

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Ahmad, Sohail, Muhammad Tariq, Mian Ahmad Jan, and Houbing Song. "Reconfigurable Intelligent Surfaces Assisted 6G Communications for Internet of Everything." IEEE Internet of Things Journal, 2023, 1–1. <https://doi.org/10.1109/JIOT.2023.3339531>.

<https://doi.org/10.1109/JIOT.2023.3339531>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Reconfigurable Intelligent Surfaces Assisted 6G Communications for Internet of Everything

Sohail Ahmad, *Member, IEEE*, Muhammad Tariq, *Senior Member, IEEE*, Mian Ahmad Jan, *Senior Member, IEEE*, Houbing Song, *Fellow, IEEE* .

Abstract—The dynamic evolution of wireless communication, driven by the Internet of Everything (IoE) and the envisioned 6G networks, presents both challenges and opportunities. IoE expands beyond IoT, encompassing diverse devices, human interactions, and environmental elements. To unlock IoE's vast potential and harness the capabilities of 6G, we propose a framework named ISRiD, which assimilates three key components: Integrated Sensing and Communication (ISAC), Reconfigurable Intelligent Surfaces (RIS), and Deep Deterministic Policy Gradient (DDPG). This creates an adaptive, data-driven communication framework for IoE and 6G. ISAC enhances situational awareness, RIS optimizes signal paths, and DDPG adds intelligence. This empowers devices to collect environmental data for optimization and intelligent decisions. By leveraging real-time sensor data, ISRiD optimizes communication protocols, significantly improving efficiency and reliability in wireless networks. Despite challenges, this approach equips IoE, including the advancements brought by 6G, to meet evolving demands, as validated by empirical experiments and simulations.

Index Terms—Integrated sensing and communications, reconfigurable intelligent surfaces, deep deterministic policy gradient, internet of everything, 6G.

I. INTRODUCTION

The landscape of wireless communication is undergoing a profound transformation due to the emergence of the Internet of Everything (IoE). IoE signifies the culmination of the digital revolution, where every facet of our lives is intricately interconnected through a vast and dynamic network. It transcends the domain of the Internet of Things (IoT), encompassing not only smart devices but also the synergy between human interaction, devices, systems, and the environment itself. IoE envisions a world where data flows continuously, intelligence is ubiquitous, and every aspect of our daily lives benefits from the symbiotic relationship between the physical and digital realms [1]–[4]. However, realizing IoE's promise is a complex endeavor. IoE networks are inherently heterogeneous, accommodating a diverse array of devices,

each with unique communication requirements. Moreover, the wireless environment in which these devices operate is far from static. It's characterized by ever-changing channel conditions, the presence of interference sources, and the continuous ingress and egress of devices.

Traditional communication systems, originally designed for static and homogeneous environments, falter in the face of the multifaceted challenges posed by IoE [5]. Numerous traditional optimization techniques [6]–[11] have been presented in prior works, primarily designed for static Reconfigurable Intelligent Surfaces (RIS) networks, assuming well-known channel conditions. These methods exhibit lower computational demands and provide effective solutions in such scenarios. However, their suitability diminishes when applied to dynamic environments with Integrated Sensing and Communication (ISAC), where the complexity of these techniques escalates significantly. In contrast, Deep Deterministic Policy Gradient (DDPG) stands out as a remarkably versatile approach, ideally suited for dynamic and ever-changing settings like ISAC-assisted RIS networks. It excels in real-time decision-making and proves its mettle in managing intricate, large-scale networks. Despite the potential for higher training complexity, DDPG remains a robust choice in such scenarios.

To address these challenges and unlock the full potential of IoE, the integration of ISAC, RIS, and the intelligent decision-making capabilities of DDPG becomes crucial. ISAC [12] integrates real-time sensing capabilities with communication systems, enhancing the network's situational awareness by collecting and processing data about the wireless environment. RIS [2], [13] dynamically optimizes signal paths by manipulating the phase of electromagnetic waves, enabling adaptability to changing channel conditions and interference mitigation. DDPG [14] brings intelligence to the decision-making process by optimizing communication channels using sensor data in real-time.

This paper embarks on a comprehensive exploration of an architectural framework in which ISAC, RIS, and DDPG converge to create an intelligent, adaptive, and data-driven IoE communication ecosystem known as ISRiD. This framework enhances the performance and efficiency of wireless communication systems by blending sensing and communication functionalities. Sensing devices collect environmental data, which is used to optimize network operations, manage resources, and make intelligent decisions. The integration of sensing capabilities into the communication system enables

Manuscript received on Aug. 1, 2023. In the revised form on Oct. 6, and Nov. 09, 2023.

Sohail Ahmad and Muhammad Tariq (Corresponding Author) are with the Department of Electrical Engineering, National University of Computer and Emerging Sciences, Islamabad 44000, Pakistan. (E-mail: p191601@nu.edu.pk, mtariq@princeton.edu).

Mian Ahmad Jan is with the Department of Computer Science, College of Computing and Informatics, University of Sharjah, Sharjah, 27272, United Arab Emirates (e-mail: mjan@sharjah.ac.ae).

Houbing Song is with the University of Maryland, Baltimore County, USA (email: h.song@ieee.org)

the network to adapt and optimize its operation based on real-time environmental data. This leads to improved efficiency, reliability, and data rates in wireless communication. The aim is to optimize wireless communication systems by leveraging the capabilities of ISAC, RIS, and DDPG in the context of 6G IoE scenarios [15], [16]. However, autonomously controlling the phase shifts of numerous RIS elements independently, based on ISAC data, presents a formidable challenge.

