Evolving Cooperation in the N-player Prisoner's Dilemma: A Social Network Model

Golriz Rezaei, Michael Kirley, and Jens Pfau

Department of Computer Science and Software Engineering
The University of Melbourne, Australia
{grezaei,mkirley,jens.pfau}@csse.unimelb.edu.au

Abstract. We introduce a social network based model to investigate the evolution of cooperation in the N-player prisoner's dilemma game. Agents who play cooperatively form social links, which are reinforced by subsequent cooperative actions. Agents tend to interact with players from their social network. However, when an agent defects, the links with its opponents in that game are broken. We examine two different scenarios: (a) where all agents are equipped with a pure strategy, and (b) where some agents play with a mixed strategy. In the mixed case, agents base their decision on a function of the weighted links within their social network. Detailed simulation experiments show that the proposed model is able to promote cooperation. Social networks play an increasingly important role in promoting and sustaining cooperation in the mixed strategy case. An analysis of the emergent social networks shows that they are characterized by high average clustering and broad-scale heterogeneity, especially for a relatively small number of players per game.

1 Introduction

Social dilemma games such as the prisoner's dilemma [1, 13] have provided important insights into the emergent properties of interactions in multi-agent systems. However, in order to investigate cooperation within a social group, games consisting of more than two players must be considered. In the N-player prisoner's dilemma [3, 16, 15], multiple agents ($N \geq 2$) interact within their designated group and must choose to *cooperate* or *defect*. Any benefit or payoff is received by all participants; any cost is borne by the cooperators only. Hardin [8] describes the N-player prisoner's dilemma as a "tragedy of the commons" game in which the players are worse off acting according to their self interests than if they were cooperating and coordinating their actions.

Recently, a number of studies have investigated the evolution of cooperation in 2×2 games on dynamical networks, where the interaction links between agents playing the game varied over time. The models have ranged from comparative studies examining the level of cooperation on different base network models [11] to endogenous network formation models based on local interactions [2]. In Tanimoto [17], individuals were able to self-organize both their strategy and their social ties throughout evolution, based exclusively on their self-interest. It was reported that the clustering coefficient of the network (see Section 4) affects the emergence of cooperation in the games.

In this paper, we extend this line of research by focussing on the N-player prisoner's dilemma game based on the formalism of Boyd and Richerson [3]. We propose an endogenous network formation model. Successful strategies spread via a form of cultural evolution based on imitation. When agents play cooperatively, they form a social link that is reinforced each time the action is repeated. This weighted link provides an estimate of the "reliability" of the agent. When an agent defects, all links with its opponents in that game are dissolved. We examine different levels of agent cognitive ability: agents with a pure strategy always play cooperate or defect. In contrast, an agent equipped with a mixed strategy plays a particular action based on a function of the reliability of other agents in its group in a given game. Thus, individual agents are able to adjust both these links and the action they play based on interactions with other agents.

Detailed simulation experiments show that our model is able to promote higher levels of cooperation as compared with panmictic populations. An analysis of the social interactions between cooperative agents reveals high average clustering and associated single-to-broad-scale heterogeneity for relatively small values of N. When agents are equipped with mixed strategies, social networks play a more significant role.

The remainder of this paper is organized as follows: In Section 2 we present background material related to N-player prisoner's dilemma game. In Section 3 our model is described in detail. This is followed by a description of the simulation experiments and results in Section 4. We conclude the paper in Section 5 and identify future research directions.

2 Background and related work

2.1 N-player prisoner's dilemma game

In the 2×2 prisoner's dilemma game, two players interact with each other by simultaneously choosing to *cooperate* or to *defect*. Based on their joint actions, each individual receives a specific payoff or utility value, U [1]. The altruistic act consists of conferring a benefit b on the recipient at a cost c to the donor. Here, $b > c$. If both players cooperate, each receives $b - c$, which is better than what they would obtain by both defecting (the Nash equilibrium of the game). But a unilateral defector would earn b, which is the highest payoff, and the exploited cooperator would pay the cost c without receiving any benefit.

