Sustainable Diffusion-based Incentive Mechanism for Generative AI-driven Digital Twins in Industrial Cyber-Physical Systems

Jinbo Wen, Jiawen Kang, Dusit Niyato, *Fellow, IEEE*, Yang Zhang, and Shiwen Mao, *Fellow, IEEE*

Abstract—**Industrial Cyber-Physical Systems (ICPSs) are an integral component of modern manufacturing and industries. By digitizing data throughout the product life cycle, Digital Twins (DTs) in ICPSs enable a shift from current industrial infrastructures to intelligent and adaptive infrastructures. Thanks to data process capability, Generative Artificial Intelligence (GAI) can drive the construction and update of DTs to improve predictive accuracy and prepare for diverse smart manufacturing. However, mechanisms that leverage sensing Industrial Internet of Things (IIoT) devices to share data for the construction of DTs are susceptible to adverse selection problems. In this paper, we first develop a GAI-driven DT architecture for ICPSs. To address the adverse selection problem caused by information asymmetry, we propose a contract theory model and develop the sustainable diffusion-based soft actor-critic algorithm to identify the optimal feasible contract. Specifically, we leverage the dynamic structured pruning technique to reduce parameter numbers of actor networks, allowing sustainability and efficient implementation of the proposed algorithm. Finally, numerical results demonstrate the effectiveness of the proposed scheme.**

Index Terms—**Industrial cyber-physical systems, generative AI, digital twins, contract theory, sustainable diffusion models.**

I. INTRODUCTION

With the advancement of industrial technologies, such as the Industrial Internet of Things (IIoT) and information communication technology, the convergence of physical and cyber spaces gives rise to a new paradigm called Industrial Cyber-Physical Systems (ICPSs) [1]. ICPSs are intricate and intelligent systems that seamlessly integrate physical and computational components, enabling real-time data exchange and decision-making. To design integrated control and monitoring systems in ICPSs, Digital Twins (DTs) have received significant attention from academia and industry [1]. DTs refer to virtual replicas that cover the life cycle of physical entities. By simulating the behavior and performance of physical objects based on real-time sensing data, DTs can possess predictive capabilities and provide insights to eliminate physical mistakes and attacks [1], thus optimizing manufacturing processes and improving ICPS performance.

Generative Artificial Intelligence (GAI) is a branch of AI technology that identifies the structures and patterns from existing data to generate various and original content [2]. The popular class of GAI applications has emerged from foundation models, such as GPT-3 and Stable Diffusion, which are trained on vast quantities of data by leveraging different learning approaches. For example, ChatGPT is trained on a large corpus of text from diverse sources. This process enables it to acquire knowledge of linguistic patterns and structures, thus automatically generating valuable content based on the prompts provided by users. Relying on its incredible data processing and generation capabilities, GAI technology has great potential to revolutionize various domains. For example, GAI can drive the progression of modern IoT and enable more adaptive and intelligent IoT applications [2], such as virtual assistants and smart surveillance.

GAI also provides a novel tool for DT innovation [3], [4]. With the help of GAI technology, the construction, maintenance, and optimization of DTs can be facilitated. In addition, researchers used GAI to enhance DT status emulation, feature abstraction, and decision-making modules [4], driving innovation across diverse applications, especially ICPSs. For instance, in smart manufacturing, GAI is capable of aiding the creation of adaptive DTs that are specifically designed for the manufacturing environment, which enhances efficiency in both production scheduling and real-time control. Although GAI holds the potential to herald a new era for DTs in ICPSs, there exist several challenges to future development:

C1) Through extensive literature review, there has been limited research conducted on the utilization of GAI to drive DTs within ICPSs. Hence, it becomes imperative to actively explore the development of DT architecture driven by GAI to improve the performance of ICPSs.

C2) GAI can synthesize supplementary data based on real-world sensing data for DT construction. Due to adverse selection problems caused by information asymmetry, IIoT devices may not be willing to share high-quality data for DT construction without reliable incentives [5].

C3) Incentivizing IIoT devices for DT construction is

J. Wen and Y. Zhang are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China (e-mails: jinbo1608@nuaa.edu.cn; yangzhang@nuaa.edu.cn).

J. Kang is with the School of Automation, Guangdong University of Technology, China (e-mail: kavinkang@gdut.edu.cn).

D. Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore (e-mail: dniyato@ntu.edu.sg).

S. Mao is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA (e-mail: smao@ieee.org).

Corresponding authors: Yang Zhang, Jiawen Kang.

a complex and high-dimensional problem [1], [5]. As an advanced GAI model, Generative Diffusion Models (GDMs) show superior performance in solving high-dimensional decision-making problems [6], however, often at the cost of substantial computational overhead during training [7]. Therefore, it is necessary to develop a sustainable GDM framework for ICPSs.

Some studies have been conducted to design incentive mechanisms for DT construction [5], [8]. However, there is no work that utilizes the sustainable GAI technique to find the optimal incentive design, thus adapting to dynamic scenarios and mitigating the environmental impact of model training.

To address the above challenges, in this paper, we first design a GAI-driven DT architecture for ICPSs. To address information asymmetry, we utilize contract theory to motivate IIoT devices to share sensing data for DT construction. Furthermore, we develop a sustainable diffusion model to find the optimal feasible contract while migrating the environmental impact for ICPSs. The main contributions of this paper are summarized as follows:

- We propose a GAI-driven DT architecture for ICPSs. In particular, we study how GAI drives the pipeline for creating DTs, including the collection of real-time physical data, communications among DTs and between DTs and physical counterparts, DT modeling and maintaining on the servers, and DT decision-making (for **C1**).
- To effectively alleviate the adverse selection problem caused by information asymmetry, we propose a contract theory model to motivate IIoT devices to participate in data sharing for DT construction, and the optimal feasible contract is achieved by maximizing the expected utility of the DT server (for **C2**).
- To achieve sustainability for DT construction within ICPSs, we develop the sustainable diffusion-based Soft Actor-Critic (SAC) algorithm to generate the optimal contract under information asymmetry, where we apply the dynamic structured pruning technique to GDM-based networks. *To the best of our knowledge, this is the first work that leverages the sustainable diffusion model for incentive design* (for **C3**).

