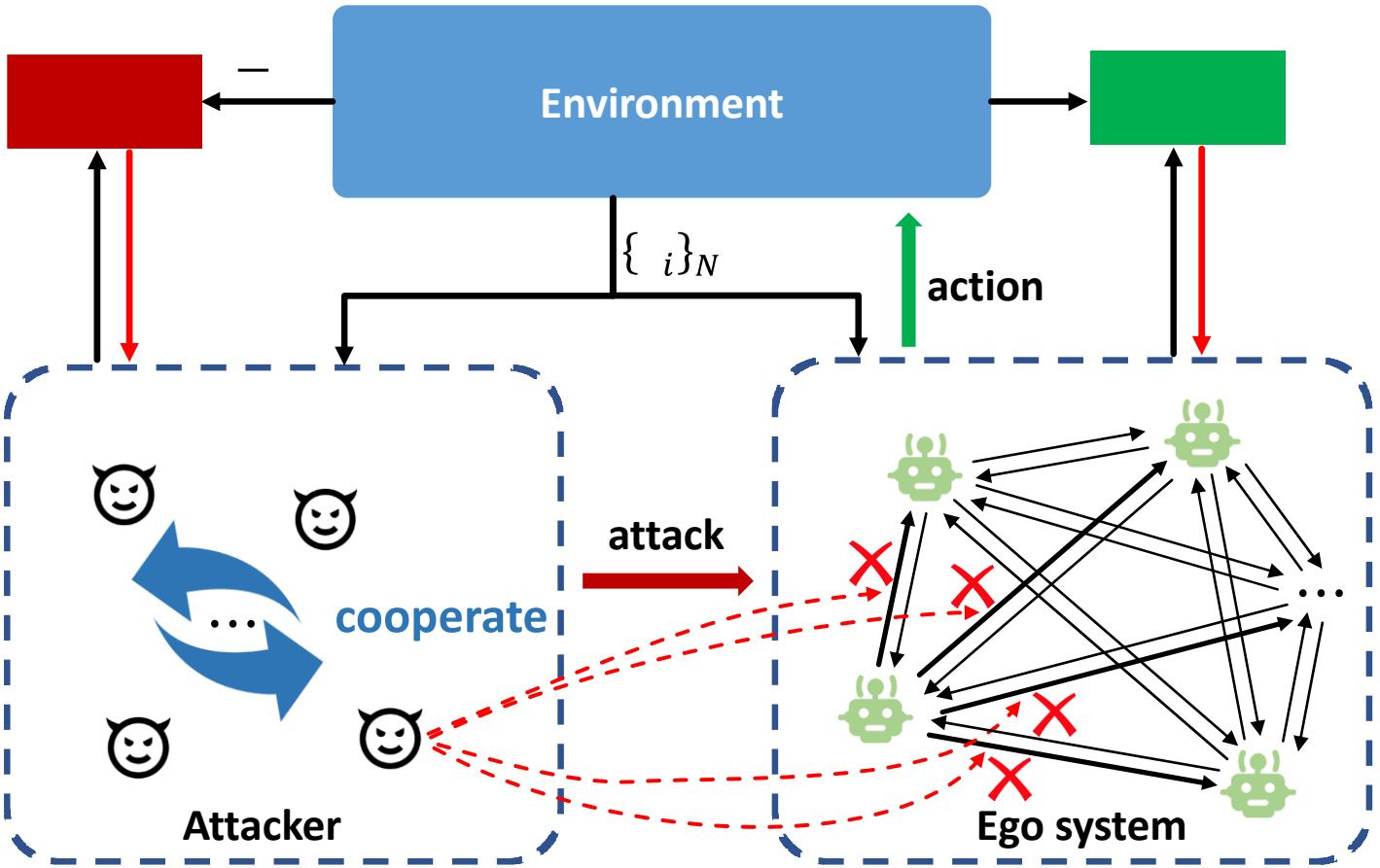
Communication-Robust Multi-Agent Learning by Adaptable Auxiliary Multi-Agent Adversary Generation

