Journal of King Saud University - Computer and Information Sciences 36 (2024) 102215

HOSTED BY

Contents lists available at ScienceDirect

Journal of King Saud University - Computer and Information Sciences

journal homepage: www.sciencedirect.com

Full length article

Advanced security measures in coupled phase-shift STAR-RIS networks: A DRL approach

Abdul Wahid [a], Syed Zain Ul Abideen [a], Manzoor Ahmed [b,*], Wali Ullah Khan [c], Muhammad Sheraz [d], Teong Chee Chuah [d,*], Ying Loong Lee [e]

[a] The College of Computer Science and Technology, Qingdao University, Qingdao, 266071, China
[b] School of Computer and Information Science and also with Institute for AI Industrial Technology Research, Hubei Engineering University, Xiaogan, 432000, China
[c] Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg City, L-1359, Luxembourg
[d] The Centre for Wireless Technology, Faculty of Engineering, Multimedia University, Cyberjaya, Selangor, 63100, Malaysia
[e] Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Kajang, 43000, Malaysia

ARTICLE INFO

Keywords:
Simultaneous transmitting and reflecting RIS (STAR-RIS)
Physical layer security (PLS)
Deep reinforcement learning (DRL)

ABSTRACT

The rapid development of next-generation wireless networks has intensified the need for robust security measures, particularly in environments susceptible to eavesdropping. Simultaneous transmitting and reflecting reconfigurable intelligent surfaces (STAR-RIS) have emerged as a transformative technology, offering full-space coverage by manipulating electromagnetic wave propagation. However, the inherent flexibility of STAR-RIS introduces new vulnerabilities, making secure communication a significant challenge. To overcome these challenges, we propose a deep reinforcement learning (DRL) based secure and efficient beamforming optimization strategy, leveraging the deep deterministic policy gradient (DDPG) algorithm. By framing the problem as a Markov decision process (MDP), our approach enables the DDPG algorithm to learn optimal strategies for beamforming and transmission and reflection coefficients (TARCs) configurations. This method is specifically designed to optimize phase-shift coefficients within the STAR-RIS environment, effectively managing the coupled phase shifts and complex interactions that are critical for enhancing physical layer security (PLS). Through extensive simulations, we demonstrate that our DRL-based strategy not only outperforms traditional optimization techniques but also achieves real-time adaptive optimization, significantly improving both confidentiality and network efficiency. This research addresses key gaps in secure wireless communication and sets a new standard for future advancements in intelligent, adaptable network technologies.

1. Introduction

To address the increasing demands of multiple novel applications and an increasing demand for wireless communications, reconfigurable intelligent surfaces (RIS) (Elmossallamy et al., 2020) came to be recognized as a pivotal technological advancement poised to revolutionize next-generation wireless networks (Basar et al., 2019; Ahmed et al., 2022). RIS technology, characterized by its ability to dynamically alter electromagnetic wave propagation, offers a tailored and efficient communication environment thereby optimizing signal transmission and reception. This adaptability ensures that RIS can increase the performance of wireless networks catering to the burgeoning needs for higher data rates and ubiquitous connectivity (Mu et al., 2020; Ahmed et al., 2023a). Despite the promise of RIS in enhancing network capabilities conventional RIS systems limited to reflective functionalities offer only half-space coverage. This limitation narrows their application potential and restricts the full exploitation of electromagnetic wave manipulation.

Innovatively the emergence of simultaneous transmitting and reflecting reconfigurable intelligent surfaces (STAR-RIS) has provided an advanced solution to the limitations of traditional RIS technology.

* Corresponding authors.
E-mail addresses: wahidjan999@gmail.com (A. Wahid), zain208shah@gmail.com (S.Z.U. Abideen), manzoor.achakzai@gmail.com (M. Ahmed), waliullah.khan@uni.lu (W.U. Khan), engr.msheraz@gmail.com (M. Sheraz), tcchuah@mmu.edu.my (T.C. Chuah), leeyingl@utar.edu.my (Y.L. Lee).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2024.102215
Received 11 June 2024; Received in revised form 1 September 2024; Accepted 3 October 2024
Available online 9 October 2024

Ahmed et al. (2023c), STAR-RIS extends the capabilities of RIS by not only reflecting but also transmitting signals thereby providing full space coverage and creating a more versatile and smart radio environment (SRE) (Zhang et al., 2022c). STAR-RIS can change incoming signals by modifying the transmission and reflection coefficients (TARCs) of its elements. This means that the STAR-RIS can either reflect or refract signals based on these adjustments. This dual functionality enables unprecedented flexibility in signal propagation management allowing for comprehensive coverage and enhanced network performance (Xu et al., 2022). However, the introduction of STAR-RIS brings to light significant security challenges notably the risk of eavesdropping across its full space operational domain Mozaffari-Kermani et al. (2014). The inherent flexibility of STAR-RIS although advantageous for optimizing network operations unintentionally gives rise to potential vulnerabilities that permit unauthorized user to intercept sensitive communications (Kermani et al., 2013).

As security attacks increase, robust cryptographic algorithms are essential but often lead to larger packet sizes, higher bandwidth demands, and reduced transmission reliability (Nia et al., 2015; Koziel et al., 2016; Bisheh-Niasar et al., 2021). Cryptography-based confidentiality may not be ideal for next-generation networks due to vulnerabilities from advanced adversaries and challenges in key management for distributed systems like IoT, UAVs, and vehicular networks (Dhanda et al., 2020; Ahmed et al., 2018). PLS offers a complementary solution, using wireless medium randomness to prevent eavesdropping while maintaining transmission reliability (Niu et al., 2021). PLS leverages the inherent properties of wireless channels such as fading, noise and interference to secure data transmission against eavesdropping attempts (Wang et al., 2022c). Nonetheless, the complexity of ensuring confidentiality in STAR-RIS networks is accentuated by the intricate nature of these systems especially in configurations where transmission and reflection capabilities are inherently linked. This coupling effect necessitates a reevaluation of traditional secrecy measures and the development of novel strategies tailored to the unique operational dynamics of STAR-RIS ensuring robust confidentiality while maintaining network performance (Han et al., 2022).

As the technological landscape evolves, the integration of machine learning (ML) algorithms with STAR-RIS and RIS systems opens new avenues for optimizing and securing wireless communications (Wang et al., 2020; Mirza et al., 2023), ML algorithms can analyze vast datasets to predict network behaviors identify potential secrecy threats and dynamically adjust RIS configurations to maximize efficiency and security (Mao et al., 2022; Xu et al., 2023; Ahmed et al., 2023b). This integration promises to enhance the adaptability of RIS technologies enabling them to autonomously respond to changing network conditions and emerging security challenges. In addition to the integration of deep reinforcement learning (DRL) for optimizing STAR-RIS systems, dimensionality reduction techniques can also play a crucial role in enhancing computational efficiency. For instance, Rashid et al. (2022) discuss various dimensionality reduction methods applied to IoT data, highlighting their potential to streamline data processing without significant information loss. These techniques could be valuable in further optimizing the performance of STAR-RIS systems. A new era of intelligent and secure wireless communication networks capable of satisfying the needs of future generations while protecting the confidentiality and integrity of data is ushered in by the combination of ML and PLS for STAR-RIS.

1.1. Related works

RIS have emerged as a focal point in current research heralded for their affordability, energy efficiency and versatile configurability (Wu and Zhang, 2019; Basar et al., 2019). Strategic manipulation of wireless environments is made possible by RIS creating opportunities for improved anti-jamming and anti-eavesdropping capabilities. Through precise control over signal reflection RIS can diminish the strength of eavesdropping channels while bolstering legitimate communication channels thereby elevating the secrecy rate (Yu et al., 2019; Cui et al., 2019; Hong et al., 2021; Guan et al., 2020; Niu et al., 2023). Concurrently, RISs possess the ability to amplify signals for authorized transmissions while mitigating interference from jamming sources thereby bolstering anti-jamming efficacy (Tang et al., 2021; Yang et al., 2021a; Sun et al., 2022a). Studies (Sun et al., 2022b,c) have explored scenarios where RIS-assisted networks empower legitimate users with enhanced capabilities to counteract eavesdropping and jamming attempts. Traditional RIS technology, however, is constrained to either signal reflection or transmission, necessitating that both transmitter and receiver align on the same side of the RIS for reflection-based communication or on opposite sides for transmission. This limitation confines RIS-enabled communication enhancement to a mere half-space.

Addressing this limitation, an innovative RIS architecture, known as a STAR-RIS (Liu et al., 2021; Xu et al., 2021a; Mu et al., 2022) or an intelligent omni surface (IOS) (Xu et al., 2022, 2021b), has been proposed. Unlike its conventional counterparts, STAR-RIS can manipulate both the transmission and reflection properties of incoming electromagnetic waves in unison (Liu et al., 2021). By meticulously tailoring the reflection and transmission parameters of each element within a STAR-RIS, it is possible to direct electromagnetic waves precisely along desired paths (Xu et al., 2021a; Mu et al., 2022; Xu et al., 2022), achieving passive beamforming across the entire space. This advancement from traditional RIS technologies allows for comprehensive spatial reconfigurability in communication, a leap from semi-spatial to full-spatial communication. A conventional reflection-only RIS, when receiving a signal from a base station (BS), is limited to reflecting signal to users positioned on same side as the BS, leaving those on the opposite side unreachable while STAR-RIS can provide full space coverage.

The field of STAR-RIS is still emerging, with foundational research laying the groundwork for this advanced technology. The introduction of metasurface designs capable of modulating both the amplitude and phase shifts for electromagnetic wave transmission and reflection, setting the stage for the development of STAR-RIS (Zhu et al., 2014; Zhu and Feng, 2015). To developed a conceptual physical model for STAR-RIS, showing the potential to evolve from solely reflective RIS to more versatile STAR-RIS is presented in Xu et al. (2021a). Notably, three operational protocols for (Xu et al., 2021a) were proposed, energy splitting (ES), time switching (TS) and mode switching (MS), as detailed in Mu et al. (2022). The challenge of optimizing phase shifts within STAR-RIS to enhance spectral efficiency was tackled in Zhang et al. (2020). The authors further explored the coverage capabilities of STAR-RIS (Wu et al., 2021), devising a one-dimensional search algorithm aimed at maximizing coverage area, with their simulations confirming the superior coverage benefits of STAR-RIS over traditional RIS models. In an effort to reduce transmission power within STAR-RIS-supported non-orthogonal multiple access (NOMA), networks, In Liu et al. (2022), the authors proposed a unique element-wise alternating optimization (AO) strategy, taking into account the intricacies of coupled transmission and reflection processes. Despite these advancements, it is important to recognize the limitations highlighted in Abeywickrama et al. (2020) regarding the assumed capabilities of STAR elements to provide arbitrary TARC, a scenario unlikely for passive devices, pointing to the necessity for ongoing research and development in this area.

