Journal of Autonomous Intelligence (2023) Volume 6 Issue 2
doi: 10.32629/jai.v6i2.652

ORIGINAL RESEARCH ARTICLE

Hybrid Chaos Particle Swarm Optimization algorithm for smart Cloud Service System based on optimization resource scheduling and allocation

Víctor P. Gil Jiménez[1,*], Abdulmajeed Al-Jumaily[1,2], A. Sali[2], Dhiya Al-Jumeily[3]

[1] *Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Leganés 28911, Madrid, Spain*

[2] *Department of Computer & Communication System Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia*

[3] *School of Computer Science & Mathematics, Liverpool John Moores University, Liverpool 565161, UK*

*** Corresponding author:** Víctor P. Gil Jiménez, vgil@ing.uc3m.es

ABSTRACT

To enhance the smart Cloud Service System for diverse user requirements in 5G and other service networks, this study leverages resource utilization and multi-tenancy network slicing operation costs. Specifically, we propose a multi-tenancy network resource allocation strategy based on the Chaos Particle Swarm Optimization (CPSO) algorithm. In a multi-tenancy network (MTN), we lease the wireless spectrum resources of the infrastructure provider's base station, construct access service slices as network slice services, and offer network access services to users. Introduce detailed formulation of the relationship between MTN and users, represented as a multi-master and multi-slave construct that defines the strategy space and profit function after MTN decision-making. Reverse induction is used to analyze the proposed model, and a distributed iterative algorithm is proposed to obtain the optimal throughput demand of users and the optimal slice cost of MTN. Simulation results demonstrate that the proposed strategy can effectively enhance resource utilization and user satisfaction while reducing energy consumption.

Keywords: multi-tenancy network; Particle Swarm Optimization; resource allocation; Cloud Service System

ARTICLE INFO

Received: 23 May 2023
Accepted: 19 June 2023
Available online: 31 July 2023

1. Introduction

With the increasing demand for cloud services and the advent of 5G networks, there is a growing need to develop efficient resource scheduling and allocation strategies to ensure optimal performance of smart Cloud Service Systems. Traditional network architectures are limited in their ability to cater to the diverse service demands of different vertical industries and users, leading to low resource utilization and significant service response delays[1,2]. To address these issues, the concept of network slicing was introduced in 5G networks, allowing for more effective infrastructure use and individualized service delivery without compromising service quality. The 5G network was to cater to the diverse service demands of various vertical industries and users. In traditional networks, a dedicated network can be designed for specific services and requirements, which is simple to deploy and has strong isolation. However, it suffers from issues such as low resource utilization and significant service response delays. To address these issues, the solution for 5G networks is network slicing, which enables more efficient infrastructure utilization and personalized service delivery without compromising service quality[1–3].

Furthermore, it provides dedicated networks for different requests, which can further optimize the waste of resources and network deployment costs caused by request differentiation and service diversification. In the paper of Han et al.[4], the successful deployment of flexible network slicing hinges on efficient slicing resource allocation and orchestration management is described. A well-designed resource allocation plan can significantly reduce network deployment costs, while simultaneously improving resource utilization and user satisfaction. To this end, many researchers have focused on optimizing resource allocation and plan deployment in network slicing. Typically, such plans begin by addressing two key areas: enhancing resource utilization and reducing network energy consumption. On the other hand, the paper[5] introduces slice as a service and develop a content distribution network as a service platform based on a multi-cloud domain, improving the deployment cost and service quality of content distribution network slice. Then, Dong et al.[6] investigate bandwidth resource allocation in wireless virtualized environments and achieves efficient resource scheduling through a greedy search, dynamically allocating bandwidth resources for users while optimizing system performance. Resource allocation in multi-service wireless networks jointly considers resource allocation between slices and resource scheduling within slices[4]. This scheme can balance allocation efficiency and service delay while ensuring slice isolation. Spectrum resource allocation in heterogeneous cellular networks optimizes the spectrum allocation and user access, increasing the network terminal rate[6]. The computational algorithms provide a systematic theoretical tool for analysing and predicting the strategic interaction behaviour of the group and are widely used to solve the problem of network slicing resource allocation and optimization. However, this work mainly focuses on optimizing the cost of network slicing and service delay but does not consider network energy consumption. An automated mechanism based on a genetic algorithm to assist tenants in making decisions is proposed[7] for the dynamic resource allocation problem of wireless access network slices. Halabian[8] modeled the resource allocation problem of end-to-end network slicing as a means to maximize system revenue using the ANT algorithm. While the mechanism mentioned can promote fairness in resource allocation, it may not adequately address the issue of competition between slices. However, Zheng et al.[9] proposed an optimization strategy for allocating 5G network cache resources that aimed to jointly optimize the revenue of network providers and network energy consumption. This approach led to improved cache hit rate and system energy consumption, although it did not consider cache resource allocation specifically. Finally, Raveendran et al.[10] tackled the resource allocation problem through efficient three-party matching between wireless spectrum resources, wireless network infrastructure, and mobile users using matching games. This approach reduced response delay and system cost, while enhancing user satisfaction. It does not consider the key indicators of service slicing in actual network deployments, such as network energy consumption and resource utilization. Realizing the reasonable allocation of spectrum resources in a multi-tenancy network through a parallel computational algorithm can improve the utilization of spectrum resources, reduce the cost of slice deployment, increase network revenue, and improve user satisfaction.

The aim of this paper is to present a multi-tenancy network resource allocation strategy that utilizes the Chaos Particle Swarm Optimization algorithm. In this approach, the infrastructure provider allocates the base station spectrum resources required to construct access service slices for MTN. The relationship between MTN and users is modelled using a multi-master and multi-slave model. By introducing slicing popularity and service hit rate indicators, a profit function is constructed for both parties involved in the game. Based on this model, the reverse induction method is applied. The CPSO algorithm is used to determine the user's optimal throughput demand, followed by a distributed iterative algorithm to determine the optimal access service slice cost of MTN. After two iterations, the system converges to the only Nash equilibrium point. The proposed strategy's effectiveness and feasibility are compared with other optimization strategies that consider the allocation of slicing resources to validate its performance.

