

Reidentify: Context-Aware Identity Generation for Contextual Multi-Agent Reinforcement Learning

Zhiwei Xu · Kun Hu · Xin Xin · Weiliang Meng · Yiwei Shi ·

Hangyu Mao · Bin Zhang · Dapeng Li · Jiangjin Yin

Shandong University

National University of Defense Technology

Institute of Automation, Chinese Academy of Sciences University of Bristol

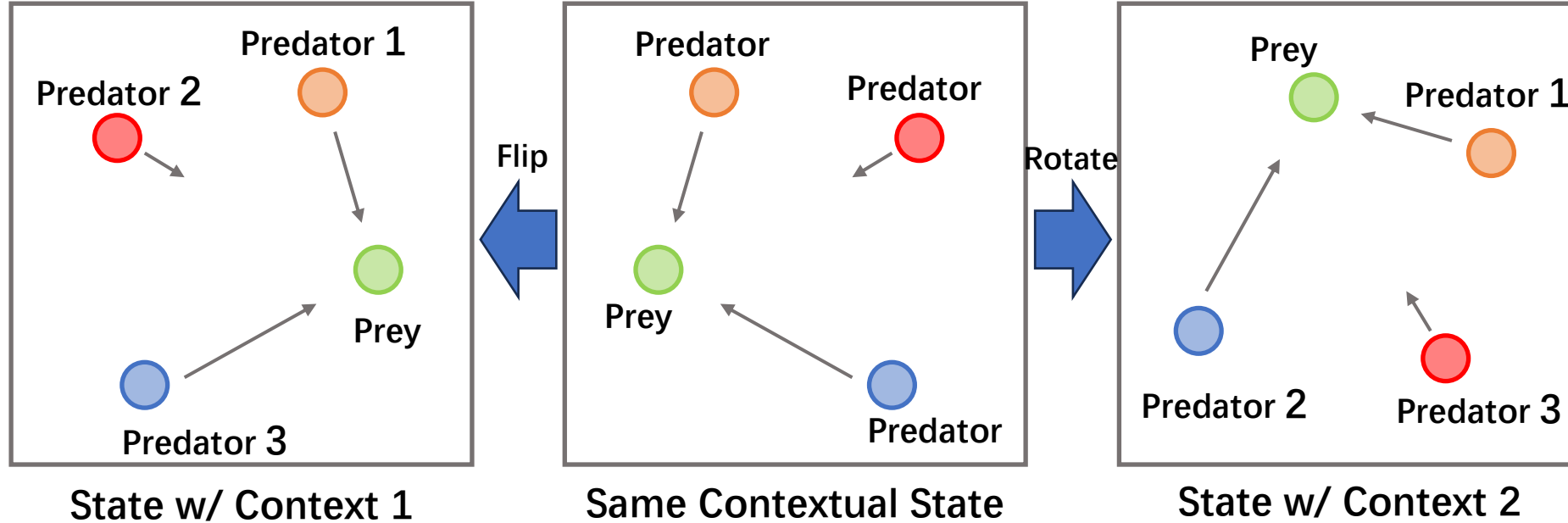
Kuaishou Technology

Huazhong Agricultural University



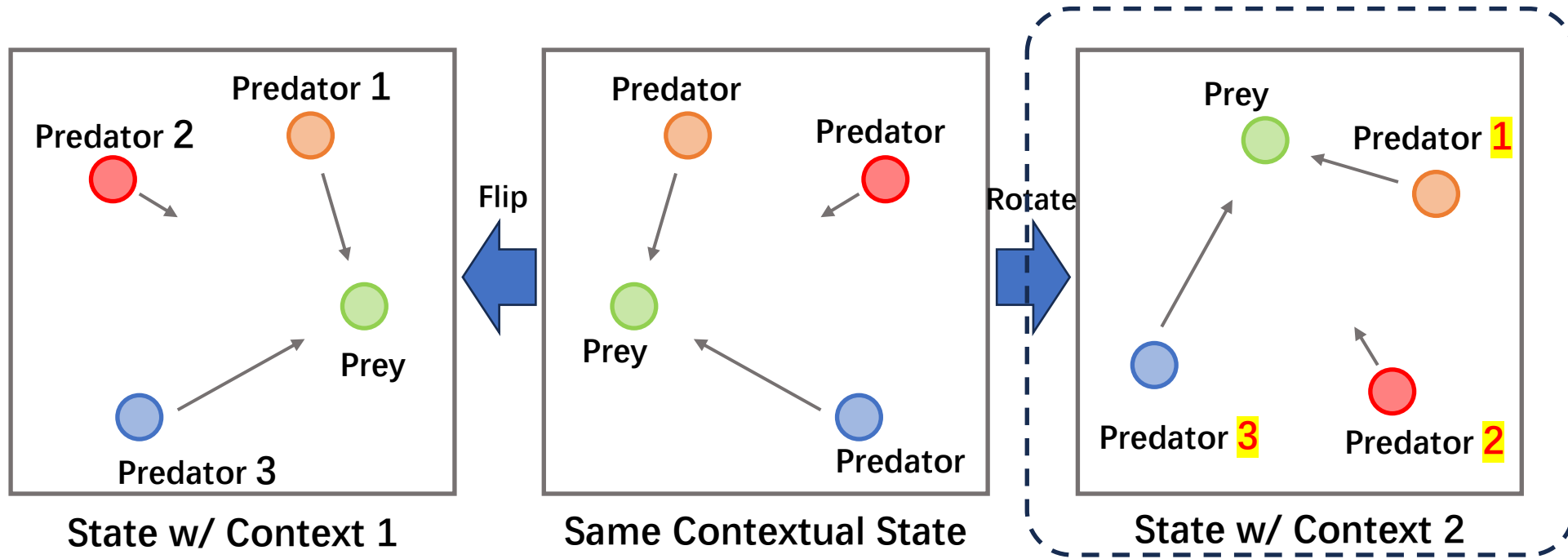
ICML
International Conference
On Machine Learning

Motivation: Generalization in CMARL is Hard



- Real-world MARL faces task variations across deployments
- Policies fail to generalize due to misaligned agent roles
- Current methods lack consistent identity alignment

