重慶郵電大學
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

IEEE GLOBECOM®

IEEE Global Communications Conference
4–8 December 2023 // Kuala Lumpur, Malaysia

DeepISL: Joint Optimization of LEO Inter-Satellite Link Planning and Power Allocation via Parameterized Deep Reinforcement Learning

Authors : Yue Li, **Jiangtao Luo***, Yongyi Ran, Jiahao Pi
Presenter: **Jiangtao Luo** (CQUPT, China)

Dec 6, 2023

About the Conference

CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

- One of the Flagship or Top Conferences (旗舰或顶会) in Communications.
  - IEEE Global Communications Conference (**GLOBECOM**)
  - IEEE International Conference on Communications （**ICC**）
- Conference Site: Kuala Lumpur, Malaysia (马来西亚吉隆坡)

PETRONAS Twin Towers

Kuala Lumpur Convention Centre

Oral Presentation

About my team

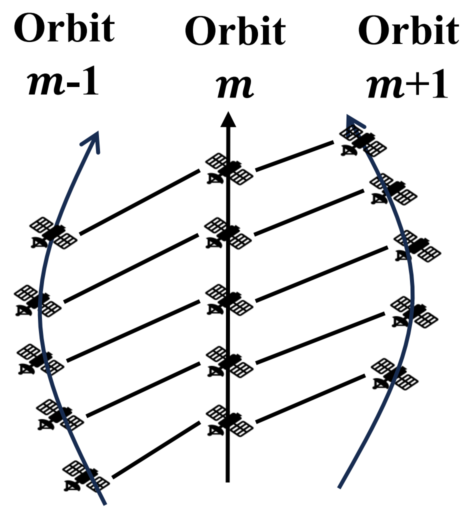
	**Dr. Prof. LUO, Jiangtao**	**Vehicular networks, Satellite Internet** (Nankai Univ., CAS, Shanghai Jiao Tong Univ.)
	**Prof. PENG, Daqin**	**Mobile Communications, Mobile IoT** one of the earliest TD-SCDMA specialist
	**Dr. Prof. XU, Guoliang**	**Mobile Big Data, Computer Vision, Fiber sensing** (Tsinghua Univ., CAS)
	**Prof. WANG, Qianzhu**	**Mobile Communications, Mobile IoT**
	**Dr. ZHANG, Changhong**	**Antenna Design** (UESTC, CAS, Shanghai Jiao Tong Univ.)
	**Dr. RAN, Yongyi**	**Data Center Networking、Satellite Internet** (USTC/USTC/USTC, NTU)

Contents

1. Background and Motivation

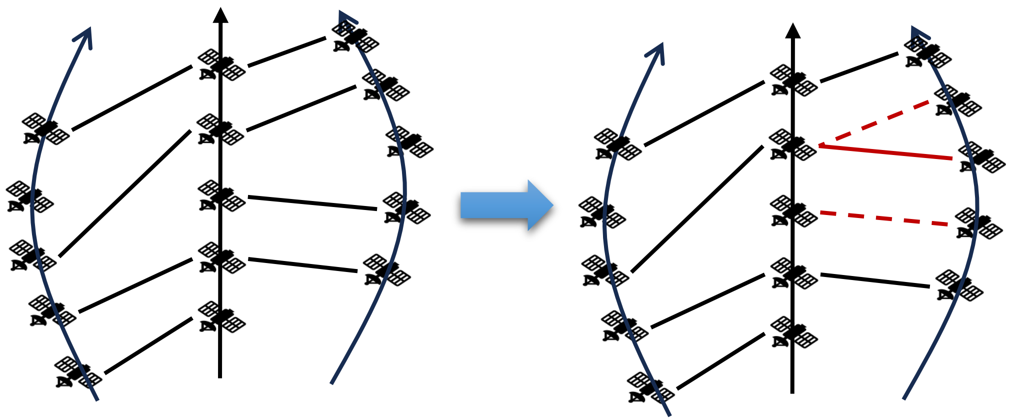
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

- Mega LEO Constellations are being built all over the world
- Inter-Satellite Links (ISLs) offer **relays**, **collaboration**, **path redundancy** and **resilience**.
- Static scheme is simple but excessively redundant and needs more energy!



