Joint Optimization of Satellite Beam Hopping Scheduling and Time-Frequency-Space Resource Allocation under Dynamic Service Demands

Jiang Tian
School of Communications and Information Engineering
Chongqing University of Posts and Telecommunications
Chongqing, China
s210101125@stu.cqupt.edu.cn

Jiangtao Luo*
School of Communications and Information Engineering
Chongqing University of Posts and Telecommunications
Chongqing, China
Luojt@cqupt.edu.cn

Yongyi Ran
School of Communications and Information Engineering
Chongqing University of Posts and Telecommunications
Chongqing, China
ranyy@cqupt.edu.cn

Ruili Zhao
School of Communications and Information Engineering
Chongqing University of Posts and Telecommunications
Chongqing, China
D200101029@stu.cqupt.edu.cn

2024 IEEE/CIC International Conference on Communications in China (ICCC) | 979-8-3503-7841-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCC62479.2024.10681784

***Abstract*—In multibeam satellite (MBS) systems, a joint optimization algorithm is proposed to tackle the challenges of beam hopping and time-frequency-space resource allocation. The method aims to reduce the supply-demand gap between ground service communication needs and the capacity provided by the beams, while ensuring the quality of ground service communication. This enhances the efficiency of satellite communication systems. Firstly, from the perspective of time-frequency-space interference isolation, multidimensional resources are designed. Bandwidth resources are partitioned and power resources are allocated based on demand. The multidimensional resource allocation problem is formalized as a multi-objective optimization problem to maximize system throughput and minimize the supply-demand gap. Secondly, by representing the state of the MBS as a multidimensional matrix and considering dynamic service demands, the target problem is modeled as a Markov decision process. Finally, a deep reinforcement learning algorithm is employed to solve the joint optimization problem of satellite beam hopping scheduling and multidimensional resource allocation. Simulation results show that the algorithm improves system average throughput by approximately 3–59% and reduces the supply-demand gap by about 49–66%.**

***Index Terms*—Multibeam satellite, beam hopping, time-frequency-space resources, supply-demand gap**

I. Introduction

In recent years, the resource allocation optimization problem of multibeam satellite systems (MBS) has become a major research hotspot in the field of space communication technology due to the rapid growth in the demand for satellite communication services and the scarcity of satellite communication resources. The power of each beam are coupled with each other, and the bandwidth power allocation is constrained by the total resources of the satellite payload, so the joint bandwidth and power allocation problem of MBS is a key issue to be considered [1], [2]. Therefore, it is particularly important to target time-frequency-space resource allocation to meet the operational requirements and improve resource utilization.

Research on the optimization of resource allocation in MBS can currently be divided into two directions: single-objective optimization and multi-objective optimization. For single-objective optimization, Han et al. in [3] proposed a beam hopping resource allocation algorithm based on deep reinforcement learning with the objective of minimizing transmission delay in MBS, but without considering optimal bandwidth allocation. Wang et al. in [4] proposed an offline beam hopping resource allocation algorithm based on maximizing the total system throughput, which requires predicting future channel information, making it difficult to implement in practice. Guo et al. in [5] proposed three user selection schemes and an accumulated delay-aware power allocation algorithm with the objective of minimizing the maximum queueing delay between users. Zhang et al. in [6] studied resource allocation in low Earth orbit (LEO) satellite systems based on beam hopping with the objective of maximizing cell throughput. It decomposed the resource allocation problem into time slot allocation sub-problems and transmit power optimization (TPO) sub-problems, solving the TPO sub-problem using convex optimization theory. Xu et al. in [7] proposed a deep reinforcement learning-based algorithm to maximize system throughput, which flexibly utilizes three degrees of freedom: time, space, and beam coverage radius, but does not fully utilize frequency resources. For multi-objective optimization, Lin et al. in [8] proposed a joint beam hopping mode and bandwidth allocation scheme to increase data throughput while

*The corresponding author is Jiangtao Luo.

*This work was supported by the National Natural Science Foundation of China (No.U23A20275, 62171072, 62172064, 62003067), and the Natural Science Foundation of Chongqing (cstc2021jcyj-msxmX0586).

reducing latency. However, this approach does not involve power, which is a critical resource in MBS.

The above studies have not fully considered frequency allocation, and the existing algorithms ignore long-term benefits and are weakly adaptable. In order to improve the performance of MBS, it is necessary to fully optimize beam hopping, power and bandwidth. However, this increases the search space and leads to “dimensional disaster”. Therefore, resource optimization algorithms with high adaptability and learning capability are urgently needed to improve throughput and reduce the supply-demand gap.

