Stable Configurations with (Meta)Punishing Agents

Nathaniel Beckemeyer, William Macke, and Sandip Sen

{nate,william-macke,sandip}@utulsa.edu,
The Tandy School of Computer Science
The University of Tulsa
Tulsa, OK, United States

Abstract. We consider an adaptation of Axelrod's metanorm model, where a population of agents choose between cooperating and defecting in bilateral interactions. Because punishing incurs an enforcement cost, Axelrod proposes using metanorms, to facilitate the stability of a norm of punishing defectors, where those who do not punish defectors can themselves be punished. We present two approaches to study the social effects of such metanorms when agents can choose their interaction partners: (a) a theoretical study, when agent behaviors are static, showing stable social configurations, under all possible relationships between system parameters representing agent payoffs with or without defection, punishment, and meta-punishment, and (b) an experimental evaluation of emergent social configurations when agents choose behaviors to maximize expected utility. We highlight emergent social configurations, including anarchy, a "police" state with cooperating agents who enforce, and a unique "corrupt police" state where one enforcer penalizes all defectors but defects on others!

Keywords: MABS workshop, multi-agent systems, cooperation, norm emergence, network topologies, metanorm, metapunishment, punishment

1 Introduction

With the burgeoning of participation and activities in online social networks, there is increasing interest in understanding how interactions between individuals can give rise to emergent social structure and phenomena [4,3,10], such as information cascades [7], as well as the influence individuals have on others [9]. Concomitantly, researchers have used agent-based models and simulations to study how behavioral traits and interaction decisions can shape the dynamics of social networks. The goal of these research is to understand the dynamics of network connections and topologies [15,18,13], information flow [5,19], or to characterize the emergence of conventions or norms [1,11,12,17,21] or cooperative behavior [14,16]. While some of these research analytically prove convergence or equilibrium or formally derive rational agent behaviors [8,16,18], others use extensive experimental evaluations to understand the nature of emerging behaviors and topologies in networks of self-interested agents [1,14,15].

A number of these studies investigate scenarios where the network topology changes based on strategic or exploratory rewiring of connections by agents seeking more beneficial partnerships [15,17,18]. Interaction between neighbors on networks are often represented as a stage game [11,12]. Some of these studies on norm emergence have also considered agents who use punishments and sanctions to facilitate convergence to social welfare maximizing outcomes for these games [14,20]. The use of punishments to facilitate norm emergence goes back to the work of Axelrod [2] who observes *"A norm exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way."* Axelrod observes that punishing norm violators can be costly and hence free riders, who do not punish violators but rely on others to do so, may proliferate. He then suggested the use of a *metanorm*, a norm to punish those who do not punish norm violators (we refer to this as *metapunishment*)! Mahmoud *et al.* [14] have used resource-aware, adaptive use of metanorms to promote cooperation in peer-to-peer resource sharing networks, when individuals may have incentives to defect.

Our goal in this paper is to investigate how the ability to rewire as well as the use of punishment and metapunishment can result in the emergence of different network topologies between different types of agents. We consider the following agent types connected in a network: *cooperators* who always cooperate with their neighbors, *defectors* who defect against all neighbors, *punishers (corrupt)* who are cooperating (defecting) agents that also punish, and metapunish if that option is available. A link is created between two agents if any one of them wants to interact with the other. Each agent interaction is represented as a stage game with a payoff matrix representing a social dilemma: mutual cooperation is preferable to mutual defection but there is incentive to defect against a cooperator. When punishment is allowed, the situation corresponds to an extensive form game, where an agent has the option to punish a defecting neighbor. When metapunishment is allowed, an agent can metapunish a neighbor who do not punish its defecting neighbors. Punishment and metapunishment have costs to the enforcer, which are less than the corresponding costs to the recipient.

In the present study, we make the two following assumptions: Only one agent is necessary to choose another as a neighbor, or, equivalently, both agents must agree to cease interacting; and agents, once having selected a strategy, do not change their behavior. We assume the former following initial work in [6], wherein only one agent must choose to interact in order to connect to the other agents. Additionally, one can imagine a variety of real world scenarios corresponding to bilateral agreement, such as a group in a social network where leaving brings a substantial cost to the user, reputation or otherwise, which forces the user to interact with others he or she may not like. This formulation of the problem also allows for an interesting new aspect of the game: oppression. With mutual consent required to terminate a link, one party can defect and enforce norms upon another without this parties permission. Additionally, we find the choice of static strategies a reasonable formation because people tend to maintain a mostly constant persona when interacting with their neighbors.