2. Multi-tenancy network model

2.1. Wireless access network

The wireless access network structure of the multi-tenancy network is illustrated in **Figure 1**, comprising four modules: physical infrastructure layer, virtual network resource layer, network slice instance layer, and network service instance layer[11]. To address the reasonable allocation of spectrum resources in the access service slice communication resources, this paper proposes a game optimization strategy for multi-tenancy network resource allocation. This strategy is utilized in the network slice management and orchestration module to ensure the proper allocation of access service slice instance resources, which provide access service instances to users.

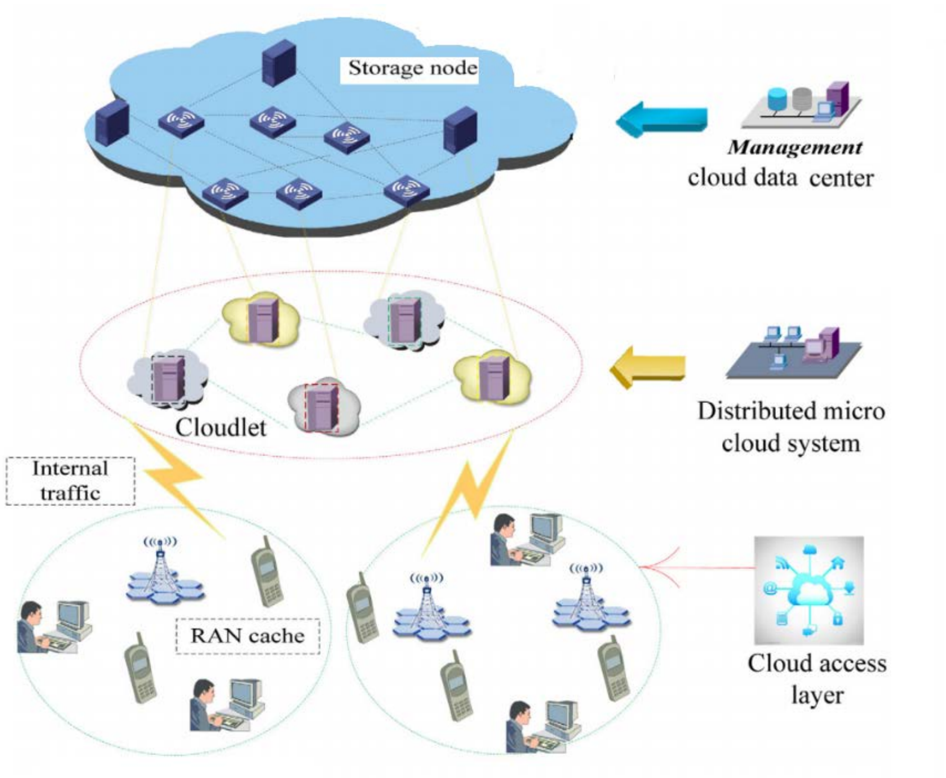

Figure 1. The structure of wireless access network.

The specific functions of each module are as follows:

1) The physical network infrastructure includes access facilities such as base stations, computing servers, and storage units, which different providers manage.

2) Virtual network resources include three major network resources: resource blocks integrated by physical network infrastructure through virtualization technology and provided to MTN.

3) MTN analyses service demand mapping, leases virtual network resources, and builds network-slicing instances.

4) The network service instance layer provides user and business services. Each service includes different use cases, and a service slice instance represents each use case.

Spectrum sharing relies on multiple access technology. The network model consists of an infrastructure provider, M, $MTN[1, 2, \ldots, M]$, and N users. Among them, the users include K internal access users $[1, 2, \ldots, K]$ and L external access user's $[1, 2, \ldots, L]$. All users can access all MTN.

2.2. The specific process of service model construction

Due to the limited physical resources in the network, MTN will dynamically adjust the slice pricing according to the user's needs and service capabilities to maximize their revenue and thus maximize the network revenue.

1) To provide an infrastructure of base station spectrum resources to use physical facilities reasonably and efficiently and increase their revenue, they lease virtual network resources to MTN.

2) MTN constructs access service slices according to user needs and slice's popularity to provide network access services.

3) The user purchases access service slices and access the network to obtain the required content.

2.3. The spectrum resource model

Virtual spectrum resources are leased from infrastructure providers to allocate a certain percentage of spectrum bandwidth MTNm, [1, 2, ..., M] ∊ m. The spectrum resource of the base station of the infrastructure provider is abstracted as the maximum bandwidth spectrum resource is B, the downlink transmission power is p, and it is assumed that B and p are constants. By assuming the access service slices are physically isolated, there is no interference between MTN. The network access throughput R_i that can be directly obtained is related to the subscribed slice type for internal users. It is related to the spectrum resources reallocated by MTNm in the access service as Equation (1).

$$R_i = x_{i,m} W_m \cdot log_2\left(1 + \frac{p.h_i^2}{N_0}\right) \tag{1}$$

where $x_{i,m}$ is the allocation Proportion; W_m is the maximum spectral bandwidth of MTNm; h_i^2 is the channel gain, and; N_0 is the noise power. Assume that the user requests a total of C types of access service slices $C = [S_1, S_2, \ldots, S_C]$. The relationship between the access service slice ranking and its access probability[12]. To determine the access probability D_c as Equation (2).

$$D_c = \frac{(c+\omega)^{-\upsilon}}{\sum_{i=1}^{C}(c+\omega)^{-\upsilon}}, \forall c \tag{2}$$

where ω is the flatness factor, and when $\omega = 0$, it satisfies the general power-law distribution; $\upsilon > 0$ is the skewness factor, also known as the popularity index. The larger υ, the higher the probability of the slice being accessed. The higher the slice popularity, the more popular S_1, and a small number of slice types have a higher request probability.