STAR-RIS or IOS presents enhanced opportunities for bolstering PLS across the overall communication space compared to traditional RIS. This expanded coverage however introduces greater exposure to potential eavesdropper (Eve) and jamming threats from additional third party entities elevating the risk of confidentiality breaches within the network. Unlike conventional RIS systems that manage either reflection or transmission STAR-RIS requires intricate design considerations for both functions complicating the system's design process. Consequently

designing PLS solutions for STAR-RIS-assisted networks is becoming increasingly crucial. The industry growing recognition of this importance has spurred various studies aimed at enhancing confidentiality through STAR-RIS.

In the context of secure STAR-RIS networks, various strategies have been proposed to address challenges such as PLS, beamforming, and phase shift optimization. Table 1 provides a comparative overview of these approaches, highlighting their unique contributions and limitations. The comparison particularly emphasizes the role of AI-based techniques, the handling of coupled phase shifts, and the primary objectives of each study.

For instance, one study (Niu et al., 2021) employed an AO technique to enhance both the beamforming at the BS and the passive beamforming vectors at the STAR-RIS, resulting in an improved secure sum rate in a multiple input single output (MISO) system. Another work (Fang et al., 2022) explored secure transmission in an IOS-supported multiple input multiple output (MIMO) system, utilizing Lagrangian-dual and quadratic-constrained quadratic-programming techniques to jointly optimize AP beamforming and IOS phase shifts for maximum secure rates. Additionally, research by Wang et al. (2022a) focused on improving secrecy energy efficiency (SEE) in aerial secure offloading through an AO method, while (Wang et al., 2022b) emphasized the role of STAR-RIS in enhancing network secrecy within an internet-of-medical-things network by optimizing SEE.

Further studies have tackled the design of secure coupled phase-shift transmission in STAR-RIS, such as the approach presented by Zhang et al. (2023). In addition, investigations into STAR-RIS based NOMA systems have explored the use of artificial noise to enhance security, as demonstrated in works like (Wang et al., 2022c; Zhang et al., 2022a; Han et al., 2022). Interference mitigation has also been a focus, with (Hou et al., 2022) addressing signal enhancement and interference cancellation in a NOMA coordinated multi-point network. Research into IOS-assisted systems includes (Zhang et al., 2022b), which proposed a distributed hybrid beamforming algorithm for MIMO systems to boost sum rates, and Fang et al. (2023), which examined full-duplex MIMO systems with an emphasis on reducing self-interference and maximizing data transmission through enhanced IOS beamforming.

The integration of deep learning (DL) and DRL offers a forward-thinking approach to managing RISs, leading to significant advancements in channel estimation, beamforming, and security enhancements for RIS-enabled networks. DL techniques, particularly, have shown effectiveness in estimating channel state information (CSI), a critical factor in passive beamforming with RIS. For example, Gao et al. (2021) utilized a deep neural network (DNN) to precisely estimate BS-RIS and RIS-user channels, reducing pilot overhead while maintaining high accuracy. Additionally, DL has been applied to phase-shift optimization, with models like the unsupervised DL approach in Song et al. (2021) targeting the simultaneous optimization of active and passive beamforming.

DRL has further expanded the possibilities in RIS-assisted networks by enabling the joint optimization of key network components. For example, Huang et al. (2020) applied a deep deterministic policy gradient (DDPG) method to simultaneously optimize RIS phase-shifts and BS beamforming, enhancing overall network efficiency. In a related study, Zhong et al. (2022) introduced a hybrid DRL framework combining DDPG and deep Q network (DQN) algorithms to optimize both active and passive beamforming in STAR-RIS-assisted MISO systems, achieving improved energy efficiency and performance. Moreover, Yang et al. (2021b) explored the use of DQN for developing secure beamforming strategies in dynamic environments, demonstrating the approach's effectiveness in increasing secrecy rates and enhancing user satisfaction.

In addition, a new DRL framework was presented in Non-orthogonal (2020) to manage RIS in NOMA networks. This system combines a long–short-term-memory (LSTM) based echo-state-network (ESN) with a decaying double-deep Q-network (D3QN) to effectively manage the changing information needs of users. The introduction of a federated learning method for STAR-RIS in Ni et al. (2021) aimed at maximizing data rates within a heterogeneous NOMA network supported by STAR-RIS, underscores the significant potential of integrating advanced machine learning methods with RIS technologies to optimize network outcomes.

Table 1 provides an overview of the methodologies explored in recent studies on STAR-RIS networks. The proposed DRL-based approach distinguishes itself by integrating AI-based strategies with the management of coupled phase shifts, which are essential for optimizing STAR-RIS performance. While techniques like AO and other traditional methods have been used to improve secrecy rates or reduce power consumption, they often fall short in offering the real-time adaptability and continuous control needed in dynamic environments. The DRL approach, leveraging the DDPG algorithm, not only improves the minimum secrecy capacity for legitimate users but also facilitates efficient, real-time optimization of beamforming and TARCs, making it well-suited for the complex demands of STAR-RIS networks.

1.2. Motivation and contribution

The complexity of next-generation wireless networks has heightened the need for robust security, particularly in eavesdropping-prone environments. STAR-RIS technology, with its ability to control electromagnetic waves for full-space coverage, offers great potential. However, the dual role of RIS elements in both reflecting and transmitting signals results in coupled phase shifts, where changes in one process directly influence the other. This interdependence makes it challenging to optimize STAR-RIS systems effectively, and traditional methods often struggle with this complexity.

Our analysis (Table 1) shows that existing approaches frequently overlook these critical phase-shift complexities, even when leveraging AI-based methods. This oversight can lead to vulnerabilities and suboptimal network performance. To bridge this gap, we propose a DRL framework designed specifically for STAR-RIS. This approach not only optimizes phase-shift coefficients but also effectively manages the challenges posed by coupled phase shifts, enabling real-time adaptive optimization. As a result, our method significantly enhances both security and overall network performance, addressing the limitations of existing techniques. Following are the salient contributions:

- We propose a DRL framework that optimizes phase-shift coefficients in STAR-RIS networks, which is crucial for enhancing PLS against eavesdroppers.
- Our approach, utilizing the DDPG algorithm, effectively manages the complexities of coupled phase shifts and the intricate interactions within the STAR-RIS environment, resulting in more efficient management of STAR-RIS elements and enhanced security performance.
- Extensive simulations validate that our DRL-based strategy significantly outperforms traditional AO and other machine learning approaches, demonstrating substantial improvements in both secrecy and network efficiency in practical scenarios.

1.3. Organization of the article

This remaining paper is organized as follows: Section 2 (System Model) outlines the communication framework, covering the scenario, channel model, STAR-RIS design, and signal model, establishing a foundation for secure STAR-RIS-supported wireless communications. Section 3 (Problem Formulation and Proposed DRL Solution) addresses the key challenge of enhancing minimum secrecy capacity for legitimate users through transmit beamforming and TARCs interaction at STAR-RIS, introducing a novel DRL approach using the DDPG algorithm. Section 4 (Simulation) presents the simulation setup and evaluates the proposed DRL technique across various conditions, such as the

Table 1
Comparison with existing work.

Reference	Methodology	PLS	AI-based	Coupled phase shifts	Objective
Niu et al. (2021)	AO	✓	✗	✗	Maximize WSSR in a STAR-RIS-assisted MISO network by optimizing BF and TARCs
Yang et al. (2021b)	DRL with PDS and PER	✓	✓	✗	Maximize system secrecy rate in an IRS-aided wireless communication system with multiple Eves
Liu et al. (2022)	AO	✗	✗	✓	Minimize total power consumption while satisfying user rate requirements
Zhong et al. (2022)	Hybrid DRL (DDPG-DQN)	✗	✓	✓	Minimize power consumption while ensuring QoS in a STAR-RIS network
Han et al. (2022)	AO	✓	✗	✗	Maximize sum secrecy rate in a STAR-RIS assisted NOMA system
Wang et al. (2022b)	AO	✓	✗	✗	Maximize SEE in a STAR-RIS-assisted internet of medical things network
Zhang et al. (2022a)	Alternating hybrid-BF (AHB)	✓	✗	✗	Maximize minimum secrecy capacity and minimize maximum SOP in a STAR-RIS-assisted uplink NOMA network
Fang et al. (2022)	Block coordinate descent	✓	✗	✗	Maximize secrecy rate in an IOS-assisted MIMO communication network by optimizing BF and phase shifts
Zhang et al. (2023)	Penalty based secrecy BF	✓	✗	✓	Maximize minimum secrecy capacity by jointly optimizing transmit BF and phase shifts
Our study	Proposed DRL	✓	✓	✓	Enhance minimum secrecy capacity for legitimate users by jointly optimizing transmit BF and TARCs in STAR-RIS

BF—Beamforming, PDS—Prioritized experience replay, PER—post-decision state

number of STAR-RIS elements and transmit power levels, demonstrating the method's effectiveness with visual evidence. Finally, Section 5 (Conclusions) summarizes the findings, underscoring the DRL method's role in improving confidentiality and efficiency in STAR-RIS-aided communications and suggesting avenues for future research.

2. System model

2.1. Scenario

We consider a downlink wireless communication system, as illustrated in Fig. 1, where an access point (AP) with multiple antennas transmits confidential information to users via a strategically positioned STAR-RIS. This setup divides the service area into two zones: a reflective (R) zone and a transmissive (T) zone. For simplicity, we assume that each zone contains a single user with a single antenna. Additionally, an eavesdropper (Eve) with a single antenna is present in each zone, attempting to intercept the transmitted signals. The STAR-RIS overcomes physical barriers by redirecting and penetrating signals, ensuring secure and efficient communication. By simultaneously reflecting and transmitting signals, the STAR-RIS employs advanced spatial strategies to enhance both connectivity and confidentiality, underscoring the significance of innovative techniques in modern communication systems.

