Reinforcement Learning based Latency Minimization in Secure NOMA-MEC Systems with Hybrid SIC

Kaidi Wang, *Member, IEEE,* Haodong Li, *Student Member, IEEE,* Zhiguo Ding, *Fellow, IEEE,* and Pei Xiao, *Senior Member, IEEE*

***Abstract*—In this paper, physical layer security (PLS) in a non-orthogonal multiple access (NOMA)-based mobile edge computing (MEC) system is investigated, where hybrid successive interference cancellation (SIC) decoding is considered. Specifically, users intend to complete confidential tasks with the help of the MEC server, while an eavesdropper attempts to intercept the offloaded tasks. By jointly designing computational resource allocation, task assignment, and power allocation, a latency minimization problem is formulated. Based on the interactions between local computing time and MEC processing time, the closed-from solutions of computational resource allocation and task assignment are derived. After that, a strategy selection mechanism is established to select offloading strategies based on the corresponding conditions. Moreover, according to the analysis of hybrid SIC decoding, the conditions of different decoding orders in secure NOMA networks are derived. Furthermore, a reinforcement learning based algorithm is proposed to solve the power allocation problems for NOMA and OMA offloading strategies. This work is extended to a multi-user scenario, in which a matching-based algorithm is proposed to solve the formulated sub-channel assignment problem. Simulation results indicate that: i) the proposed solution can significantly reduce the latency and provide dynamic strategy selection for various scenarios; ii) the NOMA offloading strategy with hybrid SIC decoding can outperform other strategies in the considered system.**

***Index Terms*—Mobile edge computing (MEC), Non-orthogonal multiple access (NOMA), physical layer security (PLS), reinforcement learning, sub-channel assignment.**

I. Introduction

With the explosive development of mobile applications, the demand for computational capacity of mobile equipments has significantly increased in recent years [1]. In this context, mobile edge computing (MEC) has been emerged as a promising technology to provide the real-time computational service [2]. Specifically, by equipping the high-performance central processing units (CPUs) at the base station (BS), the computation-intensive and latency-critical tasks of the mobile users can be partially or fully offloaded to the MEC server for processing [3]. As a result, the sophisticated applications are executable at the smart device with the limited computational capacity and power [4]. Moreover, in order to achieve the high-speed and low-latency task offloading, the non-orthogonal multiple access (NOMA) transmission scheme is extensively employed in MEC systems [5], [6]. With NOMA schemes, multiple users are able to offload tasks to the MEC server at the same time and frequency band [7]. By adopting successive interference cancellation (SIC) techniques, part of co-channel interference can be decoded and removed at the BS based on the channel state information (CSI) or quality of service (QoS) [8], [9]. In previous research, the advantages of adopting NOMA schemes in MEC systems have been demonstrated, and compared with conventional orthogonal multiple access (OMA) schemes [10], [11].

The work of Z. Ding was supported by H2020-MSCA-RISE-2020 under grant number 101006411.

K. Wang and P. Xiao are with the Institute for Communication Systems (ICS), Home for 5GIC & 6GIC, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom (email: {kaidi.wang, p.xiao}@surrey.ac.uk).

H. Li and Z. Ding are with the Department of Electrical and Electronic Engineering, the University of Manchester, Manchester, M13 9PL, U.K (email: {haodong.li, zhiguo.ding}@manchester.ac.uk).

A. Related Works

In existing works on NOMA-MEC systems, user clustering and resource allocation have been extensively studied in order to minimize the latency, energy consumption, or their combination [12]–[18]. However, due to the broadcast nature of the multiple access networks, the offloaded tasks have a high risk of being intercepted by external eavesdroppers [19]. Therefore, the research on physical layer security (PLS) in NOMA-based MEC systems has emerged recently [20]–[26]. In [20], by introducing an energy weight for each user, a weighted energy consumption minimization problem was investigated in a two-user NOMA-MEC system. In order to further improve the aforementioned system, a secrecy outage probability minimization problem was considered. The energy consumption minimization problem was also studied in [21], where a group of wireless devices was considered as the jammer to help the eavesdropped edge computing device. For efficiently solving the formulated problem, a three-layered algorithm was proposed with the help of vertical decomposition. The authors in [22] focused on maximizing the minimum anti-eavesdropping ability of the proposed NOMA-MEC system. It was indicated that users tend to compute all tasks locally, and the offloading strategy is selected only if the energy for computing is insufficient. By jointly designing task assignment and power allocation, a latency minimization problem was studied in [23], where a bisection searching based algorithm was developed to solve the simplified problem. A novel multi-user MEC system with one unmanned aerial vehicle (UAV) server and one UAV jammer was proposed in [24], where both time division multiple access (TDMA) and NOMA transmission schemes were utilized. By maximizing the minimum secure computing capacity in those two schemes, the superiority of NOMA schemes compared to TDMA schemes was proved. By defining the energy efficiency as the ratio of the sum secure

rate to the sum power consumption, an energy efficiency maximization problem with imperfect and perfect CSI was studied in [25] and [26], respectively. Specifically, in order to mitigate the impact of the eavesdropper, in [25], a full-duplex MEC server was introduced in a multi-subcarrier scenario to generate artificial noise, and hence, sub-channel allocation and power control for the MEC server were taken into consideration. In [26], the formulated problem was divided into two sub-problems, and the closed-form solutions of local CPU-cycle frequency scheduling and the transmit power allocation were provided.

B. Motivation and Contribution

Although the optimization for NOMA-based MEC systems in PLS scenario has been extensively investigated, the in-depth research and analysis in this field is still limited. First, since the size of offloaded tasks is optimized, it may lead to different strategies to complete the given tasks. For example, some users may not be able, or need, to offload tasks to the MEC server. In this case, it is necessary to establish a dynamic strategy selection mechanism based on task assignment and power allocation. Second, in secure NOMA-MEC systems, the signals are decoded at both the BS and the eavesdropper, and hence, the SIC decoding order plays an important role in determining system performance. In existing research, the SIC decoding order is considered to be fixed [21], [24]–[26], or a strong assumption that eavesdroppers can cancel all interference is adopted [20], [22], [23]. For the sake of efficiency and practicality, at the BS and the eavesdropper, a more flexible SIC decoding order and an uncontrollable SIC decoding order should be employed, respectively. Third, even though the computational capacity of the MEC server is obviously larger than that of the user, it is still limited, and then the MEC computing time should be taken into consideration. For the multi-user scenario, computational resource allocation for processing offloaded tasks is required.

Against this background, this paper aims to find the possible task offloading strategies and derive the corresponding conditions. Moreover, it is considered that the SIC decoding order can be dynamically switched at the BS, and randomly selected at the eavesdropper. At the MEC server, computational resource allocation is included to further improve the proposed system. The main contributions of this paper are listed as follows:

- A novel secure NOMA-MEC system is proposed, where an eavesdropper intends to overhear the offloaded tasks. The users can partially offload tasks to the MEC server via the NOMA or OMA schemes, or completely compute all confidential tasks locally. At the BS and the eavesdropper, the hybrid and random SIC decoding orders are utilized, respectively. A latency minimization problem is formulated by jointly designing power allocation, task assignment, and computational resource allocation.
- The computational resource allocation and task assignment problems are solved by investigating the interactions between local computing time and MEC processing time, where the closed-form solutions are derived and expressed in terms of the secrecy rate. The offloading strategies and the corresponding conditions for switching among those strategies are derived to improve the performance of task processing. Furthermore, the conditions for achieving different SIC decoding orders are analyzed.
- The formulated power allocation problem is divided into three sub-problems according to the possible strategies, including NOMA offloading, OMA offloading, and local computing. With the help of reinforcement learning, a deep deterministic policy gradient (DDPG)-based algorithm is developed, which can optimize power allocation coefficients for the NOMA and OMA offloading strategies. The complexity of the proposed DDPG-based algorithm is also presented.
- In order to extend the work to a multi-user scenario, sub-channel assignment is studied, where users are paired and assigned to different sub-channels. The formulated sub-channel assignment problem is considered as a two-to-one matching, and then a matching-based algorithm is proposed, in which the DDPG-based power allocation algorithm can be iteratively performed. The property of the matching-based algorithm is also analyzed.

The effectiveness of the proposed solution is verified by simulations, where the appropriate strategies are selected based on various situations. Moreover, it is demonstrated that for both two-user and multi-user scenarios, hybrid SIC decoding can outperform fixed SIC decoding in the considered NOMA-MEC system. Furthermore, the correctness of the provided insights is verified.

II. System Model

Consider a MEC system with one BS, one eavesdropper and two users, where the MEC server is equipped at the BS, and all nodes are equipped with single-antennas. In order to complete the confidential tasks, users tend to partially offload tasks to the trusted MEC server by utilizing NOMA transmission schemes, and compute the rest of the tasks locally in the meantime. At the MEC server, the available computational resource, i.e., CPUs cycles, is dynamically allocated to process the offloaded tasks. The notations used in this paper are listed in Table I.

A. Signal Model

Based on the uplink NOMA scheme, users occupy the same time and frequency to offload tasks to the MEC server. At the BS and the eavesdropper, the received signals can be respectively expressed as

$$y = \sum_{i=1}^{2} \hat{h}_i \sqrt{p_i} s_i + n_0, \tag{1}$$

and

$$y_e = \sum_{i=1}^{2} \hat{h}_{i,e} \sqrt{p_i} s_i + n_0. \tag{2}$$

It is assumed that perfect channel state information (CSI) of all nodes is available at the BS, and the channel gain is constant during offloading.

TABLE I: Summary of Notations

Notation	Description
y	Received signal at the BS
y_e	Received signal at the eavesdropper
$\hat{h}_i$	Channel between user i and the BS
$\hat{h}_{i,e}$	Channel between user i and the eavesdropper
p_i	Power allocation coefficient of user i
s_i	Transmitted signal
n_0	Additive noise
$R_{i,\mathrm{e}}$	Data rate of user i at the eavesdropper
B	Bandwidth of each sub-channel
P_t	Available transmit power
σ^2	The variance of additive noise
$R_{i,\mathrm{s}}$	Secrecy rate of user i
β_i	Task assignment coefficient of user i
D_i	The size of user i's tasks in bits
T_i^{off}	Offloading time of user i
$T_{0,i}^{\mathrm{com}}$	MEC computing time for user i's tasks
μ	Required cycles to complete one bit of tasks
τ_i	Computational resource allocation coefficient of user i
C_0	Available CPU cycles at the MEC server
T_i^{com}	Local computing time at user i
C_i	Available CPU cycles at user i

In order to decode the received signals, the SIC technique is utilized at the eavesdropper and the BS [27]. At the eavesdropper, the eavesdropping process can be considered as an uplink NOMA system, whose sum rate is not affected by SIC decoding orders. Therefore, in this paper, an arbitrary decoding order is adopted at the eavesdropper, since it does not change the size of intercepted tasks. The denotation of users is based on the SIC decoding order at the eavesdropper. Without loss of generality, the user whose signal is decoded firstly is denoted by user 1, and the user whose signal is decoded afterwards is denoted by user 2. Therefore, at the eavesdropper, the data rate of user 1 and user 2 can be respectively expressed as follows:

$$R_{1,\mathrm{e}} = B\log_2\left(1+\frac{p_1|h_{1,\mathrm{e}}|^2}{p_2|h_{2,\mathrm{e}}|^2+1}\right), \tag{3}$$

and

$$R_{2,\mathrm{e}} = B\log_2\left(1+p_2|h_{2,\mathrm{e}}|^2\right), \tag{4}$$

where $|h_{i,\mathrm{e}}|^2 = P_t|\hat{h}_{i,\mathrm{e}}|^2\sigma^{-2}$ is the normalized channel gain. At the BS, the hybrid SIC decoding is employed, where the decoding orders can be further optimized in order to improve users' individual performance [28], [29]. Due to the fact that there are two users in the proposed system, two different decoding orders should be considered, as shown in follows.

