(t), \mathbf{S}(t+1)\}$, and updates an experience replay pool $\mathcal{U}_i(t)$.

Learning and Network Training: The controller randomly samples a set $\tilde{\mathcal{U}}_i(t)$ from $\mathcal{U}_i(t)$, and calculates the loss function of the main network. The loss function of the main network includes two parts, where the first part reflects the degree of confusion discrimination, and the second part reflects the training accuracy of the main network. The loss function of the main network is represented as

$$\eta_i^G(t) = -\mathbb{E}_{\tilde{\mathcal{U}}_i(t)}\left\{\ln D_{\boldsymbol{\vartheta}_i^D(t)}\left(G_{\boldsymbol{\vartheta}_i^G(t)}\left(\mathbf{S}(t), \mathbf{a}_i(t)\right)\right)\right\} + \varpi_i \mathbb{E}_{\tilde{\mathcal{U}}_i(t)}\left\{\chi_i(t)^2\right\}, \tag{20}$$

where ϖ_i is the adjusting weight. $D_{\boldsymbol{\vartheta}_i^D(t)}(.)$ is the function of discrimination network with parameter $\boldsymbol{\vartheta}_i^D(t)$. $G_{\boldsymbol{\vartheta}_i^G(t)}(.)$ is the function of main network with parameter $\boldsymbol{\vartheta}_i^G(t)$. $\chi_i(t)$ is the temporal difference (TD) error, which is calculated as

$$\chi_i(t) = \Omega(t) + \kappa \max_{\mathbf{a}_i(t+1)} G_{\boldsymbol{\vartheta}_i^{tarG}(t)}\left(\mathbf{S}(t+1), \mathbf{a}_i(t+1)\right) - G_{\boldsymbol{\vartheta}_i^G(t)}\left(\mathbf{S}(t), \mathbf{a}_i(t)\right), \tag{21}$$

where κ is the discount factor. $G_{\boldsymbol{\vartheta}_i^{tarG}(t)}(.)$ is the function of target network with parameter $\boldsymbol{\vartheta}_i^{tarG}(t)$.

Finally, the main network parameter $\boldsymbol{\vartheta}_i^G(t)$ is updated based on the gradient descent approach, i.e.,

$$\boldsymbol{\vartheta}_i^G(t+1) = \boldsymbol{\vartheta}_i^G(t) - \iota^G \nabla_{\boldsymbol{\vartheta}_i^G(t)} \eta_i^G(t)^2, \tag{22}$$

where ι^G is the learning step of the generator. The target network is updated as $\boldsymbol{\vartheta}_i^{tarG}(t) = \boldsymbol{\vartheta}_i^G(t)$ every T_0 iterations.

3) Discriminator Training: The training purpose of the discriminator is to improve its ability to distinguish the expert resource scheduling policy from the resource scheduling strategy generated by the generator. Thus, the loss function of the discriminator network is defined as

$$\eta_i^D(t) = \mathbb{E}_{\tilde{\mathcal{U}}_i(t)}\left\{\ln D_{\boldsymbol{\vartheta}_i^D(t)}\left(\mathbf{S}(t), \boldsymbol{\pi}_i^*(t)\right)\right\} - \mathbb{E}_{\tilde{\mathcal{U}}_i(t)}\left\{\ln D_{\boldsymbol{\vartheta}_i^D(t)}\left(G_{\boldsymbol{\vartheta}_i^G(t)}\left(\mathbf{S}(t), \mathbf{a}_i(t)\right)\right)\right\}, \tag{23}$$

where $\boldsymbol{\pi}_i^*(t)$ is the expert resource scheduling policy of m_i from the expert memory database.

The first term in (23) represents the output of the discriminator network with the input of the resource scheduling strategy generated by the generator, while the second term represents the output of the discriminator network with the input of the corresponding expert policy. The larger $\eta_i^D(t)$ is, the stronger the discriminator is at distinguishing between the expert policy and the strategy generated by the generator. Thus, the discriminator network parameter $\boldsymbol{\vartheta}_i^D(t+1)$ is updated based on the gradient ascent approach, i.e.,

$$\boldsymbol{\vartheta}_i^D(t+1) = \boldsymbol{\vartheta}_i^D(t) + \iota^D \nabla_{\boldsymbol{\vartheta}_i^D(t)} \eta_i^D(t)^2, \tag{24}$$

where ι^D is the learning step of the discriminator.