Through rigorous experimentation and meticulous simulations, this paper demonstrates how the proposed framework enhances wireless communication performance, improves adaptability, and optimizes resource allocation. DDPG is trained using collected data and advanced techniques, enhancing the stability, convergence, and exploration capabilities of the learning process. The trained DDPG models are seamlessly integrated into the ISAC-assisted RIS system for real-time intelligent decision-making. This system uses sensor data to make predictions on wireless signal strength, channel conditions, and other relevant parameters. These predictions are then used to dynamically adjust the reflection coefficients of the RIS elements or optimize communication protocols, resulting in improved wireless communication performance in 6G IoE scenarios. The main contributions of this work involve:

- enhancing adaptability and optimization in wireless communication systems through the integration of DDPG into the ISAC-assisted RIS as ISRD framework,
- leveraging real-time sensor data for training DDPG models and seamlessly incorporating these models into the system, and
- improving wireless communication performance in 6G IoE, adapting to changing environments, and efficiently allocating network resources, thereby validating the effectiveness and advantages of the proposed approach.

The remainder of this paper is organized as follows: Section II describes ISAC-assisted RIS modeling for the IoE environment. Section III aims at the problem formulation. Section IV details the methodology used in our study, including the data collection, RIS phase shift, and DDPG model design. Section V discusses the interaction and working of ISRD. Section VI presents the simulation results and performance evaluation. Finally, Section VII concludes the paper.

II. ISAC-ASSISTED RIS MODELING FOR IOE ENVIRONMENT

A. Network Model

In the context of this investigation, we are exploring a particular scenario where the direct line-of-sight (LoS) connection between the base station (BS) and the users experiences signal weakening due to fading effects, resulting in a non-line-of-sight (NLoS) condition. Consequently, users receive the signal with the assistance of a RIS. The adjustment of the phase shift by each discrete element within the RIS is utilized to alter the path of electromagnetic (EM) waves. This phase shift, represented as θ , is determined by the following relationship:

$$\theta = 2\pi(\Delta d/\lambda) \quad (1)$$

Here, Δd represents the displacement of the reflecting element from its original position, while λ corresponds to the wavelength of the incident signal. The combined strength of the reflected signal at a specific spatial point is the sum of contributions from all individual elements, defined as:

$$E = \sum (\beta e^{j\theta_N}) \quad (2)$$

In this context, β , θ_N , and $e^{j\theta_N}$ denote the reflection amplitude (either 0 or 1), the phase shift of the N reflecting elements, and the reflection coefficients, respectively. The phase shifts associated with the N RIS elements are concisely represented using a diagonal matrix Φ :

$$\Phi = \text{diag}(\beta e^{j\theta_1}, \beta e^{j\theta_2}, \dots, \beta e^{j\theta_N}) \quad (3)$$

This explanation emphasizes that achieving the optimal configuration of the RIS involves individually adjusting phase shifts across its reflective components.

B. Channel Model

The channel modeling involves capturing the interactions between the RIS elements, the BS, and the users. The channel model typically considers the channel gain and phase shifts introduced by the RIS elements. The channel gain from the BS to a user through the RIS is affected by the phase shifts introduced by the RIS elements. Let $h_{\text{BS-RIS}}$ be the complex channel gain between the BS and RIS and the user and can be modeled as follows:

$$h_{\text{BS-RIS}} = \sqrt{\beta_{\text{BS-RIS}}} \cdot g_{\text{BS-RIS}} \quad (4)$$

where $\beta_{\text{BS-RIS}}$ is the path loss factor and $g_{\text{BS-RIS}}$ represents the small-scale fading i.e rayleigh or rician fading.

For k users, (4) is formulated as:

$$h_{\text{BS-RIS}_k} = \sqrt{\beta_{\text{BS-RIS}}} \cdot g_{\text{BS-RIS}_k} \quad (5)$$

Similarly, for the RIS-Users channel, (5) is formulated as:

$$h_{\text{RIS-User}_k} = \sqrt{\beta_{\text{RIS-User}_k}} \cdot g_{\text{RIS-User}_k} \quad (6)$$

The path loss factors can be defined as:

$$\beta_{\text{BS-RIS}_k} = \frac{C_{\text{BS-RIS}}}{d_{\text{BS-RIS}_k}^{\gamma_n}} \quad (7)$$

and

$$\beta_{\text{RIS-User}_k} = \frac{C_{\text{RIS-User}}}{d_{\text{RIS-User}_k}^{\gamma_n}} \quad (8)$$

where $C_{\text{BS-RIS}}$ and $C_{\text{RIS-User}}$ are constants, $d_{\text{BS-RIS}_k}$ and $d_{\text{RIS-User}_k}$ are distances between the BS, RIS and user k , while γ_n is the path loss exponent.

The overall channel gain $h_{\text{BS-User}_k}$ for k users can be expressed as:

$$h_{\text{BS-User}_k} = h_{\text{BS-RIS}_k} \cdot \Phi \cdot h_{\text{RIS-User}_k} \quad (9)$$

where Φ represents the diagonal matrix of phase shifts introduced by the RIS elements as in (3).

Then the received signal y_k at user k can be represented as:

$$y_k = h_{\text{BS-RIS},k}^H \cdot \Phi \cdot h_{\text{RIS-User},k}^H \cdot s_k + n_k \quad (10)$$

where s_k is the transmitted signal from the BS and n_k is the additive white Gaussian noise.