2.2. Channel model

For clarity, we assume that the T user (U_t) is positioned on the transmission side of the STAR-RIS, while the R user (U_r) is located on

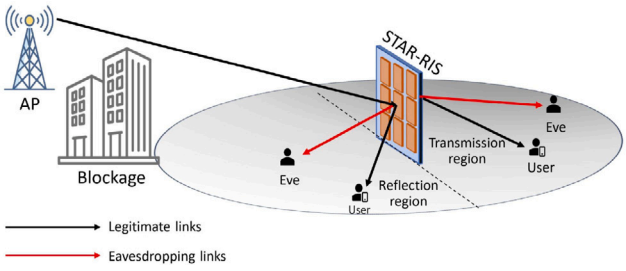

Fig. 1. STAR-RIS-assisted secure downlink scenario.

the reflection side. The eavesdroppers, T-Eve (E_{et}) and R-Eve (E_{er}), are assumed to be near U_t and U_r, respectively, to facilitate the interception of wiretapped signals. There is no direct connection between the AP and U_t, U_r, E_{et}, and E_{er} due to the barriers. From STAR-RIS to AP the baseband equivalent channels, U_t, U_r, E_{et}, and E_{er}, are represented by the complex matrix $\mathbf{G} \in \mathbb{C}^{N\times M}$, $h_{U_t,S} \in \mathbb{C}^{N\times 1}$, $h_{U_r,S} \in \mathbb{C}^{N\times 1}$, $h_{E_{et},S} \in \mathbb{C}^{N\times 1}$, and $h_{E_{er},S} \in \mathbb{C}^{N\times 1}$. These matrices and vectors represent the equivalent baseband channels from the STAR-RIS to the AP, transmission user, reflection user, and Eves, respectively. The channel matrix $\mathbf{G}$ follows a Rician fading model, while the other channels follow a Rayleigh fading model. Our primary objective is to improve secrecy by protecting against internal Eves. In this context, E_{et} and E_{er} represent other active users who have confidentiality clearance only for their own information. Nevertheless, users U_t and U_r do not

have confidence in them based on a data perspective (Niu et al., 2021; Wang et al., 2022c; Han et al., 2022; Zhang et al., 2022a). Thus, by employing the estimation techniques of parallel factor decomposition the CSI of legitimate/eavesdropping can be precisely obtained (Wei et al., 2022). Nevertheless, the Eves are the wiretapping network's exterior nodes, the AP can still utilize the unintentional leakage of local oscillator power from the Eves' radio frequency front-end to estimate CSI (Mukherjee and Swindlehurst, 2012).

2.3. Design framework for STAR-RIS

The STAR-RIS consists of N elements, each half a wavelength in size, capable of operating in various modes such as MS mode and TS mode (Liu et al., 2021). However, our focus is on the ES mode, where each element simultaneously transmits and reflects, dividing the incoming signal into two distinct components based on the amplitude coefficients β_n^{r} and β_n^{t}. The corresponding phase shifts for transmission and reflection, denoted as θ_n^{t} and θ_n^{r}, respectively, are constrained within $(0, 2\pi]$ Mu et al., 2022. Our discussion centers on the non-powered, lossless design of STAR-RIS systems, where the elements are activated solely by external signals. These signals induce magnetization and electric polarization currents, which generate the respective transmission and reflection waves. Crucially, the generation of these currents must adhere to energy conservation principles and established boundary conditions, resulting in a direct and interdependent relationship between the transmission and reflection coefficients.

Accordingly, the relationship between the TARCs for any element, labeled as n, within an array of N elements, can be mathematically represented as:

$$\beta_n^{\mathrm{t}} = 1 - \beta_n^{\mathrm{r}}, \text{for } 1 \le n \le N, \tag{1}$$

$$\left|\theta_n^{\mathrm{t}} - \theta_n^{\mathrm{r}}\right| = \frac{\pi}{2} \text{ or } \frac{3\pi}{2}, \text{for } 1 \le n \le N. \tag{2}$$

These equations define the relationship between the transmission and reflection coefficients for each element n of the STAR-RIS. The energy conservation principle dictates that the sum of the transmission and reflection amplitudes must equal one. The phase shifts are constrained to specific values to ensure proper wave propagation.

For a system composed of N elements, the matrix representing the coefficients of transmission and reflection is structured as follows:

$$\Theta_{\mathrm{t/r}} = \mathrm{diag}\left(\left[\sqrt{\beta_1^{\mathrm{t/r}}}e^{j\theta_1^{\mathrm{t/r}}}, \ldots, \sqrt{\beta_N^{\mathrm{t/r}}}e^{j\theta_N^{\mathrm{t/r}}}\right]\right), \tag{3}$$

where $\Theta_{t/r}$ is diagonal matrix, which represents the TARCs for the STAR-RIS elements. Each element of the matrix is characterized by its amplitude coefficient $\beta_{t/rn}$ and phase shift $\theta_{t/rn}$. Due to the energy conservation and coupled phase-shift constraints, the transmission and reflection coefficients are interdependent. Specifically, for each element n, the relationship between β_{t_n} and β_{r_n} is given by $\beta_{t_n} = 1 - \beta_{r_n}$, and the phase shifts θ_{t_n} and θ_{r_n} are related by $|\theta_{t_n} - \theta_{r_n}| = \frac{\pi}{2}$ or $\frac{3\pi}{2}$.

2.4. Signal model

The AP uses several beamforming vectors $\mathbf{w}_R, \mathbf{w}_T \in \mathbb{C}^{M\times 1}$ to transmit the secret signals s_R and s_T ($E\{|sU_t|^2\} = E\{|sU_r|^2\} = 1$) to the reflection zone and transmissive zone, respectively. However, the unique ES functionality of the STAR-RIS ensures that each incoming signal is evenly split between a reflected and a transmitted signal. This distribution mechanism suggests that Eves have the capability to capture signals intended to users. Consequently, the signals received at the U_t and E_{et} are described by.

$$y_\delta = \mathbf{h}_{\delta,S}^H \Theta_t^H \mathbf{G}\left(\sum_{\zeta\in\{U_t,U_r\}} \mathbf{w}_\zeta s_\zeta\right) + n_\delta, \quad \delta \in \{U_t, E_{et}\}, \tag{4}$$

where $n_{U_t}, n_{E_{et}} \sim \mathcal{CN}(0,\sigma^2)$ denote additive white Gaussian noise (AWGN) at the U_t and E_{et}, respectively. The received signal at the U_r and E_{er} are given by:

$$y_\delta = \mathbf{h}_{\delta,S}^H \Theta_r^H \mathbf{G}\left(\sum_{\zeta\in\{U_t,U_r\}} \mathbf{w}_\zeta s_\zeta\right) + n_\delta, \quad \delta \in \{U_r, E_{er}\}. \tag{5}$$

Every legitimate user decodes their own signal exclusively while regarding signal of the other user as interference when they receive superimposed signals. For legitimate users and Eves the achievable rates are given by:

$$R_\zeta = \log_2\left(1 + \frac{|\mathbf{h}_{\zeta,S}^H \Theta_t^H \mathbf{G}\mathbf{w}_\zeta|^2}{|\mathbf{h}_{\zeta,S}^H \Theta_t^H \mathbf{G}\mathbf{w}_{\bar\zeta}|^2 + \sigma^2}\right), \tag{6}$$

$$R_{E_k,\zeta} = \log_2\left(1 + \frac{|\mathbf{h}_{E_k,S}^H \Theta_t^H \mathbf{G}\mathbf{w}_\zeta|^2}{|\mathbf{h}_{E_k,S}^H \Theta_t^H \mathbf{G}\mathbf{w}_{\bar\zeta}|^2 + \sigma^2}\right), \tag{7}$$

where $k \in \{et, er\}$, $\zeta \in \{U_t, U_r\}$, and $\bar\zeta \in \{U_t, U_r\}$ with $\bar\zeta \ne \zeta$.

The secrecy capacity of the U_t or U_r is given by:

$$R_{s,\zeta} = \left[R_\zeta - \max\{R_{E_{et},\zeta}, R_{E_{er},\zeta}\}\right]^+, \tag{8}$$

where $R_{s,\zeta}$, the secrecy capacity for each user, is defined as the difference between the achievable rate for the legitimate user R_ζ and the maximum achievable rate for the Eves. Additionally, $[x]^+ = \max\{x, 0\}$.

3. Problem formulation and proposed DRL solution

3.1. Problem formulation

In our research, our primary objective is to maximize the minimum secrecy capacity for legitimate users. This is accomplished through optimizing the coupled of TARCs at the STAR-RIS and the transmit beamforming method at the AP simultaneously. This optimization takes into consideration constraints related to the budget for transmit power as well as the coupling between amplitude and phase shifts. We structure the problem as outlined below.

$$\max_{\mathbf{w}_{U_t},\mathbf{w}_{U_r},\boldsymbol{\Theta}^{\mathrm{T}},\boldsymbol{\Theta}^{\mathrm{R}}} \quad \min_{\zeta\in\{U_t,U_r\}} R_{s,e} \tag{9a}$$

$$\text{s.t.} \quad \sum_{\zeta\in\{U_t,U_r\}} \|\mathbf{w}_\zeta\|^2 \le P_{\max}, \tag{9b}$$

$$\beta_n^{\mathrm{R}} + \beta_n^{\mathrm{T}} = 1, \quad \forall n, \tag{9c}$$

$$|\theta_n^{\mathrm{T}} - \theta_n^{\mathrm{R}}| = \frac{\pi}{2} \text{ or } \frac{3\pi}{2}, \quad \theta_n^{\mathrm{T}}, \theta_n^{\mathrm{R}} \in [0, 2\pi), \quad \forall n. \tag{9d}$$

Where $P_{\max}$ shows the maximum allowable transmit power at the AP, acting as a boundary for the power resources. The constraint labeled as (9b) is indicative of the overarching limitation on power usage at the AP. Concurrently, constraints (9c) and (9d) are dedicated to specifying the relationships between amplitude and phase shifts for the STAR-RIS which is characterized by its passive and loss-free design within our study.

3.2. Proposed DRL solution

Our research aims to enhance the secrecy capacity for authorized users in wireless communications by optimizing transmit beamforming at the AP and the coupled TARCs at STAR-RIS. This involves addressing the challenges posed by transmit power constraints and the intricate interplay between amplitude and phase shifts. To tackle these issues, we introduce a novel DRL-based secure and efficient beamforming optimization strategy, leveraging the DDPG algorithm. By framing the problem as a Markov decision process (MDP), this approach allows the DDPG algorithm to learn optimal strategies for beamforming and TARC configurations within these constraints. Our goal is to significantly boost secrecy capacity while ensuring power efficiency and adhering to operational constraints, thereby setting a new benchmark in secure wireless communication.