FREEDOM, through the dynamic adjustment of the resource scheduling strategy based on the differential privacy queue deficit information, ensures differential privacy awareness. In particular, when the differential privacy deficit escalates significantly and the privacy budget deviates from

TABLE I
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
T	100	I	10
$B_i(t)$	0.1 MHz	$P_i(t)$	0.1 W
ζ_i	10^6 cycles	e_i	10^{-27} Watt $\cdot$ s^3/cycle^3
W_E	$[0.01, 0.06]$	W_C	$[1, 6] \times 10^{-4}$
W_H	$[1, 2] \times 10^{-4}$	ϵ_1, ϵ_K	$1, 30$
$\xi_i(t)$	1.5 GHz	$\lvert\mathcal{D}_i\rvert$	$[100, 500]$
$\epsilon_{i,\max}$	$[2200, 2600]$	K	10
μ_i^P	$[0.5, 2]$	σ_0	-114 dBm
λ	0.01	κ	0.99
ι^G	0.02	ι^D	0.02

the corresponding constraint, the reward decreases while the TD error increases. This encourages the controller to train the generator and discriminator, and draw another resource scheduling action, thus enabling differential privacy awareness. The assistance of GAN enables FREEDOM to achieve more accurate optimization and rapid convergence speed. When the discriminator can easily distinguish the generated strategy and expert policy, the controller adjusts the training direction of the generator to confuse the discriminator better, and vice versa. When the game between the generator and discriminator reaches the Nash equilibrium, the generator can be utilized to make the actual resource scheduling strategy without relying on expert experience.

V. PERFORMANCE ANALYSIS AND SIMULATIONS

We consider an energy management scenario featuring 10 PIoT devices to evaluate resource scheduling and privacy preservation performances. The training data of each device are randomly sampled from a distribution station area dataset including sample data such as renewable energy production, load demands, and energy storage status. An α-stable symmetric (SαS) distribution is utilized to model the power of electromagnetic interference. The specific simulation parameters are delineated in Table I [29], [35], [37]–[39]. To verify the effectiveness of differential privacy against attacks, we compare the real data with the data recovered by attackers. Define the normalized error between the real data and the data recovered by the attacker as $Error(D_{att})$, which is given by

$$Error(D_{att}) = \frac{D_{real} - D_{att}}{D_{real}}, \tag{25}$$

where D_{real} are the real data and D_{att} are the data recovered by attackers.

A comparative analysis of two cutting-edge algorithms is conducted. The first one, federated multi-agent reinforcement learning (FMARL)-based resource allocation algorithm [40], aims to optimize the FL loss function by employing the largest differential privacy budget to guarantee model accuracy. The long-term differential privacy constraint is not considered. The second one is the resource scheduling algorithm based on local differential privacy-empowered concurrent federated reinforcement learning (LDP-CFRL) [41], the objective of which is to optimize the loss function of FL, taking into account the long-term differential privacy constraint. Neither algorithm incorporates energy consumption optimization.

Fig. 4 shows the global loss function versus iterations. The loss function optimized by FREEDOM converges at the 20-th iteration, which is much faster than those of FMARL and LDF-CFRL. When $t = 100$, FREEDOM achieves a reduction of 14.4% and 22.9% in the global loss function compared with LDP-CFRL and FMARL, respectively. The light-colored part represents the global loss function of each algorithm has some fluctuations under multiple repeated experiments. FREEDOM achieves the best loss function performance by jointly optimizing the scheduling of training batch size and privacy budgets. Through the integration of DQN with GAN, FREEDOM can fit the complex relationship between state and action spaces and capitalize on the generative adversarial mechanism of GAN to attain more precise optimization and expedite convergence.

Fig. 5 shows the total energy consumption versus iterations. The total energy consumption gradually increases with the iterations. When $t = 100$, the total energy consumption of FREEDOM is reduced by 10.9% and 11.5% compared with LDP-CFRL and FMARL. The reason is that FREEDOM implements the joint optimization of global loss function and total energy consumption. Through reasonable scheduling of training batch size, the global loss function and total energy consumption performances are ensured simultaneously. FREEDOM learns the optimal resource scheduling strategy with the assistance of GAN. Both LDP-CFRL and FMARL only consider the minimization of the loss function, leading to larger total energy consumption.

