Maximal Cooperation in Repeated Games on Social Networks

Catherine Moon
Duke University
Durham, NC 27708, USA
csm17@duke.edu

Vincent Conitzer
Duke University
Durham, NC 27708, USA
conitzer@cs.duke.edu

Abstract

Standard results on and algorithms for repeated games assume that defections are instantly observable. In reality, it may take some time for the knowledge that a defection has occurred to propagate through the social network. How does this affect the structure of equilibria and algorithms for computing them? In this paper, we consider games with cooperation and defection. We prove that there exists a unique maximal set of forever-cooperating agents in equilibrium and give an efficient algorithm for computing it. We then evaluate this algorithm on random graphs and find experimentally that there appears to be a phase transition between cooperation everywhere and defection everywhere, based on the value of cooperation and the discount factor. Finally, we provide a condition for when the equilibrium found is credible, in the sense that agents are in fact motivated to punish deviating agents. We find that this condition always holds in our experiments, provided the graphs are sufficiently large.

1 Introduction

In systems of multiple self-interested agents, we cannot impose behavior on the agents. For desired behavior to take place, it will have to be in equilibrium, but this is not necessarily the case, especially in one-shot games. It is well known in game theory that a much greater variety of behavior can be obtained in the equilibria of infinitely repeated games than in the equilibria of one-shot games. The standard example is that of the prisoner's dilemma. Defecting is a dominant strategy for both players in the one-shot version. In the repeated version, cooperation between the players can be sustained due to the threat of the loss of future cooperation. Consequently, modeling repeated play and solving for equilibria of the resulting games are crucial to the design of systems of multiple self-interested agents.

The well-known *folk theorem* in game theory characterizes the payoffs that agents can obtain in the equilibria of repeated games. This theorem also serves as the basis of algorithms that compute equilibria of repeated games, which can be done in polynomial time for 2-player games [Littman and Stone, 2005]. (This does not extend to games with more than 2 players [Borgs *et al.*, 2010], unless correlated punishment is possible [Kontogiannis and Spirakis, 2008]. Recent work shows that heuristic algorithms can nevertheless be effective [Andersen and Conitzer, 2013].)

These results operate under the assumption that an agent's behavior is instantly observable to all other agents. For reasonably large multiagent systems, this can be a very limiting restriction. When an agent does not interact with another agent, it may take some time before one finds out about the other's defection. In such cases, it is more difficult to sustain cooperative behavior in equilibrium, because the punishment for defection will arrive later in the future and therefore be more heavily discounted. Then, under what conditions can we still sustain cooperation, and can we compute the resulting equilibria? These are the questions we set out to answer in this paper.

Graphical games [Kearns *et al.*, 2001] constitute a natural model for the interaction structure. One shortcoming of graphical games is that typically, the graph is undirected. Here, we are also interested in modeling directed relationships, where b is affected by a's actions but not vice versa. Of course, this can be represented using an undirected edge as well, with some duplication of values in the utility table. However, besides being concerned with computational efficiency, we also want the edges to capture the flow of information. For example, suppose b is affected by a's actions, c is affected by b's actions, and a is affected by c's actions (but not vice versa). If a defects, c will initially not be aware of this. However, b will be, causing b to defect in the second round. At that point c will become aware of a defection, resulting in a defection by c in the third round, perhaps corresponding to a receiving a delayed punishment for defecting in the first round. In any case, allowing for directed edges in graphical games is a trivial modification. (There are other graphical models that are used to represent games, such as MAIDs [Koller and Milch, 2003] and action-graph games [Jiang *et al.*, 2011], but graphical games provide the structure we need.)

2 Related Literature

Many papers on repeated public good games show full cooperation is possible when players become arbitrarily patient [Kandori, 1992; Ellison, 1994; Deb, 2008; Takahashi, 2010], even with delayed monitoring [Kinateder, 2008], imperfect monitoring [Laclau, 2012], and local monitoring [Nava and Piccione, 2014] in graphical games. However, there is not much work characterizing the maximum level of cooperation sustainable in equilibrium. As an exception, Wolitzky (2013) explores the maximum level of contributions sustainable in equilibrium for a fixed discount factor for public good games under private "all or nothing" monitoring (player i changes her monitoring neighborhoods every period, and perfectly observes contributions made within the monitoring neighborhood) and shows that grim-trigger strategies provide the strongest possible incentives for cooperation on-path. This work relates to ours as cooperation in our framework can be interpreted as providing a local public good, where the access to the "local public good" is asymmetric and only the player who benefits from the cooperation observes the cooperation. Given these complications, our model's decision variables are discrete (cooperate or defect), unlike Wolitzky's which are continuous. Also closely related is work by Mihm et al. [Mihm *et al.*, 2010], which also concerns sustaining cooperation on social networks in repeated play. An important technical difference is that in their model, edges necessarily transmit information in both directions. If we were to consider the special case of our model where every edge transmits information in both directions, our results would become rather trivial because a deviating agent would be instantly punished by all its neighbors; on the other hand, Mihm et al. allow agents to play different actions for each of their neighbors, which we do not allow. Mihm et al. also require the game played to be the same on every edge. In evolutionary game theory, repeated games on graphs are also sometimes considered, but here the network structure is used to let agents copy each others strategies (e.g., imitate the best-performing neighbor).

