Industrial Engineering and Applications
L.-C. Tang (Ed.)

doi:10.3233/ATDE230060

Multi-Dimensional Resource Allocation Algorithm for End-to-End Network Slicing

Jihong Zhao [a, b], Zhaoyang Zhu [a, 1], Zihao Huang [a], Beibei Wang [a]
[a] *Xi'an University of Posts and Telecommunications*
[b] *Xi'an Jiaotong University*

Abstract. With the rapid development of 5G networks, the emerging 5G mobile system is expected to serve a large number of users with differentiated performance requirements. Different application scenarios have differentiated resource and quality of service requirements. To optimize the allocation of network multidimensional resources, an allocation algorithm to maximize multi-dimensional resource utilization is proposed for the resource optimization problem of Enhanced Mobile Broadband (eMBB) and Ultra-Reliable and Low Latency Communication (URLLC) network slices. The algorithm combines communication, computing and storage resources to provide a solution for multi-dimensional resource allocation. And it is able to guarantee the high speed demand of users in eMBB slices and the low latency demand of users in URLLC slices. The resource problem for multiservice slicing is modelled as a non-linear mixed integer programming problem. Combining the augmented Lagrange algorithm and the branch-and-bound method to solve the problem, the optimal resource allocation strategy is obtained. It is demonstrated that the proposed algorithm can improve the system throughput, increase the resource utilization, improve the rate of eMBB class services and reduce the delay of URLLC class services compared with other algorithms.

Keywords. 5G, network slicing, multi-dimensional resources, resource allocation

1. Introduction

With the development of 5G technology and the introduction of network slicing technology, the resources to be managed by 5G networks are no longer limited to traditional communication resources such as power and spectrum, but are gradually evolving to pre-distribution for multi-dimensional resource management, which includes caches, computing resources, etc. Typical application scenarios include three categories, namely Enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low Latency Communication (URLLC) and massive machine type communication (MMTC). In addition, as the allocation of resources across dimensions can interact with each other, optimizing a resource in isolation does not give the best results for the user. Therefore, it is essential to combine multi-dimensional resources for dynamic allocation.

To accommodate the multi-service requirements of the current 5G verticals, network slicing techniques are introduced. Splitting a physical network into multiple logical networks based on different needs is more adaptable to current developments. There have been many achievements in previous research on network slicing. Currently, network slicing is one of the most cost-effective ways to meet the varying demands of

[1] Corresponding Author: Zhaoyang Zhu, 13164336721@163.com.

multi-logical network services and is a key driver for the flexibility and versatility of 5G to deliver a variety of service approaches [1]. In the literature [2], the authors investigate how network slicing and fog nodes can be combined to securely access remote service data while ensuring low latency. A network orchestration architecture for dynamic network slicing is considered in literature [3] and demonstrates how to provide dynamic network slicing to enterprise services. The above literatures only consider the traditional communication resources, don't consider the optimization of multi-dimensional resources for resource allocation. In view of this situation, in multi-user networks, a multi-dimensional resource management scheme is proposed in the literature [4], The scheme minimizes user latency by jointly optimizing communication and computing resources, and derives an expression for optimal resource allocation. To meet the demand for low-latency communication. In the literature [5], the authors assume that offloading users are known and propose a combined bidding-based service provider selection scheme to allocate spectrum and computing resources to users to improve the effectiveness of service providers. However, the impact of storage resources on system resource allocation is not considered in the above-mentioned literature for multi-dimensional resource allocation schemes, and adding consideration of storage resources can optimize the system resource allocation scheme.

In this paper, the resource optimization of eMBB and URLLC network slices is studied, and an allocation algorithm to maximize multi-dimensional resource utilization is proposed. Different from the previous resource allocation schemes, it is the consideration of adding storage resources. A multi-dimensional resource allocation scheme for joint communication, computing and storage resources is proposed. So as to improve the overall system resource utilization and rate, and to reduce the service response latency. The resource problem of multiservice slicing is modelled as a non-linear mixed integer programming problem. Linear programming is then used to solve the overall resource allocation problem.

