Multi-Satellite Beam Hopping Based on Deep Reinforcement Learning for LEO Satellite Systems

Authors: Yifan Xu, Ruili Zhao, Yongyi Ran, **Jiangtao Luo**

Presenter: Dr. Prof. Jiangtao Luo*

（Chongqing University of Posts and Telecommunications, China）

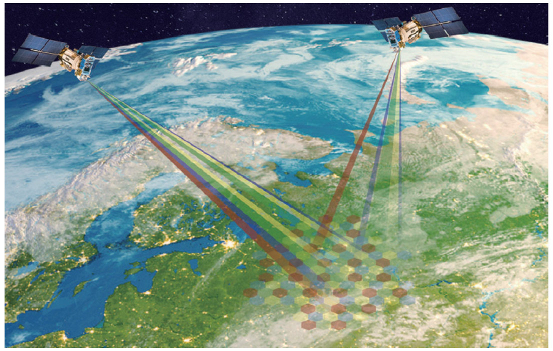
Dec 12, 2024 · Cape Town, South Africa

Outline

01 Research Motivation

Background and Challenges

重慶郵電大學
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

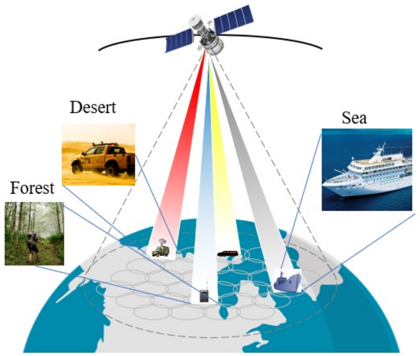


Mega LEO constellation

Features

Compared to GEO

- Shorter delay
- Higher rate
- Denser coverage
- Limited on-board resources



Beam hopping Satellite communication System

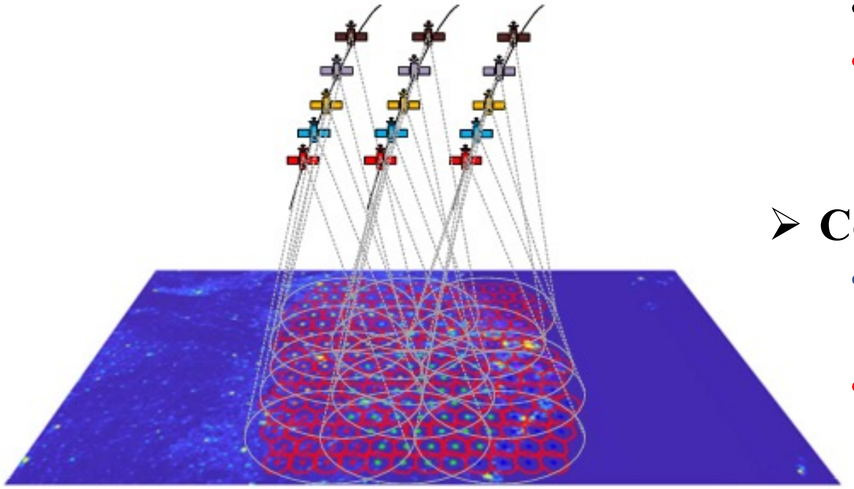
Challenges

Intra-satellite and inter-satellite interference cannot be neglected.

Problem

How to obtain load balance with optimized resource allocation for a multi-beam mega LEO constellation?

Common Approaches



Existing Methods:

- **Borrow single-satellite scheme from GSP system:**
  - Independent decision-making by different satellites
  - seriously beam interference neglected, low resource utilization and load imbalance

- **Consider inter-satellite interferences**:
  - **Classical approaches:**
    - ✓ **Dynamic programming [1]**, **Convex optimization[2]**
  - **Solution space increases drastically** with the number of satellites and beams, leading to local optima.
  - **Poor timeliness and inflexible:** As cover area changes, it requires **re-modeling** and iterative solving.

