

# 面向多智能体协作的注意力意图与交流学习方法

俞文武, 杨晓亚, 李海昌, 王瑞, 胡晓惠

自动化学报, 2021, 47:1-16.

联系人: 俞文武 联系方式: 15652326966

(wenwu2016@iscas.ac.cn)

## 背景

对于部分可观测环境下的多智能体交流协作任务, 现有工作大多只利用了当前时刻的网络隐藏层信息, 限制了信息来源. 面临的挑战有:

- 复杂度: 随着智能体的数量指数级增长;
- 局部信息: 只能访问其本地的观察结果;
- 非平稳性: 个体最优策略会随着其它智能体的行为策略的改变而变化。

## 方法

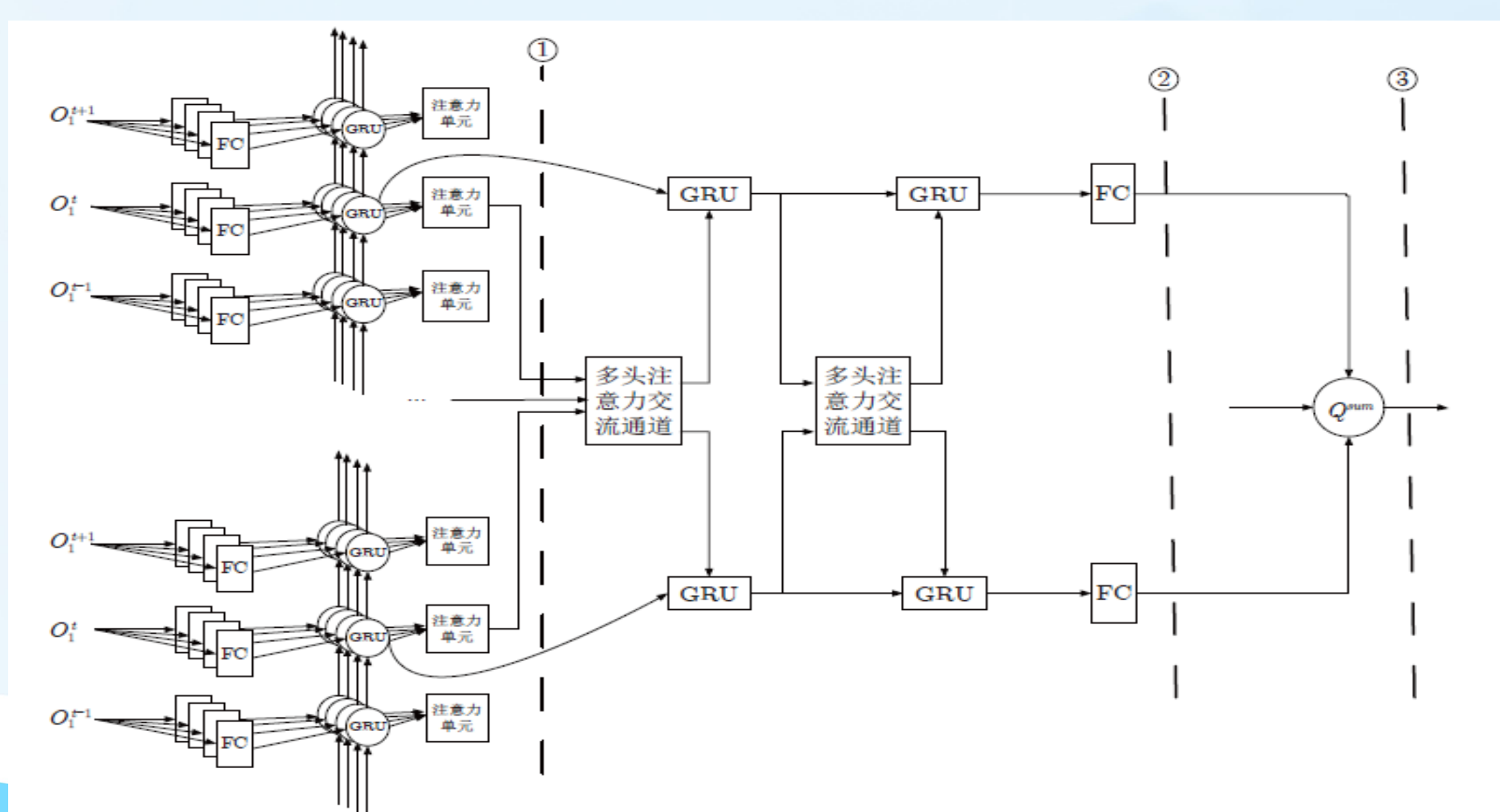
针对部分可观测的多智能体合作任务, 基于值函数分解系列的方法, 本文采用集中式训练, 分布式执行的训练框架, 提出了面向部分可观测合作问题的多智能体注意力意图交流学习算法 (Multi-Agent Attentional Intention and Communication, MAAIC):

