Proceedings of the 2nd GCC International Conference on Industrial Engineering and Operations Management Muscat, Oman, December 1-3, 2024

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/GC02.20240108
Published: December 01, 2024

Multi-Agent for QoS-Aware Flying-IRS Deployment in 6G THz Networks

Nahla Nur Elmadina
Department of Electrical Engineering, Alzaiem Alazhari University
Khartoum, Sudan
nahlla.awadnoor@gmail.com

Nour Eldin Osman
Head of the Department of Information Studies at Sultan Qaboos University
Muscat, Oman
n.osman@aqu.edu.om

Rashid A. Saeed
Electronics Department, Sudan University of Science and Technology
Khartoum, Sudan
eng_rashid@hotmail.com

Elsadig Saeid
Department of Electrical Engineering, Alzaiem Alazhari University
Khartoum, Sudan
els197778@gmail.com

Abstract

Intelligent reflecting surfaces (IRS) and terahertz (THz) communication networks are proven to continue to help shape the advancement of the Sixth generation wireless communication. Next technology enables very high-speed data transfer and vastly increases communications network capacity. Motivated by these advancements, this work is aimed at investigating them. In this paper, the feasibility of flying -IRS in THz communication paradigms for revolutionizing wireless services, particularly in remote areas is explored. This paper introduces the Fly-IRS algorithm as a new approach to regulate the choice of user devices, IRS phase shifts, and UAV positions. The Fly-IRS algorithm is designed to improve data rates, increase network capacity, and minimize disruptive services that have a detrimental effect on users. To perform an optimal user classification, the method integrates Multi-Agent Deep Reinforcement Learning (MADRL) and Particle Swarm Optimization (PSO) to optimize both the IRS phase configurations and paths. The simulations show that the Fly-IRS algorithm improves offered data rates by 10% compared to the MADRL. Therefore, these results emphasize the role of THz communication and IRS technology in the development of future wireless network generations.

Keywords

6G, Terahertz (THz), UAV-RIS, Multi-Agent -DRL, Particle Swarm Optimization (PSO), QoS

1. Introduction

Several emerging new applications require wireless networks to be very reliable and provide high-speed connectivity such as augmented reality, auto-mobiles, smart cities, etc. Having been considered the next generation of communication system after 5G, the sixth generation (6G) is expected to provide data rates of more than 100 Gbps per user, together with spectral efficiency. Terahertz (THz) communication, defined as operating in the frequency range from 0.1 to 10 THz, has become regarded as the backbone of such innovations, which could not exist under conventional microwave and millimeter-wave communication technologies (Miyamoto et al., 2024; Zawish et al., 2024).

To properly use terahertz communications in the 6G context, the IRS and unmanned aerial UAV is a good solution. IRS technology distorts electromagnetic waves to increase signal strength and increase network coverage. This feature is very useful for eliminating bottlenecks and increasing visual acuity. The mobility and adaptability of UAVs, in conjunction with IRS, create an innovative approach for optimizing wireless network performance (Amodu et al., 2023; Chen et al., 2022). This combination has the potential to significantly enhance energy efficiency and overall quality of service (QoS) in 6G THz networks (Mustari et al., 2024; Ranjha, Naboulsi, et al., 2024). Despite these advantages, the deployment of UAV-IRS systems in THz-enabled 6G networks introduces several challenges, particularly regarding the optimization of UAV trajectories, resource allocation, and interference management. The higher frequencies utilized in THz communication are more susceptible to propagation losses and blockages, necessitating efficient algorithms for real-time trajectory and resource optimization (Ranjha et al., 2023; Ranjha, Javed, et al., 2024).

However, the implementation of THz-enabled 6G brings several challenges, as follows: THz bands, and especially the high frequencies, have higher propagation loss and are prone to blockage by obstacles within urban environments. Due to these limitations, different solutions such as Intelligent Reflecting Surfaces (IRS) and Unmanned Aerial Vehicles (UAVs) have been introduced. IRS improves the signal intensity and provides higher coverage for the electromagnetic signal by effectively reflecting these signals, and the UAV technology provides flexibility and quick deployment in a dynamic scenario. The integration of UAVs and IRS, forming UAV-IRS systems, is a new concept in wireless communication that is capable of providing adaptive and energy-efficient solutions.

Such Fly-IRS systems come with unique critical challenges, among which are the following: UAV trajectory planning, IRS beam forming, resource management, and energy consumption. Problem-solving for these issues in real-time dynamic conditions may be unsatisfactory with the help of some existing solutions. Inspired by these gaps, this work studies the application of UAV-based IRS units in a THz-based 6G networking environment by developing a novel PSO-MADRL framework. In delivering reliability to THz communication, this research envisions improving the data rates and guaranteeing the Quality of Service (QoS) for various applications.

1.1 Proposed Problem Statement

The integration of Fly with Intelligent Reflecting Surfaces (IRS) in the THz-enabled 6G network paradigm can be a game changer for various performance aspects of the network. However, due to the high-frequency usage of THz communication, there is a huge path loss and vulnerability to obstacles within an urban area. Most of the current techniques developed for determining the UAV trajectory and IRS phase shifts are off-line and thus inadequate for dynamic environments, which implies poor resource utilization and QoS. Solving these challenges involves the application of new optimization methods that can guarantee the availability to meet various QoS demands.

1.2 Objectives

In this paper, we introduce the PSO-MADRL framework, specifically designed for QoS-aware deployment of Fly-IRS systems in 6G THz networks. Contributions to our work include:

- We propose an advanced framework that integrates Particle Swarm Optimization (PSO) with Multi-Agent Deep Reinforcement Learning (MADRL) to optimize the positioning of Fly-mounted IRS units, accounting for varying user QoS requirements and dynamic network conditions.
- By leveraging the PSO-MADRL approach, we effectively address the complex decision-making processes required for optimal deployment, trajectory optimization, and resource allocation within Fly-IRS systems.
- The PSO-MADRL framework enhances the data rate of the network, successfully tackling the challenges associated with the high-frequency nature of THz communications.