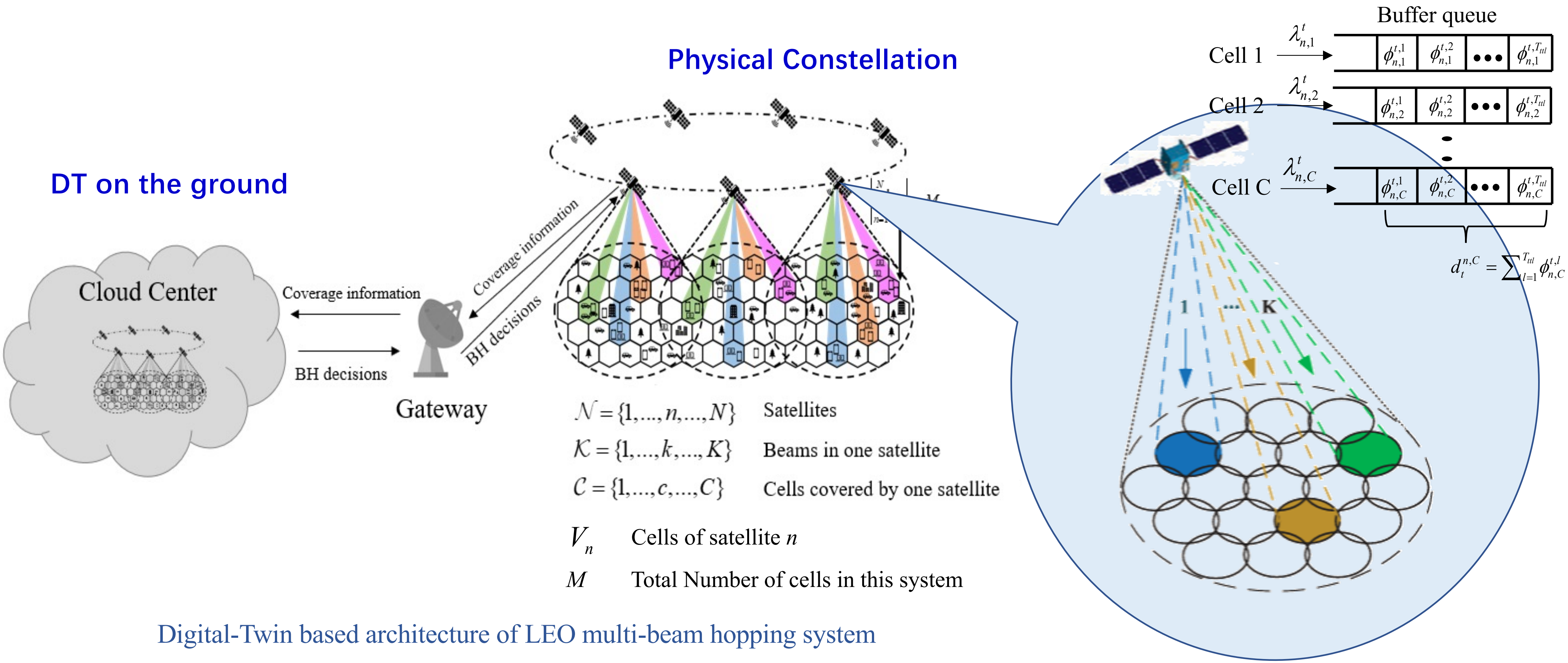
Our approach: Deep Reinforcement Learning + Digital Twin

1. Y. Wang, M. Zeng, and Z. Fei, "Efficient resource allocation for beam-hopping based multi-satellite communication systems" Electronics, vol. 12, no. 11, 2023.
2. Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, "Multi-satellite beam hopping based on load balancing and interference avoidance for NGSO satellite communication systems," IEEE Transactions on Communications, vol. 71, pp. 282–295, Jan 2023.

