# What's a Little Collusion Between Friends?

Edmund L. Wong and Lorenzo Alvisi
Department of Computer Science
The University of Texas at Austin
Austin, TX, USA
{elwong,lorenzo}@cs.utexas.edu

## ABSTRACT

This paper proposes a fundamentally different approach to addressing the challenge posed by colluding nodes to the sustainability of cooperative services. Departing from previous work that tries to address the threat by disincentivizing collusion or by modeling colluding nodes as faulty, this paper describes two new notions of equilibrium, $k$-indistinguishability and $k$-stability, that allow coalitions to leverage their associations without harming the stability of the service.

## Categories and Subject Descriptors

C.2.4 [**Computer-Communication Networks**]: Distributed Systems; H.1.1 [**Models and Principles**]: Systems and Information Theory—*General systems theory*; K.6.0 [**Management of Computing and Information Systems**]: General—*Economics*

## Keywords

Game theory, collusion, cooperative services, P2P

## 1. INTRODUCTION

This paper proposes a new approach to address the challenge posed by collusion to the sustainability of peer-to-peer (P2P) cooperative services. These services rely on resources offered by their participants to implement popular applications, including content distribution (e.g., [1]), file backup (e.g., [5]), and BGP routing [38]. When resources are not under the control of a single administrative domain, the necessary cooperation cannot simply be achieved by diktat. Instead, the service must be structured so that participants have an incentive to help sustain it. Practitioners (e.g., [1]) and researchers (e.g., [2, 18]) alike have recognized that game theory can provide a rigorous basis for designing and analyzing the incentive mechanisms behind cooperative services (e.g., [5, 17, 20, 24, 25, 26, 28, 32, 35, 36, 42]). These mechanisms typically aim to ensure that a service's protocol is a best response for every individual so that no individual can profitably deviate from the service, making such a protocol, or strategy, an equilibrium strategy.

Preventing *individual* deviations, however, is unlikely to be sufficient to build robust cooperative services. The social nature of these services suggests that participants will develop, or may have already established, a rich web of relationships (based, for instance, on friendship or on belonging to the same organization), which may cause *coalitions* of participants to collude and deviate together [29]. Participants may even be able to fabricate colluders by launching Sybil attacks [13]. We submit that cooperative services that ignore the possibility of collusion do so at their own peril. That most cooperative services still choose to do so is a testament to how hard it is to address the threat posed by collusion to the stability of an equilibrium.

The literature offers two approaches to address this threat. The first is to model collusion as a fault and colluding participants as Byzantine [5, 15, 32]. The limitations of this approach are obvious: since basic distributed computing primitives such as consensus and reliable broadcast cannot be implemented if more than one third of the participants are Byzantine [23], modeling colluding participants as Byzantine imposes a cap on the number and size of coalitions that is both artificial (since it lacks a game theoretic basis) and dangerously low.

The second approach is to deny any benefit to colluders. If the equilibrium is a best response not just to every individual, but also to every possible coalition, then collusion poses no harm to the equilibrium's stability, since participants gain no benefit by colluding. This is the aim of solution concepts such as strong Nash [7] and $k$-resilient equilibria [3, 4], which offer this guarantee, respectively, for all conceivable coalitions and for arbitrary coalitions of size at most $k$. Coalition-proof Nash equilibria [8] similarly ensure that participants cannot gain any benefit from colluding and deviating in a self-enforcing way (such that there cannot be further profitable deviations from sub-coalitions).

Our work is motivated by what we believe to be a critical flaw of the second approach: its inability to account for the role played by social factors that are impossible to completely capture a priori (such as friendships or shared participation in social groups) in determining whether a participant will consider a strategy to be a best response. Intuitively, participants in coalitions formed on the basis of social "side channels" are likely to know more about each other, trust each other more, and in general be able to hold stronger assumptions about one another than about non-coalition members. Since stronger assumptions typically lead to more


*PODC'13,* July 22–24, 2013, Montréal, Québec, Canada.

efficient protocols, techniques that aim to deny benefits to coalitions face a fundamentally uphill battle: as we show in Section 3, identifying a single strategy that is a best response both inside and outside every possible coalition is very hard.

To overcome this impasse, this paper introduces and begins to explore a fundamentally different approach to dealing with coalitions. The key observation is that the fundamental property provided by an equilibrium is *stability*—in that participants do not want to deviate—and that while finding a single best response between all participants is sufficient to achieve stability, it is not *necessary*: insisting on this requirement as the means to providing stability puts the cart (i.e., best responding) before the horse (i.e., stability). As a first concrete step in this new direction, we introduce two new solution concepts that do not require fighting the strong headwinds of social relationships to guarantee stable cooperative services; instead, they explicitly model the advantages that coalition members have while ensuring that participants do not want to deviate from the specified equilibrium. Both solution concepts achieve stability through a simple observation: coalitions (including the trivial singleton coalition of one non-colluding participant) will not deviate from an equilibrium as long as the equilibrium specifies a best-response strategy for every *coalition*. Thus, the strategy a participant follows depends on whom the participant is colluding with, allowing the equilibrium to specify how participants can benefit from their coalitions.

The first solution concept, *k-indistinguishability*, achieves stability through a guarantee that, while stronger than necessary, is attractively simple. In a $k$-indistinguishable equilibrium, the actions performed by a participant within its coalition may depend on who belongs to the coalition, but the actions towards those with whom that participant is not colluding are unaffected. Thus, in a $k$-indistinguishable equilibrium, participants cannot tell whether another participant, with whom they are not colluding, is itself part of some other coalition (of at most $k$ participants). The second solution concept, *k-stability*, instead adheres to the conditions necessary for stability: like $k$-indistinguishability, $k$-stable equilibria specify a strategy per coalition that is a best response to the strategies played by all other possible coalitions; unlike $k$-indistinguishability, the actions that a participant takes as a part of a $k$-stable equilibrium may be informative about whether it is colluding and with whom. Finally, because $k$-stability and $k$-indistinguishability allow participants to change their strategies depending on whom they are colluding with, strategy profiles—traditionally used by equilibria to specify a single best-response strategy per participant—cannot capture the range of strategies that a participant may play. Instead, we use *strategy functions*, a new construct that lets us express a participant's strategy as a function of the coalition the participant belongs to.

In summary, our contributions are as follows:

- We illustrate the limits of generalizing Nash equilibria that prevent colluding participants from receiving any benefit. Specifically, we show that requiring that a single strategy be a best response for every participant, regardless of whether it is colluding, does not admit an equilibrium in several scenarios that commonly arise in cooperative services.
- We decouple the fundamental property that defines an equilibrium—stability—from the requirement that a single strategy be a best-response. This requirement, while sufficient, is not necessary when participants may collude. We take a first step at leveraging this separation by introducing (1) a new construct, strategy functions, that allows us to describe, for each participant and each possible coalition it may be part of, the strategies the participant will play, and (2) two new solution concepts, $k$-indistinguishability and $k$-stability, that admit a strategy function as an equilibrium if no coalition wants to deviate from its specified strategy.
- We demonstrate the applicability and utility of specifying a strategy per coalition by showing how our solution concepts admit useful equilibria in the same scenarios where traditional solution concepts could not.

We proceed as follows. We describe our model and setup in Section 2. In Section 3, we demonstrate the limits of generalizing traditional equilibria in the context of several common scenarios encountered in many cooperative services. In Section 4, we define two new solution concepts—$k$-indistinguishability and $k$-stability—and demonstrate how they overcome challenges faced by traditional approaches. Finally, we discuss related work and conclude in Sections 5 and 6.