The N-player game can be thought as a natural extension of the 2-player game. Boyd and Richerson [3] define the payoff values as follows:

$$U = \begin{cases} \frac{b \times i}{N} - c & \text{if the agent cooperated,} \\ \frac{b \times i}{N} & \text{if the agent defected.} \end{cases} \quad (1)$$

with i being the number of cooperators in the game, and N the number of players. The following conditions must also hold for a valid multi-player prisoner's dilemma game [16]: $c > \frac{b}{N}$ (defection is preferred for the individual) and $b > c > 0$ (contribution to social welfare is beneficial for the group).

Conventional evolutionary game theory predicts that cooperation is unlikely to emerge in the N-player prisoner's dilemma, and if it does emerge, then the cooperation levels are unlikely to be stable [3]. As the number of players per game increases, there should be a decrease on the level of cooperation [16].

2.2 Reciprocity in the spatial prisoner's dilemma

It is well known that spatial structure enhances the cooperation levels in the prisoner's dilemma game [9, 13]. It has also been shown that there is a direct correlation between the relative costs and benefits of cooperating and defecting and the underlying connectivity of agents playing the game [14]. High levels of cooperation can be attributed to "network reciprocity". That is, by limiting the number of game opponents (based on a pre-defined local neighborhood) and employing a local adaptation mechanism in which an agent copies a strategy from a neighbor linked by a network, higher levels of cooperation typically emerge.

There have been relatively few papers investigating spatial versions of the N-player prisoner's dilemma game. One recent notable example is the work of Santos et al. [15] who have shown that heterogenous graphs and the corresponding diversity in the number and size of the public goods game in which each individual participates (represented via an N-player prisoner's dilemma game) helps to promote cooperation. Here, the heterogenous graphs enable cooperators to form clusters in some instances, thereby reducing exploitation.

2.3 Reputation systems and indirect reciprocity mechanisms

Reputation systems are a common method of gaining information about an agent's behavioral history [10]. Approaches based on reciprocity and image scoring [12] have been applied as a solution for "tragedy of the commons" problems in open peer-to-peer networks.

The use of "shared history" where all nodes have access to the behavior of all other nodes, provides a mechanism for nodes to adapt their behavior [6]. A disadvantage of this approach is that maintaining a shared history increases overhead [7]. Consequently, Hales and Arteconi introduced a computational sociology tagging approach to support high levels of cooperation. It was achieved without central control or reciprocity based on a reputation mechanism in prisoner's dilemma like games. This approach was based on simple rules of social behavior observed in human societies and emergent networks. As such it was consistent with the work of Suzuki and Akiyama [16] who examined the relationship between cooperation and the average reputation of opponents.

Ellis and Yao [5] introduced a distributed reputation system for the prisoner's dilemma game using the notion of "indirect reciprocity". In their model, individuals do not store their own or their opponents' image/reputation scores. Instead, these scores are embedded in a social network in the form of mutually established links. However, this model is potentially vulnerable to collusion, as

Algorithm 1 Social network based N-player prisoner's dilemma model

Require: Population of agents $\mathcal{P}$, evolutionary rate $e \in [0, 1]$, number of iterations i_{max}, number of players per game $N \geq 2$.

1: **for** $i = 0$ to i_{max} **do**
2: $\quad \mathcal{G} = \emptyset$
3: $\quad$ **while** $g = \textsc{NextGame}(\mathcal{P}, \mathcal{G}, N)$ **do**
4: $\quad\quad \mathcal{G} = \mathcal{G} \cup \{g\}$
5: $\quad\quad \textsc{PlayGame}(g)$
6: $\quad\quad \textsc{AdaptLinks}(g)$
7: $\quad$ **end while**
8: $\quad$ **for** $i = 0$ to $|\mathcal{P}| \times e$ **do**
9: $\quad\quad a, b = \textsc{Sample}(\mathcal{P})$
10: $\quad\quad \textsc{CompareUtilityAndSelect}(a, b)$
11: $\quad$ **end for**
12: **end for**

agents may leave unjustified positive feedback in order to boost each other's scores.