To address the above issues, this paper proposes a deep reinforcement learning-based joint optimization algorithm for beam hopping and resource allocation (D-BHRA), which aims to rationally allocate power and bandwidth resources of satellite beams to maximize throughput and reduce the supply-demand gap, while achieving spatial interference isolation.

The contributions of this paper are summarized as follows:

- We propose an allocation scheme for multidimensional resources considering time, frequency, and space aspects, leading to improvements in energy efficiency and isolation of spatial interference.
- We incorporate the supply-demand gap into the model for the first time, which improves the energy efficiency and fairness.
- The experimental results illustrates that compared with the baseline algorithm, the D-BHRA significantly improves throughput and reduces the supply-demand gap by about 8.77% and 66.16%, respectively, when compared with greedy-based beam hopping and resource allocation.

II. System Model and Problem Formulation

A. System Model

This paper addresses a satellite operating in the K_a band and orbiting in geostationary Earth orbit (GEO) [9], with the system model shown in Fig. 1. It is assumed that the satellite carries K controllable multibeam antennas and one fixed signal antenna. For the forward link, the multibeam satellite system provides K beams to cover N cells, with the satellite carrying a buffer to record the communication demands of each cell. The K beams serve the N cells in a time-division multiplexing manner, where the multibeam antennas can adjust their pointing and transmission power in real-time, and the fixed signal antenna is used for receiving ground communication demands and channel state information. The system link adopts a Gaussian white noise channel.

The satellite beam refers to the signal transmitted by the satellite's onboard transmitter, denoted as $\mathcal{K} = \{k|k = 1,2,...,K\}$. This paper divides the entire area covered by the satellite into multiple equally sized cells, denoted as $\mathcal{N} = \{n|n = 1,2,...,N\}$, satisfying $K \ll N$. The number of data packets requested by cell n at time t is ϕ_t^n, and the total number of data packets requested by all cells at time t can be represented as $\Lambda_t = [\phi_t^n|n = 1,2,..,N]$. The vector of data packets awaiting service in the buffer for cell n at time

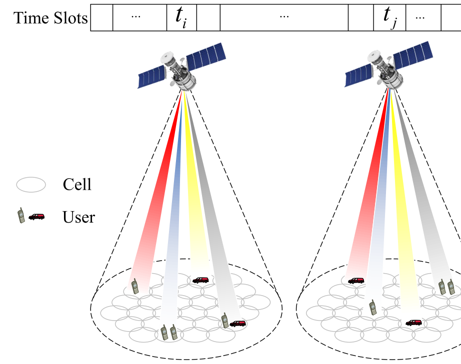

Fig. 1. Based on beam hopping satellite communication architecture.

t is represented as $\boldsymbol{E}_{t,n}^T = [\phi_{t,l}^n|l = 0,1,...,l_{th+1}]$, and the matrix of data packets cached in the buffer is represented as $\boldsymbol{\Phi}_t = [\boldsymbol{E}_{t,1}, \boldsymbol{E}_{t,2}, ..., \boldsymbol{E}_{t,N}]$. Here, l is the queueing delay of the data packet, and l_{th} is the maximum allowable data packet queueing delay. If the data packet queueing delay exceeds the maximum allowable delay, i.e., $l = l_{th+1}$, the buffer will discard the unsent data request. The number of data packets awaiting service in cell n at time t is represented as λ_t^n, $\lambda_t^n = \sum_{l=0}^{l=l_{th}} \phi_{t,l}$.

The total bandwidth of the satellite is denoted as B_{tot}, which is divided into M frequency blocks. Therefore, the bandwidth of each frequency block is $B_{block} = B_{tot}/M$. In this paper, it is assumed that each beam can only be assigned contiguous frequency blocks. Based on the number and location of occupied contiguous blocks, there are a total of $\frac{M(M+1)}{2}$ available bandwidth allocation schemes. When bandwidth blocks assigned to different beams overlap, this results in co-channel interference in the overlapping frequency bands. Referring to [8], the overlap factor $\alpha_t^{i,j} = \frac{B_t^i \cap B_t^j}{|B_t^i|}$ is defined to represent the co-channel interference from beam i to beam j within time slot t, where B_t^i denotes the frequency block used by beam i in time slot t.