## 2. SETUP

We model cooperative services as a game played by a set of $n$ players $N = \{1, \ldots, n\}$ that represent the nodes participating in the service. Each node $x$ follows some protocol or *strategy* $\sigma_x$, which specifies the actions $x$ takes at any point in the game. A *strategy profile* $\sigma$ assigns a single strategy $\sigma_x$ per node $x \in N$. A *utility function* $U$ defines every node's preferences by mapping a strategy profile $\sigma$ to a per-node *payoff*. Rational nodes prefer and select strategies that increase their payoffs as specified by the utility function, which we instantiate when discussing specific games in subsequent sections. We denote "everyone but $x$" as $-x$; indicate the combination of multiple strategies into a strategy profile using parentheses, e.g., $\sigma = (\sigma_x, \sigma_{-x})$; and drop parentheses when the meaning is obvious, e.g., $U_x(\sigma'_x, \sigma_{-x})$ denotes $x$'s payoff when $x$ plays $\sigma'_x$ and everyone else plays $\sigma_{-x}$. We use the same notation for sets of nodes as well, e.g., for some set of nodes $K$, $-K$ represents "everyone but nodes in $K$."

Our goal is to find an *equilibrium*, which typically consist of a set of strategies in which no node deviates from its assigned strategy. For example, the celebrated Nash equilibrium achieves this stability by ensuring that the strategy $\sigma^*_x$ of any given node $x$ is a *best response* (i.e., it maximizes $x$'s payoff) to everyone else following $\sigma^*_{-x}$. Thus, no node has any incentive to unilaterally deviate.

Definition 1. *A strategy profile $\sigma^*$ is a Nash equilibrium if for all $x \in N$, there does not exist some strategy $\sigma'_x$ such that*

$$U_x(\sigma'_x, \sigma^*_{-x}) > U_x(\sigma^*)$$

A *solution concept* defines a set of conditions (e.g., Definitions 1, 2, 3, 9, and 10) that describe when a set of strategies is considered an equilibrium.

## 3. DISINCENTIVIZING COALITIONS

Solution concepts such as strong Nash equilibria and $k$-resilience specify, for each node, a single best response in

which a node's actions towards a peer do not depend on whether the two are colluding. However, if coalition members trust each other more than other nodes, the practical applicability of these solution concepts are fundamentally limited. To illustrate this point, we describe techniques and scenarios likely to occur in cooperative services where the stronger assumptions that insiders can rely on when dealing with one another hamper the ability to achieve $k$-resilience. These examples are by no means comprehensive; rather, our goal is to provide a taste of the larger challenges faced by solution concepts that aim to discourage coalition formation.

Before we proceed, we first formally define $k$-resilience [3, 4], which generalizes the Nash equilibrium by requiring that the strategy profile be a best response (i.e., admit no profitable deviations) not only for every individual node (as required by a Nash equilibrium) but also for any coalition of up to size $k$. As a Nash equilibrium is simply a 1-resilient equilibrium, we generally focus on $k$-resilient equilibria where $k \geq 2$. Note that a strong Nash equilibrium is a $n$-resilient equilibrium.

DEFINITION 2. *A strategy profile $\sigma^*$ is a $k$-resilient equilibrium if, for all $K \subseteq N$ such that $|K| \leq k$, there does not exist some strategy $\sigma'_K$ such that for all $x \in K$,*

$$U_x(\sigma'_K, \sigma^*_{-K}) > U_x(\sigma^*)$$

We use this version of $k$-resilience to prove our negative results; our results therefore apply to stronger notions of $k$-resilience that guarantee stability even if coalitions are willing to deviate for less [3, 4]. Our negative results also do not rely on coalition members being able to "cheap talk", i.e., communicate at no cost, during the game [12, 16].

When there is randomness in the game, a node's best response and expected payoff depend on its *beliefs*, which represent the likelihood, from this node's viewpoint, of said random events occurring. Given every node's beliefs, we can define a Bayesian notion of $k$-resilient and strong Nash equilibrium similar to a Bayes (Nash) equilibrium.

DEFINITION 3. *A strategy profile and set of beliefs $(\sigma^*, \mu^*)$ is a $k$-resilient Bayes equilibrium if for all $K \subseteq N$ such that $|K| \leq k$, there does not exist some strategy $\sigma'_K$ such that for all $x \in K$,*

$$\mathrm{E}^{(\sigma'_K, \sigma^*_{-K}), \mu^*}[U_x] > \mathrm{E}^{\sigma^*, \mu^*}[U_x]$$

*where $\mathrm{E}^{\sigma,\mu}[U_x]$ represents $x$'s expected payoff from the strategy profile $\sigma$ with belief $\mu_x$, given that $x \in K$.*

It is important to note that all the solution concepts and equilibria we discuss in this paper are notions from *non-cooperative* game theory. There has also been extensive work in *cooperative* game theory (see any game theory text, e.g., [33], for a survey of related work) that explicitly studies the formation of coalitions in games where players are trying to work together. Cooperative and non-cooperative game theory significantly differ in focus: cooperative game theory focuses on interactions *within* a coalition—how and which coalitions form (players join a coalition based on the benefit the coalition offers) and how payoffs are allocated among coalition members (based on each member's value to the coalition)—whereas non-cooperative game theory focuses instead on the interactions *between* competing players (which, in our case, consist of exogenously-determined coalitions and non-colluding nodes).

## 3.1 Can trusted third parties limit equilibria?

Cooperative services often rely on a trusted third party to incentivize cooperation among nodes. This type of trust, which in some cases is indispensable (e.g., to implement fair exchange [22, 34]), is unnecessary among coalition members; indeed, perhaps surprisingly, it can actually render $k$-resilient equilibria impossible to achieve.

We illustrate this point through the following game, which models the fundamental choice that each node makes in P2P cooperative services: should I contribute my fair share?

DEFINITION 4. *The mediated pairwise-exchange game is a $R$-repeated game where, in each round $r \in \{1, \ldots, R\}$, each node $x \in N$:*

1. *Decides (simultaneously) on some set of peers $M^r_x \subseteq N \setminus \{x\}$ to use a mediator with.*
2. *Observes which peers are using a mediator with $x$.*
3. *Decides on some set of peers $\Gamma^r_x \subseteq N \setminus \{x\}$ to contribute to; any other peer is snubbed.*
4. *Receives a contribution from a peer $y$ if $y$ contributed to $x$ and either (a) $y$ did not use a mediator with $x$, or (b) $x$ contributed to $y$. Denote the set of all such $y$ as $C^r_x$, i.e., $y \in C^r_x$ iff $x \in \Gamma^r_y \wedge (x \notin M^r_y \vee y \in \Gamma^r_x)$.*

*$x$ pays $\gamma$ per peer that $x$ contributes to and $\epsilon$ per peer that $x$ uses a mediator with. $x$ earns $b > 2\gamma + \epsilon$ per received contribution, for a round payoff of $v^r_x = |C^r_x|b - |\Gamma^r_x|\gamma - |M^r_x|\epsilon$. A node's total payoff is the sum of all round payoffs: $\sum_{r=1}^{R} v^r_x$.*

While this game resembles a finitely-repeated prisoner's dilemma, the mediator, who can serve as a trusted third party and ensure a fair pairwise exchange, enables the existence of Nash equilibria in which contribution occurs (without the mediator, no such equilibrium exists).

THEOREM 1. *Let $\sigma^*$ be a strategy profile in the mediated pairwise-exchange game in which a node $x$, following $\sigma^*_x$, contributes to a peer $y$, using a mediator only in round $R$, iff (1) $x$ and $y$ have never snubbed each other in the past and (2) $x$ and $y$ have not used a mediator in any round other than $R$; otherwise, $x$ snubs $y$ without a mediator. Then $\sigma^*$ is a Nash equilibrium.*

PROOF. Same as the backwards-induction half of the proof of Theorem 5. □

The Nash equilibrium in Theorem 1 uses the mediator to ensure cooperation in the last round, which encourages cooperation in prior rounds without the mediator. We now prove that this same mediator precludes the existence of $k$-resilient equilibria. The reason, essentially, is that using the mediator, which incurs cost, is undesirable between colluding nodes (Lemma 2) but necessary to ensure cooperation between two non-colluding nodes (Lemma 1). This tension makes it impossible for a single strategy to be a node's best response regardless of how it colludes (Theorem 2).