1) SIC Decoding Order 1: If user 1's signal is decoded first at the BS, the data rate of user 1 is given by

$$R_1 = B\log_2\left(1+\frac{p_1|h_1|^2}{p_2|h_2|^2+1}\right), \tag{5}$$

where $|h_i|^2 = P_t|\hat{h}_i|^2\sigma^{-2}$ is the normalized channel gain. After removing the interference caused by user 1, the data rate of user 2 can be presented as follows:

$$R_2 = B\log_2(1+p_2|h_2|^2). \tag{6}$$

2) SIC Decoding Order 2: If user 2's signal is decoded before user 1's, the following data rate can be respectively obtained at the BS:

$$R_1 = B\log_2(1+p_1|h_1|^2), \tag{7}$$

and

$$R_2 = B\log_2\left(1+\frac{p_2|h_2|^2}{p_1|h_1|^2+1}\right). \tag{8}$$

Note that the individual secrecy rate of any user should be non-negative, otherwise this user would stop offloading tasks to the MEC server. Therefore, the secrecy rate of any user i can be expressed as

$$R_{i,\mathrm{s}} = \max\{0, R_i - R_{i,\mathrm{e}}\}. \tag{9}$$

If the secrecy rate of any user is zero, the local computing strategy is employed to process all tasks. In this case, the other user is allowed to offload tasks to the MEC server by utilizing OMA transmission schemes, where the data rate of the OMA user can be calculated by removing the interference caused by the local computing user.

B. MEC Model

Suppose that $\beta_i D_i$ bits of tasks are offloaded from user i to the MEC server. The required offloading time is given by

$$T_i^{\mathrm{off}} = \frac{\beta_i D_i}{R_{i,\mathrm{s}}}. \tag{10}$$

In this paper, the pure NOMA transmission scheme is employed, which implies that the offloading time of all users should be the same[1]. That is, if both users decide to offload tasks to the MEC server, the following condition should be satisfied:

$$T^{\mathrm{off}} = T_1^{\mathrm{off}} = T_2^{\mathrm{off}}. \tag{11}$$

This condition can be achieved by adjusting task assignment coefficients β_i and power allocation coefficients p_i. On the other hand, if any user decides to compute all tasks locally, the offloading time of this user is zero. Therefore, the offloading time of any user i can be presented as follows:

$$T_i^{\mathrm{off}} \in \{0, T^{\mathrm{off}}\}. \tag{12}$$

At the MEC server, the computational resource is allocated to process the offloaded tasks. For user i's offloaded tasks, the computing time is given by

$$T_{0,i}^{\mathrm{com}} = \frac{\mu\beta_i D_i}{\tau_i C_0}. \tag{13}$$

After offloading $\beta_i D_i$ bits of tasks to the MEC server, the remaining tasks, i.e., $(1-\beta_i)D_i$, are processed by the user. The local computing time of user i is given by

$$T_i^{\mathrm{com}} = \frac{\mu(1-\beta_i)D_i}{C_i}. \tag{14}$$

[1]It is revealed in [30] that with sufficient energy, the lower latency can be achieved by the pure NOMA scheme, compared to the hybrid NOMA scheme. In this paper, in order to minimize the latency, the pure NOMA scheme is performed, which can be implemented based on power allocation and task assignment.

III. Problem Formulation

In order to explore the security issue and improve the performance, latency minimization is investigated in the proposed NOMA-MEC system. Due to the fact that users can compute and offload tasks simultaneously, while the MEC server can only compute tasks after offloading, the latency of any user i can be expressed as

$$T_i = \max\{T_i^{\text{com}}, T_i^{\text{off}} + T_{0,i}^{\text{com}}\}. \tag{15}$$

By jointly designing power allocation, task assignment, and computational resource allocation, the latency minimization problem is formulated as follows:

$$\min_{\boldsymbol{p},\boldsymbol{\beta},\boldsymbol{\tau}} \quad \max_{i\in\{1,2\}} \{T_i^{\text{com}}, T_i^{\text{off}} + T_{0,i}^{\text{com}}\} \tag{16}$$

$$\text{s.t.} \quad 0 \le p_i \le 1, \forall i \in \{1,2\}, \tag{16a}$$

$$0 \le \beta_i \le 1, \forall i \in \{1,2\}, \tag{16b}$$

$$0 \le \tau_i, \forall i \in \{1,2\}, \tag{16c}$$

$$\tau_1 + \tau_2 \le 1, \tag{16d}$$

$$T_i^{\text{off}} \in \{0, T^{\text{off}}\}, \forall i \in \{1,2\}, \tag{16e}$$

where $\boldsymbol{p}$, $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$ are the collections of all power allocation coefficients, task assignment coefficients, and computational resource allocation coefficients, respectively. In the formulated problem, the objective function is the maximum time consumption of both users, including the local computing time T_i^{com} and the MEC processing time $T_i^{\text{off}} + T_{0,i}^{\text{com}}$. In constraints (16a) and (16b), the ranges of users' power allocation coefficients and task assignment coefficients are defined. Constraints (16c) and (16d) state the condition of the allocated computational resource at the MEC server for computing each user's offloaded tasks. Constraint (16e) indicates that the offloading time of any user is zero with the local computing strategy, or equals to another user's offloading time with the NOMA offloading strategy.

The formulated problem is non-convex and difficult to be transformed into a convex problem. However, this problem can be solved by analyzing the interactions between all terms in the objective function. Specifically, there is a trade-off between users' MEC computing time, i.e, $T_{0,1}^{\text{com}}$ and $T_{0,2}^{\text{com}}$, which is decided by the computational resource allocation coefficients. Moreover, the balance between each user's local computing time T_i^{com} and MEC processing time $T_i^{\text{off}} + T_{0,i}^{\text{com}}$ is determined by the task assignment coefficient. Based on different strategies, including NOMA offloading, OMA offloading and local computing, the formulated problem can be divided into three sub-problems and solved separately.

IV. Computational Resource Allocation and Task Assignment

In this section, by analyzing the aforementioned interactions, the closed-form expressions of computational resource allocation coefficients and task assignment coefficients are derived, and the formulated problem is transformed according to different strategies.

A. Optimal Computational Resource Allocation Coefficients

In the formulated latency minimization problem, the computational resource allocation coefficients only involve the MEC computing time. That is, if power allocation coefficients and task assignment coefficients are fixed, the computational resource allocation strategy can be obtained by balancing the MEC computing time of both users.

If both users offload tasks to the MEC server by utilizing NOMA schemes, i.e., $\beta_i > 0$ and $R_{i,\text{s}} > 0$, $\forall i \in \{1,2\}$, the users' offloading time is the same and equals to T^{off}, as shown in (11). In this case, the original objective function (16) can be transformed as

$$\max_{i\in\{1,2\}} \{T_i^{\text{com}}, T^{\text{off}} + T_{0,i}^{\text{com}}\}. \tag{17}$$

With the given secrecy rate and task assignment coefficients, the local computing time T_i^{com} and offloading time T^{off} can be regarded as constants and removed[2]. Based on (13), the following problem can be obtained:

$$\min_{\boldsymbol{\tau}} \quad \max_{i\in\{1,2\}} \left\{ \frac{\mu\beta_i D_i}{\tau_i C_0} \right\} \tag{18}$$

$$\text{s.t.} \quad \text{(16c), (16d).}$$

That is, by minimizing the MEC computing time, the optimal computational resource allocation coefficients can be obtained. It is indicated by (18) that any user's MEC computing time is monotonically decreasing with the increasing computational resource allocation coefficient. Therefore, in problem (18), all available computational resources at the MEC server should be utilized, i.e., $\tau_1 + \tau_2 = 1$. On the other hand, there is a trade-off between users' computing time at the MEC server. Specifically, the increase in any user's computational resource allocation coefficient will lead to the increase in another user's MEC computing time. According to [12], $T_{0,1}^{\text{com}} = T_{0,2}^{\text{com}}$ is satisfied by the optimal computational resource allocation coefficients. As a result, the following condition can be obtained:

$$\begin{cases} \tau_1^* + \tau_2^* = 1, \\ \dfrac{\mu\beta_1 D_1}{\tau_1^* C_0} = \dfrac{\mu\beta_2 D_2}{\tau_2^* C_0}. \end{cases} \tag{19}$$

Hence, the optimal computational resource allocation coefficients can be expressed as

$$\tau_i^* = \frac{\beta_i D_i}{\beta_1 D_1 + \beta_2 D_2}. \tag{20}$$

Note that the derived optimal computational resource coefficients always satisfy constraints (16c) and (16d) with any task assignment coefficients. Moreover, with the OMA offloading strategy in which any user does not offload tasks to the MEC server, i.e., $\beta_i = 0, \exists i \in \{1,2\}$, the derived optimal solution in (20) still holds. In this case, all available computational resources at the MEC server is allocated to compute the offloaded tasks of the OMA user. Furthermore, with the local computing strategy in which both users complete all tasks

[2]Note that if the local computing time of any user, i.e., T_i^{com}, is significantly greater than other terms, the computational resource allocation strategy will not affect the latency of the proposed system. However, in this case, a computational resource allocation strategy is still required.

locally, there is no offloaded tasks at the MEC server, and hence, the computational resource allocation coefficients of both users are zero. Therefore, the optimal computational resource allocation coefficients with all strategies can be presented as follows:

$$\tau_i^* = \begin{cases} \dfrac{\beta_i D_i}{\beta_1 D_1 + \beta_2 D_2}, & \text{if } \beta_i > 0, \exists i \in \{1,2\}, \\ 0, & \text{if } \beta_i = 0, \forall i \in \{1,2\}. \end{cases} \tag{21}$$

B. Optimal Task Assignment Coefficients

In this subsection, with the derived optimal computational resource allocation coefficients, the optimal task assignment coefficients can be obtained according to the fixed secrecy rate. At this stage, the terms in the objective function (16) can be expressed as functions of task assignment coefficients, as shown below:

$$\max_{i\in\{1,2\}} \left\{ f(\beta_i), g(\beta_i) \right\}, \tag{22}$$

where

$$f(\beta_i) = \frac{\mu(1-\beta_i)D_i}{C_i}, \tag{23}$$

and

$$g(\beta_i) = \left(\frac{1}{R_{i,\mathrm{s}}} + \frac{\mu}{\tau_i C_0} \right) \beta_i D_i, \tag{24}$$

are local computing time and MEC processing time (including the offloading time and MEC computing time), respectively. For any user i, there is a trade-off between $f(\beta_i)$ and $g(\beta_i)$, which is determined by task assignment coefficient β_i. For example, if β_i grows, local computing time $f(\beta_i)$ is decreased, while MEC processing time $g(\beta_i)$ is increased. Therefore, in order to minimize the latency, the following condition should be satisfied:

$$f(\beta_i^*) = g(\beta_i^*), \tag{25}$$

if user i can offload tasks to the MEC server, i.e., $R_{i,\mathrm{s}} > 0$. Moreover, if the NOMA offloading strategy is adopted, i.e., $R_{i,\mathrm{s}} > 0, \forall i \in \{1,2\}$, the following condition can be obtained from (11):

$$\beta_1^* D_1 R_{2,\mathrm{s}} = \beta_2^* D_2 R_{1,\mathrm{s}}. \tag{26}$$

As a result, with the fixed secrecy rate and the optimal task assignment coefficients, the optimal task assignment coefficients can be obtained[3].