- 增加额外的公共意图网络, 保存其他智能体的历史时刻最优的策略, 为智能体的决策提供了额外的信息来源;
- 采用注意力机制对意图信息与交流信息的进行处理, 从而能够提取到更为有效的信息。

## 算法框架

- 信息处理模块:

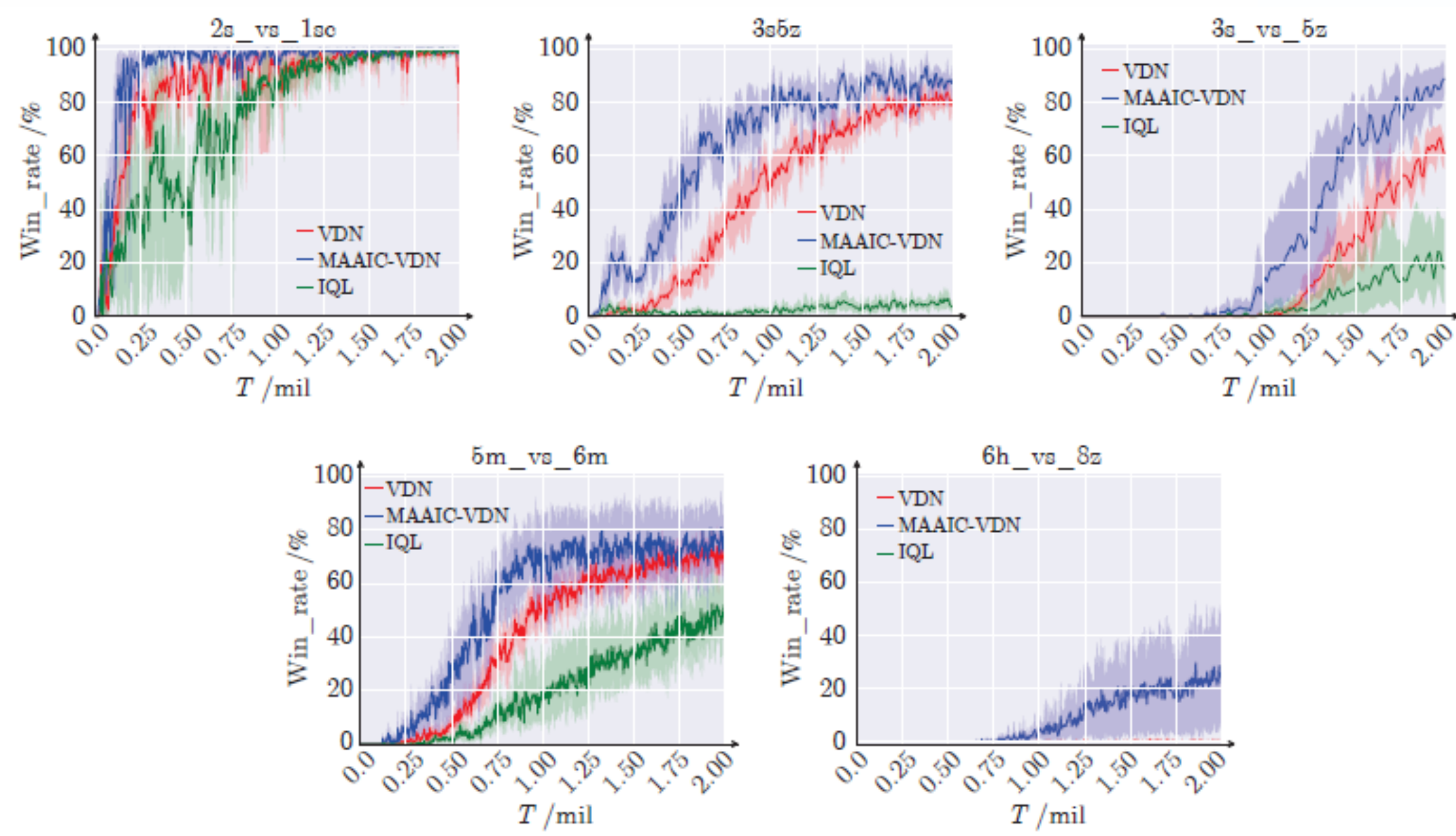
增加公共意图网络与注意力机制;



