ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

RIS-Assisted Heterogeneous Collaborative Secure Offloading for Multi-Server MEC Systems based on HATD3

Xiaojuan Bai*, Ao Gao, Tianxiang Liu, and Xiao Ma

Northwest Normal University, Lanzhou, Gansu, 730070, China

*Corresponding author: Xiaojuan Bai (Email: baixiaojuan@nwnu.edu.cn)

Abstract

To address physical-layer eavesdropping threats and heterogeneous hybrid decision complexity in Reconfigurable Intelligent Surface (RIS)-assisted multi-server Multi-access Edge Computing (MEC) systems, this paper proposes a heterogeneous collaborative secure offloading method based on HATD3. First, a secure offloading model involving user equipment (UE), edge servers, RIS, and an eavesdropper is established for a multi-server MEC system under communication, computation, energy consumption, and secrecy constraints. The resulting problem is then formulated as a Decentralized Partially Observable Markov Decision Process (DEC-POMDP) for heterogeneous multi-agent coordination. Then, to handle the hybrid action space composed of continuous power control, discrete server selection, and RIS resource allocation, a Heterogeneous-Agent Twin Delayed Deep Deterministic Policy Gradient (HATD3)-based heterogeneous action decoupling and mapping mechanism is developed. Simulation results demonstrate that, compared with MATD3 and the no-RIS baseline, the proposed method achieves superior secure offloading performance in complex wireless environments, maintaining the secrecy rate around 7.0 Mbps, reducing the security violation rate to below 10%, and achieving a better trade-off among latency, energy consumption, and security.

Keywords

Multi-access Edge Computing; Secure Computation Offloading; Resource Allocation; Heterogeneous Multi-Agent Deep Reinforcement Learning; Physical Layer Security; Reconfigurable Intelligent Surface.

1. Introduction

With the expansion of Multi-access Edge Computing (MEC) network topologies, multi-server collaborative computing has emerged as an important deployment paradigm [1, 2]. In such scenarios, the secure offloading problem becomes significantly more challenging because the system must jointly consider open-channel transmission security and the high-dimensional decision complexity caused by heterogeneous node coordination [3, 4]. Because of the broadcast nature of wireless communication, offloaded task data are highly susceptible to interception by malicious eavesdroppers [5].

To address this physical-layer security issue, the reconfigurable intelligent surface (RIS) has emerged as a promising wireless enhancement technology [6]. By reconfiguring the wireless propagation environment, an RIS can improve the signal quality of legitimate receivers while reducing the eavesdropper's reception capability [7]. However, introducing RIS into multi-server MEC systems also enlarges the decision space, since the resulting joint optimization problem involves both

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

continuous variables, such as transmission power and RIS phase configuration, and discrete variables, such as server selection and RIS subarray allocation.

Existing multi-agent deep reinforcement learning (MADRL) methods often rely on parameter sharing, which is more suitable for homogeneous agents than for the considered system with heterogeneous user devices, edge servers, and RIS control entities. In RIS-assisted multi-server MEC secure offloading, these agents differ substantially in observation spaces, action structures, and physical decision roles, making conventional homogeneous MADRL schemes difficult to apply effectively. Although heterogeneous MARL (HARL) provides a useful theoretical basis [8], designing a practical action decoupling and mapping mechanism that can coordinate heterogeneous agents while handling coupled continuous and discrete decision variables remains a critical challenge.

To address these challenges, this paper proposes a heterogeneous collaborative secure offloading framework, termed CHESS. The main contributions of this paper are summarized as follows: (1) A RIS-assisted multi-server MEC secure offloading system model is established by jointly considering communication, computation, energy consumption, and security constraints in the presence of an eavesdropper. (2) The secure offloading problem is formulated as a decentralized partially observable Markov decision process (DEC-POMDP). Based on this formulation, a HATD3-based action decoupling and mapping mechanism is designed to handle heterogeneous agents and the coupled continuous-discrete hybrid action space. (3) Extensive simulations are conducted to validate the proposed framework, and the results demonstrate its superiority over baseline schemes in secure offloading performance and overall system efficiency.

2. System Model

2.1 Network Architecture

We consider a RIS-assisted multi-server MEC system within an $L \times L$ square area, as shown in Fig. 1. The system comprises a set of M edge base stations (BSs) denoted as $\mathcal{M} = \{1, \ldots, M\}$ each equipped with an MEC server having U_m parallel computation units operating at a computing frequency f_m^{unit}. Additionally, each BS can transmit a cooperative jamming signal with a maximum power $p_m^{\max}$ to suppress the eavesdropper. A set of N single-antenna User Equipments (UEs) denoted as $\mathcal{N} = \{1, \ldots, N\}$ are distributed within the coverage area, each possessing a maximum CPU frequency $f_n^{\max}$, maximum transmission power $p_n^{\max}$, and a battery with capacity $b_n^{\max}$.The system also includes an RIS with N_{RIS} passive reflection elements divided into K_{RIS} sub-arrays, where each sub-array can serve at most one user per time slot, a single-antenna passive eavesdropper (Eve), and an SDN controller responsible for collecting global channel state information and coordinating resources.