The rest of the paper is organized as follows: Section II reviews the related work. In Section III, we introduce the GAI-driven DT architecture for ICPSs. In Section IV, we propose the contract model to incentivize IIoT devices to share sensing data for DT construction. In Section V, we propose the sustainable diffusion-based SAC algorithm to sustainably generate the optimal contract. In Section VI, we present numerical results to demonstrate the effectiveness of the proposed algorithm. Section VII concludes the paper.

II. RELATED WORK

In this section, we discuss several related works, including GAI-driven DTs, incentive mechanism design for DT construction, and structured pruning techniques.

A. Generative AI-driven Digital Twins

GAI possesses formidable capabilities to drive DTs from various domains [3], [4], [9]. In [3], the authors explored the potential of GAI-driven human DTs, including GAI-enabled data acquisition, communication, data management, digital modeling, and data analytics. The authors in [4] proposed a GAI-driven DT network architecture to realize intelligent external and internal closed-loop network management, where GAI models can drive DT status emulation, feature abstraction, and network decision-making. In [9], the authors introduced how GAI facilitates DT modeling and provided existing implementations of GAI for DTs in drug discovery and clinical trials. While the integration of GAI and DTs is capable of revolutionizing various sectors, there is no work systematically studying how GAI can drive DT in ICPSs.

B. Incentive Mechanism Design for Digital Twin Construction

A few works have been conducted to design incentive mechanisms for DT construction [5], [8]. In [5], the authors proposed an iterative contract design to motivate IoT devices to share data for DT construction and used a multi-agent reinforcement learning algorithm to solve the formulated contract problem. In [8], the authors proposed a DT edge network framework for flexible and secure DT construction. To efficiently construct DTs, the authors also proposed an iterative double auction-based joint cooperative federated learning and local model update verification scheme. Incentive mechanism design for ensuring high-quality DT construction is critical, and this topic is still worth investigating. However, none of the existing studies have employed GAI techniques to discover the optimal incentive design for DT construction.

C. Structured Pruning Techniques

Structured pruning focuses on eliminating parameters and substructures from networks [7]. The authors in [10] proposed a tiny multi-agent Deep Reinforcement Learning (DRL) algorithm to obtain the optimal game solution, which leverages the structured pruning technique to reduce the parameter number of the actor-critic network. The authors in [11] proposed a dynamic structured pruning approach to remove the unimportant neurons of DRL models during the training stage. In [7], the authors proposed a dedicated method for compressing diffusion models by utilizing Taylor expansion. Motivated by the above works, we, for the first time, apply structured pruning techniques to GDMs for optimal incentive design, making them more efficient and sustainable.

III. GENERATIVE AI-DRIVEN DIGITAL TWIN ARCHITECTURE

In this section, we introduce the proposed architecture of GAI-driven DTs in ICPSs from the pipeline of DT construction, as shown in Fig. 1.
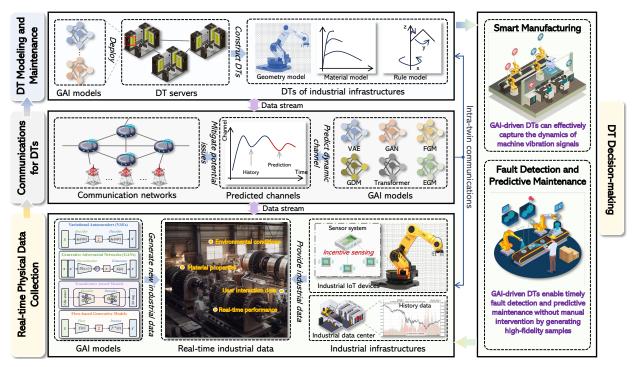
Fig. 1: The architecture of GAI-driven DTs. We study how GAI drives DTs in ICPSs from the pipeline of DT construction, i.e., real-time physical data collection, communications for DTs, DT modeling and maintenance, and DT decision-making.

A. Real-time Physical Data Collection

Real-time data collection from the physical space is the first step in constructing DTs in ICPSs [12]. The required data, such as equipment parameters and real-time operation, are collected by industrial infrastructures [1]. However, due to the heterogeneity of diverse sensing sources, some real-world data may be insufficient and confidential [1], affecting the rendering quality of DTs. Fortunately, GAI aids in data generation [2], which can provide ultra-realistic industrial data based on the collected data to improve DT training.

B. Communications for Digital Twins

Sending the collected data to servers of DTs in edge computing is critical for DT construction [12]. Communications among DTs and between DTs and physical counterparts act as a bridge for data transmission between physical and virtual spaces [3], such as collected data transmission and DT feedback. However, due to the potentially harsh industrial environments, the communication channel conditions in ICPSs may dynamically change [1], causing reliability issues in data transmission. Fortunately, GAI can predict the communication channel dynamics to mitigate potential issues. For example, a recent study [13] proposed a convolutional time-series generative adversarial network to generate synthetic data based on original channel data, which can enhance the prediction accuracy of DT channels in dynamic channel conditions.