P_{tot} represents the maximum available transmission power of the satellite system, and the maximum transmission power that can be carried by a single beam is P_{max}.

According to ITU-R S.672-4 standard, the downlink path loss $\boldsymbol{H} = \{h^{n,k}|n \in \mathcal{N}, k \in \mathcal{K}\}$from the satellite transmitter to the user receiver can be calculated by [10]:

$$\boldsymbol{H} = \boldsymbol{\Theta} \cdot \boldsymbol{G_U} \cdot \boldsymbol{G_B} \tag{1}$$

where $\boldsymbol{\Theta} = \text{diag}\{\sigma_1, \sigma_2, ..., \sigma_N\}$ represents the channel gain matrix; $\boldsymbol{G_B} = \{g_{n,k}^b|n \in \mathcal{N}, k \in \mathcal{K}\}$ represents the transmit antenna gain matrix from beams to cells; $\boldsymbol{G_U} = \text{diag}\{g_1^u, g_2^u, ..., g_N^u\}$ represents the receive antenna gain matrix for the corresponding N cells. Therefore, when beam k serves cell n, the signal-to-noise ratio (SNR) of that cell is given by:

$$\Gamma_t^{n,k} = \frac{h^{n,k} \cdot P_t^k}{|B_t^k| B_{block} N_0 + \sum_{i \in \mathcal{K}, i \neq n} h^{n,i} \cdot \alpha_t^{i,j} \cdot P_t^i} \tag{2}$$

where $h^{n,k} \in \boldsymbol{H}$ denotes the loss from beam k to cell n; P_t^k is the transmit power of beam k at moment t; B_t^k denotes the bandwidth allocation scheme used for beam k at moment

t; and N_0 is the power spectral density of the noise. The channel capacity can be obtained from the DVB-S2 standard as follows:

$$C_t^{n,k} = x_t^{n,k} \cdot B_t^k \cdot f_{DVB-S2X}(\Gamma_t^{n,k}) \tag{3}$$

where $x_t^{n,k}$ denotes whether beam k covers cell n; $f_{DVB-S2X}$ is the performance mapping function [11]. The channel capacity of cell n at time t is $C_t^n = \sum_{k=1}^{K} C_t^{n,k}$. From the above, the actual amount of data transmitted by cell n in time slot t is $\Pi_t^n = \min\{C_t^n, \lambda_t^n\}$.

B. Problem Formulation

In this paper, we aim to dynamically select the beam hopping and power bandwidth allocation strategies to meet the dynamic service demands of the ground, while maximizing the system throughput and reducing the supply-demand gap.

In order to measure the degree of matching between the resources supplied by the satellite and the demand of the cell, this paper defines the supply-demand gap $\delta_t^{n,k}$ as follows:

$$\delta_t^{n,k} = |C_t^{n,k} - D_t^{n,k}| \tag{4}$$

where $C_t^{n,k}$ denotes the communication capacity allocated to cell n by beam k at moment t, and $D_t^{n,k}$ denotes the communication capacity demanded by cell n when it is illuminated by satellite beam k at moment t. The smaller supply-demand gap is, the better the matching between the communication demand of the ground cell and the transmission capacity of the beam is.

In order to maximize the data throughput while reducing the supply-demand gap, this paper defines system utility P is:

$$P = \omega \frac{\sum_{n=1}^{N} \Pi_t^n}{\Pi_{t,\max}} - (1-\omega) \frac{\sum_{n=1}^{N}\sum_{k=1}^{K} \delta_t^{n,k}}{\delta_{t,\max}}, \omega \in [0,1] \tag{5}$$

where ω denotes a predetermined weight to achieve a trade-off between throughput and supply-demand gap; $\Pi_{t,\max}$ denotes the maximum value of throughput among all cells at decision time slot t; and $\delta_{t,\max}$ denotes the maximum value of supply-demand gap between satellites and cells at decision time slot t.

Based on the optimization objective in Eq. (5), this paper establishes a dynamic beam-hopping multi-objective optimization problem, which can be modeled as:

$$\begin{aligned} &\max\ P \\ &s.t.\ \ \mathrm{C1}: \textstyle\sum_{k=1}^{K} P_k \le P_{tot} \\ &\qquad \mathrm{C2}: P_k \le P_{\max} \\ &\qquad \mathrm{C3}: |B_t^k| B_{block} \le B_{tot} \\ &\qquad \mathrm{C4}: B_t^k = 1, 2, ..., \tfrac{M(M+1)}{2}, \forall k \in K \end{aligned} \tag{6}$$

where C1 means that the sum of the assigned powers of the individual beams is less than the total transmit power; C2 means that the assigned powers of the individual beams must not exceed the maximum transmit power of the individual beams; C3 implies that the sum of the bandwidths assigned to all the beams is less than the total bandwidth; and C4 implies that the bandwidths of the beams are selected for each of the beams from $M(M+1)/2$ bandwidth assignment schemes.