Static Scheme (+Grid)

How to setup ISLs on-demand and adaptable?



Dynamic Planning

1. Background and Motivation

CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

❑ Challenges

- Beam steering and tracking
- High-speed movement of satellites
- Interference management and channels allocation
- **Power and energy management**
- Limited resources on-board

❑ Existing solutions

- Greedy Matching [2]
- ILP [3]
- MA-DRL [4]

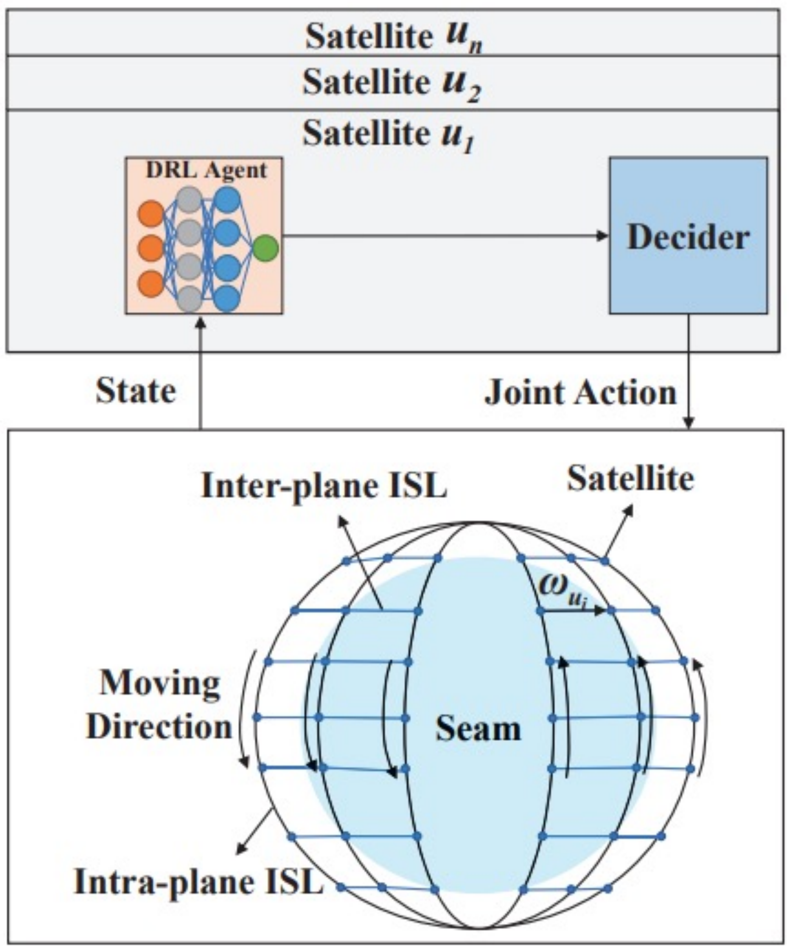
[2] I. Leyva-Mayorga, GLOBECOM 2021
[3] Z. Yan, WCL, 2020
[4] J. Pi, ICC 2022 (our previous work)

Can we combine
dynamic power allocation with dynamic setup of ISLs ?