Our work progresses the state-of-the-art in 6G networks considerably by employing new optimization techniques for performance improvement and user satisfaction.

2. Literature Review

Deploying IRS and UAVs has attracted much attention as potential candidates for signal improvement and stability in THz communication systems. THz frequencies are capable of providing ultra-high data rates, however, propagation and blockages can pose a significant issue for signal connectivity, thus requiring different solutions to be used to enhance network performance (Miyamoto et al., 2024). In addition, (Saad et al., 2020) presented an overview of key technologies for 6G, including THz communications and UAV-assisted systems, outlining the benefits and challenges of integrating these technologies. Furthermore, (Zawish et al., 2024) explored the intersection of AI, 6G, and the metaverse, emphasizing the role of AI-driven optimization in THz-enabled systems.

Work by (Mustari et al., 2024) focused on cooperative THz communication for UAVs in 6G and beyond, discussing the potential of UAV-IRS systems for enhancing coverage and signal quality. Similarly, (Pan et al., 2021) explored UAV-assisted IRS-supported THz communications, highlighting the importance of trajectory optimization to improve spectral efficiency and signal quality in THz bands. However, these studies do not usually take into consideration how the system can adjust dynamically to the needs of the user.

In terms of multi-agent learning, (Ibrahim et al., 2021) surveyed models and algorithms for multi-agent deep reinforcement learning, presenting a foundation for cooperative UAV and IRS deployments. Recent advances have also emphasized the energy efficiency and trajectory optimization of UAV-mounted IRS in THz networks. (Aziz & Girici, 2022) Addressed the deployment of UAV-mounted IRS in THz bands, pointing out the challenges of positioning and energy constraints. More recent work by (Elmadina et al., 2024) proposed a double deep reinforcement learning strategy for UAV-assisted energy harvesting optimization, further emphasizing the importance of advanced learning techniques in 6G network optimization.

To fill these gaps, this work presents a PSO-MADRL framework that extensible targets the deployment of Fly-IRS in the 6G THz network. This work aims at improving data rates whilst maximizing QoS in highly dynamic environments by using multi-agent learning and optimization.

3. System Model

The model examines a UAV-IRS-aided 6G THz network to optimize the UAV's trajectory while accounting for energy consumption limitations. The following equations outline the core components of the system model.

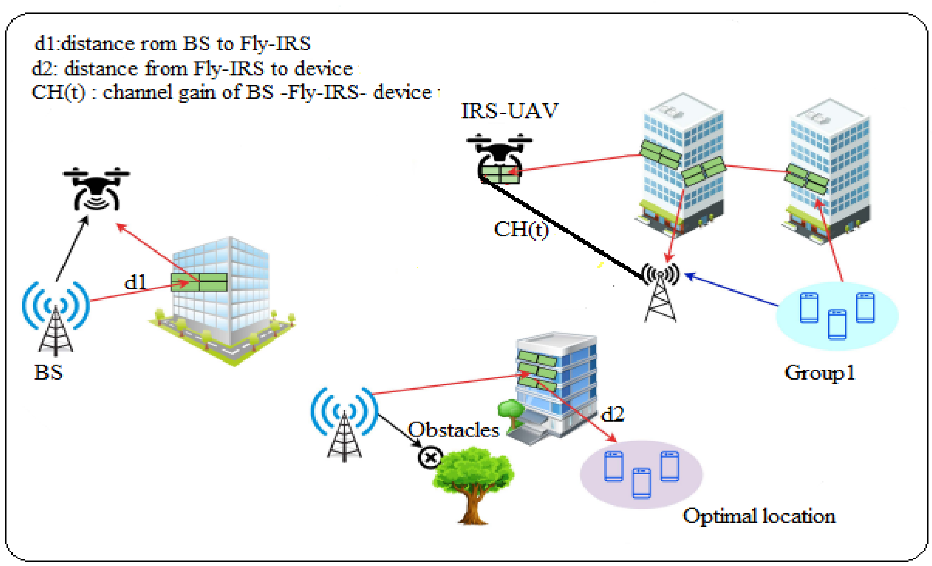

Figure 1. The Fly-IRS --assisted Terahertz (THz) communication system

The model examines a UAV-IRS-aided 6G THz network to optimize the UAV's trajectory while accounting for energy consumption limitations. The following equations outline the core components of the system modeling this work, we focus on a 6G terahertz (THz) communication network which is assisted by a UAV-IRS operating in the frequency range of 0.3 to 10 THz. The system features a base station (BS) positioned at the center of an urban area, ground devices (GDs) deployed at various locations in the area to be covered, and IRS-equipped unmanned aerial vehicle (UAV) or Fly-IRS for improving the communication link. It is assumed that the direct link between the BS and GDs is restricted by obstacles such as buildings and other structures.

3.1 Channel Model with Fly -IRS

In this framework, Fly-IRS units act as flying amplifiers to provide alternating access between the BS and the GDs while avoiding interference from the environment. A transmitted signal goes through a channel transfer function relevant to the THz band and is received by the receiver. Due to the ability of select Fly-IRS units to change their orientation with these channels, the communication quality is improved. Therefore, every IRS is provided with$I_X \times I_Y$, reflecting elements with which it can manage phase shifts and thereby form stable channels between the BS and the GDs. The channel gain for the communication path from the BS to the Fly-IRS, and from the Fly-IRS to the GDs, is represented as $CH(t)$Where are the respective channel gains for these links in Figure 1.

$$CH(f,d) = \frac{v}{4\pi(f)d} e^{-0.5k(f)d} \tag{1}$$

v is the speed of light, k(f) denotes the absorption loss coefficient (which varies with the transmission frequency f, and d represents the distance the signal travels.