III. The Proposed D-BHRA Algorithm

The framework of the joint optimization algorithm for satellite beam hopping and power bandwidth allocation based on D-BHRA is shown in Fig. 2. The action space of this algorithm consists of two parts, which on one hand is responsible for specifying the cell that will be covered by the beam, and on the other hand allocating the transmission power and bandwidth for the beam. When making a decision, the satellite will select the action with the highest probability among them for execution based on the probability values of a set of different actions. After executing an action, the satellite jumps from the current state to the next state and receives a reward value reflecting the effect of the action. This section describes how to optimize this decision-making strategy using the MDP model and the D-BHRA algorithm.

A. MDP Model

The communication capacity demand requested by the same cell follows the first-arrival-first-transmission model, so the packet matrix $\boldsymbol{\Phi}_t$ recorded in the satellite buffer at moment t is affected by $\boldsymbol{\Phi}_{t-1}$ at moment $t-1$, and also the packet matrix $\boldsymbol{\Phi}_t$ is affected by the action a_{t-1} of the beam at moment $t-1$ and by the packet arrivals Λ_t. Therefore, the packet matrix can be constructed as:

$$\boldsymbol{\Phi}_t = \boldsymbol{\Phi}_{t-1} + \Lambda_t - a_{t-1} \cdot C_{t-1} \tag{7}$$

Define state $s_i \in \boldsymbol{S}$, action $a_i \in \boldsymbol{A}$, transfer probability $p(s_i|s_i, a_i) \in \boldsymbol{P}$ and reward r_i. The details are shown below:

State Space: The state space is composed of two aspects. On the one hand, the matrix of packets to be served recorded in the satellite cache is used as a constituent of the state space, the matrix of packets being packets requested by individual cells over a period of time and not served at moment t. On the other hand, the demand of individual cells is also used as a constituent of the state space. The state space s_t is:

$$s_t = [\boldsymbol{\Phi}_t, \boldsymbol{D}_t] \tag{8}$$

where $\boldsymbol{\Phi}_t$ denotes the number of unserved packets cached in the queue at the moment t; $\boldsymbol{D}_t$ denotes the communication capacity demanded by each cell at the moment t.

Action Space: The action space consists of three aspects. On the one hand is the coverage strategy of the satellite beam hopping, which itself is the selection of the cells to be covered by the beams, i.e., $x_t^{n,k} \in (x_t^{1,k}, x_t^{2,k}, ..., x_t^{N,k})$. If beam k covers cell n at moment t then $x_t^{n,k} = 1$. On the one hand is the power allocation strategy, which allocates the transmission power of the satellite assigned to the beams according to the cell demand, i.e., $P_k \in (\ell_1, \ell_2, ..., \ell_N)$. On the other hand the bandwidth allocation strategy, which divides the bandwidth resource into M blocks of frequencies, i.e., $B_k = 1, 2, ..., \frac{M(M+1)}{2}$. The action space a_t is:

$$a_t = \{\boldsymbol{X}_t, \boldsymbol{P}_t, \boldsymbol{B}_t\} \tag{9}$$

where $\boldsymbol{X}_t$ denotes the coverage strategy of the satellite beam hopping at moment t; $\boldsymbol{P}_t$ denotes the transmission power strategy assigned to the beam by the satellite at moment t; and $\boldsymbol{B}_t$ denotes the bandwidth strategy assigned to the beam by the satellite at moment t.

