人工智慧作業報告簡報

信息工程學院 2021 級 干皓丞 2101212850

A game-theoretic analysis of networked system control for common-pool resource management using multi-agent reinforcement learning

使用多智能體強化學習進行公共池資源管理的網絡系統控制的博弈論分析

(NeurIPS) conference, 2020

https://arxiv.org/abs/2010.07777

Motivation 动机

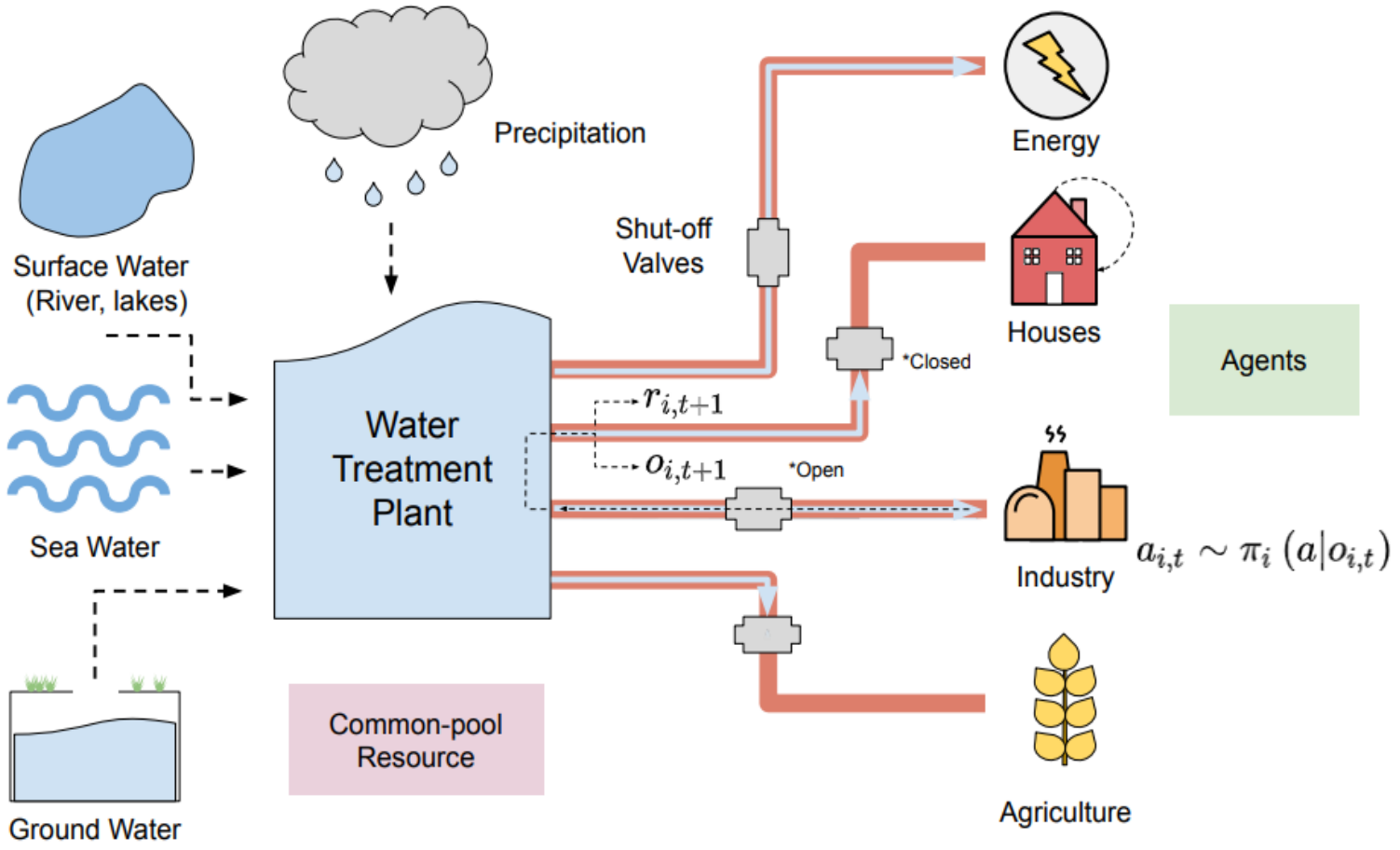
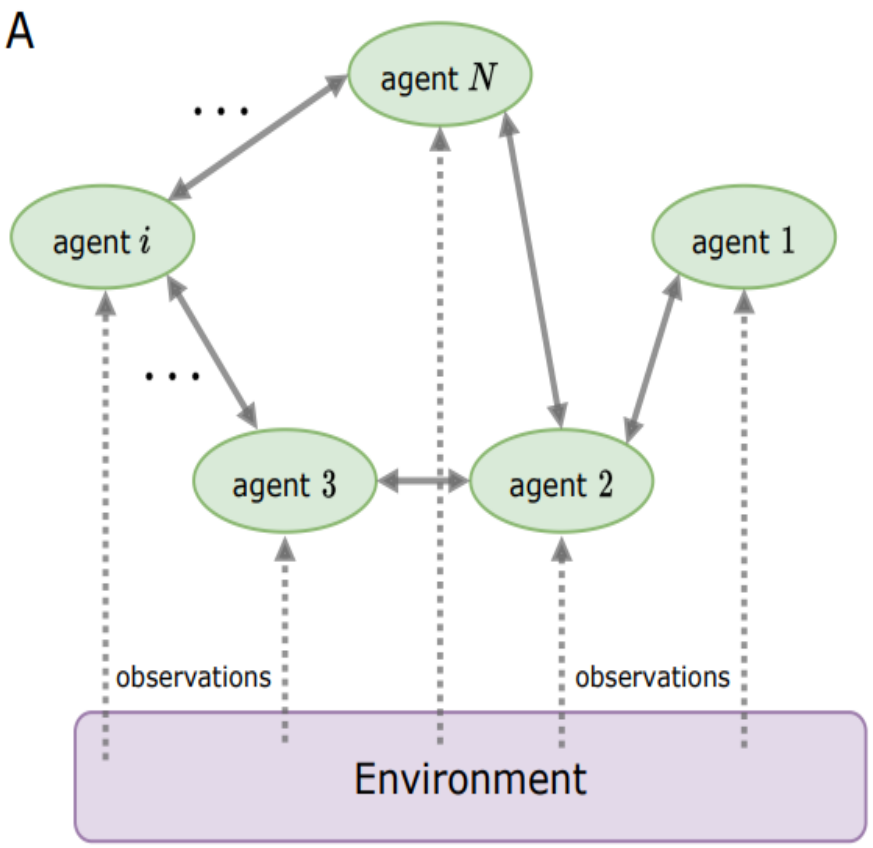