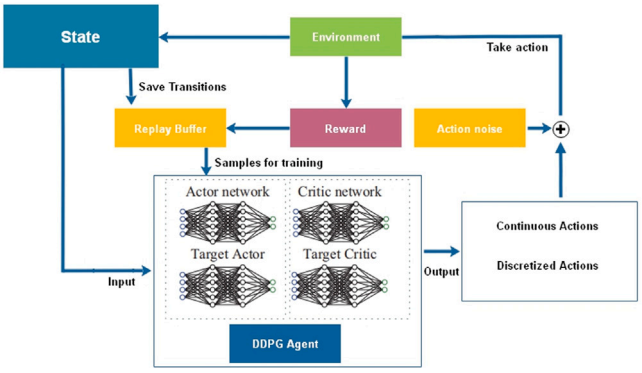

Fig. 2. Proposed DRL flow diagram illustrating the comprehensive approach of algorithm applied in STAR-RIS aided secure wireless communications.

3.3. MDP formulation

The DDPG algorithm (Lillicrap et al., 2015) has proven effective in addressing continuous control problems. To optimize STAR-RIS using DRL, we employ a MDP to model the transmission dynamics over time. The agent examines the current state $\mathbf{s}_t \in \mathbf{S}$ at each timestep $t \in T$ by analyzing the CSI of the current channels. Afterwards, it determines the appropriate action to take, denoted as $\mathbf{a}_t \in \mathbf{A}$, where $\mathbf{S}$ represent the state and $\mathbf{A}$ represent the action spaces, respectively. The action vector includes the AP and STAR-RIS beamforming coefficients, both active and passive, respectively.

- **State Space:** The MDP state at each timestep t is defined by the CSI of all channels involved in passive and active beamforming. The state $\mathbf{s}_t$ is given by:

$$\mathbf{s}_t = \left\{ \mathbf{G}, h_{U_t,S}, h_{U_r,S}, h_{E_{et},S}, h_{E_{er},S} \right\}. \tag{10}$$

  After taking action $\mathbf{a}_t$, the agent calculates the reward r_t based on the data rate and power consumption of the transceiver. The next state $\mathbf{s}_{t+1} \in \mathbf{S}$ is then determined, and the tuple $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1})$ is recorded in the replay buffer, serving as crucial data for the agent's training.
- **Action Space:** The action space for the DRL agent, $\mathbf{a}_t = \left\{ \mathbf{w}_{U_t}, \mathbf{w}_{U_r}, \boldsymbol{\Theta}_t, \boldsymbol{\Theta}_r \right\}$, includes the optimized variables $\mathbf{w}_{U_t}$ and $\mathbf{w}_{U_r}$, managed through continuous control. Assuming $\beta_n \neq 0$, the phase relationship $\theta_{t,n} = \theta_{r,n} \pm \frac{\pi}{2}$ is maintained, allowing continuous control for $\mathbf{w}_{U_t}$, $\mathbf{w}_{U_r}$, and $\boldsymbol{\Theta}_r$, while keeping $\boldsymbol{\Theta}_t$ discrete. This setup creates a hybrid action space, combining discrete and continuous elements.
  To address the optimization problem in this hybrid action space, the DDPG algorithm is applied, with a subset of continuous outputs discretized, as shown in Algorithm 1 and Fig. 2. The DDPG algorithm provides an action policy influenced by action noise, described by:

$$\mathbf{a}_t = \mu(\mathbf{s}_t|\boldsymbol{\omega}_t^{\mu}) + \mathcal{N}_{\mathrm{OU}}(0, \xi), \tag{11}$$

  where $\mu(\mathbf{s}_t|\boldsymbol{\omega}_t^{\mu})$ represents the deterministic policy and $\mathcal{N}_{\mathrm{OU}}(0, \xi)$ is the Ornstein–Uhlenbeck noise with volatility ξ.
  The continuous action space is defined as:

$$\mathbf{a}_t^c = \{\mathbf{a}_t^{\mathbf{w}_{U_r}}, \mathbf{a}_t^{\mathbf{w}_{U_t}}, \mathbf{a}_t^{\Theta_r}\}. \tag{12}$$

  The actor network's outputs are normalized and converted into actionable commands:

$$\begin{aligned} \mathbf{w}_{U_r} &\leftarrow \mathbf{a}_t^{\mathbf{w}_{U_r}}, \\ \mathbf{w}_{U_t} &\leftarrow \mathbf{a}_t^{\mathbf{w}_{U_t}}, \\ \boldsymbol{\Theta}_r &\leftarrow \mathbf{a}_t^{\Theta_r}. \end{aligned} \tag{13}$$

Algorithm 1: DRL based secure STAR-RIS beamforming

Input: Initial actor network weights, critic network weights, target actor network weights, target critic network weights, θ^{π}, θ^{Q}, $\theta^{\pi'}$, and $\theta^{Q'}$, respectively and replay buffer $\mathcal{D}$

Output: Optimized policy π^*, value function Q^*

Initialize actor network $\pi(s|\theta^{\pi})$ with weights θ^{π};
Initialize critic network $Q(s, a|\theta^{Q})$ with weights θ^{Q};
Initialize target actor network $\pi'(s|\theta^{\pi'})$ with weights $\theta^{\pi'} \leftarrow \theta^{\pi}$;
Initialize target critic network $Q'(s, a|\theta^{Q'})$ with weights $\theta^{Q'} \leftarrow \theta^{Q}$;
Initialize replay buffer $\mathcal{D}$ with size N;
Set learning rate α for actor and critic networks;
Initialize ϵ for exploration rate;
while *not converged* **do**
    Initialize a random process $\mathcal{N}$ for action exploration;
    Receive initial observation state s_1;
    for *each step of the episode* **do**
        Select action $a_t = \pi(s_t|\theta^{\pi}) + \mathcal{N}_t$;
        Execute a_t and observe r_t and s_{t+1};
        Record the transition (s_t, a_t, r_t, s_{t+1}) into $\mathcal{D}$;
        Sample a random minibatch of N transitions (s_i, a_i, r_i, s_{i+1}) from $\mathcal{D}$;
        Set $y_i = r_i + \gamma Q'(s_{i+1}, \pi'(s_{i+1}|\theta^{\pi'})|\theta^{Q'})$;
        Update Critic via minimizing the loss: $L = \frac{1}{N}\sum_i (y_i - Q(s_i, a_i|\theta^{Q}))^2$;
        Update Actor utilizing the sampled policy gradient;
        Update target networks:;
        $\theta^{\pi'} \leftarrow \tau\theta^{\pi} + (1-\tau)\theta^{\pi'}$;
        $\theta^{Q'} \leftarrow \tau\theta^{Q} + (1-\tau)\theta^{Q'}$;
    end
    Update ϵ;
end

For each STAR element n, the transmission phase $\theta_{t,n}$ is derived from the binary discretized action $a_{n,t}^d$ by:

$$\theta_{t,n} = \begin{cases} \theta_{r,n} + \frac{\pi}{2}, & \text{if } a_{n,t}^d > 0, \\ \theta_{r,n} - \frac{\pi}{2}, & \text{if } a_{n,t}^d \leq 0. \end{cases} \tag{14}$$

- **Reward:** The reward function is designed to maximize secrecy capacity while minimizing power consumption, aligning with the operational constraints of the system. The reward function $\mathbb{R}_t$ is expressed as:

$$\mathbb{R}_t = R_{s,\zeta}(t) - \lambda P_t, \tag{15}$$

  where $R_{s,\zeta}(t)$ is the secrecy rate at time t, P_t is the power consumed at time t, and λ is a weighting factor that balances the importance of power consumption relative to the secrecy rate.

3.4. Training the DDPG agent

The aim of training a DRL agent is to discover for each state $\mathbf{s}_t$ the optimal path of action $\mathbf{a}_t$ that maximizes $\mathbb{E}\left[\sum_{i=t}^{T} \gamma r_{t+1}\right]$ in terms of cumulative expected reward. γ, where $\gamma \in [0, 1]$ is the discount factor. The Bellman equation is used to determine the action value a_t for the DDPG agent, also known as Q:

$$Q^{\mu}\left(\mathbf{s}_t, \mathbf{a}_t\right) = \mathbb{E}\left[r\left(\mathbf{s}_t, \mathbf{a}_t\right) + \gamma Q^{\mu}\left(\mathbf{s}_{t+1}, \mathbf{a}_{t+1}\right)\right]. \tag{16}$$

This process excludes full reflection mode, which is less desirable. The action policy function μ maps states $\mathbf{S}$ to actions $\mathbf{A}$. The objective

DRL agent training is identify the optimal action that maximizes the Q value:

$$Q^*\left(\mathbf{s}_t,\mathbf{a}_t\right)=\mathbb{E}\left[r\left(\mathbf{s}_t,\mathbf{a}_t\right)+\max_{\mathbf{a}\in\mathbf{A}}\gamma Q^*\left(\mathbf{s}_{t+1},\mathbf{a}_{t+1}\right)\right]. \tag{17}$$

An actor function $\mu\left(\mathbf{s}\mid\boldsymbol{\omega}^{\mu}\right)$ is parameterized, and a function approximator is associated with a set of parameters ω^{Q}. The agent learns by minimizing the loss function through sampling transition experiences from the replay buffer:

$$L\left(\boldsymbol{\omega}^{Q}\right)=\frac{1}{e}\sum_{e}\left[y_t-Q\left(\mathbf{s}_t,\mathbf{a}_t\mid\boldsymbol{\omega}_t^{Q}\right)\right]^2. \tag{18}$$

Here, e represents the number of sampled transitions. To prevent training instability, the target network, mirroring the training network's structure but updating parameters more slowly, provides y_t:

$$y_t=r_t\left(\mathbf{s}_t,\mathbf{a}_t\right)+\gamma Q'\left[\mathbf{s}_t,\mu'\left(\mathbf{s}_t\mid\boldsymbol{\omega}_t^{\mu'}\right)\mid\boldsymbol{\omega}_t^{Q'}\right]. \tag{19}$$

The actor network in the DDPG algorithm is trained using a policy gradient from the critic network's:

$$\nabla_{\boldsymbol{\omega}^{\mu}}J=\frac{1}{e}\sum_{e}\nabla_{\mathbf{a}}Q\left(\mathbf{s}_e,\mathbf{a}_e\mid\boldsymbol{\omega}^{Q}\right)\Big|_{\mathbf{s}_e=\mathbf{s}_t,\mathbf{a}_e=\mu(\mathbf{s}_t)}\nabla_{\boldsymbol{\omega}^{\mu}}\mu\left(\mathbf{s}_e\mid\boldsymbol{\omega}^{\mu}\right)_{\mathbf{s}_e=\mathbf{s}_t}. \tag{20}$$

3.5. Structure of neural network

To guarantee precise model fitting, it is essential to make an acceptable selection regarding the architecture and size of both target actor networks and target critic networks. Actor networks should consist of one or more activation layers that use the 'relu' activation function in a sequential manner, an input layer, and a batch normalization (BN) layer. As an additional measure, the output layer makes use of a 'tanh' activation function in order to ensure that the outputs are within a range that is considered to be acceptable. In order to ensure that the output layer continues to receive valid input values, an additional BN layer is positioned immediately before it. The structure of critic networks consists of an input layer, a BN layer, a concatenate layer that merges state and action inputs, and many activation layers. In order to accommodate the dimensions of state and action spaces, which are determined by number of antennae and number of STAR components N, the hidden layer dimensions must be changed to ensure compatibility with the required dimensions.