Proposition 1. *In the proposed NOMA-MEC system, if $R_{i,s} > 0, \exists i \in \{1,2\}$, the optimal task assignment coefficients of problem (16) can be expressed as*

$$\begin{cases} \beta_1^* = \dfrac{\mu R_{1,s} C_0}{C_0 C_1 + \mu R_{1,s} C_1 + \mu R_{1,s} C_0 + \mu R_{2,s} C_1}, \\ \beta_2^* = \dfrac{\mu R_{2,s} C_0}{C_0 C_2 + \mu R_{2,s} C_2 + \mu R_{2,s} C_0 + \mu R_{1,s} C_2}. \end{cases} \tag{27}$$

Proof: Refer to Appendix A. ∎

Due to the fact that the above proposition is derived based on (26), constrain (16e) always holds by the optimal task assignment coefficients. That is, with any given secrecy rate, the task assignment coefficients can be dynamically adjusted to satisfy constrain (16e). Moreover, it is worth to mention that in the local computing strategy, all tasks are computed by users, and hence, the optimal task assignment coefficients can be expressed as

$$\beta_1^* = \beta_2^* = 0, \quad \text{if } R_{i,\mathrm{s}} = 0, \forall i \in \{1,2\}. \tag{28}$$

C. Strategy Selection

At this stage, the closed-form solutions of task assignment and computational resource allocation coefficients are derived, and expressed by the secrecy rate. Based on whether the secrecy rate is zero, three power allocation problems can be obtained with different offloading strategies.

1) NOMA Offloading Strategy: In this case, both users offload tasks to the MEC server, i.e., $R_{i,\mathrm{s}} > 0, \forall i \in \{1,2\}$. By substituting the derived task assignment coefficients (27) into (22), the power allocation problem can be presented as

$$\begin{aligned} \min_{\boldsymbol{p}} \quad & \max\{f_1, f_2, g_1, g_2\} \\ \text{s.t.} \quad & \text{(16a)}, \end{aligned} \tag{29}$$

where

$$f_1 = \frac{\mu D_1 (C_0 + \mu R_{1,\mathrm{s}} + \mu R_{2,\mathrm{s}})}{C_0 C_1 + \mu R_{1,\mathrm{s}} C_1 + \mu R_{1,\mathrm{s}} C_0 + \mu R_{2,\mathrm{s}} C_1}, \tag{30}$$

$$f_2 = \frac{\mu D_2 (C_0 + \mu R_{2,\mathrm{s}} + \mu R_{1,\mathrm{s}})}{C_0 C_2 + \mu R_{2,\mathrm{s}} C_2 + \mu R_{2,\mathrm{s}} C_0 + \mu R_{1,\mathrm{s}} C_2}, \tag{31}$$

$$\begin{aligned} g_1 = & \frac{\mu D_1 (C_0 + \mu R_{1,\mathrm{s}})}{C_0 C_1 + \mu R_{1,\mathrm{s}} C_1 + \mu R_{1,\mathrm{s}} C_0 + \mu R_{2,\mathrm{s}} C_1} \\ & + \frac{\mu^2 D_2 R_{2,\mathrm{s}}}{C_0 C_2 + \mu R_{2,\mathrm{s}} C_2 + \mu R_{2,\mathrm{s}} C_0 + \mu R_{1,\mathrm{s}} C_2}, \end{aligned} \tag{32}$$

and

$$\begin{aligned} g_2 = & \frac{\mu D_2 (C_0 + \mu R_{2,\mathrm{s}})}{C_0 C_2 + \mu R_{2,\mathrm{s}} C_2 + \mu R_{2,\mathrm{s}} C_0 + \mu R_{1,\mathrm{s}} C_2} \\ & + \frac{\mu^2 D_1 R_{1,\mathrm{s}}}{C_0 C_1 + \mu R_{1,\mathrm{s}} C_1 + \mu R_{1,\mathrm{s}} C_0 + \mu R_{2,\mathrm{s}} C_1}. \end{aligned} \tag{33}$$

In the NOMA offloading strategy, the offloading time of both users is the same. By including the derived task assignment coefficients (27) into (11), the condition of the optimal secrecy rate can be obtained.

Remark 1. *In the proposed NOMA-MEC system, if $R_{i,s} > 0, \forall i \in \{1,2\}$, the optimal secrecy rate satisfies the following condition:*

$$A_1 R_{2,s}^* + D_1 C_0 C_2 = A_2 R_{1,s}^* + D_2 C_0 C_1, \tag{34}$$

where

$$\begin{cases} A_1 \triangleq \mu D_1 C_2 + \mu D_1 C_0 - \mu D_2 C_1, \\ A_2 \triangleq \mu D_2 C_1 + \mu D_2 C_0 - \mu D_1 C_2. \end{cases} \tag{35}$$

Note that (34) is the condition of the theoretical optimal secrecy rate. If this equation is satisfied, the following condition can be achieved:

$$f_1 = f_2 = g_1 = g_2, \tag{36}$$

[3]In the OMA offloading strategy, although the task assignment of the OMA offloading user may not affect the latency, the derived optimal task assignment coefficient can still be utilized.

and the global optimal solution is obtained, which can achieve the minimum latency of the proposed NOMA-MEC system. However, due to the existence of interference and different SIC decoding orders, (34) is difficult to be satisfied.

2) OMA Offloading Strategy: In this case, one of the users tends to compute all tasks locally since the positive secrecy rate cannot be achieved. However, this condition is considered at the offloading phase. As shown in footnotes 2 and 3, there exists another condition for adopting the OMA offloading strategy, which is considered at the computing phase.

Proposition 2. *In the proposed NOMA-MEC system, the OMA offloading strategy will be selected if the following condition is satisfied:*

$$\frac{D_{i'}}{C_{i'}} \leq \frac{D_i}{C_0 + C_i}, \tag{37}$$

where users i and i' are the OMA offloading user and local computing user, respectively.

Proof: Refer to Appendix B. ∎

The above proposition is due to the fact that the computing time to process all tasks at user i' is less than the computing time to simultaneously process user i's tasks at user i and the MEC server, even though all available computational resources at the MEC server are allocated for user i's tasks. This situation will occur if the size of user i's tasks is significantly larger than that of user i''s, and/or the CPUs equipped at user i' are more powerful compared to user i and the MEC server. In this case, the OMA offloading strategy is selected and the latency is determined by the OMA offloading user i. Based on these conditions, the OMA offloading strategy is adopted, and the following problem is formulated:

$$\begin{aligned} \min_{\boldsymbol{p}} \quad & \max\{f_i, f_{i'}, g_i\} \\ \text{s.t.} \quad & \text{(16a)}, \end{aligned} \tag{38}$$

where

$$f_i = \frac{\mu D_i (C_0 + \mu R_{i,\mathrm{s}})}{C_0 C_i + \mu R_{i,\mathrm{s}} C_i + \mu R_{i,\mathrm{s}} C_0}, \tag{39}$$

$$f_{i'} = \frac{\mu D_{i'}}{C_{i'}}, \tag{40}$$

and

$$g_i = \frac{\mu D_i (C_0 + \mu R_{i,\mathrm{s}})}{C_0 C_i + \mu R_{i,\mathrm{s}} C_i + \mu R_{i,\mathrm{s}} C_0}. \tag{41}$$

In the OMA offloading strategy, the MEC processing time of the local computing user i' is zero, i.e., $g_{i'} = 0$, and hence, it is removed from the objective function. Moreover, constraint (16e) is removed since it is always satisfied in this case.

3) Local Computing Strategy: In this case, the positive secrecy rate cannot be achieved by both users, i.e., $R_{i,\mathrm{s}} = 0, \forall i \in \{1, 2\}$, and hence, the power allocation coefficients can be presented as follows:

$$p_1^* = p_2^* = 0. \tag{42}$$

Therefore, the latency of the proposed NOMA-MEC system is decided by the local computing time of both users, as shown below:

$$T = \max\left\{\frac{\mu D_1}{C_1}, \frac{\mu D_2}{C_2}\right\}. \tag{43}$$

D. Analysis of SIC Decoding Order

In this subsection, the NOMA offloading strategy is analyzed, where the different SIC decoding orders are investigated. From (10), it is indicated that the increasing secrecy rate can reduce the offloading time, or allows the users offload more tasks to the more efficient MEC server with the same offloading time. Therefore, in NOMA offloading strategy, under the condition in Remark 1, both users tend to maximize the secrecy rate. However, due to the interaction of users' power allocation coefficients, the optimal power allocation coefficients cannot be explicitly derived. In this in context, this subsection only compares SIC decoding orders.

1) SIC Decoding Order 1: If user 1's signal is decoded before user 2's, the secrecy rate can be presented as

$$R_{1,\mathrm{s}} = B\log_2\left[\frac{(p_1|h_1|^2 + p_2|h_2|^2 + 1)(p_2|h_{2,\mathrm{e}}|^2 + 1)}{(p_2|h_2|^2 + 1)(p_1|h_{1,\mathrm{e}}|^2 + p_2|h_{2,\mathrm{e}}|^2 + 1)}\right], \tag{44}$$

and

$$R_{2,\mathrm{s}} = B\log_2\left(\frac{p_2|h_2|^2 + 1}{p_2|h_{2,\mathrm{e}}|^2 + 1}\right). \tag{45}$$

With this decoding order, the condition for performing NOMA transmission schemes can be obtained as follows.

Proposition 3. *In the proposed NOMA-MEC system, the NOMA transmission scheme can be utilized with SIC Decoding Order 1 if the following condition is satisfied:*

$$\begin{cases} |h_1|^2 > |h_{1,e}|^2, \\ |h_2|^2 > |h_{2,e}|^2. \end{cases} \tag{46}$$

Proof: Refer to Appendix C. ∎

It is described by (96) that the feasible region of user 2's power allocation coefficient is replaced by

$$p_2 \in \left(0, \min\left\{\frac{|h_{1,e}|^2 - |h_1|^2}{|h_1|^2|h_{2,e}|^2 - |h_2|^2|h_{1,e}|^2}, 1\right\}\right), \tag{47}$$

if $|h_1|^2 > |h_{1,e}|^2$ and $|h_1|^2|h_{2,e}|^2 < |h_2|^2|h_{1,e}|^2$ hold. By eliminating normalization, the terms in (47) can be expressed as

$$\frac{|\hat{h}_{1,e}|^2 - |\hat{h}_1|^2}{P_t\sigma^{-2}(|\hat{h}_1|^2|\hat{h}_{2,e}|^2 - |\hat{h}_2|^2|\hat{h}_{1,e}|^2)}. \tag{48}$$

It can be found that the above term is monotonically decreasing with the increasing transmit power, and hence, the feasible region of user 2's power allocation coefficient p_2, i.e., (47), is shrinking. The following insight can be obtained:

Remark 2. *In the proposed NOMA-MEC system, by adopting SIC Decoding Order 1, the upper bound of user 2's power allocation coefficient decreases with the transmit power if $|h_1|^2|h_{2,e}|^2 < |h_2|^2|h_{1,e}|^2$ holds.*

2) SIC Decoding Order 2: If the SIC decoding order at the BS is swapped, the signal of user 2 is decoded firstly, and the expressions of secrecy rate can be shown as

$$R_{1,\mathrm{s}} = B\log_2\left[\frac{(p_1|h_1|^2 + 1)(p_2|h_{2,\mathrm{e}}|^2 + 1)}{p_1|h_{1,\mathrm{e}}|^2 + p_2|h_{2,\mathrm{e}}|^2 + 1}\right], \tag{49}$$

and

$$R_{2,\mathrm{s}} = B\log_2\left[\frac{(p_2|h_2|^2 + p_1|h_1|^2 + 1)}{(p_1|h_1|^2 + 1)(p_2|h_{2,\mathrm{e}}|^2 + 1)}\right]. \tag{50}$$

In this case, the condition for adopting SIC Decoding Order 2 can be presented as follows.