02 System model & Problem formulation

System model

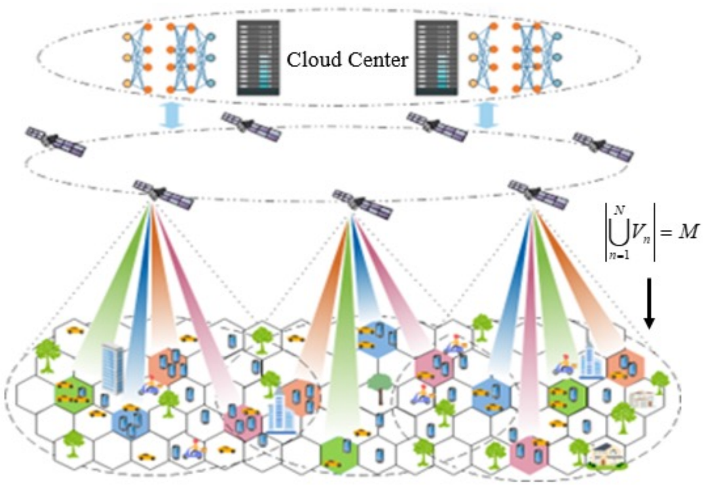
- Overall architecture



Digital-Twin based architecture of LEO multi-beam hopping system

System model

- Communication model



$\mathcal{N} = \{1,\ldots,n,\ldots,N\}$ Satellites

$\mathcal{K} = \{1,\ldots,k,\ldots,K\}$ Beams in one satellite

$\mathcal{C} = \{1,\ldots,c,\ldots,C\}$ Cells covered by one satellite

The architecture of LEO multi-satellite beam-hopping system

- Signal-to-noise ratio (SINR) [2] for a cell

$$\mathrm{SINR}_t^{n,c} = \frac{p_{n,c}^t \left|h_{c,c}\right|^2}{N_0 B + \Theta_1 + \Theta_2} \quad (1)$$

$$\Theta_1 = \sum_{e \neq c, S(e)=S(c)} p_{n,e}^t \left|h_{c,e}\right|^2$$

the **intra-satellite** beam interference

$$h_{c,e} = \frac{\sqrt{G_t(\theta_{c,e})G_r^c}}{4\pi \frac{d_{c,e}}{\lambda}}$$

$$\Theta_2 = \sum_{e \neq c, S(e) \neq S(c)} p_e^t \left|h_{c,e}\right|^2$$

the **inter-satellite beam interference**

- Throughput $Th_t^{n,c}$:

$$Th_t^{n,c} = \min\{r_t^{n,c} \cdot T_{slot}, d_t^{n,c}\} \quad (2)$$

$$r_t^{n,c} = B f_{\mathrm{DVB}}(1 + \mathrm{SINR}_t^{n,c}),$$

Channel capacity or link rate

Stored Traffic

Problem formulation

CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

Load gap:

$$G = \sum_{t=1}^{T} \left(\max_{n \in \mathcal{N}} \{I_t^n\} - \min_{n \in \mathcal{N}} \{I_t^n\} \right) \quad (3)$$

$$I_t^n = \sum_{c \in V_n} Th_t^{n,c} x_t^{n,c} \quad (4)$$

Average queue delay:

$$J = \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{c=1}^{C} \tau_t^{n,c} \quad (5)$$

Objective function:

$$P_1 : \min \left\{ \alpha \frac{G}{G_{\max}} + (1-\alpha) \frac{J}{J_{\max}} \right\} \quad (6)$$

$$\text{s.t.} \quad C1: I_t^n = \sum_{c \in V_n} Th_t^{n,c} x_t^{n,c}, n = 1..., N,$$

$$C2: \sum_{c=1}^{C} x_t^{n,c} = K, \forall n, t,$$

$$C3: x_t^{n,c} \in \{0,1\}, \forall n, c, t,$$

$$C4: \sum_{n} x_t^{n,c} \le 1, \forall c, t,$$

MDP design

State Space $s = \{\mathbf{D}_t, \mathbf{W}_t, \mathbf{U}_t\}$ (7)

Traffic matrix

$$\mathbf{D}_t = \begin{bmatrix} d_t^{1,1} & d_t^{1,2} & \dots & d_t^{1,C} \\ d_t^{2,1} & d_t^{2,2} & \dots & d_t^{2,C} \\ \dots & \dots & \dots & \dots \\ d_t^{N,1} & d_t^{N,2} & \dots & d_t^{N,C} \end{bmatrix}$$

Delay matrix

$$\mathbf{W}_t = \begin{bmatrix} \tau_t^{1,1} & \tau_t^{1,2} & \dots & \tau_t^{1,C} \\ \tau_t^{2,1} & \tau_t^{2,2} & \dots & \tau_t^{2,C} \\ \dots & \dots & \dots & \dots \\ \tau_t^{N,1} & \tau_t^{N,2} & \dots & \tau_t^{N,C} \end{bmatrix}$$