In our paper, we focus on finding the maximal set of cooperating agents in equilibrium for a given discount factor and experimentally investigating the range of discount factors under which cooperation can be sustained. We also assess the credibility of the threats used in this equilibrium: we give a condition for threats to be credibly executed and experimentally investigate when this condition holds for the maximal set of cooperating agents.

3 Model

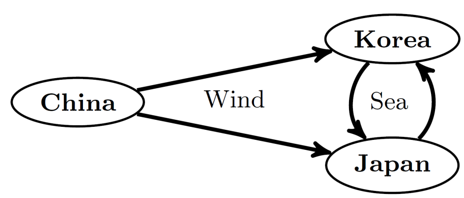
We suppose that there is a set of agents A organized in a directed graph $G = (V, E)$ where $V = A$, i.e., agents are vertices in the graph. A directed edge indicates that the agent at the tail of the edge benefits if the agent at the head of the edge cooperates. Associated with every agent (vertex) i is a cooperation cost $\kappa_i \in \mathbb{R}$. Associated with every directed edge $(i, j) \in E$ is a cooperation benefit $\beta_{i,j} \in \mathbb{R}$. Agents play a repeated game, where in each round, each agent either cooperates or defects. Note that it is not possible for an agent to cooperate for some of its outgoing edges but defect for others within the same round. In round t, i receives $-x_{i,t}\kappa_i + \sum_{j:(j,i)\in E} x_{j,t}\beta_{j,i}$, where $x_{i,t} \in \{0, 1\}$ indicates whether i cooperated in round t or not. That is, agent i experiences cost κ_i for cooperating and benefit $\beta_{j,i}$ for j cooperating. Defection is irreversible, i.e., once an agent has defected it will defect forever, thereby effectively destroying its outgoing edges. At the end of round t, agent i learns, for every j with $(j, i) \in E$, whether j cooperated or defected in round t, and nothing else. As is standard in repeated games, a *strategy* of a player maps the entire history that that player observed to an action. Payoffs in round t are discounted by a factor δ^t. We are interested in Nash equilibria of this game (and, later, in equilibrium refinements).

4 Motivation and Illustrative Examples

Our directed graph framework enables us to model asymmetry in both payoffs and information flow. (Undirected graphs can be seen as the special case of our framework where edges always point in both directions.) In this section, we illustrate our model using some examples.

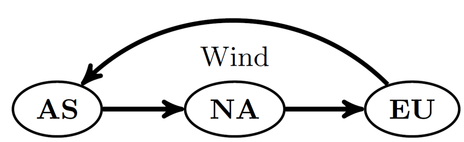
4.1 A Pollution Reduction Example

Suppose there is a set of countries, {China, South Korea, Japan}, gathering together to try to reduce pollution. Geographically, South Korea and Japan share the Sea of Japan (also called the East Sea of Korea), and China is located to the west of both countries. Consequently, if South Korea does not reduce its pollution, Japan would find out and suffer due to the pollutants in the sea, and vice versa; but China would not be affected in either case. If China does not reduce its pollution, however, both South Korea and Japan would notice and suffer, with the pollutants traveling with the dominant west wind.



In this simplified model, South Korea and Japan do not have a way to punish China even if China decides to not reduce pollution. However, Japan and South Korea may be able to negotiate and enforce an agreement between themselves, because when one of the two defects, the other can punish the former (enough) by defecting back. Thus, our model allows us to solve for the set of agents that would be able to negotiate an enforceable agreement.

4.2 Delayed Punishment due to Directionality of Information Flow



As a similar example, consider three agents, North America (NA), Europe (EU), and Asia (AS) negotiating a pollution reduction agreement. With the dominant wind from the

West, North America would notice Asia's pollution, Europe would notice North America's pollution, and Asia would notice Europe's pollution. Hence, defection in North America could trigger Europe to defect in the next period, which in turn could trigger Asia to defect in the period after, at which point North America experiences a delayed punishment.

4.3 Discussion of Assumptions in the Model

One of our assumptions is that it is not possible for an agent to defect on one neighbor but to simultaneously cooperate with another neighbor. This makes sense in the examples above—e.g., it is not possible for China to pollute only towards Japan but not towards South Korea. Moreover, the cost of reducing pollution does not depend on one's neighbors.

Another assumption is that defection is irreversible. Our results depend on this assumption: dropping it may allow arrangements with partial cooperation (where, for example, one agent is expected only to cooperate every other period) which we do not consider here. In any case, this assumption is reasonable in cases where continued cooperation requires upkeep—e.g., once you install a dirty coal-fired power plant it may be hard to shut it down; in accounting of nuclear resources, cooperation may mean following a tight protocol, and once an agent defects for a period it may become impossible to re-account for all the resources properly; etc.