The total distance from the base station (BS) to the Fly-IRS and from the Fly-IRS to the ground devices (GDs) can be described as: $d = d_1 + d_2 = \sqrt{(x^{BS} - x_i^{Fly-IRS}(t))^2 + (y^{BS} - y_i^{Fly-IRS}(t))^2 + (z^{BS} - z^{Fly-IRS})^2} + \sqrt{(x_i^{Fly-IRS}(t) - x^{GD})^2 + (y_i^{Fly-IRS}(t) - y^{GD})^2 + (z^{Fly-IRS} - z^{GD})^2}$ Where d1 represents the distance between the BS and the first reflector of the IRS, and d_2 is the distance from the first IRS reflector to the GDs. To optimize the linkages, the Fly-IRS dynamically modifies its position and phase shifts. The beamforming matrix W of the IRS is a diagonal matrix that adjusts the phase shifts of the elements (Dhuheir et al., 2024):

$$\varphi = \text{diag}[w_1 e^{j\emptyset_1},\ w_2 e^{j\emptyset_2},\ ...,\ w_n e^{j\emptyset_n M}] \tag{2}$$

where $w_n \in [0,1]$ is the amplitude reflection coefficient for the n^{th} element $\emptyset_n \in [0,2\pi]$is the phase shift of the n^{th}element and i represents the imaginary unit.

The Fly-IRS's position is updated at each time step, and the traveled distance is recalculated accordingly(Saleh et al., 2024). The Fly-IRS visits a series of stop points $Q_i$along its path, represented by:

$$Q_i(T_i^F) = \left[x_i^{Fly-IRS}(t), y_i^{Fly-IRS}(t), z_i^{Fly-IRS}(t)\right]^T \tag{3}$$

Let R_n represent the total data transmitted to the device users. The rate can be expressed as:

$$R_n = \left(\frac{BWi}{n}\right) \log_2\left(1 + \frac{p_t * CH(t)}{\vartheta + \left(d\, e^{-0.5k(f)d}\right)\sigma^2}\right) \tag{4}$$

In this equation, p_t represents the transmit power from the Fly-IRS to the ground devices (GDs), BW is the total bandwidth allocated in the THz spectrum, CH(t)is the channel gain, σ^2 is the variance of the additive white Gaussian noise (AWGN), and ϑ accounts for all-experienced interference.

3.2 Energy Consumption

The energy consumption of a UAV-IRS system is primarily determined by propulsion power for flight and hovering, along with communication power. The propulsion power, ϵ , depends on the UAV's speed v and is expressed (Dhuheir et al., 2024; Elmadina et al., 2024) as:$\epsilon = P_0\left(1 + \frac{3V^3}{U^3}\right) + P_1\left(1 + \frac{V^4}{4{V_0}^4} - \frac{V^2}{4{V_0}^2} + \frac{d_0 \rho V^3\ rot\ A}{2}\right)$ where $P_0$and P_1

are constants associated with the blade profile and induced hovering power, respectively. U is the rotor blade tip speed, v_0 represents the average rotor-induced velocity during hovering, d_0 is the fuselage drag coefficient, ρ is the air density, rot is rotor stiffness, and A is the rotor disk area. When the UAV is hovering, i.e., v_0, the power consumption is simplified P_0+ P_1 . The total energy consumption ϵ is the sum of flight, hovering, and operational energy (Hassan et al., 2023)

$$\varepsilon = \sum_{i=1}^{M} \varepsilon_i^{hov} + \varepsilon_i^{Fl} + \varepsilon_i^{oper} \tag{5}$$

Flight and hovering energy are calculated as: $\varepsilon_i^F = T_i^F . P^F$, $\varepsilon_i^{hov} = T_i^{hov} .\ P^{hov}$, Obstacle models optimize Fly-IRS paths, enhancing safety but increasing energy due to detours. Although the Fly-IRS algorithm improves data rates, a detailed analysis of the energy efficiency trade-offs is crucial for practical deployments (Bakri Hassan et al., 2022). Future enhancements could include adaptive energy-saving strategies that balance throughput and energy consumption, particularly under varying UAV altitude, speed, and hovering time. Introducing an energy-aware optimization criterion within the PSO-MADRL framework may help to better align system performance with energy efficiency (Hassan et al., 2024)

4. Optimization Problem Formulation Fly-IRS-Assisted THz Wireless Communication System

This system aims to enhance the trajectory and communication of Fly-IRS units to increase the overall data transmission rate and decrease energy consumption. The Fly-IRS functions in a network utilizing the terahertz band alongside ground devices spread throughout the coverage zone, though obstacles block direct communication between the base station and ground devices.

Maximize the sum data. R_{sum} across all devices:

$$P_1: max_{O,\emptyset_n,\omega_n} \sum_{i=1}^{N} R_{sum} \tag{6}$$

$$C1: R_i(t) \geq R_{min} \quad \forall\ [1,2\ 3,..m]$$

$$C2: d(t) \leq d_{max} \quad \forall t$$

$$C3: z_u^{Fly-IRS}(t) \geq z_{min} \quad \forall i, \forall t$$

$$C4: p_i^t \leq p_u^{max} \quad \forall i$$

$$C5: \emptyset_n \in [0,2\pi]$$

The function is intended to achieve the highest total data rate, for all the devices. The optimization problem must adhere to several important constraints: C1 demands the minimum data rate of each GD, and C2 restricts the distance that UAVs can move. Furthermore, the altitude of the Fly-IRS units should be above a certain level, the corresponding transmit power for each UAV is limited, and the phase shift of IRS should be within the given range. The system works effectively under these restrictions to satisfy both the performance and the limitation requirements.