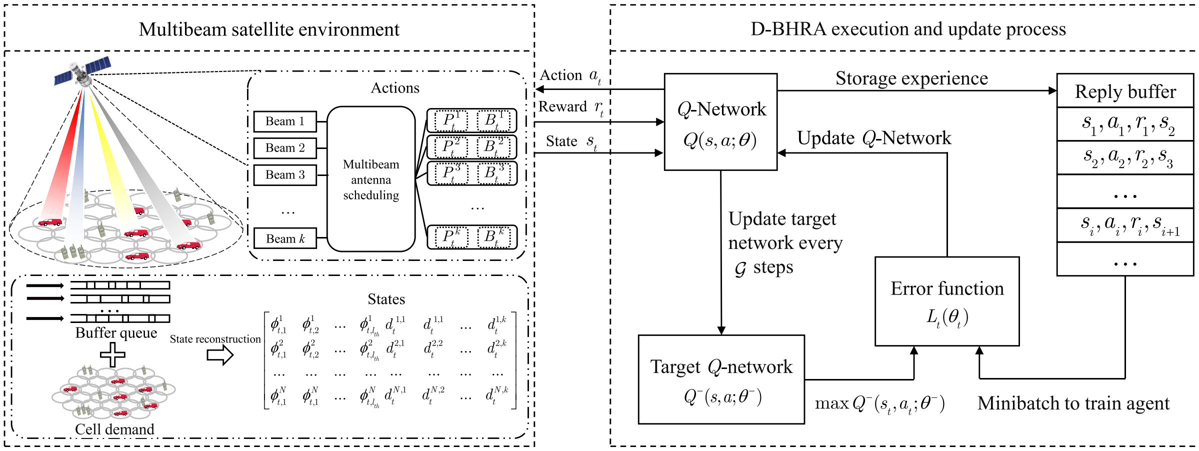

Fig. 2. Framework of the D-BHRA.

Reward function: In D-BHRA, the optimization objective is to maximize the system throughput and minimize the supply-demand gap. Therefore, the reward function is specified as:

$$r_t = \omega \frac{\sum_{n=1}^{N} \Pi_t^n}{\Pi_{t,\max}} - (1-\omega) \frac{\sum_{n=1}^{N} \sum_{k=1}^{K} \delta_t^{n,k}}{\delta_{t,\max}} \tag{10}$$

where ω is the weighting factor.

B. D-BHRA Algorithm Network Structure

This section describes the D-BHRA algorithm in terms of Deep Q-network (DQN) structure and training.

DQN structure: Since the action space defined in Eq. (9) is discrete, this paper uses the DQN learning method [12], which uses satellites as the main body of the intelligent's learning strategy. By training the convolutional neural network, the optimal action value function can be obtained, and from this, the objective value of the optimal policy network is derived as follows:

$$Q^*(s,a) = \max_{\pi} \mathbb{E}[r_t + \gamma r_{r+1} + \ldots | s_t = s, a_t = a, \pi] \tag{11}$$

where γ is the discount factor.

DQN training: The use of convolutional neural network to approximate the Q-value function may face the problems of overestimation, instability and even divergence. The D-BHRA algorithm employs a dual network technique, i.e., the strategy network Q^* and the target network Q^-, which enhances the stability and reliability of DQN training by periodically updating the target network parameters. Meanwhile, the memory replay technique is introduced to store and utilize previous experiences to improve sampling efficiency and learning robustness. The algorithm randomly extracts the experience terms from the cache and calculates the target value according to the Bellman equation, which effectively solves the overestimation and instability problems of DQN. The target value is calculated as follows:

$$y_t = r_t + \gamma \max Q^-(s_{s+1}, a; \theta^-) \tag{12}$$

where θ^- is the parameter of the target network, which is updated with the policy network parameter θ every $\mathcal{G}$ steps. Based on the target value in Q^*, the loss value $L_t(\theta_t)$ of the network at time t is:

$$L_t(\theta_t) = \mathbb{E}_{(s_t,a_t,r_t,s_t+1) \sim U(\mathfrak{R})} (y_t - Q^*(s,a;\theta_t))^2 \tag{13}$$

where θ_t is the policy network parameter at moment t.

C. D-BHRA Algorithm Steps and Process

The basic steps of the D-BHRA algorithm proposed in this paper are shown in Algorithm 1.