The system operates in discrete time slots $t \in \{1, 2, \ldots, T\}$. At the beginning of each slot, UE n generates a computation task characterized by a 4-tuple:

$$\mathcal{T}_n(t) = \{z_n(t), c_n(t), \tau_n(t), r_n^{\text{th}}(t)\} \tag{1}$$

where $z_n(t)$ denotes the input data size in bit, $c_n(t)$ denotes the required CPU cycles per bit, $\tau_n(t)$ denotes the delay constraint, and $r_n^{\text{th}}(t)$ denotes the minimum required secrecy rate.

A partial offloading model is adopted, where $\alpha_n(t) \in [0, 1]$ denotes the offloading ratio of UE n, where a fraction $\alpha_n(t)$ of the task is offloaded to the selected MEC server m_n^* and the remaining fraction $1 - \alpha_n(t)$ is processed locally.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

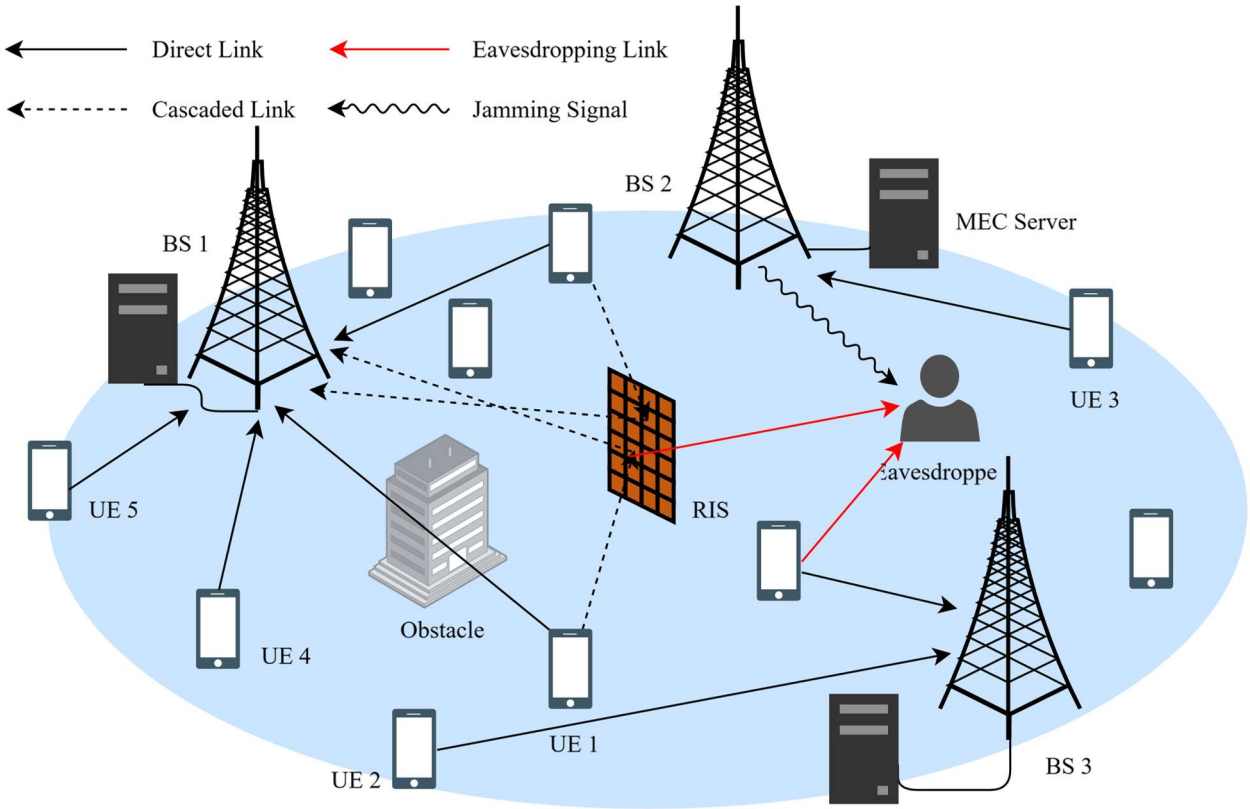

Figure 1. System Model

2.2 Channel Model

The channel coefficients are modeled following the approach in [9], which accounts for both large-scale and small-scale fading. The large-scale fading, encompassing path loss and shadowing, is characterized by a three-slope path loss model with correlated shadowing as detailed in [9], while the small-scale fading is updated using a first-order autoregressive process to capture temporal correlations.