The SINR _{k} at the user's receiver can be defined as:

$$\text{SINR}_k = \frac{|h_{\text{BS-RIS},k} \Phi h_{\text{RIS-User},k}|^2}{\sum_{j \neq k} |h_{\text{BS-RIS},k} \Phi h_{\text{RIS-User},j}|^2 + \sigma^2} \quad (11)$$

where SINR _{k} is the signal-to-interference plus noise ratio, and σ^2 is the noise variance.

Then the achievable rate for user k can be calculated using the Shannon formula:

$$R_k = \log_2(1 + \text{SINR}_k) \quad (12)$$

The RIS elements' phase shifts are crucial in optimizing signal propagation and enhancing network performance by using these mathematical representations to design and optimize the RIS configurations to achieve desired network outcomes.

III. PROBLEM FORMULATION

Consider a 6G wireless communication network with multiple users and a single BS, assisted by an RIS. The objective is to optimize the RIS phase shift configuration to enhance the network performance, including data rate, accuracy, convergence, and delay.

The optimization challenge can be expressed as:

Maximize:

$$R = \sum_{k=1}^K R_k \quad (13)$$

Subject to:

$$\begin{aligned} 0 \leq \theta_i \leq 2\pi, \forall i & \text{ (phase shift constraint)} \\ R_k \geq R_{\min}, \forall k & \text{ (data rate constraint)} \\ P \leq P_{\max} & \text{ (power constraint)} \end{aligned}$$

$$R_k = \log_2(1 + \text{SINR}_k) \quad (14)$$

$$\text{SINR}_k = \frac{|h_{\text{BS-RIS},k} \Phi h_{\text{RIS-User},k}|^2}{\sum_{j \neq k} |h_{\text{BS-RIS},k} \Phi h_{\text{RIS-User},j}|^2 + \sigma^2} \quad (15)$$

The primary goal is to identify the optimal phase shift matrix, denoted as Φ , to maximize the overall data rate through the dynamic optimization of signal paths. In the realm of theoretically-based static RIS-assisted networks, various traditional optimization techniques, including AO and Semi Definite Relaxation (SDR), have been proposed in existing literature for designing these optimal phase shift matrices. These techniques are appreciated for their simplicity and minimal computational requirements. Nonetheless, they encounter limitations when applied to dynamic and adaptable ISAC-assisted RIS networks. These networks necessitate real-time adaptability, and as their scale increases, alongside the growing number of reflecting elements in the RIS, the optimization problem becomes increasingly intricate. In such scenarios, the effectiveness of

AO diminishes. Conversely, the DDPG algorithm excels in these dynamic contexts. DDPG is a model-free approach with several advantages, such as low computational complexity, rapid decision-making, and adaptability to changing network conditions. It is ideally suited for real-time decision-making, adept at maneuvering through dynamic environments, and capable of adjusting to evolving conditions within ISAC-assisted RIS networks. Consequently, ISRiD substantially enhances the optimization process.

IV. PROPOSED ISRiD ARCHITECTURE

The proposed ISRiD architecture incorporates DDPG into the ISAC-assisted RIS communication, which represents a compelling and innovative approach to enhancing wireless communication. ISRiD amalgamates the benefits of ISAC and RIS technologies with the decision-making prowess of DDPG algorithms, ultimately leading to the optimization of wireless communication performance and the achievement of advanced connectivity in future networks. Below, we delve into the core components of this architecture, as depicted in Figure 1.

A. Integrated Sensing and Communication

The architecture capitalizes on the ISAC concept, which harmonizes sensing capabilities with communication systems. It involves the utilization of real-time data from sensors, such as IoE devices or environmental sensors, to actively gather up-to-the-minute information concerning the wireless network environment. This data encompasses details regarding channel conditions, interference levels, and other pertinent parameters. Subsequently, the collected data undergoes processing and the extraction of key features. This process yields a comprehensive understanding of the wireless environment. Once the sensor data has been adequately processed, it is integrated into the communication system via the ISAC framework. This infusion of real-time information equips the network to adapt to ever-changing surroundings and acquire situational awareness, empowering it to make intelligent decisions based on the dynamic wireless landscape. The outcomes of this integration are transformative. The network gains the ability to dynamically optimize its communication protocols, effectively respond to an adaptive configuration, deftly mitigate interference, extend its coverage, autonomously self-heal network issues, and efficiently allocate resources. This level of sophistication elevates the network from being reactive to proactive, enabling it to make intelligent decisions based on the dynamic landscape of wireless communication [17].

B. Optimization and Adaptive Control of RIS Elements

This network architecture has been devised to enhance wireless communication performance by dynamically tuning the reflection coefficients of the RIS based on sensing information. The principal function of the RIS within this network is to refine the wireless connections between the base station and user devices. This is achieved by fine-tuning the reflected signals to minimize interference, maximize signal strength,