3.5.1. Complexity analysis

The hybrid control is achieved through a combination of discrete and continuous control mechanisms. The discrete control aspect of the DRL framework involves discretizing the phase shifts of the STAR-RIS elements to maintain a balance between transmission and reflection. Specifically, the discrete control ensures that each element's phase shift difference satisfies $|\theta_n^t-\theta_n^r|=\frac{\pi}{2}$ or $\frac{3\pi}{2}$, as shown in Eq. (2). On the other hand, the continuous control mechanism continuously adjusts the amplitude coefficients β_n^t and β_n^r, ensuring the energy conservation constraint $\beta_n^t+\beta_n^r=1$.

The proposed DRL approach achieves hybrid control by utilizing a specially designed mapping function, eliminating the need for additional DNNs, thus keeping complexity low and simplifying the training process with a DNN having fewer trainable parameters. The complexity of the DDPG algorithm is determined by the specifications of the employed DNN. The complexity of propagation for an actor network with I layers, each containing $\boldsymbol{\omega}_i^{\mu}$ nodes (Qiu et al., 2019), is expressed as $\sum_{i=0}^{I}\boldsymbol{\omega}_i^{\mu}\boldsymbol{\omega}_{i+1}^{\mu}$, and the number of floating-point operations required is $5\boldsymbol{\omega}_{\mathrm{b}}^{\mu}+\boldsymbol{\omega}_{\mathrm{r}}^{\mu}+6\boldsymbol{\omega}_{\mathrm{t}}^{\mu}$.

Applying this theory to critic networks, the complexity for a single prediction and training step is given by $\mathcal{O}\left(\sum_{i=0}^{I}\boldsymbol{\omega}_i^{\mu}\boldsymbol{\omega}_{i+1}^{\mu}+\sum_{i=0}^{I}\boldsymbol{\omega}_i^{Q}\boldsymbol{\omega}_{i+1}^{Q}+5\boldsymbol{\omega}_{\mathrm{b}}^{\mu}+\boldsymbol{\omega}_{\mathrm{r}}^{\mu}+6\boldsymbol{\omega}_{\mathrm{t}}^{\mu}+5\boldsymbol{\omega}_{\mathrm{b}}^{Q}+\boldsymbol{\omega}_{\mathrm{r}}^{Q}+6\boldsymbol{\omega}_{\mathrm{t}}^{Q}\right)$, which, given that typically $\boldsymbol{\omega}_i\gg 5$, can be approximated by $\mathcal{O}\left(\sum_{i=0}^{I}\boldsymbol{\omega}_i^{\mu}\boldsymbol{\omega}_{i+1}^{\mu}+\sum_{i=0}^{I}\boldsymbol{\omega}_i^{Q}\boldsymbol{\omega}_{i+1}^{Q}\right)$. Since the proposed DRL algorithm has a smaller output and DNN scale, where $\boldsymbol{\omega}_i^{\mu,h}<\boldsymbol{\omega}_i^{\mu}$, it is consequently less complex than the conventional DDPG algorithm.

Table 2
Simulations Parameters.

Parameters	Description	Value
L_0	Ref. path loss	−30 dB
α_{AP}	Path loss exp. AP	2.2
$\alpha_{U_tS},\alpha_{U_rS},$ $\alpha_{E_{et}S},\alpha_{E_{er}S}$	Path loss exp.	2.5
σ^2	Noise power density	−105 dBm
M	Antenna no.	4
N	Element no.	12
B	Bandwidth	1 MHz
U_r	R-zone User	1
U_t	T-zone User	1
E_{er}	R-zone Eve	1
E_{et}	T-zone Eve	1
$P_{\max}$	Max power/antenna	29 dBm
κ	Rician factors	5
κ	Rayleigh factors	0
r	Replay buffer size	10 000
γ	Discount factor	1
DDPG	Batch size	32 samples
DDPG	Target update rate	0.002

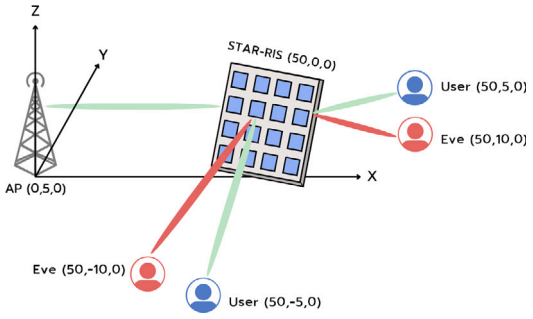

Fig. 3. The considered network simulations setup.

4. Simulation results

The study examines simulation results that confirm the algorithm's efficacy. As seen in Fig. 3, our investigation unfolds within a three-dimensional coordinate framework. At its core, the AP anchors the network from the (0,0,0) meter (m) mark, establishing the foundational node of our communication model. The deployment of the STAR-RIS at (50,0,0) m and the strategic positioning of the U_t and U_r at (50,5,0) m and (50,−5,0) m, accordingly. This configuration is purposefully designed to simulate a diverse range of user environments, reflecting the algorithm's adaptability and robustness in real-world applications. To assess the networks secrecy against potential eavesdropping hypothetical adversaries namely Eves E_{et} and E_{er} are posited at (50,10,0) m and (50,−10,0) m. This aspect of the simulation is crucial for evaluating the algorithms ability to maintain secrecy in the presence of Eves. The proposed DRL simulation parameters are listed in Table 2.

The channel model is meticulously formulated to encompass both large-scale path loss and small-scale fading offering a comprehensive view of the signal propagation dynamics. The channel is mathematically characterized as follows:

$$\mathbf{h}=\sqrt{L_0 d^{-\alpha}}\left(\sqrt{\frac{\kappa}{1+\kappa}}\bar{\mathbf{h}}+\sqrt{\frac{1}{1+\kappa}}\tilde{\mathbf{h}}\right). \tag{21}$$

In this equation, $\bar{\mathbf{h}}$ and $\tilde{\mathbf{h}}$ symbolize the LoS and NLoS components, respectively, critical for a holistic understanding of signal transmission. The variable L_0 denotes the reference path loss, while α is the path loss d is the transmission distance, respectively, essential for quantifying signal attenuation. The Rician factor κ distinguishes between Rician channels, indicative of a predominant direct path, and Rayleigh channels, characterized by multipath scattering and $\kappa=0$ for Rayleigh

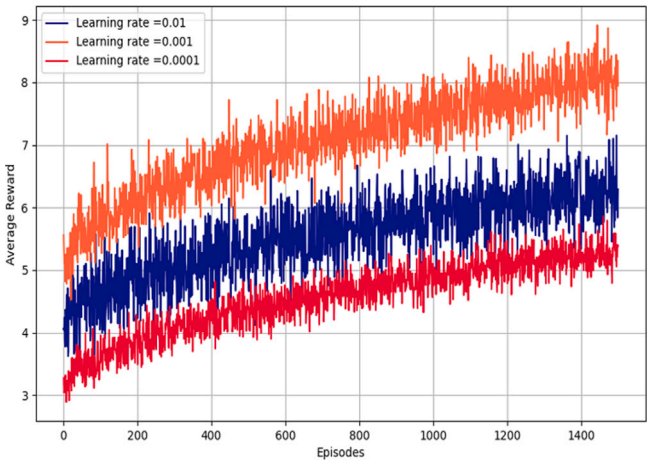

Fig. 4. Average reward per episode for different learning rates.

channels and $\kappa = 5$ for Rician channels. Simulation parameters are carefully selected to mirror real-world conditions: $L_0 = -30$ dB sets the stage for initial path loss exponents, $\alpha_{AP} = 2.2$ and $\alpha_{U_iS} = \alpha_{U_rS} = \alpha_{E_{et}S} = \alpha_{E_{et}S} = 2.5$ adjust for various propagation environments. The DDPG replay buffer size, batch size and target update rate is 10000, 32 samples, and 0.002 respectively, are standard settings for stable and effective DRL training. The noise power is established at $\sigma^2 = -105$ dBm, providing a realistic noise scenario.

Fig. 4 shows the average reward per episode for different learning rates ($\alpha = \{0.01, 0.001, 0.0001\}$) in the context of our DRL approach applied to STAR-RIS networks. The choice of learning rate significantly impacts the performance of the DRL algorithm: A high learning rate ($\alpha = 0.01$) leads to erratic and oscillatory behavior, resulting in significantly lower reward performance. This is likely due to the learning rate being too high, causing the model to overshoot the optimal policy frequently and fail to converge to a stable solution. An optimal learning rate ($\alpha = 0.001$) strikes a balance, allowing the model to effectively learn and converge to an optimal policy within a reasonable timeframe. This rate facilitates a steady increase in the average reward, indicating stable and efficient learning. Conversely, a low learning rate ($\alpha = 0.0001$) ensures stable learning but necessitates a much longer duration for the model to converge to an optimal policy. The convergence is slow, which can be impractical for real-time applications where quick adaptation is required. In summary, the average reward changes with varying learning rates, with $\alpha = 0.001$ identified as the optimal rate for achieving efficient learning and high performance in the proposed DRL approach. This balance ensures that the model neither overshoots the optimal policy nor takes too long to converge, making it suitable for dynamic and complex environments like STAR-RIS-aided wireless communications.