Fig. 6 shows the global loss function versus average batch size and $\epsilon_{i,\max}$. As the average batch size and $\epsilon_{i,\max}$ increases, the global loss function decreases. This phenomenon can be attributed to the fact that the larger the average batch size is, the smaller the difference between the protected model parameters and the actual model parameters is, thus reducing the global loss function. Moreover, with $\epsilon_{i,\max}$ increasing, FREEDOM can schedule more privacy budgets to guarantee the precision of the global model.

Fig. 7 shows the differential privacy cost versus iterations. The differential privacy cost of FREEDOM decreases with the increase of iterations. Compared with LDP-CFRL and FMARL, the average differential privacy cost of the algorithm is reduced by 22.1% and 36.3%, respectively. FREEDOM needs to reasonably arrange its privacy budget to meet long-term differential privacy constraints. Specifically, the global loss function decreases faster at the start of optimization. To ensure the convergence of the loss function, more privacy budgets need to be arranged, resulting in greater differential privacy costs. With the continuous improvement of the convergence of loss function, the demand for privacy preservation increases, thus reducing the differential privacy cost. FREEDOM integrates GAN to develop optimal scheduling policies and achieve better cost performance.

Fig. 8 shows total energy consumption and global loss

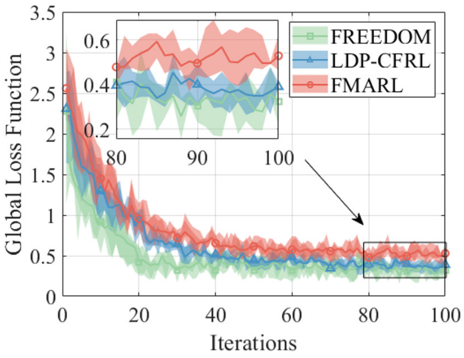

Fig. 4. Global loss function versus iterations.

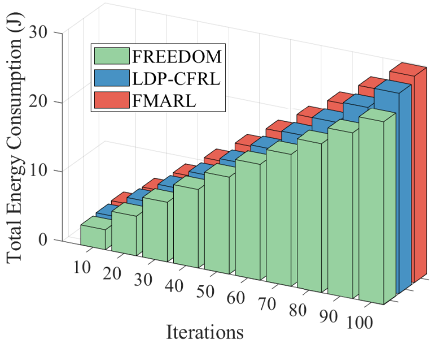

Fig. 5. Total energy consumption versus iterations.

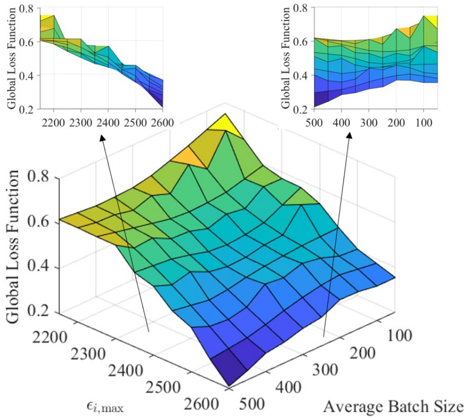

Fig. 6. Global loss function versus average batch size and $\epsilon_{i,\max}$.

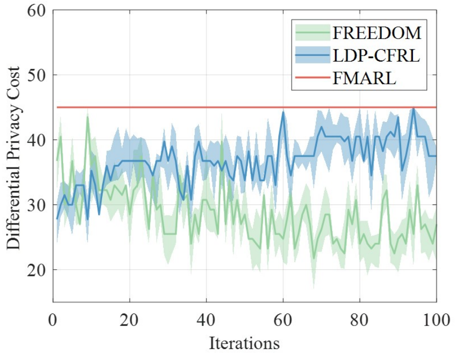

Fig. 7. Differential privacy cost versus iterations.