Proposition 4. *In the proposed NOMA-MEC system, NOMA transmission schemes with SIC Decoding Order 2 can be employed if the following condition is satisfied:*

$$\begin{cases} \dfrac{|h_1|^2}{|h_{1,e}|^2} > \dfrac{1}{|h_{2,e}|^2+1}, \\ |h_2|^2 > |h_{2,e}|^2. \end{cases} \tag{51}$$

Proof: Refer to Appendix D. ∎

From (104), the feasible region of user 1's power allocation coefficient is shown as follows:

$$p_1 \in \left(0, \min\left\{\frac{|h_2|^2-|h_{2,e}|^2}{|h_1|^2|h_{2,e}|^2}, 1\right\}\right). \tag{52}$$

For the unnormalized channel condition, the term in the above function is given by

$$\frac{|\hat{h}_2|^2-|\hat{h}_{2,e}|^2}{P_t\sigma^{-2}|\hat{h}_1|^2|\hat{h}_{2,e}|^2}. \tag{53}$$

It is indicated that with the increasing transmit power, the feasible region of p_1 tends be small. The following conclusion can be drawn.

Remark 3. *In the proposed NOMA-MEC system, by adopting SIC Decoding Order 2, the upper bound of user* 1*'s power allocation coefficient decreases with the transmit power.*

V. DDPG-based Power Allocation

In this section, power allocation with both NOMA and OMA offloading strategies is investigated. As aforementioned, the formulated power allocation problems (29) and (38) are non-convex and difficult to be solved. To tackle this issue, the DDPG scheme is adopted to minimize the latency, where the power allocation solution is obtained based on the DDPG decision-making strategy. The DDPG framework and the training algorithm are discussed in the sequel.

A reinforcement learning problem can be described by a 4-tuple (s_t, a_t, r_t, s_{t+1}) at any time step t, where s_t denotes the state, a_t denotes the action, r_t denotes the immediate reward of action a_t in state s_t, and s_{t+1} denotes the state at the next step [31]. Specifically, in the proposed DDPG-based algorithm, the following elements are defined:

- **State Space** $\mathcal{S}$: The state space $\mathcal{S}$ is the collection of all states, i.e., $s_t \in \mathcal{S}, \forall t$, where any state at step t can be expressed as

$$s_t = \{h_1(t), h_2(t), h_{1,\mathrm{e}}(t), h_{2,\mathrm{e}}(t)\}. \tag{54}$$

  Note that all channel conditions are available at the BS, and the channel gains are different among time steps.
- **Action Space** $\mathcal{A}$: The action space $\mathcal{A}$ is the collection of all actions, where any action at step t contains the corresponding power allocation coefficients, i.e.,

$$a_t = \{p_1(t), p_2(t)\}. \tag{55}$$

  It is worth mentioning that the designed DDPG-based algorithm is capable of outputting continuous actions, and hence, the action space $\mathcal{A}$ is also continuous.
- **Reward Function**: After choosing action a_t for any given state s_t, the immediate reward at step t is defined as follows:

$$r_t(s_t, a_t) = -1 \times \max\{f_1(t), f_2(t), g_1(t), g_2(t)\}. \tag{56}$$

  The term -1 is included in the reward function to minimize the objective function of (29). That is, if the DDPG network takes an action that increases the reward, the latency will be decreased. Moreover, the reward of the previous round is obtained at the beginning of any step.

The DDPG network is designed to find an optimal strategy which can maximize the discounted long term reward R_t, defined as

$$R_t = \sum_{t=0}^{T-1} \gamma^t r_t, \tag{57}$$

where $\gamma \in [0,1]$ is the discount factor which determines the balance of current and future rewards. If γ is small, the network will focus on maximizing the current reward. When γ increases, the network tends to choose the action which can maximize the future reward. In this paper, the actions between steps are independent because the power allocation constraint is only set for each step. Therefore, a relatively small γ is chosen in order to focus on the current reward.

As shown in Fig. 1, an experience memory is included in the designed DDPG network in order to avoid inefficient learning caused by the highly correlated input data. Moreover, the DDPG network adopts the actor-critic architecture, which contains an actor network and a critic network. Meanwhile, the DDPG network includes an additional neural network for both the actor network and the critic network, namely target networks.

For actor-critic schemes, the actor network takes the state s_t as the input, and then outputs the instant action a_t to the MEC-NOMA network based on the weight ω and a stochastic noise, i.e.,

$$a_t = \pi(s_t;\omega) + \mathcal{N}_0, \tag{58}$$

where $\mathcal{N}_0$ is the exploration noise, which can balance the exploration of new actions and the exploitation of previous actions. As a result, the case that the neural network stuck on local optimal decisions can be prevented. It is worth pointing out that each term in a_t is limited to $[0,1]$, and hence, the value of the noise-added action will be clipped if the result is beyond the desired range. Moreover, the decision of the actor network, i.e., $\pi(s_t;\omega)$, is also outputted to the critic network for evaluation. The critic network receives the decision, and then outputs the estimated Q-value $Q(s_t, \pi(s_t;\omega);\theta)$ with the weight θ, where the Q-value describes the expected long term reward. Based on the principle of policy gradient theorem [32], the objective of the actor network is to maximize the long term discounted reward $J(\mu) = Q(s, \pi(s;\omega);\theta)$, and hence, the actor network updates the weight ω according to the gradient:

$$\nabla_\omega J(\omega) = \nabla_\omega \pi(s;\omega) \nabla_a Q(s,a;\theta). \tag{59}$$

In terms of the target network, it is essentially a duplication of the original network with a slower update frequency. As a

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3194685

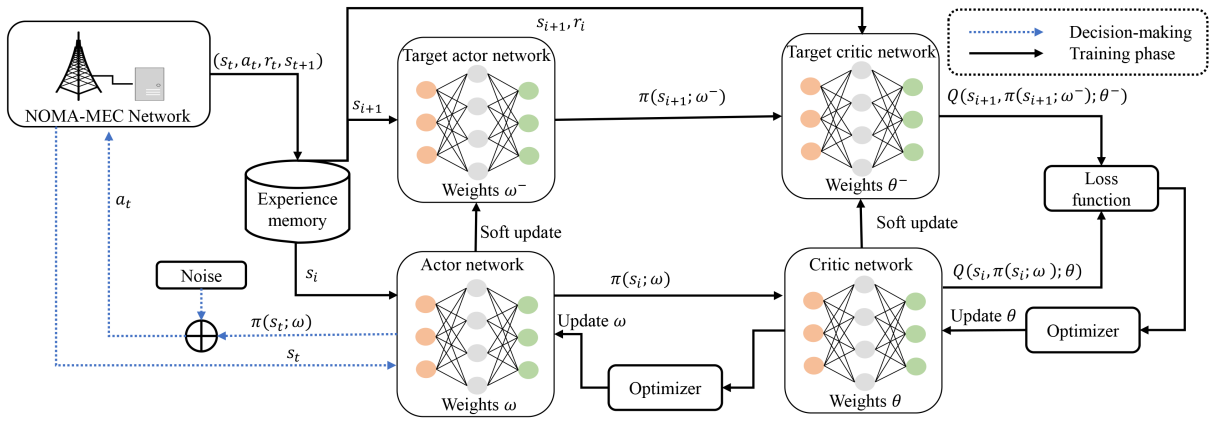


Fig. 1: An illustration of the DDPG framework.

result, more stable labels can be provided for the actor and critic networks. In the target networks, the soft update policy is adopted with the following weights:

$$\omega^- \leftarrow \rho\omega + (1-\rho)\omega^-, \tag{60}$$

$$\theta^- \leftarrow \rho\theta + (1-\rho)\theta^-, \tag{61}$$

where $0 < \rho \ll 1$ is the updating parameter. Moreover, in the target critic network, the target Q-value is estimated based on the experience tuple (s_i, a_i, r_i, s_{i+1}), as shown below:

$$y_i = r_i + \gamma Q(s_{i+1}, \pi(s_{i+1};\omega^-);\theta^-), \tag{62}$$

where i is the random sample index. Hence, the loss function of the critic network can be written as follows:

$$L(\theta) = \left(y_i - Q(s_i, a_i;\theta)\right)^2. \tag{63}$$

Here, the critic network is trained by minimizing the loss function. According to the above settings, a DDPG-based power allocation algorithm is presented in Algorithm 1.

In terms of the DDPG framework, each neural network contains one hidden layer with 100 neurons. Moreover, both the critic network and the target critic network adopt the Rectified Linear Unit (ReLU) as the activation function, while the activation function for the actor network is Sigmoid, which output the action value $a_t \in [0, 1]$ [33]. Based on the proposed DDPG-based power allocation algorithm, the formulated problems in (29) and (38) can be solved, where the strategies can be dynamically switched according to different conditions. In order to implement the hybrid SIC decoding order, the performance of different decoding orders is compared in the NOMA offloading strategy, and the best decoding order which can achieve the minimum latency is selected.

According to [34], [35], the computational complexity of the proposed DDPG-based power allocation algorithm can be expressed as $\mathcal{O}(4N_{\text{ep}}N_{\text{ts}}N_{\text{mb}}\sum_{i=1}^{L} N_i N_{i+1})$, where N_{mb} is the size of the mini-batch, L is the number of layers, and N_i is the number of neurons in the i-th layer. Specifically, for all 4 neural networks, the computational complexity of each step is $4N_{\text{mb}}\sum_{i=1}^{L} N_i N_{i+1}$. Due to the fact that there are N_{ep} episodes and each episode includes N_{ts} steps, the total computational complexity is given by $4N_{\text{ep}}N_{\text{ts}}N_{\text{mb}}\sum_{i=1}^{L} N_i N_{i+1}$.

Algorithm 1 DDPG-based Power Allocation Algorithm

Parameter initialization:
Initialize actor network $\pi(s_i;\omega)$, critic network $Q(s, a;\theta)$, target actor network $\pi(s;\omega^-)$, and target critic network $Q(s, a;\theta^-)$.
Initialize reply memory $\mathcal{R}$ with size $|\mathcal{R}|$, and memory counter.
Initialize discount factor γ, batch size, soft update coefficient ρ.
Training Phase:
for episode $= 1, 2, ..., N_{\text{ep}}$ **do**
  for step $= 1, 2, ..., N_{\text{ts}}$ **do**
    Input state s_t into actor network and obtain action $a_t = \pi(s_t;\omega) + \mathcal{N}$.
    Observe $r_t \leftarrow -1 \times \max\{f_1(t), f_2(t), g_1(t), g_2(t)\}$, and the next state s_{t+1}.
    Store experience tuple (s_t, a_t, r_t, s_{t+1}) into the memory $\mathcal{R}$.
    if memory counter $> |\mathcal{R}|$ **then**
      Remove previous experiences from the beginning.
    end if
    Randomly sample a mini-batch of experience tuple (s_t, a_t, r_t, s_{t+1}) with batch size and input DNNs.
    Update the weights of actor and critic networks based on (59) and (63).
    Update target network weights ω^- and θ^- according to (60) and (61).
  end for
end for

VI. Matching-based Sub-Channel Assignment

In order to extend the work to a multi-user scenario, a sub-channel assignment problem is considered in this section. Specifically, N users are paired and assigned to K sub-channels, where $N = 2K$. The collections of users and sub-channels are denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$ and $\mathcal{K} = \{1, 2, \ldots, K\}$, respectively. For the multi-user situation, the subscripts of variables are changed accordingly.

