RIS-Assisted Ad Hoc Edge for Optimal User Distribution in Service-Intensive Scenarios

Chenxuan Hou*, Chenyang Wang*, Kai Dong†, Xiaofei Wang* and Tarik Taleb†

* College of Intelligence and Computing, Tianjin University, Tianjin, China.

† Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland.

Email: {chenxuanhou, chenyangwang, xiaofeiwang}@tju.edu.cn, kai.dong@oulu.fi, talebtarik@gmail.com

***Abstract*—Massive device connections in upcoming 6G networks have led to a sharp increase in network traffic volume, posing significant challenges in providing reliable performance guarantees, e.g., low latency. The Computing Power Network (CPN) is a new framework for resource integration involving multiple parties. It integrates the resources of various owners via the network, providing users with efficient and adaptable services. Due to the uncertainty of the signal quality, the majority of existing studies do not adequately organize the topology of user allocation in CPNs when optimizing network resources. Reconfigurable Intelligent Surface (RIS) is a new type of network node for constructing future smart radio environments with high spectral efficiency and nearly zero energy consumption that can offer new access options for user allocation in CPNs. In this paper, we investigate the user access allocation in a RIS-assisted Ad Hoc Edge (RAHE) scenario where the users are with service-intensive demands. To maximize the overall service tasks of the system constrained by a service time threshold, we propose a RIS-assisted interval scheduler strategy (RS^3) approach to balancing the whole system service completion and total latency. Specifically, RS^3 is a graph-theoretic optimization method based on the interval scheduling problem. The numerical simulation results demonstrate that our proposed RS^3 approach is superior to commonly utilized methods in terms of the number of serves given the service time constraint.**

***Index Terms*—Computing power network, Reconfigurable intelligent surface, User allocation.**

I. Introduction

For widespread mobile edge computing (MEC) [1] applications, computing and storage resources will no longer be confined to large centralized data centers. Resource nodes appear to be ubiquitous [2]. In addition to single telecom operators or cloud service providers, suppliers of resources will also include small and medium-sized regional service providers and even individual providers. Concurrently, new types of services [3] (e.g., digital twin, virtual reality) will have higher, lower and broader requirements for computing power, latency, and coverage, respectively. As a novel solution, computing power network [4] (CPN) has emerged in recent years. In essence, the objective of CPN is to integrate computing power with network resources in order to meet the requirements of modern applications for large-scale computing and data processing. However, effectively meeting the diverse business requirements of users and maximizing the utilization of available network resources presents a significant challenge. In CPN architecture, communication transmission bridges computing resources and user requirements. However, the severe path loss in Millimeter Wave (mmWave) or sub-THz bands in the 6G networks [5] makes it difficult to optimize the resource utilization of CPN. To address this issue, the deployment of Reconfigurable Intelligent Surfaces (RIS) [6] has emerged as a promising technique for improving energy and spectral efficiency. In particular, it comprises an array of passive reflective components. Each can dynamically adjust the amplitude and phase of the impinging signal to serve the user in a dedicated direction. By reconstructing the radio propagation environment, RIS provides the new access option between edge servers and end-users during computing task transmitting [7].

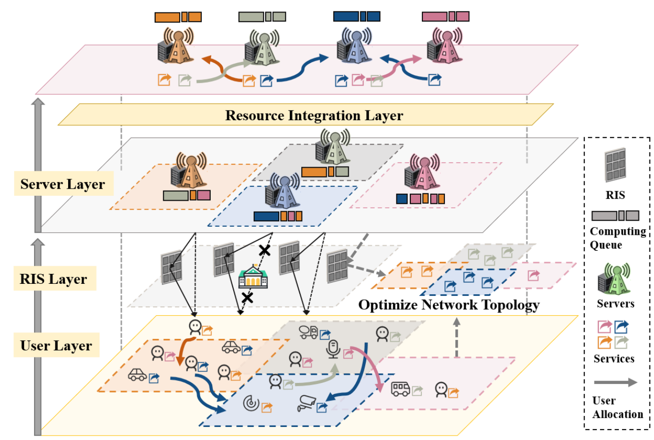

Fig. 1. RIS-assisted ad hoc edge scenario under CPN.