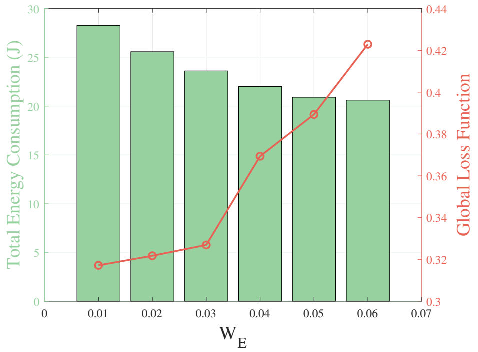

Fig. 8. Total energy consumption and global loss function versus W_E.

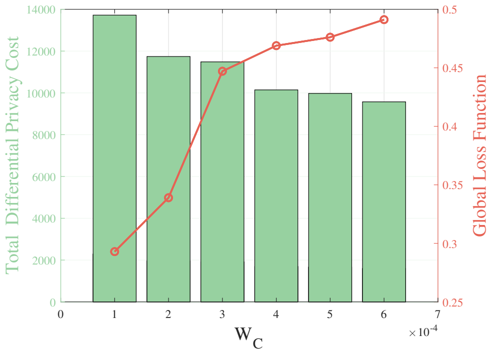

Fig. 9. Total differential privacy cost and global loss function versus W_C.

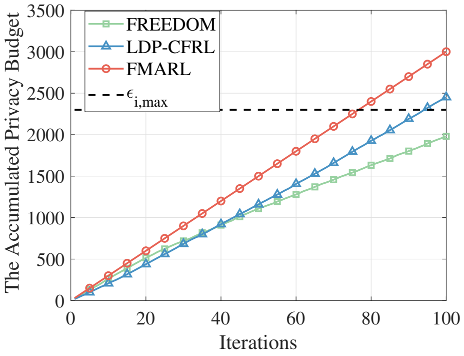

Fig. 10. The accumulated privacy budget versus iterations.

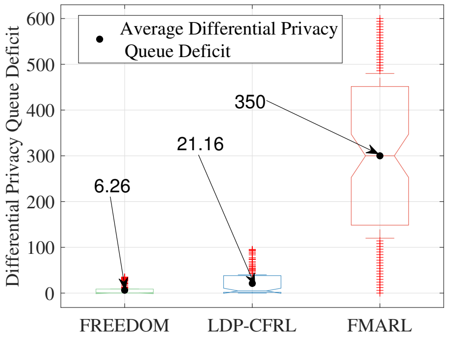

Fig. 11. Distribution of differential privacy queue deficit.

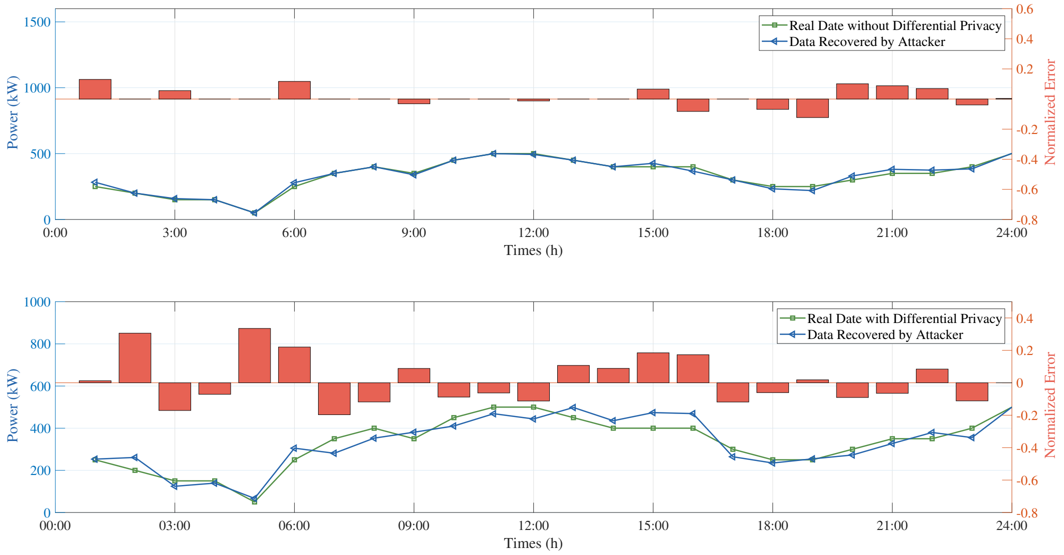

Fig. 12. Attack resistance performance of FREEDOM with differential privacy.

function versus W_E. As W_E increases from 0.01 to 0.06, the total energy consumption decreases by 55.4%, and the global loss function increases by 25.0%. The reason is that as W_E increases, FREEDOM focuses more on minimizing total energy consumption, thereby reducing the batch size used for model training, leading to an increase in the global loss function.