2.4. Service rate

Assuming that the user's request for network access service satisfies the Poisson distribution, c can be expressed as

$$H_{m,c} = 1 - e^{-\sum_{n\in N} \lambda_{n,c} T_m} \tag{3}$$

where λ is the average request rate, $\lambda_{nc} = \lambda \cdot D_c$ represents the user's request rate for c, and T_m represents the characteristic time of the MTN access service slice. According to the equivalent infinitesimal theorem, $e^{-x} \sim x (x \to 0)$, it can be approximated as

$$H_{m,c} = \sum_{n\in N} \lambda_{n,c} T_m \tag{4}$$

3. Proposed method

3.1. Resource allocation

The multi-master and multi-slave model has used the spectrum resource allocation problem of multi-tenancy network access service slicing. Furthermore, the essential elements include:

1) Set of decision-making individuals: The multi-tenancy network, MTN = $[1, 2, \ldots, M]$ lease the base station resources of the infrastructure provider. It has a first-hand advantage. Firstly, it determines the access service slice pricing and influences the decision-making of the subordinates. The subordinates include internal users $[1, 2, \ldots, K]$ and external users $[1, 2, \ldots, L]$. The users consider the slice pricing of MTN and their throughput requirements and adjust the slice ordering strategy to maximize Self-benefit.

2) MTN instantiates and resells slices of resources. Its strategy is to set the price of slices $P = [P1, P2, \ldots, Pm]$. Strategy set: the user's strategy is the service slice ratio $x_{i,m}$, $x_{j,m}$, ordered at each MTN, and its needs can be met by multiple MTNs.

3) Income function: the user's income function is represented by U_n, and the income function of MTNm is represented by U_m, where $\in [1, 2, \ldots, K, 1, 2, \ldots, L]$, $M \in [1, 2, \ldots, m]$.

Suppose there are two types of users in a multi-tenancy network, internal users and external users. This assumption is reasonable. For example, when accessing the school network through the access service slice, internal users, including students, faculty, and staff, can be directly accessed through the campus network, while external personnel needs to access indirectly through the virtual private network. The user's revenue U_n consists of two parts: the income of accessing service slices, expressed as the logarithmic form of access throughput R_n[13], and the other part is the expenditure of purchasing service slices. For internal access users, $i \in [1, 2, \ldots, K_i]$, Equation (5) represents R_i revenue related to the access throughput.

$$U_i(R_i) = ln\left(1 + R_i - b_i^0\right) \tag{5}$$

where R_i is the access throughput, b_i^0 represents the basic throughput provided by MTN to user i. Then the final benefit is when the user gets access to the service slice as Equation (6).

$$U_i^{QoS} = \delta_i \sum_{m=1}^{M} \sum_{c=1}^{C} \lambda H_{m,c} U_i(R_i) \tag{6}$$

where δ_i is the profit coefficient, and λ is the average request rate. The price of the slice should be in accordance with market regulations. That is, the slice price is proportional to the number of visits to MTNm. This relationship is represented by $\frac{n_m}{N_m}$, which is the number of users currently existing in MTNm divided by its maximum service capacity. The cost for users to purchase access service slices is shown in Equation (7), and the net income of internal access users.

$$U_i^{cost} = \eta_m \sum_{m=1}^{M} x_{i,m} \frac{n_m}{N_m} P_m \tag{7}$$

$$U_i = U_i^{QoS} - U_i^{cost} \tag{8}$$

where η_m is the cost coefficient, and $\frac{n_m}{N_m} P_m$ indicates that it is affected by market laws.

Equations (9) and (10) are the benefits of external access users related to access throughput and the cost of purchasing access service slices. The game modelling and solution of external access users and MTN are similar to internal access users. For external access users, $j \in [1, 2, \ldots, L]$, the difference from internal access users is that it is cost, and the slice price is quadratic. That is, the external access users are charged higher, and the basic throughput b_i^0 is not guaranteed.

Therefore, only internal user access is considered in the subsequent game-solving.

$$U_j(R_j) = ln(1 + R_j) \tag{9}$$

$$U_j^{cost} = \eta_m \sum_{m=1}^{M} x_{j,m} \left(\frac{n_m}{N_m} P_m\right)^2 \tag{10}$$

where t is the adjustment coefficient, which controls the change of the cost function of external user's, the benefits of MTN include the benefits of access service slice cost P_m and the cost of constructing slice instances E_m.

The revenue function of MTNm is shown in Equation (11).

$$U_m = \sum_{i=1}^{K} (P_m - E_m) x_{i,m} \eta_m \tag{11}$$

In the resource allocation system, users access the network by subscribing to access slices and obtaining the required services after paying the slice cost. At the same time, MTN obtains revenue by setting the resale price of resources and creating network slices. The resource allocation optimization problem for each MTN can be mapped to its profit maximization function, as.

$$maxU_m(p, x) \tag{12}$$

where the slice cost is a non-negative value, by substitution Equations (6) and (7) into Equation (8), the user's network revenue of internal access as the optimization objective function is also mapped to maximizing revenue for the two types of users as Equation (13).

$$U_i = \delta_i \sum_{m=1}^{M} \sum_{c=1}^{C} \lambda H_{m,c} \ln[1 + x_{i,m} W_m \cdot r_{i,m} - b_i^0] - \eta_m \sum_{m=1}^{M} x_{i,m} \frac{n_m}{N_m} P_m \tag{13}$$

where $0 \le x \le 1$ is the value of the proportion of slices ordered by the user. Moreover, the MTN user decision making as follows:

First, calculate the user profit function U_i is a concave function in its strategy space x as.

$$\frac{\partial U_m}{\partial x_{i,m}} = \frac{\delta_i \sum_{c=1}^{C} \lambda H_{m,c} W_m \cdot r_{i,m}}{1 + x_{i,m} W_m \cdot r_{i,m} - b_i^0} - \eta_m \frac{n_m}{N_m} P_m \tag{14}$$

Then calculate the user's profit function concerning the strategy space $x_{i,m}$ as.

$$\frac{\partial^2 U_j}{\partial^2 k_{i,j}} = \frac{-\delta_i \sum_{c=1}^{C} \lambda H_{m,c} \left(W_m \cdot r_{i,m}\right)^2}{\left(1 + x_{i,m} W_m \cdot r_{i,m} - b_i^0\right)^2} \tag{15}$$

The spectrum resources of the base station of the infrastructure provider are limited, and user's also have budget constraints, so the optimal user's throughput x^* is expressed as Equation (16).

$$x^* = \delta_i \sum_{m=1}^{M} \sum_{c=1}^{C} \lambda H_{m,c} U_i(B_i) - \eta_m \sum_{m=1}^{M} x_{i,m} \frac{n_m}{N_m} P_m \tag{16}$$

where the denoted B_i is the user's maximum spectrum bandwidth budget. After obtaining the user's optimal throughput requirement x^*, we use optimal access service slice pricing p^* of MTN concerning P_m as.

$$\frac{\partial U_m}{\partial P_m} = x^* \eta_m + \eta_m (P_m - E_m) \partial x^* / \partial P_m^t \tag{17}$$

The optimal slice cost of MTN needs to be solved iteratively as.

$$P_m^{t+1} = E_m - \frac{\lambda_k \partial x}{\partial x^* / \partial P_m^t} \tag{18}$$

P_m^t is the slice pricing of MTNm at the t iteration, and λ_k is the iteration step length, which is inversely proportional to the total number of iterations.

3.2. Chaos Particle Swarm Optimization algorithm

This algorithm makes full use of the group search characteristics of the PSO algorithm and the intelligent search characteristics of the particles to strengthen the internal search of the algorithm. The local search strategy improves the intelligence of the particles and global searchability. Assuming that the controllable resource allocation i is located at the point (m, n) of the lattice point expansion map, $A_{m,n} = (x_{A1}^t, x_{A2}^t, \dots, x_{AD}^t)$ represents the point (m, n). Where is the vector in the search space, which represents the output power of the controllable unit in a specific time, where $i = [1, 2, \dots, m]$; $n = [1, 2, \dots, j]$, the controllable unit with 8 neighbours of i in the lattice point expansion graph can be expressed as Equation (19).

$$\mathrm{N_{m,n}} = \{\mathrm{A_{m1,n1}}, \mathrm{A_{m1,n}}, \mathrm{A_{m1,n2}}, \mathrm{A_{m,n1}}, \mathrm{A_{m,n2}}, \mathrm{A_{m2,n1}}, \mathrm{A_{m2,n}}, \mathrm{A_{m2,n2}}\} \tag{19}$$

The optimal vector of the eight neighbours of vector $A_{m,n}$, as determined by Equation (20).

$$F(A_{m,n}) \le F(V_{m,n}) \tag{20}$$

Then the position of the controllable unit intelligent particles in the solution space will be retained. Otherwise, the position of the controllable unit intelligent particles will change according to the change of the service request, and the controllable unit in Equation (21) can be updated according to Li et al.[14].

$$x_{Ad}^{t+1} = x_{Nd}^t + (2r - 1)(x_{Nd}^t - x_{Ad}^t) \tag{21}$$

where $d = [1, 2, \dots, D, D]$ represents the dimension of the problem to be solved; r is a random number between [0, 1]. Each controllable unit in the algorithm updates its speed and position through the speed and position update formula of the PSO algorithm. It is assumed that the particles speed information can be expressed as $V_k^t = (v_{k1}^t, v_{k2}^t, \dots, v_{kD}^t)$, in the optimization process. each particle can record the optimal position $P_g = (p_{k1}^t, p_{k2}^t, \dots, p_{kD}^t)$, in the entire group. The optimal global position of the particles can be expressed as $P_g = [p_{g1}, p_{g2}, \dots, p_{gD}]$ to update the speed and position information[13–16], as in Equation (22).

$$\begin{gathered} V_{kd}^{t+1} = \omega V_{kd}^{t+1} + c_1 r_1 (P_{kd}^t - x_{kd}^t) + c_2 r_2 (P_{gd} - x_{kd}^t) \\ x_{kd}^{t+1} = x_{kd}^t + v_{kd}^{t+1} \end{gathered} \tag{22}$$

Assuming that the corresponding vector $P_g = [x_{p1}^t, x_{p2}^t, \dots, x_{pD}^t]$ of the optimal controllable unit particles position as Equation (23). Where $k = [1, 2, \dots, Pop_{max}, Pop_{max}]$ represent the total number of particle swarms; r_1 and r_2 are random numbers between [0, 1]; c_1 and c_2 are learning factors; ω is inertia Weight.

