Connecting the World from the Skies | ITU Space Sustainability Forum

Organizers | CST هيئة الاتصالات والفضاء والتقنية Communications, Space & Technology Commission | ITU

Carleton University

EQROC:

Hierarchical Reinforcement Learning for Energy- and QoE-Aware Satellite Resource Orchestration

Dhiraj Bhattacharjee · Pablo G. Madoery · Halim Yanikomeroglu

Non-Terrestrial Network (NTN) Lab, Department of Systems and Computer Engineering
Carleton University, Ottawa, Canada

What is innovative

QoE yield as objective
QoE per joule, $\eta_f = \Omega_f / E_f$, steers admission, bandwidth and routing jointly.

Two-tier hierarchical RL
A high-level admission agent sits on a hop-by-hop routing agent.

Router as an oracle
The frozen router previews path quality at each candidate bandwidth before admitting.

Smart rejection
APO training plus congestion-gated, sensitivity-shaped rejection costs.

How it relates to sustainability

Energy-aware by design
Transmit power scales with bandwidth: each decision choice is an energy choice.

Fewer launches
More QoE from satellites already in orbit: less cost, embodied carbon and debris.

Scales with demand
Yield rises 16 % with load while the baselines decline or collapse.

Connectivity for all
Efficient LEO capacity helps 6G bridge the digital divide.

Problem formulation · maximize long-run QoE yield with one joint decision per arriving flow

$$\max_{\alpha_f, B_f, r_f} \lim_{N\to\infty} \frac{1}{N} \sum_{T=1}^{N} \sum_{f \in \mathcal{F}_T} \left[\underbrace{\alpha_f \cdot \frac{\Omega_f}{E_f}}_{\text{QoE yield}} - \underbrace{(1-\alpha_f) \cdot \frac{\mu}{\zeta_f^\tau}}_{\text{rejection penalty}} \right]$$

$$\text{s.t.} \quad \alpha_f \in \{0,1\}, \ \forall f \in \mathcal{F}; \qquad B_f \in \mathcal{B}_f, \quad r_f \in \mathcal{P}(f), \ \forall f \in \mathcal{F} : \alpha_f = 1;$$

$$\sum_{f \in \mathcal{A}_T : \{i,j\} \in r_f} \alpha_f \cdot B_f \ \le \ R_l, \qquad \forall T, \ \forall \{i,j\} \in \mathcal{E}$$

$\alpha_f \in \{0,1\}$: admit or reject · B_f: bandwidth level from the candidate set · r_f: route among the feasible paths · μ/ζ_f^τ: rejection price, higher for throughput-insensitive classes · R_l: link capacity

THE OBJECTIVE

QoE yield

$$\eta_f = \frac{\Omega_f}{E_f}$$

QoE delivered per unit of energy: the sensitivity-weighted QoS score Ω_f of a flow divided by the energy E_f spent on the bandwidth it occupies along its route.

Per arriving flow, decide jointly:

admit? α_f · bandwidth B_f · route r_f

to maximize long-run yield, with a penalty for every rejected flow

$\Omega_f = \zeta_f^\delta \Omega_f^\delta + \zeta_f^\tau \Omega_f^\tau + \zeta_f^l \Omega_f^l$ sensitivity-weighted latency / throughput / drop-rate score

$E_f \propto B_f \cdot |r_f|$ energy per flow: allocated bandwidth and route hop count as proxy

How EQROC works · one arriving flow, two learned decisions

Incoming flow
class · QoE bounds · max requested bandwidth

arrives

High-level admission agent
average-reward PPO (APO) · per flow

CANDIDATE BANDWIDTH LEVELS: L1 L2 L3 L4 …

each level is previewed through the routing policy: feasibility, spare capacity, utilisation, hop count

decides: **admit** at level b* · or **reject**

admit

Admitted
routed at level b*

Rejected
forced or proactive

query levels ↓ · ↑ route preview

Low-level routing policy
episodic PPO · one hop at a time

2-hop look-ahead · latency budget · masked hops

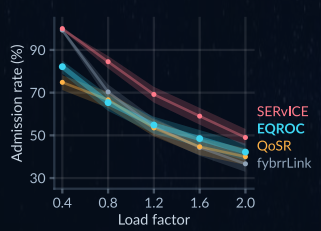
frozen after training · queried as an oracle



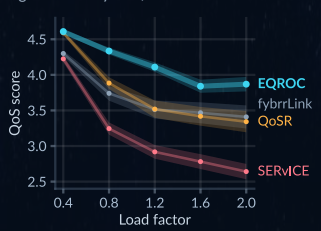
route built hop by hop by the frozen low-level policy; congested links are avoided thanks to the two-hop look-ahead

Results · 4×4 LEO grid · 10 Gbps ISLs · 7 application types · load factor 0.4–2.0 · mean ± 95 % CI · baselines [1]–[3]

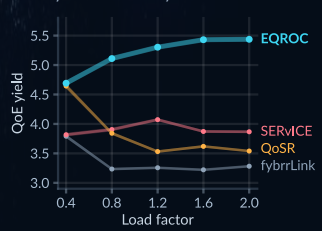
Admission rate
fewer flows admitted: 42 % vs 49 % at load 2.0



QoS Score of admitted flows
highest at every load, +14 % at load 2.0



QoE yield
the only scheme whose yield rises with load



+41 %
QoE yield vs the best baseline at load 2.0

+16 %
EQROC yield from load 0.4 to 2.0, while QoSR drops 24 %

+14 %
QoS score of admitted flows vs the best baseline

Contact: NTN Lab · Carleton University

Dhiraj Bhattacharjee
dhirajbhattacharjee@sce.carleton.ca

REFERENCES

[1] P. Kumar et al., "fybrrLink: efficient QoS-aware routing in SDN enabled future satellite networks," IEEE TNSM, 2022.

[2] T. Li et al., "SERVICE: a software defined framework for integrated space-terrestrial satellite communication," IEEE TMC, 2018.

[3] Y. Huang et al., "An efficient differentiated routing scheme for MEO/LEO-based multi-layer satellite networks," IEEE TNSE, 2024.

[4] X. Ma et al., "Average-reward reinforcement learning with trust region methods," IJCAI, 2021.

[5] Y. Yang et al., "Towards energy-efficient routing in satellite networks," IEEE JSAC, 2016.

[6] D. Bhattacharjee et al., "DSROQ: dynamic scheduling and routing for QoE management in LEO satellite networks," IEEE GLOBECOM, 2025.