Fig. 5 compares the minimum secrecy capacities of different STAR-RIS configurations across varying levels of transmit power at the AP. The configurations examined include coupled phase-shift, which uses simultaneous transmission and reflection with coupled phases; independent phase-shift, allowing separate control over transmission and reflection phases; and random phase-shift, where phase settings are randomized. The results show that both the coupled and independent phase-shift configurations surpass the random phase-shift in terms of secrecy capacity. The independent phase-shift configuration, due to its flexible phase control, achieves the highest secrecy capacity, especially at higher transmit powers, reflecting its superior time resource utilization. These findings highlight the importance of strategic phase control in enhancing the secrecy performance of STAR-RIS systems in wireless communications.

While Fig. 6 shows the minimum secrecy capacity vs the number of STAR-RIS elements. For all phase configurations, a clear correlation has been identified between the minimum secrecy capacity and the number of elements. The coupled phase configuration exhibits enhanced performance relative to the other two configurations, as the number of elements increases with a more notable rise in secrecy capacity. The capacity of the independent phase configuration increases gradually as the number of elements increases, whereas the random phase configuration demonstrates the least improvement and, generally, lower secrecy capacity. As a result, Figs. 5 and 6 highlight an important realization of coupled phase configuration, which leverages both the increased transmit power and the number of STAR-RIS elements more effectively than the independent or random configurations, emphasizing its importance role in enhancing the secrecy performance of wireless networks enabled by STAR-RIS.

Fig. 7 compares three different algorithms to compare the minimum secrecy capacity vs transmit power that includes AO, DQN, and a specially developed DRL approach. As the transmit power increases, all algorithms achieve higher secrecy capacity, benefiting from the higher signal strength which can be exploited to make the signal more distinguishable from noise and potential eavesdropping. However, the proposed DRL algorithm demonstrates the most efficient use of additional power to improve secrecy capacity, indicating its potential for energy efficient secure wireless communication approach. While Fig. 8 presents a comparison of various algorithms against the number of STAR-RIS elements, illustrating how all algorithms improve secrecy capacity with an increasing the number of STAR-RIS elements, but the proposed DRL approach consistently outperforms the others, suggesting that can better exploit the capabilities of STAR-RIS. DQN also improves secrecy capacity with more elements but reach the performance of the Proposed DRL, while AO shows the least improvement, indicating its limitations in dynamic environments.

4.1. Discussion and analysis

The simulation results offer a comprehensive evaluation of the proposed scheme in the context of STAR-RIS-aided wireless communication networks, particularly focusing on improving secrecy capacity—a vital metric for secure communication systems. The interconnected insights provided by various figures collectively enhance the understanding of how to optimize network security.

Beginning with the analysis of learning rates on the DRL approach, Fig. 4 reveals the average reward per episode for different learning rates. The results show that the optimal learning rate ($\alpha = 0.001$) effectively balances learning efficiency and stability. A higher learning rate ($\alpha = 0.01$) leads to unstable learning, characterized by erratic behavior and reduced rewards due to overshooting the optimal policy. Conversely, a lower learning rate ($\alpha = 0.0001$) ensures stability but slows down convergence, which is not ideal for applications requiring real-time performance. This balance is crucial in environments like STAR-RIS-aided wireless networks, where rapid and stable learning is necessary.

Following this, Fig. 5 compares the minimum secrecy capacities of different STAR-RIS configurations as a function of transmit power. The independent phase-shift configuration is shown to be the most effective, particularly at higher transmit powers, due to its ability to independently manage transmission and reflection phases. This control allows for better resource utilization and enhanced security. The coupled phase configuration also performs well, while the random phase-shift configuration consistently shows lower performance, highlighting the need for precise phase control in these systems.

The impact of the number of STAR-RIS elements on secrecy capacity is further explored in Fig. 6. As the number of elements increases, all phase configurations show improved performance. The coupled phase configuration, in particular, benefits the most, leveraging the additional elements to significantly boost security. The independent phase configuration also shows gradual improvements, while the random phase configuration demonstrates the least enhancement, emphasizing the

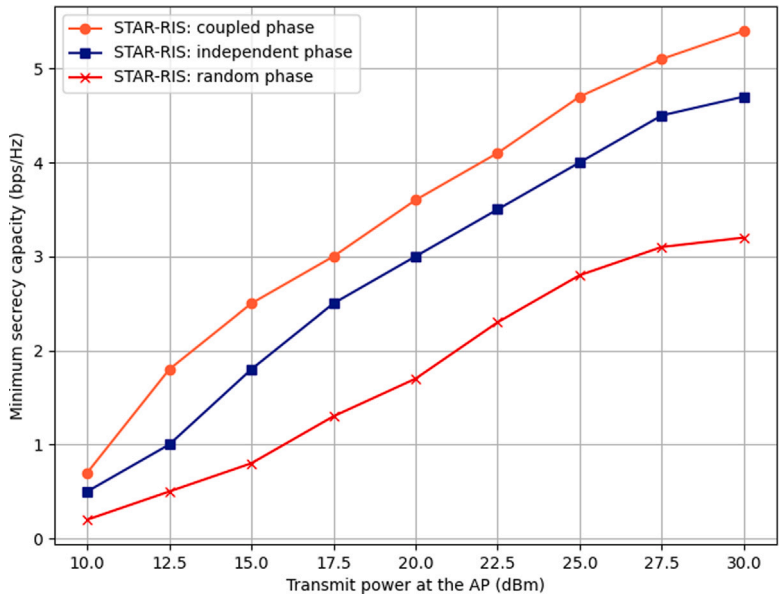

Fig. 5. Minimum secrecy capacity vs transmit power.

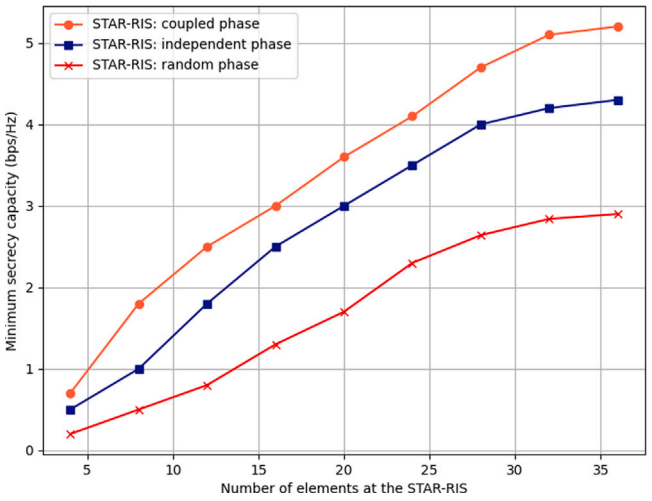

Fig. 6. Minimum secrecy capacity vs number of STAR-RIS elements.

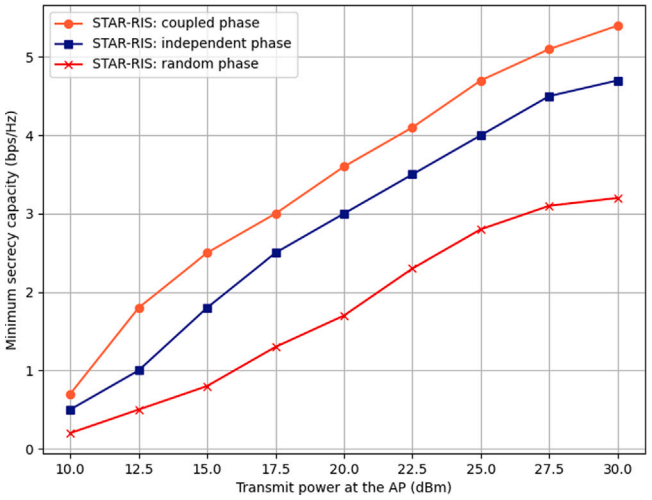

Fig. 7. Minimum secrecy capacity against transmit power.

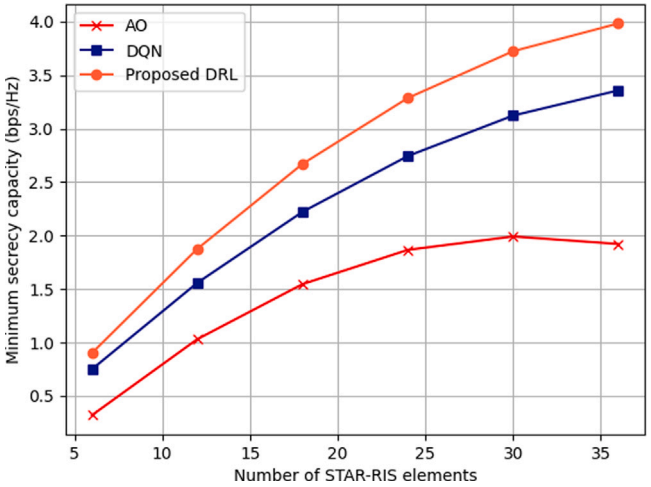

Fig. 8. Minimum secrecy capacity vs increasing number of STAR-RIS elements.

importance of coherent phase management as the complexity of the network increases.

Lastly, Figs. 7 and 8 compare the effectiveness of different algorithms — AO, DQN, and the proposed DRL approach — under varying conditions of transmit power and the number of STAR-RIS elements. The DRL approach consistently outperforms the others, particularly as transmit power increases and more STAR-RIS elements are incorporated, demonstrating its superior adaptability and efficiency. This makes the DRL method a promising solution for improving secrecy capacity in complex and dynamic wireless environments.

The discussion also considers the impact of uncertainties, such as imperfect CSI and environmental variability. These uncertainties can challenge the accuracy of phase shift configurations and reduce the effectiveness of learning algorithms. However, the DRL approach, with its ability to adapt continuously to changing conditions, proves to be robust even under uncertain conditions—a flexibility that more static methods like AO and DQN struggle to achieve.