For the RIS-assisted links, the equivalent composite channel gain for n user to m BS is given by the superposition of the direct path and the reflected paths via RIS subarrays. This can be expressed as:

$$G_{n,m}(t) = H^{d}_{n,m}(t) + \sum_{k=1}^{K_{\mathrm{RIS}}} I_{n,k}(t)(\mathbf{H}^{rb}_{k,m}(t))^{H}\boldsymbol{\Theta}_{k}(t)\mathbf{H}^{ur}_{n,k}(t) \tag{2}$$

where $H^{d}_{n,m}(t)$ is the direct UE-BS channel, $\mathbf{H}^{ur}_{n,k}(t)$ is the UE-RIS channel, $\mathbf{H}^{rb}_{k,m}(t)$ is the RIS-BS channel, and $\boldsymbol{\Theta}_{k}(t)$ is the diagonal phase shift matrix of subarray k .The indicator $I_{n,k}(t)$ denotes whether subarray k serves user n. Because the phases are optimized for the target BS, the signal at the eavesdropper, denoted as $G_{n,e}(t)$ combines incoherently, thereby providing a physical layer security advantage.

2.3 Secure Communication Model

Using Orthogonal Frequency Division Multiple Access (OFDMA), each sub-channel has bandwidth W. The signal-to-noise ratio at the legitimate BS and the signal-to-interference-plus-noise ratio at Eve can be respectively written as:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

$$\mathrm{SNR}_{n,m}^{\mathrm{leg}}(t)=\frac{p_n(t)\,|G_{n,m}(t)|^2}{\sigma^2} \tag{3}$$

$$\mathrm{SINR}_{n,e}^{\mathrm{eve}}(t)=\frac{p_n(t)|G_{n,e}(t)|^2}{\sigma^2+\sum_{m\in\mathcal{M}} p_m^{\mathrm{jam}}(t)|G_{m,e}(t)|^2}, \tag{4}$$

$$C_S \tag{5}$$

where $p_n(t)\le p_n^{\max}$ is user n's transmission power, $p_j^{\mathrm{jam}}(t)\le p_m^{\max}$ is BS m's jamming power, and $\sigma^2=N_0W$ is the noise power.

The achievable rates for the legitimate and the eavesdropping links are given by:

$$R_{n,m}(t)=W\log_2(1+\mathrm{SNR}_{n,m}^{\mathrm{leg}}(t)) \tag{6}$$

$$R_{n,e}(t)=W\log_2(1+\mathrm{SINR}_{n,e}^{\mathrm{eve}}(t)) \tag{7}$$

respectively. Following the foundational works on physical layer security [10, 11], the instantaneous secrecy rate is defined as the non-negative part of their difference, expressed as:

$$R_n^{\mathrm{sec}}(t)=[R_{n,m}(t)-R_{n,e}(t)]^+ \tag{8}$$

where $[x]^+=\max(0,x)$.

2.4 Computation and Energy Models

For local computing, the portion of the task processed locally is $z_n^{\mathrm{loc}}(t)=(1-\alpha_n(t))z_n(t)$. The corresponding delay and energy consumption follow the Dynamic Voltage and Frequency Scaling model, given by:

$$T_n^{\mathrm{loc}}(t)=\frac{z_n^{\mathrm{loc}}(t)c_n(t)}{f_n(t)} \tag{9}$$

$$E_n^{\mathrm{loc}}(t)=\kappa[f_n(t)]^2 z_n^{\mathrm{loc}}(t)c_n(t) \tag{10}$$

where κ is the effective capacitance coefficient.

For edge offloading, the offloaded portion $\alpha_n(t)$ is transmitted to the target m^*. The transmission delay and energy are:

$$T_n^{\mathrm{trans}}(t)=\frac{\alpha_n(t)z_n(t)}{R_{n,m^*}(t)} \tag{11}$$

$$E_n^{\mathrm{trans}}(t)=p_n(t)T_n^{\mathrm{trans}}(t) \tag{12}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

Upon reaching the server, tasks wait in a parallel queue. The total offloading delay, considering the earliest server availability $T_m^{\text{ear}}(t)$, is:

$$T_n^{\text{off}}(t) = \max(T_n^{\text{trans}}(t), T_m^{\text{ear}}(t)) + \frac{\alpha_n(t) z_n(t) c_n(t)}{f_m^{\text{unit}}}. \tag{13}$$

Since the UE and server process in parallel, the total task delay and energy are:

$$T_n^{\text{total}}(t) = \max(T_n^{\text{loc}}(t), T_n^{\text{off}}(t)) \tag{14}$$

$$E_n^{\text{total}}(t) = E_n^{\text{loc}}(t) + E_n^{\text{trans}}(t) \tag{15}$$

respectively.

2.5 Joint Optimization Problem Formulation

The objective is to minimize the long-term weighted system cost. The per-slot cost consists of normalized latency cost $C_T(t)$, energy cost $C_E(t)$, and security cost $C_S(t)$, combined as:

$$Cost(t) = \omega_T C_T(t) + \omega_E C_E(t) + \omega_S C_S(t). \tag{16}$$

These components are defined as:

$$C_T(t) = \frac{1}{N} \sum_{n=1}^{N} \frac{\min(T_n^{\text{total}}(t), \tau_{\max})}{\tau_{\max}}, \tag{17}$$

$$C_E(t) = \frac{1}{N} \sum_{n=1}^{N} \frac{E_n^{\text{total}}(t)}{E_{n,\max}^{\text{total}}} + \frac{1}{M} \sum_{m=1}^{M} \frac{E_m^{\text{jam}}(t)}{E_{m,\max}^{\text{jam}}}, \tag{18}$$

$$C_S(t) = \frac{1}{N} \sum_{n=1}^{N} \begin{cases} \dfrac{r_n^{\text{th}} - R_n^{\text{sec}}(t)}{r_n^{\text{th}}}, & \text{if } \alpha_n > 0 \text{ and } R_n^{\text{sec}}(t) < r_n^{\text{th}} \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