5 Theoretical Analysis for Cooperation in Nash Equilibrium

We will show that we can without loss of generality restrict our attention to *grim-trigger* equilibria, in which there is some subset of agents $S \subseteq A$ that each cooperate until another agent in S defects on them.

Definition 1. *Player i's strategy* grim trigger for S *consists of i cooperating until some player $j \in S$ with $(j,i) \in E$ defects, which is followed by i defecting (forever).*

Definition 2. *The* grim trigger profile $T[S]$ *consists of all the agents in S playing the grim trigger strategy for S, and all other players always defecting.*

When one agent's defection triggers another's defection, the latter can set off further defections. This cascade of defections is what lets the information that there has been a defection travel through the graph. Accordingly, to determine how long it takes an agent to find out about the original defection, we need a definition of distance that takes into account that information is transmitted only through the (originally) cooperating agents.

Definition 3. *For any subset S of the agents and any $i, j \in S$, the* distance from i to j through S*, denoted $d(i,j,S)$, is the length of the shortest path from i to j that uses only agents in S. For a set of agents $G \subseteq S$, $d(G,j,S) = \min_{i \in G} d(i,j,S)$.*

The next proposition establishes a necessary and sufficient condition for $T[S]$ to be a Nash equilibrium.

Proposition 1. (Equilibrium) $T[S]$ is an equilibrium if and only if for all $i \in S$, $\sum_{j \in S:(j,i)\in E} \delta^{d(i,j,S)} \beta_{j,i} \geq \kappa_i$.

Proof. First observe that every player outside S is best-responding, because her actions do not affect any other players' future actions and within any single round, defecting is a strictly dominant strategy. For a player i in S, without loss of generality we can focus on whether she would prefer to defect (starting) in the first round. If i were to defect, the total utility gain from reduced effort are exactly $\sum_{t=0}^{\infty} \delta^t \kappa_i = \kappa_i/(1-\delta)$, and the total utility loss from neighbors eventually defecting as a result of i's defection is $\sum_{j \in S:(j,i)\in E} \sum_{t=d(i,j,S)}^{\infty} \delta^t \beta_{j,i} = \sum_{j \in S:(j,i)\in E} \delta^{d(i,j,S)} \beta_{j,i}/(1-\delta)$. The latter follows from the fact that i's defection will cause j to defect exactly $d(i,j,S)$ rounds later. Thus, i has no incentive to defect if the former is no greater than the latter. Multiplying both sides by $1-\delta$ gives the desired inequality. □

The next proposition shows that no equilibria can have more agents cooperate forever than the grim-trigger equilibria. Intuitively, this is because grim-trigger profiles provide the maximum possible punishment for deviating agents.

Proposition 2. (Grim Trigger is WLOG) Suppose there exists a pure-strategy equilibrium in which S is the set of players that cooperates forever. Then $T[S]$ is also an equilibrium.

Proof. For an arbitrary player $i \in S$, we must prove the inequality in Proposition 1. For the given equilibrium, consider some period τ at which every player outside S has defected (on the path of play). If player i considers defecting at this point, the total utility gain from reduced effort would be exactly $\sum_{t=\tau}^{\infty} \delta^t \kappa_i = \delta^\tau \kappa_i/(1-\delta)$. On the other hand, the total utility loss from neighbors defecting as a result of i's defection is at most $\sum_{j \in S:(j,i)\in E} \sum_{t=\tau+d(i,j,S)} \delta^t \beta_{j,i} = \delta^\tau \sum_{j \in S:(j,i)\in E} \delta^{d(i,j,S)} \beta_{j,i}/(1-\delta)$. The latter follows from the fact that i's defection can only cause changes in j's behavior $d(i,j,S)$ rounds later, because no information can pass through nodes that have already defected. By the equilibrium assumption, the latter expression is at least as large as the former. Multiplying both by $(1-\delta)/\delta^\tau$ gives the desired inequality. □

The next lemma shows that the more other agents cooperate, the greater the incentive to cooperate.

Lemma 3. (Monotonicity) If $S \subseteq S'$ and the incentive constraint from Proposition 1 holds for i relative to S (so $i \in S$, $\sum_{j \in S:(j,i)\in E} \delta^{d(i,j,S)} \beta_{j,i} \geq \kappa_i$), then it also holds for i relative to S'.

Proof. We argue that $\sum_{j \in S':(j,i)\in E} \delta^{d(i,j,S')} \beta_{j,i} \geq \sum_{j \in S:(j,i)\in E} \delta^{d(i,j,S)} \beta_{j,i}$, which establishes that the former is also at least κ_i. First, all summands are nonnegative, and the former expression has a summand for every j for which the latter expression has a summand. Second, for any i, j, we have that $d(i,j,S') \leq d(i,j,S)$, because having additional agents cannot make the distance greater. Because $\delta < 1$, it follows that $\delta^{d(i,j,S')} \geq \delta^{d(i,j,S)}$. Hence, for j for which both expressions have a summand, the summand is at least as large in the former expression (corresponding to S') as in the latter (corresponding to S). This establishes that the former expression is at least as large. □

The next proposition shows that we cannot have multiple distinct maximal grim-trigger equilibria.