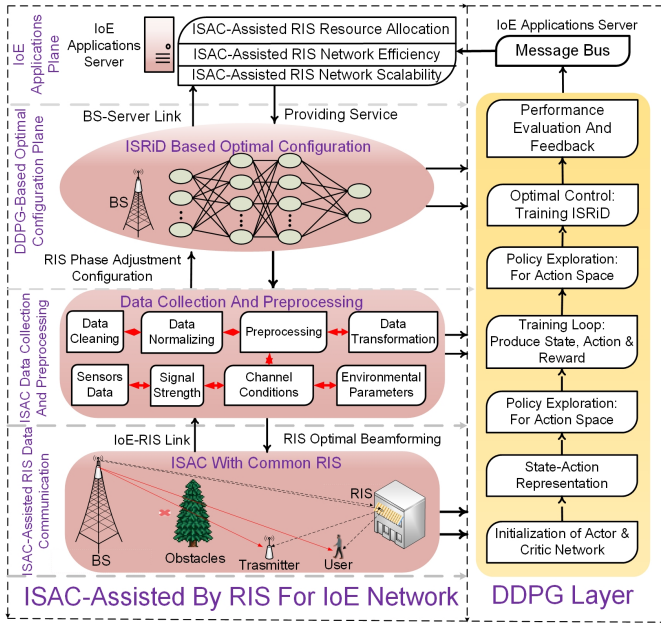


Figure 1: ISAC-Assisted RIS for 6G IoE

and accommodate fluctuations in the channel. It dictates how the RIS should configure its reflection coefficients to amplify wireless communication. The proposed approach is capable of real-time adaptation to shifting conditions; for instance, when a new user device joins the network or if the environment encounters interference, the system promptly adjusts the RIS's behavior to optimize the communication link [18].

C. Real-time Decision-Making Using DDPG

The ISRID framework uses DDPG, employing deep neural networks (DNNs) to optimize RIS reflection coefficients based on sensor data. DDPG adapts over time, suitable for tasks like robotic control and autonomous vehicles. RIS dynamically adjusts reflection coefficients, optimizing the wireless channel for better signal quality and less interference. This integration achieves higher data rates, adaptability, and resource allocation. The DDPG model is seamlessly integrated, making real-time decisions based on wireless conditions. It continuously refines its policies, forming a feedback loop with the wireless environment to enhance communication. The ISAC-assisted RIS system facilitates real-time intelligent decision-making.

V. INTERACTION AND WORKING OF ISRID LEVERAGING DDPG

A. Key Idea

A mathematical framework is established as in equation (10), by setting the foundation, upon which the ISAC functions to describe signal reflection by simulating how varying phase shifts (θ_i) within the diagonal matrix (Φ) influence signals of multiple sensors, implantable devices, to integrates various sensor data. The primary objective involves identifying the optimal phase shifts (θ_i) that maximize the desired outcome. To

address this optimization task, a DDPG network is employed to predict the optimal phase shifts based on input parameters. Let x denote the input data encompassing the ISAC sensor's position and channel conditions. The DDPG learns the intricate relationship between x and the optimal phase shifts (θ_i^*), yielding the expression:

$$\theta_i^* = DDPG(x) \quad (16)$$

The ISAC, in a continuous manner, updates its emulated RIS configuration utilizing the anticipated optimal phase shifts (θ_i^*) from the DDPG. This dynamic process empowers the ISAC to furnish real-time insights into the implications of diverse RIS configurations on network performance. As the DDPG is exposed to additional real-world data and refines its internal parameters, it progressively improves its predictions concerning optimal phase shifts. The ISAC then adapts to these refined forecasts, resulting in simulations that align more closely with real-world scenarios, ultimately enhancing their accuracy. The proposed ISRID model comprises dynamic equations that simulate the behavior of the IoE system, providing insights into changing conditions. The input sensed data and output data can be represented in equation form as follows:

Let I_t represents the sensed input at a time t , collected from various sensors deployed in the network and can be mathematically expressed as:

$$I(t) = f_{\text{sensing}}(S_1(t), S_2(t), \dots, S_n(t)) \quad (17)$$

where $(S_1(t), S_2(t), \dots, S_n(t))$ are the individual sensor readings at time t , and f_{sensing} is a function that processes and integrates these sensor inputs into a comprehensive sensed data vector $I(t)$.

Let $O(t)$ represents the output data or actions taken by the ISAC network in response to the sensed data. This can include communication decisions, control actions, or any other network responses. In equation form:

$$O(t) = f_{\text{processing}}(I(t)) \quad (18)$$

where $f_{\text{processing}}$ is a function that processes the sensed data $I(t)$ to determine the appropriate network actions or responses, resulting in the output data $O(t)$. These equations capture the relationship between input sensed data and the network's output data within an ISAC system. The architecture features an array of reconfigurable intelligent reflecting elements forming RIS. Each element can dynamically adjust its phase shift to influence signal propagation paths. The phase shifts of the RIS elements are represented by a vector $\theta = [\theta_1, \theta_2, \dots, \theta_N]$ where N is the number of elements. The RIS manipulates the electromagnetic wave propagation using these phase shifts. The total reflection coefficient Γ for a given incident wave ϕ_{inc} can be expressed as:

$$\Gamma = \sum_{n=1}^N r_{\text{element}} e^{j\theta_n} \quad (19)$$

Where r_{element} is the reflection coefficient of each RIS element and j is the imaginary unit. The DDPG algorithm [14] is

employed to optimize the phase shifts of the RIS elements in real-time. The actor-network $\mu(s|\theta^\mu)$ takes the current state from the ISAC-assisted RIS network and outputs optimal phase shifts. The critic network $Q(s, a|\theta^Q)$ evaluates the Q-value of state-action pairs to guide optimization. The actor network maps the state s to the action a , while the critic network estimates the Q-value of the state-action pair. The actor network is updated based on the deterministic policy gradient:

$$\nabla_{\theta^\mu} J \approx \mathbb{E}_{s_t \sim \rho} \left[\nabla_a Q(s, a|\theta^Q) \Big|_{s=s_t, a=\mu(s_t)} \nabla_{\theta^\mu} \mu(s|\theta^\mu) \Big|_{s_t} \right] \quad (20)$$