LEMMA 1. *In any $k$-resilient equilibrium of the mediated pairwise-exchange game where some node contributes, the last time in the game that any node contributes with positive probability to a peer must always involve a mediator.*

PROOF. By contradiction. Fix some $k$-resilient equilibrium $\sigma^*$, where the last time that any node contributes with positive probability does not involve a mediator with positive probability (if there exist multiple such node/peer pairings, choose one arbitrarily). During this "last contribution," let $x$ be the node that contributes, $y$ be the receiving peer, and $\alpha$ be the probability that $x$ contributes to $y$ after deciding not to use a mediator with $y$. By assumption, $\alpha > 0$.

Since $\sigma^*$ must be a best response regardless of who is colluding, suppose $x$ and $y$ are not colluding. Then it must be the case that, in $\sigma^*$, $y$ snubs $x$ during the last contribution if $x$ does not use a mediator: $y$ expects to earn, from $x$'s contribution, $\alpha b$ without incurring the cost of contributing; moreover, since this is the last time a contribution occurs with positive probability, $y$'s choice of whether to snub $x$ does not negatively impact $y$'s *continuation* (i.e., subsequent) payoff. It follows that $x$ could profitably deviate from $\sigma^*$ by always snubbing $y$ during the last contribution if $x$ does not use a mediator: doing so would save $x$ an expected cost of $\alpha\gamma$ with no negative effect on $x$'s continuation payoff. Contradiction. □

LEMMA 2. *In any $k$-resilient equilibrium of the mediated pairwise exchange game where some node contributes, the last time in the game that any node contributes with positive probability to a peer must never involve a mediator.*

PROOF. By contradiction. Fix some $k$-resilient equilibrium $\sigma^*$ where the last time that any node contributes with positive probability also involves a mediator with positive probability (if there exist multiple such node/peer pairings, choose one arbitrarily). During this "last contribution," let $x$ be the node that contributes; $y$ be the peer; $\alpha > 0$ be the probability that $x$ decides *to use* a mediator with $y$; and $p_x$ ($p_y$) be the probability that $y$ ($x$) observes a contribution from $x$ ($y$) in expectation over all possible combinations of $x$ and $y$'s choices regarding using a mediator and contributing with one another.

Since $\sigma^*$ must be a best response regardless of who is colluding, suppose $x$ and $y$ *are* colluding. Consider an alternate strategy profile $\sigma'$ in which all nodes play the same actions with the same probabilities as in $\sigma^*$, except, during the last contribution, $x$ and $y$ do not use a mediator with one another, $x$ ($y$) contributes to $y$ ($x$) with probability $p_x$ ($p_y$), and $x$ and $y$ subsequently play actions as if $x$ and $y$ had instead followed $\sigma^*$. It follows that the payoffs for $x$ and $y$ are exactly the same, with the exception of the payoffs that $x$ and $y$ receive from one another during the last contribution, where (1) $x$ and $y$'s expected benefit remains the same, (2) $x$'s expected cost is strictly lower since $x$ contributes with the same probability in expectation without the cost of a mediator ($\alpha\epsilon > 0$), and (3) $y$'s expected cost is no higher (and is lower if $y$ was using a mediator in $\sigma^*$). Thus, $x$ is better off and $y$ is no worse off. Contradiction. □

THEOREM 2. *There exists no $k$-resilient equilibrium in the mediated pairwise-exchange game.*

PROOF. Lemmas 1 and 2 imply that there exists no $k$-resilient equilibrium where nodes contribute. Further, a strategy profile $\sigma$ in which all nodes snub and earn 0, while a Nash equilibrium, is not a $k$-resilient equilibrium. To see why, consider an alternate strategy profile $\sigma'$ and some coalition $K$ (such that $|K| \geq 2$) where no one uses mediators and only members of $K$ contribute to one another. $\sigma'$ earns $K$'s members payoffs of $(|K|-1)R(b-\gamma) > 0$ each, making it a profitable deviation from $\sigma$. □

## 3.2 What if nodes may fail?

When nodes may fail, a node's best response will generally depend on the probability with which it expects other nodes may fail. Greater trust and access to more information (e.g., concerning the frequency with which fellow coalition nodes are patched) may allow nodes within a coalition to reasonably believe that fellow coalition members have a lower probability of failing than outsiders. Unfortunately, even a slightly lower failure probability can make $k$-resilience practically unachievable.

We illustrate this point using a simple single-shot simultaneous game that models a simplified version of secret-sharing [3, 39]. In this game, each node wants to reconstruct a secret that requires the node to request shares from its peers. These peers deliver the requested shares unless they fail (e.g., by crashing). Each node must then decide how many shares to request: requesting more shares incurs more cost, but requesting fewer shares may result in the node being unable to reconstruct the secret because of peer failures.

DEFINITION 5. *The simple secret-sharing game is a single-shot, simultaneous game in which every node $x \in N$:*

1. *Selects a set $\Gamma_x \subseteq N \setminus \{x\}$ of nodes to request shares from.*
2. *Pays $|\Gamma_x|\gamma$ for this request.*
3. *Receives shares from some set $C_x \subseteq \Gamma_x$.*
4. *Earns benefit $b > |N|\gamma$ iff $|C_x| \geq m$, where $m$ is the number of shares that $x$ must gather from its peers before being able to reconstruct the secret.*

The simple secret-sharing game is a decision theory problem: a node's choice does not affect its peers' outcomes.[1] This is intentional: our goal is to show that, despite the game's simplicity, it is often impossible to find $k$-resilient equilibria. To account for a node's beliefs regarding how likely its peers are to fail, we use $k$-resilient Bayes equilibria (Definition 3). In this game, a strategy profile $\Gamma$ represents the peers that each node requests from. A set of beliefs $\mu$ represents the view of each node, given the set of peers it is colluding with, regarding the likelihood that any peer will successfully deliver its share if requested. In other words, $\mu$ represents each node $x$'s view of the likelihood that a peer in $\Gamma_x$ will also be in $C_x$. An equilibrium in the simple secret-sharing game is some $(\Gamma^*, \mu^*)$ where no node $x$, colluding with any $(k-1)$ peers, could do any better in expectation requesting shares from some set $\Gamma'_x \neq \Gamma^*_x$. More formally, for all $K \subseteq N$ such that $|K| \leq k$, there is no $\Gamma'_x$ such that for all $x \in K$,

$$H[|C'_x| - m]b - |\Gamma'_x|\gamma > H[|C^*_x| - m]b - |\Gamma^*_x|\gamma$$

where $H[i]$ is the discrete unit step function ($H[i] = 1$ if $i \geq 0$; otherwise $H[i] = 0$).

[1] If the game were sequential, the choice of some node $x$ to request a share from some peer $y$ could inform $y$ of whether $x$ has failed. However, finding $k$-resilient equilibria is no less challenging, since (1) there is at least one node (the first node to move) that will never have such a signal and (2) even if $x$ successfully requests a share from $y$, $x$ could subsequently fail before $y$'s turn.