Proposition 4. (Maximality) If $T[S]$ and $T[S']$ are both equilibria, then so is $T[S \cup S']$.

Proof. Consider some $i \in S \cup S'$; without loss of generality, suppose $i \in S$. We must show that the incentive constraint from Proposition 1 holds for i relative to $S \cup S'$. But this follows from Lemma 3 and the facts that it holds for i relative to S and that $S \subseteq S \cup S'$. □

Proposition 4 implies that there exists a unique maximal set S^* of forever-cooperating agents such that $T[S^*]$ is an equilibrium. We can use the following algorithm for finding the unique maximal set S^*. Initialize S^{current} to include all agents. Then, in each iteration, check, for each $i \in S$, whether the incentive constraint holds for i relative to S^{current}. Remove those i for which it does not hold from S^{current}. Repeat this until convergence; then, return S^{current}. Call this Algorithm 1, presented formally below.

Algorithm 1

$D_{\text{elimination}} \leftarrow$ true
while $D_{\text{elimination}} =$ true **do**
    $D_{\text{elimination}} \leftarrow$ false
    $L_{\text{defectors}} \leftarrow \emptyset$
    $I \leftarrow$ IncomingEdges(E)
    $O \leftarrow$ OutgoingEdges(E)
    for all $i \in S^{\text{current}}$ **do**
        $L^i \leftarrow$ ShortestPath(i, j, E) ($\forall j \in S^{\text{current}}$)
        $C^i \leftarrow$ IncentiveConstraint($L^i, \kappa_i, \{\beta_{j,i}\}_{j:(j,i)\in E}, \delta$)
        if C^i = false **then**
            add i to $L_{\text{defectors}}$
        end if
    end for
    if $L_{\text{defectors}} \neq \emptyset$ **then**
        $D_{\text{elimination}} \leftarrow$ true
        for all $i \in L_{\text{defectors}}$ **do**
            remove (i, j) from E for all j
            remove i from S^{current}
        end for
    end if
end while

IncentiveConstraint($L^i, \kappa_i, \{\beta_{j,i}\}_{j:(j,i)\in E}, \delta$) checks the incentive constraint inequality in Proposition 1. If it is satisfied, it returns true, indicating i's willingness to cooperate.

Proposition 5. (Correctness) Algorithm 1 returns the unique maximal set S^* such that $T[S^*]$ is an equilibrium.

Proof. It suffices to show that if i is at some point eliminated from S^{current} in Algorithm 1, then there exists no set S such that $i \in S$ and $T[S]$ is an equilibrium. We prove this by induction on the round in which i is eliminated. Suppose it holds for all rounds before t; we prove it holds when i is eliminated in round t. Let S^t denote the set of agents that have not yet been eliminated in round t (including i). By the induction assumption, for any S such that $T[S]$ is an equilibrium, we have $S \subseteq S^t$. But the incentive constraint from Proposition 1 does not hold for i relative to S^t, because i is eliminated in this round. But then, by Lemma 3, it also does not hold for i relative to any $S \subseteq S^t$. Hence, there is no S such that $i \in S$ and $T[S]$ is an equilibrium. □

Proposition 6. (Runtime) The runtime of Algorithm 1 is $O(|V|^2 |E| \log |V|)$.

Proof. Because at least one agent is removed in every iteration before the last, there can be at most $|V| + 1$ iterations. In each iteration, we solve all-pairs shortest paths, which can be done in O($|V| |E| \log |V|$) time [Cormen *et al.*, 2001]. Also, in each iteration, for every agent, we need to evaluate whether the incentive constraint holds, requiring us to sum over all its incoming edges. Because each edge has only one vertex for which it is incoming, we end up considering each edge at most once per iteration for this step, so that its runtime is dominated by the shortest-paths runtime. □

6 Experimental Analysis on Random Graphs

In this section, we evaluate the techniques developed so far experimentally. After this, we continue with further theoretical development, which we subsequently experimentally evaluate as well.

6.1 Assumptions for Simulation

For our experimental analysis, we make the following additional assumptions on the cost and benefit structure. First, i's cost of cooperation is proportional to the number of outgoing edges i initially has: $\kappa_i = \sum_{i \in S:(i,j)\in E} 1$, normalizing the per-edge cost to 1. Also, we assume a constant benefit to having an incoming edge, that is, $\beta_{j,i} = \beta$ for all $(j, i) \in E$. This implies that the total benefit i receives is proportional to the number of incoming edges i has (from cooperating agents).