The chaotic local search algorithm improves the algorithm's search performance and optimization accuracy.

$$y_d^t = \frac{x_{pd}^t - x_{min}^t}{x_{max}^t - x_{min}^t} \tag{23}$$

where x_{pd}^t is the element in the vector P_g, x_{min}^t, x_{max}^t represents the minimum and maximum range of the solution space position of the optimal global particles. The elements in the vector are converted into chaotic variables between [0, 1], and the logistic equation $y_{n+1}^t, = \mu y_n^t, (1 - y_n^t)$ is used to obtain the value after n iterations. The chaotic sequence $y^t = (y_1^t, y_2^t, \dots, y_n^t,)$ where $n = 1, 2, \dots, T_{max}$. According to the following calculation expression, the chaotic sequence is inversely mapped back to the original solution space as.

$$x_{pd}^{t\prime} = x_{min}^{t} + \left(x_{min}^{t} - x_{max}^{t}\right)y_{d}^{t} \tag{24}$$

where $x_{pd}^{t\prime}$ is a global solution sequence of chaotic variables.

4. Experimental verification

A multi-tenancy network environment and access service are constructed and analysed using MATLAB tools. The performance of the proposed resource allocation optimization strategy is simulated and analysed to validate the strategy. The proposed strategy is compared with other algorithms in terms of maximum and minimum fairness strategy[16]. The priority-based dynamic resource allocation strategy[17]. Furthermore, delay minimization resource allocation strategy[18] regarding resource utilization, network revenue, and energy consumption reduction rate.

At the same time, the first maximum and minimum fairness algorithm meets the minimum needs of all users[14]. And then, the remaining resources are evenly distributed; the priority-based dynamic resource allocation algorithm considers the needs and priorities of users[19,20] and dynamically for resource allocation. The time delay minimization resource allocation algorithm takes the priority of the slice into account and minimizes the slice delay[16]. The network simulation environment includes 100 internal access users, the largest bandwidth spectrum resource in the network B = 500, and each MTN charges $W_m = 250$, and other parameters are shown in **Table 1**.

Table 1. Simulation parameters and values.

Simulation parameter	**Value**
Spectral efficiency $r_{i,m}$	70
Popularity factor α	0.6
Flatness factor β	1.0
Average request arrival rate λ	0.8
Slice feature time T_m	5
Iteration step λ_k	0.01

4.1. MTN-based resource scheduling verification

Figure 2 shows the relationship between network slicing providers MTN1 and MTN2 slice pricing. The points on the two curves represent the current MTN best response pricing strategy to another MTN. It can be seen that the cost equilibrium point is at (2.09, 2.32) for both cases. At this point, MTN1 and MTN2 are not interested in changing their pricing strategies to reduce their revenue, which is the best pricing combination that maximizes the revenue of both parties.

Figure 3 shows the relationship between MTN revenue and its slice popularity. It can be seen that the slice popularity is related to the popularity index, υ and the flatness factor ω. It can be seen from the figure that when υ is constant, the more oversised ω, the more popular the slice, the greater the access probability, and the higher the corresponding MTN revenue. When ω is constant, the minor ω is, the higher the access probability of famous slices is, and the higher the profit of the corresponding MTN is. When υ and ω are constant, the MTN revenue value gradually increases as the number of iterations increases. When the number of iterations is about 70, the MTN revenue converges to the maximum value, and the larger υ, the minor ω, the faster the convergence.

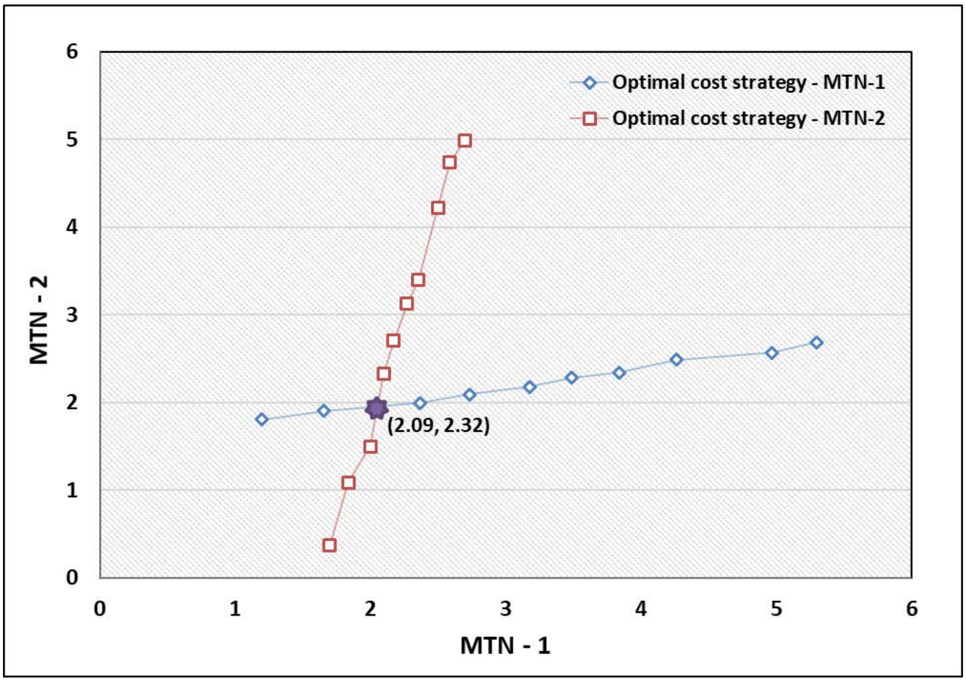

Figure 2. The slice pricing relationship between MTN1 and MTN2.