4.1 Framework PSO-MADRL

The framework PSO-MADRL could improve network performance in settings with unexpected mobility and fluctuating user density. In order to maximize the overall data rate in 6G THz networks for UAV-IRS operations,

we frame the issue as a Markov Decision Process (MDP) with crucial elements such as states, actions, rewards, and transition probabilities (Hassan et al., 2022). In our framework, the state s_t represents the Fly-IRS's current position and environmental context at time t, including user locations, signal quality, and interference levels. The action a_t indicates the trajectory decisions made by the Fly-IRS, such as moving to a new position or adjusting the IRS angle to optimize signal strength. The PSO-MADRL framework comprises a state-set s_t, an action set a_t, a reward set $r_t$and the Fly-IRS agent, which performs actions to achieve rewards while updating states. Given that IRS phase shifts in the THz band are continuous variables with high quantization levels, employing the PSO-MADRL frame network effectively addresses the challenge of continuous solutions as Figure 2.

- The state at time step t is defined as:

$$s_t^j = (x_i^{Fly-IRS}\,(t), y_i^{Fly-IRS}\,(t), \emptyset_n, x_i^{GD}\,(t), y_i^{GD}\,(t)) \quad (7)$$

Where $x_i^{Fly-IRS}\,(t), y_i^{Fly-IRS}$represents the optimal location of the UAV at time t,$\emptyset_n$ denotes the IRS phase shifts, and $x_i^{GD}\,(t), y_i^{GD}(t)$ represents the optimal location of the GD.

The action space comprises the options available to an agent as it transitions from its current state to a subsequent state. At times t the action taken by the agent, which encompasses IRS phase adjustments and UAV movement, directly relates to two optimization parameters Optimal Deployment Position (Elfatih et al., 2023; Saleh et al., 2024) and IRS Phase Adjustments (Hassan et al., 2022, Omar et al., 2023). In this context, the agent determines the UAV's next position at each time step. The proposed method allows the agent to ascertain the most effective movement strategy while considering long-term rewards. For every component in the system, the agent identifies the optimal phase shift at that moment, without accounting for the time required to adjust the angles of the reflective elements. Each Fly-IRS agent consists of discrete actions that dictate movement direction.

- The action set can be defined as:

$$a_t = (w_n\,, x_i^{Fly-IRS}(t),\, y_i^{Fly-IRS}(t)\ ) \quad (8)$$

- $r_t = \sum_{i=1}^{N} R_{sum}$ (9)

The agent receives a reward $r_t\ (s_t, a_t)$ upon executing the action a_t in state s_t, at time t. In alignment with our objective, the reward is defined as the total data rate for each group from equation (6).

To utilize the PSO-DRL algorithm, the Q-value update for each UAV is a critical step in Deep Reinforcement Learning (DRL), as it enables the UAV to learn the optimal policy for maximizing data rates over time. The Q-value reflects the expected cumulative reward for taking a specific action. $a_i(t)$ in state $s_i(t)$ and it is updated iteratively based on the Bellman equation (Bilal Ur Rehman et al., 2023). The MADRL employs double Q-learning to ensure more stable training and reduce overestimation issues (Alatabani et al., 2023).

- The Q-value update for UAV i is given by:

$$Q^{DRL}_{(S_i(t),a_i(t))} \leftarrow Q(s_i(t), a_i(t) + \eta[(r_i(t) + \gamma Q(\,s'_{\mathrm{i}}(t+1), argmax_a Q(s_i(t+1), a';\ \Theta)) - Q(s_i(t), a_i(t), \Theta)] \quad (10)$$

Where η is the learning rate, γ is the discount factor, and $S'_{\mathrm{i}}\ (t+1)$ represents the next state after taking action a_i.The gradient updates for the MADRL algorithm follow the standard deep reinforcement learning approach, with the loss function being computed using Temporal Difference (TD) error, which quantifies the gap between the current and target Q-values. For each UAV agent i, the loss function $\boldsymbol{L_i}\ (\boldsymbol{\theta}) =$ is defined as:

$$\boldsymbol{L_i}\ (\boldsymbol{\theta}) = \boldsymbol{E}[(y_i(t) - Q_i(s_i(t), a_i(t), \Theta))^2] \quad (11)$$

The target Q-value $Y_i(t)$ is calculated as:$Y_i(t) = r_i(t) + \gamma Q(s'_i(t+1), argmax_{a'}\ Q(S'_i(t+1), a'_i;\ \Theta); \Theta')$

Here, θ represents the current network parameters for agent i, and θ′ denotes the parameters of the target network, which are updated periodically to stabilize learning. The parameters θ are updated using gradient descent, θ← θ + η·$\nabla_\Theta \boldsymbol{L_i}\ (\boldsymbol{\theta})$Here, η is the learning rate that controls the step size during the parameter update process.

In a multi-agent setting, each Fly-IRS computes its gradients independently, but the environment is shared among all agents. Each agent's parameter updates take into account the influence of other agents' actions. The use of DRL, along with separate target networks, enhances the stability of the learning process by mitigating overestimation risks. PSO optimizes the phase shifts of the IRS, while MADRL focuses on optimizing UAV trajectories. These two processes work iteratively to maximize total data rates throughout the Fly-RIS-user system. This integrated PSO-MADRL approach significantly boosts overall data rates while adhering to communication quality constraints, ensuring reliable performance in dynamic environments and enhancing overall network efficiency in 6G THz systems.