Algorithm 1 The training of D-BHRA

Input: States s_t
Output: Actions a_t

Initialize cell communication demands $\boldsymbol{D}_t$, delay tolerance matrix $\boldsymbol{\Phi}_t$.
Initialize replay buffer $\mathfrak{R}$, exploration parameter ϵ, minibatch size B.
Initialize target network Q^- with weights $\theta^- \leftarrow \theta$.
for episode = 1 to $episode_{\max}$ **do**
  Collect current satellite state s_t.
  for k = 1 to K **do**
    Determine the action $a_{k,t}$ according to the satellite state s_t.
  end for
  Get all beam actions $a_t = (a_{1,t}, \ldots, a_{K,t})$.
  Allocate power and frequency blocks according to a_t.
  Calculate the reward r_t and get next real state s_{t+1}.
  Store transition (s_t, a_t, r_t, s_{t+1}) in $\mathfrak{R}$.
  Samples a mini-batch of (s_i, a_i, r_i, s_{i+1}) from $\mathfrak{R}$ and calculates the loss $L_t(\theta_t)$.
  Train the Q-network.
  Update the target network θ^- of agent every $\mathcal{G}$ steps.
end for

IV. Evaluation

In this section, we describe the experiment setup and analyze the experimental results.

A. Experiment Setup

The simulations were performed based on the Python 3.10 platform, and all simulations were performed on AMD Ryzen 7 4800H, 16 GB RAM, and Radeon Graphics. In a simulation scenario of a GEO multibeam satellite operating in K_a band GEO, this paper creates 30 ground communities in the system, with each satellite carrying 7 beams. Table I summarizes the simulation parameters of the system and algorithm [7].

The communication demand required by the terrestrial base station follows a Poisson distribution with parameter $\mu = 1$. During training, the demand for a single cell is controlled to be between 0 and 30 Mbps. During testing, the service demand of individual cells and the total service demand of each cell are varied. The statistical average of 300 test results is taken as the evaluation metric in this paper.

TABLE I
THE SETTING OF EXPERIMENTAL PARAMETERS

PARAMETERS	VALUES
Satellite altitude h	36786 km
K_a Band f_c	20 GHz
Total available bandwidth B_{tot}	18000 KHz
Total satellite power P_{tot}	23 dBW
Single-beam power threshold $P_{\max}$	20 dBW
Number of beams K	7
Number of cells N	30
Maximum transmit antenna gain G_m	40.3 dBi
Free space loss L_f	209.6 dB
Maximum receiving antenna gain G_r	31.6 dBi
Replay memory capacity $\mathfrak{R}$	100000
Minibatch size B	128
Discount factor γ	0.9
Target network update frequency $\mathcal{G}$	100
Learning rate α	0.0001

B. Performance Metrics

To evaluate the performance of D-BHRA algorithm, the following performance evaluation metrics are defined in this paper:

- System throughput: the total number of packets transmitted by the system per unit time.
- Supply-demand gap: the absolute value of the communication capacity provided by the illumination of the satellite beam and the demand of the cell.
- Delay: the sum of queuing delay and propagation delay of packets in the cache queue.

To verify the impact of the proposed D-BHRA algorithm on the performance of multibeam satellite systems, we compare the proposed method with different schemes as follows:

- Greedy-Based Beam Hopping and Resource Allocation (G-BHRA) [14]: each decision time slot and the algorithm selects the seven cells with the highest cell demand for service and uses a fixed allocation power of 20W and a bandwidth of 18000KHz.
- Genetic Algorithm-Based Beam Hopping and Resource Allocation (GA-BHRA) [10]: candidate solutions evolve as individuals. Through selection, crossover, and mutation, less fit solutions are eliminated until the most fit solution is selected.
- Polling-Based Beam Hopping and Resource Allocation (P-BHRA) [7]: according to the order of cell location, each decision time slot selects 7 cells in turn for beam coverage and allocates transmission power in proportion to demand.

C. Results Analysis

In order to assess the effectiveness of the proposed algorithm, the convergence process of the algorithm is firstly demonstrated and analyzed for the convergence process of the algorithm, as shown in Fig. 3. The convergence process of the target reward value of the D-BHRA algorithm can be seen that the horizontal coordinate is the number of iterations, and the vertical coordinate is the normalized reward value, and the normalized reward value tends to be stable and finally converges better when trained up to 10,000 episodes, whose reward value is the normalized value of system throughput and supply-demand gap.

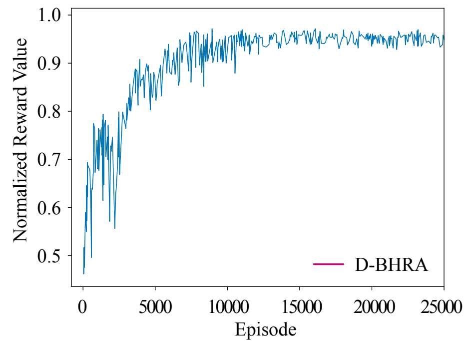

Fig. 3. The normalized reward value.