A. Sub-Channel Assignment Problem Formulation

By considering the maximum value of all users' time consumptions as the latency of the system, the formulated sub-channel assignment problem is given by

$$\min_{\boldsymbol{X}} \quad \max_{n\in\mathcal{N}} \left\{ \sum_{k=1}^{K} x_{k,n} T_{k,n}^{\text{com}}, \sum_{k+1}^{K} x_{k,n} (T_{k,n}^{\text{off}} + T_{0,k,n}^{\text{com}}) \right\} \tag{64}$$

$$\text{s.t.} \quad x_{k,n} = \{0,1\}, \forall k \in \mathcal{K}, \forall n \in \mathcal{N}, \tag{64a}$$

$$\sum\nolimits_{n=1}^{N} x_{k,n} = 2, \forall k \in \mathcal{K}, \tag{64b}$$

$$\sum\nolimits_{k=1}^{K} x_{k,n} = 1, \forall n \in \mathcal{N}, \tag{64c}$$

where $\boldsymbol{X}$ is a matrix of all sub-channel assignment indicators $x_{k,n}$. Constraint (64a) indicates two possible situation of each user, where $x_{k,n} = 1$ means user n is assigned to sub-channel k; otherwise $x_{k,n} = 0$. Constraints (64b) and (64c) show that each sub-channel is occupied by two users, and each user is assigned to one sub-channel, respectively.

B. Matching-based Sub-Channel Assignment Algorithm

Due to the fact that there exists a binary constraint, matching is utilized to solve the formulated sub-channel assignment problem. By treating users and sub-channels as two disjoint sets of players, a two-to-one matching can be defined as follows:

Definition 1. *Given two disjoint sets $\mathcal{N}$ and $\mathcal{K}$, a two-to-one matching ψ denotes the mapping from $\mathcal{N}$ to $\mathcal{K}$, which satisfies*

- $\psi(n) \in \mathcal{K}, \forall n \in \mathcal{N}$, $\psi(k) \in \mathcal{N}, \forall k \in \mathcal{K}$;
- $|\psi(n)| = 1, \forall n \in \mathcal{N}$, $|\psi(k)| = 2, \forall k \in \mathcal{K}$;
- $n \in \psi(k) \Leftrightarrow \psi(n) = k$.

It is assumed that all players are selfish. Therefore, the case that any user n is willing to be assigned to sub-channel k' rather than k only depends on its utility, as follows:

$$(k, \psi) \succ_n (k', \psi') \Leftrightarrow U_n(\psi) < U_n(\psi'), \tag{65}$$

where user n's utility with sub-channel k is given by

$$U_n(\psi) = \max\{T_{k,i}^{\text{com}}, T_{k,i}^{\text{off}} + T_{0,k,i}^{\text{com}} | \forall i \in \psi(k)\}. \tag{66}$$

Moreover, the utility of any sub-channel k, i.e., $U_k(\psi)$, is also decided by the above function.

In the proposed matching-based algorithm, if any user intends to use any sub-channel, it needs to exchange with one of the users assigned to that sub-channel. The swap matching $\psi_n^m = \{\psi \backslash \{(k,m),(k',n)\} \cup \{(k,n),(k',m)\}\}$, $m \in \psi(k)$, $n \in \psi(k')$, $m \in \psi_n^m(k')$, and $n \in \psi_n^m(k)$ denotes the sub-channels of users m and n are exchanged, and the matching is transformed from ψ to ψ_n^m. This exchange operation indicates that (m,n) is a swap-blocking pair, which is defined as follows:

Definition 2. *A swap-blocking pair (m,n) can be confirmed if and only if the following conditions are satisfied*

1) $\forall i \in \{m, n, \psi(m), \psi(n)\}$, $U_i(\psi_n^m) \leq U_i(\psi)$;
2) $\exists i \in \{m, n, \psi(m), \psi(n)\}$, $U_i(\psi_n^m) < U_i(\psi)$.

Algorithm 2 Sub-Channel Assignment Algorithm

Step 1: Initialization phase

1) Randomly match all users and sub-channels.
2) Record current matching as ψ_{init}.

Step 2: Swap matching phase

Any user n searches another user m.
if $\psi(m) \neq \psi(n)$ holds, and (m,n) is a swap-blocking pair
1) User n and user m exchange sub-channels.
2) Set $\psi = \psi_n^m$.
end if
Step 2 is repeated until no swap-blocking pair can be obtained in a complete cycle.

Based on the definition of swap-blocking pairs, a matching-based sub-channel assignment algorithm is proposed in Algorithm 2. In the algorithm, the DDPG-based power allocation algorithm is implemented at each iteration in order to calculate the utility and find the swap-blocking pair. That is, in the multi-user case, Algorithm 1 and Algorithm 2 are iteratively performed to minimize the latency.

C. Property Analysis

1) Complexity: The computational complexity of the proposed matching-based algorithm is $\mathcal{O}(CN^2)$. Specifically, in the worse case, each user needs to exchange with all other users, except the one with the same sub-channel. Hence, for N users, $N(N-2)$ calculations are performed. Given a cycle number C, the complexity of the worse case can be expressed by $CN(N-2)$.

2) Convergence: It is indicated by the definition of swap-blocking pairs that at least one pair can achieve less utility (latency) during the algorithm, and none of the pairs can increase the utility. With a finite number of users and sub-channels, the number of swap-blocking pairs is limited. As a result, the proposed sub-channel assignment algorithm is guaranteed to converge to a stable matching.

3) Stability: The stability of the proposed algorithm follows the definition of two-side exchange-stable (2ES) matching, as shown below:

Definition 3. *A matching is 2ES if and only if there are no swap-blocking pairs.*

In Algorithm 2, the condition for finalization is that no swap-blocking pair can be obtained in a complete cycle. Therefore, the final matching obtained from the proposed sub-channel assignment algorithm is always 2ES.

VII. Simulation Results

In this section, the simulation results are presented to demonstrate the effectiveness of the proposed NOMA-MEC system. It is considered that the BS is located at the centre of a disc with radius r, and the users are randomly distributed within the disc. The distance between the eavesdropper and users is fixed as $D_{1,e}$ and $D_{2,e}$. In the DDPG-based algorithm, each point at the x-axis includes 300 episode, each episode

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3194685

TABLE II: Table of Parameters

Radius of the disc	1000 m
Carrier frequency	$f = 1$ GHz
AWGN spectral density	$N_0 = -174$ dBm/Hz
Path loss exponent	$\alpha = 3.76$
Bandwidth for each sub-channel	$B = 1$ MHz
CPU cycles for each bit of tasks	$\mu = 10^3$
Data of tasks for each user	$D = 0.2$ Mbits
Computation capacity of the MEC server	$C_0 = 10$ GHz
Learning rate of DDPG	0.001
Experience memory size	$\|\mathcal{R}\| = 10000$
Batch size	$S = 32$, $\rho = 0.01$, $\gamma = 0.6$
Exploration noise variance	2
Discount factor	0.995

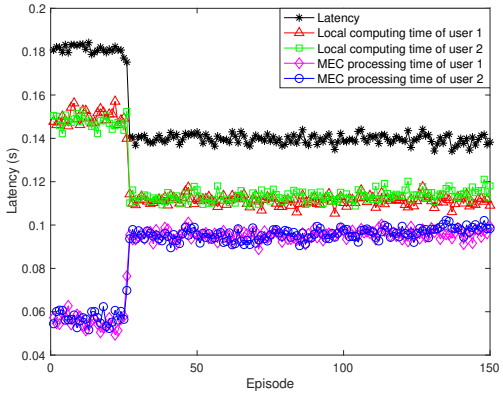


Fig. 2: The convergence of the proposed DDPG-based algorithm in the two-user scenario. $C_1 = C_2 = 1$ GHz, $Pt = 30$ dBm, and $D_{1,e} = D_{2,e} = 1500$ m.

includes 400 steps, and the neural networks are randomly initialized at each point of the x-axis. In order to compare the performance of the proposed DDPG-based algorithm, the full local computing scheme is included as the benchmark, where all tasks are computed at users. The simulation parameters are shown in Table II.

The convergence of the proposed power allocation algorithm is examined in Fig. 2, which shows that the DDPG-based algorithm is able to converge to a stable structure within around 30 episodes. Particularly, during power allocation, the secrecy rate of both users is increased, and then MEC processing time g_1 and g_2 is increased. On the other hand, local computing time is decreased since more tasks are offloaded to the MEC server. As a result, the latency of the investigated system is reduced. It can thus be concluded that there is a trade-off between local computing time and MEC processing time. Moreover, it can be inferred that the optimal solution can be obtained when the values of all four terms are the same, which is around 0.11 s. Therefore, the proposed DDPG-based algorithm can achieve about 90% of the global optimal solution.

In Fig. 3, the impact of transmit power on various aspects is demonstrated. With the increasing of transmit power, the secrecy rate of users is increased, and then the task assignment coefficients β_i is increased accordingly. In this case, the time consumption for both local computing and MEC processing is reduced, and hence, the latency of the considered system is significantly reduced compared to the full local computing scheme. Moreover, it can be observed that hybrid SIC decoding can outperform fixed SIC decoding in terms of latency. By

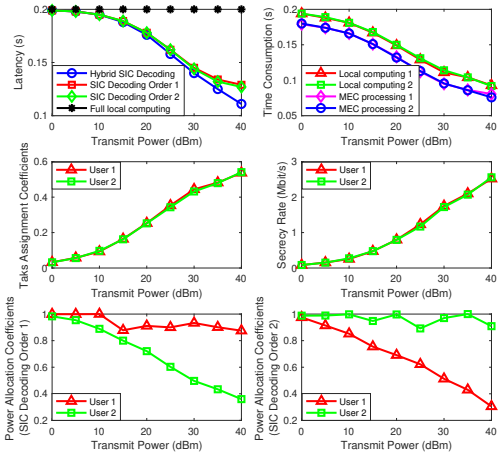


Fig. 3: The impact of transmit power in the two-user scenario. $C_1 = C_2 = 1$ GHz, and $D_{1,e} = D_{2,e} = 1500$ m.