Fig. 9 shows the total differential privacy cost and global loss function versus W_C. As W_C increases from 1×10^{-4} to 6×10^{-4}, the total differential privacy cost decreases by 30.2% and the global loss function increases by 40.3%. The reason is that as W_C increases, FREEDOM focuses more on minimizing the differential privacy cost of all devices, which results in the introduction of substantial disturbance noise during model training, contributing to the increase in the global loss function.

Fig. 10 shows the accumulated privacy budget versus iterations. The accumulated privacy budget of FREEDOM satisfies the differential privacy constraint, while that of FMARL and LDP-CFRL do not. When $t = 100$, the accumulated privacy budget of FREEDOM is reduced by 16.5% and 36.2% compared with LDP-CFRL and FMARL. FREEDOM can realize the dynamic adjustment of the privacy budget with the assistance of GAN to better meet the long-term differential privacy constraint.

Fig. 11 shows the distribution of differential privacy queue deficit. It can be seen that FREEDOM has the lowest average differential privacy queue deficit. Compared with LDP-CFRL and FMARL, the average differential privacy queue deficit of FREEDOM is reduced by 70.4% and 98.2%. The reason is that FREEDOM can achieve differential privacy awareness through the dynamical learning of resource scheduling strategy to effectively reduce the differential privacy deficit.

Fig. 12 shows the attack resistance performance of FREEDOM with differential privacy. When there is no differential privacy, the attacker can recover the real data more accurately with less normalization error. When differential privacy exists, the normalization error between the data recovered by the attacker and the real data is larger. Compared with FREEDOM without differential privacy, the normalized error of FREEDOM with differential privacy improves by 74.3%. The differential privacy mechanism can defend against model inversion attacks by adding differential privacy noise into the protected model.

VI. Conclusion

In this paper, we addressed the joint resource scheduling optimization problem for green multi-mode PIoT under the model inversion attacks. We designed FREEDOM to dynamically learn the scheduling of training batch sizes and privacy budgets with differential privacy awareness and the assistance of GAN. Compared with LDP-CFRL and FMARL, FREEDOM improves the global loss function by 14.4% and 22.9%, total energy consumption by 10.9% and 11.5%, and differential privacy cost by 22.1% and 36.3%. FREEDOM also achieves the least differential privacy queue deficit and excellent privacy preservation performance due to differential privacy awareness. Future work will focus on the optimization of authentication and isolation for raw data and model parameters in industrial and smart park scenarios to further improve the robustness of model training. Besides, the integration of other communication modes such as wireless local area network (WLAN) and ZigBee will also be considered.

References

[1] W. Lu, P. Si, Y. Gao, H. Han, Z. Liu, Y. Wu, and Y. Gong, "Trajectory and resource optimization in OFDM-based UAV-powered IoT network," *IEEE Trans. Green Commun. Networking*, vol. 5, no. 3, pp. 1259–1270, Sept. 2021.

[2] X. Wang, M. Umehira, M. Akimoto, B. Han, and H. Zhou, "Green spectrum sharing framework in B5G era by exploiting crowdsensing," *IEEE Trans. Green Commun. Networking*, vol. 7, no. 2, pp. 916–927, Jun. 2023.

[3] H. Al Haj Hassan, D. Renga, M. Meo, and L. Nuaymi, "A novel energy model for renewable energy-enabled cellular networks providing ancillary services to the smart grid," *IEEE Trans. Green Commun. Networking*, vol. 3, no. 2, pp. 381–396, Jun. 2019.

