# Robust Cooperative Multi-agent Reinforcement Learning via Multi-view Message Certification

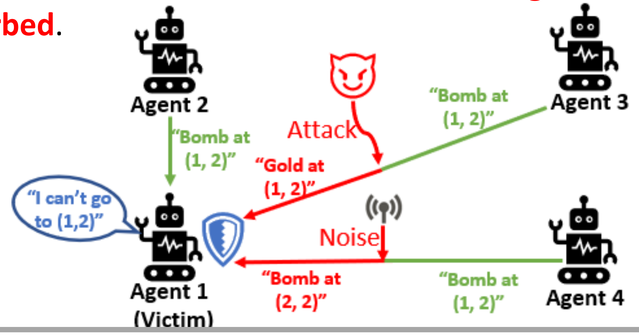
Lei Yuan, Tao Jiang, Lihe Li, Feng Chen, Zongzhang Zhang, Yang Yu
Nanjing University, Polixir Technologies
Science China Information Sciences (SCIS)

## Background & Motivation

**Message-adversarial** Decentralized Partially Observable Markov Decision Process **under Communication**:

$$\mathcal{M} = \langle \mathcal{N}, S, A, P, \{O_i\}_{i=1}^n, \Omega, R, \gamma, M, V \rangle$$

We can make agents learn a robust policy via **Multi-view Message Certification** so that the agents can still choose the optimal action when the **received messages are perturbed**.




## Method

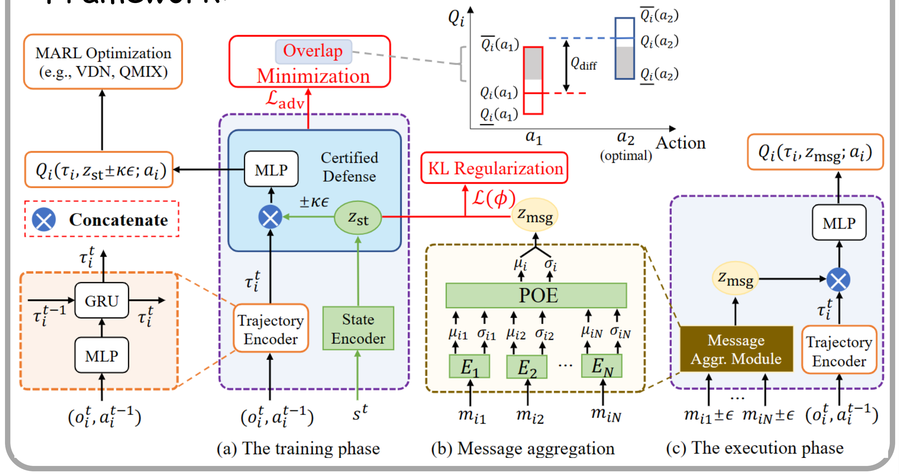
### Certified Defense Module:

- **Centralized training phase:** train the state representation **robust against perturbations by minimizing overlap of Q-values,** distill knowledge to the joint message representation obtained by **Multi-view VAE**.
- **Decentralized execution phase:** encode local observation history and aggregate messages to execute a robust policy.

### Message aggregation Module:

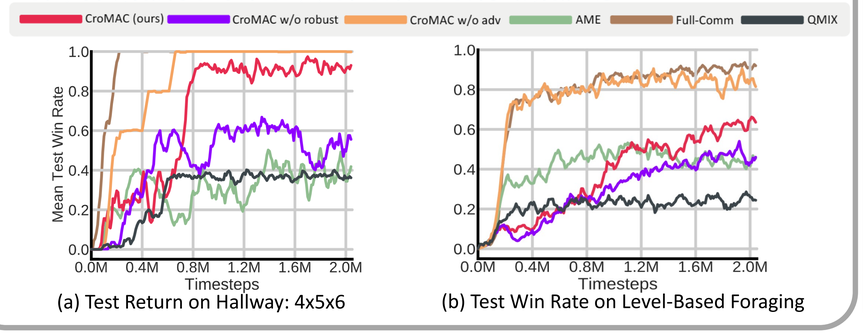
- To aggregate information from different views and calculate the quantitative relationship between $z_{\text{st}}$ and $z_{\text{msg}}$.

### Framework:

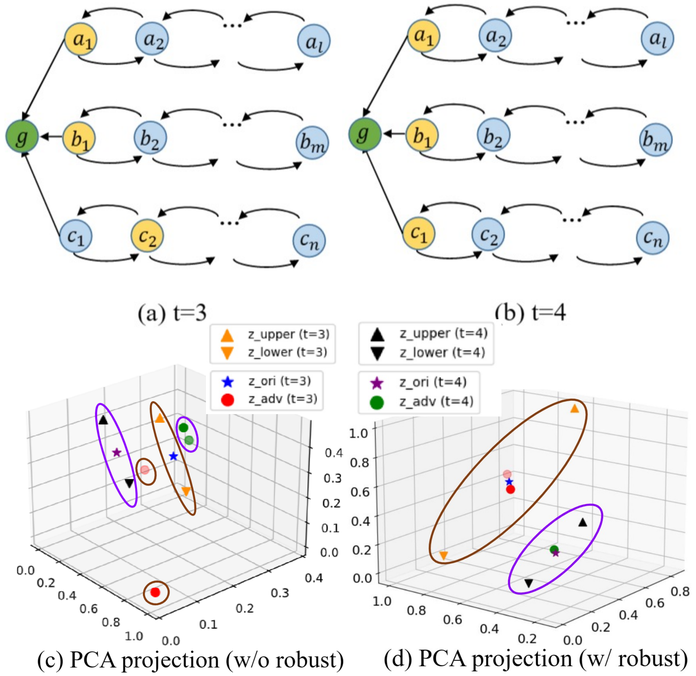



## Experiments

### Performance on different benchmarks




## Visualization Analysis




- When suffering from perturbations, the joint message representation **jumps out of the normal range** and the agents choose unexpected actions.
- With our robustness scheme, the joint message representation can **be bounded in a reasonable range**, leading to a robust action selection.