Satellite coverage matrix

$$\mathbf{U}_n^t = \begin{bmatrix} u_t^{1,1} & u_t^{1,2} & \dots & u_t^{1,C} \\ u_t^{2,1} & u_t^{2,2} & \dots & u_t^{2,C} \\ \dots & \dots & \dots & \dots \\ u_t^{N,1} & u_t^{N,2} & \dots & u_t^{N,C} \end{bmatrix}$$

$u_t^{n,c} \in \{0, 1\}$

$u_t^{i,j} = 1 \rightarrow$ Cell j of satellite n is covered by satellite i

Action Space $a_n = \{x_t^{n,1}, x_t^{n,2}, \dots x_t^{n,C}\}, \sum_{c=1}^{C} x_t^{n,c} = K$ (8)

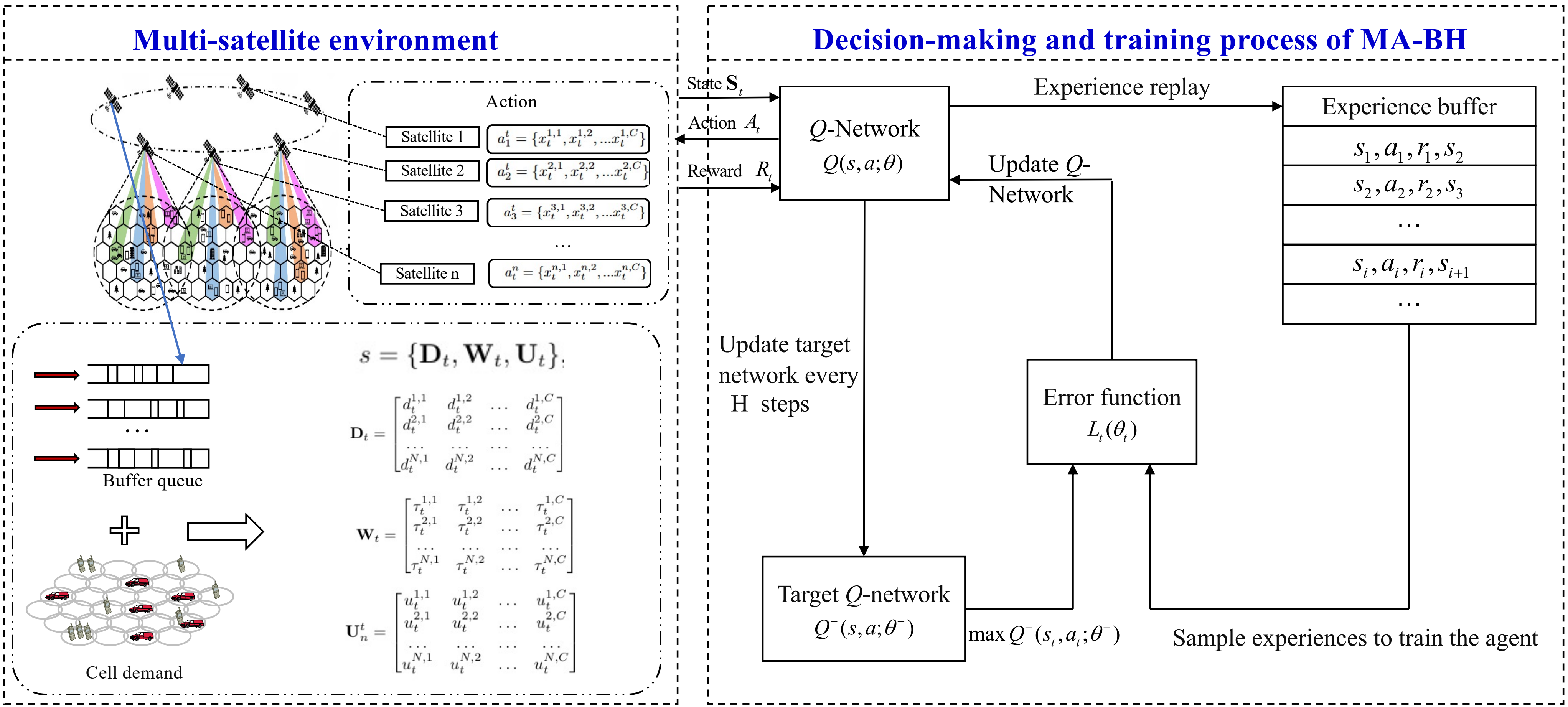
Reward Function

$$F_t = \sum_{n=1}^{N} \sum_{c=1}^{C} \tau_t^{n,c}.$$

$$r^t = -\left\{ \alpha \frac{Q_t}{Q_{\text{norm}}} + (1-\alpha) \frac{F_t}{F_{\text{norm}}} \right\} \quad (9)$$