Key Idea: Identity as Context

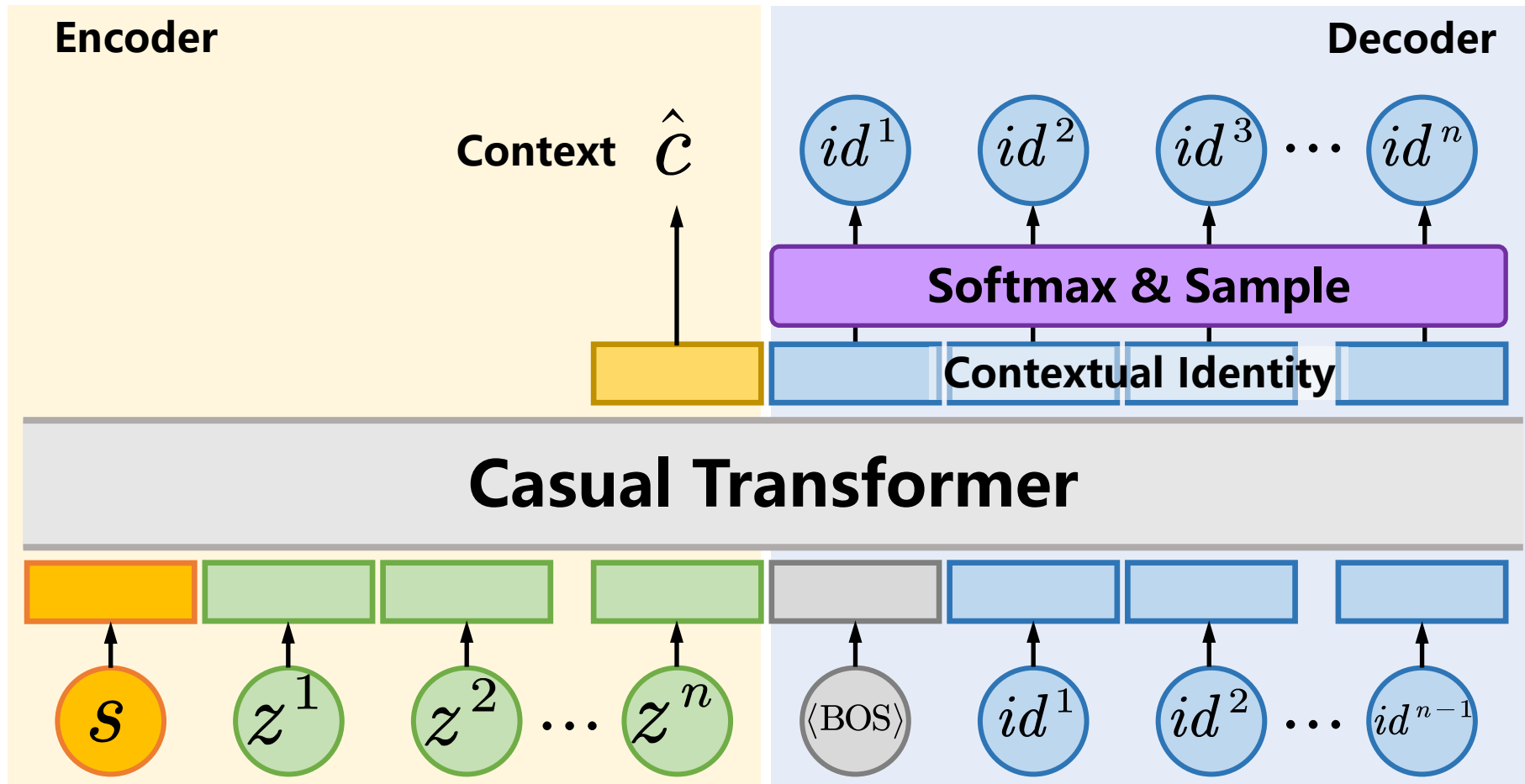


- So we introduce **CAID**: Context-Aware Identity Generation
- Dynamically assign unique identities based on context