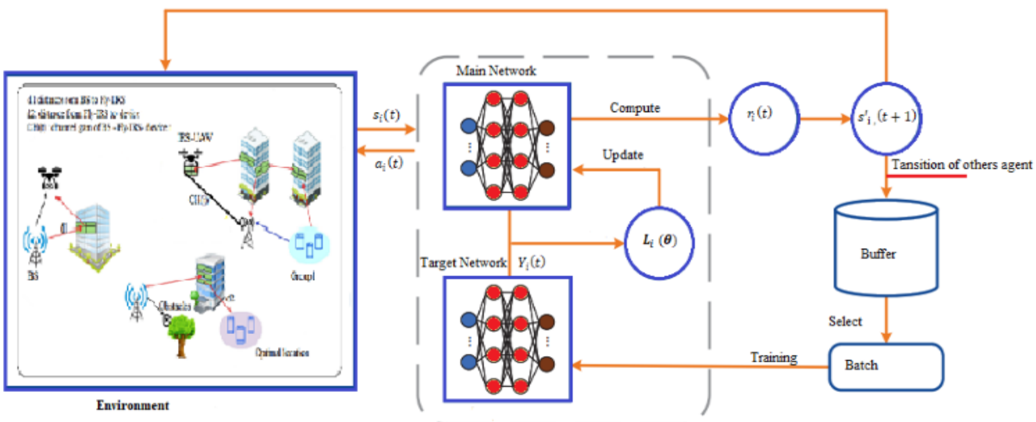

Figure 2. Training Frame PSO-MADRL

Algorithm 1: PSO-MADRL

Part1 MADRL: Input: number of Fly-IRS, N, iter=max iteration PSO, popu=size of population in PSO, ϵ: Exploration rate, γ: Discount factor

Output: Optimized UAV trajectories, IRS phase shifts, Optimal user device

Initialization: Define the environment parameters: Fly positions for each UAV (1, 2,..U), initialize ground devices (GDs) based on user position, and randomly initialize IRS phase φ.

For each time slot t ∈ T:
For each Fly -IRS ∈ U:

- **State Representation**
  - Define the current s_t For each UAV-IRS agent as:
  S_t^j=($x_i^{Fly-IRS}(t), y_i^{Fly-IRS}(t), \emptyset_n$)
- **Action Space on the ϵ -greedy policy,** a_t
- *Check constraints*
  - *If the constraint is satisfied, increment reward*
- **Reward Updating:**
  - If all constraints are satisfied:$r_i(t) \leftarrow r_i(t+1)$
- **Observe next state s_{t+1} and reward r_t**
  - Update new Fly UAV locations
  - Save (s_t, a_t, r_t, s_{t+1}) in replay memory
- **Minibatch Learning**

- Sample a minibatch from replay memory (s_t, a_t, r_t, s_{t+1})
- Updating the value function set $\Theta_{old} \leftarrow \Theta$
- calculate error r_t+ γV (s_{t+1})-V (s_t)
- Update the Q-network parameters

Part2: PSO for IRS phase shift optimization

- **Initialize PSO parameter:**
  - Set initialize weight w, cognitive factor $b_1(self\ learning)$,social factor b_2 (swarm influence)
  - Randomly initialize the **best** phase shift φ_{best} and each particle's phase shift $\varphi(i)$
- For each iteration t=1to max -iter
  For each particle's i=1 to popu
  - update velocity V_i(t+1)=w.V_i(t)+ b_1. r_1 .($\varphi_{best} - \varphi_i(t)$)+ r_2. b_2.($\varphi_{best} - \varphi_i(t)$), r_1, r_2*radamly[0,1]*
  - Updating the phase shift
    $\varphi_i(t+1) = \varphi_i(t) + V_i(t+1)$
- Evaluate cost function
  Compute cost function φ_{cost}=evaluate (constraints)
- Updating the best shift if necessary
  If new cost $\varphi_{cost} > \varphi_{best}$ then $\varphi_{best} = \varphi_{cost}$
- Convergence Criteria
  - If convergence is achieved either by
    Change$\left|\varphi_{i+1} - \varphi_i\right| < \epsilon$

5. Results and Discussion

We examine a communication scenario within an urban area spanning 300 m², where a base station (BS) is centrally positioned and serves 200 ground users by MATLAB and Python 3.7.4 with TensorFlow 2.10. The UAV-IRS maintains a fixed altitude of 80 meters, traveling at a speed of 15 m/s. Further simulation parameters are outlined in Table 1 providing a comprehensive overview of the environmental setup used for the simulation.

Table 1. Parameter of modeling

Parameter	symbol	Value
Number of UAV-IRS Units	M	5
Carrier frequency	f	0.3THz
Absorption coefficient	K(f)	0.005
Bandwidth	BW	0.2 THz
Max power transmit	P	5W
Min Rate	R_{min}	0.01bps/Hz

Parameter	symbol	Value
Noise power	σ^2	-174dBm
Discount Factor	γ	0.9
Experience Replay Memory Size	C	10000
Hovering Power Consumption	P_{hoov}	50 W

5.1 Graphical Results

Figure 3 illustrates the improvement in data performance as the number of RIS elements increases. The addition of more RIS elements leads to a significant enhancement in the overall data transmission rate. The proposed PSO-MADRL outperforms the MADRL, achieving a 10% increase in throughput by effectively optimizing the placement and configuration of the RIS elements to enhance signal reflection and data transmission. Although the MADRL approach yields competitive results, it slightly underperforms in maximizing throughput due to a less comprehensive optimization strategy (Saeed et al., 2023). In contrast, the QN method, which lacks advanced optimization capabilities, consistently produces the lowest throughput across all scenarios, highlighting its ineffectiveness in utilizing additional RIS elements for enhanced performance.