THEOREM 3. *Let $(\Gamma^*, \mu^*)$ be a $k$-resilient Bayes equilibrium of the simple secret-sharing game in which some node $x \in N$ believes that a peer $y$ will fail with probability $\mu_x^*$ if $x$ and $y$ are not colluding and $\mu_x^* - \epsilon$ if $x$ and $y$ are, where $\epsilon > 0$. Then either $x$ requests secrets from no one or everyone, i.e., $\Gamma_x^* \in \{\emptyset, N \setminus \{x\}\}$.*

PROOF. Suppose $k = 2$ and $K = \{x, y\}$, i.e., $x$ and $y$ are colluding. If $x$ incurs more cost requesting shares than it earns in expectation from reconstructing the secret (e.g., because of high rates of failure), then $\Gamma_x^* = \emptyset$. Otherwise, suppose $x$ requests shares from peers in $\Gamma_x^* \neq \emptyset$. It is obvious that since $x$ believes that $y$ will fail with probability $\mu_x^* - \epsilon$, which is lower than the probability of any other peer $z \neq y$ failing ($\mu_x^*$), $x$ should always request shares from $y$, so any 2-resilient $\Gamma_x^*$ must contain $y$. However, as $y$ can be any peer, the only $\Gamma_x^*$ that is guaranteed to contain all possible $y$ is $\Gamma_x^* = \{y \mid y \in N \wedge y \neq x\} = N \setminus \{x\}$. Finally, as $k$-resilience implies 2-resilience, this result applies to $k$-resilience for $k \geq 2$. □

A node that wants to reconstruct the secret rarely wants to request shares from all of its peers, since the cost of these additional requests is not worth the slight insurance that redundant shares provide. However, in such cases, it follows from Theorem 3 that no $k$-resilient Bayes equilibrium exists. Therefore, the only scenarios in which a node wants to reconstruct the secret as a part of a $k$-resilient Bayes equilibrium are those in which the secret's value is sufficiently high to justify requesting shares from all peers to maximize the likelihood of success.

Figure 1 quantifies what this value must be, using example numbers based on a movie-streaming context: $n = 100$ nodes; each node expects that coalition members never fail[2] and that non-coalition members fail with independent probability $\beta$; $m$ is set such that, given an independent failure probability of $\beta$, there is at least a 0.99999 chance that at least $m$ peers, out of $n - 1$ possible peers, will not fail; and $\gamma$ is set to (1500 Kbps) × (2 hours) × ($1/GB)/$(m + 1)$. As $k$ increases, the expected probability of a coalition member reconstructing the secret increases, thus making it more difficult to convince such a node to request shares from every other peer. Note that while Figure 1 implies that the minimum required benefit goes up as probability of failure goes up, this is an artifact of how we define $m$; in reality, the minimum required benefit goes up as the probability of failure goes down, as expected.

As Figure 1 shows, even with coalitions of at most two nodes and beliefs that non-coalition nodes fail with probability 0.01, a 2-resilient equilibrium exists only if a node values a two-hour movie, which incurs $\gamma(n - 1) > $1.37 in communication costs, at over $268.95!

## 3.3 Do nodes want to punish one another?

Cooperative services often incentivize nodes not to deviate by relying on the threat of punishment. In this section, we show that punishments that hurt both the enforcing and receiving nodes are never used within a coalition, and other forms of punishment will be difficult, if not impossible, to achieve in real-world scenarios. We illustrate this through a simplified version of the mediated pairwise-exchange game.[3]

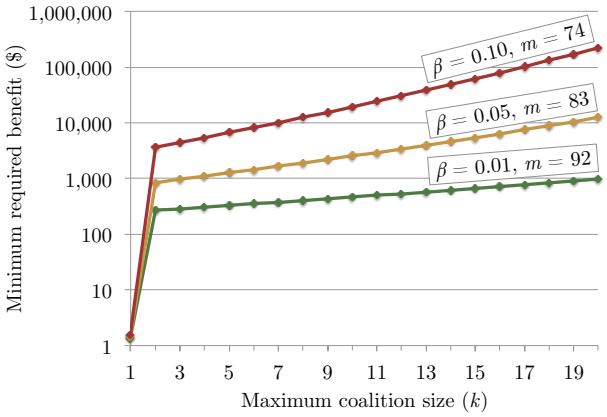


**Figure 1:** In the simple secret-sharing game (Definition 5), the minimum benefit needed for a $k$-resilient equilibrium where nodes attempt to reconstruct the secret.

DEFINITION 6. *The simple pairwise-exchange game is an infinitely-repeated game where, in each round $r \geq 0$, every node $x \in N$:*

1. *Simultaneously decides on some set of nodes $\Gamma_x^r \subseteq N \setminus \{x\}$ that it will contribute to; any node not in $\Gamma_x^r$ is snubbed.*
2. *Observes which peer $y \neq x$ contributed to it; let $C_x^r$ denote the set of all such $y$, i.e., $y \in C_x^r$ iff $x \in \Gamma_y^r$.*

*$x$'s round payoff is $v_x^r = |C_x^r|b - |\Gamma_x^r|\gamma$, where $b > \gamma$. $x$'s total payoff is the $\delta$-weighted sum of the round payoffs: $\sum_{r=0}^{\infty} \delta^r v_x^r$.*

THEOREM 4. *Let $\sigma^*$ be a $k$-resilient equilibrium in the simple pairwise-exchange game in which some contribution occurs. In other words, $\sigma^*$ specifies that, at some point in the game, a node $y$ contributes to some node $x$, who "rewards" $y$ if $y$ contributes and "punishes" $y$ if $y$ snubs $x$. Then either (1) $x$ must prefer punishing to rewarding $y$, and/or (2) $x$ punishing $y$ is not a $k$-resilient best response (i.e., $x$ may threaten to punish $y$, but, given the opportunity, $x$ and $y$ can profitably deviate by not following through).*

PROOF. Fix some $k$-resilient equilibrium $\sigma^*$ in which contribution occurs and, unlike condition (1) above, $x$ is no worse off rewarding $y$. We prove that condition (2) follows: $x$ punishing $y$ is not a $k$-resilient best response. Let $r$ be the round in which this contribution occurs, $U_x(\sigma^* | (\Gamma_x^r, C_x^r \cup \{y\}))$ denote $x$'s continuation payoff from rewarding $y$, and $U_x(\sigma^* | (\Gamma_x^r, C_x^r \setminus \{y\}))$ denote $x$'s continuation payoff for punishing $y$. We have:

$$U_x(\sigma^* | (\Gamma_x^r, C_x^r \cup \{y\})) \geq U_x(\sigma^* | (\Gamma_x^r, C_x^r \setminus \{y\})) \quad (1)$$

Denote $y$'s continuation payoff from contributing to and snubbing $x$ as $U_y(\sigma^* | (\Gamma_y^r \cup \{x\}, C_y^r))$ and $U_y(\sigma^* | (\Gamma_y^r \setminus \{x\}, C_y^r))$,

[2]While this may seem extreme, note that this is exactly what failure-aware $k$-resilient solution concepts, such as $(k, t)$-robustness [3, 4], require: nodes do not deviate assuming that the coalition and set of faulty nodes do not overlap.

[3]While we could use the mediated pairwise-exchange game to illustrate this point, we instead use a game with an infinite horizon (which enables the existence of Nash equilibria where contribution occurs) and no mediator as the mediator already makes $k$-resilient equilibria impossible to achieve.

respectively. As $y$ contributes to $x$ as a part of a $k$-resilient equilibrium, $y$ must be no worse off doing so:

$$|C_y^r|b - |\Gamma_y^r \cup \{x\}|\gamma + U_y(\sigma^*|(\Gamma_y^r \cup \{x\}, C_y^r)) \geq$$
$$|C_y^r|b - |\Gamma_y^r \setminus \{x\}|\gamma + U_y(\sigma^*|(\Gamma_y^r \setminus \{x\}, C_y^r))$$

Unsurprisingly, it follows that $y$, in continuation, is worse off being punished than being rewarded:

$$\begin{aligned} U_y(\sigma^*|(\Gamma_y^r \cup \{x\}, C_y^r)) &\geq \gamma + U_y(\sigma^*|(\Gamma_y^r \setminus \{x\}, C_y^r)) \\ &> U_y(\sigma^*|(\Gamma_y^r \setminus \{x\}, C_y^r)) \end{aligned} \quad (2)$$