Similar work was performed by Galán *et al.* in [13]. We note, however, that their work focused on stable norms resulting from static topologies; our paper considers the converse question of the stable topologies that result from rewiring connections while agents follow static behaviors. The network characterizations that they present are also unsuitable for our model due to the fact that, in our work, the networks either initialize as fully connected or links can be added as agents deem rational, as opposed to constant topologies. For example, since all agents of a particular type behave in the same manner in our model, they will all make the same decisions as to which other agents to connect to or attempt to disconnect from—contrasting the probabilistic behaviors used in their work. Consequently, analyses of the resultant clustering coefficients, numbers of triples, or other metrics are uninteresting. Another key difference between the two works is that the agents in their work change strategies by the genetic forces of selection and mutation; in our work, however, behaviors only change in the experimental analysis due to rational choice, and are constant in the theoretical analysis.

The paper is organized as follows. In Section 2, we present the configurations that will result when agents cannot change their type but can change their connections. These situations are amenable to algebraic solutions and we can precisely derive the network topologies that will arise by the rewiring process. We consider all possible game scenarios conforming with the social dilemma mentioned above and for various cost of making a new connection. We highlight interesting resultant networks for situations where there is (a) no punishment, (b) punishment but not metapunishment, and (c) metapunishment. In Section 3, we present experimental results showing converged network topologies where in addition to rewiring their connections, agents can also myopically change their types to maximize the utility they expect to receive given their current neighbors (these scenarios do not lend them to similar algebraic analysis as in the case of fixed agent types). We find interesting converged topologies such as a *police state* where few punishing agents keep other agents from defecting, as well as an oddball *corrupt police state* where a lone (meta)punishing agent prevents others from defecting but itself defects against all others! An associated interesting observation is the relative frequency with which the different converged topologies result when punishment is used with or without metapunishment. We conclude with a brief discussion of future work.

2 Theoretical Analysis

2.1 Specification

Game Mechanics Starting with an initial network of fully connected agents, the game proceeds in many rounds. In each round, an agent interacts with each of its neighbors. An agent, *Player A*, can either choose to cooperate or defect against its neighbor, *Player B*. Choosing to defect gives *Player A* the temptation reward and *Player B* the hurt value, and choosing to cooperate gives the baseline reward to both players. When the punishment option is present, each interaction has a second stage, wherein, if *Player A* chooses to defect against *Player B*, then *Player B* has the opportunity to punish *Player A*.

Finally, if the metapunishment option is present, each round has a second phase. Each player, *Player A*, observes the interactions of each other agent, *Player B*—specifically, whether *Player B* chose to punish. If *Player B* chose not to punish a defector, then *Player A* has the opportunity to metapunish *Player B*. Metapunishment enables agents to encourage other agents to punish those agents who defect.

An agent has to pay a linking cost r for each of its link to a neighbor. If a link to a neighbor brings negative utility, then an the agent will try to cut that link at the end of a round. If both agents in a linked pair attempt to cut a link, the link will be eliminated. If only one agent, however, attempts to cut that link, then the link will remain.

Agent Strategies For a description of the payoffs used in this game, see Table 1.

Table 1. Glossary of Payoffs. If a payoff contains the letter on the left, then the payoff includes the reward for the interaction on the right (the payoffs are additive). For instance, *dh* indicates that the agent both defected and was defected against.

b	The baseline—the reward for cooperation on both sides.
d	Defecting.
h	Being defected against (harmed).
dp	Defecting and being punished.
he	Being defected against and enforcing.
m	Being metapunished.
M	Metaenforcing.

Each agent type in the population has a type or strategy which cannot be changed. Without punishment, there are two agent types: *cooperator* types always cooperate and *defector* types always defect.

In the case of basic punishment, the *cooperator* type agents cooperate but do not enforce punishment. The *defector* type agents defect but do not enforce. There are two additional types: The *punisher* and *corrupt*. The *punisher* type agents cooperate and enforce punishment. The *corrupt* type agents defect and enforce punishment.

In the case of metapunishment, the agent types are the same as those in the basic punishment case, but the *punisher* and *corrupt* types both metapunish as well while other agent types do not.

2.2 Payoff Topologies

No punishment We first examine the case of no punishment. Table 2 represents the payoff matrix for this scenario.