C. Digital Twin Modeling and Maintenance

Based on collected data, high-fidelity DTs can be constructed by reproducing the geometry and physical properties of industrial infrastructures [12]. These virtual models are maintained on the edge servers to ensure real-time updates [12]. However, due to the complicated environments of ICPSs [1], traditional DT modeling methods, such as structural modeling or behavioral modeling [12], rely on accurate simulation parameters [3], making it challenging to continuously guarantee the quality of DTs. Thanks to the robust data generation capability, GAI can provide supplementary data for DT modeling in ICPSs, including environmental conditions and operational parameters, and even allow DTs to update and refine their knowledge base over time. Moreover, the augmented data can preserve all the patterns from the original data [3], thus protecting the private data of ICPSs.

D. Digital Twin Decision-Making

After achieving the virtual-real mapping, DTs can timely analyze real-world data and provide decision-making results to ICPSs [12]. GAI-driven DTs enable intelligent and reliable applications, such as smart manufacturing, fault detection, and predictive maintenance. For smart manufacturing, GAI-driven DTs can timely forecast product testing and evaluate manufacturing strategies based on generated fault diagnosis. Then, manufacturers can react according to the feedback conveyed by DTs.

IV. Contract Modeling

Sensing IIoT devices equipped with a set of sensors can collect geospatial data from the surrounding environment and send the data to the DT server for DT construction [5]. If the data quality is low, the created DTs may not accurately reflect the real-time dynamics of industrial infrastructures. Thus, High-quality sensing data is critical for DT construction.

However, due to the high cost of collecting data in ICPSs [1], IIoT devices may not actively send high-quality sensing data to the DT server, or even maliciously provide harmful data [14], leading to adverse selection problems caused by information asymmetry [5]. To this end, we use contract theory to motivate IIoT devices to provide high-quality sensing data for DT construction, where contract theory as a powerful economy tool aims at addressing information asymmetry [5], [14]. In this paper, we consider a DT server and M IIoT devices for DT construction.

A. Utility of Sensing IIoT Devices

While the sensing and communication levels of IIoT devices are not fully disclosed due to information asymmetry [14], the levels of IIoT devices can be divided into different discrete types, which are sorted in ascending order, i.e., $0 < \psi_1 \leq \cdots \leq \psi_K$. Especially, IIoT devices with higher sensing and communication levels, such as Unmanned Aerial Vehicles (UAVs), are characterized as higher types and can provide better sensing data. Motivated by [5], the utility of a sensing IIoT device with type ψ_k to send sensing data to the DT server with size $\hat{s}_k$ is defined as

$$\tilde{u}_k(\hat{s}_k, r_k) = \rho\psi_k r_k - c\hat{s}_k - c_0, \qquad (1)$$

where ρ is a pre-defined weight parameter about the incentive, r_k is the received reward associated with the provided data $\hat{s}_k$, c is the unit cost related to data collection, computation, and transmission for the data $\hat{s}_k$ [14], and c_0 is an additional cost involving energy consumption [5].

B. Utility of Digital Twin Server

Considering the more high-quality data provided by IIoT devices, the higher satisfaction of the DT server, we adopt the β-fairness function $g(\hat{s}_k)$ to quantify the DT server's satisfaction [5], given by

$$g(\hat{s}_k) = \frac{1}{1-\beta}\hat{s}_k^{1-\beta}, \qquad (2)$$

where $0 \leq \beta < 1$ is a pre-defined constant. Since the DT server just knows the number of IIoT devices and the distribution of each type due to information asymmetry, the overall utility of the DT server is given by

$$U(\hat{\boldsymbol{s}}, \boldsymbol{r}) = M\sum_{k=1}^{K} q_k(\vartheta g(\hat{s}_k) - r_k), \qquad (3)$$

where $\vartheta > 0$ is the unit revenue for the satisfaction of the DT server [14], q_k represents the probability that IIoT devices belong to type ψ_k, and $\hat{\boldsymbol{s}} = [\hat{s}_k]_{1\times K}$ and $\boldsymbol{r} = [r_k]_{1\times K}$ represent the vectors of high-quality data volume sizes and rewards, respectively.

C. Contract Formulation

To maximize the overall utility, the DT server designs a contract consisting of a group of contract items. The contract item is denoted as $\Omega = \{(\hat{s}_k, r_k), k \in \{1,\ldots,K\}\}$, where the more $\hat{s}_k$, the higher r_k. To ensure that each IIoT device
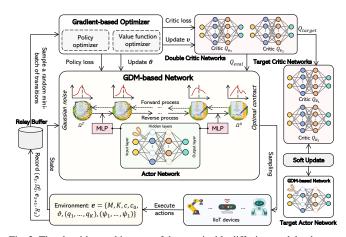

Fig. 2: The algorithm architecture of the sustainable diffusion model, where we utilize dynamic structured pruning techniques to sparsify the actor networks of the diffusion model.

selects the most suitable contract item to maximize its utility, the feasible contract should satisfy the Individual Rationality (IR) and Incentive Compatibility (IC) constraints [5], [14]:

$$\boldsymbol{IR}: \tilde{u}_k(\hat{s}_k, r_k) = \rho\psi_k r_k - c\hat{s}_k - c_0 \geq 0, \forall k \in \{1,\ldots,K\}, \quad (4)$$

$$\boldsymbol{IC}: \rho\psi_k r_k - c\hat{s}_k \geq \rho\psi_k r_n - c\hat{s}_n, \forall k, n \in \{1,\ldots,K\}, k \neq n. \quad (5)$$

The DT server aims to maximize its overall utility. Based on the above IC and IR constraints, the optimization problem is expressed as

$$\begin{aligned} \max_{\hat{\boldsymbol{s}}, \boldsymbol{r}} \ & U(\hat{\boldsymbol{s}}, \boldsymbol{r}) \\ \text{s.t. } & \text{IR (4) and IC (5),} \\ & \hat{s}_k \geq 0, r_k \geq 0, \psi_k > 0, \ \forall k \in \{1,\ldots,K\}. \end{aligned} \qquad (6)$$

Due to the heterogeneity and dynamics of ICPSs [1], the states of IIoT devices and the DT server change dynamically. Thus, traditional mathematical techniques may not effectively face the change in practice. Fortunately, GDMs have capabilities for solving dynamic and high-dimensional optimization problems [6]. For efficient implementation in ICPSs, we propose a sustainable GDM and use it to solve (6).