Based on the preceding discussion, the issue of joint allocation optimization in CPNs introduces new challenges that must be addressed. In particular, it entails determining how to effectively leverage RIS to dynamically adjust resource allocation in order to provide users with optimal network connectivity options.

Sun et al. [8] propose the need for integrating resources in CPNs to cope with user allocation requirements. Yang [8] and Sun et al. [9] introduced RIS as a communication resource in multi-user MEC systems and formulated offloading energy optimization problems. The deployment of RIS aims to achieve a trade-off between the limited energy and computational resources while meeting strict latency constraints. Nevertheless, RIS is primarily utilized as a signal enhancement facility and does not actively collaborate with the user allocation function. Zhuang et al. [10] proposed a collaborative edge scenario among multiple RISs, employing multi-agent reinforcement learning to solve complex RIS configuration and allocation problems across multiple regions. Most of these studies neglect RIS's crucial role as a link within a CPN that can have a significant effect on the network's topology.

Therefore, we concentrate on the ad hoc edge scenario under CPN. Ad hoc edge (AHE) is a self-organizing management paradigm within the computing power network aimed at addressing the chaotic and diverse nature of user demands. Positioned at the underlying layer, AHE dynamically perceives and optimizes user requirements, enabling the optimal matching and allocation of computational resources. Through self-organizing management, intelligent decision-making, and resource coordination, AHE adapts to varying demands, ensures efficient resource utilization, and provides reliable computational support for edge users.

Creating a reasonable network topology in AHE presents several obstacles: (1) The complex wireless network typologies in practical and severe path loss of the electromagnetic wave signal propagation in the mmWave/sub-THz bands make it challenging to guarantee the critical quality-of-service requirements in the beyond 5G and 6G networks. (2) How to allocate users under dense heterogeneous services and increase system throughput brings new challenges in the optimization of network resource allocation.

RIS-assisted Ad Hoc Edge (RAHE) is an AHE scenario for RIS-assisted network topology. The architecture of the entire RAHE system is depicted in Figure 1. Without prior planning, this network can facilitate temporary communication and data exchange between servers and users via the new transmission path established by RIS. The neighboring RIS serves as a relay to the corresponding server, aiding it in improving communication performance and offering an alternative access option for users at their designated location in the network. By optimizing the network's topology, redundant nodes and links can be eliminated, resulting in reduced network construction and maintenance costs. Therefore, it demonstrates enhanced resilience and dependability, allowing for smooth operations in complex and unpredictable environments.

The main contributions are listed in the following:

- To achieve significant resource utilization and meet the computing power demands of digital transformation, we implement RAHE in CPN to maximize the use of available resources in the edge-end layered architecture.
- As part of our graph-based approach, we propose a RIS-assisted interval scheduling strategy (RS^3) to solve the user allocation problem. By utilizing graph structures, our proposed approach facilitates the seamless integration of new users and resources into the system, thereby enabling optimization across a wide range of scenarios and meeting diverse requirements.
- The experimental results illustrate the effectiveness of the proposed RS^3 compared to some benchmarks in edge scenarios. With more services, our system achieves a request completion rate of up to 94%.

The subsequent sections of this paper are structured as follows. Section II presents the system model. The graph-based approach to interval scheduling strategy is discussed in Section III. Section IV presents the simulation results, and Section V summarizes the paper.

II. System Model

We consider a RAHE scenario aided by multiple RISs. As shown in Figure 1, this service placement architecture consists of a RIS-assisted communication layer and an edge computing layer supported by graph topology. The placement of servers within the system is arbitrary, and each RIS only assists a single server. We specify that the servers and devices are respectively denoted as $\mathbb{M} = \{1, 2, ..., M\}$ and $\mathbb{N} = \{1, 2, ..., N\}$. Denote $\mathbb{I} = \{1, 2, ..., I\}$ as the set of RISs. Each user will randomly generate a service request, with the service type indicated by the set $\mathbb{S} = \{1, 2, ..., S\}$. Assume the user has one computing service and each service is processed by a specific dominant server. Let $J_{n,s} = (L_{n,s}, Q_{n,s})$ denote the service of user n, where $L_{n,s}$ represents the service upload data size and $Q_{n,s}$ is the overall number of the CPU cycles to complete the service. Here, we assume that all the servers and devices are equipped with a single antenna, and the RIS is with K reflection elements. These assumptions will not affect the performance of our proposed algorithms.