Table 2. Payoffs without punishment

	$cooperator$	$defector$
$cooperator$	(b, b)	(h, d)
$defector$	(d, h)	(dh, dh)

Because there is no punishment, the only options are passivity and defection. b is simply the baseline. d is the baseline plus the temptation reward, which is included to incentivize agents to defect. h is b plus the hurt value, included to incentivize agents to punish. So, we make the following assumptions:

1. The temptation reward is greater than 0, or equivalently, $d > b$
2. The hurt value is less than 0.

From these assumptions, we can conclude that $d > b > h$ and, furthermore, that $d > dh > h$, since dh is simply $b + hurt\ value + temptation\ reward$

These conditions lead to six meaningful placements of the linking cost, r, and five unique topologies:

1. $r > d$: The network is empty because the linking cost is higher than the maximum possible reward from a link.
2. $d > r \geq dh, b$: The defecting agents form links with the passive agents in order to gain the temptation reward, d
3. $d, dh > r > b$: The defecting agents form links with themselves (for dh) and the passive agents (for d).
4. $b > r \geq dh, h$: The defecting agents connect to the passive agents, and the passive agents connect to themselves.
5. $b, dh > r > h$: A complete network is formed (the defecting agents will forcibly connect to the passive agents).
6. $h > r$: A complete network is formed.

Punishment In this section, we examine the case of basic punishment. Table 3 represents the payoff matrix for this scenario.

Table 3. Payoffs with basic punishment

	$cooperate$	$punish$	$defect$	$corrupt$
$cooperate$	(b, b)	(b, b)	(h, d)	(h, d)
$punisher$	(b, b)	(b, b)	(he, dp)	(he, dp)
$defector$	(d, h)	(dp, he)	(dh, dh)	(dhp, dhe)
$corrupt$	(d, h)	(dp, he)	(dhe, dhp)	$(dhpe, dhpe)$

In addition to the assumptions made in the previous section, we assume that enforcing and being punished cost the agent, and that it is worse for an agent to be punished after defecting than for an agent to enforce after being defected against:

1. The enforcement cost is less than 0.
2. The punishment cost is less than the enforcement cost.
3. $he > dp$: Total payoff for the punisher is greater than that of the punished.

From these assumptions, we can conclude that $d > b > h > he > dp > dhp > dhpe$, that $d > dh > h$, and that $dh > dhe > he$. These orderings suggest 13 possible placements for the linking cost, which lead to 10 different topologies. An interesting few selected results follow.

Agents who punish can, in some configurations, prevent defecting agents from connecting to themselves. Figure 1(a) shows a sample configuration wherein the *punisher* agents are not connected to defecting agents, but the *cooperator* agents are. An interesting note about Figure 1(a) is its similarity to a hub network, where the *cooperator* agents are the hub, and the other agents do not interact outside of their own groups.

In general, punishment is a highly effective method for agents to defend themselves against defection. Figure 1(b) represents the most connected network wherein agents who defect, *corrupt* and *defector* agents, still connect to the *punisher* agents.

(a) $d, b, dh > r \geq dhe, he, dp, dhp, dhpe$
This is an example of a network where punishers are safe from defection.

(b) $d, b, dh, h, dhe, he > r \geq dp, dhp, dhpe$
This topology is the most connected network wherein agents still defect against punishers.

Fig. 1. Interesting topologies from basic punishment.

Metapunishment In this section, we examine the case of metapunishment. Table 4 represents the payoff matrix for this scenario. In this section, similarly to the case of basic punishment, we assume additionally that it costs to meta-enforce and to be metapunished, and that being metapunished for neglecting to punish is worse for an agent than for an agent's meta-enforcing. That is,

1. The meta-enforcement cost is less than 0.
2. The metapunishment cost is less than the meta-enforcement cost.

Table 4. Payoffs with metapunishment

	cooperate	*punish*	*defect*	*corrupt*
cooperate	(b, b)	(m, M)	(h, d)	(hm, dM)
punish	(M, m)	(b, b)	(eM, dpm)	(he, dp)
defect	(d, h)	(dpm, eM)	(dh, dh)	$(dhpm, dheM)$
corrupt	(dM, hm)	(dp, he)	$(dheM, dhpm)$	$(dhpe, dhpe)$

3. $M > m$: Being metapunished is worse than meta-enforcing.