We generate random graphs based on two models: the Erdős-Rényi random graph model (ER) and the Barabási-Albert preferential-attachment random graph model (PA), with some modifications to generate directed rather than undirected graphs. The ER model requires n, the number of nodes, and p, the probability of a directed edge between two nodes. The PA model requires n, the number of nodes, e, the number of edges per node, and γ, a parameter determining the degree of "preference" given to nodes with more edges when a new node enters the network in the formation process ($\gamma = 0$ makes the probability of an existing node being selected for edge generation directly proportional to its current degree, whereas $\gamma = 1$ makes the probability of being selected equal among all the currently existing nodes). To obtain directed graphs from PA, we first turn any undirected edge into two directed edges (directed both ways). We then add $p_{\text{degeneration}}$, the probability of removing a directed edge between any two nodes (we calibrate e and $p_{\text{degeneration}}$ to keep p, the probability of a directed edge between two nodes, comparable to the ER cases and the analytical expression). All the graphs presented in this paper are with $p = 0.3$, and with each point averaged over 100 graphs. This probability is chosen as a representative value for presentation, but the patterns shown are consistent with graphs generated using different p values.

6.2 Equilibrium Defection Phase Transition

Given a graph, there are two parameters that can further affect i's incentive to cooperate or defect. Varying β will change the value of incoming edges (relative to the cost of outgoing edges). On the other hand, varying δ will affect how an agent trades off current and future payoffs. Thus, we apply Algorithm 1 for varying values of δ and β.

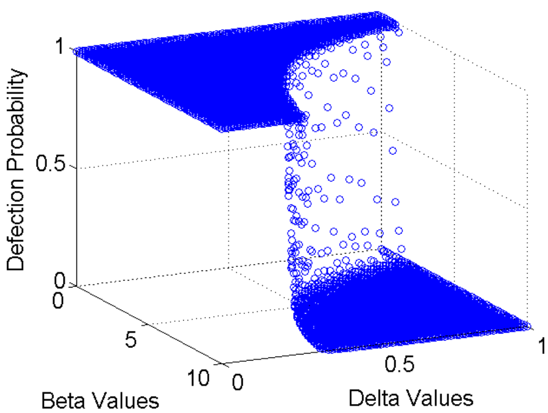


Figure 1: Nodes' defection probabilities—the fractions of nodes that fail to cooperate in the maximally cooperative equilibrium—for values of β and δ (ER: $n = 100$, $p = 0.3$).

Figure 2 shows the resulting equilibrium cooperation and defection patterns for representative ER and PA specifications for different values of n. These are top-down views of graphs of the form in Figure 1. For all the gradient graphs, the transition pattern is similar to that of Figure 1. We see an apparent phase transition, with a sudden sharp drop in the defection probability, indicating a transition from everyone defecting to everyone cooperating. This implies that β and δ suffice to predict whether a node will cooperate; in particular, knowing a node's centrality in the graph does little to help predict cooperation of that node, because it tends to be the case that the whole graph cooperates or defects together.

Intuitively, we might expect a phase transition for the following reason. For high values of β and δ, it will be rare that any agents have an incentive to defect when all other agents play the grim trigger strategy. However, once we drop to values where some agents have an incentive to defect, we can no longer have these agents participate in the grim trigger strategy. This then increases the incentives for other agents to defect, so some of them no longer participate either, etc., quickly leading to a collapse of all cooperation.

As pointed out in the proof of Proposition 6, the algorithm returns the equilibrium after at most $|V| + 1$ iterations, but this is a very pessimistic upper bound, assuming exactly one agent is eliminated each iteration. Experimentally, significantly fewer iterations are required for convergence, as indicated in Figure 3. On average, the algorithm converges within at most 12 iterations. Moreover, we see the typical pattern where runtime is greatest around where the phase transition happens. This pattern confirms our intuition that the cascade of defection set off by the initial defection results in the observed phase transition. Further away from the band, the algorithm converges within an average of 2 iterations.

7 Credibility of Threats in Equilibrium

So far, we have focused on the maximal set of cooperating agents, S^*, in Nash equilibrium. However, these equilibria involve threats of grim-trigger defection that may not be credible. Suppose agent 1 has one otherwise isolated neighbor, agent 2, that has defected. If agent 1 defects, as she is supposed to in the grim trigger equilibrium, then she will also set off defections in the rest of the network, which may come at a significant cost to her. On the other hand, if she ignores the defection and continues to cooperate, then the other agents will never learn of the defection (since agent 2 is otherwise isolated) and continue to cooperate. The latter may be better for agent 1, in which case the grim trigger strategy is not credible. Hence, if we impose an additional credibility condition on the equilibrium, the maximal set of nodes cooperating in equilibrium may be strictly smaller than S^* (it will always be a subset).

In this section, we will give a condition for the grim trigger strategy to be credible for a given set of agents in the network ($S \subseteq A$). Hence, when $T[S]$ is a Nash equilibrium and the condition holds, then $T[S]$ is also a credible equilibrium. As we will later show, for sufficiently large networks, this condition always holds in our experiments for S^*.