The critic network is updated using the Bellman equation:

$$y_t = r_t + \gamma Q' \left(s_{t+1}, \mu' \left(s_{t+1} | \theta^{\mu'} \right) \Big| \theta^{Q'} \right) \quad (21)$$

$$\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'} \quad (22)$$

$$\theta^Q \leftarrow \theta^Q - \eta \sum_i \nabla_{\theta^Q} [Q(s_t, a_t | \theta^Q) - y_t]^2 \quad (23)$$

Where J is the performance metric to be optimized, ρ represents the state distribution, γ is the discount factor, and α is the learning rate. DDPG employs target policy and target value networks to ensure stability during updates. To encourage exploration and enhance the policy's ability to discover improved actions and policies, exploration noise is incorporated into the policy's actions during training.

The policy network undergoes updates through deterministic policy gradient ascent, as formulated by the subsequent equation.

$$\nabla_{\theta_i} J_i(\theta_i, \phi_i) = \mathbb{E} \left[\nabla_{\theta_i} \mu_{\theta_i}(S_i(t)) \nabla_a Q_{\phi_i}(S_i(t), a) \Big| a = \mu_{\theta_i}(S_i(t)) \right] \quad (24)$$

For the value network, updates are performed using a temporal difference (TD) error. The target networks are updated gradually through soft updates to ensure they closely track the learned networks:

$$\theta'_i \leftarrow \tau \theta_i + (1 - \tau) \theta'_i \quad (25)$$

$$\phi'_i \leftarrow \tau \phi_i + (1 - \tau) \phi'_i \quad (26)$$

Training continues across multiple episodes until the policy network converges to an optimal policy.

B. Training

The framework relies on real-world network data, especially the quality of the training dataset representing network behavior. ISAC should cover scenarios with adverse impacts like link failures, misconfigurations, and congestion. The training dataset is generated within specific network environments. Training involves iterative dataset processing, forward-backward propagation, and gradient descent to minimize loss and improve predictions. During training, the DNN model predicts dynamic network attributes (e.g., user locations and

channel conditions). The model's architecture is detailed in Table 1. Further details on training and learning are in the following sections.

C. Prediction

Throughout this procedure, the training network effectively assimilates training data sourced from the physical model, and the pre-existing network configuration. By leveraging these diverse data streams from the database, the training network dynamically tunes its weights and biases, employing stochastic gradient descent and back-propagation. When the network achieves full training, the framework transitions to the prediction phase. The trained control policies (ϕ_i) are subsequently employed in real-time for making adaptive and optimal control decisions. These decisions are based on the observed ISAC states and the predicted RIS configurations. This process can be expressed as:

$$\nabla_{\phi_i} J_i(\theta_i, \phi_i) = E [Q_{\phi_i}(S_i(t), C_i(t)) - (R_i(t) + \gamma Q_{\phi_i}(S_i(t+1), \mu_{\theta_i}(S_i(t+1))))] \nabla_{\phi_i} Q_{\phi_i}(S_i(t), C_i(t)) \quad (27)$$

However, when any updates pertaining to the network environment occur, the proposed solution seamlessly shifts back to the learning phase, detailed in Algorithm 1. The computational

Algorithm 1 ISRiD: ISAC Assisted RIS Leveraging DDPG

```

Initialize ISRiD environment, actor-network  $\theta_\mu$ , critic network  $\theta_Q$  with random weights target actor network  $\theta_{\mu'}$ , target critic network  $\theta_{Q'}$  with same weights as  $\theta_\mu$  and  $\theta_Q$ , replay buffer  $D$  exploration noise  $\epsilon$ , learning rates  $\alpha$  (actor),  $\beta$  (critic), discount factor  $\gamma$ , and soft update parameter  $\tau$ 
for episode = 1 to Max Episodes do
    Initialize environment and set initial state  $s_0$ 
    for  $t = 1$  to Max Steps do
        Select action  $a_t = \mu(s_t | \theta_\mu) + \epsilon$ 
        Apply action  $a_t$  to sensed data environment, receive reward  $r_t$  and next state  $s_{t+1}$ 
        Store transition  $(s_t, a_t, r_t, s_{t+1})$  in replay buffer  $D$ 
        if  $D$  contains enough samples then
            Sample random mini-batch of transitions  $(s_i, a_i, r_i, s_{i+1})$  from  $D$ 
            Compute target Q-value:
                 $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} | \theta_{\mu'}) | \theta_{Q'})$ 
            Compute critic loss:  $L = (Q(s_i, a_i | \theta_Q) - y_i)^2$ 
            Update  $\theta_Q$ :  $\theta_Q \leftarrow \theta_Q - \beta \cdot \nabla_{\theta_Q} L$ 
            Compute actor loss:  $J = -Q(s_i, \mu(s_i | \theta_\mu) | \theta_Q)$ 
            Update  $\theta_\mu$ :  $\theta_\mu \leftarrow \theta_\mu + \alpha \cdot \nabla_{\theta_\mu} J$ 
            Soft update target networks:
                 $\theta_{\mu'} \leftarrow \tau \theta_\mu + (1 - \tau) \theta_{\mu'}$   $\theta_{Q'} \leftarrow \tau \theta_Q + (1 - \tau) \theta_{Q'}$ 
        end if
        Update state:  $s_t = s_{t+1}$ 
    end for
end for