V. SUSTAINABLE DIFFUSION-BASED SOFT ACTOR-CRITIC FOR OPTIMAL CONTRACT DESIGN

In this section, we propose the sustainable diffusion-based SAC algorithm to efficiently generate the optimal feasible contract under information asymmetry, where we apply the dynamic structured pruning technique to GDM-based networks to reduce the training cost and carbon emissions. The architecture of the proposed algorithm is shown in Fig. 2.

A. Generative Diffusion Models for Optimal Contract Design

Diffusion models work by corrupting the training data by continuously adding Gaussian noise and then learning to recover the data by reversing this noise process [7]. Through an iterative process of denoising the initial distribution, the GDM can generate the optimal contract $\Omega^0 = (\hat{s}_k^*, r_k^*), k \in \{1,\ldots,K\}$. During iterations of T, Gaussian

noise is gradually added to the initial contract Ω, emerging a series of contract samples $(\Omega^1, \Omega^2, \ldots, \Omega^T)$. We define the environment during the optimal contract design as

$$e \triangleq \{M, K, c, c_0, \vartheta, (q_1, \ldots, q_K), (\psi_1, \ldots, \psi_K)\}. \quad (7)$$

We denote the contract design policy as $\pi_{\boldsymbol{\theta}}(\Omega|e)$ with parameters $\boldsymbol{\theta}$, which is constituted by the GDM-based network, as shown in Fig. 2. Through the reverse process of a conditional diffusion model, the policy $\pi_{\boldsymbol{\theta}}(\Omega|e)$ can be expressed as [15]

$$\pi_{\boldsymbol{\theta}}(\Omega|e) = p_{\boldsymbol{\theta}}(\Omega^{0:T}|e) = \mathcal{N}(\Omega^T; \mathbf{0}, \mathbf{I}) \prod_{t=1}^{T} p_{\boldsymbol{\theta}}(\Omega^{t-1}|\Omega^t, e), \quad (8)$$

where the end sample of the reverse chain is the selected contract Ω^0. $p_{\boldsymbol{\theta}}(\Omega^{t-1}|\Omega^t, e)$ as a noise prediction model is modeled as a Gaussian distribution $\mathcal{N}(\Omega^{t-1}; \boldsymbol{\mu}_{\boldsymbol{\theta}}(\Omega^t, e, t), \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\Omega^t, e, t))$ with the covariance matrix expressed as $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\Omega^t, e, t) = \delta_t \mathbf{I}$ [15], and the mean $\boldsymbol{\mu}_{\boldsymbol{\theta}}(\Omega^t, e, t)$ is given by

$$\boldsymbol{\mu}_{\boldsymbol{\theta}}(\Omega^t, e, t) = \frac{1}{\sqrt{\alpha_t}} \Big(\Omega^t - \frac{\delta_t}{\sqrt{1 - \bar{\alpha}_t}} \boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\Omega^t, e, t) \Big), \quad (9)$$

where $\delta_t \in (0, 1)$ is a hyperparameter for model training, $\alpha_t = 1 - \delta_t$, and $\bar{\alpha}_t = \prod_{j=0}^{t} \delta_j$ [6], [15]. $\boldsymbol{\epsilon}_{\boldsymbol{\theta}}$ is a deep model that generates contracts conditioned on the environment e, as determined by the policy. We first sample $\Omega^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and then the selected contract can be sampled via the reverse diffusion chain parameterized by $\boldsymbol{\theta}$, given by

$$\Omega^{t-1}|\Omega^t = \frac{\Omega^t}{\sqrt{\alpha_t}} - \frac{\delta_t}{\sqrt{\alpha_t(1 - \bar{\alpha}_t)}} \boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\Omega^t, e, t) + \sqrt{\delta_t}\boldsymbol{\epsilon}. \quad (10)$$

To improve the policy $\pi_{\boldsymbol{\theta}}(\Omega|e)$, we introduce the Q-function [15], which can guide the reverse diffusion chain to preferentially sample contracts with high values. We build two critic networks $Q_{\boldsymbol{v}_1}, Q_{\boldsymbol{v}_2}$, target critic networks $Q_{\hat{\boldsymbol{v}}_1}, Q_{\hat{\boldsymbol{v}}_2}$, and target policy $\hat{\pi}_{\hat{\boldsymbol{\theta}}}$. Then, we define the contract quality network as $Q_{\boldsymbol{v}}(e, \Omega) = \min\{Q_{\boldsymbol{v}_1}(e, \Omega), Q_{\boldsymbol{v}_2}(e, \Omega)\}$. Thus, the optimal contract design policy aims at maximizing the expected cumulative reward, expressed as [6]

$$\pi = \arg\max_{\pi_{\boldsymbol{\theta}}} \mathbb{E}\left[\sum_{z=0}^{Z} \gamma^z (R(e_z, \Omega_z) + \varsigma H(\pi_{\boldsymbol{\theta}}(e_z))) \right], \quad (11)$$