So far, we have avoided a precise definition of when equilibria are credible. Of course, the notion of threats that are not credible is common in game theory, motivating refinements of Nash equilibrium such as subgame-perfect Nash equilibrium. Subgame-perfect equilibrium will not suffice for our purposes, because generally our game has no proper subgames: the acting agent does not know which moves were just taken by agents at distance 2 or more. Hence, subgame perfection does not add any constraints. We need a stronger refinement. Common stronger equilibrium refinements (such as perfect Bayesian equilibrium or sequential equilibrium) generally require specifying the beliefs that the players have at information sets that are off the path of play. In fact, in our game, we can obtain the following very strong refinement, which we will call *credible equilibrium*: if an agent learns that some deviation from the equilibrium has taken place, then she will be best off following her equilibrium strategy (e.g., grim trigger) *regardless of her beliefs about which deviations have taken place*, assuming that the other agents also follow their equilibrium strategies from this round on.[1] We

[1] A similar notion from the literature on repeated games with imperfect private monitoring is that of *belief-free equilibrium* [Ely *et al.*, 2005], in which an agent's continuation strategy is optimal regardless of her beliefs about opponents' histories. This, however, is not true for the strategies we study: specifically, if agent i is in a situation where all agents other than i's own neighbors have defected, and all i's neighbors will defect in the next period, then i would be better off defecting this period—but she will not do so un-

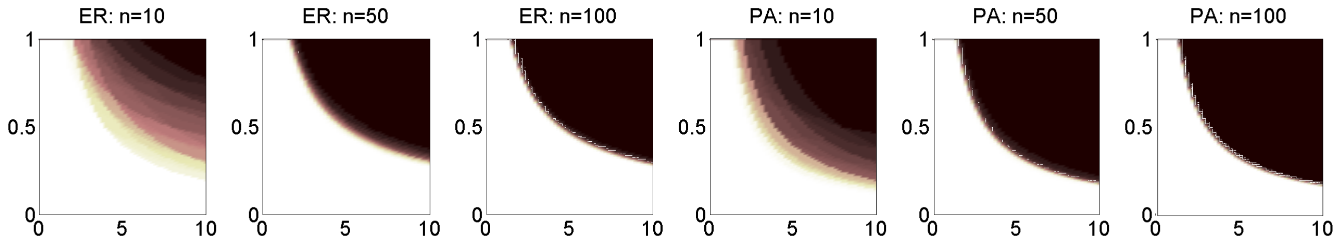


Figure 2: Gradient figures (top-down view of graphs of the form in Figure 1) showing nodes' defection probabilities for different values of β and δ, for both ER and PA graphs with various numbers of nodes n. The x-axis shows the β values, ranging from 0 to 10, and the y-axis shows the δ values, ranging from 0 to 1. The white section indicates complete defection, while the black section indicates complete cooperation. Intermediate shades indicate partial cooperation.

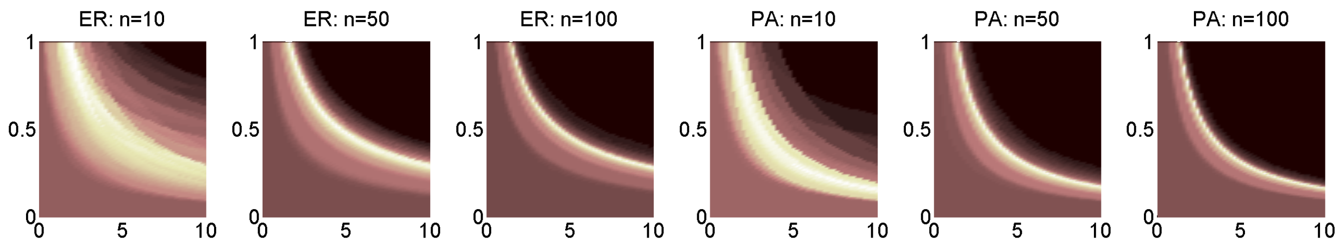


Figure 3: Gradient figures showing the average number of iterations needed until convergence of the algorithm, for various values of β and γ. The specifications of each panel are the same as in Figure 2. In the bottom left, where we see total collapse of cooperation, Algorithm 1 converges within 2 iterations on average. In the top right, where we see total cooperation, Algorithm 1 (of course) converges within 1 iteration. The middle band requires more iterations (maximally 9 on average).

now proceed towards deriving our condition for $T[S^*]$ to be a credible equilibrium. The first lemma shows that we can restrict our attention to only a single neighbor defecting.

Lemma 7. (Sufficiency of Singleton Deviations) Suppose that, for agent i, there is a set of agents K (with $i \notin K$ and $K \cap \{k : (k,i) \in E\} \neq \emptyset$) such that if i believes that K is the set of all agents that have defected so far, and all other agents will play grim trigger from now on, then i prefers postponing defection for r ($1 \leq r \leq \infty$) rounds to defecting immediately (i.e., grim trigger is not credible). Then, for any $k \in K \cap \{k : (k,i) \in E\}$, if i believes that $\{k\}$ is the set of all agents that have defected so far, and all other agents will play grim trigger from now on, i prefers postponing defection for r rounds to defecting immediately as well.

Proof Sketch: The intuition for the lemma is that, in both scenarios, the cost of cooperating for r more rounds is the same, but the benefit from postponing defection—which is that other nodes will cooperate longer—is always at least as large in the second case. □

Lemma 8. (One-Round Postponement) Suppose that the following holds: if i believes that $\{k\}$ (with $(k,i) \in E$) is the set of all agents that have defected so far, and all other agents will play grim trigger from now on, then i prefers postponing defection for r ($1 \leq r \leq \infty$) rounds to defecting immediately (i.e., grim trigger is not credible). Then, in these circumstances, i will also prefer postponing defection for exactly 1 round to defecting immediately.