```

complexity of the DDPG within the ISAC-assisted RIS network can be estimated as $O(S + A + B + I)$, where $O()$ represents

Table I: Details of hyper and system parameters

Parameters	Values
Discount rate, γ	0.99
Decay rate (DR) of actor-network, λ_a	0.00001
DR of critic network, λ_c	0.00001
Learning rate (LR) of actor target network, τ_a	0.001
LR of critic target network, τ_c	0.001
LR of actor training network, μ_a	0.001
LR of critic training network, μ_c	0.001
Episodes no., S	5000
Number of iterations, T	1000
Buffer size, D	100000
Realizations	500
Transmit Power P_t ,	5-20 mW
BS antenna elements, M	30
Number of IoE devices, I	10
Maximum time steps, Z	64
Bandwidth, B in MHz	100
Experiences number in the mini-batch	16
Reflecting elements N of RIS,	10-100

the Big O notation, and S , A , B , and T refer to the dimensions of the state space, action space, experience replay buffer, and the number of training iterations, respectively.

VI. SIMULATION RESULTS AND ANALYSIS

In this section, we delve into simulation results that scrutinize the computational complexity, resource allocation efficiency, and maximum achievable rate. We conduct this analysis by making performance comparisons with the AO scheme [6], [7], [13], [18], [19], and the ISAC (without RIS) scenario. The simulations are conducted using experimental data, and the setup encompasses a BS and an RIS arranged in a uniform linear array and a rectangular array, respectively. The vertical reflecting elements in the RIS array (N_y) remain fixed at 10, while the number of horizontal reflecting elements (N_x) varies. The signal attenuation for the BS-RIS, RIS-IoE, and BS-IoE channels is consistently set at 30dB, with a reference distance denoted as $D_0 = 1m$. In this setup, the BS-IoE and BS-RIS channels are characterized as Rayleigh faded and Rician faded, respectively, thereby replicating real-world wireless propagation conditions. The horizontal distance between the BS and IoE devices, referred to as BS-IoE, is defined as $d = 50m$. For our analysis, we employ a 64-QAM modulation scheme, implemented using the Keras library with the TensorFlow backend. Our objective is to evaluate the robustness of the trained DDPG policy. We subject it to various tests by altering wireless channel conditions. This assessment aims to verify that the policy consistently maintains stable and high performance. We scrutinize how performance scales with the expansion of sensor data size, the number of IoE devices, variations in transmit power levels, and changes in the number of elements of RIS and BS.

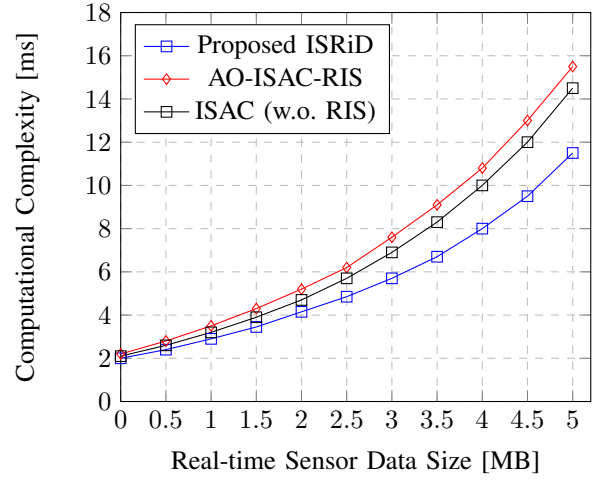


Figure 2: Computational Complexity Vs. Data Size

A. Impact of Increasing Real-time Sensor Data on Computational Complexity

In Fig. 2, we investigate how increasing real-time sensor data impacts computational complexity. We compare the proposed ISRiD with AO and ISAC (without RIS) to understand the advantages of using DDPG for optimizing the network. The results demonstrate that as real-time sensor data increases, both DDPG and AO algorithms experience gradually rising computational complexity, leading to longer execution times. Notably, DDPG exhibits slightly better computational efficiency than AO in dynamic, real-world environments with factors like more reflecting elements, increased resource needs, and more complex optimization problems. This suggests that DDPG excels in managing larger data sizes due to its real-time adaptability and learning capabilities, making it well-suited for such dynamic networks. Moreover, the computational complexity in the ISAC (without RIS) scenario is significantly lower compared to ISRiD. This is because optimizing a large number of RIS elements in real time to adapt to changing environmental conditions requires more computational resources. Hence, DDPG offers computational advantages in such dynamic scenarios with increasing real-time sensor data, highlighting its potential to enhance ISAC-assisted RIS networks.

B. Impact of Increasing IoE Devices on Resource Allocation

In Fig. 3, the impact of the ISRiD framework is demonstrated, providing real-time insights into the IoE environment. This integration empowers the decision-making process for resource allocation by optimizing phase shifts in the RIS. The figure shows an interesting trend: as the count of IoE devices increases, there is a corresponding improvement in resource allocation efficiency. However, this upward trend gradually levels off as the number of IoE devices reaches higher levels. This deceleration is a result of the increasing complexity and heightened demands placed on the network. In comparison to scenarios without RIS and those with AO-ISAC-RIS, the rate of decline in resource allocation efficiency is more gradual