- Enable policy reuse across task variants

CAID Framework Overview

- Contextual State Encoder
- Agent Identity Decoder
- Action Regulator

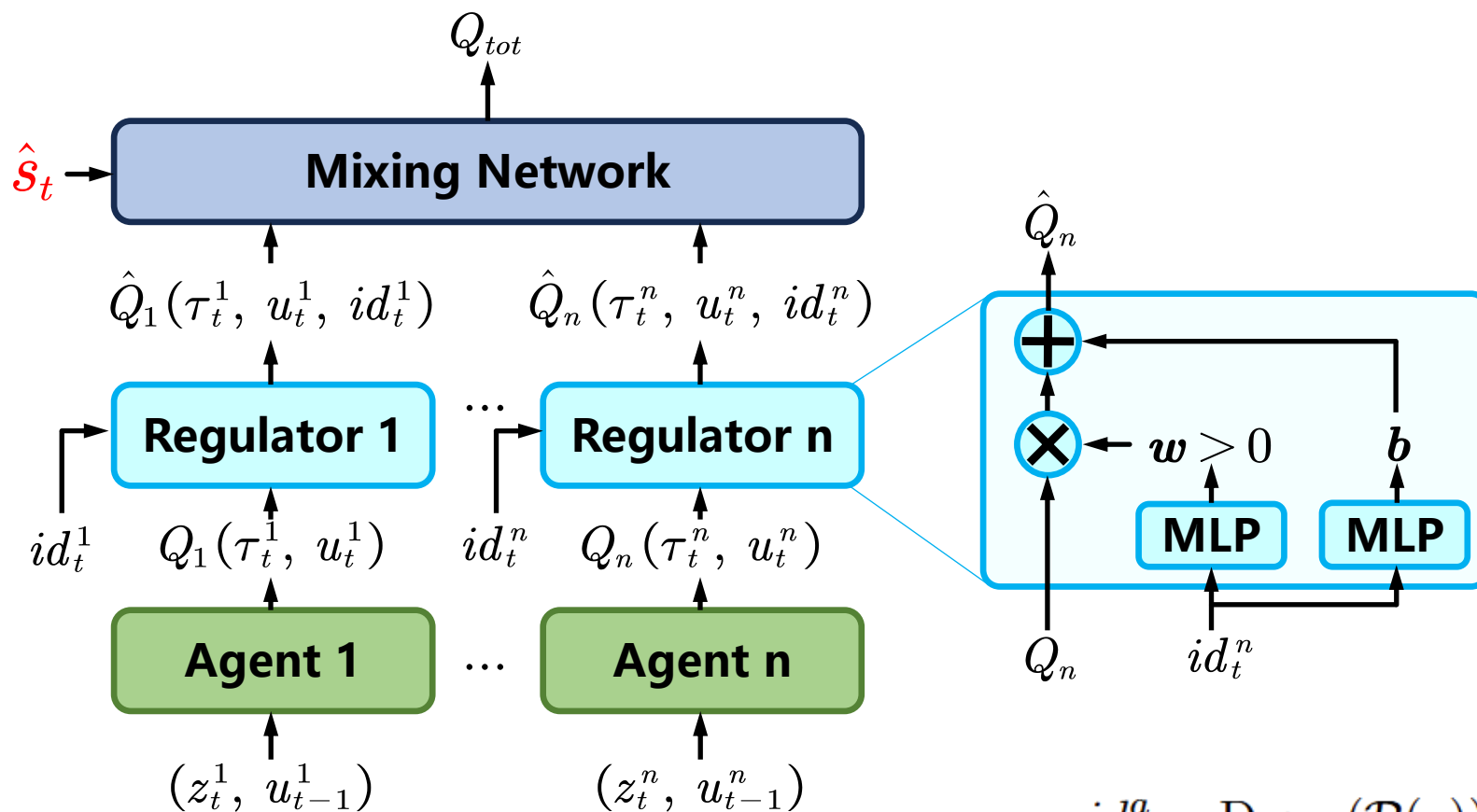
Contextual State Encoder & Agent Identity Decoder