Higher trans. rates | Higher energy efficiency | Lowest switching costs

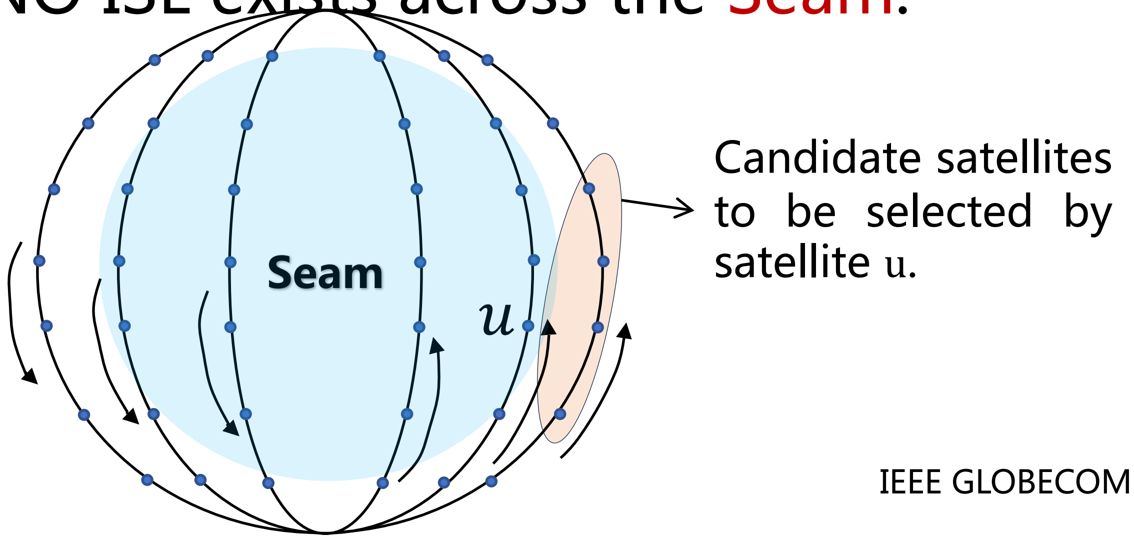
2. Problem and Approach

System Model



❑ Basic ideas

- Every satellite has an agent.
- Every satellite selects its target peer on the right orbit to establish an ISL, based on the observed system states.
- Intra-plane ISLs do not change.
- NO ISL exists across the Seam.



2. Problem and Approach

Problem Statement

Normalized Utility function：

energy efficiency

transmission rate

Steering angle

$$\varphi(t)=\sum_{e_{uv}}\left(\alpha_1 E^{uv}_{eff,t}+\alpha_2 R_{e_{uv},t}\right)-\sum_{e_{uv}}\alpha_3\theta_{uv,t}$$

Benefits per unit of energy efficiency

Benefits per unit transmission rate

Cost per unit antenna steering angle

Optimization Problem：

$$\max\sum_{t=1}^{N_d}\varphi(t)$$

$$s.t.\begin{cases} uv\in Y_t \\ e_{uv}\in E_t \\ P_{e_{uv},t}\le P^{\max}_{u,t}-P_0 \\ \lambda\omega_{u,t}/\delta(t)\ \le\ R_{e_{uv},t}\ \le\ \omega_{u,t}/\delta(t) \end{cases}$$

2. Problem and Approach

重庆邮电大学
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

Related Models

Energy efficiency model：

$$E_{eff,t}^{uv} = \omega_{e_{uv},t} / E_{e_{uv},t} = R_{e_{uv},t} / P_{e_{uv},t}$$

Data to transmit

Energy consumed

Doubly constrained of rate：

$$\lambda\omega_{u,t} / \delta(t) \le R_{e_{uv},t} \le \omega_{u,t} / \delta(t)$$

Satisfaction factor: minimum ratio of data transmitted

[8] M. Marchese, IEEE Tran. on Green Comm. and Net., 4(3), 2020.

Energy model：

Solar energy harvested：

$$E_{u,t}^{harvest} = P_{u,t}^{harvest} \cdot \delta(t)$$

$$P_{u,t}^{harvest} = \tau \cdot \varphi \cdot \varpi \cdot A_e \cdot \sin\sigma$$

Area of the solar panel

collection constant, 0 or 1

conversion efficiency

Solar irradiance per unit area

Incident angle

Energy consumed：

$$P_{u,t}^{consume} = P_{e_{uv},t} + P_0 \le P_{u,t}^{max}$$

for Intra-plane ISLs and others

for Inter-plane ISLs

2. Problem and Approach

Related Models

Switching cost model :

The steering angle of satellite u:

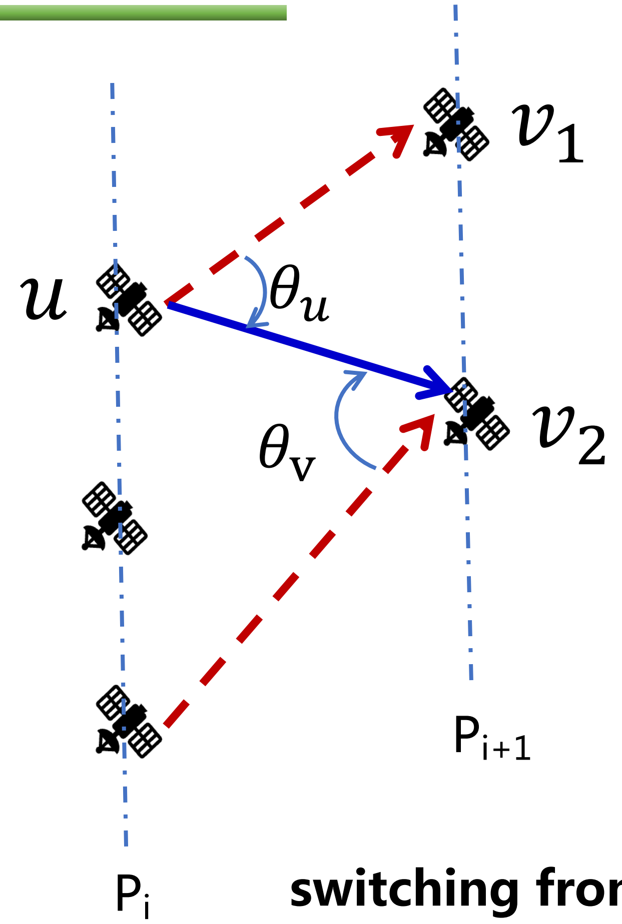
$$\theta_u = \arccos\left(\frac{\left(d_{|uv_1|}\right)^2 + \left(d_{|uv_2|}\right)^2 - \left(d_{|v_1 v_2|}\right)^2}{2 \cdot d_{|uv_1|} \cdot d_{|uv_2|}}\right)$$

Average antenna steering angle

$$\hat{\theta}_{u,t} = \frac{\sum_{v_1 \neq v_2 \in Y_{u,t}^{+}} \theta_u + \sum_{v_1 \neq v_2 \in Y_{u,t}^{-}} \theta_u}{\binom{N_{u,t}^{+}}{2} + \binom{N_{u,t}^{-}}{2}}$$

Mean antenna steering angle

$$\theta_{uv,t} = \begin{cases} 0, & e_{uv} \in E_{t-1} \\ \hat{\theta}_{u,t} + \hat{\theta}_{v,t}, & e_{uv} \notin E_{t-1} \end{cases}$$



switching from uv_1 to uv_2

The switching costs are assumed proportional to the steering angle,

Switching cost $\propto \theta_{uv}$

[4] J. Pi, ICC 2022.

2. Problem and Approach

重庆邮电大学
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

MA-DRL based Approach

State Space ：

$$S_i = \{\boldsymbol{D}_{i,t}, C_{i,t}, \omega_{i,t}\}$$

{ Distance to target satellites; Battery capacity; Data bytes to transmit }

Action :

$$a_{i,t} = \{v_{i,t}, p_{i,t}\}$$

{Selected target satellites (discrete); Allocated power (continuous)}

A **hybrid action apace** calls for **parameterized DRL approach**.

Reward Function ：

$$\text{RWD} = \sum_{i=1}^{N_n} r_{i,t},$$

$$r_{i,t} = \underbrace{\kappa_i}_{\text{Conflict factor}}\left(\alpha_1 \underbrace{E_{eff,t}^{iv_{i,t}}}_{\text{EE Reward}} + \alpha_2 \underbrace{R_{e_{iv_{i,t}},t}}_{\text{TR reward}}\right) - \underbrace{\alpha_3 \theta_{iv_{i,t},t}}_{\text{Switching cost}}$$

Procedure of DeepISL algor.