$$Q_t = \max_{n \in \mathcal{N}} \{I_t^n\} - \min_{n \in \mathcal{N}} \{I_t^n\}$$

MA-BH



Multi-satellite environment
Decision-making and training process of MA-BH
Action
Satellite 1
Satellite 2
Satellite 3
Satellite n
Buffer queue
Cell demand
State S_t
Action A_t
Reward R_t
Q-Network
$Q(s,a;\theta)$
Experience replay
Experience buffer
s_1,a_1,r_1,s_2
s_2,a_2,r_2,s_3
s_i,a_i,r_i,s_{i+1}
Update Q-Network
Update target network every H steps
Error function
$L_t(\theta_t)$
Target Q-network
$Q^{-}(s,a;\theta^{-})$
$\max Q^{-}(s_t,a_t;\theta^{-})$
Sample experiences to train the agent

03 Performance evaluation

Performance Evaluation

重庆邮电大学
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

Simulation Parameters Settings

PARAMETERS	VALUES
Satellite altitude	700 km
Downlink frequency	11.7 GHz
Number of satellites, N_0	12
System bandwidth, B	500 MHz
Total number of cells, M	168
Number of cells covered by each satellite, C	19
Number of spotbeams, K	4
Satellite payload capacity	2000 Mbps
Poisson arrival rate of cell traffic	50 Mbps $\sim$ 150 Mbps
Length of queue buffer	10 timeslots
Timeslot duration	2 ms
Weight factor α	0.5
Noise power spectral density N_0	-171.6 dBm/Hz
Satellite 3 dB beamwidth θ_{3dB}	3°
Training Episodes	6000
Timeslots every episode	100
Learning rate, η	0.001
Replay buffer size, $\lvert R\rvert$	10000
Target network update frequency, $\lvert H\rvert$	200
Minibatch size	128
Discount factor, γ	0.95
Initial exploration rate	0.5
Final exploration rate	0.02
Optimizer	Adam

Baseline Algorithms

- **Random Beam Hopping** [4]**(R-BH)**: In the R-BH method, each satellite randomly selects 4 cells from the 19 available cells to serve in each time slot.
- **Greedy Beam Hopping** [5]**(G-BH)**: In the G-BH method, each satellite selects the 4 cells with the highest traffic volume from its coverage area to serve in each time slot.
- **Periodic Beam Hopping** [2]**(P-BH)** : In the P-BH method, each satellite periodically serves different cells according to a predetermined beam-hopping sequence.

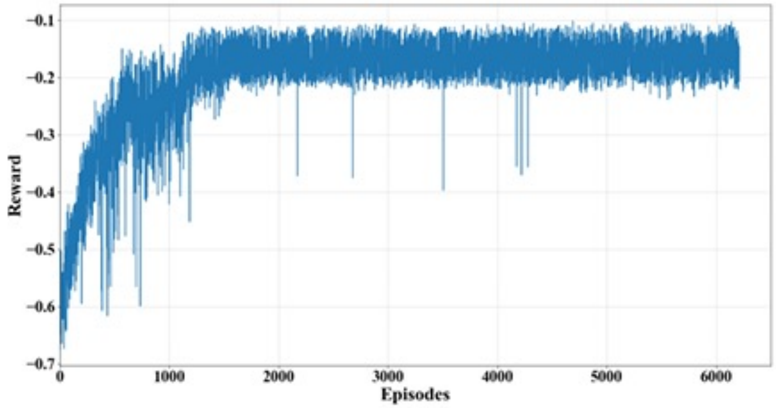
[4]. Z. Lin, Z. Ni, L. Kuang, C. Jiang and Z. Huang, "Dynamic Beam Pattern and Bandwidth Allocation Based on Multi-Agent Deep Reinforcement Learning for Beam Hopping Satellite Systems," in IEEE Transactions on Vehicular Technology, vol. 71, no. 4, pp. 3917-3930, April 2022.