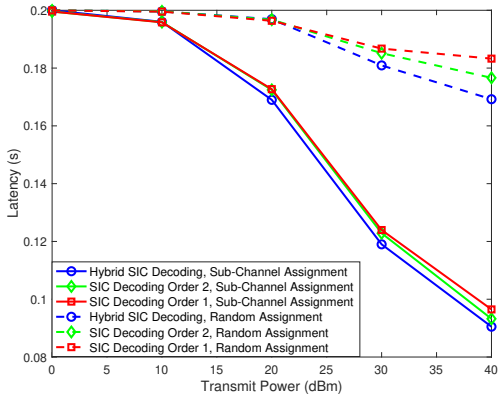
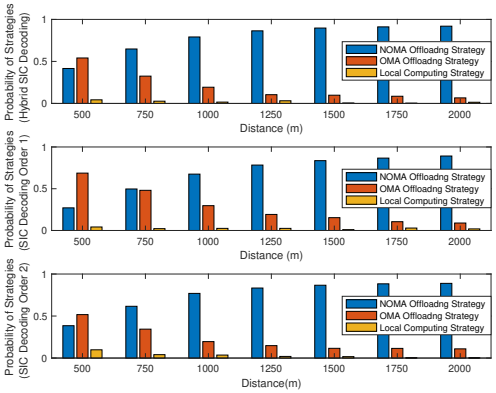


Fig. 4: The impact of transmit power in the multi-user scenario. $N = 10$, $K = 5$, $C_i = 1$ GHz, and $D_{i,e} = 1500$ m, $\forall i \in \mathcal{N}$.

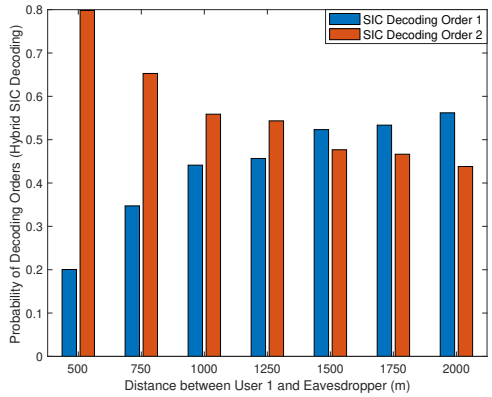
comparing fixed SIC decoding, one can see that the latency with SIC Decoding Order 2 is slightly lower than that with SIC Decoding Order 1. This is due to the fact that the condition of SIC Decoding Order 2 is more readily achieved, as presented in Proposition 3 and Proposition 4. Furthermore, it is worth pointing out that the power allocation coefficients of user 2 and user 1 are respectively reduced in SIC Decoding Order 1 and SIC Decoding Order 2, which confirms the insights in Remark 2 and Remark 3.

The performance of the considered system in the multi-user scenario is shown in Fig. 4, where a scheme with random assignment is selected as the benchmark. It can be found that the matching-based sub-channel assignment algorithm can significantly reduce the latency. This is due to the fact that based on sub-channel assignment, each user can select not only a better channel to the BS, but also a worse channel to the eavesdropper. The matching-based sub-channel assignment is dynamically combined with the DDPG-based power allocation scheme, and a better performance can thus be achieved. Furthermore, it is worth pointing out that the hybrid SIC decoding can achieve the best performance compared to fixed SIC decoding in the multi-user scenario.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3194685



(a)



(b)

Fig. 5: The impact of the distance between user 1 and the eavesdropper in the two-user scenario. $C_1 = C_2 = 1$ GHz, $Pt = 40$ dBm, and $D_{2,e} = 1500$ m. (a) Probability of strategies. (b) Probability of SIC decoding orders.

The probability of the offloading strategies in Section IV-C is shown in Fig. 5(a), where both hybrid and fixed SIC decoding schemes are included. With the increasing distance between user 1 and the eavesdropper, the probability for performing NOMA offloading strategy is increased, while the probability for adopting OMA offloading strategy and local computing strategy is decreased. Due to the fact that the channel condition $|h_{1,e}|^2$ is significant when the distance between user 1 and the eavesdropper is less than 1000 m, the probability for adopting NOMA offloading strategy in SIC Decoding Order 2 is greater than that in SIC Decoding Order 1. Moreover, as shown in Fig. 5(b), SIC Decoding Order 2 plays a dominant role when the distance is less than 1000 m. These results confirms the conditions for adopting different SIC decoding orders, i.e., Proposition 3 and Proposition 4. Therefore, it can be claimed that SIC Decoding Order 2 has advantages in practical systems. When the distance between user 1 and the eavesdropper is greater than 1250, there is no significant difference in strategy selection, and the probability of these two SIC decoding orders is comparable.

In Fig. 6, the scenario when distances between the eavesdropper and both users are simultaneously increased is studied. Due to the deterioration of the channel conditions between

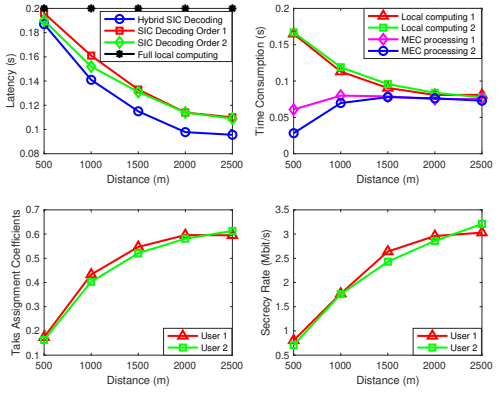


Fig. 6: The impact of the distance between the eavesdropper and users in the two-user scenario. $C_1 = C_2 = 1$ GHz, and $Pt = 40$ dBm.

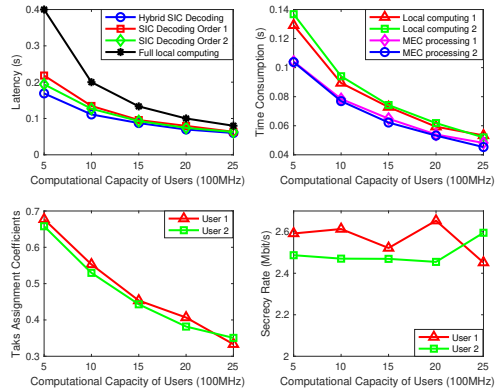


Fig. 7: The impact of the computational capacity of users in the two-user scenario. $Pt = 40$ dBm, and $D_{1,e} = D_{2,e} = 1500$ m.

users and the eavesdropper, the secrecy rate is improved, and the size of offloaded tasks is increased accordingly. Hence, although the MEC processing time is slightly increased, the local computing time is significantly reduced. As a result, the latency of the considered system with both hybrid and fixed SIC decoding can be reduced. It is worth to highlight that the gap between SIC Decoding Order 1 and SIC Decoding Order 2 in terms of latency is decreasing with the increasing distance between users and the eavesdropper. This is due to the fact that compared to SIC Decoding Order 1, it is easier for SIC Decoding Order 2 to achieve the NOMA offloading strategy when the channel conditions are relatively poor. Moreover, it shows that the improvement caused by hybrid SIC decoding is more prominent when the channel conditions are good.

In Fig. 7, the computational capacity of both users simultaneously increases from 500 MHz to 2500 MHz. It is indicated that the latency of all schemes is reduced. In terms of the local computing scheme, the increase of computational capacity can directly reduce the computing time at users. In the considered NOMA-MEC system, with the increasing computational capacity, more tasks can be allocated for local computing, as shown by the decreasing task assignment coefficients. For this reason, the MEC processing time is reduced, although the secrecy rate still remains at the same level (between 2.4 Mbit/s and 2.7 Mbit/s). Since most of the

tasks are completed locally with the increasing computational capacity, the gap between the proposed solution and the full local computing scheme becomes small. However, the hybrid SIC decoding still achieves the best performance, and SIC Decoding Order 2 can outperform SIC Decoding Order 1.

VIII. Conclusions

This paper investigated a secure NOMA-based MEC system with hybrid SIC decoding, where users tend to simultaneously offload confidential tasks to the MEC server in the presence of an eavesdropper. In order to minimize the latency of the considered system, power allocation, task assignment, and computational resource allocation were jointly studied. It was revealed that there is a trade-off between local computing time and MEC processing time, and hence, the closed-form solutions of computational resource allocation and task assignment were derived. By obtaining the conditions of different strategies, a strategy selection mechanism was established, and the corresponding power allocation problems were formulated. Based on the reinforcement learning, a DDPG-based power allocation algorithm was proposed, which can dynamically select the appropriate strategy and provide the near-optimal power allocation solution. Moreover, by comparing the conditions of different decoding orders, it was indicated that the NOMA offloading strategy is easier to be implemented if the decoding order at the BS is in the opposite order of that at the eavesdropper. The considered system was also investigated in a multi-user scenario, where a matching-based sub-channel assignment problem was developed in conjunction with the DDPG-based algorithm. The performance of the proposed scheme was demonstrated and validity of the provided insights verified by the simulation results. It is worth mentioning that designing the system for specific applications and simulating it with real-world datasets is a promising research direction.

Appendix A: Proof of Proposition 1

With the NOMA offloading strategy, both users offload tasks to the MEC server, i.e., $\beta_i > 0, \forall i \in \{1,2\}$. By substituting (21) into (13), the MEC computing time of any user i can be presented as

$$T_{0,i}^{\text{com}} = \frac{\mu(\beta_1 D_1 + \beta_2 D_2)}{C_0}. \tag{67}$$

The MEC processing time can be expressed as

$$g(\beta_i) = \frac{\beta_i D_i}{R_{i,\text{s}}} + \frac{\mu(\beta_1 D_1 + \beta_2 D_2)}{C_0}. \tag{68}$$

From (25), the following condition can be obtained:

$$\begin{cases} \dfrac{\mu(1-\beta_1^*)D_1}{C_1} = \dfrac{\beta_1^* D_1}{R_{1,\text{s}}} + \dfrac{\mu(\beta_1^* D_1 + \beta_2^* D_2)}{C_0}, \\ \dfrac{\mu(1-\beta_2^*)D_2}{C_2} = \dfrac{\beta_2^* D_2}{R_{2,\text{s}}} + \dfrac{\mu(\beta_1^* D_1 + \beta_2^* D_2)}{C_0}. \end{cases} \tag{69}$$

Hence, the optimal task assignment coefficients can be presented as

$$\begin{cases} \beta_1^* = \dfrac{\mu D_1 R_{1,\text{s}} C_0 - \mu\beta_2^* D_2 R_{1,\text{s}} C_1}{D_1 C_0 C_1 + \mu D_1 R_{1,\text{s}} C_1 + \mu D_1 R_{1,\text{s}} C_0}, \\ \beta_2^* = \dfrac{\mu D_2 R_{2,\text{s}} C_0 - \mu\beta_1^* D_1 R_{2,\text{s}} C_2}{D_2 C_0 C_2 + \mu D_2 R_{2,\text{s}} C_2 + \mu D_2 R_{2,\text{s}} C_0}. \end{cases} \tag{70}$$

By substituting (26), the optimal task assignment coefficient of user 1 can be presented as follows:

$$\begin{aligned} \beta_1^* &= \frac{\mu D_1 R_{1,\text{s}} C_0 - \mu\beta_1^* D_1 R_{2,\text{s}} C_1}{D_1 C_0 C_1 + \mu D_1 R_{1,\text{s}} C_1 + \mu D_1 R_{1,\text{s}} C_0} \\ \Rightarrow \beta_1^* &= \frac{\mu R_{1,\text{s}} C_0}{C_0 C_1 + \mu R_{1,\text{s}} C_1 + \mu R_{1,\text{s}} C_0 + \mu R_{2,\text{s}} C_1}. \end{aligned} \tag{71}$$

Similarly, the expression of user 2's optimal task assignment coefficient is given by

$$\beta_2^* = \frac{\mu R_{2,\text{s}} C_0}{C_0 C_2 + \mu R_{2,\text{s}} C_2 + \mu R_{2,\text{s}} C_0 + \mu R_{1,\text{s}} C_2}. \tag{72}$$

Note that constraint (16b) should be satisfied by the derived optimal task assignment coefficients. Take user 1's task assignment coefficient as an example, by including (16b), the following inequalities can be obtained:

$$\begin{cases} \mu R_{1,\text{s}} C_0 \geq 0, \\ C_0 C_1 + \mu R_{1,\text{s}} C_1 + \mu R_{2,\text{s}} C_1 \geq 0. \end{cases} \tag{73}$$

Due to the fact that the above condition always holds with any non-negative secrecy rate, constraint (16b) is satisfied by the derived task assignment coefficients.