[4] S. Zhang, Z. Wang, Z. Zhou, Y. Wang, H. Zhang, G. Zhang, H. Ding, S. Mumtaz, and M. Guizani, "Blockchain and federated deep reinforcement learning based secure cloud-edge-end collaboration in power IoT," *IEEE Wireless Commun.*, vol. 29, no. 2, pp. 84–91, Apr. 2022.

[5] H. Zhou, X. Wang, M. Umehira, B. Han, and H. Zhou, "Energy efficient beamforming for small cell systems: A distributed learning and multicell coordination approach," *ACM Trans. Sen. Netw.*, vol. pp, no. 99, pp. 1–21, Sept. 2023.

[6] S. F. Abedin, A. Mahmood, N. H. Tran, Z. Han, and M. Gidlund, "Elastic O-RAN slicing for industrial monitoring and control: A distributed matching game and deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 71, no. 10, pp. 10 808–10 822, Oct. 2022.

[7] A. Ndikumana, K. K. Nguyen, and M. Cheriet, "Federated learning assisted deep Q-learning for joint task offloading and fronthaul segment routing in open RAN," *IEEE Trans. Netw. Serv. Manage.*, vol. 20, no. 3, pp. 3261–3273, Sept. 2023.

[8] Y. Qian, L. Shi, L. Shi, K. Cai, J. Li, and F. Shu, "Cache-enabled power line communication networks: Caching node selection and backhaul energy optimization," *IEEE Trans. Green Commun. Networking*, vol. 4, no. 2, pp. 606–615, Jun. 2020.

[9] Z. Zhou, X. Chen, H. Liao *et al.*, "Collaborative learning-based network resource scheduling and route management for multi-mode green IoT," *IEEE Trans. Green Commun. Networking*, vol. 7, no. 2, pp. 928–939, Jun. 2023.

[10] J. Liu, X. Zhao, P. Qin, S. Geng, and S. Meng, "Joint dynamic task offloading and resource scheduling for WPT enabled space-air-ground power internet of things," *IEEE Trans. Network Sci. Eng.*, vol. 9, no. 2, pp. 660–677, Mar. 2022.

[11] H. Liao, Z. Zhou, N. Liu, Y. Zhang, G. Xu, Z. Wang, and S. Mumtaz, "Cloud-edge-device collaborative reliable and communication-efficient digital twin for low-carbon electrical equipment management," *IEEE Trans. Ind. Inf.*, vol. 19, no. 2, pp. 1715–1724, Feb. 2023.

[12] C.-Y. Chen, C.-A. Sung, and H.-H. Chen, "Capacity maximization based on optimal mode selection in multi-mode and multi-pair D2D communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6524–6534, Jul. 2019.

[13] X. Yan, Y. Miao, X. Li, K. K. R. Choo, X. Meng, and R. H. Deng, "Privacy-preserving asynchronous federated learning framework in distributed IoT," *IEEE Internet Things J.*, vol. 10, no. 15, pp. 13 281–13 291, Aug. 2023.

[14] L. Zhang, T. Zhu, P. Xiong, W. Zhou, and P. S. Yu, "A robust game-theoretical federated learning framework with joint differential privacy," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 3333–3346, Apr. 2023.

[15] M. Yang, I. Tjuawinata, K. Y. Lam, J. Zhao, and L. Sun, "Secure hot path crowdsourcing with local differential privacy under fog computing architecture," *IEEE Trans. Serv. Comput.*, vol. 15, no. 4, pp. 2188–2201, Aug. 2022.

[16] M. Iqbal, A. Tariq, M. Adnan, I. Ud Din, and T. Qayyum, "FL-ODP: An optimized differential privacy enabled privacy preserving federated learning," *IEEE Access*, vol. 11, no. 99, pp. 116 674–116 683, Oct. 2023.

[17] B. Wang, Y. Chen, H. Jiang, and Z. Zhao, "PPeFL: privacy-preserving edge federated learning with local differential privacy," *IEEE Internet Things J.*, vol. 10, no. 17, pp. 15 488–15 500, Sept. 2023.

[18] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Differentially private asynchronous federated learning for mobile edge computing in urban informatics," *IEEE Trans. Ind. Inf.*, vol. 16, no. 3, pp. 2134–2143, Mar. 2020.