[5]. J. Zhang et al., "System-Level Evaluation of Beam Hopping in NR-Based LEO Satellite Communication System," 2023 IEEE Wireless Communications and Networking Conference (WCNC), Glasgow, United Kingdom, 2023.

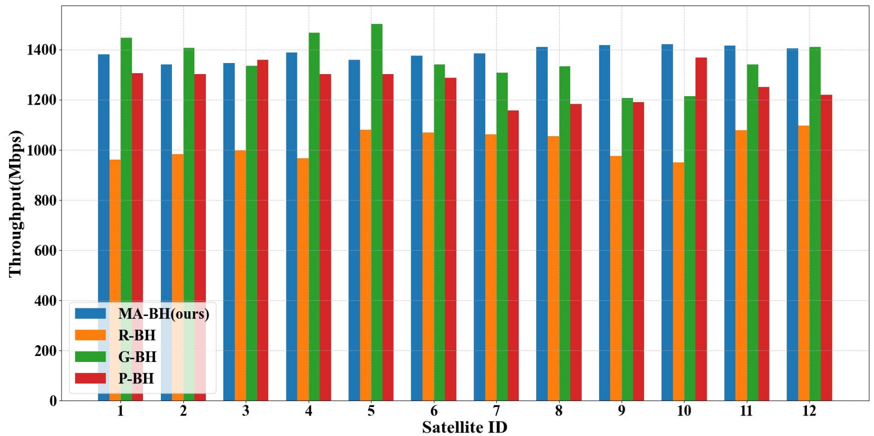
Performance Metrics

- ***System throughput*****:** the total number of packets transmitted by the system per unit time.
- ***Load gap* between satellites:** the difference between the load of the largest and least load satellite in a multi-satellite system
- ***Delay*****:** the average queuing delay of packets in the cache queue for cells.

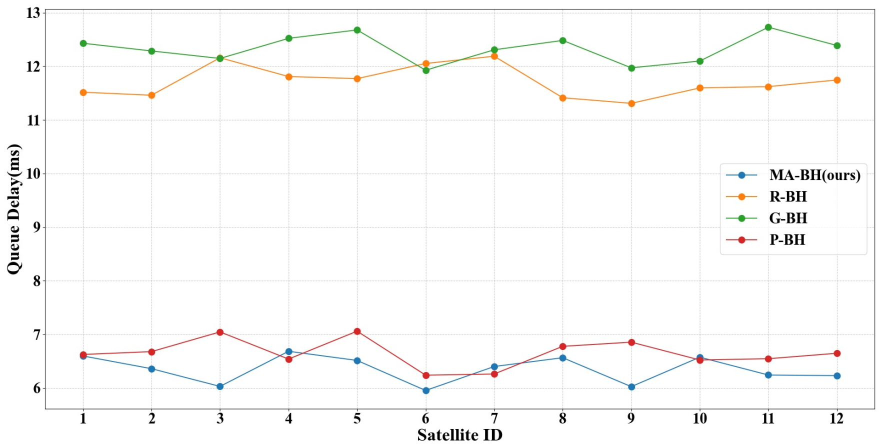
Performance Evaluation



The normalized reward value.



Throughput of each satellite



Queue delay of each satellite.

	Throughput (Mbps)	Queue delay (ms)	Load gap (Mbps)
G-BH	16319 0.79%↑	12.33 48.4% ↓	294.6 70.1%↓
R-BH	12281 33.9%↑	11.72 45.8% ↓	144.3 44.2%↓
P-BH	15233 7.9% ↑	6.5 2.3% ↓	210.6 61.3%↓
MA-BH(ours)	16449	6.35	80.4

04 Conclusions

Conclusion

重慶郵電大學
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

- Propose a **multi-agent beam hopping algorithm** for a LEO satellite system, based on **a digital-twin architecture**.
- Consider both **inter-satellite interference** and **intra-satellite ones**.
- Optimize jointly the system load gap and queue delay to achieve **load balance** between satellites while keeping **high overall throughout**.

Any comments and discussion is welcome

My Email:
Luojt2008@gmail.com ;
Luojt@cqupt.edu.cn