Figure 4 presents the simulation results illustrating the percentage of service satisfaction among user devices as the UAV transmit power increases. The proposed PSO-MADRL, which integrates a genetic algorithm with double deep reinforcement learning, effectively addresses both objective functions. As the transmit power of the UAVs rises, the quality of service (QoS) improves, leading to more efficient resource allocation and increased throughput for the RIS. The PSO-MADRL approach consistently outperforms both the MADRL and baseline algorithms in terms of QoS. However, it requires more computational resources due to the larger search space and the complex optimization processes involved in identifying the optimal solution for maximizing QoS in 6G Fly-IRS systems.

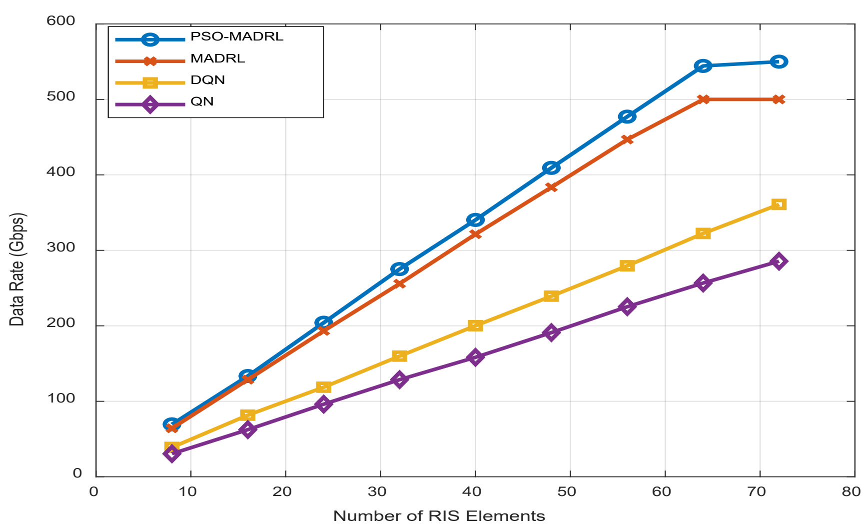

Figure 3. illustrates the data performance improvement as the number of RIS elements increases.

Figure 5 illustrates the average sum data rate as a function of the number of users, with a fixed configuration of 64 IRS elements. It also includes comparisons with a randomly configured IRS. Across all four scenarios, the sum data rate initially increases with the number of users, reaching a peak before declining. This trend can be attributed to the significantly higher SINR at THz frequencies, which results from the unique high path-loss characteristics of terahertz channels, leading to minimal user interference. The PSO-MADRL consistently outperforms all other configurations, demonstrating superior performance in every scenario. The advantage of a well-optimized IRS configuration over a randomly set one is evident, particularly in terms of maximizing throughput. Additionally, once the user count exceeds 75, the level of user interference increases. In this scenario, the PSO-MADRL approach excels, providing a higher sum data rate compared to the other algorithms due to its effective management of interference and optimization of resource allocation. Another significant observation from this Figure 5 is that when the user count reaches 60, the PSO-MADRL achieves an impressive sum data rate of 300 Gbps. This performance surpasses that of the other approaches, underscoring the effectiveness of the proposed algorithm in handling user dynamics and optimizing the communication environment (Table 2).

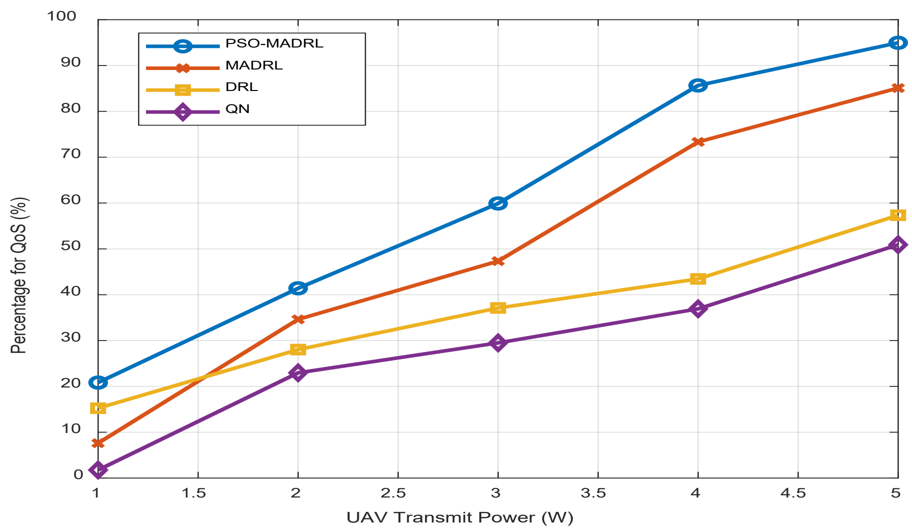

Figure 4. presents the simulation results illustrating the percentage of service satisfaction among user devices (UDs) as the UAVs transmit power increases.

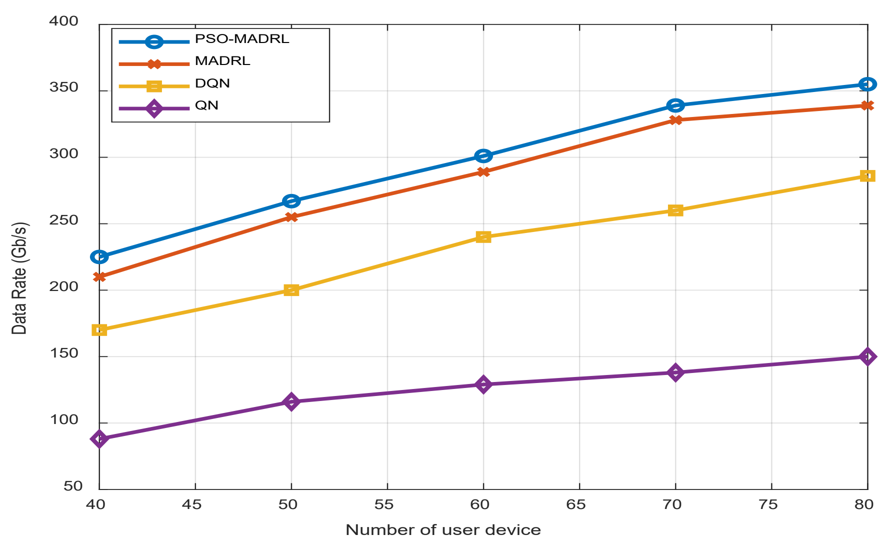

Figure 5. illustrates the average sum data rate as a function of the number of users