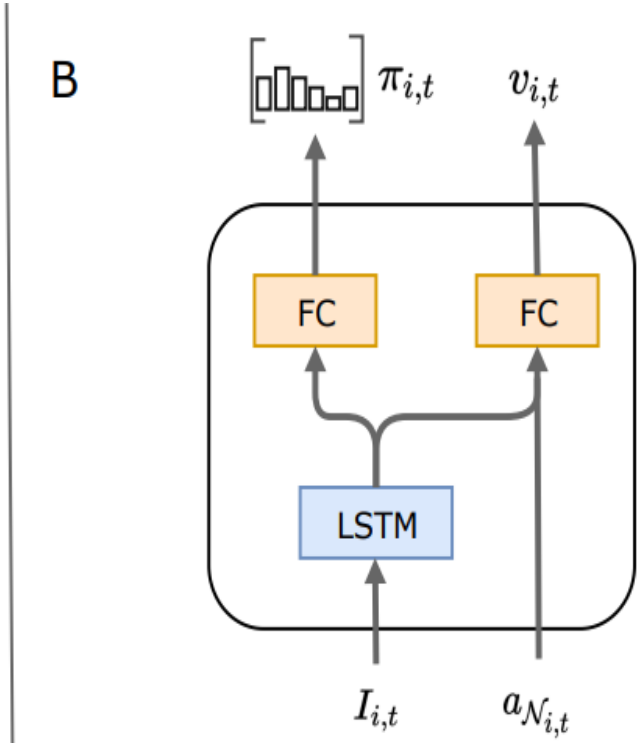


Figure 1: *Water management system as a common-pool resource environment for networked multi-agent reinforcement learning.* Shut-off valve controllers are treated as agents in the system and are responsible for providing water to different sectors of society, such as industry and agriculture.