In the OMA offloading strategy, one of the users computes all tasks locally, and hence, the task assignment coefficient and secrecy rate of this user are zero. For the OMA offloading user, denoted by user i, the optimal computational resource allocation coefficient can be obtained from (21) as $\tau_i^* = 1$. From (25), the following equation can be obtained:

$$\frac{\mu(1-\beta_i^*)D_i}{C_i} = \frac{\beta_i^* D_i}{R_{i,\text{s}}} + \frac{\mu\beta_i^* D_i}{C_0}. \tag{74}$$

The optimal task assignment coefficient of the OMA offloading user can be presented as

$$\beta_i^* = \frac{\mu R_{i,\text{s}} C_0}{C_0 C_i + \mu R_{i,\text{s}} C_i + \mu R_{i,\text{s}} C_0}. \tag{75}$$

It can be shown that the same solution can be obtained from (71) and (72) by setting the secrecy rate of the local computing user as zero. Hence, the obtained solutions in (71) and (72) can also be utilized for the OMA offloading strategy. Moreover, the obtained solution in (75) always satisfies constraint (16b) with any non-negative secrecy rate of the OMA user. The proof of this proposition is completed.

Appendix B: Proof of Proposition 2

This proposition can be proved by assuming the NOMA offloading strategy can be selected, that is, condition $R_{i,\text{s}} > 0, \forall i \in \{1,2\}$ can be achieved. In this case, the secrecy data of both users should satisfy the condition in Remark 1. From (34), the following expression can be obtained:

$$R_{2,\text{s}} = \frac{A_2 R_{1,\text{s}} + D_2 C_0 C_1 - D_1 C_0 C_2}{A_1}. \tag{76}$$

By substituting the above equation into (30), the local computing time of user 1 can be transformed as follows:

$$\begin{aligned}F_1 &= \frac{\mu D_1[\mu(A_1+A_2)R_{1,\mathrm{s}}+C_0(A_1+\mu D_2C_1-\mu D_1C_2)]}{\mu(A_1C_1+A_2C_1+A_1C_0)R_{1,\mathrm{s}}+C_0C_1(A_1+\mu D_2C_1-\mu D_1C_2)}\\ &= \frac{\mu D_1[(A_1+A_2)R_{1,\mathrm{s}}+\mu D_1C_0^2]}{\mu(A_1C_1+A_2C_1+A_1C_0)R_{1,\mathrm{s}}+\mu D_1C_0^2C_1}\\ &= \frac{\mu^2(D_1+D_2)R_{1,\mathrm{s}}+\mu D_1C_0}{\mu(C_1+C_2+C_0)R_{1,\mathrm{s}}+C_0C_1}. \end{aligned} \tag{77}$$

The derivative of F_1 can be presented as follows:

$$\begin{aligned}\frac{\partial F_1}{\partial R_{1,\mathrm{s}}} &= \frac{\mu^2C_0(D_2C_1-D_1C_2-D_1C_0)}{[\mu(C_1+C_2+C_0)R_{1,\mathrm{s}}+C_0C_1]^2}\\ &= \frac{-\mu A_1C_0}{[\mu(C_1+C_2+C_0)R_{1,\mathrm{s}}+C_0C_1]^2}. \end{aligned} \tag{78}$$

It is indicated that the monotonicity of user 1's local computing time F_1 is decided by term A_1. Specifically, if $A_1>0$ holds, F_1 is monotonically decreasing with the increasing secrecy rate; otherwise, F_1 is monotonically increasing. Therefore, in the case that $A_1\leq 0$, in order to minimize the latency, the secrecy rate of user 1 tends to be zero. This condition can be transformed as follows:

$$A_1\leq 0 \Rightarrow \frac{\mu D_1}{C_1}\leq \frac{\mu D_2}{C_0+C_2}. \tag{79}$$

Similarly, based on (34) and (31), the local computing time of user 2 can be expressed as

$$F_2 = \frac{\mu^2(D_1+D_2)R_{2,\mathrm{s}}+\mu D_2C_0}{\mu(C_1+C_2+C_0)R_{2,\mathrm{s}}+C_0C_2}. \tag{80}$$

The derivative of the above function is

$$\frac{\partial F_2}{\partial R_{2,\mathrm{s}}} = \frac{-\mu A_2C_0}{[\mu(C_1+C_2+C_0)R_{2,\mathrm{s}}+C_0C_2]^2}. \tag{81}$$

The monotonicity of user 2's local computing time can be revealed. Particularly, if $A_2\leq 0$, user 2's secrecy rate tends to be the minimum, and hence, user 2 will compute all tasks locally. The following condition can be obtained:

$$A_2\leq 0 \Rightarrow \frac{\mu D_2}{C_2}\leq \frac{\mu D_1}{C_0+C_1}. \tag{82}$$

It shows that the OMA offloading strategy will be adopted if (79) or (82) is satisfied. Moreover, it is worth to mention that conditions (79) and (82) cannot be satisfied simultaneously, and therefore, this case will not lead to the local computing strategy. This proposition is proved.

Appendix C: Proof of Proposition 3

With respect of p_1, the partial derivative of user 1's secrecy rate in (44) can be presented as

$$\frac{\partial R_{1,\mathrm{s}}}{\partial p_1} = \frac{B[|h_1|^2(p_2|h_{2,e}|^2+1)-|h_{1,e}|^2(p_2|h_2|^2+1)]}{\ln(2)(p_1|h_{1,e}|^2+p_2|h_{2,e}|^2+1)(p_1|h_1|^2+p_2|h_2|^2+1)}. \tag{83}$$

It is indicated that the monotonicity of $R_{1,\mathrm{s}}$ is decided by the term $|h_1|^2(p_2|h_{2,e}|^2+1)-|h_{1,e}|^2(p_2|h_2|^2+1)$. That is, if $|h_1|^2(p_2|h_{2,e}|^2+1)>|h_{1,e}|^2(p_2|h_2|^2+1)$, (83) is always greater than zero, and $R_{1,\mathrm{s}}$ is monotonically increasing with p_1. If $|h_1|^2(p_2|h_{2,e}|^2+1)<|h_{1,e}|^2(p_2|h_2|^2+1)$ holds, $R_{1,\mathrm{s}}$ is monotonically decreasing with p_1. Moreover, due to the fact that the secrecy data rate of user 1 is zero when $p_1=0$, the secrecy rate $R_{1,\mathrm{s}}$ can be positive only if $|h_1|^2(p_2|h_{2,e}|^2+1)>|h_{1,e}|^2(p_2|h_2|^2+1)$. In terms of user 2, the partial derivative of the secrecy rate is given by

$$\frac{\partial R_{2,\mathrm{s}}}{\partial p_2} = \frac{B(|h_2|^2-|h_{2,e}|^2)}{\ln(2)(p_2|h_{2,e}|^2+1)(p_2|h_2|^2+1)}. \tag{84}$$

It can be shown that the secrecy rate of user 2 can be positive only if $|h_2|^2>|h_{2,e}|^2$. As a result, the condition for adopting SIC Decoding Order 1 can be presented as

$$\begin{cases} |h_1|^2(p_2|h_{2,e}|^2+1)>|h_{1,e}|^2(p_2|h_2|^2+1),\\ |h_2|^2>|h_{2,e}|^2. \end{cases} \tag{85}$$

The first inequality in (85) can be expressed as

$$p_2(|h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2)>|h_{1,e}|^2-|h_1|^2. \tag{86}$$

The above inequality can be respectively presented as

$$\begin{cases} p_2>\dfrac{|h_{1,e}|^2-|h_1|^2}{|h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2},\\ |h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2>0, \end{cases} \tag{87}$$

and

$$\begin{cases} p_2<\dfrac{|h_{1,e}|^2-|h_1|^2}{|h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2},\\ |h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2<0. \end{cases} \tag{88}$$

In the first inequality of (87), the condition of user 2's power allocation coefficient is derived. Since $p_2\leq 1$, the following condition should be satisfied:

$$\begin{cases} 1>\dfrac{|h_{1,e}|^2-|h_1|^2}{|h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2},\\ |h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2>0. \end{cases} \tag{89}$$

The above inequalities can be expressed as

$$\begin{cases} \dfrac{|h_1|^2}{|h_{1,e}|^2}>\dfrac{|h_2|^2+1}{|h_{2,e}|^2+1},\\ \dfrac{|h_1|^2}{|h_{1,e}|^2}>\dfrac{|h_2|^2}{|h_{2,e}|^2}. \end{cases} \tag{90}$$

By including the second inequality of (85), the following condition can be obtained:

$$\frac{|h_1|^2}{|h_{1,e}|^2}>\frac{|h_2|^2}{|h_{2,e}|^2}>1, \tag{91}$$

where the first inequality in (90) is removed since it is always satisfied with the above condition. In this case, the following inequality can be obtained:

$$\frac{|h_{1,e}|^2-|h_1|^2}{|h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2}<0. \tag{92}$$

Hence, the feasible region of p_2 is $p_2\in(0,1]$.

Similarly, according to $p_2\geq 0$, the following inequalities can be obtained from (88):

$$\begin{cases} 0<\dfrac{|h_{1,e}|^2-|h_1|^2}{|h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2},\\ |h_1|^2|h_{2,e}|^2-|h_2|^2|h_{1,e}|^2<0. \end{cases} \tag{93}$$

The above condition can be transformed as

$$\begin{cases} |h_1|^2 > |h_{1,e}|^2, \\ \dfrac{|h_1|^2}{|h_{1,e}|^2} < \dfrac{|h_2|^2}{|h_{2,e}|^2}, \end{cases} \tag{94}$$

and the following condition can be obtained:

$$\frac{|h_2|^2}{|h_{2,e}|^2} > \frac{|h_1|^2}{|h_{1,e}|^2} > 1, \tag{95}$$

where the second inequality of (85) is already contained. In this case, the feasible region of p_2 can be expressed as

$$p_2 \in \left(0, \min\left\{\frac{|h_{1,e}|^2 - |h_1|^2}{|h_1|^2|h_{2,e}|^2 - |h_2|^2|h_{1,e}|^2}, 1\right\}\right). \tag{96}$$

By including the conditions in (91) and (95), the conditions for adopting SIC Decoding Order 1 can be presented as

$$\begin{cases} |h_1|^2 > |h_{1,e}|^2, \\ |h_2|^2 > |h_{2,e}|^2. \end{cases} \tag{97}$$

This proposition is proved.