Suppose $K = \{x, y\}$, and let $\sigma'_K$ specify the same actions as in $\sigma^*$, except $x$ and $y$ play $\sigma^*$ as if $y$ contributed even if $y$ snubbed $x$. We can see that by inequality (1),

$$\begin{aligned} U_x((\sigma'_K, \sigma^*_{-K})|(\Gamma_x^r, C_x^r \setminus \{y\})) &= U_x(\sigma^*|(\Gamma_x^r, C_x^r \cup \{y\})) \\ &\geq U_x(\sigma^*|(\Gamma_x^r, C_x^r \setminus \{y\})) \end{aligned}$$

and, by inequality (2),

$$\begin{aligned} U_y((\sigma'_K, \sigma^*_{-K})|(\Gamma_y^r \setminus \{x\}, C_y^r)) &= U_y(\sigma^*|(\Gamma_y^r \cup \{x\}, C_y^r)) \\ &> U_y(\sigma^*|(\Gamma_y^r \setminus \{x\}, C_y^r)) \end{aligned}$$

Thus, $x$ punishing $y$ is not a $k$-resilient best response. □

Theorem 4 applies to many forms of punishment, including various flavors of grim trigger, forgiving trigger, and tit-for-tat (if $b/\gamma > 1/\delta$). A $k$-resilient equilibrium can still use these punishments as a non-credible threat and hope that such bluffs are not called in practice. Alternatively, any punishment in which nodes strictly prefer to punish than reward can be part of a $k$-resilient equilibrium. However, if network loss is a possibility (as in real-world environments), (1) behaviors that are not part of a $k$-resilient equilibrium (e.g., snubbing when only contribution is supposed to be played) may be observed, resulting in a node (rationally) reneging on its non-credible threat and the collapse of any $k$-resilient equilibrium that encourages contribution using such threats; and (2) the inability to observe what a node has observed (as in [42]) may result in nodes feigning being snubbed and frivolously punishing their peers under false pretenses.[4]

## 3.4 What other issues are there?

Finally, we briefly describe two commonly-used techniques that are often not $k$-resilient. *Digital signatures*, with their guarantee of non-repudiation, are useful in adversarial environments, but their bandwidth and computational costs are hard to justify within a coalition where members trust each other. Generally, digitally signing messages is part of a $k$-resilient protocol only if not signing may affect the protocol's outcome, e.g., if this message is passed around to more than $k$ nodes that check the signature, and coalition members cannot sign for each other.

*Junk*, i.e., semantically meaningless data, has been used (e.g., [5, 28, 44]) as a form of payment to ensure that nodes contribute their fair share to the cooperative service. For instance, if a node is required to send data but has nothing useful to send, it may instead send protocol-specified "junk." By making junk more expensive to transfer than useful content, junk transfers discourage free-riding by incentivizing nodes to send real content whenever possible. However, junk transfers incur bandwidth costs on the sender and receiver while providing no benefit to the receiver; nodes that trust each other have no incentive to perform them. It follows that no protocol that relies on junk transfers is $k$-resilient.

[4]We omit further discussion due to lack of space; see [41].

# 4. ACCEPTING COALITIONS

The scenarios in Section 3 suggest that it is difficult for a single strategy profile to specify strategies that a node, colluding with up to $(k-1)$ peers, prefers over all possible deviations, as required by $k$-resilience. Yet, we believe these scenarios are symptomatic of a more general problem: the ability for colluding nodes to hold stronger beliefs and assumptions about fellow coalition members (and potentially about the system as a whole) often results in more efficient protocols. As a result, we believe there are likely very few scenarios in which $k$-resilience will bear fruit.

In this section, we show that the insight to overcome this impasse is to recognize that denying benefits to nodes that belong to a coalition, while sufficient for stability, is not necessary. We propose a fundamentally different notion of equilibrium: instead of specifying a single best-response strategy to each node, our equilibria map each node to possibly multiple strategies, depending on whom it colludes with. By effectively mapping each possible coalition to a strategy, our equilibria can specify, as a part of the strategy, the efficiencies that a coalition can leverage among its members. Despite this flexibility, our equilibria guarantee that the strategies specified for every coalition is a best response to what other nodes play, despite how they collude.

**Specifying coalitional strategies.** Because our equilibria specify a strategy per coalition, the strategies that the nodes, within each coalition, follow may depend on whom they are colluding with. Our equilibria cannot use strategy profiles used by traditional equilibria because they specify only a single strategy per node. Our equilibria instead use a novel construct, a *strategy function*, to specify a node's strategy based on whom the node is colluding with. We formally represent how nodes collude by a partition $P$ of $N$, in which two nodes $x$ and $y$ are colluding if there exists some element (a coalition) $K \in P$ such that $x, y \in K$. Intuitively, each partition represents one way that nodes can collude. We use $\mathbb{P}^k = \{P : \forall K \in P, |K| \leq k\}$ to denote the space of all partitions that contains no coalition larger than size $k$.

DEFINITION 7. *A strategy function $\mathcal{S}$ is a mapping from a partition (representing a particular way that nodes have chosen to collude) to a strategy profile (which specifies the strategies that these nodes will play as a result) such that if there exists some coalition $K$ that is in $P$ and $P'$, $\mathcal{S}$ maps the same strategy to $K$ in $P$ and $P'$, i.e., if $K \in P$ and $K \in P'$, $\mathcal{S}_K(P) = \mathcal{S}_K(P')$, where $\mathcal{S}_K(P)$ and $\mathcal{S}_K(P')$ denote the strategies deployed by $K$ given partitions $P$ and $P'$.*

Note that a node's strategy does not depend on how nodes outside of its coalition collude, which a node may not know. We define $\mathcal{M}$ as the membership function: $\mathcal{M}(x, P) = K$ if, in partition $P$, $K$ is the coalition that $x$ is a part of, i.e., $K \in P$ and $x \in K$. With respect to a node $x$ in coalition $K$, all nodes in $K$ are *insiders*, and all others are *outsiders*.

## 4.1 Coalition-indistinguishable equilibria

Where $k$-resilience makes coalitions futile, $k$-indistinguishability makes them invisible; where $k$-resilience fundamentally aims to deny coalitions any claim of exceptionalism

and sees a system as a collection of individual nodes, $k$-indistinguishability sees a system as a collection of coalitions, some of which may contain a single node; where $k$-resilience ensures that every node best responds to every other node, $k$-indistinguishability ensures that every coalition best responds to every other coalition: in both equilibria, nodes that belong to different coalitions interact with each other as if no coalition existed.

DEFINITION 8. *Two strategy profiles $\sigma$ and $\sigma'$ are indistinguishable with respect to some node $x$, denoted as $\sigma \overset{x}{=} \sigma'$, if all histories resulting from $\sigma$ and $\sigma'$, as observed by $x$, occur with equal probability and $U_x(\sigma) = U_x(\sigma')$.*

DEFINITION 9. *$\mathcal{S}^*$ is a $k$-indistinguishable equilibrium if:*

- *For any $P, P' \in \mathbb{P}^k$, any coalition $K$ such that $K \in P$ and $K \in P'$, and any $x \in K$, $\mathcal{S}^*(P) \overset{x}{=} \mathcal{S}^*(P')$.*
- *For all $P \in \mathbb{P}^k$ and all $K \in P$, there does not exist a strategy $\sigma'_K$ such that for all $x \in K$,*

  $$U_x(\sigma'_K, \mathcal{S}^*_{-K}(P)) \geq U_x(\mathcal{S}^*(P))$$

  *and, for some $y \in K$, the inequality is strict.*

The first condition (indistinguishability) requires that a node cannot distinguish whether an outsider is itself colluding with others; the second condition (best response) requires that in any partition, there exists some node in every coalition that prefers the equilibrium-specified strategy to any coalitional deviation. Note that while we defined best response to be consistent with the definition of $k$-resilience, weaker or stronger notions could have been used instead. Also, observe that the best-response condition of $k$-indistinguishable equilibria must hold for all possible partitions. Therefore, like $k$-resilient equilibria, a $k$-indistinguishable equilibrium consists of strategies that make up a best response for *all* possible coalitions of up to size $k$, not just one particular coalition or set of coalitions.