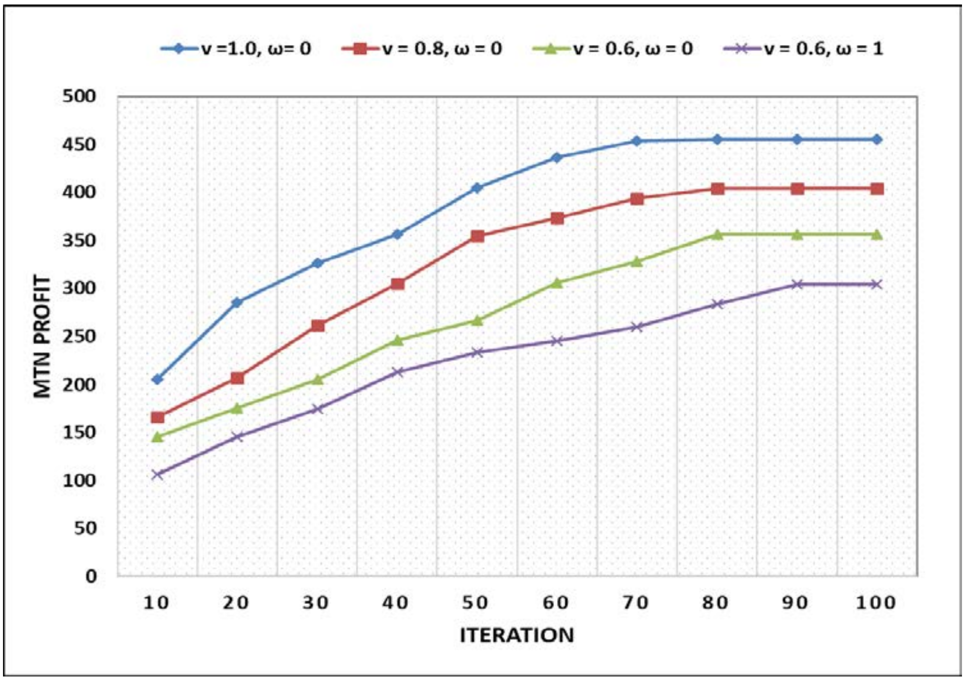

Figure 3. The relationship between MTN revenue and the number of iterations.

Figure 4 shows the relationship between the revenue of access service users and the number of iterations under different strategies. It can be seen from the figure that when the number of iterations is the same, the game-based resource allocation strategy proposed in this paper has the most significant revenue value for access service users. As the number of iterations increases, the revenue value of access service users gradually increases. The priority-based dynamic allocation strategy only considers maximum slicing service quality and fails effectively guarantee user benefits. In contrast, the time delay minimization resource allocation strategy considers the slicing service delay and sets the slicing preference, and the user benefits are relatively higher.

The resource allocation strategy speeds up the convergence through the CPSO algorithm. When the number of iterations is about 70, the convergence speed is second only to the maximum and minimum fairness strategy.

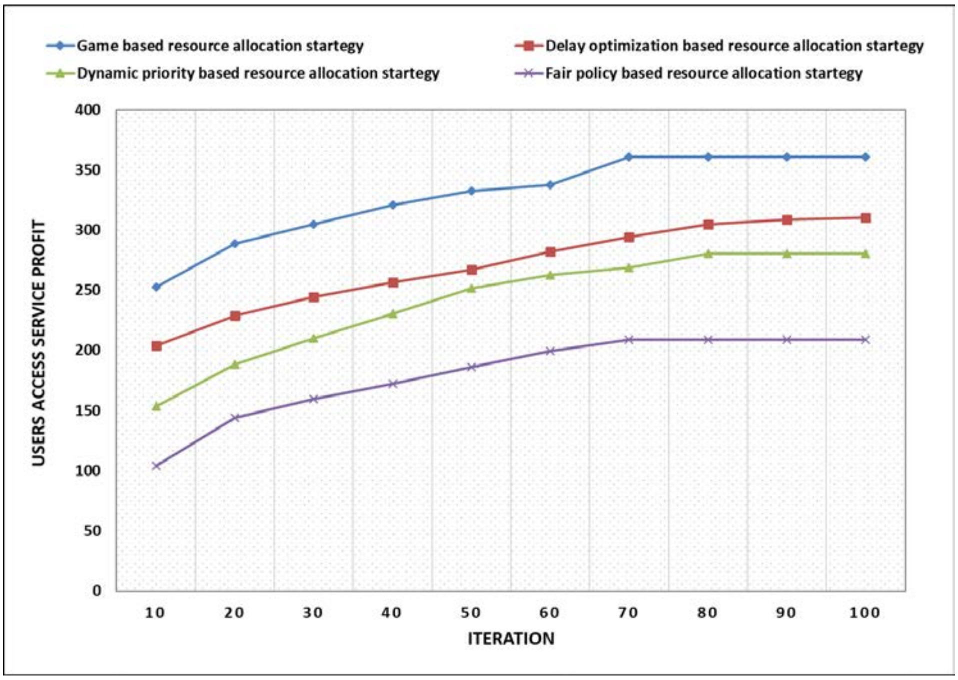

Figure 4. The relationship between user revenue and the number of iterations.

Figure 5 shows the network's spectrum bandwidth utilization rate and the user access service slice request arrival rate under different strategies. The initial conditions for user access are the same, and the resource utilization rates of the four strategies are the same. It can be seen that the arrival rate of user slice requests increases and the utilization rate gradually increases. The proposed resource allocation strategy can converge to the maximum resource utilization rate. The priority-based dynamic resource allocation strategy dynamically allocates resources to the slice by maintaining the priority and demand of the slice. Compared with the time delay minimization resource allocation strategy, only the on-demand allocation strategy of computing and communication resources is considered, achieving greater resource utilization. Based on the game-based resource allocation strategy, the revenue value of access service users reached a convergence of about 70 iterations. The maximum and minimum fairness algorithms converged the fastest, but the resource utilization was the lowest.