Table 2 summarizes the percentage improvements of PSO-MADRL compared to the other algorithms (MADRL, DRL, and RL)

Table 2. The percentage improvements of PSO-MADRL compared to the other algorithms (MADRL, DRL, and RL)

Metric	**Proposed vs MADRL**	**Proposed vs DRL**	**Proposed vs QN**
Data Improvement (%)	10%	22.45%	122.22%
QoS Improvement (%)	11.76%	63.79%	86.27%

The PSO-MADRL emerges as a highly effective approach, demonstrating superior performance in terms of both throughput and service quality compared to other contemporary reinforcement learning-based algorithms.

6. Conclusion and Future Work

This research shows that the proposed PSO-MADRL improves data transmission rates and service quality in Fly-IRS-aided THz communication systems. The results show that increasing the number of RIS elements improves data performance significantly, with the PSO-MADRL achieving a 10% increase in throughput over the MADRL method. This improvement is primarily due to the optimized placement and configuration of RIS elements, which results in better signal reflection and data transmission efficiency.

The PSO-MADRL consistently outperforms other configurations in terms of quality of service (QoS) and total data rate, reaching an impressive 300 Gbps when the user count reaches 60. It excels at managing user interference and optimizing resource allocation, making it a top contender in complex communication environments. From this research, it is evident that the PSO-MADRL framework can enhance the data transmission rates and services of the THz-aided Fly-IRS communication system. Several potential research areas can be investigated based on these findings to improve the framework's capabilities and adapt it to future scenarios.

References

Alatabani, L. E., Saeed, R. A., Ali, E. S., Mokhtar, R. A., Khalifa, O. O., & Hayder, G. (2023). Vehicular Network Spectrum Allocation Using Hybrid NOMA and Multi-agent Reinforcement Learning. In G. H. A. Salih & R. A. Saeed (Eds.), *Sustainability Challenges and Delivering Practical Engineering Solutions* (pp. 151–158). Springer International Publishing. https://doi.org/10.1007/978-3-031-26580-8_23

Amodu, O. A., Jarray, C., Busari, S. A., & Othman, M. (2023). THz-enabled UAV communications: Motivations, results, applications, challenges, and future considerations. *Ad Hoc Networks*, *140*, 103073. https://doi.org/10.1016/j.adhoc.2022.103073

Aziz, R., & Girici, T. (2022). Deployment of a UAV-Mounted Intelligent Reflecting Surface in the THz Band. *2022 International Balkan Conference on Communications and Networking (BalkanCom)*, 168–172. https://doi.org/10.1109/BalkanCom55633.2022.9900625

Bakri Hassan, M., Saeed, R., Khalifa, O., Sayed Ali Ahmed, E., Mokhtar, R., & Hashim, A. (2022). *Green Machine Learning for Green Cloud Energy Efficiency*. 288–294. https://doi.org/10.1109/MI-STA54861.2022.9837531

Chen, H., Sarieddeen, H., Ballal, T., Wymeersch, H., Alouini, M.-S., & Al-Naffouri, T. Y. (2022). A Tutorial on Terahertz-Band Localization for 6G Communication Systems. *IEEE Communications Surveys & Tutorials*, *24*(3), 1780–1815. IEEE Communications Surveys & Tutorials. https://doi.org/10.1109/COMST.2022.3178209

Dhuheir, M., Erbad, A., Al-Fuqaha, A., & Guizani, M. (2024). Multi-UAV Multi-RIS QoS-Aware Aerial Communication Systems Using DRL and PSO. *ICC 2024 - IEEE International Conference on Communications*, 654–659. https://doi.org/10.1109/ICC51166.2024.10622205

Elfatih, N. M., Ali, E. S., & Saeed, R. A. (2023). Navigation and Trajectory Planning Techniques for Unmanned Aerial Vehicles Swarm. In A. T. Azar & A. Koubaa (Eds.), *Artificial Intelligence for Robotics and Autonomous Systems Applications* (pp. 369–404). Springer International Publishing. https://doi.org/10.1007/978-3-031-28715-2_12

Elmadina, N. N., Saeed, R. A., Saeid, E., Ali, E. S., Nafea, I., Ahmed, M. A., Mokhtar, R. A., & Khalifa, O. O. (2024). Double Deep RL-Based Strategy for UAV-Assisted Energy Harvesting Optimization in Disaster-Resilient

IoT Networks. *2024 9th International Conference on Mechatronics Engineering (ICOM)*, 411–416. https://doi.org/10.1109/ICOM61675.2024.10652500

Elmadina, N. N., Saeed, R., Saeid, E., Ali, E. S., Abdelhaq, M., Alsaqour, R., & Alharbe, N. (2023). Downlink Power Allocation for CR-NOMA-Based Femtocell D2D Using Greedy Asynchronous Distributed Interference Avoidance Algorithm. *Computers*, *12*(8), 158. https://www.mdpi.com/2073-431X/12/8/158