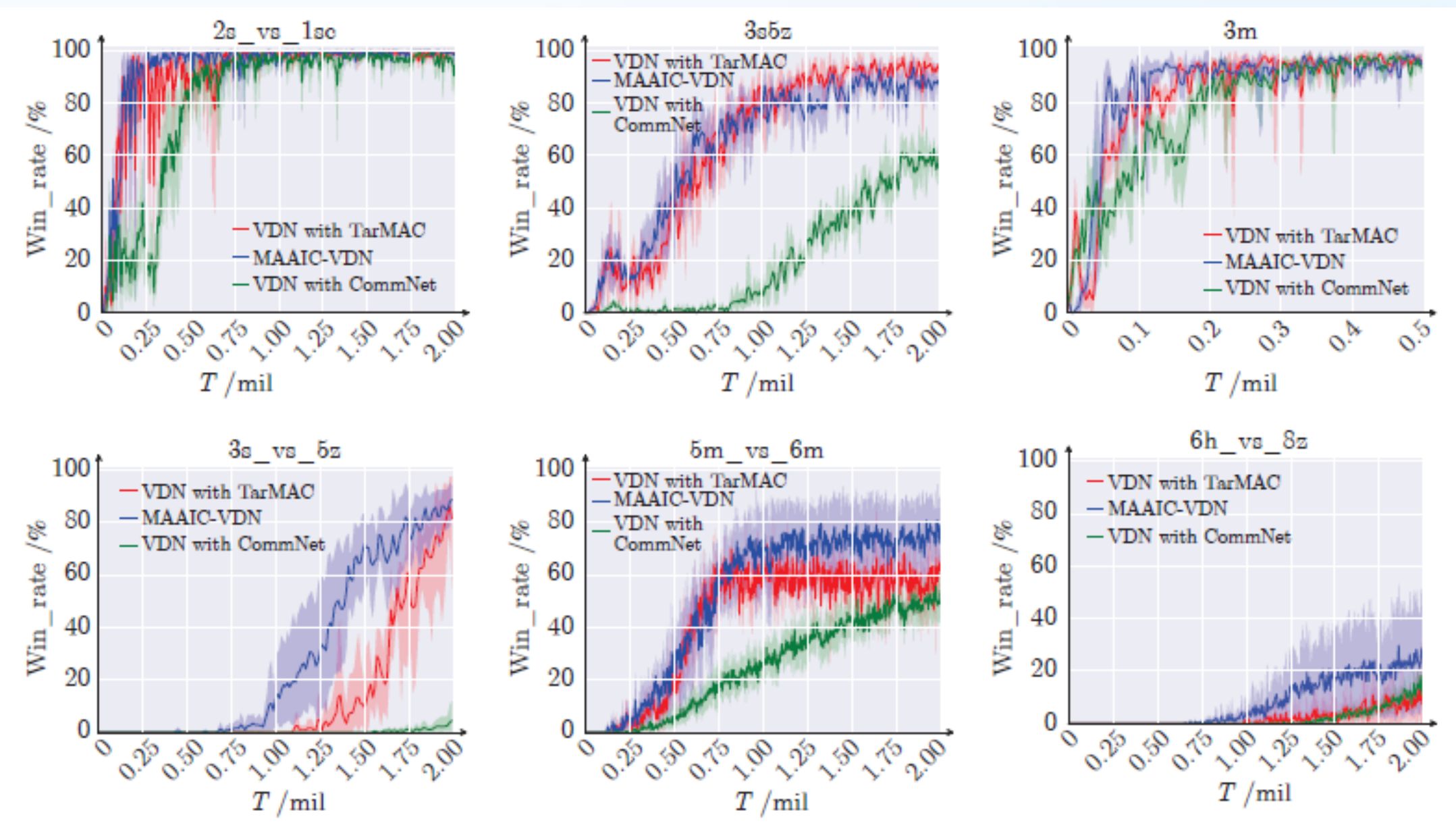
- 交流模块: 智能体之间增加多头注意力机制的交流通道; GRU结构单元;
- 整合模块: 得到整体Q 值;

## 实验

- MAAIC-VDN与VDN比较性能分析



- 对交流结构的消融实验分析



- 不同算法测试过程中最终性能

场景	MAAIC-VDN	VDN	IQL	MAAIC-QMIX	QMIX	Heuristic	MAAIC-QTRAN	QTRAN
2a_vs_1cc	100	100	100	100	100	0	100	100
3o5z	90	87	9	97	91	42	31	20
5m_vs_6m	87	78	59	74	75	0	67	68
3a_vs_5z	98	73	46	98	97	0	97	15
6h_vs_8z	55	0	0	31	3	0	22	0

## 总结

- 提出了一种意图存储机制, 引入额外的公共网络来保存历史表现最好的策略网络, 以此来建模其他智能体的意图信息.
- 引入交流模块, 采用多头注意力机制来提高整个网络的表示能力. 本文的多智能体意图交流算法可以充分利用过往的局部最优网络的信息, 扩大了信息的来源渠道, 不再局限当前时刻策略网络的信息.