Algorithm 1: Training process of DeepISL

for *agent* $i = 1, N_n$ **do**
  Initialize deterministic strategy network $\mu_{v_i}(\theta_i)$ and value network $Q_i(w_i)$, learning rate α,β and probability ξ . Initialize the experience pool $\boldsymbol{\Gamma}$
end
for *episode* $= 1$ *to* M' **do**
  for *agent* $i = 1, N_n$ **do**
    Observe the state $s_{i,t}$
    Obtain continuous parameter $p_{v_{i,t}} \leftarrow \mu_{v_i}(\theta_i)$.
    Obtain discrete action by $v_{i,t} = argmax_{v_i \in V_i} Q\left(s_{i,t}, (V_i, p_{V_i}); w_i\right)$
    Select action $a_{i,t}$ according to ξ-greedy strategy
    Execute $a_{i,t}$ and observe $r_{i,t}$ and $s_{i,t+1}$
    Store transition $[s_{i,t}, s_{i,t+1}, a_{i,t}, r_{i,t}]$ into $\boldsymbol{\Gamma}$
  end
end
for *agent* $i = 1, N_n$ **do**
  Randomly draw a batch of $[s_b, s_{b+1}, a_b, r_b]_{b \in \bar{B}}$ from $\boldsymbol{\Gamma}$ $y_b = r_b + \gamma \max_{v \in V} Q\left(s_{b+1}, v, \mu_{v_i}(s_{b+1}; \theta_t); w_t\right)$ Calculate $\ell_t(w_i)$ and $\ell_t(\theta_i)$ according to Equations (25) and (26)
  Update the network parameters w_i and θ_i according to Equations (28) and (29)
end

3. Evaluation and Results

重庆邮电大学
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

Metrics

- Mean energy efficiency:
  - the ratio of the sum of the energy efficiency of each inter-plane ISL to the total number of inter-plane ISLs.
- Mean throughput:
  - the ratio of the sum of the throughput of each inter-plane ISL to the total number of inter-plane ISLs.
- Switching ratio:
  - the ratio of switched inter-plane ISLs to the total inter-plane ISLs.

Comparison Algorithms

a) GIEM: Greedy Independent Experiments Matching [2]

From [2] I. Leyva-Mayorga, TWC, 20(6), 2021.