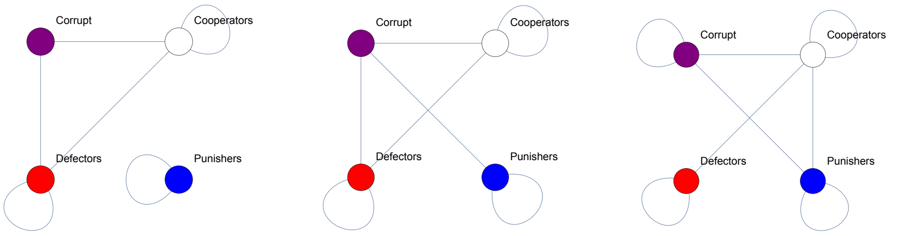
From these assumptions, we can conclude that $d > b > h > he > dp > dhp > dhpe$, that $d > dM > M > m > hm > dpm > dhpm$, that $M > eM > dpm > hm$, that $d > dh > h$, that $dh > dhe > he$, that $dhe > dheM$, and that $dM > dheM > dpm$.

These constraints imply 113 possible placements for the linking cost, which lead to 73 unique topologies. In the following paragraphs, we highlight notable results.

Metapunishment can destabilize previously stable topologies. Figure 2(a) shows one circumstance in which *punisher* agents will not connect to *cooperator* agents in contrast to the case of basic punishment. Specifically, the *cooperator* and *punisher* agents used to receive b when they interacted; however, in this case, metapunishment reduces the payoffs below the linking cost.

Additionally, metapunishment can entirely cease interactions between *punisher* agents and nonpunishing agents. As an example, figure 2(b) contains no connections between the *punisher* agents and the *defector* agents nor the *cooperator* agents. This topology is also remarkable because the temptation reward is sufficient to offset the meta-enforcement cost, as evidenced by the connection from the *corrupt* agents to the nonpunishing ones. This phenomena is interesting because the *corrupt* agents are punishing agents for not punishing the *corrupt* agents.

An interesting side effect of metapunishment is that the *defector* strategy may actually present a way for agents to defend themselves. In figure 2(c), the *defector* agents are not connected to the *punisher* agents. The *cooperator* agents also are connected to the *punisher* agents. This connections implies that the meta-enforcement cost is, alone, insufficient to prevent *punisher* agents from linking to *cooperator* agents. Additionally, the *corrupt* agents are, connected to the *punisher* agents. This connection implies that the hurt value and enforcement cost are insufficient to prevent a link from forming. Therefore, it is the combination of hurt value, enforcement cost, and meta-enforcement cost that does prevent the link from the *punisher* to the *defector* agents from forming—a combination that can only occur with agents using the *defector* strategy.



(a) d, b, dh, dhe, dM, $dheM > r > h$, he, dp, dhp, $dhpe$, M, m, hm, eM, dpm, $dhpm$
This topology demonstrates metapunishment can halt agents connected with only basic punishment from connecting now.

(b) d, b, dh, h, dhe, he, dp, dhp, dM, $dheM$, $> r >$ $dhpe$, M, m, hm, eM, dpm, $dhpm$
A topology wherein the meta-enforcement cost prevents *punisher* agents from connecting to nonpunishing agents.

(c) d, b, dh, h, dhe, he, dp, dhp, $dhpe$, dM, $M > r >$ m, hm, eM, $dheM$, dpm, $dhpm$
This topology shows the power of $defector$ agents to thwart metapunishment.

Fig. 2. Interesting topologies from including metapunishment

3 Experimental Analysis

The above analysis assumed agent types were static. To understand the emergent topologies when agents could myopically adapt their types to optimize payoff given their neighbors types, we ran simulations varying various parameters. During rounds, agents would follow this algorithm:

```
procedure Agent_Behavior()
maxUtility = Utility(currentStrategy)
maxStrategy = currentStrategy
for strategy in Strategies do
  if Utility(strategy) > maxUtility then
    maxUtility = Utility(strategy)
    maxStrategy = strategy
  end if
end for
if maxStrategy != currentStrategy then
  currentStrategy = maxStrategy
  return
end if
for link in CurrentLinks do
  if Utility(link)<0 then
    removeLink()
    return
  end if
end for
LinkToRandomAgent()
```

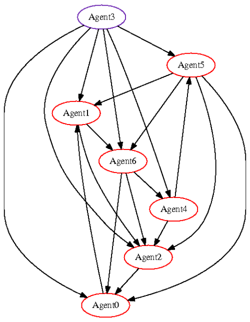


Fig. 3. One Way Corrupt Police (Red-*cooperator*, Violet-*corrupt*).