Typically, system servers cache only the programs required for a particular service type. Due to loading, processing new services can result in lengthy delays. Consequently, servers have defining characteristics for particular services. Forwarding the user's service to the dominant server in this situation is preferable. Even though this process incurs additional communication delays due to forwarding, RIS can significantly reduce the communication burden. This paper investigates network configuration and selection methods for RIS-assisted services.

A. Channel Model

In this paper, we assume that all the communication channels are block-fading based, i.e., the channel coefficients remain constant during current block data transmission but may change over different time slots. Therefore, the channel coefficients from the n-the device to the associated m-th server, from the n-th device to the i-th RIS, from the i-th RIS to the associated m-th server are denoted by $h_{n,m}$, $\mathbf{h}_{i,n} \in \mathbb{C}^{K\times 1}$, and $\mathbf{h}_{i,m} \in \mathbb{C}^{K\times 1}$, respectively. The Channel State Information (CSI) estimation is out of the scope of this paper. For more details, refer to [11]–[13] for CSI acquisition between the server and device, server and RIS, as well as the link between RIS and device. With the full knowledge of the CSI among all the communication links, the server can perform resource management strategies to maximize the computation and communication efficiencies.

B. Computation Model

For the direct link between the m-th server and the associated n-th device, the offloading rate is given by

$$R_n^d = B \log_2 \left(1 + \frac{P_n |h_{n,m}|^2}{\sigma^2}\right) \tag{1}$$

where B denotes the allocated bandwidth, P_n is the transmit power of the n-th device, σ^2 represents the variance of the Gaussian noise, respectively.

If the direct link between the server and the device is too weak due to the long propagation distance (i.e., severe path loss in mmWave or sub-THz bands) or the blockage such as surrounding vehicles, buildings, etc., the RIS is scheduled to enhance the end-to-end signal quality. In such case, the rate between server and device assisted by the RIS is given by [14]

$$R_n^r = B\log_2\left(1+\frac{P_n|\mathbf{h}_{i,m}^H\mathbf{\Theta}_i\mathbf{h}_{i,n}+h_{n,m}|^2}{\sigma^2}\right) \tag{2}$$

where

$$\mathbf{\Theta}_i = \mathrm{diag}\{\eta_1 e^{j\phi_{i,1}},\cdots,\eta_k e^{j\phi_{i,k}},\cdots,\eta_K e^{j\phi_{i,K}}\} \tag{3}$$

are the matrix of reflection amplitude and phase shift adjustment at the associated i-th RIS. In particular, $\eta_k \in [0,1]$ and $\phi_{i,k} \in [0,2\pi)$. Here, we assume all the RIS elements are with the ideal reflection amplitudes, i.e., $\eta_1 = \eta_2 = \cdots = \eta_K = 1$. In addition, an ideal continuous phase shift $\phi_{i,k}$ (i.e., accurate phase shift adjustment for signal reflection to the direction of the scheduled user) is supported here by the hardware for simplicity. Therefore, the decision of user allocation determines the specific optimization analysis.

We assume that each user generates only one service. The service j upload time can be obtained as:

$$t_{tr}(j) = \frac{L_{n,s}(j)}{b_m(j)\cdot R_n^r + (1-b_m(j))\cdot R_n^d} \tag{4}$$

where $b_m(j) = \{0,1\}$ denotes the indicator of the reflecting link through the RIS between the server and scheduled user. Specifically, $b_m(j) = 1$ indicates that RIS is scheduled to serve the user, otherwise $b_m(j) = 0$.

Each user's computational duties can be delegated to either the dominant server or the nearest non-dominant server. Assume that each server has the essential program to execute its function. The non-dominant server should have a configuration delay, such as a cache delay or waiting period. Due to the difficulty of specifying the configuration time in actuality, we set the processing rates differently for dominant and non-dominant servers.