b) DY-DQN : relax continuous power allocation to discrete power allocation

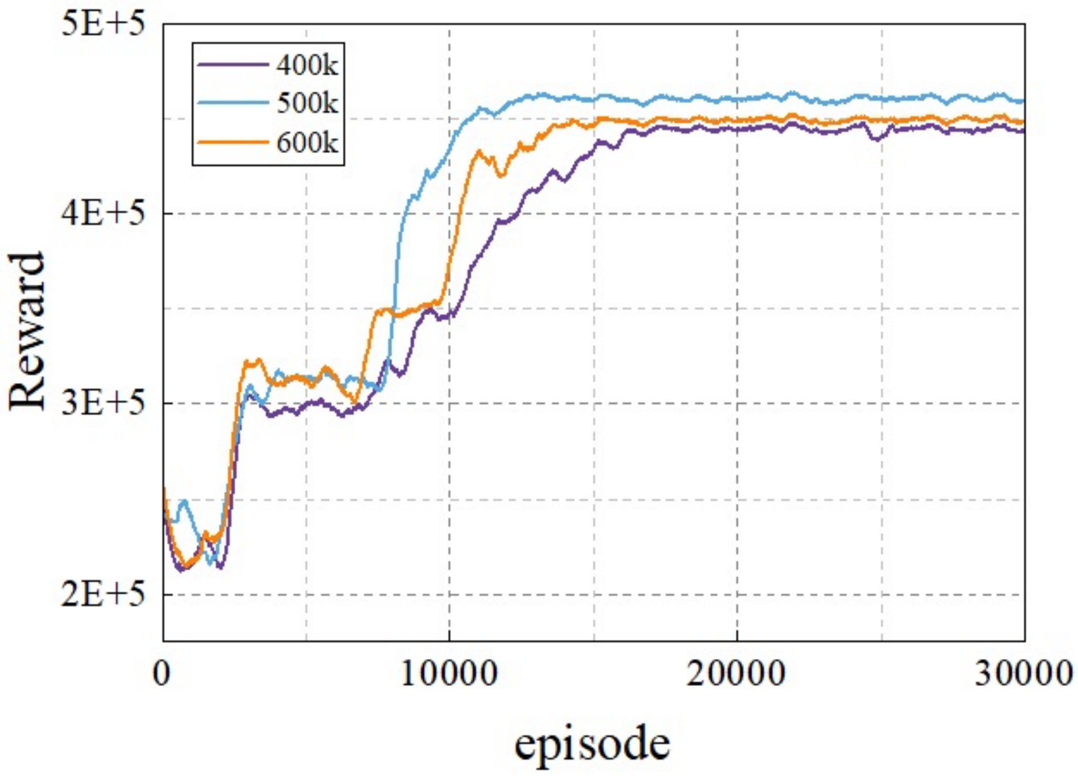
c) FP-DQN : fixed power allocation

Derived from DeepISL.

3. Evaluation and Results

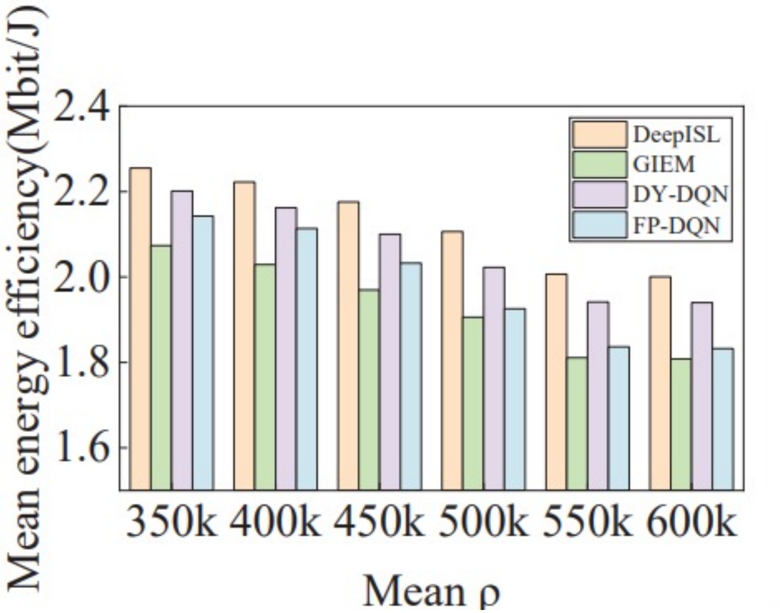
Parameter	Symbol	Value
Number of satellites	N	66
Number of orbital planes	M	6
Altitude of orbital planes	H	780 Km
Inclination of orbital planes	ϵ_m	86.4 deg
Carrier frequency in the Ka-band	f	23.28 GHz
Carrier bandwidth	B	15 MHz
Quality factors	G_{rec}/T_e	8 dB/K
The size of each packet	F_f	1500B
The duration of the time slot	$\delta(t)$	300 s
Number of inter-plane transceivers	Q	2
Satisfaction factor	λ	$\{0.85, 0.9, 0.95\}$
Probability of greedy strategy	ξ	0.8
Size of the Mini-batch	$\bar{B}$	1024
Capacity of the experience memory	$Memory$	10000
Lerning rate	α,β	0.0095
Discount factor	γ	0.95
Weight factors	α_1,α_2,α_3	1, 0.1, 1