where γ denotes the discount factor for future rewards, $R(e_z, \Omega_z)$ represents the immediate reward upon executing action Ω_z in state e_z, ς is the temperature coefficient controlling the strength of the entropy, and $H(\pi_{\boldsymbol{\theta}}(e_z)) = -\pi_{\boldsymbol{\theta}}(e_z)\log\pi_{\boldsymbol{\theta}}(e_z)$ is the entropy of the policy $\pi_{\boldsymbol{\theta}}(e_z)$ [16]. To update the Q-function, the optimization of $\boldsymbol{v}_m$ for $m = \{1, 2\}$ aims at minimizing the objective function as [15]

$$\mathbb{E}_{(e_z, \Omega_z, e_{z+1}, R_z) \sim \mathcal{B}_z, \Phi_{z+1}^0 \sim \hat{\pi}_{\hat{\boldsymbol{\theta}}}} \Big[\sum_{m=1,2} (R(e_z, \Omega_z)$$
$$+ \gamma^z (1 - d_{z+1})\hat{\pi}_{\hat{\boldsymbol{\theta}}}(e_{z+1})Q_{\hat{\boldsymbol{v}}}(e_{z+1}) - Q_{\boldsymbol{v}_m}(e_z, \Omega_z))^2 \Big], \quad (12)$$

where $\mathcal{B}_z$ is a mini-batch of transitions with a size b sampled from the experience replay memory $\mathcal{D}$ and d_{z+1} is a 0-1 variable that represents the terminated flag [6].

B. Dynamic Structured Pruning

In terms of the network structure, both the policy and critic networks are multilayer perception networks, which consist of an input layer, multiple hidden layers, and an output layer. Considering a L-layer policy network, where parameters in the l-th fully-connected layer are denoted by $\boldsymbol{\theta}^{(l)}$, $l \in \{1, \ldots, L\}$, the output of the l-th layer is expressed as

$$\boldsymbol{h}^{(l)} = f^{(l)}(\boldsymbol{\theta}^{(l)}\boldsymbol{h}^{(l-1)} + b^{(l)}), \quad (13)$$

where $f^{(l)}$ is the activation function of the l-th layer and $b^{(l)}$ is the deviation at the l-th layer. We focus on pruning the redundant neurons and connected weights of policy networks without affecting the performance, which enhances the training efficiency of the policy network. To this end, we introduce a binary mask $m_i^{(l)}$ to represent the pruning state of each neuron $o_i^{(l)}$ at the l-th layer [11], where $m_i^{(l)} = 0$ indicates that the neuron $o_i^{(l)}$ should be pruned, and $m_i^{(l)} = 1$ indicates that the neuron $o_i^{(l)}$ should be reserved. Based on the binarized mask $\boldsymbol{m}^{(l)}$, the output of the l-th layer is rewritten as

$$\boldsymbol{h}^{(l)} = f^{(l)}(\boldsymbol{\theta}^{(l)}\boldsymbol{h}^{(l-1)} \odot \boldsymbol{m}^{(l)} + b^{(l)}), \quad (14)$$

where $\odot$ represents the Hadamard product. We use the policy gradient algorithm to update the contract design policy [6]. Specifically, the policy gradient with respect to the policy parameters $\boldsymbol{\theta}$ can be computed as the expectation over $\mathcal{B}_z$. Thus, the gradient is given by

$$\nabla_{\boldsymbol{\theta}_z} J(\boldsymbol{\theta}) = \mathbb{E}_{e_z \sim \mathcal{B}_z}[-\nabla_{\boldsymbol{\theta}_z}\pi_{\boldsymbol{\theta}_z}(e_z)Q_{\boldsymbol{v}_z}(e_z)], \quad (15)$$

where $\boldsymbol{\theta}_z$ and $\boldsymbol{v}_z$ are the policy and critic parameters at the z-th training step, respectively. Thus, the policy parameters $\boldsymbol{\theta}$ are updated by performing gradient descent based on (15), which is expressed as [6], [11]

$$\boldsymbol{\theta}_{z+1}^{(l)} \leftarrow \boldsymbol{\theta}_z^{(l)} - \eta \nabla_{\boldsymbol{h}_z^{(l)} \odot \boldsymbol{m}_z^{(l)}} J(\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}_z^{(l)}}(\boldsymbol{h}_z^{(l)} \odot \boldsymbol{m}_z^{(l)}), \quad (16)$$

where $\eta \in (0, 1]$ is the learning rate of the actor. For the target networks that have the same network structure as the online networks [6], the parameters of the target policy are also performed dynamic structured pruning, and their parameters are updated by using a soft update mechanism, given by [15]

$$\hat{\boldsymbol{\theta}}_{z+1} \leftarrow \varepsilon\boldsymbol{\theta}_z + (1 - \varepsilon)\hat{\boldsymbol{\theta}}_z,$$
$$\hat{\boldsymbol{v}}_{m,z+1} \leftarrow \varepsilon\boldsymbol{v}_{m,z} + (1 - \varepsilon)\hat{\boldsymbol{v}}_{m,z}, \text{ for } m = \{1, 2\}, \quad (17)$$

where $\varepsilon \in (0, 1]$ is the update rate of the target network.

The pruning threshold plays an important role in the pruning decisions of network parameters [10], [11]. We adopt a dynamic pruning threshold Υ, which is given by [10]

$$\Upsilon_z = \sum_{l=1}^{L} \sum_{i=1}^{I} \phi_i^{(l)} \omega_z,$$
$$\omega_z = \hat{\omega} - \hat{\omega}\Big(1 - \frac{z}{N\triangle}\Big)^3, \quad (18)$$

where $\phi_i^{(l)}$ and N represent the importance of the i-th neuron of the l-th layer and the total number of pruning steps, respectively. $\triangle$ denotes the pruning frequency, and ω_z and $\hat{\omega}$ represent the current sparsity at the z-th training step and the target sparsity, respectively.
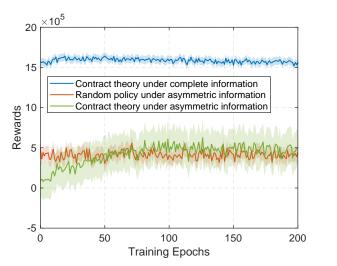
Fig. 3: Test reward comparison of the proposed scheme with the random scheme and contract theory under complete information.