Intuition 直觉

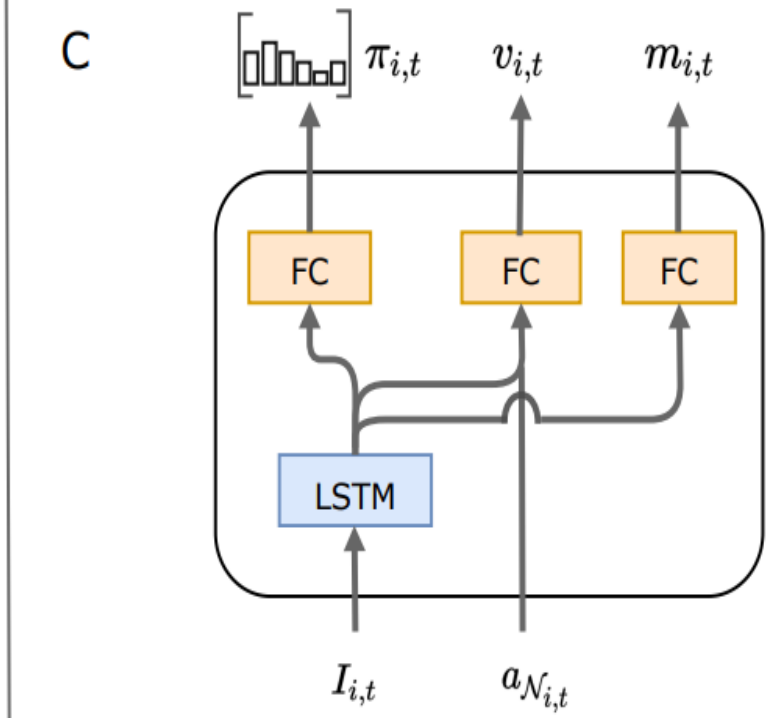




NA2C: $I_{i,t} = \{o_{\nu_{i,t}}\}$

FPrint: $I_{i,t} = \{o_{\nu_{i,t}}, \pi_{\mathcal{N}_{i,t-1}}\}$

ConseNet: $I_{i,t} = \{o_{\nu_{i,t}}\}$



DIAL: $I_{i,t} = \{o_{\nu_{i,t}}, m_{\mathcal{N}_{i,t}}, a_{i,t-1}\}$

CommNet: $I_{i,t} = \{o_{\nu_{i,t}}, \overline{m}_{\mathcal{N}_{i,t}}\}$

NeurComm: $I_{i,t} = \{o_{\nu_{i,t}}, \pi_{\mathcal{N}_{i,t-1}}, m_{\mathcal{N}_{i,t}}\}$