Let $a_m(j) = \{0,1\}$ denote the edge execution indicator, when $a_m(j) = 1$ indicates that the service j is executed on dominant server m, and 0 is otherwise. In this way, the edge execution time can be expressed as:

$$t_s(j) = \frac{L_{n,s}(j)Q_s(j)}{a_m(j)f_{m,1} + (1-a_m(j))\,f_{m,0}} \tag{5}$$

where Q_s indicates the number of CPU cycles required to process each unit byte, while $f_{m,0}$, $f_{m,1}$ represents the dominant and non-dominant server's computing capacity to distinguish the server's service configuration time. The downlink transmission time can be neglected due to the small size of the computing results of j.

C. Problem Formulation

According to the model presented above, the allocated service of each server in the RAHE scenario has two parts, one for dominant services and the other for non-dominant services. The latency for completing the service of server m is

$$T_{m,n,s}(j) = \sum\left(t_{tr}(j) + t_s(j)\right) \tag{6}$$

Therefore, the service completion time for each server m can be expressed as

$$T_t(j) = \sum\left(x_n * T_{m,n,s}(j)\right) \tag{7}$$

where x_n takes the value 0 or 1 to indicate whether the service was completed or not. The value of 1 denotes successful completion of the service, whereas the value of 0 signifies that the service was not completed. Then, the total number of services processed by server m can be determined.

$$U_{m,n} = \sum\left(x_n | T_t \le T_{max}\right) \tag{8}$$

Our objective is to maximize the size of the subset of services U while minimizing the total completion time, which can be formulated as

$$\begin{aligned}
P1 : \max_{U,\mathbf{\Theta}_i} \; & \sum_{m,n=1}^{M,N} U_{m,n} \\
s.t. \quad & \text{C1:} \quad T_t \le T_{\max} \\
& \text{C2:} \quad \phi_{i,k} \in [0,2\pi), \quad \forall i \in I, k \in K \\
& \text{C3:} \quad a_m, b_m \in \{0,1\}, \quad \forall m \in M \\
& \text{C4:} \quad x_n \in \{0,1\}, \quad \forall n \in N
\end{aligned}$$

Here, C1 ensures that the size of each selected element is less than T_{max}, under the assumption of minimizing the sum of selected elements. T_{max} denotes the upper bound on the total completion time of the subset of services U. C2 is the phase shift adjustment constraint, i.e., the phase shift adjustment of each element of the RIS within the range of 0 and 2π. C3 denotes the allocated location of each service and can only be allocated to one server. C4 is a decision variable that indicates whether the n-th service allocated to the corresponding server has been completed. Here, the user's position change determines the phase shift adjustment of the associated RIS.

III. User Allocation Algorithm based on Maximum Independent Set

A. Combinatorial optimization

Depending on the server's properties, we can transform the problem into optimal interval scheduling [15]. Consider services (users) that require a specific machine (server) to be used for a period of time. The tolerance threshold $T_{m,n,s}$ of delay for each user becomes the start time $ST_{m,n,s}$ and end time $ET_{m,n,s}$ specified by a service. These services should be processed without interruption. Thus, the problem becomes the need to find a solution for selecting a subset of servers such that (1) there is no time conflict between user services

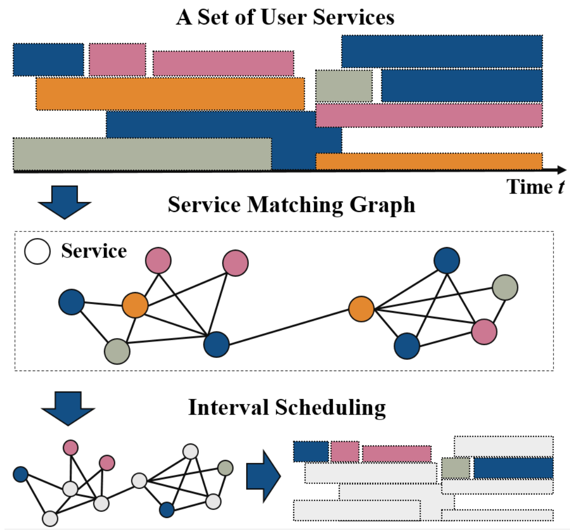

Fig. 2. An example for scheduling services in one server: the top diagram shows the set of user services for scheduling, the middle diagram shows the service placement of interval scheduling problems, and the bottom diagram shows the corresponding service association diagram.

in using the servers; (2) all services are processed as much as possible; (3) the total user latency is as low as possible.