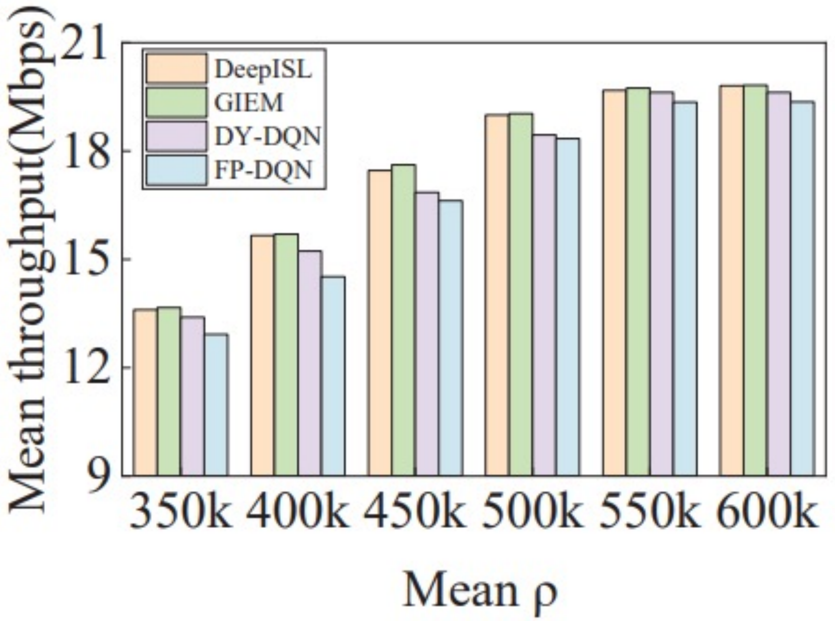
Evaluation Parameter Setting

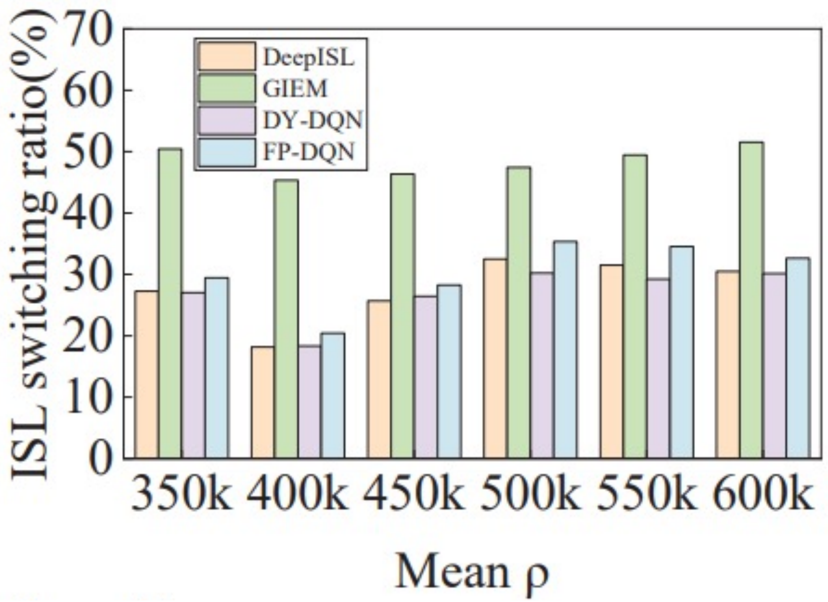


Convergence of DeepISL for training with different amounts of data packets.

3. Evaluation and Results







Energy efficiency

1. DeepISL outperforms others in EE.
2. As traffic load increases, EE decreases.
3. DY-DQN performs close to DeepISL.

Throughput

1. Throughput increases as traffic load increases.
2. DeepISL and GIEM performs close, but better than others.

Switching Ratio

1. GIEM needs maximum switching counts.
2. DeepISL and DY-DQN, FP-DQN performs close.

4. Conclusions

重庆邮电大学
CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

- Discussed the joint problem of dynamic ISL setup with dynamic power allocation.
- Formulated it into a joint optimization problem about target satellite selection and transmission power allocation to maximize energy efficiency and transmission rate with minimum switching costs.
- Solve it using a parameterized deep reinforcement learning method, called *DeepISL*.
- In the future, integrate routing with DeepISL to optimize its end-to-end performance.

Thanks for Your Attention

Email: Luojt@cqupt.edu.cn

Original Paper:

DeepISL: Joint Optimization of LEO Inter-Satellite Link Planning and Power Allocation via Parameterized Deep Reinforcement Learning