After determining the pruning threshold, neurons ranked below the threshold will be pruned [10], [11], and the binary masks used for pruning are updated, given by [10]

$$m_i^{(l)} = \begin{cases} 1, & \text{if } |m_i^{(l)} \theta_i^{(l)}| \geq \Upsilon, \\ 0, & \text{otherwise.} \end{cases} \quad (19)$$

Finally, we build a compact policy network according to the redundancy of the sparse policy network. The procedure of the proposed sustainable diffusion model is shown in Algorithm 1. The computational complexity of the proposed algorithm consists of three parts. In the initialization part, the computational complexity is $\mathcal{O}(|\boldsymbol{\theta}| + |\boldsymbol{v}|)$. In the training part, the computational complexity is $\mathcal{O}(Z(T \sum_{l=1}^{L-1} |\boldsymbol{\theta}^{(l)}| + \sum_{l=1}^{L-1} |\boldsymbol{h}^{(l)}|))$ [6], [10]. In the inference part, the computational complexity is $\mathcal{O}(|\hat{\boldsymbol{\theta}}|)$. Thus, the computational complexity of the proposed algorithm is around $\mathcal{O}(|\boldsymbol{\theta}| + |\boldsymbol{v}| + |\hat{\boldsymbol{\theta}}| + Z(T \sum_{l=1}^{L-1} |\boldsymbol{\theta}^{(l)}| + \sum_{l=1}^{L-1} |\boldsymbol{h}^{(l)}|))$.

VI. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed scheme and algorithm. We consider $M = 10$ sensing IIoT devices and divide them into $K = 2$ types.

For the settings of experimental parameters, ψ_1 and ψ_2 are randomly sampled within $(50, 100)$ and $(200, 250)$, respectively. q_1 and q_2 are randomly generated following the Dirichlet distribution [2]. Considering the dynamic environment of ICPSs, the unit cost c is randomly sampled within $(25, 35)$, and the unit revenue ϑ is randomly sampled within $(10, 15)$. In addition, the pre-defined weight parameter ρ is set to 0.6, the additional cost c_0 is set to 0.01, and the pre-defined constant β is set to 0.5. Note that our experiments are conducted using PyTorch with CUDA 12.0 on NVIDIA GeForce RTX 3080 Laptop GPU.

Figure 3 presents the test reward comparison of the proposed scheme under different schemes and different scenarios. Specifically, we compare the performance of the proposed scheme under asymmetric information and complete information. The contract theory under complete information does

---

Algorithm 1: Sustainable Diffusion Models based on Dynamic Structured Pruning

---

Input: Environment $\boldsymbol{e}$ and diffusion parameters.
Output: The optimal contract design Ω^0.

1 ## *Initialize*
2 Initialize relay buffer $\mathcal{D}$, policy network $\pi_{\boldsymbol{\theta}}$, critic networks $Q_{\boldsymbol{v}_1}, Q_{\boldsymbol{v}_2}$, target networks $\hat{\pi}_{\hat{\boldsymbol{\theta}}}, Q_{\hat{\boldsymbol{v}}_1}, Q_{\hat{\boldsymbol{v}}_2}$, and binary masks $\boldsymbol{m}$.
3 ## *Training*
4 **for** the training step $z = 1$ to Z **do**
5   ### *Generating contracts*
6   Observe environment $\boldsymbol{e}_z$ and initialize a random process $\mathcal{N}$ for contract design exploration.
7   Set Ω_z^T as Gaussian noise and generate contract design Ω_z^0 by (10).
8   Execute Ω_z^0, observe the next environment $\boldsymbol{e}_{z+1}$, and obtain reward R_z.
9   Store record $(\boldsymbol{e}_z, \Omega_z^0, \boldsymbol{e}_{z+1}, R_z)$ into $\mathcal{D}$.
10   Sample a random mini-batch of transitions $\mathcal{B}_z$ from replay buffer $\mathcal{D}$.
11   Update $Q_{\boldsymbol{v}_1}, Q_{\boldsymbol{v}_2}$ using $\mathcal{B}_z$ by (12).
12   ### *Dynamic pruning and fine-tuning*
13   Compute neuron importance $\boldsymbol{\phi}_z$ based on [11].
14   Update the policy $\pi_{\boldsymbol{\theta}}$ using $\mathcal{B}_z$ by (16).
15   Update the target networks $\hat{\pi}_{\hat{\boldsymbol{\theta}}}, Q_{\hat{\boldsymbol{v}}_1}, Q_{\hat{\boldsymbol{v}}_2}$ by (17).
16   Compute the dynamic pruning threshold Υ for the policy network parameters $\boldsymbol{\theta}_z$ by (18).
17   Update binary masks $\boldsymbol{m}_z$ for $\boldsymbol{\theta}_z$ by (19).
18   **if** $\phi_i^{(l)} < \Upsilon$ **then**
19   Remove the i-th neuron at the l-th layer with the mask $m_i^{(l)}$ and associated parameters $\boldsymbol{\theta}_z^{(l)}$ from the policy network.
20   **end**
21 **end**
22 Reconstruct the compact policy networks.
23 ## *Inference*
24 Input the environment vector $\boldsymbol{e}$.
25 Generate the optimal contract design Ω^0 based on the target policy $\hat{\pi}_{\hat{\boldsymbol{\theta}}}$ by (10).
26 **return** $\Omega^0 = \{(\hat{s}_k^*, r_k^*), 1 \leq k \leq K\}$.