Lei YUAN, Feng CHEN, Zongzhang ZHANG, Yang YU

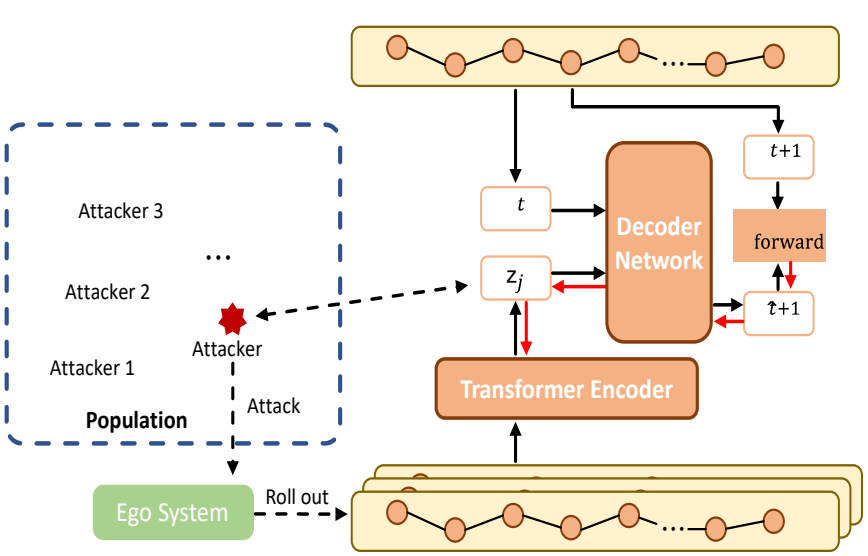
Frontiers of Computer Science, DOI: 10.1007/s11704-023-2733-5

Problems & Ideas

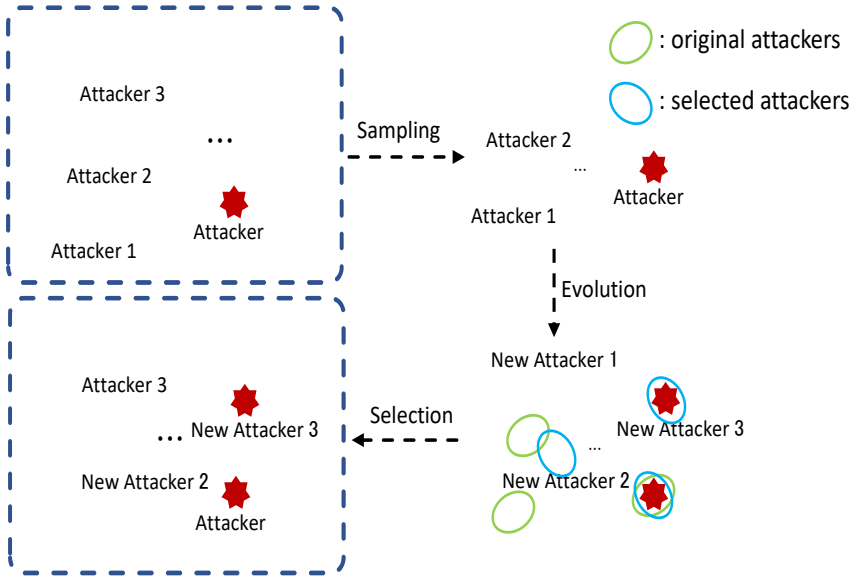


We train an auxiliary to attack every messaging channel at a different degree at any time, and the ego system is adversarially trained with the attacker population. The solid black arrows indicate the direction of data flow, the red solid ones indicate the direction of gradient flow, and the red dotted ones mean the attack actions from the attacker onto specific communication channels.

Main Contributions



(a) Utilizing trajectory representation as identification for attacker instance.



(b) The process of the population updating.

(a) An encoder-decoder architecture to learn the trajectory representation, which is then used to represent different attackers. The black solid arrows indicate the direction *of data flow, and the red solid ones imply the direction of gradient flow. (b) The attcker population is updated based on evolutionary learning. The* locations *of points imply the distances of representations, and the color shades indicate the attack ability, i.e., the attackers corresponding to deeper points* are stronger *attackers. For example, new attacker 3 is accepted as it is distant enough from other attackers, and the oldest attacker 1 is removed.*

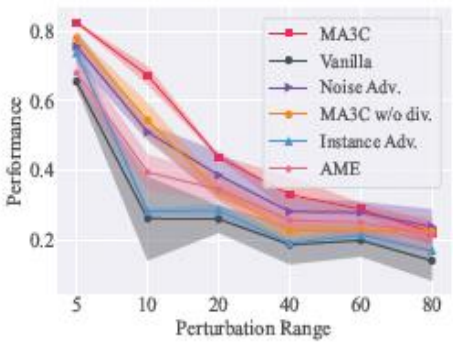
Main Results

Table 1 Performance comparison under different attack modes.