Justification 理由

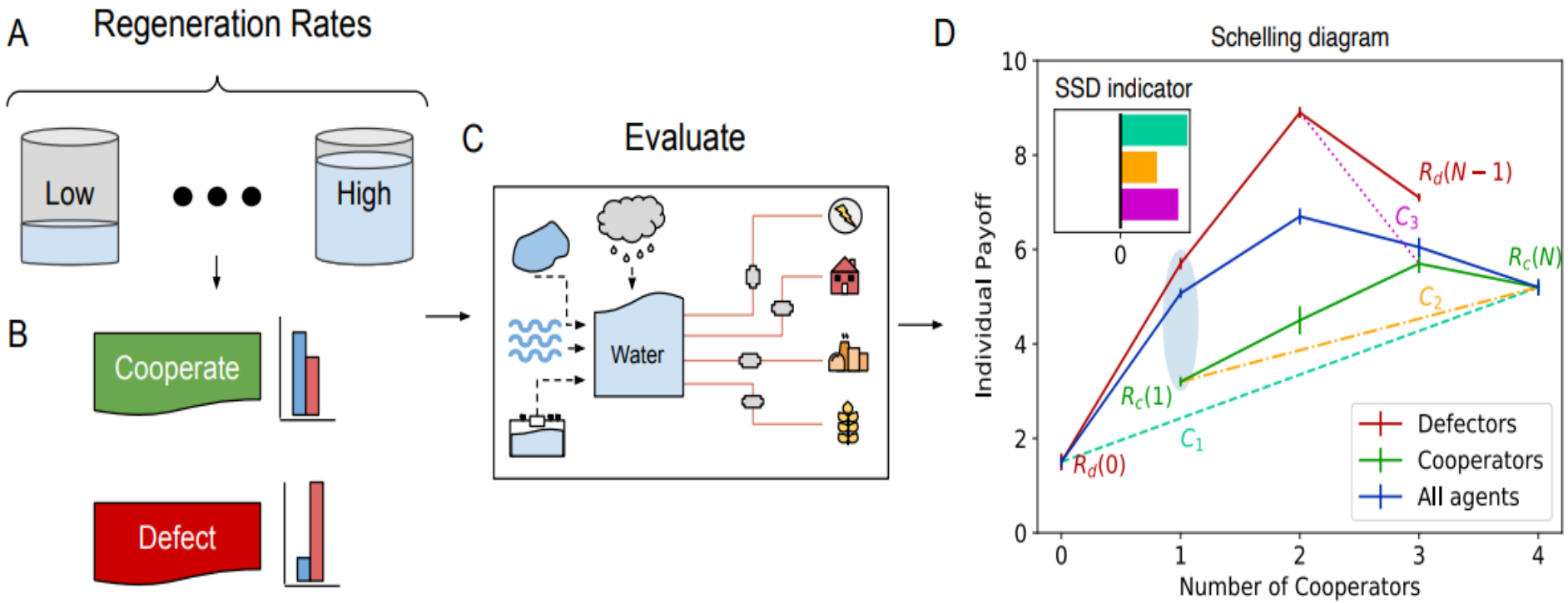


Figure 3: *Empirical game-theoretic analysis pipeline*. **(A)** Agents train on a common resource pool under different regeneration rates, where they may act greedily or show restraint. **(B)** Cooperative policies show restraint, while defecting policies act greedily. **(C)** Evaluate sampled policies on the environment. **(D)** Plot the Schelling diagram, which can be used to visually classify systems.