- Input: Global state and agent observations
- Output: Low-dimensional context embedding $\hat{c}$ and ids for agents

Noted: Sampling-based identity assignment avoids collapse

Action Regulator & End-to-End Optimization



➤ Loss Function:

$$\mathcal{L} = (y_{tot} - Q_{tot}(\tau, u, \hat{s}))^2$$

➤ Straight-Through Gradients

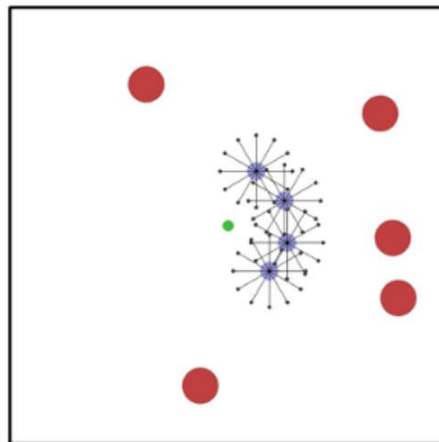
$$id^a = \text{Draw}(\mathcal{P}(a)) + \mathcal{P}(a) - \text{StopGrad}(\mathcal{P}(a)), \forall a \in \mathcal{A}$$

- Uses identity to transform Q-values via hypernetwork
- Ensures alignment under the **IGM** principle