Ibrahim, A. M., Yau, K.-L. A., Chong, Y.-W., & Wu, C. (2021). Applications of Multi-Agent Deep Reinforcement Learning: Models and Algorithms. *Applied Sciences*, *11*(22), Article 22. https://doi.org/10.3390/app112210870

Miyamoto, M., Kobayashi, R., Kuwano, G., Tsujimoto, M., & Kakeya, I. (2024). Wide-band frequency modulation of a terahertz intrinsic Josephson junction emitter of a cuprate superconductor. *Nature Photonics*, *18*(3), 267–275. https://doi.org/10.1038/s41566-023-01348-0

Mustari, N., Karabulut, M. A., Shahen Shah, A. F. M., & Tureli, U. (2024). Cooperative THz communication for UAVs in 6G and beyond. *Green Energy and Intelligent Transportation*, *3*(1), 100131. https://doi.org/10.1016/j.geits.2023.100131

Omar, S. S., Abd El-Haleem, A. M., Ibrahim, I. I., & Saleh, A. M. (2023). Capacity Enhancement of Flying-IRS-Assisted 6G THz Network using Deep Reinforcement Learning. *IEEE Access*. https://ieeexplore.ieee.org/abstract/document/10251426/

Pan, Y., Wang, K., Pan, C., Zhu, H., & Wang, J. (2021). UAV-Assisted and Intelligent Reflecting Surfaces-Supported Terahertz Communications. *IEEE Wireless Communications Letters*, *10*(6), 1256–1260. IEEE Wireless Communications Letters. https://doi.org/10.1109/LWC.2021.3063365

Ranjha, A., Javed, M. A., Piran, Md. J., Asif, M., Hussien, M., Zeadally, S., & Frnda, J. (2024). Toward Facilitating Power Efficient URLLC Systems in UAV Networks Under Jittering. *IEEE Transactions on Consumer Electronics*, *70*(1), 3031–3041. IEEE Transactions on Consumer Electronics. https://doi.org/10.1109/TCE.2023.3305550

Ranjha, A., Javed, M. A., Srivastava, G., & Asif, M. (2023). Quasi-Optimization of Resource Allocation and Positioning for Solar-Powered UAVs. *IEEE Transactions on Network Science and Engineering*, *10*(6), 4071–4081. IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/TNSE.2023.3282870

Ranjha, A., Naboulsi, D., Emary, M. E., & Gagnon, F. (2024). Facilitating URLLC vis-á-vis UAV-Enabled Relaying for MEC Systems in 6-G Networks. *IEEE Transactions on Reliability*, 1–15. IEEE Transactions on Reliability. https://doi.org/10.1109/TR.2024.3357356

Saad, W., Bennis, M., & Chen, M. (2020). A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems. *IEEE Network*, *34*(3), 134–142. IEEE Network. https://doi.org/10.1109/MNET.001.1900287

Saleh, A. M., Omar, S. S., Abd El-Haleem, A. M., Ibrahim, I. I., & Abdelhakam, M. M. (2024). Trajectory optimization of UAV-IRS assisted 6G THz network using deep reinforcement learning approach. *Scientific Reports*, *14*(1), 18501. https://doi.org/10.1038/s41598-024-68459-8

Zawish, M., Dharejo, F. A., Khowaja, S. A., Raza, S., Davy, S., Dev, K., & Bellavista, P. (2024). AI and 6G Into the Metaverse: Fundamentals, Challenges and Future Research Trends. *IEEE Open Journal of the Communications Society*, *5*, 730–778. IEEE Open Journal of the Communications Society. https://doi.org/10.1109/OJCOMS.2024.3349465

Biographies

Nahla Nur Elmadina received her B.Sc. and M.Sc degrees from Sudan University of Science and Technology College of Electronic Engineering Dept. Telecommunication 2007, 2013. She was a lecturer in the Future University Dept. Telecommunication and Space Technology 2013_2014.she was a lecturer at Taibah University Dept. College of Computer Science and Engineering 2020 in KSA. She is pursuing her PhD with Dept. Electrical engineering, telecommunications engineering. She published several papers related to IoT, ML, AI, Next-generation wireless networks, NOMA, and optimization power.

Nour Eldin Osman is currently an Associate Professor and Head of the Department of Information Studies at Sultan Qaboos University. He has a significant academic background in Knowledge Management, Informatics, Project Management, and Information Technology. He has authored over 70 research papers and secured funding for four projects from Oman's Research Council (TRC) and three from Sultan Qaboos University. His research interests include Knowledge Management, Computer Applications, Health and Translational Informatics, Digitalization, E-learning, and Project Management. Nour has supervised more than 50 MSc and PhD students. He is a member of the

College of Arts Board and the Executive and Advisory Committee at SQU, serves as a reviewer for four international journals, and chaired the 2021 International Arab Conference on Information Technology (ACIT) at SQU.

Rashid A. Saeed (M'03, SM08) Currently he is a professor in the Computer Engineering Department, at Taif University. He is also working in the Electronics Department, at the Sudan University of Science and Technology (SUST). He was a senior researcher in Telekom Malaysia™ R&D and MIMOS. His areas of research interest include ML in wireless protocols. He has been successfully awarded 3 U.S patents in these areas. He supervised more than 50 MSc/PhD students. Rashid is a Senior member of IEEE, Member of IEM (I.E.M), SigmaXi, and SEC.

Elsadig Saeid received his B.Sc. degree from the Sudan University of Science and Technology of Sudan in 2003 and his M.Sc. degree from the same university in 2006, both electronics engineering (Communications). In 2013, he received his Ph.D. in Electrical and Electronics Engineering from University Technology PETRONAS (UTP). His research interests include Signal Processing, Image Processing, Information theory and coding, probability and Stochastic processes, and wireless communications.