Framework 框架

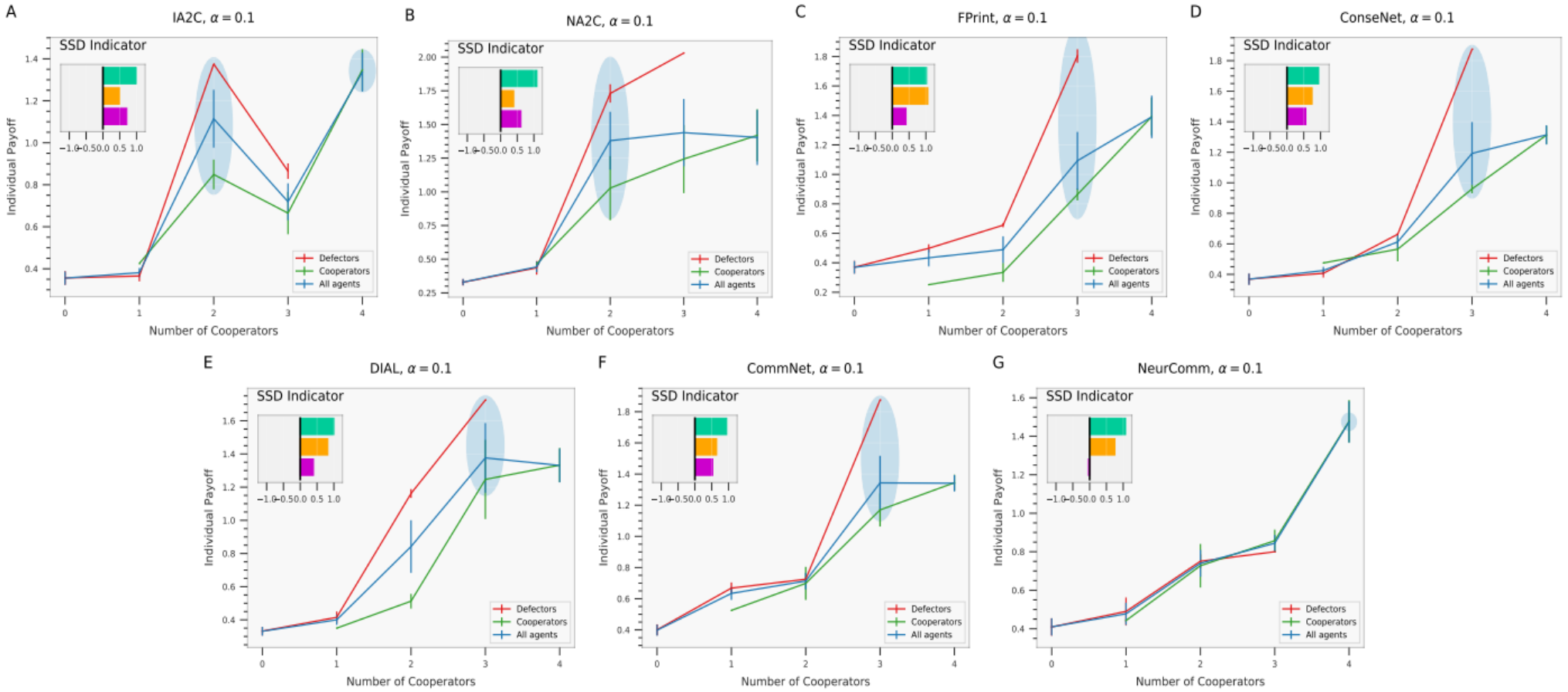


Figure 5: *EGTA for networked system control, with* $\alpha = 0.1$.**(A-G)** Schelling diagrams for each approach with sequential social dilemma (SSD) indicators given as insets. Potential equilibria are shaded in blue.