Every $k$-resilient and Nash equilibrium $\sigma^*$ has an equivalent $k$-indistinguishable equilibrium $\mathcal{S}^*$ in which $\mathcal{S}^*(P) = \sigma^*$ for all $P$. However, by allowing nodes to base their strategies on whom they collude with, $k$-indistinguishable equilibria circumvent the challenges described in Section 3 while ensuring that no coalition will deviate from its specified strategy (Section 4.3). Moreover, similar to $k$-resilience, any service that uses a protocol which is the non-colluding strategy in a $k$-indistinguishable equilibrium is guaranteed to be supported and maintained by nodes, even if they may collude. Although $k$-indistinguishability cannot guarantee that the exact protocol will be followed to the letter by a node when interacting with a fellow insider, $k$-indistinguishability does guarantee that any actions that a node takes when interacting with an outsider is the same as those specified by the service's protocol. Thus, from the service's perspective, every node is effectively running the service's protocol and supporting the service.

## 4.2 From indistinguishability to stability

Although indistinguishability is an attractive guarantee, it may in practice prove too stringent for some applications. For example, a content-distribution service in which colluding nodes freely exchange content with one another may not be $k$-indistinguishable because non-colluding nodes may be able to detect the presence of a coalition simply by observing that colluding nodes statistically have more content at any given time than everyone else. $k$-stable equilibria do away with indistinguishability, focusing only on the conditions necessary for stability.

DEFINITION 10. *$\mathcal{S}^*$ is a $k$-stable equilibrium if for all $P \in \mathbb{P}^k$ and all $K \in P$, there does not exist a strategy $\sigma'_K$ such that for all $x \in K$,*

$$U_x(\sigma'_K, \mathcal{S}^*_{-K}(P)) \geq U_x(\mathcal{S}^*(P))$$

*and, for some $y \in K$, the inequality is strict.*

As in $k$-indistinguishable equilibria, a $k$-stable equilibrium requires a best response for all possible coalitions of up to size $k$, and every $k$-resilient and Nash equilibrium has a $k$-stable equivalent. Moreover, every $k$-indistinguishable equilibrium is also $k$-stable. However, $k$-stable equilibria do not guarantee that a colluding node's strategy is indistinguishable from that of a non-colluding node. In other words, it is possible that the strategy of a colluding node $x$ provides outsiders with information about whether $x$ is colluding, with whom $x$ is colluding, etc. In addition, if $x$ chooses to collude, $x$'s coalition may affect the payoffs of peers both inside and outside of $x$'s coalition. Nevertheless, a $k$-stable equilibrium still guarantees that, for any coalition, the specified strategy is a best response to the strategies played by all outsiders, regardless of how these other nodes may collude.

**Other $k$-stable solution concepts.** $k$-stability is a very general notion that, we believe, provides a useful basis for developing new solution concepts that guarantee stability in the presence of collusion. $k$-indistinguishability is one such solution concept, the result of adding indistinguishability to $k$-stability. Another requirement that one may desire is some notion of self-enforcement (no profitable deviation by sub-coalitions), e.g., a solution concept could require that, in equilibrium, nodes prefer to be with their respective coalitions over working alone ($k$-stability and $k$-indistinguishability do not have any such requirement). Alternatively, one could devise a Bayesian version of $k$-stability that guarantees an expected best response for each coalition based on the likelihood that certain coalitions will form. Yet another interesting direction would be to devise a version of $k$-stability that bounds the "price of collusion," i.e., how much a node's payoff is affected when outsiders choose to collude (similar to the notion of a safety-net guarantee used in [44]). We leave exploring these and other notions of equilibrium to future work.

## 4.3 Examples of equilibria

In this section, we show the applicability of $k$-stability and $k$-indistinguishability by showing that such equilibria exist in the scenarios described in Section 3, where $k$-resilient equilibria did not exist before.

**$k$-stability and $k$-indistinguishability in the mediated pairwise-exchange game.** It is simple to prove that there exists a $k$-indistinguishable equilibrium in the mediated pairwise-exchange game (Definition 4). Because $k$-indistinguishable and $k$-stable equilibria allow nodes to base their play on whom they are colluding with, a node, as a part of a $k$-indistinguishable equilibrium, can use the medi-

ator with outsiders (as in Theorem 1) and leverage the trust provided by the coalition with insiders.

THEOREM 5. *Let $\mathcal{S}^*$ be a strategy function such that, for any partition $P \in \mathbb{P}^k$ and any $x \in N$, $\mathcal{S}^*_x(P)$ specifies that*

- *For $y \in \mathcal{M}(x, P)$ such that $y \neq x$, $x$ never uses a mediator and always contributes.*
- *For $y \notin \mathcal{M}(x, P)$, $x$ contributes to $y$, using a mediator only in round $R$, iff (1) $x$ and $y$ have never snubbed each other in the past and (2) $x$ and $y$ have not used a mediator in any round other than $R$. Otherwise, $x$ snubs $y$ without a mediator.*

*Then $\mathcal{S}^*$ is a $k$-indistinguishable equilibrium.*

PROOF. Without loss of generality, fix some partition $P$,[5] and consider the interactions of some node $x$ with some peer $y$.[6] Suppose that $y$ is an insider, i.e., $y \in \mathcal{M}(x, P) = K$. Let $R_s$ be the set of rounds in which $x$ snubs $y$ and $R_m$ be the set of rounds in which $x$ uses a mediator with $y$. In each round in $R_s$, $x$ gains $\gamma$, but $y$ loses $b$. In each round in $R_m$, $x$ loses $\epsilon$; $y$'s payoff is unaffected. Any deviation in which $R_s \neq \emptyset$ or $R_m \neq \emptyset$ is then not in $K$'s best interest.

Suppose instead that $y$ is an outsider, i.e., $y \notin \mathcal{M}(x, P)$. We can show that by following $\mathcal{S}^*_x(P)$ with respect to $y$ is $x$'s best response by backwards induction.

*Base case*: round $R$ (the last round). We first show that $\mathcal{S}^*_x(P)$ is a best response for $x$ with respect to $y$ by considering the following two cases:

- $x$ and $y$ have always contributed to one another. If $x$ deviates by snubbing and/or not using a mediator, $x$ saves at most $\gamma + \epsilon$. However, since $y$ is using a mediator, $x$ loses benefit $b$ it would have received from $y$ otherwise. Since $b > 2\gamma + \epsilon > \gamma + \epsilon$ by assumption (Definition 4), $x$ is clearly worse off.
- $x$ and/or $y$ have snubbed one another in the past. If $x$ deviates by contributing to $y$ or using a mediator, $x$ is obviously worse off: $x$ must pay at least $\min(\gamma, \epsilon) > 0$ but receives no additional benefit.

*Inductive step.* Assume that for all rounds following some round $r_0 > 1$, $\mathcal{S}^*_x(P)$ is a best response for $x$ with respect to $y$. We now prove the inductive step—$\mathcal{S}^*_x(P)$ is a best response for $x$ with respect to $y$ in round $r_0$—in a similar fashion by considering the following two cases:

- $y$ has always contributed to $x$. If $x$ deviates by using a mediator, $x$ is at least $\epsilon$ worse off in round $r_0$. If $x$ deviates by snubbing $y$, $x$ saves $\gamma$ in round $r_0$. Regardless, $y$ will snub $x$ in every subsequent round, resulting in $x$ losing at least $b - (\gamma + \epsilon)$ per round. $x$ is then worse off since the net change in $x$'s payoff is at least $\gamma - (b - (\gamma + \epsilon)) = -b + 2\gamma + \epsilon < 0$.
- $y$ has snubbed $x$. If $x$ deviates by contributing or using a mediator, $x$ is worse off, as argued in the base case.