The overall optimization problem is subject to several constraints. The offloading ratio, UE transmission power, and UE CPU frequency must each remain within their respective bounds of $[0, 1]$, $[0, p_n^{\max}]$, and $[0, f_n^{\max}]$. BS jamming power is constrained to $[0, p_m^{\max}]$. UE battery levels must stay within $[0, b_n^{\max}]$, and task latency cannot exceed the maximum tolerable delay $\tau_n(t)$. The RIS allocation indicators $I_{n,k}(t)$ are binary, with each sub-array serving at most one user and the total number of served users not exceeding K_{RIS}. Finally, the selected BS for each user must belong to the set $\mathcal{M}$.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

$$
\begin{aligned}
(\mathbf{P1}) \quad & \min_{\mathbf{A}^{UE},\mathbf{A}^{BS},\mathbf{A}^{RIS}} \sum_{t=0}^{T-1} Cost(t) \\
\text{s.t.} \quad & \mathrm{C1}: \alpha_n(t) \in [0,1], \forall n \in \mathcal{N} \\
& \mathrm{C2}: p_n(t) \in [0, p_n^{max}], \forall n \in \mathcal{N} \\
& \mathrm{C3}: f_n(t) \in [0, f_n^{max}], \forall n \in \mathcal{N} \\
& \mathrm{C4}: p_m^{jam}(t) \in [0, p_m^{max}], \forall m \in \mathcal{M} \\
& \mathrm{C5}: b_n(t) \in [b_n^{min}, b_n^{max}], \forall n \in \mathcal{N} \\
& \mathrm{C6}: T_n(t) \le \tau_n(t), \forall n \in \mathcal{N} \\
& \mathrm{C7}: I_n(t) \in \{0,1\}, \ \textstyle\sum_{n \in \mathcal{N}} I_n(t) \le K_{RIS} \\
& \mathrm{C8}: m_n^*(t) \in \mathcal{M}, \forall n \in \mathcal{N}
\end{aligned}
\tag{20}
$$

3. Methodology

3.1 DEC-POMDP Modeling

To solve the optimization problem using multi-agent reinforcement learning, we model it as a Decentralized Partially Observable Markov Decision Process (DEC-POMDP), defined by the tuple: $\langle S, \{\mathcal{O}_i\}, \{\mathcal{A}_i\}, \mathcal{P}, \mathcal{R}, \gamma \rangle$. The system has three types of heterogeneous agents: N UE agents, M BS agents, and one RIS agent. The global state $S(t)$ contains complete information including task queues, battery levels, and full channel state information, but is used only during centralized training. Each agent perceives the environment through a unique observation space tailored to its role. For each UE agent n, the local observation is defined as:

$$
\mathbf{o}_n^{\mathrm{UE}}(t) = [\hat{r}_n^{\mathrm{th}}, \hat{z}_n, \hat{c}_n, \hat{\tau}_n, \hat{b}_n, \hat{f}_n^{\max}, \hat{p}_n^{\max}, \mathbf{g}_n^{\mathrm{leg}}(t)] \tag{21}
$$

where $\mathbf{g}_n^{\mathrm{leg}}(t)$ are channel gains to all M BSs. For each BS agent m, the observation is

$$
\mathbf{o}_m^{\mathrm{BS}}(t) = [\hat{U}_m, \hat{f}_m^{\mathrm{unit}}, \hat{p}_m^{\max}, \bar{G}_m^{\mathrm{leg}}(t), \bar{G}_m^{\mathrm{eve}}(t)] \tag{22}
$$

encompassing its physical properties and perceived average channel qualities.

The RIS agent observes an $N \times (5 + 2M)$ feature matrix $\mathbf{O}^{\mathrm{RIS}}(t)$, where each row:

$$
\mathbf{x}_n^{\mathrm{RIS}}(t) = \left[\hat{u}_n(t), \hat{b}_n(t), \hat{r}_n^{\mathrm{th}}(t), \bar{G}_n^{ru}(t), \bar{G}_n^{eu}(t), \mathbf{g}_n^{ub}(t), \mathbf{g}^{rb}(t)\right] \in \mathbb{R}^{5+2\mathrm{M}} \tag{23}
$$

combining task urgency, battery, security threshold, and various channel information for user n.