Where the utilities of links are defined by the values in the payoff matrices given in section 2, and the utility of a strategy is simply the sum of all of the links of an agent, assuming that the agent adopts that strategy. In a round agents make their decisions sequentially. The order of turns was decided randomly at the beginning of each round. Simulations were run with both simple punishment and meta-punishment and with various numbers of agents. Each simulation ran for 1000 rounds. Only simple graphs were used; i.e., if Agent 1 connected to Agent 2, then Agent 2 could not connect to Agent 1. Each agent was assigned a random strategy at the beginning of the game.

3.1 Observed Stable Configurations

All experiments produced one of three stable configurations: *Anarchy* indicates all agents are defecting, *Police State* refers to a few punishing agents and the rest neutral, and *Corrupt Police State* refers to exactly one agent defecting and punishing while the rest are neutral.

The three stable configurations mentioned above could form different topologies: *Complete Network, Empty Anarchy, One way corrupt police.* In the complete network, all agents linked with all other agents. Any of the three configurations could form with this topology. Empty anarchy was an anarchy network without any agent linking to any other agent. The one way Corrupt Police was the most interesting of the three topologies. It was a corrupt police state, but none of the cooperators were willingly linked to the corrupt police officer. Thus we had one group of agents that would link to only agents of their own type, but were being stabilized and exploited simultaneously by an outside agent. See Figure 3.

3.2 Conditions for network development

An important goal of the experimental analysis was to observe what conditions were required for each of the three stable configurations to emerge.

Figure 4 shows the relative frequency of emergence of different stable configurations as we vary the number of agents in the network. Without metapunishment, as the number of agents increases, the number of configurations that result

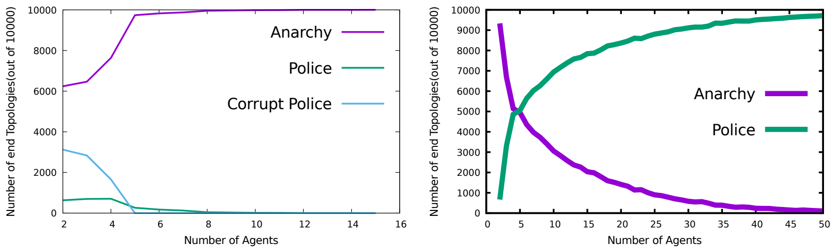


Fig. 4. End Topologies with different # of Agents: Punishment only (Left), Metapunishment included (Right). Parameters: Base 0, Defection Reward 3, Defection Hurt 1, Punishment Cost 2, Punishment Hurt 9, Linking Cost 0.

in anarchy also increases. We will discuss this phenomena in detail below. With metapunishment, increasing the number of agents increases the likelihood of a police state emerging. Presence of more metapunishers force non-punishers to start punishing; thus with more agents present there is an increase in frequency of the emergence of police states.

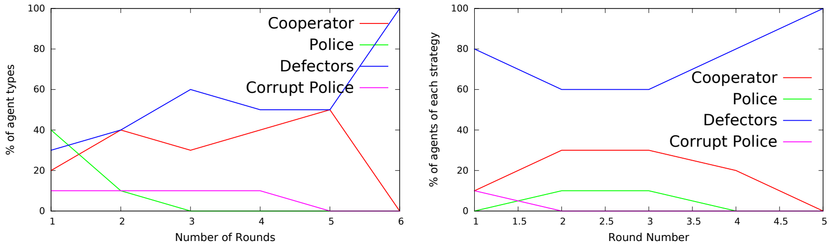


Fig. 5. % of agent types as Anarchy develops, 10 agents: punishment only (Left), with Metapunishment (Right). **Parameters: Base 10, Defection Reward 1, Defection Hurt 3, Punishment Cost 3, Punishment Hurt 12, Linking Cost 0.**