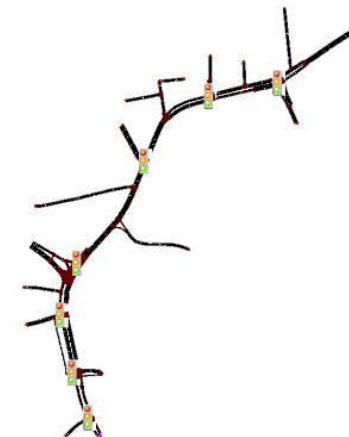
Experiment



➤ SMACv2

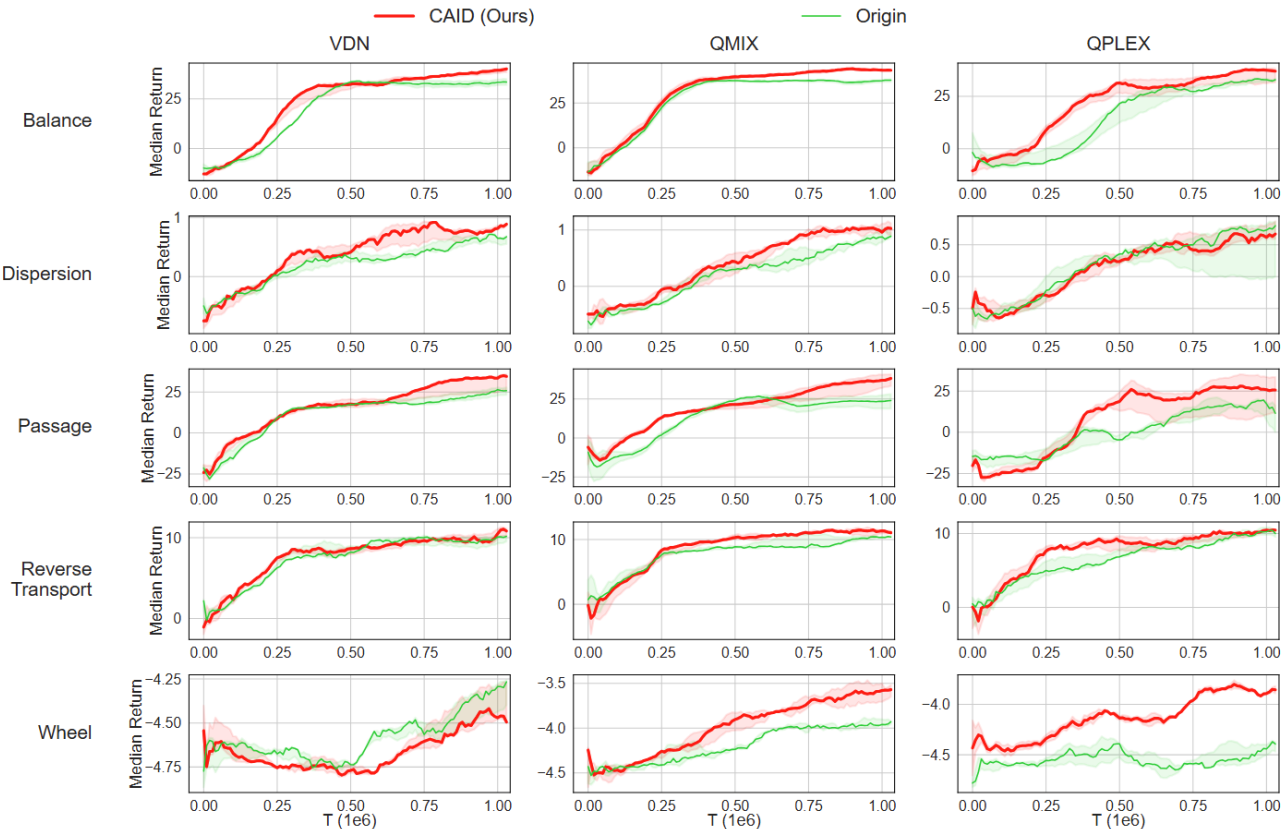
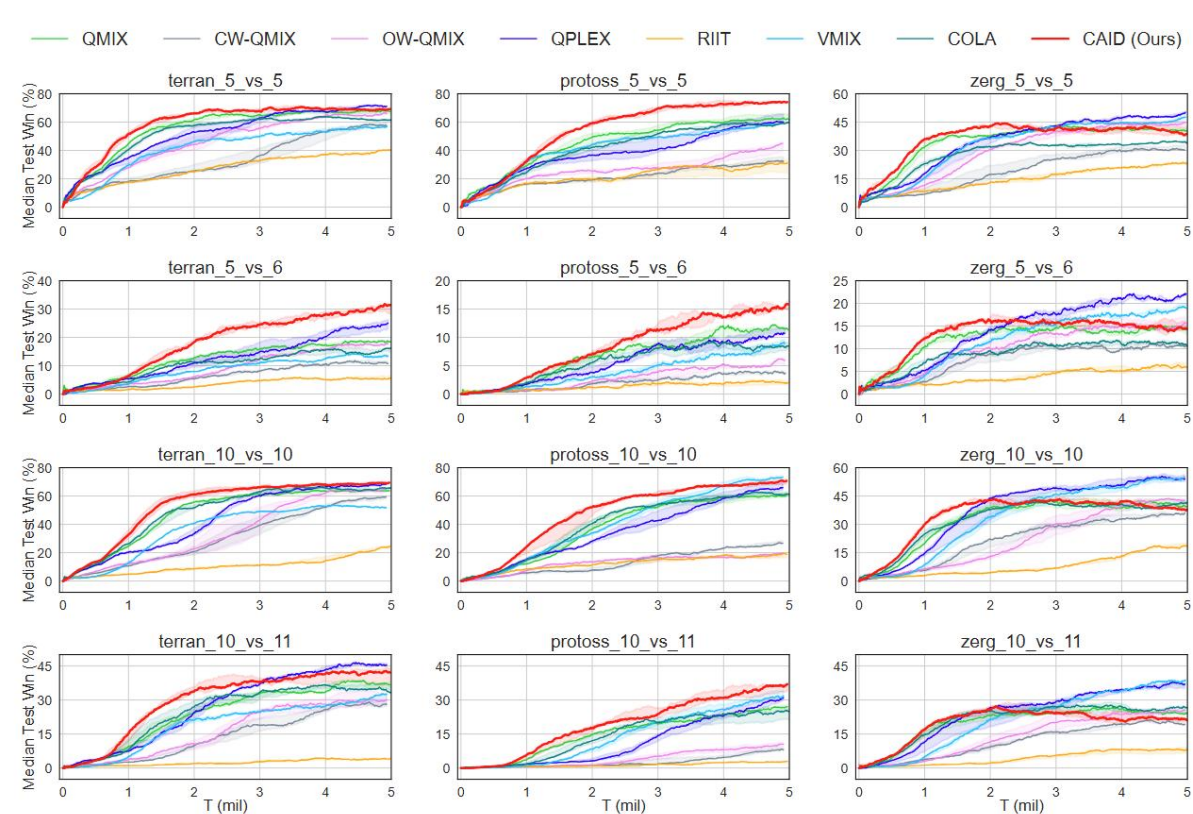