The hybrid action space is handled through a preference mapping mechanism. Each UE agent n outputs a continuous vector:

$$
\mathbf{a}_n^{\mathrm{UE}}(t) = [\alpha_n, f_n^{\mathrm{ratio}}, p_n^{\mathrm{ratio}}, s_{n,1}, \ldots, s_{n,M}] \tag{24}
$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

where the first three components map directly to physical resources, and the last M values represent continuous preferences for BS selection. The physical action $m_n^* = \text{argmax}_j\{s_{n,j}\}$ is derived by the environment, while the Critic evaluates the original continuous vector to maintain gradient flow. Each BS agent m outputs a continuous scala $a_m^{\text{jam}}(t) \in [0, 1]$, which is mapped to jamming power:

$$p_m^{\text{jam}}(t) = a_m^{\text{jam}}(t) \cdot p_m^{\max} \tag{25}$$

The RIS agent outputs a continuous priority vector $\mathbf{a}^{\text{RIS}}(t) = \mathbf{w}(t) \in [0, 1]^N$ The environment selects the top K_{RIS} users based on the priority weights for subarray allocation, and uses expert knowledge of channel state information to perform deterministic user-centric phase alignment, thereby avoiding high-dimensional phase control.

A shared team reward is defined to guide all agents toward the optimization goal:

$$Reward(t) = -Cost(t) - \omega_P Penalty(t) \tag{26}$$

where the penalty term discourages deadline violations and battery depletion. The state transition probability $\mathcal{P}$ is determined by the random generation of new tasks, the AR-based update of small-scale fading, and the dynamic update of battery levels based on total energy consumption.

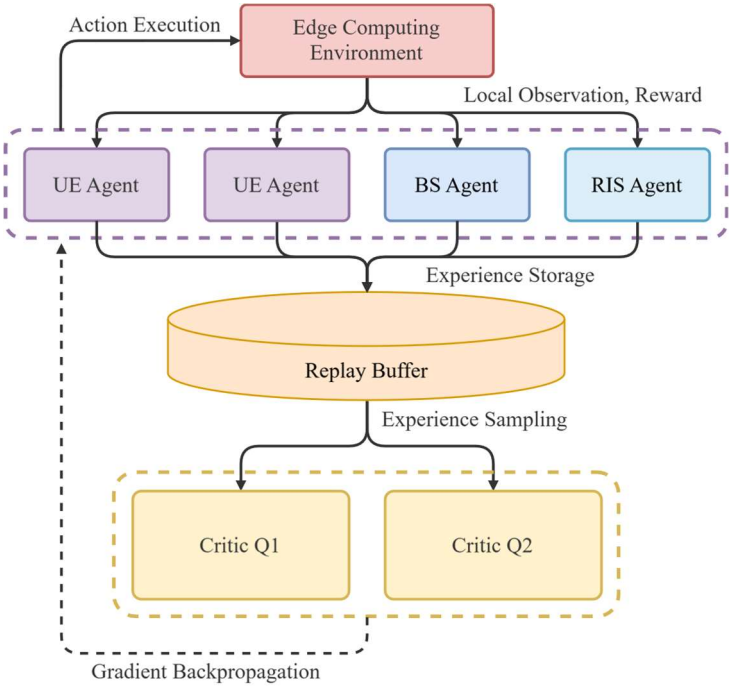

Figure 2. Flowchart of the proposed algorithm

3.2 The HATD3-based CHESS Framework

We employ the HATD3 algorithm [8] within a Centralized Training with Decentralized Execution framework to solve the DEC-POMDP. HATD3 is particularly well-suited for this problem due to its heterogeneous architecture, which avoids parameter sharing and allows each UE, BS, and the RIS agent to maintain independent Actor networks tailored to their specific observation and action spaces. This architectural choice respects the fundamental differences in the physical attributes and decision roles of each agent type.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

Furthermore, HATD3 effectively handles the hybrid action space through its Critic design. By having the centralized Critic evaluate the original continuous action vectors-including the BS selection preferences-before environment discretization, gradient information can flow back through the otherwise non-differentiable argmax operation. The incorporation of Twin Critics and Delayed Updates [12] mitigates the overestimation bias common in actor-critic methods and ensures stable learning in this complex hybrid action space, enabling effective collaboration among heterogeneous agents.

4. Experiments and Analysis of Results

4.1 Experimental Setup

Simulations were conducted on a server equipped with an NVIDIA GTX 1660 Ti GPU featuring 6GB of memory. The environment was implemented using Python 3.13 and PyTorch 2.6.0, building upon the HARL library for multi-agent reinforcement learning. The simulation area is configured as an $L \times L$ square containing $M = 3$ BSs and $N = 20$ UEs. Key system parameters, listed in Table 1, follow values established in [9, 13], while the deep reinforcement learning hyperparameters are detailed in Table 2. The proposed CHESS framework utilizing HATD3 was compared against two baselines: a homogeneous MATD3 algorithm and a version without RIS assistance denoted as HATD3-no-RIS. Each experiment was repeated with 5 random seeds to ensure statistical reliability.