2. System Model and Problem Description

2.1. System Model

In this paper, an end-to-end network slicing model for 5G networks is considered. The users U = {1,2...U} are randomly distributed in the coverage area of the base station. The total bandwidth of the system is equally divided into N sub-bandwidths B, whose corresponding sub-carrier set is Q_n= {1,2...., N}. The user set in the eMBB slice is set to U_1, $|U_1| = U_1$, and the user set in the URLLC slice is set to U_2, $|U_2| = U_2$, where U=$U_1 U U_2$, $U_1 \cap U_2$ = ø. When building a system model, the system model incorporates three major capabilities within the network: communication, computation and storage.

- Network model

Consider a network topology g = {v, e}, where v denotes the set of network nodes. Nodes can be routers, users or servers. e denotes the links between physical network nodes. M = $\{M_A, M_B, o\}$ denotes the set of some nodes in the model with service capabilities. Where $M_A \subset$ V denotes nodes in the model with caching capabilities and $M_B \subset$ V denotes nodes in the model with computing capabilities. A source server is represented by a source node o which satisfies all service requests in the network.

S = {S_A, S_B} denotes the set of service requests from users, s ∈ S denotes the

services requested by users, $s^a \in S_A$ denotes the content services requested by users, and $s^b \in S_B$ denotes the computation services requested by users.

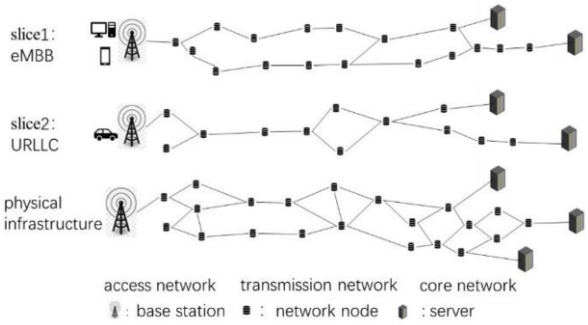

Figure 1. Network slicing model.

- Computation and Storage Model

In the compute and store model, h_m^s is a binary variable to represent the compute/cache policy on network node m. For user u, $h_m^s = 1$ if network node m can satisfy the user's service request; conversely $h_m^s = 0$. Here, it is assumed that the service node has a limited service capacity within the network. c_m^a is applied to denote the size of the cache capacity of network node m and c_m^b is applied to denote the computational capacity of network node m. Based on the above expression, the total amount of all content stored on a network node must be less than the maximum capacity it can cache.

$$\sum_s h_m^s o_s \leq c_m^a, \quad \forall \mathrm{m} \in M_A \tag{1}$$

Moreover, the sum of the computational tasks of the network nodes must be less than the maximum amount of computation that their computational resources can provide.

$$\sum_s h_m^s o_s \leq c_m^b, \quad \forall \mathrm{m} \in M_B \tag{2}$$

For the source server o, it is assumed that there is no upper limit on its capacity. So, for all service requests from user's s, $h_o^s = 1$.

2.2. Problem Description

For end-to-end network slicing, this paper applies network slicing techniques to solve the problem of having two types of services eMBB and URLLC requirements.

Slice 1: eMBB (throughput priority)

When user u in slice 1 sends a service request, the corresponding rate is given by the following equation.

$$r_u = \sum_{n=1}^{N} a_{n,u} r_{n,u} \geq \sigma_0, \quad u \in U_1, \tag{3}$$

$$a_{n,u} \begin{cases} 1, & \text{assign subcarrier N to user U} \\ 0, & \text{others} \end{cases} \tag{4}$$

$$r_{n,u} = B\log\left(1 + \frac{g_{n,u}p_n}{N_0 B}\right) \tag{5}$$

$r_{n,k}$ is the data rate of user u subcarrier n. N_0 is the noise power spectral density. $g_{n,u}$ and p_n denote the channel gain and transmit power of the associated subcarrier n of user u, respectively. σ_0 is the throughput requirement threshold of user u.

