SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Semantic-aware coordinated transmission in cohesive clustered satellites: Utility of information perspective

Liang Xu[1], Jian Jiao[1,2*], Shiyao Jiang[1], Ye Wang[2] & Qinyu Zhang[1,2*]

[1]*Harbin Institute of Technology (Shenzhen), Shenzhen* 518055*, China;*
[2]*Pengcheng Laboratory, Shenzhen* 518055*, China*

Appendix A The simulation parameters of the SCT scheme

The simulation parameters of the SCT scheme are presented in Table A1.

Table A1 Simulation Parameters.

Parameters	**Value**
Transmit Power	100 mW
Transmit Distance	600 km
Transmit Frequency	30 GHz
Transmit Antenna Gain	20 dBi
Satellite Antenna Gain	43.3 dBi
Thermal Noise Density	-174 dBm/Hz
Average Power of Scatter Component b_0 in SR Channel [1]	0.063
Average Power of LOS Component Ω in SR Channel [1]	0.000897
Nakagami-m Parameter m in SR Channel [1]	1
SNR Region	$[0, 20]$ dB
Sliding Window Size	[5,100] time slots
M	6
$\mathscr{C}$	3
B	4
$\mathcal{V}_t$	$[[1,0,0,1],[1,1,0,0],[1,1,0,0],[0,0,1,1],[0,0,1,1],[1,1,0,0]]$
Capacity of Experience Memory	1000
Discount Factor	0.95
Update Target Network Interval	50
Batch size	32
Learning rate	10^{-3}
D3QL Network	LSTM and fully connected layer

Appendix B The training procedure of MAD3QL algorithm and related simulation results

The proposed MAD3QL algorithm is based on the Value Decomposition Network (VDN) framework [2], therefore deployed in a Centralized Training and Decentralized Execution (CTDE) manner. Specifically, in the training process at the GS, the environment can obtain complete observations of the multi-agent state to train the model. Note that each satellite (agent) can only access to local observed information and part of feedback information from GS to infer other agents' states or channel states, and then make distributed decisions.

Moreover, in the M3DQL algorithm, S_i uses individual Q-function $\mathcal{Q}_i$ to compute the total Q-value, denoted by $Q_G(\mathcal{S}_t, \mathscr{A}_t) = \sum_{i \in \mathbb{M}} \mathcal{Q}_i(s_t^i, a_t)$ so as to train the network parameters, denoted by $\mathbb{P} = \{\mathcal{P}_i | i \in \mathbb{M}\}$. At time slot t, each satellite gets an action based on individual observation and acts to obtain a reward r_t^i. The transitions $(q_t^i, \mathscr{A}_t^i, r_t^i, q_{t+1}^i)$ are restored into a replay buffer. The controller periodically samples a random mini-batch transitions from the replay buffer to update $\mathcal{P}_i$ of each satellite. Above all, the parameters update process at the GS can be expressed as [3]:

$$\mathbb{P}_{t+1} = \mathbb{P}_t + \delta(\dagger_t - Q_G(\mathcal{S}_t, \mathscr{A}_t; \mathbb{P}_t))\nabla_{\mathbb{P}_t}, \tag{B1}$$

* Corresponding author (email: jiaojian@hit.edu.cn, zqy@hit.edu.cn)

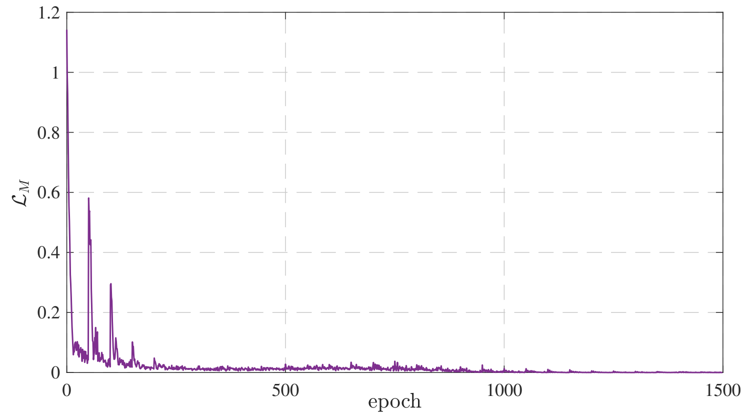

Figure B1 The training MSE loss versus the number of epochs.

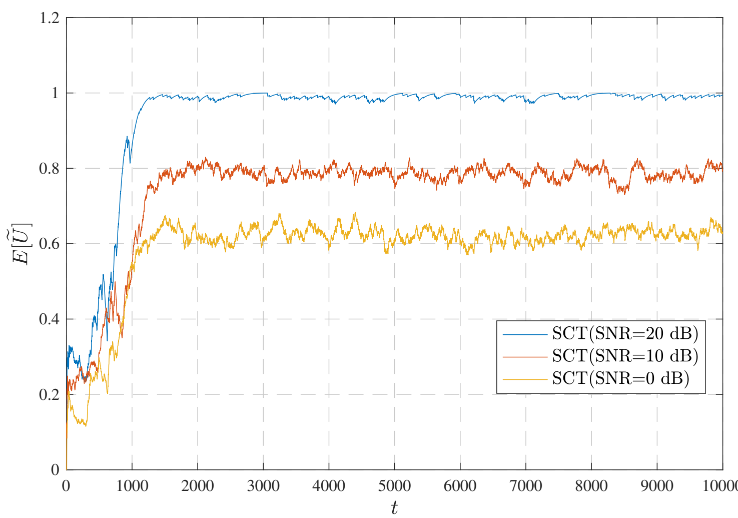

Figure B2 The UoI per unit energy consumption versus t.

where $\dagger_t$ is the target value determined by the dueling and double mechanisms.

Once the scheduling policies are trained successfully, the trained model can be updated to the clustered satellites and perform autonomous semantic-aware coordinated transmission without feedback on their individual state information to the GS. However, when the relevance between SBI of clustered satellites changed a lot, it is necessary for the GS to periodically retrain the model and update it to the clustered satellites to guarantee the long-term average UoI per unit energy consumption.

In the training process, we adopt the mean-squared error (MSE) as the loss function $\mathcal{L}_M(\mathbb{P})$ and the Adam optimizer with an initial learning rate of 10^{-3} and weight decay 10^{-3} to reduce averaged MSE loss. The MSE loss of the training process is presented in Figure B1. It shows that the MSE loss gradually decreases and converges after about 1000 epochs, where the fluctuation is caused by periodical random sampling in the training process.

In addition, we present $\mathbb{E}[\widetilde{U}]$ performance under different channel conditions in Figure B2. It can be observed that the reduction of signal-to-noise ratio (SNR) can only reduce $\mathbb{E}[\widetilde{U}]$, and will not affect the convergence performance of semantic-aware MAD3QL algorithm, which verifies the robustness of the SCT scheme to perform semantic-aware coordinated transmission in dynamic channel conditions.

References

1 Huang J, Jiao J, Wang Y, et al. Age-critical long erasure coding-CCSDS file delivery protocol for dual-hop S-IoT. IEEE Internet Things J, 2023, 10: 17070-17084
2 Sunehag P, Lever G, et al. Value-decomposition networks for cooperative multi-agent learning. arXiv preprint, arXiv:2309.10185, 2023
3 Mazandarani H, Shokrnezhad M, Taleb T. A semantic-aware multiple access scheme for distributed, dynamic 6G-based applications. arXiv preprint, arXiv:2401.06308, 2024