Then $T_{m,n,s} = ST_{m,n,s} - ET_{m,n,s}$ is the service length of user n. In general, user scheduling can be viewed as the process of selecting the appropriate server for each user. With hard deadlines, the following equation should be satisfied for each user's service: $\cap_{n \in J_{n,s}} [ST_{m,n,s}, ET_{m,n,s}) = \varnothing$. Suppose the server wants to allocate users in service set $J_{n,s}$. This means that $J = \bigcup_{n=1}^{N} J_{n,s}$. Each user can only be allocated to one server, so $\cap_{J_{n,s} \in J} J_{n,s} = \varnothing$. Consequently, the completion time of the problem can be expressed as

$$R_{m,n,s} = \sum_{n=1}^{N} \sum_{n \in J_{n,s}} T_{m,n,s} \tag{9}$$

In $P1$, there are two different optimization objectives: maximizing the size of the service subset U_n and minimizing the total completion time $T_{m,n,s}$. Subject to the constraint that the services do not overlap, these two objectives may conflict, requiring trade-off considerations. Therefore, we propose a weight parameter β to balance the trade-off between the two objectives based on the idea of weighting. The optimization problem $P1$ can be transformed into

$$\begin{aligned} P2: \max \ & \beta U_n - (1-\beta) \sum_{n=1}^{N} \Phi\left(J_{n,s}\right) R_n \\ s.t. \quad & \text{C1:} \quad J = \textstyle\bigcup_{n=1}^{N} J_{n,s} \\ & \text{C2:} \quad \cap_{n \in J_{n,s}} [ST_{m,n,s}, ET_{m,n,s}) = \varnothing \\ & \text{C3:} \quad \cap_{J_{n,s} \in J} J_{n,s} = \varnothing \end{aligned}$$

In $P2$, $\Phi\left(J_{n,s}\right)$ indicates the server location indication for service completion, which is represented as:

$$\Phi\left(J_{n,s}\right) = \begin{cases} 0 & J_{n,s} = \varnothing \\ 1 & J_{n,s} \neq \varnothing \end{cases} \tag{10}$$

B. Integer Linear Programming for MIS Problem

This interval scheduling problem involves finding the largest possible set of non-overlapping services. By transforming this problem into a graph model, the objective becomes finding the largest subset of vertices with no edges between them, which is equivalent to the maximum independent set problem.

A conflict diagram is used to illustrate the conflict between services in Figure 2. The problem can be represented by an undirected graph $G(V, E)$, called an allocation graph. Each vertex $V = \{v_1, v_2, ..., v_n\}$ is a set of n network nodes, where each node denotes the user service. An edge will connect the affected nodes if there is a temporal conflict between two services. Therefore, the maximum number of services that a server can fulfil corresponds to the maximum number of independent edges in the corresponding undirected graph. According to the definition of the maximum independent set, the number of nodes (services) is selected as the optimal solution within the constraints specified.

We utilize an Integer Linear Programming (ILP) approach to address the weighted maximum independent set problem, incorporating the $P2$ objective function. The ILP framework easily accepts different constraints and objective functions within the graph structure. The mathematical formulation of the problem is expressed as:

$$\begin{aligned} P3: \max & \sum_{i \in T} [\beta \cdot x_i - (1-\beta) \cdot T_{m,n,s} \cdot x_i] \\ s.t. \quad & \text{C1:} \quad x_i \in \{0, 1\}, \quad \forall i \in V \\ & \text{C2:} \quad x_i + x_j \leq 1, \quad \forall (i,j) \in E \end{aligned}$$

Algorithm 1 Algorithm for interval scheduling problem

Input:

The initial parameter settings, e.g. N users, M servers, S service types, and the allocation user service, including an undirected graph $G = (V, E)$

Output:

Optimal interval service scheduling *(MIS problem)*

A maximum independent set $L \subseteq V$

Define 0-1 variables $x_i \in 0, 1$, representing whether node i is in the independent set, where $i \in V$.