[19] D. Kwon, J. Jeon, S. Park, J. Kim, and S. Cho, "Multiagent DDPG-based deep learning for smart ocean federated learning IoT networks," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9895–9903, Oct. 2020.

[20] T. Zhao, F. Li, and L. He, "DRL-based joint resource allocation and device orchestration for hierarchical federated learning in NOMA-enabled industrial IoT," *IEEE Trans. Ind. Inf.*, vol. 19, no. 6, pp. 7468–7479, Jun. 2023.

[21] J. Zheng, K. Li, N. Mhaisen, W. Ni, E. Tovar, and M. Guizani, "Exploring deep-reinforcement-learning-assisted federated learning for online resource allocation in privacy-preserving EdgeIoT," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21 099–21 110, Nov. 2022.

[22] S. D. Okegbile, J. Cai, H. Zheng, J. Chen, and C. Yi, "Differentially private federated multi-task learning framework for enhancing human-to-virtual connectivity in human digital twin," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 11, pp. 3533–3547, Nov. 2023.

[23] B. Ghimire and D. B. Rawat, "Recent advances on federated learning for cybersecurity and cybersecurity for federated learning for internet of things," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8229–8249, Jun. 2022.

[24] P. Qin, Y. Fu, Y. Xie, K. Wu, X. Zhang, and X. Zhao, "Multi-agent learning-based optimal task offloading and UAV trajectory planning for AGIN-power IoT," *IEEE Trans. Commun*, vol. 71, no. 7, pp. 4005–4017, Jul. 2023.

[25] H. Liao, Z. Zhou, X. Zhao, and Y. Wang, "Learning-based queue-aware task offloading and resource allocation for space–air–ground-integrated power IoT," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5250–5263, Apr. 2021.

[26] Z. Zhang, Q. Liu, Z. Huang, H. Wang, C.-K. Lee, and E. Chen, "Model inversion attacks against graph neural networks," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 9, pp. 8729–8741, Sept. 2023.

[27] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning," in *IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA, 2019, pp. 739–753.

[28] X. Shen, Y. Liu, and Z. Zhang, "Performance-enhanced federated learning with differential privacy for internet of things," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 24 079–24 094, Dec. 2022.

[29] S. A. Alvi, Y. Hong, and S. Durrani, "Utility fairness for the differentially private federated-learning-based wireless IoT networks," *IEEE Internet Things J.*, vol. 9, no. 19, pp. 19 398–19 413, Oct. 2022.

[30] Y. Xu, M. Xiao, H. Tan, A. Liu, G. Gao, and Z. Yan, "Incentive mechanism for differentially private federated learning in industrial internet of things," *IEEE Trans. Ind. Inf.*, vol. 18, no. 10, pp. 6927–6939, Oct. 2022.

[31] G. Muthukrishnan and S. Kalyani, "Grafting Laplace and Gaussian distributions: A new noise mechanism for differential privacy," *IEEE Trans. Inf. Forensics Secur.*, vol. 18, no. 99, pp. 5359–5374, Aug. 2023.

[32] Y.-T. Tsou, H.-L. Chen, and J.-Y. Chen, "RoD: Evaluating the risk of data disclosure using noise estimation for differential privacy," *IEEE TBD*, vol. 7, no. 1, pp. 214–226, Mar. 2021.

[33] J. Zhang, Z. Zhang, X. Xiao, Y. Yang, and M. Winslett, "Functional mechanism: Regression analysis under differential privacy," *Proc. VLDB Endow.*, vol. 5, no. 11, pp. 1364–1375, Jul. 2012.

[34] A. Ghosh and A. Roth, "Selling privacy at auction," *Games Econ. Behav.*, vol. 91, no. 99, pp. 334–346, May 2015.

[35] K. Wei, J. Li, M. Ding *et al.*, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, no. 99, pp. 3454–3469, Apr. 2020.

[36] H. Liao, Z. Zhou, Z. Jia *et al.*, "Ultra-low AoI digital twin-assisted resource allocation for multi-mode power IoT in distribution grid energy management," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 10, pp. 3122–3132, Oct. 2023.