---

not consider IC constraints. Although the performance of this scheme is better than that of the proposed scheme, it is not practical since the environment of complete information is not realistic. Then, we compare the performance of the proposed scheme with the random scheme in which the DT server randomly designs contracts. We can observe that the test reward of the proposed scheme is higher than that of the random scheme, which indicates that the proposed scheme can effectively mitigate the effect of information asymmetry by utilizing contract theory. Overall, the proposed scheme is reliable in practice.

Figure 4 shows the performance of the proposed algorithm and other DRL algorithms in optimal contract design. We can observe that although the proposed algorithm does not converge quickly due to the influence of denoising, it can
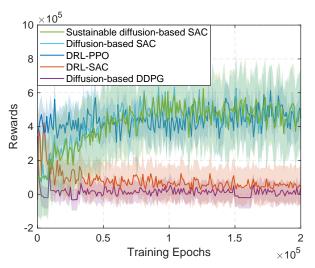
Fig. 4: Performance comparison of the proposed algorithm with other DRL algorithms in optimal contract design. For the parameter settings of the proposed algorithm, we set the pruning rate to 10%, the diffusion step to 6, the learning rate of actor networks to 2×10^{-7}, and the learning rate of critic networks to 2×10^{-6}.
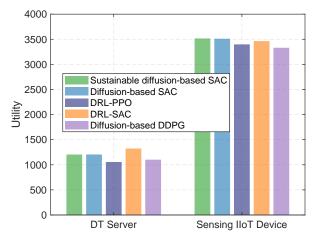

Fig. 5: The utility of the DT server and the average utility of sensing IIoT devices under five DRL algorithms.
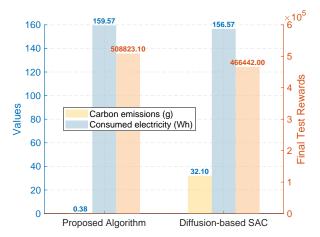

Fig. 6: Comparison of the environmental impacts of the proposed algorithm with the diffusion-based SAC algorithm.

stabilize the highest final reward compared with other DRL algorithms. The reason is that our algorithm optimizes a

stochastic policy in an entropy-augmented reward framework, encouraging exploration and robustness [16]. Moreover, based on diffusion models, our algorithm can generate samples with higher quality by multiple fine-tuning [6], enhancing the sampling accuracy and reducing the effect of uncertainty and noise from the environment. Besides, by pruning the unimportant neurons and connected weights of actor networks, the performance of the diffusion-based SAC algorithm is not affected but improved. The possible reason is that the pruning technique can reduce the complexity of GDM networks and improve model generalization to unseen states.

Figure 5 illustrates the utility of the DT server and the average utility of sensing IIoT devices under our algorithm and other DRL algorithms. Due to the dynamic environment of ICPSs, we consider that the unit revenue of the satisfaction of the DT server is dynamically changing, and it is reasonable that the proposed algorithm achieves a higher utility of the DT server rather than the highest. However, the proposed algorithm can achieve the highest average utility of sensing IIoT devices, indicating that the proposed algorithm can generate more reasonable contracts that better incentivize sensing IIoT devices to contribute data for DT construction. In summary, based on the analyses of Fig. 4 and Fig. 5, the proposed algorithm can design feasible contracts to achieve the highest utility of IIoT devices under asymmetric information.

Figure 6 evaluates the environmental impact of the proposed algorithm. We use a Python package named CodeCarbon[1] to estimate the carbon emissions and electricity consumption of these algorithms in optimal contract design. Compared with the diffusion-based SAC algorithm, which also performs well in optimal contract design, we can observe that the proposed algorithm not only has better performance, i.e., a higher final test reward but also produces lower carbon emissions of about 0.38g during model training for optimal contract design. The reason is that pruning techniques can remove excess neurons that are useless to the performance of actor networks [11], which is beneficial for decreasing model size and the need for multiple iterations, thus reducing carbon emissions. Therefore, the above numerical results demonstrate that the proposed algorithm is sustainable and effective.

VII. CONCLUSION

In this paper, we studied how GAI empowers DTs in ICPSs. Specifically, we designed a GAI-driven DT architecture in ICPSs from the pipeline of DT construction, including real-time physical data collection, communications for DTs, DT modeling and maintenance, and DT decision-making. To motivate IIoT devices to contribute sensing data for GAI-empowered DT construction, we proposed a contract theory model under information asymmetry. Furthermore, we proposed a sustainable diffusion-based SAC algorithm to generate the optimal feasible contract, where we utilized the dynamic structured pruning technique to sparsify actor networks of GDMs, allowing efficient implementation of the proposed algorithm in ICPSs. Finally, numerical results demonstrate the effectiveness and sustainability of the proposed approaches.

[1]https://github.com/mlco2/codecarbon

For future work, we will propose a more general and systematic architecture of GAI-empowered DT, and explore combining state-of-the-art techniques with DRL algorithms for optimal contract design.

References

[1] K. Zhang, Y. Shi, S. Karnouskos, T. Sauter, H. Fang, and A. W. Colombo, "Advancements in industrial cyber-physical systems: An overview and perspectives," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 716–729, 2022.

[2] J. Wen, J. Nie, J. Kang, D. Niyato, H. Du, Y. Zhang, and M. Guizani, "From generative AI to generative Internet of Things: Fundamentals, framework, and outlooks," *IEEE Internet of Things Magazine*, vol. 7, no. 3, pp. 30–37, 2024.