Define the objective function $f(x) = \sum_{i \in V} x_i$, which maximizes the size of the independent set.

Add the constraint $x_i + x_j \leq 1$ for each edge $(i, j) \in E$, which ensures that two adjacent nodes cannot both be in the independent set.

Set the lower and upper bounds for all variables as 0 and 1, respectively, i.e., $0 \leq x_i \leq 1$, where $i \in V$.

Solve the linear program to obtain a maximum.

Update the MIS L as the best solution

Update interval service scheduling

A binary variable $x_i = 1$ in C1 means service i is included in the optimal subset U, while a value of 0 indicates that it does not. C2 indicates that two adjacent nodes cannot simultaneously appear in an independent set.

The initial step of Algorithm 1 is designed for configuring the environment based on the system model presented in Section III. Determine the computation time for the RIS-assisted edge system once the number of users and service types has been initialized. Then, the ILP problem is formulated with the objective function consisting of the weighted sum of the service subset size and the total completion time (lines 4-6) and the constraints ensuring that no services overlap in the solution. After determining the upper and lower limits of L, we perform a linear programming relaxation on the objective function values L (lines 7-10) to identify the largest independent set. Therefore, for each examined value of L, we execute at most $N(N-1)$ ILP programs.

IV. Simulation Results and Discussions

A. Simulation Settings

We evaluate the performance of our algorithm using the Public EUA dataset [16]. In the simulation, users categorized randomly into four distinct labels are randomly distributed in a square area of 1200m × 1200m. The distribution of base stations is shown in Figure 3. In addition, the eight RISs are distributed randomly on the circumference that corresponds to the coverage limits of the eight base stations. These RISs with fixed altitude $H = 5$ m provide communication forwarding functions for all requirements within the set area. Each edge server needs one of the programs to execute the service, and it is assumed that the server caches programs for the same amount of time.

In the simulation, we consider the realistic scenario calculation configurations, which means all the sizes of services and user tolerance limits are set randomly. Direct links between servers and devices will be randomly blocked.

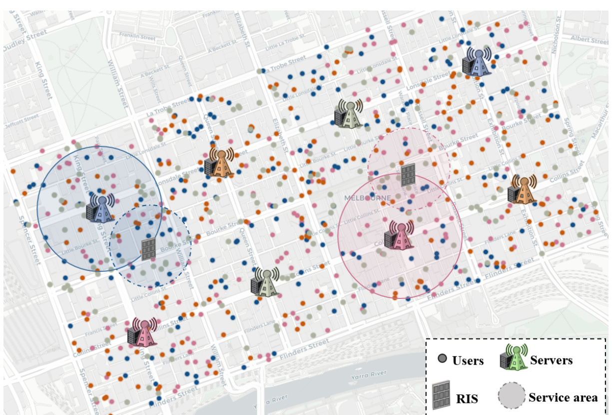

Fig. 3. Diagram of the experimental setup.

To demonstrate the benefits of our proposed algorithm, we compare it with the following baseline algorithms:

1) **Random Selection:** The system randomly selects servers in response to user services.

2) **Greedy Selection:** The system selects the server closest to users. Alternatively, the system selects the dominant server.

3) **GraphDP Selection** [17]**:** The system employs a graph-based dynamic programming approach to optimize service packing near edges with a high overlap ratio.

4) **The proposed RS^3:** Consider both the dominant server and the nearest server with the assistance of RIS.

B. Simulation Result

1) **System throughput simulation:** We compared the amount of system completion for user services at different time thresholds. Figure 4 shows that the ad hoc edge assisted by RIS can still effectively allocate user services when the time threshold is small. When the time threshold of the system is about 200s, the allocation efficiency of our system is about one and a half times that of greedy and random. With the help of RIS, the system's throughput increases significantly as the time threshold is gradually increased. The server's completion rate of user services is maintained at about 80%, reaching a maximum of 94%.