These findings highlight the critical role of advanced algorithmic strategies and optimized configurations in enhancing the security of STAR-RIS-aided networks, especially in uncertain environments. The

ability to precisely control phase shifts and utilize advanced DRL techniques significantly improves secrecy capacity, ensuring robust security even in challenging scenarios. This study confirms the effectiveness of the proposed methodologies and paves the way for future research aimed at further optimizing secure communication in next-generation wireless networks.

5. Conclusions

This study represents a significant advancement in secure wireless communication by employing STAR-RIS technology and leveraging the DDPG algorithm to address the complexities of coupled phase shifts and enhance secrecy in full-space communications. By demonstrating a method that dynamically adapts to varying network conditions and significantly outperforms traditional optimization techniques, this research underscores the potential of STAR-RIS to elevate both confidentiality and network efficiency. The results from our extensive simulations showcase the superior capability of our DDPG-based framework in real-time, adaptive optimization, setting a new benchmark for integrating machine learning with STAR-RIS in next-generation wireless networks. Looking forward, future research should aim to overcome the challenges posed by imperfect or delayed CSI by developing robust frameworks that can adapt to CSI uncertainties. This could involve integrating Bayesian reinforcement learning or implementing real-time feedback mechanisms to dynamically adjust to changing channel conditions. Additionally, exploring the deployment of multiple STAR-RIS units in dense networks could provide valuable insights into scalability and interference management. Expanding the current approach to multi-agent scenarios, where multiple STAR-RIS elements and users collaborate, could further enhance system robustness and efficiency. These advancements would not only address the limitations related to CSI assumptions in this study but also pave the way for more resilient and adaptable wireless communication systems.

CRediT authorship contribution statement

Abdul Wahid: Writing – original draft, Methodology, Investigation. **Syed Zain Ul Abideen:** Writing – review & editing, Formal analysis. **Manzoor Ahmed:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Wali Ullah Khan:** Writing – review & editing, Investigation. **Muhammad Sheraz:** Writing – review & editing, Resources. **Teong Chee Chuah:** Writing – review & editing, Investigation, Funding acquisition. **Ying Loong Lee:** Writing – review & editing.

Declaration of competing interest

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

Acknowledgment

This work was supported by the Multimedia University Research Fellow under Grant MMUI/240021.

References

Abeywickrama, S., Zhang, R., Wu, Q., Yuen, C., 2020. Intelligent reflecting surface: Practical phase shift model and beamforming optimization. IEEE Trans. Commun. 68 (9), 5849–5863. http://dx.doi.org/10.1109/TCOMM.2020.3001125.

Ahmed, M., Ali, Z., Khan, W.U., Waqar, O., Asif, M., Javed, M.A., Al-Wesabi, F.N., et al., 2022. Cooperative backscatter NOMA with imperfect SIC: Towards energy efficient sum rate maximization in sustainable 6G networks. J. King Saud Univ., Comput. Inf. Sci. 34 (10), 7940–7947.

Ahmed, M., Alshahrani, H.M., Alruwais, N., Asiri, M.M., Al Duhayyim, M., Khan, W.U., Nauman, A., et al., 2023a. Joint optimization of UAV-IRS placement and resource allocation for wireless powered mobile edge computing networks. J. King Saud Univ., Comput. Inf. Sci. 35 (8), 101646.

Ahmed, M., Liu, J., Mirza, M.A., Khan, W.U., Al-Wesabi, F.N., 2023b. MARL based resource allocation scheme leveraging vehicular cloudlet in automotive-industry 5.0. J. King Saud Univ., Comput. Inf. Sci. 35 (6), 101420.

Ahmed, M., Shi, H., Chen, X., Li, Y., Waqas, M., Jin, D., 2018. Socially aware secrecy-ensured resource allocation in D2D underlay communication: An overlapping coalitional game scheme. IEEE Trans. Wireless Commun. 17 (6), 4118–4133. http://dx.doi.org/10.1109/TWC.2018.2820693.

Ahmed, M., Wahid, A., Laique, S.S., Khan, W.U., Ihsan, A., Xu, F., Chatzinotas, S., Han, Z., 2023c. A survey on STAR-RIS: Use cases, recent advances, and future research challenges. IEEE Internet Things J. 1. http://dx.doi.org/10.1109/JIOT.2023.3279357.

Basar, E., Di Renzo, M., De Rosny, J., Debbah, M., Alouini, M.-S., Zhang, R., 2019. Wireless communications through reconfigurable intelligent surfaces. IEEE Access 7, 116753–116773. http://dx.doi.org/10.1109/ACCESS.2019.2935192.

Bisheh-Niasar, M., Azarderakhsh, R., Mozaffari-Kermani, M., 2021. High-speed NTT-based polynomial multiplication accelerator for post-quantum cryptography. In: 2021 IEEE 28th Symposium on Computer Arithmetic. ARITH, IEEE, Lyngby, Denmark, pp. 94–101.

Cui, M., Zhang, G., Zhang, R., 2019. Secure wireless communication via intelligent reflecting surface. IEEE Wirel. Commun. Lett. 8 (5), 1410–1414. http://dx.doi.org/10.1109/LWC.2019.2919685.

Dhanda, S.S., Singh, B., Jindal, P., 2020. Lightweight cryptography: A solution to secure IoT. Wirel. Pers. Commun. 112 (3), 1947–1980.

Elmossallamy, M.A., Zhang, H., Song, L., Seddik, K.G., Han, Z., Li, G.Y., 2020. Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities. IEEE Trans. Cogn. Commun. Netw. 6, 990–1002.

Fang, S., Chen, G., Abdullah, Z., Li, Y., 2022. Intelligent omni surface-assisted secure MIMO communication networks with artificial noise. IEEE Commun. Lett. 26 (6), 1231–1235. http://dx.doi.org/10.1109/LCOMM.2022.3159575.

Fang, S., Chen, G., Xiao, P., Wong, K.-K., Tafazolli, R., 2023. Intelligent omni surface-assisted self-interference cancellation for full-duplex MISO system. IEEE Trans. Wireless Commun. 1. http://dx.doi.org/10.1109/TWC.2023.3297071.

Gao, S., Dong, P., Pan, Z., Li, G.Y., 2021. Deep multi-stage CSI acquisition for reconfigurable intelligent surface aided MIMO systems. IEEE Commun. Lett. 25 (6), 2024–2028. http://dx.doi.org/10.1109/LCOMM.2021.3063464.

Guan, X., Wu, Q., Zhang, R., 2020. Intelligent reflecting surface assisted secrecy communication: Is artificial noise helpful or not? IEEE Wirel. Commun. Lett. 9 (6), 778–782. http://dx.doi.org/10.1109/LWC.2020.2969629.

Han, Y., Li, N., Liu, Y., Zhang, T., Tao, X., 2022. Artificial noise aided secure NOMA communications in STAR-ris networks. IEEE Wirel. Commun. Lett. 11 (6), 1191–1195. http://dx.doi.org/10.1109/LWC.2022.3161020.

Hong, S., Pan, C., Ren, H., Wang, K., Chai, K.K., Nallanathan, A., 2021. Robust transmission design for intelligent reflecting surface-aided secure communication systems with imperfect cascaded CSI. IEEE Trans. Wireless Commun. 20 (4), 2487–2501. http://dx.doi.org/10.1109/TWC.2020.3042828.

Hou, T., Wang, J., Liu, Y., Sun, X., Li, A., Ai, B., 2022. A joint design for STAR-RIS enhanced NOMA-CoMP networks: A simultaneous-signal-enhancement-and-cancellation-based (SSECB) design. IEEE Trans. Veh. Technol. 71 (1), 1043–1048. http://dx.doi.org/10.1109/TVT.2021.3129178.

Huang, C., Mo, R., Yuen, C., 2020. Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning. IEEE J. Sel. Areas Commun. 38 (8), 1839–1850. http://dx.doi.org/10.1109/JSAC.2020.3000835.

Kermani, M.M., Zhang, M., Raghunathan, A., Jha, N.K., 2013. Emerging frontiers in embedded security. In: 2013 26th International Conference on VLSI Design and 2013 12th International Conference on Embedded Systems. IEEE, pp. 203–208.

Koziel, B., Azarderakhsh, R., Kermani, M.M., Jao, D., 2016. Post-quantum cryptography on FPGA based on isogenies on elliptic curves. IEEE Trans. Circuits Syst. I. Regul. Pap. 64 (1), 86–99.

Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D., 2015. Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971.

Liu, Y., Mu, X., Schober, R., Poor, H.V., 2022. Simultaneously transmitting and reflecting (STAR)-RISs: A coupled phase-shift model. In: ICC 2022 - IEEE International Conference on Communications. Seoul, Korea, Republic of.

Liu, Y., Mu, X., Xu, J., Schober, R., Hao, Y., Poor, H.V., Hanzo, L., 2021. STAR: Simultaneous transmission and reflection for 360° coverage by intelligent surfaces. IEEE Wirel. Commun. 28 (6), 102–109. http://dx.doi.org/10.1109/MWC.001.2100191.

Mao, Y., Pranolo, A., Hernandez, L., Wibawa, A.P., Nuryana, Z., 2022. Artificial intelligence in mobile communication: A survey. IOP Conf. Ser. Mater. Sci. Eng. 1212 (1), 012046. http://dx.doi.org/10.1088/1757-899x/1212/1/012046.

Mirza, M.A., Yu, J., Raza, S., Krichen, M., Ahmed, M., Khan, W.U., Rabie, K., Shongwe, T., 2023. DRL-assisted delay optimized task offloading in automotive-industry 5.0 based VECNs. J. King Saud Univ., Comput. Inf. Sci. 35 (6), 101512.

Mozaffari-Kermani, M., Sur-Kolay, S., Raghunathan, A., Jha, N.K., 2014. Systematic poisoning attacks on and defenses for machine learning in healthcare. IEEE J. Biomed. Health Inform. 19 (6), 1893–1905.

Mu, X., Liu, Y., Guo, L., Lin, J., Al-Dhahir, N., 2020. Exploiting intelligent reflecting surfaces in NOMA networks: Joint beamforming optimization. IEEE Trans. Wireless Commun. 19 (10), 6884–6898. http://dx.doi.org/10.1109/TWC.2020.3006915.

Mu, X., Liu, Y., Guo, L., Lin, J., Schober, R., 2022. Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications. IEEE Trans. Wireless Commun. 21 (5), 3083–3098. http://dx.doi.org/10.1109/TWC.2021.3118225.