Proof Sketch: The intuition for the lemma is that the longer agent i waits to defect, the fewer other agents will be informationally affected—in the sense that they learn about a defection having taken place later—by i deciding to wait one additional round. This is because increasingly many agents will learn about the defection via paths not involving i. Hence, the incentive to wait is strongest in the first round. □

Theorem 9. (Credible Equilibrium) Suppose that for some set $S \subseteq A$, $T[S]$ is a Nash equilibrium. Then $T[S]$ is a credible equilibrium if and only if for any $k, i \in S$ with $(k,i) \in E$, it holds that $\kappa_i - \delta^{d(k,j,S)} \sum_{j \in D} \beta_{ji} \geq 0$, where $D = \{j \in I : d(k,j,S) + 1 \leq d(k,j,S_{-i})\}$.

Proof. Because $T[S]$ is a Nash equilibrium, it is credible if and only if immediate defection upon learning of a deviation is credible. The condition in the theorem is equivalent to saying that when a single neighbor has deviated alone, defecting immediately is better than waiting one round to defect. By Lemma 8, this implies that under these circumstances, defecting immediately is optimal. By Lemma 7, this implies that defecting immediately is optimal whenever the agent believes some nonempty subset has defected. □

8 Extension: Experimental Analysis of the Credible Equilibrium

In our previous experimental section, we did not consider whether $T[S^*]$ was a credible equilibrium. In this experimental section, we use the condition in Theorem 9 to determine

der the grim-trigger strategy, as she is not yet aware of any defection having taken place. In contrast, our notion of credible equilibrium conditions on the agent knowing that a defection has taken place.

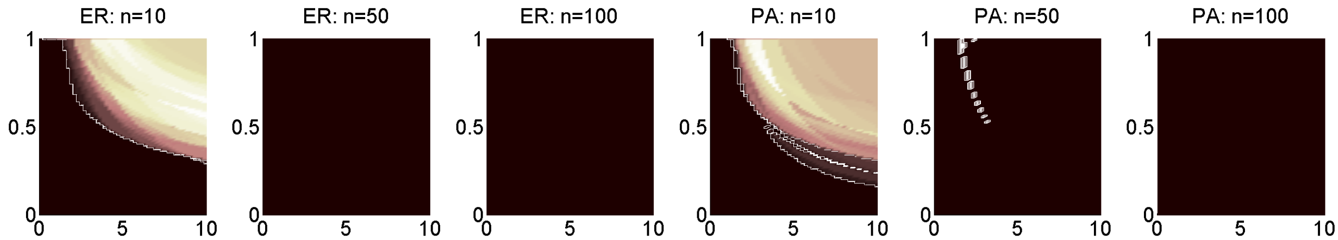


Figure 4: Gradient figures showing the fraction of cases in which $T[S^*]$ is a credible equilibrium, for different values of β and δ. The specifications of each panel are the same as in Figure 2. In the bottom left of each panel (black section), $T[S^*]$ is always credible; in the top right it is not always (when it is brighter), though as n grows quickly the condition starts holding everywhere. In the $n = 10$ cases, even for the combination of β and δ values where grim trigger is least credible, $T[S^*]$ is a credible equilibrium about 60 percent of the time.

when this is the case. Figure 4 shows the fraction of instances for which $T[S^*]$ is a credible equilibrium. Of course, this is always true when $S^* = \emptyset$. Generally, when the credibility condition fails, it is in the region where both the discount factor δ and the benefit multiplier β are high. This makes intuitive sense: if these parameters are low, there is little reason to postpone defection. More significantly, we see that as n increases, the fraction of cases where the condition holds quickly converges to 1 everywhere. This, too, makes intuitive sense: the main reason to postpone defection is to slow down the spread of the information that a defection has taken place. However, the larger the network, the less likely it is that an individual node can do much to keep this information from spreading.

9 Conclusion

In this paper, we considered the following problem. In repeated games in which the agents are organized in a social network, it may take more than one round for an agent to find out about a defection that happened elsewhere in the graph. If so, it may increase incentives for agents to defect, because the losses from resulting punishment will be delayed and therefore discounted more heavily. We restricted our attention to games in which the agents can either cooperate or defect. We proved that there exists a unique maximal set of forever-cooperating agents in such games. We also gave an efficient algorithm for computing this set, which relies on iteratively removing agents from the set that cannot possibly be incentivized to cooperate forever, based on which agents are still in the set. We evaluated this algorithm on randomly generated graphs and found an apparent phase transition: when the relative cooperation benefit β and the discount factor δ are high enough, all agents can cooperate forever, but once these are lowered beyond a threshold, we get a total collapse, with no agents cooperating. Lastly, we gave an easy-to-check condition for when the threats in the equilibrium are credible, and found experimentally that for large graphs this condition always holds.

