© ETSI 2020. All rights reserved



ENI PoC #17: Intelligent Satellite-Terrestrial Integration Network Architecture Progress Update

Rapporteur: (Tsinghua University) Zhiyuan Lin

Co-Rapporteur: China Telecom, Asiainfo, Huawei, CAICT, CNIT, CNR ISTI

Spatial-temporal uneven distribution of satellite traffic



© ETSI 2020

[1] I. del Portillo, B.G. Cameron, E.F. Crawley, "A technical comparison of three low earth orbit satellite constellation systems to provide global broadband," in Acta Astronautica, vol. 159, pp. 123-135, Jun. 2019.

Beam hopping can achieve on-demand coverage



Some satellite systems with phased array antennas

Figure C.3a: Demanded and served traffic in a conventional system

DVB standard supports BH

© ETSI 2020 [1] Digital Video Broadcasting (DVB); Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Servicess, News Gathering and Other Broadband Satellite Applications; Part 2: DVB-S2 Extensions (DVB-S2X), ETSI Std., Aug. 2020.

System Model: Traffic Driven BH scheduling



□ Traffic Driven BH scheduling

Design BH patterns and beam bandwidth allocation slot by slot relying on real-time status of traffic queues and channel states.



- □ Challenges
- > Real-time scheduling:
- Scheduling execution time should be less than timeslot duration (ms)
- > Multi-dimensional scheduling:
- Multi-beam time-space-frequency three-dimensional decision-making

Modeled as Markov decision processing (MDP)



- State s: traffic queue information of different cells : ϕ_n^t
- Action **a**: beam location, beam bandwidth
- Reward r: defined as optimization target:

$$C = \alpha \frac{TR_total}{TR_{max}} - (1 - \alpha) \frac{Fairness}{Fairness_{max}}$$

© ETSI 202 • How to get the optimal action: multi-agent deep reinforcement learning

Multi-agent DRL BH Architecture



- ✓ Each Beam is regarded as two types of agents
 - Beam irradiation agent: beam location
 - Beam bandwidth agent: beam bandwidth
- $\checkmark\,$ Each agent has small action space

Global State:

• At each time slot, the system can observe the whole traffic demand as the global state and each agent can obtain this state.

ETS

□ Action Space:

- The BH pattern and bandwidth allocation for K beams.
- **D** Reward:
- The data throughput and delay fairness is adopted as the immediate reward.
- All agents share a global reward which can evaluate the global performance and promote collaboration between agents.

DDQN Training:

• Since the action spaces are discrete, the well-performed Double Deep Q-Learning (DDQN) method is adopted for the agent to learn the policy.