Mukherjee, A., Swindlehurst, A.L., 2012. Detecting passive eavesdroppers in the MIMO wiretap channel. In: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, pp. 2809–2812. http://dx.doi.org/10.1109/ICASSP.2012.6288501.

Ni, W., Liu, Y., Eldar, Y.C., Yang, Z., Tian, H., 2021. STAR-RIS enabled heterogeneous networks: Ubiquitous NOMA communication and pervasive federated learning. ArXiv. org..

Nia, A.M., Mozaffari-Kermani, M., Sur-Kolay, S., Raghunathan, A., Jha, N.K., 2015. Energy-efficient long-term continuous personal health monitoring. IEEE Trans. Multi-Scale Comput. Syst. 1 (2), 85–98.

Niu, H., Chu, Z., Zhou, F., Zhu, Z., 2021. Simultaneous transmission and reflection reconfigurable intelligent surface assisted secrecy MISO networks. IEEE Commun. Lett. 25 (11), 3498–3502. http://dx.doi.org/10.1109/LCOMM.2021.3109164.

Niu, H., Lin, Z., Chu, Z., Zhu, Z., Xiao, P., Nguyen, H.X., Lee, I., Al-Dhahir, N., 2023. Joint beamforming design for secure RIS-assisted IoT networks. IEEE Internet Things J. 10 (2), 1628–1641. http://dx.doi.org/10.1109/JIOT.2022.3210115.

Non-orthogonal, R.E.M., 2020. Multiple access networks: Deployment and passive beamforming design. arXiv preprint arXiv:2001.10363.

Qiu, C., Hu, Y., Chen, Y., Zeng, B., 2019. Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications. IEEE Internet Things J. 6 (5), 8577–8588. http://dx.doi.org/10.1109/JIOT.2019.2921159.

Rashid, L., Rubab, S., Alhaisoni, M., Alqahtani, A., Alsubai, S., Binbusayyis, A., Bukhari, S.A.C., 2022. Analysis of dimensionality reduction techniques on Internet of Things data using machine learning. Sustain. Energy Technol. Assess. 52, 102304.

Song, H., Zhang, M., Gao, J., Zhong, C., 2021. Unsupervised learning-based joint active and passive beamforming design for reconfigurable intelligent surfaces aided wireless networks. IEEE Commun. Lett. 25 (3), 892–896. http://dx.doi.org/10.1109/LCOMM.2020.3041510.

Sun, Y., An, K., Luo, J., Zhu, Y., Zheng, G., Chatzinotas, S., 2022a. Outage constrained robust beamforming optimization for multiuser IRS-assisted anti-jamming communications with incomplete information. IEEE Internet Things J. 9 (15), 13298–13314. http://dx.doi.org/10.1109/JIOT.2022.3140752.

Sun, Y., An, K., Zhu, Y., Zheng, G., Wong, K.-K., Chatzinotas, S., Ng, D.W.K., Guan, D., 2022b. Energy-efficient hybrid beamforming for multilayer RIS-assisted secure integrated terrestrial-aerial networks. IEEE Trans. Commun. 70 (6), 4189–4210. http://dx.doi.org/10.1109/TCOMM.2022.3170632.

Sun, Y., An, K., Zhu, Y., Zheng, G., Wong, K.-K., Chatzinotas, S., Yin, H., Liu, P., 2022c. RIS-assisted robust hybrid beamforming against simultaneous jamming and eavesdropping attacks. IEEE Trans. Wireless Commun. 21 (11), 9212–9231. http://dx.doi.org/10.1109/TWC.2022.3174629.

Tang, X., Wang, D., Zhang, R., Chu, Z., Han, Z., 2021. Jamming mitigation via aerial reconfigurable intelligent surface: Passive beamforming and deployment optimization. IEEE Trans. Veh. Technol. 70 (6), 6232–6237. http://dx.doi.org/10.1109/TVT.2021.3077662.

Wang, C.-X., Di Renzo, M., Stańczak, S., Wang, S., Larsson, E.G., 2020. Artificial intelligence enabled wireless networking for 5G and beyond: Recent advances and future challenges. arXiv (Cornell University), 10.48550/arxiv.2001.08159.

Wang, W., Ni, W., Tian, H., Song, L., 2022a. Intelligent omni-surface enhanced aerial secure offloading. IEEE Trans. Veh. Technol. 71 (5), 5007–5022. http://dx.doi.org/10.1109/TVT.2022.3150769.

Wang, W., Ni, W., Tian, H., Yang, Z., Huang, C., Wong, K.-K., 2022b. Robust design for STAR-RIS secured internet of medical things. In: 2022 IEEE International Conference on Communications Workshops, ICC Workshops. Seoul, Korea, Republic of.

Wang, W., Ni, W., Tian, H., Yang, Z., Huang, C., Wong, K.-K., 2022c. Safeguarding NOMA networks via reconfigurable dual-functional surface under imperfect CSI. IEEE J. Sel. Top. Sign. Proces. 16 (5), 950–966. http://dx.doi.org/10.1109/JSTSP.2022.3175013.

Wei, L., Huang, C., Guo, Q., Yang, Z., Zhang, Z., Alexandropoulos, G.C., Debbah, M., Yuen, C., 2022. Joint channel estimation and signal recovery for RIS-empowered multiuser communications. IEEE Trans. Commun. 70 (7), 4640–4655. http://dx.doi.org/10.1109/TCOMM.2022.3179771.

Wu, C., Liu, Y., Mu, X., Gu, X., Dobre, O.A., 2021. Coverage characterization of STAR-RIS networks: NOMA and OMA. IEEE Commun. Lett. 25 (9), 3036–3040. http://dx.doi.org/10.1109/LCOMM.2021.3091807.

Wu, Q., Zhang, R., 2019. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. IEEE Trans. Wireless Commun. 18 (11), 5394–5409. http://dx.doi.org/10.1109/TWC.2019.2936025.

Xu, F., Ahmad, S., Ahmed, M., Raza, S., Khan, F., Ma, Y., Khan, W.U., et al., 2023. Beyond encryption: Exploring the potential of physical layer security in UAV networks. J. King Saud Univ., Comput. Inf. Sci. 101717.

Xu, J., Liu, Y., Mu, X., Dobre, O.A., 2021a. STAR-RISs: Simultaneous transmitting and reflecting reconfigurable intelligent surfaces. IEEE Commun. Lett. 25 (9), 3134–3138. http://dx.doi.org/10.1109/LCOMM.2021.3082214.

Xu, J., Liu, Y., Mu, X., Zhou, J.T., Song, L., Poor, H.V., Hanzo, L., 2021b. Simultaneously transmitting and reflecting (STAR) intelligent omni-surfaces, their modeling and implementation. arXiv preprint arXiv:2108.06233.

Xu, J., Liu, Y., Mu, X., Zhou, J.T., Song, L., Poor, H.V., Hanzo, L., 2022. Simultaneously transmitting and reflecting intelligent omni-surfaces: Modeling and implementation. IEEE Veh. Technol. Mag. 17 (2), 46–54. http://dx.doi.org/10.1109/MVT.2022.3157069.

Yang, H., Xiong, Z., Zhao, J., Niyato, D., Wu, Q., Poor, H.V., Tornatore, M., 2021a. Intelligent reflecting surface assisted anti-jamming communications: A fast reinforcement learning approach. IEEE Trans. Wireless Commun. 20 (3), 1963–1974. http://dx.doi.org/10.1109/TWC.2020.3037767.

Yang, H., Xiong, Z., Zhao, J., Niyato, D., Xiao, L., Wu, Q., 2021b. Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications. IEEE Trans. Wireless Commun. 20 (1), 375–388. http://dx.doi.org/10.1109/TWC.2020.3024860.

Yu, X., Xu, D., Schober, R., 2019. Enabling secure wireless communications via intelligent reflecting surfaces. In: 2019 IEEE Global Communications Conference. GLOBECOM, Waikoloa, HI, USA, pp. 1–6. http://dx.doi.org/10.1109/GLOBECOM38437.2019.9014322.

Zhang, Z., Chen, J., Liu, Y., Wu, Q., He, B., Yang, L., 2022a. On the secrecy design of STAR-RIS assisted uplink NOMA networks. IEEE Trans. Wireless Commun. 21 (12), 11207–11221. http://dx.doi.org/10.1109/TWC.2022.3190563.

Zhang, Y., Di, B., Zhang, H., Han, Z., Poor, H.V., Song, L., 2022b. Meta-wall: Intelligent omni-surfaces aided multi-cell MIMO communications. IEEE Trans. Wireless Commun. 21 (9), 7026–7039. http://dx.doi.org/10.1109/TWC.2022.3154041.

Zhang, Z., Wang, Z., Liu, Y., He, B., Lv, L., Chen, J., 2023. Security enhancement for coupled phase-shift STAR-RIS networks. IEEE Trans. Veh. Technol. 72 (6), 8210–8215. http://dx.doi.org/10.1109/TVT.2023.3243545.

Zhang, H., Zeng, S., Di, B., Tan, Y., Di Renzo, M., Debbah, M., Han, Z., Poor, H.V., Song, L., 2022c. Intelligent omni-surfaces for full-dimensional wireless communications: Principles, technology, and implementation. IEEE Commun. Mag. 60, 39–45.

Zhang, S., Zhang, H., Di, B., Tan, Y., Han, Z., Song, L., 2020. Beyond intelligent reflecting surfaces: Reflective-transmissive metasurface aided communications for full-dimensional coverage extension. IEEE Trans. Veh. Technol. 69 (11), 13905–13909. http://dx.doi.org/10.1109/TVT.2020.3024756.

Zhong, R., Liu, Y., Mu, X., Chen, Y., Wang, X., Hanzo, L., 2022. Hybrid reinforcement learning for STAR-RISs: A coupled phase-shift model based beamformer. IEEE J. Sel. Areas Commun. 40 (9), 2556–2569. http://dx.doi.org/10.1109/JSAC.2022.3192053.

Zhu, B.O., Chen, K., Jia, N., Sun, L., Zhao, J., Jiang, T., Feng, Y., 2014. Dynamic control of electromagnetic wave propagation with the equivalent principle inspired tunable metasurface. Sci. Rep. 4 (1), 4971.

Zhu, B.O., Feng, Y., 2015. Passive metasurface for reflectionless and arbitary control of electromagnetic wave transmission. IEEE Trans. Antennas and Propagation 63 (12), 5500–5511. http://dx.doi.org/10.1109/TAP.2015.2481479.