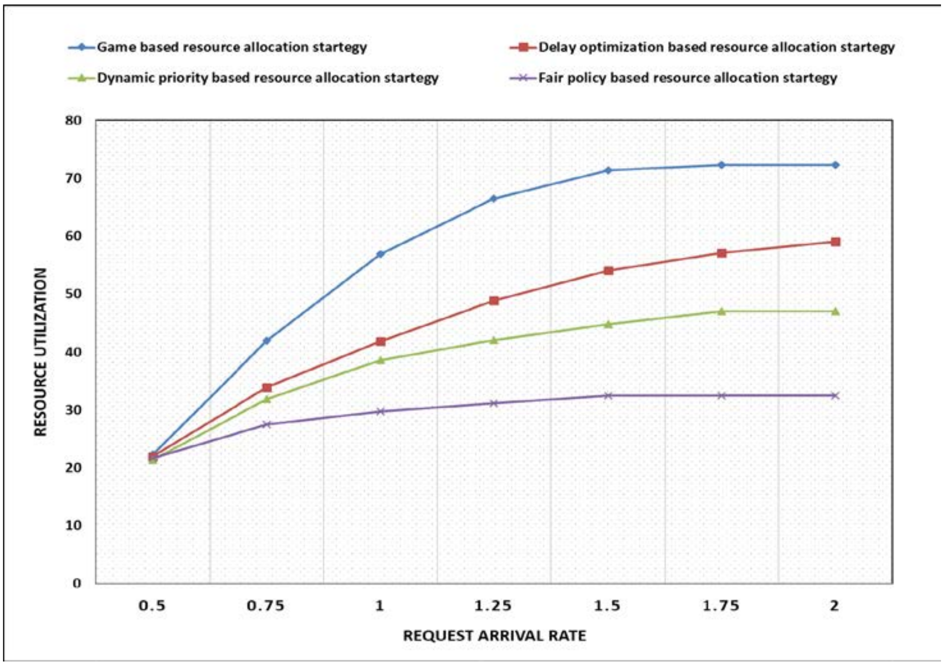

Figure 5. The relationship between resource utilization and request arrival rate.

Figure 6 shows the relationship between the system's network energy consumption reduction rate and the user request rate under different strategies. For comparison purposes, it is assumed that the user request patterns under different strategies are the same. They are all positioned on distributions to reflect the energy under different consumption situations. It can be seen from the figure that as the request arrival rate increases, the network energy consumption reduction rate continues to increase. The optimal resource allocation strategy proposed can achieve the highest network energy consumption reduction rate. The resource allocation strategy realizes a more reasonable allocation of spectrum resources in a multi-tenancy network through the CPSO algorithm between users and MTN. It can better respond to green communications to achieve a high network energy consumption reduction rate. Compared with the other three strategies, the cost of slice deployment is lower, and the utilization rate of resources is extremely high.

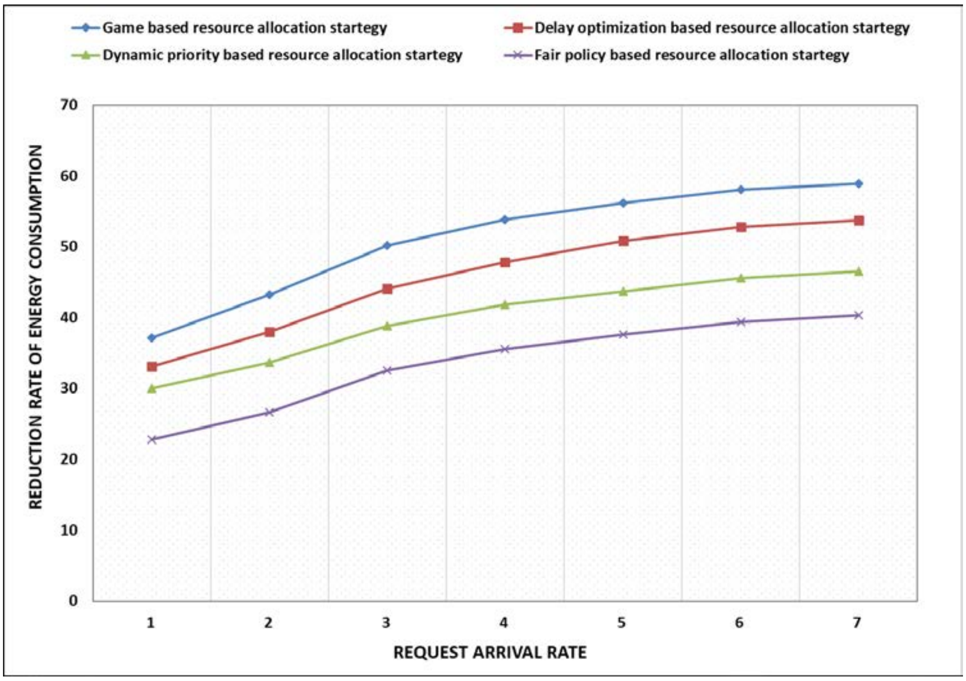

Figure 6. The relationship between network energy consumption reduction rate and request arrival rate.

5. Conclusion

In this paper, proposed a multi-tenancy network resource allocation strategy based on a Hybrid Chaos Particle Swarm Optimization algorithm for smart Cloud Service Systems. The main objective of this strategy is to solve the problem of reasonable allocation of spectrum resources in the access service slice communication resources of MTN. To achieve this objective, we first modeled the relationship between MTN and users as a multi-master and multi-slave model, and introduced slicing popularity and service hit rate indicators to construct the profit function of both parties in the game. We then applied the reverse induction method to solve the user's optimal throughput demand and used a distributed iterative algorithm to solve the optimal access service slice cost of MTN. The simulation results showed that the proposed strategy effectively improved resource utilization and user satisfaction while reducing energy consumption. The proposed strategy has significant implications for the deployment and implementation of 5G network slicing. To optimizing the allocation of network resources, it can achieve individualized service delivery without compromising service quality, ultimately benefiting both the infrastructure provider and end-users. In future work, to extend our approach to consider the resource allocation optimization plan of end-to-end network slicing and reinforcement learning to realize the active deployment. Overall, the Hybrid Chaos Particle Swarm Optimization algorithm has great potential for improving the efficiency and effectiveness of smart Cloud Service Systems in the 5G and other service networks.