3 Model

In our model of the N-player prisoner's dilemma game, a population of agents playing a series of independent games self-organize their social ties based exclusively on their self-interest. Interactions can take place between any group of agents, however, they tend to happen between those that have cooperated previously. This reflects the notion that agents prefer to seek interaction with partners that have already proven to be reliable. Agents who play cooperatively form social links, which are reinforced by subsequent cooperative actions. However, when an agent defects, the links with its opponents in this game are broken. As agents not participating in this game are unaware of the defective action, their links with the defective player, however, are retained.

The execution cycle of the model is sketched in Algorithm 1. Every iteration involves the forming of a number of *games* $\mathcal{G}$ of size N from the population (line 3); the *execution* of each game and the calculation of its outcome (line 5); the *adaptation* of the links of the agents in the game based on the actions played (line 6); and finally the *selection* process on a subpopulation, whose size is determined by the evolutionary rate e (lines 8-11). In the following sections, these steps are explained in detail.

3.1 The formation of games

The population of agents $\mathcal{P}$ is partitioned into disjoint sets of size N, each of which forms a game. While the first agent for every game is selected randomly amongst those that have not yet been assigned to any game in this iteration, the other $N - 1$ slots of the game are filled as follows: With probability ϵ, the slot is filled randomly with an agent from the first agent's neighborhood. With probability $1-\epsilon$, or if all agents in the neighborhood have already been assigned to a game in this iteration, the slot is filled randomly with an arbitrary agent from the remaining population. In that, ϵ regulates how often the agent explores cooperative play with unknown members of the population or exploits its current local neighborhood. Usually every agent plays exactly one game per iteration. However, depending on the size of $\mathcal{P}$ and N, a single agent might not play at all or the last game might not reach size N.

3.2 The execution of games

The outcome of every game depends on the strategies of its players. We consider two different scenarios that host agents with varying cognitive abilities. The first one involves only players with *pure strategies*. That is, each agent always plays the cooperate or the defect action. In the second scenario, we also consider a third type of agent that follows a *mixed strategy*. A mixed strategy is an assignment of a probability to each pure strategy (see details below).

The action taken by a mixed strategy agent i depends on the weights of the links w_{ij} it has established with each of its opponents $j \in g$ for the current game g. The average link weight for player i in game g is then defined as:

$$\overline{w}_i(g) = \frac{1}{|g|}\sum_{j\in g} w_{ij}$$

A mixed strategy i plays cooperatively in game g with probability:

$$P_i(g) = \frac{\overline{w}_i(g)^\alpha + \beta}{\overline{w}_i(g)^\alpha + \beta + 1}$$

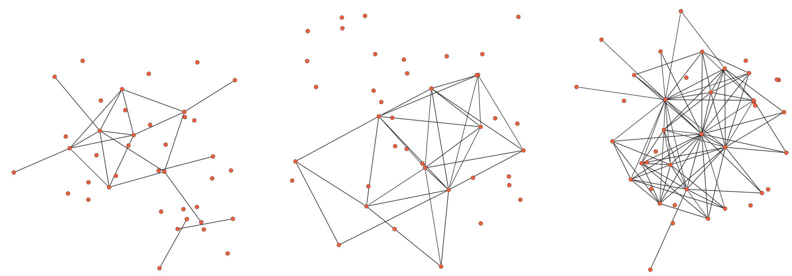
With probability $1 - P_i(g)$, it plays defectively. Basically, β determines the probability of playing cooperatively if the agent does not have any link with its opponents. It allows agents to play generously, which was found by [1] to be a feature of successful iterated prisoner's dilemma stategies. The gradient of the probability density function is determined by α. Higher values correspond to an agent being satisfied with less links with its opponents to decide on cooperative play. This decision rule was introduced by [5].

Based on the game's outcome every agent receives a payoff or utility. This payoff is based on the functions described in Boyd and Richerson [3] and Suzuki and Akiyama [16] (see Equation 1 in Section 2 for details).