Slice 2: URLLC (Latency Priority)

For users with low latency requirements, the packet length of each user follows an exponential distribution, assuming that the maximum data arrival rate of the users in this slice obeys a Poisson distribution. The probability of disruption for user u in slice 2 with a delay greater than the delay threshold is given by the following equation.

$$P_r\left\{D_u \geq D_{u,max}\right\} = e^{-(R_u - d_{u,max})D_{u,max}}, u \in U_2 \tag{6}$$

where D_u and $D_{u,max}$ are the latency of user u and the maximum latency threshold that user u can tolerate, respectively. $d_{u,max}$ is the maximum data arrival rate for user u. In addition, this paper assumes that the maximum transmit power of the base station is P_0. Combining the above two slicing constraints, the system throughput.

$$R = \sum_{u=1}^{U}\sum_{n=1}^{N} a_{n,u}\, B\log(1 + \frac{g_{n,u}p_n}{N_0 B}) \tag{7}$$

With the objective of maximizing system throughput, the original problem is transformed into a mathematical model solved as follows:

$$\max_{\{p_n, a_{n,u}\}} \sum_{u=1}^{U}\sum_{n=1}^{N} a_{n,u}\, B\log(1 + \frac{g_{n,u}p_n}{N_0 B}) \tag{8}$$

$$s.t. \quad \sum_{n=1}^{N} a_{n,u} r_{n,u} \geq \sigma_0, u \in U_1 \tag{c1}$$

$$P_r\left\{D_u \geq D_{u,max}\right\} \leq \varepsilon, u \in U_2 \tag{c2}$$

$$\sum_{u=1}^{U}\sum_{n=1}^{N} a_{n,u}\, p_n = P_0, \quad p_n \in R^+ \tag{c3}$$

$$\sum_{u=1}^{U}\sum_{n=1}^{N} a_{n,u} = N, \quad a_{n,u} \in \{0,1\} \tag{c4}$$

$$\sum_{n=1}^{N} a_{n,u} \leq 1, n \in Q_N, \tag{c5}$$

$$h_m^s o_s \leq c_s^a, \quad \forall \mathrm{m} \in M_A \tag{c6}$$

$$\sum_s h_m^s o_s \leq c_s^b, \quad \forall \mathrm{m} \in M_B \tag{c7}$$

where, (11) is intended to maximize the throughput of the system; (c1) is to ensure that the transmission rate received by user u does not fall below the minimum rate demand threshold; and (c2) ensures that the probability of interruption of user u in slice 2 with a delay greater than the maximum delay que is less than the threshold e. For (c3), (c4) and (c5), assuming that all users make full use of the subcarrier and power resources, each subcarrier at most one user is associated. (c6) and (c7) indicate the caching/computing resource capacity limits of service nodes within the network, respectively.

3. Allocation Algorithm to Maximize Multi-dimensional Resource Utilization

Solve the multidimensional resource allocation problem of the whole system network based on the above optimal service response strategy. Since the original problem (11) is a difficult NP-hard problem to solve. In this paper, an approximate method is used to solve this problem by combining the branch-and-bound method. Relaxing the variable N_u and expanding its range of values: $N_u \in$ A, Where A= $\{N_u \in R^+ | \sum_{u\in U} N_u = N\}$, the relaxed optimization problem can be obtained as follows:

$$\min_{\{p_u, N_u\}} -\sum_{u=1}^{U} N_u \text{Blog}(1+\frac{y_u p_u}{N_u}) \tag{9}$$

$$\sigma_0 - BN_u \log(1+\frac{y_u p_u}{N_u}) \le 0, u \in U_1 \tag{c1}$$

$$e^{-(R_u - d_u)D_{u.max}} - \varepsilon \le 0\ , u \in U_2 \tag{c2}$$

$$\sum_{u=1}^{U} p_u - P_0 = 0, \quad p_u \in R^+, \tag{c3}$$

$$\sum_s h_n^s o_s \ \le\ c_s^a, \quad \forall \text{n} \in N_A \tag{c4}$$

$$\sum_s h_n^s o_s \ \le\ c_s^b, \quad \forall \text{n} \in N_B \tag{c5}$$

$$N_u \in A, \tag{c6}$$

where N is the number of subcarriers allocated to user u;$R_u = N_u \text{Blog}(1+\frac{y_u p_u}{N_u})$; y_u is the channel gain-to-noise ratio (CNR) of user u. p_u is the transmission power of user u; (c3) is to make full use of the wireless resources. (c4) and (c5) indicate the caching/computing resource capacity limits of service nodes within the network, respectively.