At the beginning of the figure, the difference in system service completion for each algorithm is very small. This is due to the high volume of user services and the fact that communication delays are not the primary cause of inter-service conflicts. As the time threshold increases, the advantages of RIS-assisted communication become gradually more evident. Due to the shorter service delivery time in the RIS-assisted edge system, more services can be completed within the allotted time. Additionally, RIS expands server coverage so that additional users can be allocated to the dominant server. RIS offers more user allocation access options in the AHE scenario. Experiments indicate that the server processes more dominant services with the assistance of RIS, reducing the amount of time wasted on cache loading. This significantly increased the throughput of the system.

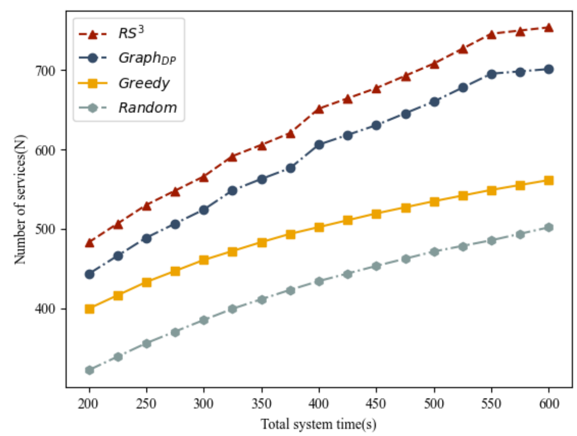

Fig. 4. Comparative analysis of user service completion.

2) **User allocation simulation:** We compare the distribution of users on the server under different methods. Figure 3 depicts the information regarding user locations for the practical dataset. Due to the extremely uneven distribution of user locations and their service attributes in the dataset, servers in certain regions are subject to a substantial influx of service requests. Figure 5 demonstrates that Edge Server 3 experiences significantly greater load pressure than the other servers. This server handles two to three times as many requests for service as the other servers. The results of the experiment indicate that

our proposed RS^3 scheme can have a significant impact on the load balancing of services. In the RAHE scenario, the server with the highest service pressure receives at least 30% relief. Although the GraphDP algorithm can also increase the system's service throughput, it is evident that its load-balancing capabilities are insufficient. Our approach can improve the system's service completion while managing the user layer's cluttered topology. This enhancement significantly reduces the server stress imbalance resulting from the density of service distribution in real datasets.

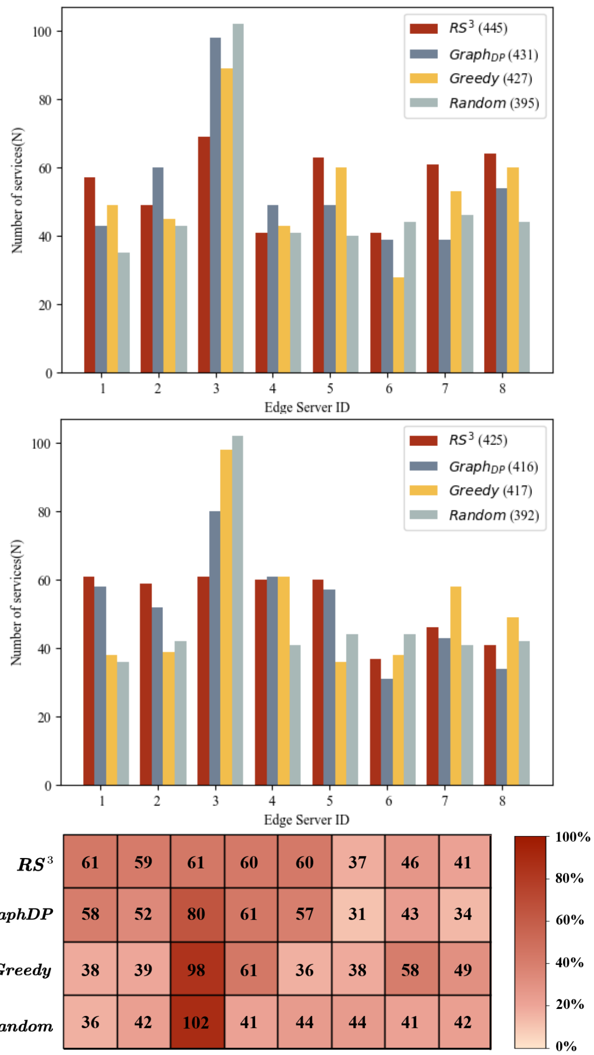

Fig. 5. Results of user allocation under different intervals.