3.3 Link adjustment

For every pair of agents i and j in the game, link weights w_{ij} are changed as follows:

$$w_{ij} = \begin{cases} w_{ij}+1 & \text{if both } i \text{ and } j \text{ played cooperatively,} \\ 0 & \text{otherwise.} \end{cases}$$



(a) Iteration 5 (b) Iteration 100 (c) Iteration 1000

Fig. 1: Three snap shots of the emerging social network at different iterations of the simulation. In this sample, $|\mathcal{P}| = 40$ and $N = 3$.

In Figure 1 three stages in the evolution of a sample network are shown, illustrating the dynamic nature of link adjustment in the model. We assume that links can only be formed by mutual consent, which prevents any defector from influencing the selection of its opponents to its own advantage. Furthermore, we assume that the actions of other agents are observable. Otherwise, establishing links as described here would be impossible. Defective actions cannot lead to negative link weights in order to encourage agents to be forgiving. This is another requirement for successful iterated prisoner's dilemma strategies according to [1].

3.4 Strategy update mechanism

A form of cultural evolution based on imitation is used for strategy update. At each time step, $|P| \times e$ pairs of agents are drawn randomly from the population for update. The accumulated utility of the agent pair is compared. The agent with the lower utility is replaced by a copy of the agent with the higher utility. The new agent copies the strategy and the links of the winning agent, forms a link with weight one to this agent and initializes its own utility. This models the propagation of successful strategies and trust within the growing network.

4 Simulations

A systematic Monte Carlo simulation study was carried out to investigate the system dynamics of our model. The underlying hypothesis tested was that the introduction of agents equipped with mixed strategies would promote higher levels of cooperation in the N-Player Prisoner Dilemma game. Our social network model should encourage high levels of cooperation to persist for longer, even when all agents played with a pure strategy.

4.1 Parameters

The following parameter settings were used in all simulations: population size $|\mathcal{P}| = 1000$ with $\epsilon = 0.9$ and strategy update probability $e = 0.001$. Payoff values $b = 5$ and $c = 3$ were used for the benefit and costs of cooperation. In the pure strategy scenario, the population was initialized with 50% cooperators and 50% defectors. In the mixed strategy scenario, the population was initialized with 33.3% pure cooperators, 33.3% pure defectors and 33.3% mixed strategy agents with $\alpha = 1.5$ and $\beta = 0.1$. We report results averaged over 20 independent trials with the number of iterations $i_{max} = 40000$.

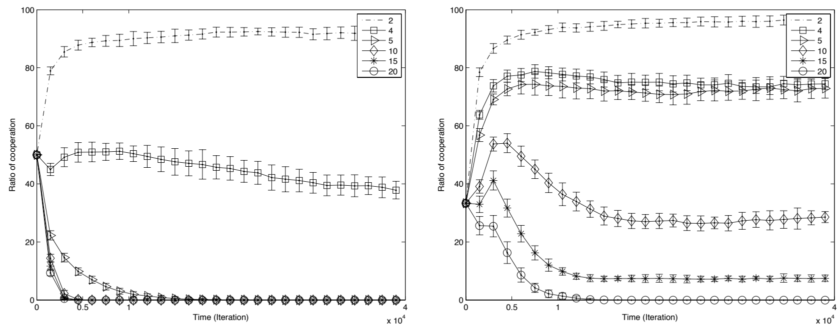
4.2 Results

Group size vs. strategy Figure 2(a) plots results for the proportion of cooperation in the population vs. time for increasing values of N when all agents play with a pure strategy. When $N = 2$, the high proportion of cooperation at equilibrium is consistent with results reported for the standard $2{\times}2$ prisoner's dilemma game using a spatial structure. As N increases, the equilibrium cooperation levels drop off. At $N = 4$, approximately 50% of the population is playing cooperatively. However, as expected, for larger values of N defection takes over the population at a rate proportional to N.

Figure 2(b) plots results for the proportion of cooperation in the population vs. time for increasing values of N when mixed strategy agents have been introduced. A comparison of the plots clearly shows that the introduction of mixed strategy agents promotes higher levels of cooperation. Decreasing trends in cooperation levels consistent with the number of players are still apparent. However, the equilibrium levels for the proportion of agents who have cooperated are significantly higher. For example, when $N = 10$, the proportion of cooperation initially increased as the social network began to form. However, after approximately 5000 time steps this level gradually decreases to an equilibrium level of approximately 30%. In the pure strategy case, when $N = 10$, defection was the dominant action after 5000 time steps.