Thus, $\mathcal{S}^*_x(P)$ is a best response for $x$. □

The mediated pairwise-exchange game, as defined in Definition 4, involves every node $x$ *privately* observing which peers use a mediator with or contribute to $x$; $x$ does not know what other peers have chosen with respect to one another. If such choices were publicly observable (e.g., if the mediator published a list describing which pairs of nodes it would mediate for), $\mathcal{S}^*$ would no longer be a $k$-indistinguishable equilibrium, since non-colluding nodes would be able to observe that coalition members never use a mediator with one another. However, because nodes, regardless of whom they collude with, are still better off following the strategies specified in $\mathcal{S}^*$, $\mathcal{S}^*$ would remain a $k$-stable equilibrium.

**$k$-stability in the simple secret-sharing game.** Likewise, it is straightforward to show that the simple secret-sharing game (Definition 5) has a $k$-stable equilibrium. In particular, a node, depending on whom it is colluding with, can choose the exact set of peers to request secrets from that the node expects will maximize its payoff.

**$k$-stability and $k$-indistinguishability in the simple pairwise-exchange game.** We can incorporate the punishments in Section 3.3 into a protocol that is $k$-stable or $k$-indistinguishable in the simple pairwise-exchange game (Definition 6). As an example, we demonstrate how a local grim-trigger punishment can be used here.

THEOREM 6. *Let $\mathcal{S}^*$ be the following strategy function: for any partition $P \in \mathbb{P}^k$ and for any $x \in N$, $\mathcal{S}^*_x(P)$ specifies the following action for $x$:*

- *For $y \in \mathcal{M}(x, P)$ such that $y \neq x$, contribute to $y$.*
- *For $y \notin \mathcal{M}(x, P)$, contribute to $y$ iff $r = 0$ or $x$ and $y$ have always contributed to one another.*

*Then $\mathcal{S}^*$ is a subgame-perfect $k$-indistinguishable equilibrium (i.e., at every point in the game, nodes play a $k$-indistinguishable best response) if $b/\gamma \geq 1/\delta$.*

PROOF. Without loss of generality, fix $P$. For any $K \in P$ in which $|K| > 1$, $\mathcal{S}^*_K(P)$ is a best response when interacting with fellow insiders. To see why, observe that following $\mathcal{S}^*_K(P)$ in each round earns a round payoff of $(n-1)(b-\gamma)$. Deviating by snubbing an insider improves one node's payoff by $\gamma$ but causes a loss of $b > \gamma$ to another's; the coalition as a whole earns $(n-2)(b-\gamma) < (n-1)(b-\gamma)$ as a result in that round, so someone in the coalition must be worse off.

Now consider any two nodes $x$, $y$ that are not colluding, i.e., $y \notin \mathcal{M}(x, P)$. If $x$ and $y$ have always contributed to each other and $x$ snubs $y$, $x$ gains $\gamma$ in the current round but loses at least $(b-\gamma)$ in every subsequent round. This is profitable only if $\delta(b-\gamma)/(1-\delta) < \gamma$, which is never the case. Finally, if $y$ has snubbed $x$ and $x$ deviates by contributing to (rather than snubbing) $y$, $x$ incurs an additional cost of $\gamma$; this is clearly not in $x$'s best interest. □

Similar to the previous example, $\mathcal{S}^*$ as defined in Theorem 6 would remain a $k$-stable equilibrium (but would not be indistinguishable at every point in the game) if a node's choices of whom to contribute to were publicly observable.

**$k$-stability and $k$-indistinguishability with digital signatures and junk.** Mechanisms such as digital signatures or junk transfers fit naturally within a $k$-stable or $k$-indistinguishable equilibrium. The equilibrium may specify

[5] As our proof makes no assumptions about $P$, it follows that our proof holds for all possible partitions $P \in \mathbb{P}^k$.

[6] We can safely do this because each interaction between any two pairs of nodes in $\mathcal{S}^*$ is independent.

that these mechanisms are used between outsiders and bypassed between insiders when unneeded.

# 5. RELATED WORK

We have seen how hard it is to achieve useful equilibria in cooperative services using strong Nash equilibrium [7], which requires a strategy profile be Pareto optimal, and $k$-resilience [3, 4], which has weaker but similar requirements. As previously mentioned, Bernheim et al. [8] describe coalition-proof Nash equilibria, which weaken strong Nash equilibria by requiring that the equilibrium be preferable only to self-enforcing deviations, i.e., a deviation by a coalition in which no subset of this coalition can further deviate and profit. Considering only self-enforcing deviations provides little benefit when coalitions have exogenous means to ensure that coalition members deviate together, which we argue is often the case in cooperative services. For instance, a set of Sybil nodes [13] controlled by a single entity will not deviate within their coalition, and friends may avoid hurting each other because of social repercussions (which can be formally modeled using notions of binding commitments or multimarket contact [9]). Finally, there has been work in defining correlated versions of strong Nash and coalition-proof equilibria (e.g., [10, 14, 31]); like their non-correlated counterparts, these equilibria require a best response despite how nodes collude and thus have similar shortcomings.

Along with [3, 4], there has been much work in providing incentives when some nodes may arbitrarily fail (e.g., [5, 15, 27, 28, 43]). Although collusion can be modeled as an arbitrary failure, these approaches typically only handle a bounded number of failures and thus a limited amount of collusion, which is further restricted if failure can occur. Moreover, failure and collusion are fundamentally separate concerns, and $k$-indistinguishability and $k$-stability can be augmented to require a best response despite failure.

In the context of mechanism design and auctions, Chen et al. [11] describe rationally-robust implementation, an interesting non-equilibrium-based solution concept that primarily aims to ensure that even if every individual or coalition is given no initial hint of what to play, the underlying mechanism induces individuals or coalitions to choose strategies that ultimately preserve some desired system property. As a result, players may play multiple strategies, as in our equilibria; unlike equilibrium-based approaches, rationally-robust implementation does not predict the exact strategies that will be used, which ultimately may be any undominated strategy. It is unclear whether rationally-robust implementation's notion of dominance can remove enough strategies in games based on cooperative services to enable the existence of useful properties that hold for all surviving strategies.

Another way to deal with collusion is by aiming for an approximate best response or $\epsilon$-equilibrium (e.g., [25, 27]), which guarantees that deviations only provide minimum benefit. This approach could be used to disincentivize coalitions if colluding provides limited benefit (which, as seen in Section 3, may not be the case) and is largely complementary to our approach. Similarly, DCast [44] is an overlay multicast protocol that guarantees each node that follows the protocol some baseline payoff, even if others may collude. However, DCast does not aim to be an equilibrium and thus provides no guarantees that nodes will actually follow the protocol.

In some cases (e.g., in a multicast cost-sharing game [6]), mechanisms can be designed that are robust to coalitional deviations. However, since it is difficult to devise such mechanisms, many systems focus instead on detecting or reducing the effects of collusion. Several content distribution systems (e.g., [35, 36]) use incentives that attempt to reduce the benefits of collusion. Lian et al. [29] use a variety of techniques to detect collusion in a popular P2P service. Reiter et al. [37] design a reputation mechanism that require nodes to solve puzzles to prove they have the content in question. Tran et al. [40] develop a credit-based system in which a node's reputation is based on the number of distinct credit issuers it has received credit from and filters out those issuers that have issued excessive credits. EigenTrust [21] uses trusted peers to provide reputations that are robust against limited misbehavior (due to coalitions or failure). Similarly, Feldman et al. [19] and Marti et al. [30] describe reputation systems that place more trust and weight in certain peers' opinions. Finally, Zhang et al. [45] describe a heuristic for preventing colluding administrators from using links to increase the ranks of their pages in Google's PageRank algorithm. These systems can only ameliorate, not eliminate, the effects of collusion and provide no rigorous assurance that rational nodes will not deviate.