Table 1. System Environment Parameters

Parameter Description	Symbol	Value
Area side length	L	500 m
Number of edge BSs/servers	M	3
Number of user equipments	N	20
Task data size	z_n	[1, 50] KB
Task computation density	c_n	[500, 1000] cycles/bit
Maximum tolerable task delay	τ_n	[0.3, 1] s
Minimum task Secrecy Rate Threshold	r_n^{th}	[1, 5] Mbps
UE transmission power	$p_n^{\max}$	[20, 24] dBm
Maximum UE CPU frequency	$f_n^{\max}$	[0.4, 1.5] GHz
UE battery capacity	$b_n^{\max}$	[2, 10] J
Effective capacitance coefficient	κ	1×10^{-28}
Number of server computation cores	U_m	[4, 8]
Server single-core computation frequency	f_m^{unit}	[2, 6] GHz
BS jamming power	p_m^{jam}	[30, 37] dBm
Total number of RIS reflection elements	N_{RIS}	256
Number of RIS sub-arrays	K_{RIS}	8
Wireless channel bandwidth	W	1 MHz

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

Table 2. Training Hyperparameters

Parameter Description	Value
Episode length	100
Total training steps	1,000,000
UE Actor network hidden sizes	[128, 128]
BS Actor network hidden sizes	[32, 32]
RIS Actor network hidden sizes	[512, 512]
Critic network hidden sizes	[1024, 512]
Actor learning rate	1.0×10^{-4}
Critic learning rate	1.0×10^{-3}
Replay buffer size	100,000
Batch size	256
Discount factor	0.99
Target network soft update coefficient	0.005
Policy update delay	2
Exploration noise	0.1
Policy noise	0.5

4.2 Results and Analysis

The convergence behavior and system reward are illustrated in Figure 3. Following an initial exploration phase characterized by significant oscillations due to the high-dimensional hybrid action space, all algorithms gradually converge after approximately 400,000 training steps. HATD3 achieves the highest and most stable average episode reward upon convergence, with MATD3 performing second and the HATD3-no-RIS baseline lagging behind. This demonstrates the effectiveness of HATD3's heterogeneous decoupling architecture and confirms the critical role of RIS assistance in enhancing overall system performance.

Figure 4,5 presents the physical layer security performance metrics. With RIS assistance and heterogeneous algorithm scheduling, HATD3 achieves a legitimate link rate that steadily climbs and stabilizes above 7.5 Mbps, with the system secrecy rate firmly maintained around 7.0 Mbps. In contrast, both MATD3 and the no-RIS baseline see their secrecy rates decline to below 6.5 Mbps in the later stages of training. The secrecy violation rate for HATD3 drops below 10%, significantly outperforming the two comparison algorithms and validating the effectiveness of the proposed collaborative security framework.

Figures 6~8 depicts the latency and energy consumption metrics. HATD3 consistently maintains the lowest average UE energy consumption, stabilizing at approximately 0.008 J. Without RIS assistance, UEs must increase transmission power to meet the same secrecy rate threshold, causing energy consumption in the HATD3-no-RIS baseline to rise to approximately 0.010 J. MATD3's energy performance falls in between, constrained by the limitations of homogeneous network parameter sharing that prevent fine-grained physical isolation of heterogeneous terminal power control. Regarding latency and deadline violation rates, all three algorithms perform comparably, with average task delay converging below 0.2 seconds and violation rates around 2%. This indicates that the proposed framework successfully achieves a balanced optimization among energy consumption, latency, and security requirements.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

