Learning to Control Network Contagion
Dylan Quigley, Diego Benalcazar, Chinwendu Enyioha, Mubarak Shah

Abstract:
• Distributed, decentralized resource allocation strategy to stop a contagion’s spread.
• Makes use of the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm [Lowe et al.].

Approach:
• Train agents using reinforcement learning to control the epidemic’s spread by allocating vaccines and antidotes.

Nonlinear system
\[ \frac{dp}{dt} = (BA - D)p(t) - P(t)BAp(t) \]

Conclusions:
• Changes to reward function and MADDPG structure result in strong performance.

Agent structure of the MADDPG algorithm, with each agent calculating its action \( a \) using its policy \( \pi \) and observations \( o \). During training, agents are given access to more information about other agents’ policies through their Q-value function \( Q \).

Final reward function used with infection and recovery rates \( \beta \) and \( \delta \) for a node, the total cost \( C(x) \), budget \( b \), and state information \( x \).

References:

This graph shows that the stability quickly converges to roughly -0.3; this value being negative indicates that the epidemic’s spread is being reduced.

This graph shows that the total cost (the green line) stays well within the budget that the nodes are under (the orange line).