# 6. CONCLUSION

Trying to identify strategies that eliminate all incentives to collude, as traditional approaches attempt to do, is difficult, possibly futile, and fundamentally unnecessary. This paper introduces a new approach to handle the challenge posed by collusion: accept that coalitions will form, allow coalitions to benefit among themselves, and aim for stability by ensuring that the strategies or protocols specified for every *coalition*, not just every *node*, are best responses. While we are only beginning to explore the space of solution concepts and equilibria allowed by this new approach, we believe our initial results are encouraging: our proposed framework offers rigorous guarantees to both colluding and non-colluding nodes in cooperative services where traditional approaches are often provably unable to yield an equilibrium.

**Acknowledgments.** We are grateful to Tom Wiseman for many illuminating discussions. This material is based upon work supported by the National Science Foundation under grant 0905625.

# 7. REFERENCES

[1] BitTorrent. `http://bittorrent.com`.

[2] Abraham, I., Alvisi, L., and Halpern, J. Y. Distributed computing meets game theory: Combining insights from two fields. *SIGACT News 42*, 2, 69–76.

[3] Abraham, I., Dolev, D., Gonen, R., and Halpern, J. Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation. In *PODC 2006*.

[4] Abraham, I., Dolev, D., and Halpern, J. Y. Lower bounds on implementing robust and resilient mediators. In *TCC 2008*.

[5] Aiyer, A. S., Alvisi, L., Clement, A., Dahlin, M., Martin, J.-P., and Porth, C. BAR fault tolerance for cooperative services. In *SOSP 2005*.

[6] Archer, A., Feigenbaum, J., Krishnamurthy, A., Sami, R., and Shenker, S. Approximation and

collusion in multicast cost sharing. *Games and Economic Behavior 47*, 1, 36–71.

[7] Aumann, R. J. Acceptable points in general cooperative $n$-person games. *Annals of Mathematics Study 40 4*, 287–324.

[8] Bernheim, B. D., Peleg, B., and Whinston, M. D. Coalition-proof Nash equilibria, I. Concepts. *Journal of Economic Theory 42*, 1, 1–12.

[9] Bernheim, B. D., and Whinston, M. D. Multimarket contact and collusive behavior. *The RAND Journal of Economics 21*, 1, 1–26.

[10] Bloch, F., and Dutta, B. Correlated equilibria, incomplete information and coalitional deviations. *Games and Economic Behavior 66*, 2, 721–728.

[11] Chen, J., Micali, S., and Valiant, P. Robustly leveraging collusion in combinatorial auctions. In *ICS 2010*.

[12] Crawford, V. P., and Sobel, J. Strategic information transmission. *Econometrica 50*, 6, 1431–1451.

[13] Douceur, J. R. The Sybil attack. In *IPTPS 2002*.

[14] Einy, E., and Peleg, B. Coalition-proof communication equilibria. In *Social Choice, Welfare, and Ethics: Proceedings of the Eighth International Symposium in Economic Theory and Econometrics*.

[15] Eliaz, K. Fault tolerant implementation. *Review of Economic Studies 69*, 589–610.

[16] Farrell, J., and Rabin, M. Cheap talk. *Journal of Economic Perspectives 10*, 3, 103–118.

[17] Feigenbaum, J., Ramachandran, V., and Schapira, M. Incentive-compatible interdomain routing. In *EC 2006*.

[18] Feigenbaum, J., and Shenker, S. Distributed algorithmic mechanism design: Recent results and future directions. In *DIAL-M 2002*.

[19] Feldman, M., Lai, K., Stoica, I., and Chuang, J. Robust incentive techniques for peer-to-peer networks. In *EC 2004*.

[20] Goldberg, S., Halevi, S., Jaggard, A. D., Ramachandran, V., and Wright, R. N. Rationality and traffic attraction: Incentives for honest path announcements in BGP. In *SIGCOMM 2008*.

[21] Kamvar, S. D., Schlosser, M. T., and Garcia-Molina, H. The EigenTrust algorithm for reputation management in P2P networks. In *WWW 2003*.

[22] Kremer, S., Markowitch, O., and Zhou, J. An intensive survey of fair non-repudiation protocols. *Computer Communications 25*, 1606–1621.

[23] Lamport, L., Shostak, R., and Pease, M. The Byzantine generals problem. *ACM TOPLAS 4*, 3, 382–401.

[24] Levin, D., LaCurts, K., Spring, N., and Bhattacharjee, B. BitTorrent is an auction: Analyzing and improving BitTorrent's incentives. In *SIGCOMM 2008*.

[25] Levin, D., Sherwood, R., and Bhattacharjee, B. Fair file swarming with FOX. In *IPTPS 2006*.

[26] Levin, H., Schapira, M., and Zohar, A. Interdomain routing and games. In *STOC 2008*.

[27] Li, H. C., Clement, A., Marchetti, M., Kapritsos, M., Robison, L., Alvisi, L., and Dahlin, M. FlightPath: Obedience vs. choice in cooperative services. In *OSDI 2008*.

[28] Li, H. C., Clement, A., Wong, E. L., Napper, J., Roy, I., Alvisi, L., and Dahlin, M. BAR Gossip. In *OSDI 2006*.

[29] Lian, Q., Zhang, Z., Yang, M., Zhao, B. Y., Dai, Y., and Li, X. An empirical study of collusion behavior in the Maze P2P file-sharing system. In *ICDCS 2007*.

[30] Marti, S., and Garcia-Molina, H. Limited reputation sharing in P2P systems. In *EC 2004*.

[31] Moreno, D., and Wooders, J. Coalition-proof equilibrium. *Games and Economic Behavior 17*, 1, 80–112.

[32] Moscibroda, T., Schmid, S., and Wattenhofer, R. When selfish meets evil: Byzantine players in a virus inoculation game. In *PODC 2006*.

[33] Myerson, R. B. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA.

[34] Pagnia, H., and Gärtner, F. C. On the impossibility of fair exchange without a trusted third party. Tech. Rep. TUD-BS-1999-02, Darmstadt University of Technology Department of Computer Science.

[35] Peterson, R. S., and Sirer, E. G. Antfarm: Efficient content distribution with managed swarms. In *NSDI 2009*.

[36] Piatek, M., Krishnamurthy, A., Venkataramani, A., Yang, R., Zhang, D., and Jaffe, A. Contracts: Practical contribution incentives for P2P live streaming. In *NSDI 2010*.

[37] Reiter, M. K., Sekar, V., Spensky, C., and Zhang, Z. Making peer-assisted content distribution robust to collusion using bandwidth puzzles. In *ICISS 2009*.

[38] Rekhter, Y., Li, T., and Hares, S. A Border Gateway Protocol 4 (BGP-4). `http://www.ietf.org/rfc/rfc4271.txt`.

[39] Shamir, A. How to share a secret. *CACM 22*, 11, 612–613.

[40] Tran, N., Li, J., and Subramanian, L. Collusion-resilient credit-based reputations for peer-to-peer content distribution. In *NetEcon 2010*.

[41] Wong, E. L., and Alvisi, L. What's a little collusion between friends? Tech. Rep. TR-12-03, The University of Texas at Austin Department of Computer Science.

[42] Wong, E. L., Leners, J. B., and Alvisi, L. It's on me! The benefit of altruism in BAR environments. In *DISC 2010*.

[43] Wong, E. L., Levy, I., Alvisi, L., Clement, A., and Dahlin, M. Regret freedom isn't free. In *OPODIS 2011*.

[44] Yu, H., Gibbons, P. B., and Shi, C. DCast: Sustaining collaboration in overlay multicast despite rational collusion. In *CCS 2012*.

[45] Zhang, H., Goel, A., Govindan, R., Mason, K., and Roy, B. V. Making eigenvector-based reputation systems against collusions. In *WAW 2004*.