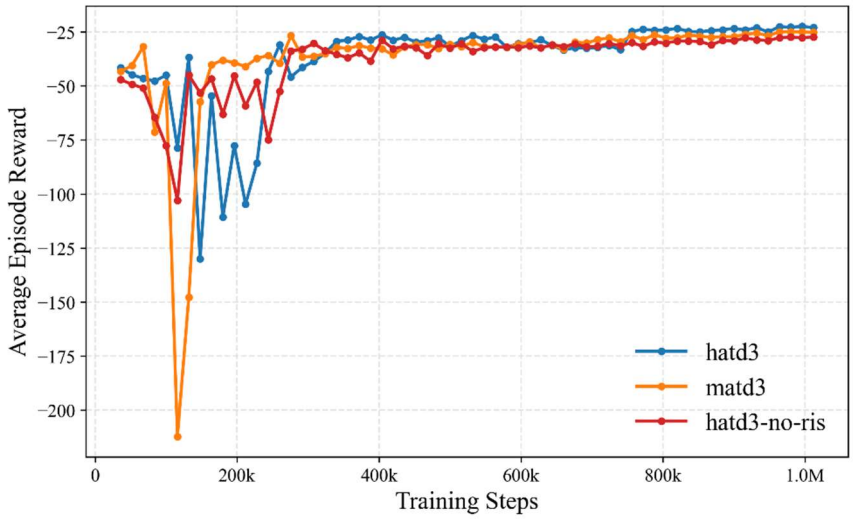

Figure 3. Average Cumulative Reward

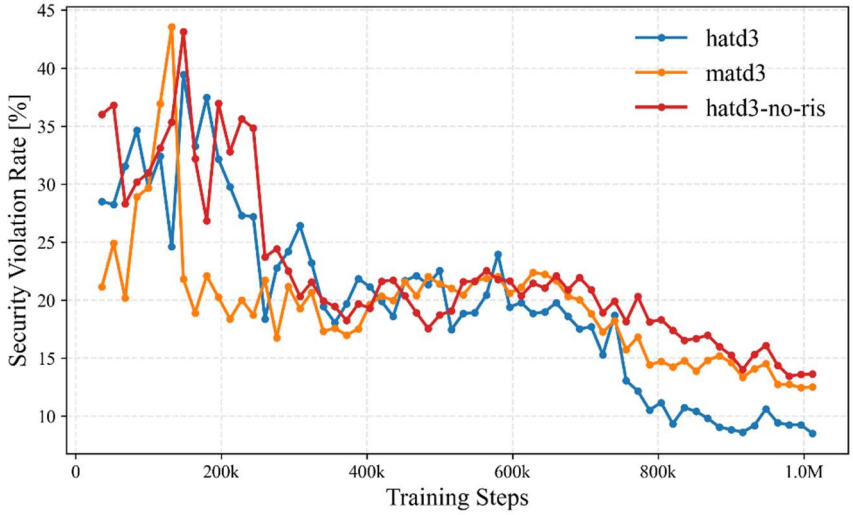

Figure 4. Security Failure Probability

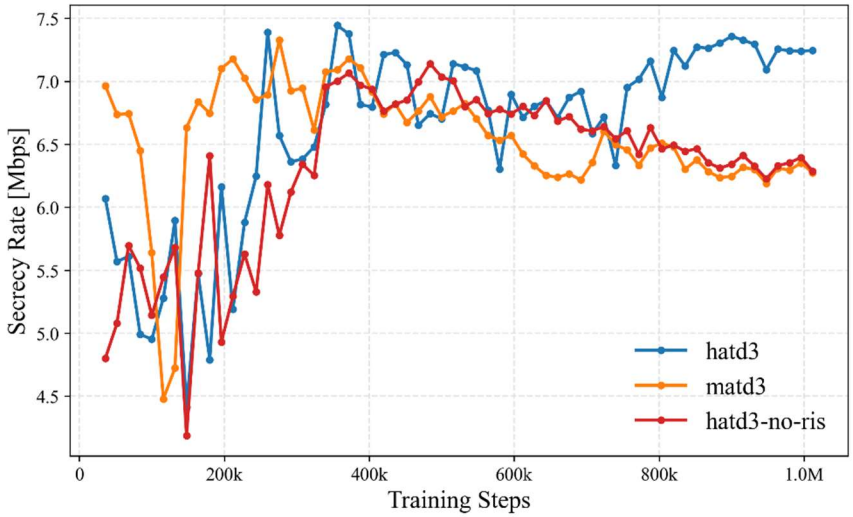

Figure 5. Average Achievable Secrecy Rate

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

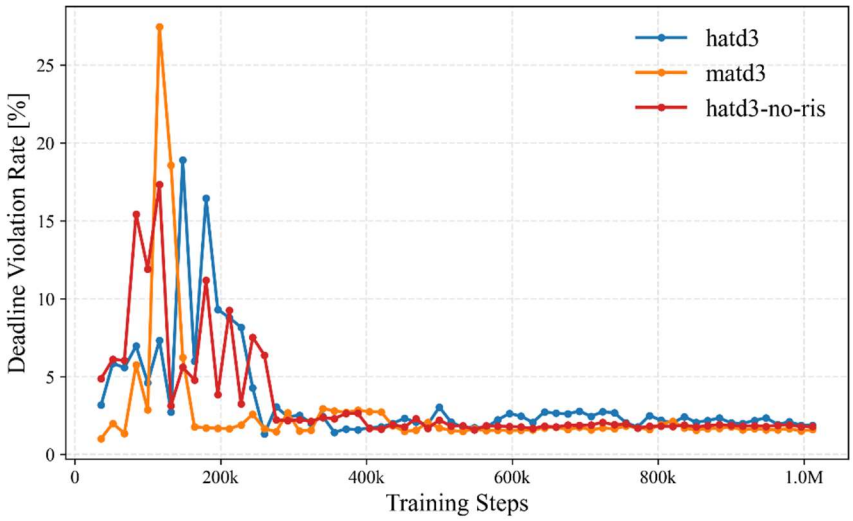

Figure 6. Deadline Violation Probability

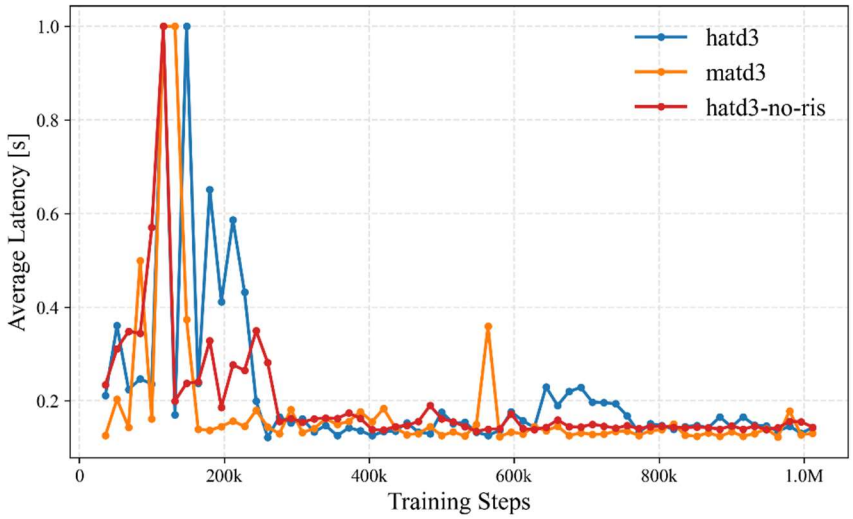

Figure 7. Average System Latency

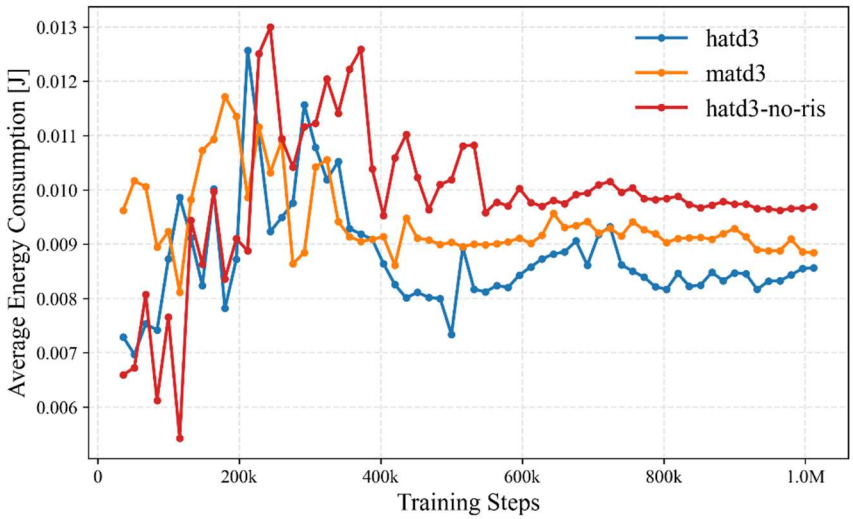

Figure 8. Average Energy Consumption

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0028

5. Conclusion

This paper has addressed the dual challenges of physical layer security threats and heterogeneous action space dimensionality in multi-server MEC systems. We proposed the CHESS framework, which models the RIS-assisted secure offloading problem as a detailed DEC-POMDP and solves it using the HATD3 algorithm with a preference-based action mapping mechanism. By decoupling continuous and discrete decision variables and leveraging HATD3's heterogeneous architecture, the framework enables effective collaboration among users, base stations, and the RIS controller. Simulation results demonstrate that CHESS significantly improves system security performance, reduces energy consumption, and achieves a superior balance among multiple performance objectives compared to baseline methods, providing a viable solution paradigm for complex heterogeneous edge networks.

Acknowledgments

This work was supported by the Northwest Normal University.

References

[1] H. Guo, J. Liu, and J. Zhang, "Computation offloading for multi-access mobile edge computing in ultra-dense networks," IEEE Communications Magazine, vol. 56, no. 8, pp. 14-19, 2018.