The augmented Lagrange algorithm is utilized to solve the optimization problem (18). First, the inequality constraint is relaxed by introducing the subsidiary variables x_k and y_k so that (c1) and (c2) are expressed in terms of the equation. Then, the augmented Lagrange function L_ρ can be obtained. as follows:

$$L_\rho(p, n, \alpha, \beta, \mu, \lambda) =$$

$$-\frac{R}{NB} + \frac{1}{2\rho}\sum_{u\in U_1} (\min\{0, \rho(R_u - \sigma_0) - \alpha_u\})^2$$

$$+\frac{1}{2\rho}\sum_{u\in U_2} (\min\{0, \rho(\varepsilon - P_r\{D_u \ge D_{u.max}\}) - \alpha_u\})^2 + \mu(P_0 - \sum_{u=1}^{U} p_u)$$

$$+\frac{\rho}{2}\left(\sum_{u=1}^{U} p_k - P_0\right)^2 + \lambda\left(N - \sum_{u=1}^{U} N_u\right) + \frac{\rho}{2}\left(\sum_{u=1}^{U} N_u - N\right)^2 + \frac{\rho}{2}\left(\sum_{u=1}^{U} N_u - N\right)^2 \tag{10}$$

where p = [P_1, P_2,…, P_u], n = [N_1, N_2,…, N_u], α = [α_1, α_2,…, α_{u_1}], β = [β_1, β_2,…, β_{u_2}]. α_u, β_u, μ and λ are Lagrange multipliers; ρ is the penalty factor.

To aid in the description of the algorithm, the following functions are defined in this paper:

$$f(p_u, N_u) = BN_u \log\left(1+\frac{y_u p_u}{N_u}\right) - \sigma_0 \tag{11}$$

$$g(p_u, N_u) = \varepsilon - e^{-(R_u - d_u)D_{u.max}} \tag{12}$$

$$h(p) = \sum_{u=1}^{U} p_u - P_0, \tag{13}$$

$$\tilde{h}(n) = \sum_{u=1}^{U} N_u - N \tag{14}$$

In addition, the gradient of the equation function and the inequality function are used in the PHR augmented Lagrange algorithm, denoted respectively as:

$$\nabla \mathrm{g}(p_u, N_u) = [\frac{\partial g}{\partial p_u}, \frac{\partial g}{\partial N_u}], u \in U \tag{15}$$

$$\nabla \mathrm{f}(p_u, N_u) = [\frac{\partial f}{\partial p_u}, \frac{\partial f}{\partial N_u}], u \in U_2 \tag{16}$$

$$\nabla \mathrm{h}(p) = \left[\frac{\partial h}{\partial p_1}, \frac{\partial h}{\partial p_2}, \dots, \frac{\partial h}{\partial p_u}\right] \tag{17}$$

$$\nabla \tilde{h}(n) = \left[\frac{\partial \tilde{h}}{\partial N_1}, \frac{\partial \tilde{h}}{\partial N_2}, \dots, \frac{\partial \tilde{h}}{\partial N_u}\right] \tag{18}$$

$$\frac{\partial f}{\partial p_u} = \frac{B N_u y_u}{(N_u + y_u p_u) \ln 2} \tag{19}$$

$$\frac{\partial f}{\partial N_u} = B \log\left(1 + \frac{y_u p_u}{N_u}\right) - \frac{B y_u p_u}{(N_u + y_u p_u) \ln 2} \tag{20}$$

The problem being relaxed will be solved by the Augmented Lagrange Algorithm, as shown in Algorithm 1.

Algorithm 1 Relaxed Optimization Problem

Initializing$\alpha, \beta, \mu, \lambda, \rho > 0, i = 0, \theta \in (0,1), i_{max} = 500, V_{i.old} = 10, \eta > 1$ $\delta \in [0,1]$.

While $V_i > \delta$ and $i < i_{max}$ do

The BFGS algorithms is used to solve the extended Lagrange function to obtain X^i, where $X^i = \{p_u^i, N_u^i | u = 1,2, \dots, U\}$.