Emergent social network When agents play cooperatively, they form a social link that is reinforced each time the action is repeated. When an agent defects, all links with its opponents in that game are dissolved. Here, we confine our discussion of the characterization of the emergent networks to the two quantities typically used in the analysis of networks: the clustering coefficient and the degree distribution [4]. The clustering coefficient is the probability that two nearest neighbors of a node are also nearest neighbors of each other. Its values range from zero to one and the latter value holds if and only if the network is globally coupled. This analysis provides important insights into the underlying social mechanisms, which promote the collective behavior within the model.

Figure 3 plots the average clustering coefficient vs. time for (a) the pure strategy model, (b) the mixed strategy model. The trend in the time-series values are consistent with the trends shown in Figure 2(a) and (b), that is, the



(a) pure strategy (b) mixed strategy

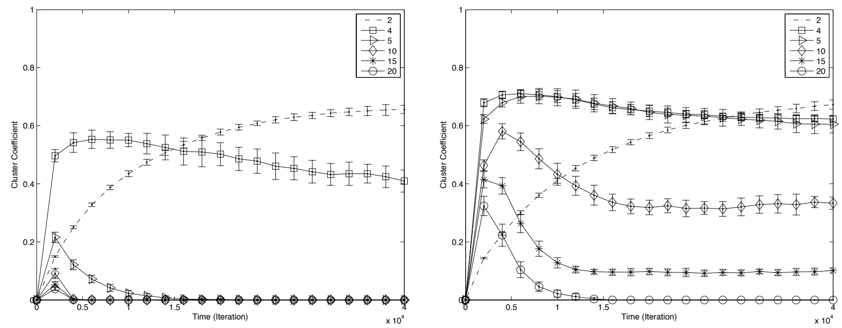
Fig. 2: The proportion of cooperation vs. time for various values of N for (a) the pure strategy model and (b) the mixed strategy model.

average clustering coefficient values are generally smaller for larger values of N. Significantly, the relative magnitudes of the average clustering coefficient values are higher in the mixed strategy model, suggesting that the social network helps to promote higher levels of cooperation. The underlying link adjustment mechanisms and mixed strategy decision-making directly favor the formation of cliques. These findings are consistent with results reported in [17]. It is also interesting to note that the values of the average clustering coefficient values when $N = 2$ are very similar in both models, confirming that the additional cognitive ability of the mixed strategy agents is not a requirement to promote cooperation in the restricted game.

In Figure 4 we provide two examples of final degree distribution plots on a log-log scale of the social network in the mixed strategy model (space constraints limit the inclusion of further plots). For all scenarios examined, a clear hierarchical structure was observed. The general trends in degree distribution plots were similar. When it was possible to sustain cooperation in a game (for $2 \leq N \leq 10$), the degree distribution was typically bimodal where the size of the gap and shape of the distribution was dependent on N. For larger values of N, the plot shows a large number of nodes with high average degrees corresponding to a very dense network. This suggests that densely knit cliques may be stable and preserve links. There are also a smaller number of nodes with lower average degrees corresponding to newly formed cooperative cliques.

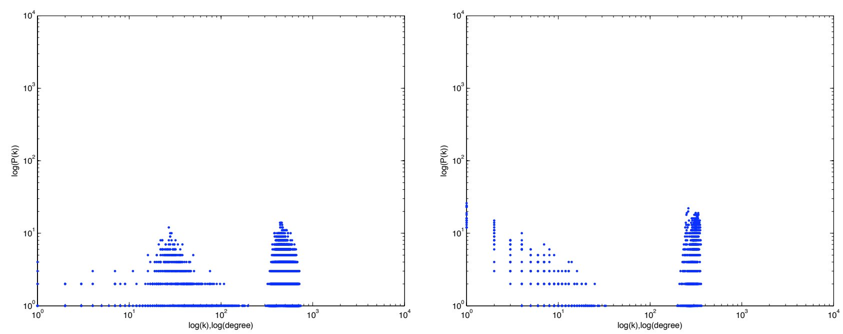
5 Summary and conclusion

In this paper, we have investigated the coevolution of strategies and social networks using a version of the N-player prisoner's dilemma game. An important component of our model is the endogenous network formation based on agent interactions. Agents who play cooperatively form social links, which are reinforced



(a) pure strategy (b) mixed strategy

Fig. 3: Average clustering coefficient vs. time for various values of N.



