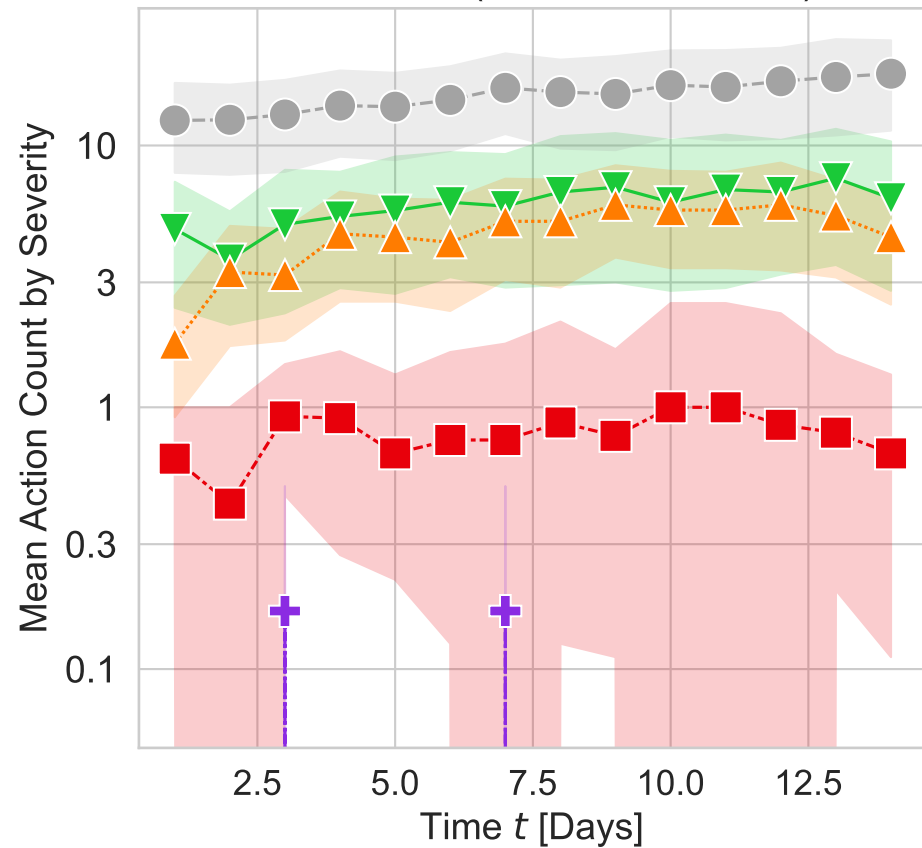
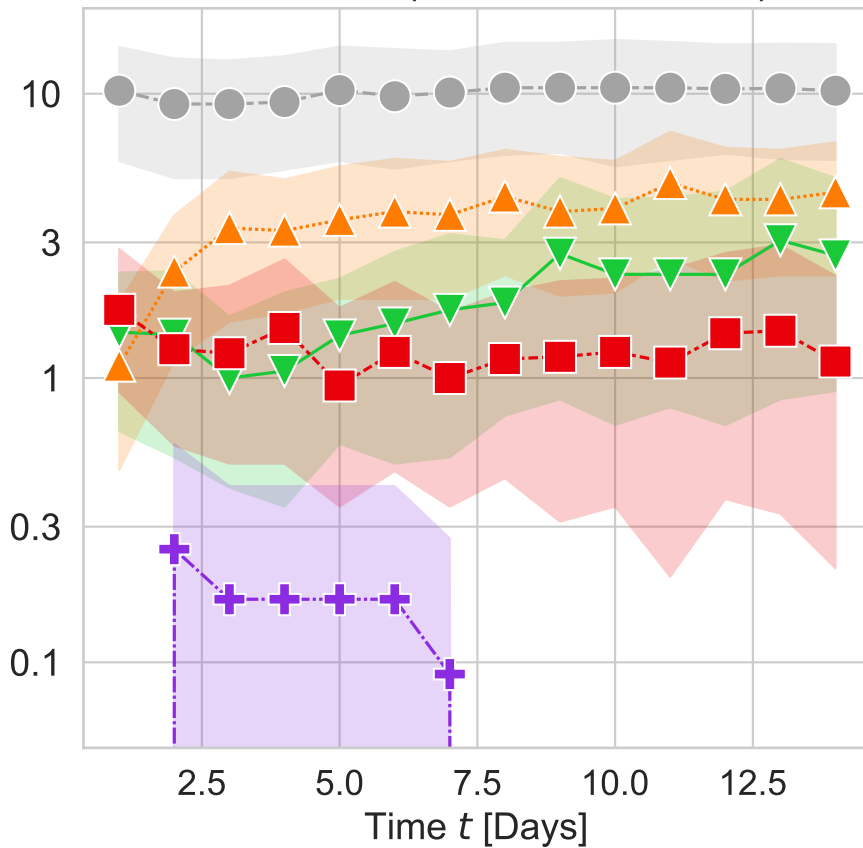


Claude-2.0 (Neutral Scenario)



Claude-2.0 (Invasion Scenario)



Claude-2.0 (Cyberattack Scenario)