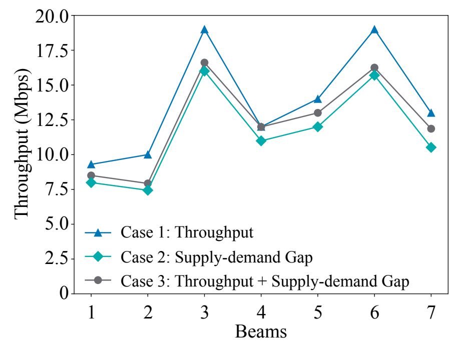

Fig. 4. Throughput comparison for three cases.

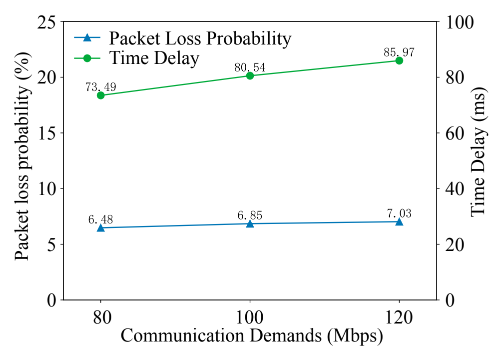

Fig. 5. The packet loss rate and delay of D-BHRA.

Fig. 4 shows the system throughput in the same decision time slot for three scenarios in an environment with a total demand of 100 Mbps for service communities. Case 1: When only considering throughput, the system prioritizes maximizing throughput, with the highest throughput in any decision time slot. During simulation, ω is set to 1; Case 2: When only considering the supply-demand gap, the system prioritizes minimizing the supply-demand gap, resulting in the lowest throughput in any decision time slot. During simulation, ω is set to 0; Case 3: Taking into account both maximizing throughput and minimizing supply-demand gap, the algorithm proposed in this article can better balance system throughput in this regard. Compared to the situation where only supply-demand gap exist, the throughput has increased by about 8.33%. Fig. 5 shows the delay and packet loss rate of the system under three demands; as the demand of the served cell increases, both the delay and packet loss rate increase.

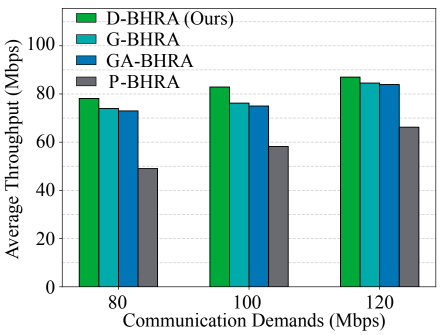

Fig. 6. Performance comparison in terms of system throughput.

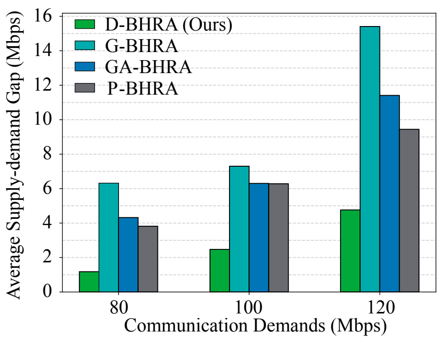

Fig. 7. Performance comparison in terms of supply-demand gap.

As shown in Fig. 6, the proposed algorithm in this paper improves the average throughput of the system by 8.77%, 10.51%, and 59.36% compared with G-BHRA, GA-BHRA, and P-BHRA. D-BHRA has a better performance in terms of average throughput mainly because D-BHRA consolidates the remaining resources and divides them to other cells twice when serving the cells with less demand, whereas G-BHRA focuses on finding the cell with the highest demand for service and ignores the secondary allocation of resources, resulting in a waste of resources. As shown in Fig. 7, D-BHRA has lower supply-demand gap and is better able to provide more satisfactory services to the ground cells, and the mean supply-demand gap is reduced by about 66.16%, 58.28%, and 49.58% compared with G-BHRA, GA-BHRA, and P-BHRA. The main reason is that D-BHRA adds the supply-demand gap in the decision making, when the cell demand is less the same will be served, while other algorithms ignore this. Hence, D-BHRA has better system throughput while taking into account minimizing the supply-demand gap.