Anarchy Due to the randomness of allocation of initial agent types, the initial number of agent types may not be equal. The initial agent type distribution is likely to be more skewed particularly for small agent populations. If there were too many defectors at the beginning, anarchy developed from large numbers of agents defecting . When a punisher links with a defector, one of two things happen: the defector stops defecting or the punisher stops punishing. When there are far more defectors than there are punishers, it becomes much more likely that the punisher will have to back down and stop punishing at some point. For small populations there are more chances of very few defectors in the initial population, whence the network may evolve to a state different from anarchy. With larger populations, there are more agents that can defect in the early rounds of the game and it becomes harder for the punishing agents to maintain order. With large enough numbers of agents, the end topology is almost always anarchy. Hence, all figures of networks developing are shown with 10 agents. Anarchy was by far the most common of the three stable configurations that formed without metapunishment. When metapunishment was included, the frequency of anarchy networks drastically decreased because of reasons listed below. The percentage of different agent types in sample runs that evolved Anarchy networks, with or without metapunishment, are shown in Figure 5. Modified parameters are used when observing network developments to reduce the anarchy development rate.

Police The convergence to a police state was facilitated by an initial state of a large number of punishers. These punishers would have to immediately link with each other in order for the police state to form, because otherwise the punishers would want to become defectors. If two punishers link with each other, neither will defect to avoid being punished by the other. However if a punisher is linked only with non-punishing agents, then it will become a defector for the Utility boost. From there they would force all defecting agents to become neutral as they connected to them. When metapunishment is included, punishers gain the ability to force other agents to become punishers. This aides the development of police networks and increases their relative frequency. Sample runs that evolved the Police state are presented in Figure 6.

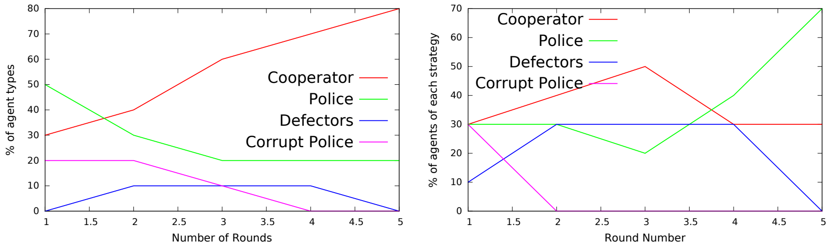


Fig. 6. % of agent types as Police state evolves, 10 agents: punishment only (Left); with Metapunishment (Right). Parameters: Base 10, Defection Reward 1, Defection Hurt 3, Punishment Cost 3, Punishment Hurt 12, Linking Cost 0.

Corrupt Police The corrupt police state developed from an initial state of a large number of agents who were defecting and punishing. As these agents linked with others, they forced those agents to become neutral to avoid punishment. When two of these agents connect, one will back down and become neutral while the other will remain a defector and punisher. A sample run that evolved a Corrupt Police network is shown in Figure 7.

This demonstrates one of the more interesting outcomes of the game: Corruption will not tolerate company while non-corruption requires it. In the corrupt police network, all corrupt police officers will eliminate each other until only one remains, while the police network requires multiple interacting officers. The corrupt police network was only stable without metapunishment. If metapunishment exists, then the corrupt police officer will have to punish neutral agents for not punishing it. This in turn forces the neutral agents to become punishers, and hence the corrupt police network does not emerge with metapunishment.

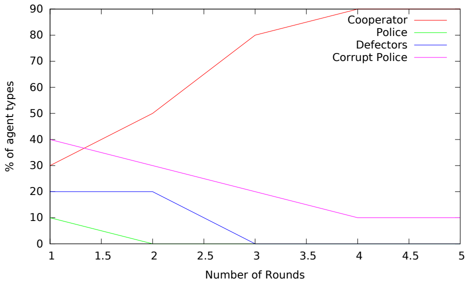


Fig. 7. % of agent types as Corrupt Police configuration develops (Simple Punishment); 10 agents. Parameters: Base 10, Defection Reward 1, Defection Hurt 3, Punishment Cost 3, Punishment Hurt 12, Linking Cost 0.

4 Conclusions

We investigated the effect of rewiring and behavior adoptions on the emergent topology of networked self-interested agents interacting in a social dilemma scenario with and without punishment and metapunishment options. When agent types were fixed, we identify, using algebraic calculations, interesting topologies that result under various relationships between agent interaction payoffs and rewiring costs. Such derivations are not forthcoming when agents can change their types myopically to maximize payoffs given their neighborhood. We run a suit of experiments and observe the emergence of different classes of network topologies. Particularly interesting are the police and corrupt police states and their relative abundance with and without the option of metapunishment.