Result 结果

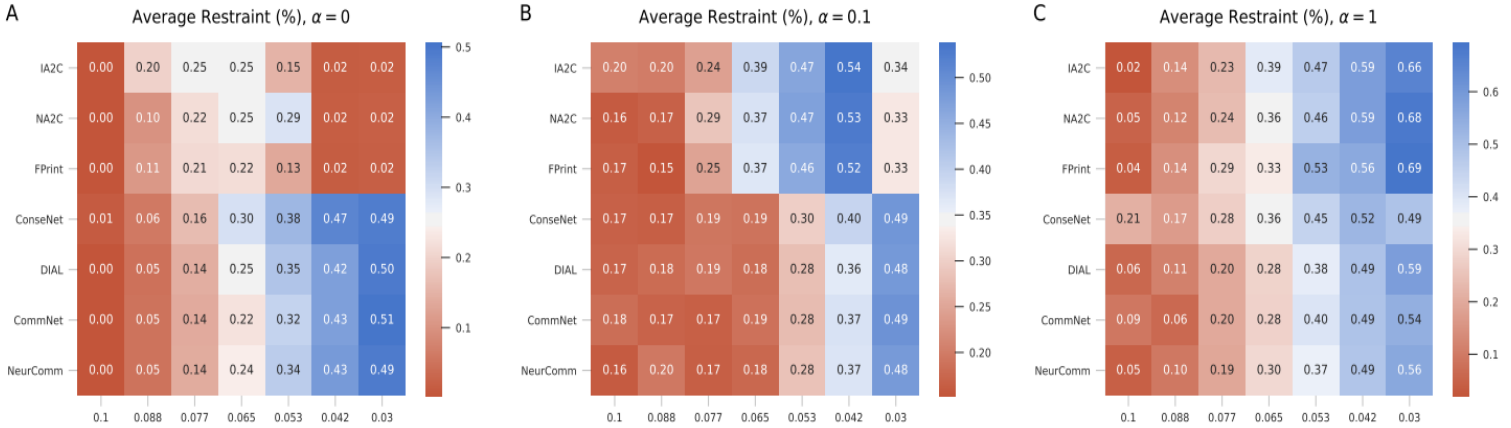


Figure 7: Heatmaps of average restraint percentage as a function of the regeneration rate for different MARL algorithms, from high (0.1) to low (0.03). **(A)** $\alpha = 0$, **(B)** $\alpha = 0.1$, **(C)** $\alpha = 1$.

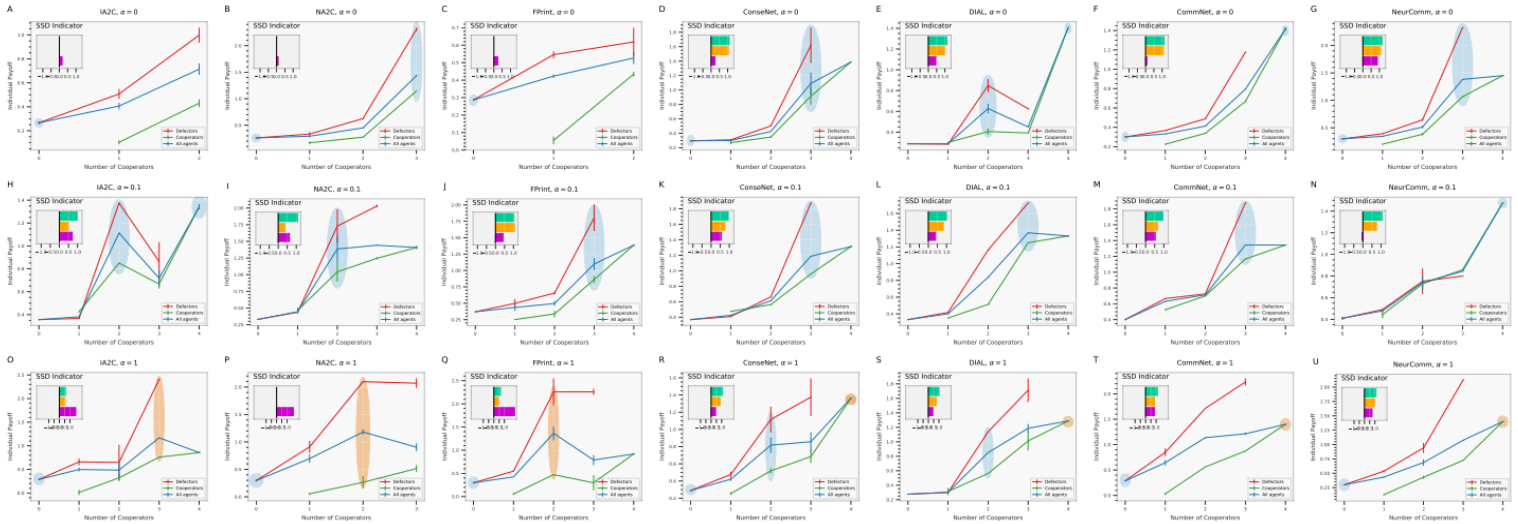


Figure 8: Schelling diagrams for each approach with network system sequential social dilemma (SSD) indicators given as insets. Potential Nash equilibria are shaded in blue. **Top row (A-G)** $\alpha = 0$, **Middle row (H-N)** $\alpha = 0.1$, **Bottom row (O-U)** $\alpha = 1$. Here we include orange shaded regions indicating configurations corresponding to the highest average payoff for all connected agents.