[2] L. Qin, H. Lu, Y. Chen, et al., "Toward decentralized task offloading and resource allocation in user-centric MEC," IEEE Transactions on Mobile Computing, vol. 23, no. 12, pp. 11807-11823, 2024.

[3] Y. Xiao, Y. Jia, C. Liu, et al., "Edge computing security: State of the art and challenges," Proceedings of the IEEE, vol. 107, no. 8, pp. 1608-1631, 2019.

[4] T. Zhang, H. Wen, J. Tang, et al., "Cooperative jamming secure scheme for IWNs random mobile users aided by edge computing intelligent node selection," IEEE Transactions on Industrial Informatics, vol. 17, no. 7, pp. 4999-5009, 2020.

[5] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," IEEE Wireless Communications Letters, vol. 8, no. 5, pp. 1410-1414, 2019.

[6] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," IEEE Transactions on Wireless Communications, vol. 18, no. 11, pp. 5394-5409, 2019.

[7] Y. Zhong, J. G. Kuba, X. Feng, et al., "Heterogeneous-agent reinforcement learning," Journal of Machine Learning Research, 2024.

[8] J. Xu, A. Xu, L. Chen, et al., "Deep reinforcement learning for RIS-aided secure mobile edge computing in industrial Internet of Things," IEEE Transactions on Industrial Informatics, vol. 20, no. 2, pp. 2455-2464, 2023.

[9] Qin L, Lu H, Chen Y, et al. Towards decentralized task offloading and resource allocation in user-centric mobile edge computing[J]. arXiv preprint arXiv:2312.01499, 2023.

[10] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," IEEE transactions on information theory, vol. 24, no. 4, pp. 451-456, 2003.

[11] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in Proc. ICML, 2018, pp. 1587-1596.

[12] J. Zhang, J. Du, Y. Shen, et al., "Dynamic Computation Offloading With Energy Harvesting Devices: A Hybrid-Decision-Based Deep Reinforcement Learning Approach," IEEE Internet Things J., vol. 7, no. 10, pp. 9303-9317, 2020.