$V_i = \omega$, where ω is shown in (21).

if $V_i > \delta$ then

if $i \geq 2$ and $V_i > \theta V_{i,old}$ then

ρ=ηρ

end *if*

$\alpha_u = max\{0, \alpha_u - \rho \mathrm{f}(p_u, N_u)\}, u \in U_1$

$\beta_u = max\{0, \beta_u - \rho \mathrm{g}(p_u, N_u)\}, u \in U_2$

$\mu = \mu - \rho\left(\sum_{u=1}^{U} - P_0\right)$,

$\lambda = \lambda - \rho\left(\sum_{u=1}^{U} N_u - N\right)$,

end *if*

$i = i + 1, x_0 = x$,

end while

$$\omega = \left(\left(\sum_{u=1}^{U} p_u - P_0 \right)^2 + \left(\sum_{u=1}^{U} N_u - N \right)^2 + \sum_{u \in U_1} \left[\min \left(\mathrm{f}(p_u, N_u), \frac{\alpha_u}{\rho} \right) \right]^2 + \sum_{u \in U_2} \left[\min \left(\mathrm{g}(p_u, N_u), \frac{\beta_u}{\rho} \right) \right]^2 \right)^{\frac{1}{2}} \tag{21}$$

For the optimal subcarrier and power allocation strategies, Algorithm 2 and the branch delimitation method are used in this chapter to obtain them. The integer constraint in the relaxed mixed integer programming problem is solved by Algorithm 2 for each relaxation problem. In Algorithm 2 power and joint subcarrier allocation is used. A combination of branch-and-bound methods is used to iterate until the resulting integer solution satisfies the iteration conditions. As a result, the optimal resource allocation strategy and spectral efficiency are obtained.

Algorithm 2 Power association subcarrier allocation algorithm

Through Algorithm 2, the nonlinear programming problem in which the variable N is relaxed in the optimization problem (9) is solved to obtain a lower bound L for (8) and initialize the resource allocation policy N_u^*, p_u^*

while ||h - l|≤ε do

$if\ N_u^* \in$ Z then

break

else the problem's upper bound is found. The number of subcarriers $N_i \in$Z and the power P_i, i = 1, 2, …, K are initialized. The feasible solution is determined and h is obtained as an upper bound for the optimization problem (9).

end if

By the branch-and-bound method, the optimization problem (9) is branched into the problem M_1 with $N_u \leq [N_u^*]$ and the problem M_2 with $N_u \geq [N_u^*] + 1$. $h_1,\ N_u^{(1)}, p_u^{(1)}$ and $h_2, N_u^{(2)}, p_u^{(2)}$ are obtained by combining Algorithm 2.

$if\ \exists h_i$, i=1,2, $\forall N_u \in$ Z then

h = min {h_i,h } or h = min {h_1, h_2,h}

$if\ h_i <$ h then

l = max {h_1, h_2,h}, $N_u^* = N_u^i$

end if

end if

s =h

end while

Output: p_u^*, N_u^* , s

4. Simulation and Result Analysis

The method in this paper is validated by simulation on matlab. The simulation parameters are shown in Table 1.

Table 1. Simulation parameters

Parameters name	value
System bandwidth	10MHz
maximum latency threshold D_{max}	0.008 s
System power P	40W
Minimum interruption probability ε	0.01
Iteration threshold i_{max}	500
Path loss factor β	2.5
Number of subcarriers N	20
Power spectral density of noise N_0	-174dBm/Hz
Minimum rate threshold σ_0	1，3，5Mbps

Through a series of experiments, the performance of the allocation algorithm to maximize multi-dimensional resource utilization proposed in this study is verified.

Figure.2. shows the curve of the relationship between average slicing rates and the number of users. The simulation results show that the average slicing rate achieved by all three algorithms tends to increase as the number of users increases. The greedy

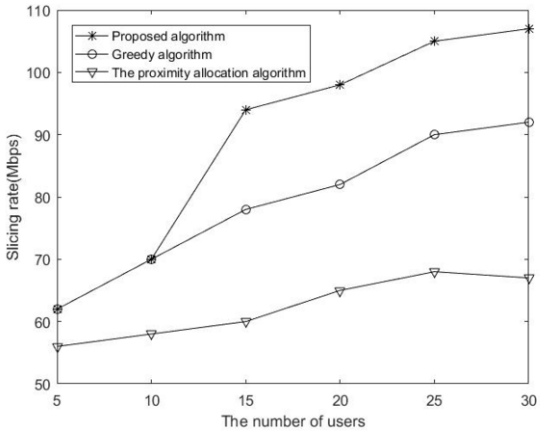

Figure 2. Relationship between average slicing rates and the number of users.

algorithm maintains essentially the same rate as the algorithm proposed in this paper when the number of users is small. However, as the total number of users continues to increase, the proposed algorithm takes into account the effect of caching resources and routing on the rate, and the rate tends to increase better than the greedy algorithm. The resource allocation algorithm proposed in this study can achieve a better system service rate. The performance is obviously better than other algorithms.