We plan to investigate unilateral elimination of links which should allow for cooperators to thrive more frequently. We will analyze mixed, rather than pure strategy types, where agents defect with some probability $0 < p < 1$. We will also study a broader class of social dilemmas, including the prisoner's dilemma and the Hawk-Dove game. In Section 2, we assumed all types are present in equal numbers; we will analyze non-uniform distribution of agent types. Finally, we intend to perform analyses similar to those done by Galán *et al.* in [13]: Allowing for nondeterministic behavior could lead to some highly intriguing resultant social networks and network properties. Combining all of these future directions, characterizing networks with unilateral links could additionally prove fascinating.

Acknowledgments

We would like to thank the University of Tulsa and in particular the Tulsa Undergraduate Research Challenge (TURC) for financial support of this project.

ORCiD

Nathaniel Beckemeyer: `https://orcid.org/0000-0002-1594-4843`

References

1. S. Airiau, S. Sen, and D. Villatoro. Emergence of conventions through social learning – heterogeneous learners in complex networks, 2014.
2. R. Axelrod. An evolutionary approach to norms. *American Political Science Review*, 80:1095–1111, 1986.
3. O. Baetz. Social activity and network formation. *Theoretical Economics*, 10(2):315–340, 2015.
4. A. Barabasi. *Network Science*. 2016.
5. F. Belardinelli and D. Grossi. On the formal verification of diffusion phenomena in open dynamic agent networks. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 237–245, Richland, SC, 2015. International Foundation for Autonomous Agents and Multiagent Systems.
6. S. Berninghaus and B. Vogt. Network formation and coordination games, March 2003.
7. J. Borge-Holthoefer, R. A. Baos, S. Gonzlez-Bailn, and Y. Moreno. Cascading behaviour in complex socio-technical networks. *Journal of Complex Networks*, 2013.
8. L. Brooks, W. Iba, and S. Sen. Modeling the emergence and convergence of norms. In *IJCAI*, pages 97–102, 2011.
9. M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi. Measuring user influence in twitter: The million follower fallacy. In *in ICWSM 10: Proceedings of international AAAI Conference on Weblogs and Social*, 2010.
10. E. David and K. Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA, 2010.
11. J. Delgado. Emergence of social conventions in complex networks. *Artificial Intelligence*, 141(1–2):171 – 185, Jan 2002.
12. J. M. Epstein. Learning to be thoughtless: Social norms and individual computation. *Computational Economics*, 18(1):9–24, 2001.
13. J. M. Galán, M. M. Łatek, and S. M. M. Rizi. Axelrod's metanorm games on networks. *PLOS ONE*, 6(5):1–11, 05 2011.

14. S. Mahmoud, S. Miles, and M. Luck. Cooperation emergence under resource-constrained peer punishment. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, AAMAS '16, pages 900–908, Richland, SC, 2016. International Foundation for Autonomous Agents and Multiagent Systems.
15. A. Peleteiro, J. C. Burguillo, and S. Y. Chong. Exploring indirect reciprocity in complex networks using coalitions and rewiring. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, pages 669–676, Richland, SC, 2014. International Foundation for Autonomous Agents and Multiagent Systems.
16. B. Ranjbar-Sahraei, H. Bou Ammar, D. Bloembergen, K. Tuyls, and G. Weiss. Evolution of cooperation in arbitrary complex networks. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, pages 677–684, Richland, SC, 2014. International Foundation for Autonomous Agents and Multiagent Systems.
17. B. T. R. Savarimuthu, S. Cranefield, M. Purvis, and M. Purvis. Norm emergence in agent societies formed by dynamically changing networks. In *Proceedings of the 2007 IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, IAT '07, pages 464–470, Washington, DC, USA, 2007. IEEE Computer Society.
18. S. Sina, N. Hazon, A. Hassidim, and S. Kraus. Adapting the social network to affect elections. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 705–713, Richland, SC, 2015. International Foundation for Autonomous Agents and Multiagent Systems.
19. A. Tsang and K. Larson. Opinion dynamics of skeptical agents. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, pages 277–284, Richland, SC, 2014. International Foundation for Autonomous Agents and Multiagent Systems.
20. D. Villatoro, G. Andrighetto, J. Sabater-Mir, and R. Conte. Dynamic sanctioning for robust and cost-efficient norm compliance. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume One*, IJCAI'11, pages 414–419. AAAI Press, 2011.
21. D. Villatoro, S. Sen, and J. Sabater-Mir. Topology and memory effect on convention emergence. In *IAT*, 2009.