(a) $N = 2$ (b) $N = 10$

Fig. 4: Final degree distributions in the mixed strategy model on a log-log scale.

by subsequent cooperative actions. However, when an agent defects, links in the social network are broken. Simulation results validate our approach and confirm that cooperation is promoted by agents exploiting the social network to guide their decision-making. However, as the number of players participating in the game increases, the proportion of cooperations within the population decreases.

We have examined the long term system dynamics of our model based on two different agent cognitive abilities: agents with pure strategies who always play one of cooperate or defect, and a mixed strategy where agents based their decision as to which action to play on the reputation or reliability of other agents. When agents have increased cognitive capacity to classify their environment, social networks play an increasingly important role in promoting and sustaining cooperation. An analysis of the emergent social networks shows that they are characterized by high average clustering and broad-scale heterogeneity, especially for a relatively small number of players. The key findings in this study suggest that the interplay between the local structure of the network and the hierarchical organization of cooperation is non-trivial.

The notion that a group of individuals can collectively do better when they cooperate, often against the self-interest of individuals, is relevant to many research domains, including social and biological systems, economics, artificial intelligence and multi-agent systems. In future work, we plan to extend this study in a number of different directions. In the present study, we have assumed that all groups were of a fixed size during a given game. It would be interesting to examine the effects of dynamic group formation on the coevolution of cooperation and social networks. A second avenue worth investigating is specific aspects of trust/reputation in groups within a peer-to-peer framework.

References

1. R. Axelrod. *The Evolution of Cooperation.* Basic Books, New York, 1984.
2. E. Bilancini and L. Boncinelli. The co-evolution of cooperation and defection under local interaction and endogenous network formation. *Journal of Economic Behavior and Organization*, 70(1–2):186–195, 2009.
3. R. Boyd and P. Richerson. The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, 132:337–356, 1988.
4. S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of Networks: From Biological Nets to the Internet and WWW.* Oxford University Press, 2003.
5. T. S. Ellis and X. Yao. Evolving cooperation in the non-iterated prisoners dilemma: A social network inspired approach. *IEEE Cong. on Evol. Comp. (CEC)*, pages 736–743, September 2007.
6. M. Feldman, K. Lai, I. Stoica, and J. Chuang. Robust incentive techniques for peer-to-peer networks. In *Proc. 5th ACM Conf. Electronic Commerce (EC 04)*, pages 102–111. ACM Press, 2004.
7. D. Hales and S. Arteconi. SLACER: A self-organizing protocol for coordination in peer-to-peer networks. *IEEE Intelligent Systems*, 21(2):29–35, 2006.
8. G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.
9. C. Hauert. Fundamental clusters in spatial 2×2 games. *Proc. R. Soc. Lond. B*, 268:761–769, 2001.
10. R. Ismail, A. Josang, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618–644, 2007.
11. A. Kun and I. Scheuring. Evolution of cooperation on dynamical graphs. *BioSystems*, 96(1):65–68, 2009.
12. M. Nowak and K. Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393:573–577, 1998.
13. M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.
14. H. Ohtsuki, C. Hauert, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, 2006.
15. F. C. Santos, M. D. Santos, and J. M. Pacheco. Social diversity promotes the emergence of cooperation in public goods games. *Nature*, 454:213–216, 2009.
16. S. Suzuki and E. Akiyama. Reputation and the evolution of cooperation in sizable groups. *Proceedings of the Royal Society B*, 272:1373–1377, 2005.
17. J. Tanimoto. Promotion of cooperation through co-evolution of networks and strategy in 2x2 games. *Physica A*, 96:953–960, 2009.