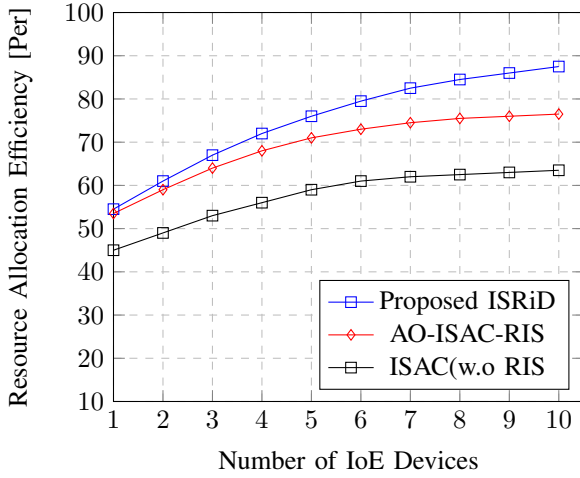


Figure 3: Resource Allocation Efficiency Vs. IoE Devices

with the proposed approach. This distinction arises from the unique capability of the proposed approach to adapt in real-time, guided by insights, achieving 87 percent (max) efficiency, derived from the ISRiD framework. Conversely, AO-ISAC-RIS and ISAC (without RIS) may face challenges in keeping pace with the rapid changes in the network, resulting in less efficient resource allocation, achieving 76.5 percent (max) and 63 percent (max) efficiency, respectively. These findings highlight the continuous refinement of resource allocation policies by the proposed approach, driven by the evolving network conditions. This process ensures that resources are distributed more effectively as the number of devices increases, showcasing the well-optimized and adaptable nature of the system, and thereby enhancing its overall efficiency.

C. Impact of Increasing Transmit Power on Max. Rate

In Fig. 4, we analyze the performance of the data rate for an increasing transmit power of the proposed ISRiD framework, and we compare it with baseline scenarios, including ISAC-RIS with AO and ISAC (without RIS) [13]. The study investigates how the data rate changes with varying transmit power at the BS and compares it with the benchmark schemes. The results clearly indicate that the proposed scheme outperforms the benchmarks significantly in terms of data rate performance, achieving 50.5 bps/Hz. In comparison, the data rates reached 40 bps/Hz for the AO-ISAC-RIS and 24.5 bps/Hz for the ISAC with RIS cases. The proposed algorithm leverages past experiences, running with the learned DDPG policy, to improve power efficiency and overall data rate performance.

D. Impact of Increasing RIS Elements (N) Versus Increasing BS Elements (M) on Max. Achievable Rate

Fig. 5 investigates the impact of increasing reflecting elements of RIS compared to the increasing number of active BS elements on the maximum achievable rate. The results use the ISAC (without RIS) as a reference point. The analysis reveals that when N is small, the ISRiD framework achieves

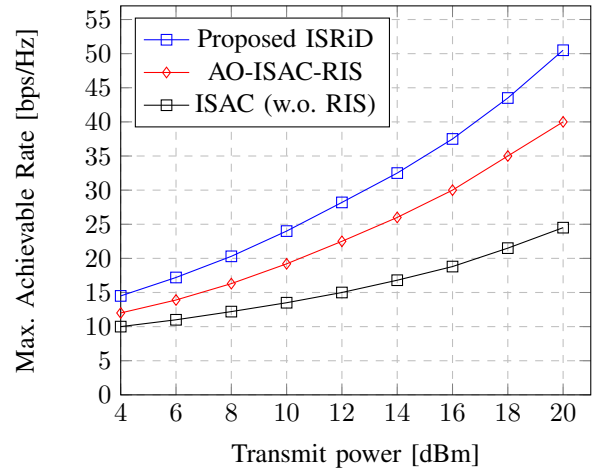


Figure 4: Max. Achievable Rate Vs. Transmit Power

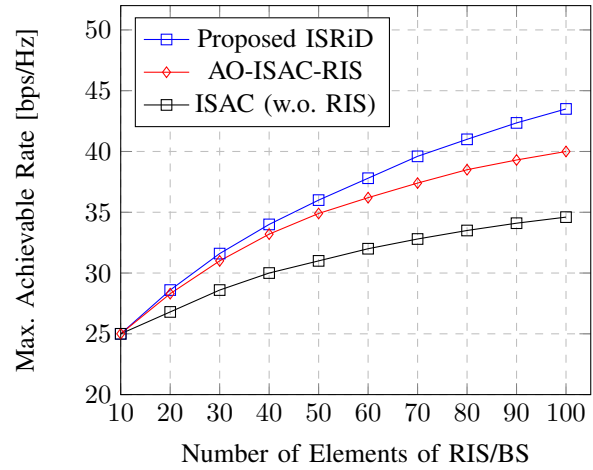


Figure 5: Max. Achievable Rate Vs. Elements of RIS/BS

a higher maximum achievable rate of 43.5 bps/Hz than the benchmark of 40 bps/Hz. This is noteworthy because, in addition to improving communication performance, a small N doesn't require a large number of active antenna elements and the associated extensive transmit RF chain. Consequently, this approach significantly reduces costs. Furthermore, the analysis indicates that the ISAC-assisted RIS scenarios exhibit considerable improvements with minor declines compared to scenarios without RIS. This can be attributed to the efficient concentration of total power towards the intended receiver. Additionally, the proposed framework demonstrates superior performance when compared to benchmark schemes. This ensures that the learned DDPG policy can adapt to different scenarios, resulting in enhanced performance across a variety of communication conditions.