Figure.3. illustrates the end-to-end delay versus the number of users in the slice. The simulation results show that the algorithm proposed in this paper can maintain a smaller end-to-end delay compared to the greedy algorithm. The simulation results show that the proposed algorithm can reduce the end-to-end delay more effectively than the greedy algorithm and the proximity allocation algorithm. The greedy algorithm obtains the maximum number of underlying physics resources from the guaranteed slices, but does not consider the processing delay and link transmission delay of each node, which leads to an increase in end-to-end delay. The proximity allocation algorithm, which reduces the delay incurred in finding mapped nodes. However, the system model

considers the entire end-to-end network slice, so the proximity allocation algorithm cannot reduce the delay of the entire end-to-end network slice if the considerations are not comprehensive enough. As the total number of users increases, the average user end-to-end latency for each algorithm is essentially the same, due to the sufficient resources on the core and access network sides under the simulation conditions. Therefore, it can be seen that the algorithm proposed in this paper outperforms the greedy algorithm and the proximity allocation algorithm in optimizing the end-to-end delay.

Figure.4. shows the curve simulation diagram of the system throughput changing with S/N ratio. From the simulation results, it can be seen that the network throughput increases with the increase of S/N ratio, and the performance improvement of the proposed algorithm is more obvious in terms of system throughput. When the signal-to-noise ratio is greater than 3 dB, the proposed algorithm outperforms the greedy algorithm and the proximity allocation algorithm in terms of system throughput performance. When the signal-to-noise ratio exceeds 15 dB, the system throughput growth rate becomes slower. It means that the throughput of the system network has already started to saturate and the increase in signal power is no longer decisive.

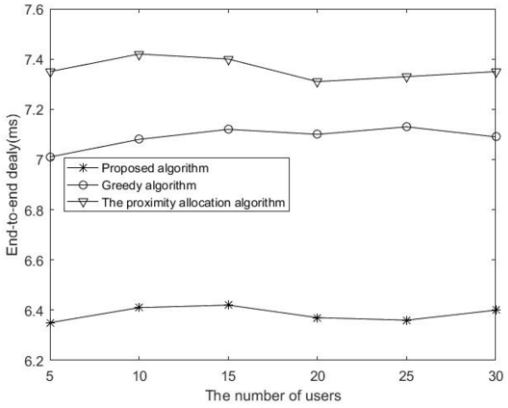

Figure 3. The change of average end-to-end delay with the number of users.

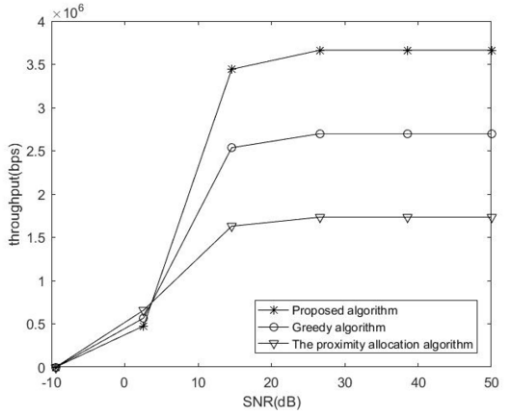

Figure 4. Relationship between system throughput and signal-to-noise ratio.

5. Conclusion

Through the study of systems sliced by eMBB and sliced by URLLC, a multidimensional resource allocation algorithm that unites communication, computation and storage resources is proposed to solve the multidimensional resource allocation problem of the system. A system model of multidimensional resources is developed, modelling the resource problem for multi-service slicing as a non-linear mixed-integer programming problem. A combination of the extended Lagrange algorithm and the branch-and-bound method is iterated until the resulting integer solution satisfies the constraints and the optimal resource allocation strategy is obtained. The algorithm is verified to improve the throughput of the system through experimental simulation comparisons. Optimizing the multi-dimensional resource allocation problem for multi-service networks improves the overall network performance. The next research direction is placed on continuing the search for other network factors in the system model, such as energy consumption, edge computing, etc. Added to the considerations of the model, the algorithm is further optimized to improve the performance of the algorithm.