[3] J. Chen, Y. Shi, C. Yi, H. Du, J. Kang, and D. Niyato, "Generative AI-driven human digital twin in IoT-healthcare: A comprehensive survey," *arXiv preprint arXiv:2401.13699*, 2024.

[4] X. Huang, H. Yang, C. Zhou, X. Shen, and W. Zhuang, "When digital twin meets generative AI: Intelligent closed-loop network management," *arXiv preprint arXiv:2404.03025*, 2024.

[5] I. Lotfi, D. Niyato, S. Sun, D. I. Kim, and X. Shen, "Semantic information marketing in the metaverse: A learning-based contract theory framework," *IEEE Journal on Selected Areas in Communications*, vol. 42, DOI 10.1109/JSAC.2023.3345402, no. 3, pp. 710–723, 2024.

[6] H. Du, Z. Li, D. Niyato, J. Kang, Z. Xiong, H. Huang, and S. Mao, "Diffusion-based reinforcement learning for edge-enabled AI-generated content services," *IEEE Transactions on Mobile Computing*, DOI 10.1109/TMC.2024.3356178, pp. 1–16, 2024.

[7] G. Fang, X. Ma, and X. Wang, "Structural pruning for diffusion models," *Advances in neural information processing systems*, vol. 36, 2024.

[8] L. Jiang, H. Zheng, H. Tian, S. Xie, and Y. Zhang, "Cooperative federated learning and model update verification in blockchain-empowered digital twin edge networks," *IEEE Internet of Things Journal*, vol. 9, no. 13, pp. 11 154–11 167, 2021.

[9] M. Bordukova, N. Makarov, R. Rodriguez-Esteban, F. Schmich, and M. P. Menden, "Generative artificial intelligence empowers digital twins in drug discovery and clinical trials," *Expert Opinion on Drug Discovery*, vol. 19, no. 1, pp. 33–42, 2024.

[10] J. Kang, Y. Zhong, M. Xu, J. Nie, J. Wen, H. Du, D. Ye, X. Huang, D. Niyato, and S. Xie, "Tiny multi-agent DRL for twins migration in UAV metaverses: A multi-leader multi-follower stackelberg game approach," *IEEE Internet of Things Journal*, DOI 10.1109/JIOT.2024.3360183, pp. 1–1, 2024.

[11] W. Su, Z. Li, M. Xu, J. Kang, D. Niyato, and S. Xie, "Compressing deep reinforcement learning networks with a dynamic structured pruning method for autonomous driving," *IEEE Transactions on Vehicular Technology*, DOI 10.1109/TVT.2024.3399826, pp. 1–15, 2024.

[12] J. Yu, A. Alhilal, P. Hui, and D. H. Tsang, "Bi-directional digital twin and edge computing in the metaverse," *IEEE Internet of Things Magazine*, vol. 7, no. 3, pp. 106–112, 2024.

[13] G. Liang, J. Hu, K. Yang, S. Song, T. Liu, N. Xie, and Y. Yu, "Data augmentation for predictive digital twin channel: Learning multi-domain correlations by convolutional timeGAN," *IEEE Journal of Selected Topics in Signal Processing*, vol. 18, DOI 10.1109/JSTSP.2024.3358980, no. 1, pp. 18–33, 2024.

[14] J. Kang, J. Wen, D. Ye, B. Lai, T. Wu, Z. Xiong, J. Nie, D. Niyato, Y. Zhang, and S. Xie, "Blockchain-empowered federated learning for healthcare metaverses: User-centric incentive mechanism with optimal data freshness," *IEEE Transactions on Cognitive Communications and Networking*, vol. 10, DOI 10.1109/TCCN.2023.3316643, no. 1, pp. 348–362, 2024.

[15] Z. Wang, J. J. Hunt, and M. Zhou, "Diffusion policies as an expressive policy class for offline reinforcement learning," *arXiv preprint arXiv:2208.06193*, 2022.

[16] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*, pp. 1861–1870. PMLR, 2018.

Jinbo Wen received the B.Eng. degree from Guangdong University of Technology, China, in 2023. He is currently pursuing an M.S. degree with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China. His research interests include Generative AI, edge intelligence, and metaverse.

Jiawen Kang received the Ph.D. degree from Guangdong University of Technology, China, in 2018. He has been a postdoc at Nanyang Technological University, Singapore from 2018 to 2021. He is currently a full professor at Guangdong University of Technology, China. His research interests mainly focus on blockchain, security, and privacy protection in wireless communications and networking.

Dusit Niyato (Fellow, IEEE) is a professor in the School of Computer Science and Engineering, at Nanyang Technological University, Singapore. He received the B.Eng. degree from King Mongkuts Institute of Technology Ladkrabang (KMITL), Thailand in 1999 and the Ph.D. in Electrical and Computer Engineering from the University of Manitoba, Canada in 2008. His research interests are in the areas of sustainability, edge intelligence, decentralized machine learning, and incentive mechanism design.

Yang Zhang is currently an associate professor in the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China. He received the B.Eng. and M.Eng. from Beihang University in 2008 and 2011, respectively. He obtained the Ph.D. degree in Computer Engineering from Nanyang Technological University, Singapore, in 2015. His current research topic is edge computing and multi-agent unmanned systems.

Shiwen Mao (Fellow, IEEE) received his Ph.D. in electrical and computer engineering from Polytechnic University, Brooklyn, NY. He is a Professor and Earle C. Williams Eminent Scholar and Director of the Wireless Engineering Research and Education Center at Auburn University. His research interests include wireless networks and multimedia communications. He is a Distinguished Lecturer of the IEEE Communications Society and IEEE Council of RFID. He is the Editor-in-Chief of IEEE Transactions on Cognitive Communications and Networking.