Appendix D: Proof of Proposition 4

The partial derivative of (49) with respect of p_1 is given by

$$\frac{\partial R_{1,\mathrm{s}}}{\partial p_1} = \frac{B[|h_1|^2(p_2|h_{2,e}|^2+1) - |h_{1,e}|^2]}{\ln(2)(p_1|h_{1,e}|^2+p_2|h_{2,e}|^2+1)(p_1|h_1|^2+1)}. \tag{98}$$

It is indicated that the secrecy rate of user 1 can be positive only if inequality $|h_1|^2(p_2|h_{2,e}|^2+1) > |h_{1,e}|^2$ holds. For user 2, the partial derivative can be presented as

$$\frac{\partial R_{2,\mathrm{s}}}{\partial p_2} = \frac{B[|h_2|^2 - |h_{2,e}|^2(p_1|h_1|^2+1)]}{\ln(2)(p_2|h_{2,e}|^2+1)(p_2|h_2|^2+p_1|h_1|^2+1)}. \tag{99}$$

Similarly, user 2 can offload tasks to the MEC server by utilizing NOMA schemes only if $|h_2|^2 > |h_{2,e}|^2(p_1|h_1|^2+1)$ is satisfied. Therefore, with this SIC decoding order, the following condition should be satisfied:

$$\begin{cases} |h_1|^2(p_2|h_{2,e}|^2 + 1) > |h_{1,e}|^2, \\ |h_2|^2 > |h_{2,e}|^2(p_1|h_1|^2 + 1). \end{cases} \tag{100}$$

From (100), the condition of the power allocation coefficients can be obtained:

$$\begin{cases} p_2 > \dfrac{|h_{1,e}|^2 - |h_1|^2}{|h_1|^2|h_{2,e}|^2}, \\ p_1 < \dfrac{|h_2|^2 - |h_{2,e}|^2}{|h_1|^2|h_{2,e}|^2}. \end{cases} \tag{101}$$

Due to the fact that $p_i \in [0, 1]$, the inequalities in (101) can be transformed as

$$\begin{cases} 1 > \dfrac{|h_{1,e}|^2 - |h_1|^2}{|h_1|^2|h_{2,e}|^2}, \\ 0 < \dfrac{|h_2|^2 - |h_{2,e}|^2}{|h_1|^2|h_{2,e}|^2}. \end{cases} \tag{102}$$

As a result, the following conditions can be derived:

$$\begin{cases} \dfrac{|h_1|^2}{|h_{1,e}|^2} > \dfrac{1}{|h_{2,e}|^2 + 1}, \\ |h_2|^2 > |h_{2,e}|^2. \end{cases} \tag{103}$$

Moreover, in this case, the feasible region of users' power allocation coefficients can be respectively expressed as

$$p_1 \in \left(0, \min\left\{\frac{|h_2|^2 - |h_{2,e}|^2}{|h_1|^2|h_{2,e}|^2}, 1\right\}\right), \tag{104}$$

and

$$p_2 \in \left(\max\left\{0, \frac{|h_{1,e}|^2 - |h_1|^2}{|h_1|^2|h_{2,e}|^2}\right\}, 1\right]. \tag{105}$$

This proposition is proved.

References

[1] A. C. Baktir, A. Ozgovde, and C. Ersoy, "How can edge computing benefit from software-defined networking: A survey, use cases, and future directions," *IEEE Commun. Surveys Tutorials*, vol. 19, no. 4, pp. 2359–2391, Fourthquarter 2017.

[2] P. Corcoran and S. K. Datta, "Mobile-edge computing and the internet of things for consumers: Extending cloud computing and services to the edge of the network," *IEEE Consum. Electron. Mag.*, vol. 5, no. 4, pp. 73–74, Oct. 2016.

[3] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surveys Tutorials*, vol. 19, no. 3, pp. 1628–1656, thirdquarter 2017.

[4] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tutorials*, vol. 19, no. 4, pp. 2322–2358, Fourthquarter 2017.

[5] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Commun. Surveys Tutorials*, vol. 19, no. 3, pp. 1657–1681, thirdquarter 2017.

[6] P. Porambage, J. Okwuibe, M. Liyanage, M. Ylianttila, and T. Taleb, "Survey on multi-access edge computing for internet of things realization," *IEEE Commun. Surveys Tutorials*, vol. 20, no. 4, pp. 2961–2991, Fourthquarter 2018.

[7] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C. L. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.

[8] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept. 2013, pp. 611–615.

[9] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[10] Z. Ding, P. Fan, and H. V. Poor, "Impact of non-orthogonal multiple access on the offloading of mobile edge computing," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 375–390, Jan. 2019.

[11] K. Wang, F. Fang, D. B. d. Costa, and Z. Ding, "Sub-channel scheduling, task assignment, and power allocation for OMA-based and NOMA-based MEC systems," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2692–2708, 2021.

[12] F. Fang, Y. Xu, Z. Ding, C. Shen, M. Peng, and G. K. Karagiannidis, "Optimal resource allocation for delay minimization in NOMA-MEC networks," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7867–7881, 2020.

[13] Z. Ding, J. Xu, O. A. Dobre, and H. V. Poor, "Joint power and time allocation for NOMA–MEC offloading," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6207–6211, 2019.

[14] L. P. Qian, A. Feng, Y. Huang, Y. Wu, B. Ji, and Z. Shi, "Optimal SIC ordering and computation resource allocation in MEC-aware NOMA NB-IoT networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2806–2816, 2019.

[15] Z. Yang, C. Pan, J. Hou, and M. Shikh-Bahaei, "Efficient resource allocation for mobile-edge computing networks with NOMA: Completion time and energy minimization," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7771–7784, Nov. 2019.

[16] H. Li, F. Fang, and Z. Ding, "Joint resource allocation for hybrid noma-assisted MEC in 6G networks," *Digital Communications and Networks*, vol. 6, no. 3, pp. 241–252, 2020.

[17] J. Zhu, J. Wang, Y. Huang, F. Fang, K. Navaie, and Z. Ding, "Resource allocation for hybrid NOMA MEC offloading," *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4964–4977, 2020.

[18] K. Wang, Z. Ding, D. K. C. So, and G. K. Karagiannidis, "Stackelberg game of energy consumption and latency in MEC systems with NOMA," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2191–2206, 2021.

[19] B. Li, W. Wu, W. Zhao, and H. Zhang, "Security enhancement with a hybrid cooperative NOMA scheme for MEC system," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2635–2648, 2021.

[20] W. Wu, F. Zhou, R. Q. Hu, and B. Wang, "Energy-efficient resource allocation for secure NOMA-enabled mobile edge computing networks," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 493–505, 2020.

[21] L. Qian, W. Wu, W. Lu, Y. Wu, B. Lin, and T. Q. S. Quek, "Secrecy-based energy-efficient mobile edge computing via cooperative non-orthogonal multiple access transmission," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4659–4677, 2021.

[22] W. Wu, X. Wang, F. Zhou, K.-K. Wong, C. Li, and B. Wang, "Resource allocation for enhancing offloading security in NOMA-enabled MEC networks," *IEEE Systems Journal*, vol. 15, no. 3, pp. 3789–3792, 2021.

[23] X. Wang, W. Wu, B. Lyu, and H. Wang, "Delay minimization for secure NOMA mobile-edge computing," in *2019 IEEE 19th International Conference on Communication Technology (ICCT)*, 2019, pp. 1529–1534.

[24] Y. Xu, T. Zhang, D. Yang, Y. Liu, and M. Tao, "Joint resource and trajectory optimization for security in UAV-assisted MEC systems," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 573–588, 2021.

[25] S. Han, X. Xu, S. Fang, Y. Sun, Y. Cao, X. Tao, and P. Zhang, "Energy efficient secure computation offloading in NOMA-based mMTC networks for IoT," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5674–5690, 2019.

[26] Q. Wang, H. Hu, H. sun, and R. Q. Hu, "Secure and energy-efficient offloading and resource allocation in a NOMA-based MEC network," in *2020 IEEE/ACM Symposium on Edge Computing (SEC)*, 2020, pp. 420–424.

[27] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *IEEE Proc. Veh. Technol. Conf. (VTC Spring)*, June 2013, pp. 1–5.

[28] Z. Ding, R. Schober, and H. V. Poor, "Unveiling the importance of sic in NOMA systems—part 1: State of the art and recent findings," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2373–2377, 2020.

[29] ——, "Unveiling the importance of sic in NOMA systems—part II: New results and future directions," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2378–2382, 2020.

[30] Z. Ding, D. W. K. Ng, R. Schober, and H. V. Poor, "Delay minimization for NOMA-MEC offloading," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1875–1879, Dec. 2018.

[31] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction," *Robotica*, vol. 17, no. 2, pp. 229–235, 1999.

[32] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *International conference on machine learning*. PMLR, 2014, pp. 387–395.

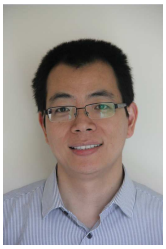
[33] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," *arXiv preprint arXiv:1811.03378*, 2018.

[34] X. Guo, Y. Chen, and Y. Wang, "Learning-based robust and secure transmission for reconfigurable intelligent surface aided millimeter wave UAV communications," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1795–1799, 2021.

[35] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 375–388, 2021.

Kaidi Wang (S'16-M'20) received the MS degree in communications and signal processing from Newcastle University in 2014, and the PhD degree in wireless communication from the University of Manchester. He is a research fellow of Wireless Communications at the Institute for Communication Systems, home of 5GIC and 6GIC at the University of Surrey. His current research interests include game theory, convex optimization, machine learning, non-orthogonal multiple access, and mobile edge computing.

Haodong Li received the M.S. in communication engineering in 2018 from The University of Manchester, U.K., where he is currently pursuing a Ph.D. degree in wireless communication from Nov. 2018. His current research interests include machine learning, convex optimization, non-orthogonal multiple access (NOMA), mobile edge computing (MEC), and physical layer security (PLS).



Zhiguo Ding (S'03-M'05-F'20) received his B.Eng in Electrical Engineering from the Beijing University of Posts and Telecommunications in 2000, and the Ph.D degree in Electrical Engineering from Imperial College London in 2005. From Jul. 2005 to Apr. 2018, he was working in Queen's University Belfast, Imperial College, Newcastle University and Lancaster University. Since Apr. 2018, he has been with the University of Manchester as a Professor in Communications. From Oct. 2012 to Sept. 2022, he has also been an academic visitor in Princeton University.

Dr Ding' research interests are B5G networks, machine learning, cooperative and energy harvesting networks and statistical signal processing. He is serving as an Area Editor for the *IEEE Open Journal of the Communications Society*, an Editor for *IEEE Transactions on Vehicular Technology*, and *IEEE Communications Surveys & Tutorials*, and was an Editor for *IEEE Wireless Communication Letters*, *IEEE Transactions on Communications*, *IEEE Communication Letters* from 2013 to 2016. He recently received the EU Marie Curie Fellowship 2012-2014, the Top IEEE TVT Editor 2017, IEEE Heinrich Hertz Award 2018, IEEE Jack Neubauer Memorial Award 2018, IEEE Best Signal Processing Letter Award 2018 , Friedrich Wilhelm Bessel Research Award 2020, and Best Paper Award at IEEE ICCC-2021. He is a Fellow of the IEEE, a Distinguished Lecturer of IEEE ComSoc, and a Web of Science Highly Cited Researcher in two categories 2020.

Pei Xiao is a professor of Wireless Communications at the Institute for Communication Systems, home of 5GIC and 6GIC at the University of Surrey. He is the technical manager of 5GIC/6GIC, leading the research team in the new physical layer work area, and coordinating/supervising research activities across all the work areas (https://www.surrey.ac.uk/institute-communication-systems/5g-6g-innovation-centre). Prior to this, he worked at Newcastle University and Queen's University Belfast. He also held positions at Nokia Networks in Finland. He has published extensively in the fields of communication theory, RF and antenna design, signal processing for wireless communications, and is an inventor on over 15 recent 5GIC patents addressing bottleneck problems in 5G systems.