References

[1] Ni J, Lin X, Shen X S. Efficient and secure service-oriented authentication supporting network slicing for 5G-enabled IoT[J]. IEEE Journal on Selected Areas in Communications, 2018, 36(3): 644-657.

[2] Alvizu R, Troia S, Maier G, et al. Network orchestration for dynamic network slicing for fixed and mobile vertical services[C]//Optical Fiber Communication Conference. Optica Publishing Group, 2018: Tu3D. 15.

[3] Ren J, Yu G, Cai Y, et al. Latency optimization for resource allocation in mobile-edge computation offloading[J]. IEEE Transactions on Wireless Communications, 2018, 17(8): 5506-5519.

[4] Zhang H, Guo F, Ji H, et al. Combinational auction-based service provider selection in mobile edge computing networks[J]. IEEE Access, 2017, 5: 13455-13464.

[5] Wang C, Liang C, Yu F R, et al. Computation offloading and resource allocation in wireless cellular networks with mobile edge computing[J]. IEEE Transactions on Wireless Communications, 2017, 16(8): 4924-4938.

[6] Guan W, Wen X, Wang L, et al. A service-oriented deployment policy of end-to-end network slicing based on complex network theory[J]. IEEE access, 2018, 6: 19691-19701.

[7] Li T, Zhu X, Liu X. An end-to-end network slicing algorithm based on deep Q-learning for 5G network[J]. IEEE Access, 2020, 8: 122229-122240.

[8] Fossati F, Moretti S, Perny P, et al. Multi-resource allocation for network slicing[J]. IEEE/ACM Transactions on Networking, 2020, 28(3): 1311-1324.

[9] Zhang P, Yao H, Liu Y. Virtual network embedding based on computing, network, and storage resource constraints[J]. IEEE Internet of Things Journal, 2017, 5(5): 3298-3304.

[10] Lin Y, Song H, Ke F, et al. Optimal caching scheme in D2D networks with multiple robot helpers[J]. Computer Communications, 2022, 181: 132-142.

[11] Miao Z, Wang Y, Han Z. A supplier-firm-buyer framework for computation and content resource assignment in wireless virtual networks[J]. IEEE Transactions on Wireless Communications, 2019, 18(8): 4116-4128.

[12] Mechtri M, Ghribi C, Soualah O, et al. NFV orchestration framework addressing SFC challenges[J]. IEEE Communications Magazine, 2017, 55(6): 16-23.

[13] Wang L, Lu Z, Wen X, et al. Joint optimization of service function chaining and resource allocation in network function virtualization[J]. IEEE Access, 2016, 4: 8084-8094.

[14] Kakkavas G, Tsitseklis K, Karyotis V, et al. A software defined radio cross-layer resource allocation approach for cognitive radio networks: From theory to practice[J]. IEEE Transactions on Cognitive Communications and Networking, 2020, 6(2): 740-755.

[15] Tran T D, Le L B. Resource allocation for multi-tenant network slicing: A multi-leader multi-follower stackelberg game approach[J]. IEEE Transactions on Vehicular Technology, 2020, 69(8): 8886-8899.

[16] Albonda H D R, Pérez-Romero J. An efficient RAN slicing strategy for a heterogeneous network with eMBB and V2X services[J]. IEEE access, 2019, 7: 44771-44782.

[17] Rafiq A, Mehmood A, Song W C. Intent-Based Slicing between Containers in SDN Overlay Network[J]. J. Commun., 2020, 15(3): 237-244.

[18] Xu C, Chen B, Qian H. Quality of Service Guaranteed Resource Management Dynamically in Software Defined Network[J]. J. Commun., 2015, 10(11): 843-850.

[19] Wang Y, Zhou J, Feng G, et al. Blockchain assisted federated learning for enabling network edge intelligence[J]. IEEE Network, 2022.

[20] Wang L, Zhou J, Wang Y, et al. Energy Conserved Computation Offloading for O-RAN based IoT systems[C]//ICC 2022-IEEE International Conference on Communications. IEEE, 2022: 4043-4048.