[37] Z. Zhou, Z. Jia, H. Liao, W. Lu, S. Mumtaz, M. Guizani, and M. Tariq, "Secure and latency-aware digital twin assisted resource scheduling for 5G edge computing-empowered distribution grids," *IEEE Trans. Ind. Inf.*, vol. 18, no. 7, pp. 4933–4943, Jul. 2022.

[38] H. Chen, S. Huang, D. Zhang, M. Xiao, M. Skoglund, and H. V. Poor, "Federated learning over wireless IoT networks with optimized communication and resources," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 16 592–16 605, Sept. 2022.

[39] S. Zhang, Z. Yao, H. Liao, Z. Zhou, Y. Chen, and Z. You, "Endogenous security-aware resource management for digital twin and 6G edge intelligence integrated smart park," *China Commun.*, vol. 20, no. 2, pp. 46–60, Feb. 2023.

[40] Y. Cui, K. Cao, and T. Wei, "Reinforcement learning-based device scheduling for renewable energy-powered federated learning," *IEEE Trans. Ind. Inf.*, vol. 19, no. 5, pp. 6264–6274, May 2023.

[41] W. Zhou, T. Zhu, D. Ye, W. Ren, and K.-K. R. Choo, "A concurrent federated reinforcement learning for IoT resources allocation with local differential privacy," *IEEE Internet Things J.*, vol. PP, no. 99, pp. 1–14, Sept. 2023.

Sunxuan Zhang is currently working toward the Ph.D degree in electrical engineering with North China Electric Power University, Beijing, China. His research interests include resource allocation and network security in smart grid communication and PIoT.

Jiapeng Xue is currently working toward the M.E. degree in information and communication engineering at North China Electric Power University. His research interests include smart grid communication and PIoT.

Jiayi Liu is currently working toward the B.S. degree in communication engineering at North China Electric Power University, Beijing, China. Her research interests include resource allocation in smart grid communication and PIoT.

Zhenyu Zhou (Senior Member, IEEE) received the M.E. and Ph.D. degrees in international information and communication studies from Waseda University, Tokyo, Japan, in 2008 and 2011, respectively. From 2012 to 2019, he was an Associate Professor with the School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China, where he has been a Full Professor since 2019. His research interests include PIoT, smart grid information and communication, communication-sensing-computing integration, and smart grid energy management. He was the recipient of the IET Premium Award in 2017, IEEE Globecom 2018 Best Paper Award, IEEE International Wireless Communications and Mobile Computing Conference 2019 Best Paper Award, and IEEE Communications Society Asia-Pacific Board Outstanding Young Researcher. He was an Associate Editor for IEEE Internet of Things Journal, IET Quantum Communication, IEEE Access, and EURASIP Journal on Wireless Communications and Networking, and the Guest Editor of IEEE Communications Magazine, IEEE Transactions on Industrial Informatics, and Transactions on Emerging Telecommunications Technologies. He is an IET Fellow and a Senior Member of the Chinese Institute of Electronics and the China Institute of Communications.

Xiaomei Chen received the PhD degree in instrument science and technology from the School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing, China, in 2010. She is currently working in Department of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China. Her research activities focus on artificial intelligence and data science, reliability analysis of complex systems, communication system design, and medical audio signal processing.

Shahid Mumtaz (Senior Member, IEEE) received the master's and Ph.D. degrees in electrical and electronic engineering from the Blekinge Institute of Technology, Karlskrona, Sweden, and University of Aveiro, Aveiro, Portugal, in 2006 and 2011, respectively. He has more than 12 years of wireless industry/academic experience. Since 2011, he has been with the Instituto de Telecomunicac¸ ˜oes, Aveiro, Portugal, where he currently holds the position of Auxiliary Researcher and adjunct positions with several universities across the Europe-Asian Region. He is currently also a Visiting Researcher with Nokia Bell Labs, Murray Hill, NJ, USA. He is the author of 4 technical books, 12 book chapters, and more than 150 technical papers in the area of mobile communications. Dr. Mumtaz is an ACM Distinguished Speaker, Editor-in-Chief for IET Journal of Quantum Communication, Vice Chair of Europe/Africa Region IEEE ComSoc: Green Communications and Computing society, and Vice Chair for IEEE standard on P1932.1, Standard for Licensed/Unlicensed Spectrum Interoperability in Wireless Mobile Networks.