V. Conclusion

In this paper, we have presented a RAHE scenario based on a computing power network. In this case, the network can improve user access by modifying the access locations of the allocated services in the event of high service demand. Then, we have designed a maximum independent set solver based on graph theory to optimize the maximum service completion among users. According to simulation results, our method has improved the unbalanced distribution of services in real-world datasets while simultaneously enhancing the system's service completion. In systems with more users and services, the service completion rate has reached 94%, relieving 40% of the load pressure.

Acknowledgment

This research work is supported by China NSFC under grant No. 62072332, China NSFC (Youth) under grant No. 62002260, the Tianjin Xinchuang Haihe Lab under Grant No. 22HHXCJC00002, and is partially supported by the Business Finland 6Bridge 6Core project under Grant No. 8410/31/2022, the Research Council of Finland (former Academy of Finland) IDEA-MILL project under Grant No. 352428, the European Union's Horizon Europe research and innovation programme under the 6GSandbox project with Grant Agreement No. 101096328, and Research Council of Finland 6G Flagship Programme under Grant No. 346208.

References

[1] T. Tan and G. Cao, "Deep learning on mobile devices through neural processing units and edge computing," in *IEEE INFOCOM*, London, United Kingdom, May 2022, pp. 1209–1218.

[2] X. Wang *et al.*, "Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9441–9455, 2020.

[3] S. Duan, D. Wang, H. Wu, and X. Shen, "Distributed artificial intelligence empowered by end-edge-cloud computing: A survey," *IEEE Commun. Surv. Tutorials*, vol. 25, no. 1, pp. 591–624, 2023.

[4] J. Liu *et al.*, "Computing power network: A testbed and applications with edge intelligence," in *IEEE INFOCOM WKSHPS*, New York, USA, May 2022, pp. 1–2.

[5] B. Ji *et al.*, "A survey of computational intelligence for 6g: Key technologies, applications and trends," *IEEE Trans. Ind. Inform.*, vol. 17, no. 10, pp. 7145–7154, 2021.

[6] Y. Liu *et al.*, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 2021.

[7] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, 2021.

[8] Y. Yang, Y. Gong, and Y. Wu, "Intelligent-reflecting-surface-aided mobile edge computing with binary offloading: Energy minimization for iot devices," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 12 973–12 983, 2022.

[9] C. Sun, W. Ni, Z. Bu, and X. Wang, "Energy minimization for intelligent reflecting surface-assisted mobile edge computing," in *IEEE ICCC*, Xiamen, China, Jul. 2021, pp. 254–259.

[10] S. Zhuang *et al.*, "When multi-access edge computing meets multi-area intelligent reflecting surface: A multi-agent reinforcement learning approach," in *IEEE IWQoS*, Oslo, Norway, Jun. 2022, pp. 1–10.

[11] A. Alkhateeb, O. El Ayach, G. Leus, and R. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

[12] X. Wei and L. Dai, "Channel estimation for RIS assisted wireless communications - part I: fundamentals, solutions, and future opportunities," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1398–1402, 2021.

[13] C. Wu, C. You, Y. Liu, X. Gu, and Y. Cai, "Channel estimation for star-ris-aided wireless communication," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 652–656, 2021.

[14] S. Mao, M. Dong, J. Zhao, J. Wu, and V. Leung, "Reconfigurable intelligent surface-assisted secure mobile edge computing networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6647–6660, 2022.

[15] W. Wu, J. Zhu, H. Huang, X. Xu, and Y. Zhang, "Dynamic idle time interval scheduling for hybrid cloud workflow management system," in *IEEE SMC*, Miyazaki, Japan, Oct. 2018, pp. 3596–3602.

[16] P. Lai *et al.*, "Optimal edge user allocation in edge computing with variable sized vector bin packing," in *Proc. ICSOC*, Hangzhou, China, Nov. 2018, pp. 230–245.

[17] P. Oikonomou *et al.*, "Graph-based approaches for the interval scheduling problem," in *IEEE ICPADS*, Hong Kong, Dec. 2020, pp. 649–654.