➤ VMAS



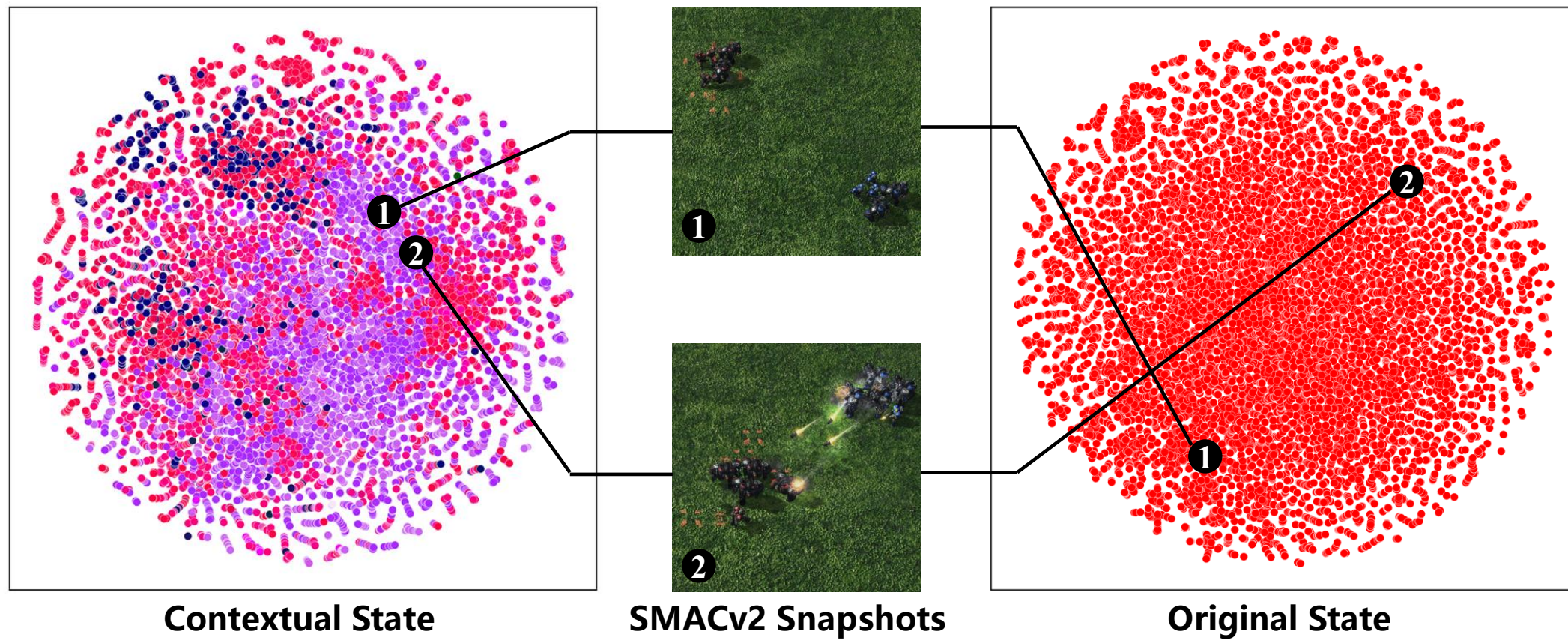
➤ PyTsc

Main Results



Algos	Jinan			Hangzhou			New York		
	Queue	Delay	Travel Time	Queue	Delay	Travel Time	Queue	Delay	Travel Time
CAID	463.39	0.471	329.28	146.50	0.651	336.52	479.48	0.901	369.05
w/o CS	476.52	0.473	326.86	155.79	0.653	340.00	488.39	0.908	372.13
w/o AI	663.59	0.498	364.05	171.37	0.659	346.77	480.41	0.902	368.38
w/o ST	490.70	0.482	332.34	150.11	0.662	330.65	487.28	0.903	371.38
QMIX	563.13	0.507	338.15	170.58	0.660	349.25	496.08	0.902	381.57

Visualization



Thanks for listening!