Author contributions

Conceptualization, VPGJ and AAJ; methodology, AAJ and VPGJ; software, AAJ; validation, VPGJ and AAJ; formal analysis, AAJ; investigation, VPGJ; resources, AAJ; data curation, AS and DAJ; writing—original draft preparation, VPGJ and AAJ; writing—review and editing, AS and DAJ; visualization, AAJ; supervision, DAJ; project administration, VPGJ; funding acquisition, VPGJ.

Conflict of interest

The authors declare no conflict of interest.

Funding

This research was partly funded by Project "IRENE" (PID2020-115323RB-C33) (MINECO/AEI/FEDER, UE).

References

1. Afolabi I, Taleb T, Samdanis K, et al. Network slicing and softwarization: A survey on principles, enabling technologies, and solutions. *IEEE Communications Surveys & Tutorials* 2018; 20(3): 2429–2453. doi: 10.1109/COMST.2018.2815638
2. Wang CJ, Wang HL, Chen JM, Jiao H. Research on power network slicing technology based on cloud-edge collaboration. In: Proceedings of 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA); 22–24 January 2021; Shenyang, China. pp. 743–753.
3. Al-Jumaily A, Sali A, Jimenez V, et al. Evaluation of 5G coexistence and interference signals in the c-band satellite earth station. *IEEE Transactions on Vehicular Technology* 2022; 71(6): 6189–6200. doi: 10.1109/TVT.2022.3158344
4. Han Y, Tao X, Zhang X, et al. Hierarchical resource allocation in multi-service wireless networks with wireless network virtualization. *IEEE Transactions on Vehicular Technology* 2020; 69(10): 11811–11827. doi: 10.1109/TVT.2020.3019217
5. Bensalem I, Taleb T, Bagua M, et al. Optimal VNFs placement in CDN slicing over multi-cloud environment. *IEEE Journal on Selected Areas in Communications* 2018; 36(3): 616–627. doi: 10.1109/JSAC.2018.2815441
6. Dong X, Cheng L, Zheng G, et al. Network access and spectrum allocation in next-generation multi-heterogeneous networks. *International Journal of Distributed Sensor Networks* 2019; 15(8): 1–14. doi: 10.1177/15501477198661
7. Lieto A, Malanchini I, Mandelli S, et al. Strategic network slicing management in radio access networks. *IEEE Transactions on Mobile Computing* 2022; 21(4): 1434–1448. doi: 10.1109/TMC.2020.3025027
8. Halabian H. Distributed resource allocation optimization in 5G virtualized networks. *IEEE Journal on Selected Areas in Communications* 2019; 37(3): 627–642. doi: 10.1109/JSAC.2019.2894305
9. Zheng Z, Song L, Han Z, et al. A stackelberg game approach to proactive caching in large-scale mobile edge networks. *IEEE Transactions on Wireless Communications* 2018; 17(8): 5198–5211. doi: 10.1109/TWC.2018.2839111
10. Raveendran N, Gu N, Jiang C, et al. Cyclic three-sided matching game inspired wireless network virtualization. *IEEE Transactions on Mobile Computing* 2019; 20(2): 416–428. doi: 10.1109/TMC.2019.2947522
11. NGMN Alliance. *Description of Network Slicing Concept*. Next Generation Mobile Networks Limited Company; 2016.
12. Li A, Wang X, Li K, et al. Collaborative multi-tier caching in heterogeneous networks: Modeling, analysis, and design. *IEEE Transactions on Wireless Communications* 2017; 16(10): 6926–6939. doi: 10.1109/TWC.2017.2734646
13. Ho TM, Tran NH, Le LB, et al. Network virtualization with energy efficiency optimization for wireless heterogeneous networks. *IEEE Transactions on Mobile Computing* 2019; 18(10): 2386–2400. doi: 10.1109/TMC.2018.2872519
14. Li J, Chen H, Chen Y, et al. Pricing and resource allocation via game theory for a small-cell video caching system. *IEEE Journal on Selected Areas in Communications* 2016; 34(8): 2155–2129. doi: 10.1109/JSAC.2016.2577278
15. Zou J, Li C, Zhai C, et al. Joint pricing and cache placement for video caching: A game theoretic approach. *IEEE Journal on Selected Areas in Communications* 2019; 37(7): 1566–1583. doi: 10.1109/JSAC.2019.2916279
16. Shi L, Ye YH, Lu G. User-centric energy efficiency fairness in backscatter-assisted wireless powered communication network. *Journal on Communications* 2020; 41(7): 84–94. doi: 10.11959/j.issn.1000-436x.2020133

17. Ko H, Lee J, Pack S. Priority-based dynamic resource allocation scheme in network slicing. In: Proceedings of 2021 International Conference on Information Networking (ICOIN); 13–16 January 2021; Jeju Island, Korea (South). pp. 62–64.
18. Zarandi S, Tabassum H. Delay minimization in sliced multi-cell mobile edge computing (mec) systems. *IEEE Communications Letters* 2021; 25(6): 1964–1968. doi: 10.1109/LCOMM.2021.3051558
19. Dorigo M, Blum C. Ant colony optimization theory: A survey. *Theoretical Computer Science* 2005; 344(2–3): 243–278. doi: 10.1016/j.tcs.2005.05.020
20. Pawar CS, Wagh RB. Priority based dynamic resource allocation in cloud computing with modified waiting queue. In: Proceedings of 2013 International Conference on Intelligent Systems and Signal Processing (ISSP); 1–2 March 2013; Vallabh Vidyanagar, India. pp. 311–316.