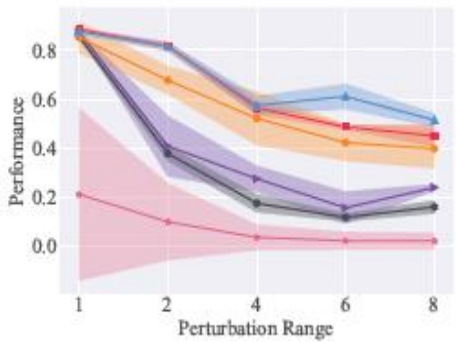
		Hallway-6x6	Hallway-4x5x9	SMAC-1o_2r_vs_4r	SMAC-1o_10b_vs_1r	GP-4r	GP-9r
Normal	MA3C	0.94±0.05	0.97±0.05	0.86±0.02	0.62±0.01	0.87±0.02	0.82±0.01
	Vanilla	**1.00**±0.00	**1.00**±0.00	0.81±0.06	**0.63**±0.04	**0.88**±0.03	0.82±0.02
	Noise Adv.	**1.00**±0.00	0.99±0.01	**0.88**±0.04	0.6±0.05	**0.88**±0.03	**0.85**±0.02
	MA3C w/o div.	0.98±0.02	0.66±0.46	0.86±0.02	0.62±0.03	0.86±0.09	0.81±0.03
	Instance Adv.	0.52±0.48	0.67±0.47	0.84±0.02	0.57±0.04	0.86±0.03	0.82±0.03
	AME	**1.00**±0.00	0.98±0.02	0.81±0.05	0.60±0.01	0.23±0.37	0.00±0.00
Random Noise	MA3C	0.91±0.07	0.79±0.18	**0.87**±0.01	**0.67**±0.03	0.88±0.01	0.80±0.07
	Vanilla	0.58±0.03	0.53±0.06	0.73±0.07	0.60±0.02	0.86±0.03	0.79±0.02
	Noise Adv.	**0.97**±0.02	**1.00**±0.00	0.82±0.02	0.56±0.02	0.88±0.01	**0.82**±0.01
	MA3C w/o div.	0.68±0.07	0.68±0.29	0.73±0.07	0.53±0.01	0.82±0.06	0.80±0.07
	Instance Adv.	0.56±0.34	0.67±0.47	0.79±0.07	0.60±0.08	**0.90**±0.03	0.81±0.02
	AME	0.61±0.06	0.79±0.03	0.71±0.13	0.59±0.08	0.22±0.37	0.00±0.00
Aggressive Attackers	MA3C	**0.91**±0.22	**0.98**±0.01	**0.67**±0.03	**0.62**±0.03	**0.81**±0.02	**0.76**±0.03
	Vanilla	0.09±0.19	0.00±0.00	0.26±0.12	0.57±0.03	0.38±0.02	0.30±0.05
	Noise Adv.	0.61±0.37	0.13±0.14	0.51±0.02	0.54±0.03	0.41±0.13	0.48±0.11
	MA3C w/o div.	0.57±0.39	0.96±0.03	0.54±0.05	0.61±0.02	0.68±0.06	0.71±0.01
	Instance Adv.	0.63±0.42	0.88±0.14	0.28±0.01	0.61±0.04	**0.81**±0.02	**0.76**±0.03
	AME	0.13±0.03	0.00±0.00	0.39±0.05	0.59±0.07	0.10±0.16	0.00±0.00

Our method MA3C *outperforms multiple baselines in different benchmarks under different communication conditions.*

Main Results

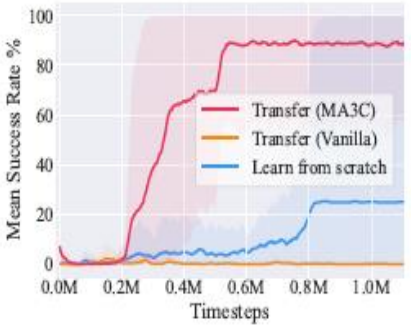


(a) SMAC-1o_2r_vs_4r

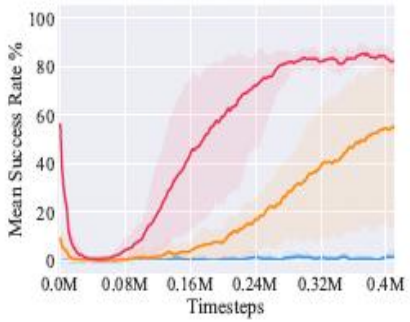


(b) GP-4r

Generalization Ability



(a) Hallway-4x5x9



(b) GP-4r

Transfer Ability

Our method MA3C enjoys high *generalization* ability to different perturbation ranges and could promote the learning phase for new tasks.