V. Conclusion

In this paper, we propose a novel joint beam hopping satellite scheduling and time-frequency-space resource allocation algorithm oriented to dynamic service demands for efficient allocation of multidimensional resources in multibeam satellite systems. In the proposed algorithm, the multidimensional resources are firstly designed in terms of time-frequency block-space interference isolation and optimized by joint beam hopping. The algorithm is modeled as a multi-objective optimization problem aiming to maximize the system throughput and minimize the supply-demand gap. The experimental results show that D-BHRA can achieve higher system throughput with the lowest supply-demand gap compared to the baseline algorithm.

References

[1] X. Du, X. Hu, Y. Wang, and W. Wang, “Dynamic Resource Allocation for Beam Hopping Satellites Communication System: An Exploration,” in 2022 IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. (TrustCom), Wuhan, China, 2022, pp. 1296–1301.

[2] A. Wang, L. Lei, E. Lagunas, A. I. Pérez-Neira, S. Chatzinotas, and B. Ottersten, “Joint Optimization of Beam-Hopping Design and NOMA-Assisted Transmission for Flexible Satellite Systems,” IEEE Trans. Wireless Commun., vol. 21, no. 10, pp. 8846–8858, Oct. 2022.

[3] Y. Han, C. Zhang, and G. Zhang, “Dynamic Beam Hopping Resource Allocation Algorithm Based on Deep Reinforcement Learning in Multi-Beam Satellite Systems,” in 2021 3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST), Guangzhou, China, 2021, pp. 68-73.

[4] Y. Wang, M. Zeng, and Z. Fei, “Efficient Resource Allocation for Beam-Hopping-Based Multi-Satellite Communication Systems,” Electronics, vol. 12, no. 11, article 2441, 2023.

[5] S. Guo, L. Zhao, and Y. Cui, “Latency Optimization of LEO Satellite Communication Systems with Beam Hopping,” in 2022 IEEE International Conference on Satellite Computing (Satellite), Shenzhen, China, 2022, pp. 37–42.

[6] M. Zhang, X. Yang, and Z. Bu, “Resource Allocation with Interference Avoidance in Beam-Hopping Based LEO Satellite Systems,” in 2023 4th Information Communication Technologies Conference (ICTC), Nanjing, China, 2023, pp. 83–88.

[7] G. Xu, F. Tan, Y. Ran, Y. Zhao, and J. Luo, “Joint Beam-Hopping Scheduling and Coverage Control in Multibeam Satellite Systems,” IEEE Wireless Commun. Lett., vol. 12, no. 2, pp. 267–271, Feb. 2023.

[8] Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, “Dynamic Beam Pattern and Bandwidth Allocation Based on Multi-Agent Deep Reinforcement Learning for Beam Hopping Satellite Systems,” IEEE Trans. Veh. Technol., vol. 71, no. 4, pp. 3917–3930, Apr. 2022.

[9] X. Liao, X. Hu, Z. Liu, S. Ma, L. Xu, X. Li, W. Wang, and F. M. Ghannouchi, “Distributed Intelligence: A Verification for Multi-Agent DRL-Based Multibeam Satellite Resource Allocation,” IEEE Commun. Lett., vol. 24, no. 12, pp. 2785–2789, 2020.

[10] X. Hu, L. Wang, Y. Wang, S. Xu, Z. Liu, and W. Wang, “Dynamic Beam Hopping for DVB-S2X GEO Satellite: A DRL-Powered GA Approach,” IEEE Commun. Lett., vol. 26, no. 4, pp. 808–812, Apr. 2022.

[11] NF Z84-307-2-2020, “Digital Video Broadcasting (DVB) - Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications - Part 2 : DVB-S2 Extensions (DVB-S2X) (V.1.2.1),” 2021.

[12] S. Liu, X. Hu, and W. Wang, “Deep Reinforcement Learning Based Dynamic Channel Allocation Algorithm in Multibeam Satellite Systems,” IEEE Access, vol. 6, pp. 15733–15742, Mar. 2018.

[13] J. P. Choi and V. W. S. Chan, “Optimum power and beam allocation based on traffic demands and channel conditions over satellite downlinks,” IEEE Trans. Wireless Commun., vol. 4, no. 6, pp. 2983–2993, Nov. 2005.

[14] F. Tian, L. Huang, G. Liang, X. Jiang, S. Sun, and J. Ma, “An Efficient Resource Allocation Mechanism for Beam-hopping Based LEO Satellite Communication System,” in 2019 IEEE Int. Symp. Broadband Multimedia Syst. Broadcasting (BMSB), Jeju, Korea (South), 2019, pp. 1–5.