10 Future research

One direction for future research is to generalize the techniques in our paper to more general games on networks, in which agents' action spaces are not restricted to cooperation and defection. We expect that many of the same phenomena would occur in such games, and similar techniques would apply, but several additional technical challenges would have to be overcome. The main issue is that in sufficiently general games, multiple agents would need to coordinate their punishment actions for them to be maximally effective (unlike in the game studied here, where defection is always maximally effective). Such coordination is known to pose computational challenges even without network structure [Kontogiannis and Spirakis, 2008; Hansen *et al.*, 2008; Borgs *et al.*, 2010; Andersen and Conitzer, 2013], and in our context there will be further challenges in coordinating punishment because information spreads slowly.

Another direction for future research would be to mathematically prove the existence of the phase transition, as well as that all maximal equilibria become credible as n grows. In both cases, it is not difficult to give rough arguments, but a careful proof will probably require the use of results in probability theory, such as concentration bounds.

Acknowledgments

We thank ARO and NSF for support under grants W911NF-12-1-0550, W911NF-11-1-0332, IIS-0953756, CCF-1101659, and CCF-1337215. We are also grateful to Rachel Kranton and Todd Sarver for helpful comments on the paper, and to participants in Duke University's Microeconomic Theory and DRIV seminars for helpful discussions. Lastly, we thank the anonymous referees for their comments and suggestions.

References

[Andersen and Conitzer, 2013] Garrett Andersen and Vincent Conitzer. Fast equilibrium computation for infinitely repeated games. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, pages 53–59, Bellevue, WA, USA, 2013.

[Borgs *et al.*, 2010] Christian Borgs, Jennifer Chayes, Nicole Immorlica, Adam Tauman Kalai, Vahab Mirrokni, and Christos Papadimitriou. The myth of the Folk Theorem. *Games and Economic Behavior*, 70(1):34–43, 2010.

[Cormen *et al.*, 2001] Thomas Cormen, Charles Leiserson, Ronald Rivest, and Clifford Stein. *Introduction to Algorithms*. MIT Press, second edition, 2001.

[Deb, 2008] Joyee Deb. Cooperation and community responsibility: A folk theorem for repeated matching games with names. NYU Working Paper Series, 2008.

[Ellison, 1994] Glenn Ellison. Cooperation in the prisoner's dilemma with anonymous random matching. *Review of Economic Studies*, 61:567–588, 1994.

[Ely *et al.*, 2005] Jeffrey C. Ely, Johannes Horner, and Wojciech Olszewski. Belief-free equilibria in repeated games. *Econometrica*, 73(2):377–415, 2005.

[Hansen *et al.*, 2008] Kristoffer Arnsfelt Hansen, Thomas Dueholm Hansen, Peter Bro Miltersen, and Troels Bjerre Sørensen. Approximability and parameterized complexity of minmax values. In *Proceedings of the Fourth Workshop on Internet and Network Economics (WINE)*, pages 684–695, Shanghai, China, 2008.

[Jiang *et al.*, 2011] Albert Xin Jiang, Kevin Leyton-Brown, and Nivan A. R. Bhat. Action-graph games. *Games and Economic Behavior*, 71(1):141–173, 2011.

[Kandori, 1992] Michihiro Kandori. Social norms and community enforcement. *Review of Economic Studies*, 59:63–80, 1992.

[Kearns *et al.*, 2001] Michael Kearns, Michael Littman, and Satinder Singh. Graphical models for game theory. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 253–260, 2001.

[Kinateder, 2008] Markus Kinateder. Repeated games played in a network. Fondazione Eni Enrico Mattei Working Papers, 2008.

[Koller and Milch, 2003] Daphne Koller and Brian Milch. Multi-agent influence diagrams for representing and solving games. *Games and Economic Behavior*, 45(1):181–221, 2003.

[Kontogiannis and Spirakis, 2008] Spyros C. Kontogiannis and Paul G. Spirakis. Equilibrium points in fear of correlated threats. In *Proceedings of the Fourth Workshop on Internet and Network Economics (WINE)*, pages 210–221, Shanghai, China, 2008.

[Laclau, 2012] Marie Laclau. A folk theorem for repeated games played on a network. *Games and Economic Behavior*, 76:711–737, 2012.

[Littman and Stone, 2005] Michael L. Littman and Peter Stone. A polynomial-time Nash equilibrium algorithm for repeated games. *Decision Support Systems*, 39:55–66, 2005.

[Mihm *et al.*, 2010] Maximilian Mihm, Russell Toth, and Corey Lang. What goes around comes around: Network structure and network enforcement, 2010. Working Paper.

[Nava and Piccione, 2014] Francesco Nava and Michele Piccione. Efficiency in repeated games with local interaction and uncertain local monitoring. *Theoretical Economics*, 9:279–312, 2014.

[Takahashi, 2010] Satoru Takahashi. Community enforcement when players observe partners' past play. *Journal of Economic Theory*, 145:42–62, 2010.

[Wolitzky, 2013] Alexander Wolitzky. Cooperation with network monitoring. *Review of Economic Studies*, 80:395–427, 2013.