VII. CONCLUSION

In this paper, we delve into IoE's transformative impact and introduce the ISRiD framework, which incorporates ISAC, RIS, DDPG to address IoE's complexities and prepare for the

envisioned 6G networks. This framework bridges sensing and communication, facilitating real-time adaptation for enhanced efficiency, reliability, and data rates. Simulations unequivocally validate the effectiveness of the ISRIID, elevating IoE networks for dynamic environments. This represents a major stride towards IoE's vision, empowering adaptable communication systems in the ever-evolving digital landscape, while also anticipating the opportunities of the envisioned 6G networks.

REFERENCES

- [1] S. Ahmad, S. Khan, K. S. Khan, F. Naeem, and M. Tariq, "Resource allocation for IRS-Assisted networks: A deep reinforcement learning approach," *IEEE Communications Standards Magazine*, vol. 7, no. 3, pp. 48–55, 2023.
- [2] A. M. Elbir, K. V. Mishra, M. B. Shankar, and S. Chatzinotas, "The rise of intelligent reflecting surfaces in integrated sensing and communications paradigms," *IEEE Network*, 2022.
- [3] M. Tariq, M. Ali, F. Naeem, and H. V. Poor, "Vulnerability assessment of 6G-enabled smart grid cyber-physical systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5468–5475, 2021.
- [4] F. Naeem, S. Seifollahi, Z. Zhou, and M. Tariq, "A generative adversarial network enabled deep distributional reinforcement learning for transmission scheduling in internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4550–4559, 2021.
- [5] C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu, Y. Pan, H. Ren, M. Di Renzo, A. L. Swindlehurst, R. Zhang *et al.*, "An overview of signal processing techniques for RIS/IRS-Aided wireless systems," *IEEE Journal of Selected Topics in Signal Processing*, 2022.
- [6] X. Song, D. Zhao, H. Hua, T. X. Han, X. Yang, and J. Xu, "Joint transmit and reflective beamforming for IRS-Assisted integrated sensing and communication," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2022, pp. 189–194.
- [7] Z. Zhu, Z. Li, Z. Chu, G. Sun, W. Hao, P. Xiao, and I. Lee, "Resource allocation for IRS Assisted mmwave integrated sensing and communication systems," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 2333–2338.
- [8] R. P. Sankar, B. Deepak, and S. P. Chepuri, "Joint communication and radar sensing with reconfigurable intelligent surfaces," in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2021, pp. 471–475.
- [9] Z.-M. Jiang, M. Rihan, P. Zhang, L. Huang, Q. Deng, J. Zhang, and E. M. Mohamed, "Intelligent reflecting surface aided dual-function radar and communication system," *IEEE Systems Journal*, vol. 16, no. 1, pp. 475–486, 2021.
- [10] T. Wei, L. Wu, K. V. Mishra, and M. B. Shankar, "Multiple IRS-Assisted wideband dual-function radar-communication," in *2022 2nd IEEE International Symposium on Joint Communications & Sensing (JC&S)*. IEEE, 2022, pp. 1–5.
- [11] X. Wang, Z. Fei, J. Huang, and H. Yu, "Joint waveform and discrete phase shift design for RIS-Assisted integrated sensing and communication system under Cramer-Rao bound constraint," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 1004–1009, 2021.
- [12] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Network*, vol. 35, no. 5, pp. 158–167, 2021.
- [13] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE transactions on wireless communications*, vol. 18, no. 11, pp. 5394–5409, 2019.
- [14] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [15] X. Liu, H. Zhang, K. Long, M. Zhou, Y. Li, and H. V. Poor, "Proximal policy optimization-based transmit beamforming and phase-shift design in an IRS-Aided ISAC system for the THz band," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2056–2069, 2022.
- [16] T. O'shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.
- [17] R. Liu, M. Li, and A. L. Swindlehurst, "Joint beamforming and reflection design for RIS-assisted ISAC systems," in *2022 30th European Signal Processing Conference (EUSIPCO)*. IEEE, 2022, pp. 997–1001.
- [18] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface-aided wireless networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3064–3076, 2020.
- [19] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–6.



Sohail Ahmad received his M.Sc. and M.Phil. degrees in Electronics engineering from the University of Peshawar in 2012 and 2017, respectively. Since 2019, he has been enrolled in Ph.D. in the Department of Electrical Engineering, National University of Computer and Emerging Sciences (NUCES), Islamabad, Pakistan. His research interests include IRS-assisted 6G communication and advanced machine learning-based optimization approaches.



Muhammad Tariq (M'12–SM'17) holds the position of Professor and Head at the Department of Electrical Engineering, NUCES, Islamabad, Pakistan. His extensive research portfolio includes publications in the internet of things, network sciences, cybersecurity, and smart grids, published in prestigious IEEE journals/transactions. He earned recognition as being among Stanford's university list of the World Top 2% Scientists.



Mian Ahmad Jan is an Assistant Professor at the Department of Computer Science, School of Computing and Informatics, University of Sharjah, United Arab Emirates. His research interests include cybersecurity, energy-efficient and secured communication for Wireless Sensor Networks, and the Internet of Things. He has published his research in the world's leading journals and conferences and has been among Stanford's university list of the World Top 2% Scientists.



Houbing Herbert Song (M'12–SM'14–F'23) received the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, in August 2012. He is currently a Professor and the Director of the NSF Center for Aviation Big Data Analytics (Planning), and the Director of the Security and Optimization for Networked Globe Laboratory, University of Maryland, Baltimore, MD. His research interests include cyber-physical systems/internet of things, cybersecurity and privacy, and AI/machine learning/big data analytics.