Strategies to Improve Reliability in Routing Overlay Networks with Selfish Nodes

Bruno Gusmão Rocha[1], **Virgilio Almeida**[1], **Dorgival Guedes**[1]

[1]Departamento de Ciência da Computação – Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte – MG – Brazil

`{gusmao, virgilio, dorgival}@dcc.ufmg.br`

Abstract. *Routing overlay networks use the current Internet infrastructure to provide support for a wide range of applications, such as content distribution, multiplayer games, search and web services. Recent studies have suggested that those networks may contain* selfish *nodes, i.e., nodes that develop their strategies considering only their own objectives and interests, regardless of the optimal global behavior. Extremely egoistic nodes may even decide not to give their own resources back to the network (free-riders), making overlay service unavailable to every node that depends on them. In this paper, we approach the problem of increasing the reliability of overlay services by detecting and removing free-riders from the selfish overlay network. We use a game-theoretic approach to evaluate the mechanism, and we show that service reliability can be increased to up 83% even when 70% of the network is formed by free-riders. Our mechanism is also fairly simple and can be efficiently distributed.*

1. Introduction

A routing overlay network is an application-layer overlay on top of the existing Internet routing substrate, that allows an alternative routing service in times of failures or degradation in the physical network [Andersen et al. 2001]. The success of those systems is based on the assumption that all participants are willing to cooperate with each other, and yield their own resources to their peers. However, in real situations, it is likely that nodes will act rationally (never taking any action that could negatively impact their self-interests) and egoistically (pursuing only to optimize their own benefits). When the behavior of such nodes are not taken into account, they may disregard any pre-defined protocol or network policy. Those selfish nodes are compelled to maximally explore all overlay services, while providing the least possible amount of resources to other members of the overlay. Actually, they may even be tempted to deny completely access to their resources, if there are no incentives against such strategy. It is clear that such behavior has a disastrous impact on the reliability of service and proper function of the system. Indeed, if too many nodes decide to adopt a free-rider behavior, the overlay service may be eventually disrupted.

In this paper, we address the problem of increasing the reliability of overlay services by detecting and removing free-riders from the selfish overlay network. We explore the concept of *reputation* [Axelrod 1984], defining it in our model as a measure of how fair and dependable a node is considered by another based on their direct interactions and also on the opinion of other peers who have interacted with it. Nodes with higher reputations have their requests served by other overlay members, while bad reputation nodes are very likely to have the access to overlay services denied. In this manner, it is in the best interest of all nodes, even the selfish ones, to achieve and maintain the best possible reputation. Our algorithm explores the intrinsic selfish and rational nature of the nodes to incentive them to become as reliable as possible, a necessary condition to attain high reputations and use overlay services. We conduct our study by modeling the interaction between nodes as a non-cooperative game, in which they choose to establish links to its neighbors seeking to maximize their own benefits (e.g., choosing neighbors with lowest latency, highest bandwidth, etc). The outcome of this game is a network topology given by the *Nash equilibrium*. The Nash equilibrium is a fundamental notion in game theory, and is characterized by the set of strategies adopted by each player such that there is no incentives to alter a strategy while all other players keep their own strategies unchanged. Our goal in designing the proposed mechanism is to define the ground rules of the game, so that free-riders are absent in the emerging topologies in the Nash equilibrium and service reliability is increased. We performed extensive experiments to study the behavior of our mechanism using different parameters, and demonstrate how fairness and reliability are reached in the overlay network.

This paper is organized as follows. In Section 2., we discuss the related work to our study. The problem formulation is given in section 3., while in section 4., we introduce the concept of reputation, the base of our approach. Section 5. explains how the game is set up and Section 6. describes the experimental evaluation. Finally, Section 7. summarizes the paper and gives directions for future work.

2. Related Work

The study of overlay routing networks have received considerable attention in the past few years [Andersen et al. 2001, Chu et al. 2000, Chawathe 2000]. These references assume that the participants are completely obedient to the pre-defined policies, regardless of the incentives to attain to those policies. However, more recent works [Fabrikant et al. 2003, Chun et al. 2004] suggest that nodes may actually behave selfishly, acting only towards their goals and self-interests. In the absence of a central authority that prescribes how nodes should set up links to each other, and the possible rewards and costs coming from the establishment of each new connection, it is up to the nodes to assign rewards and penalties to each link and then choose the best connections.

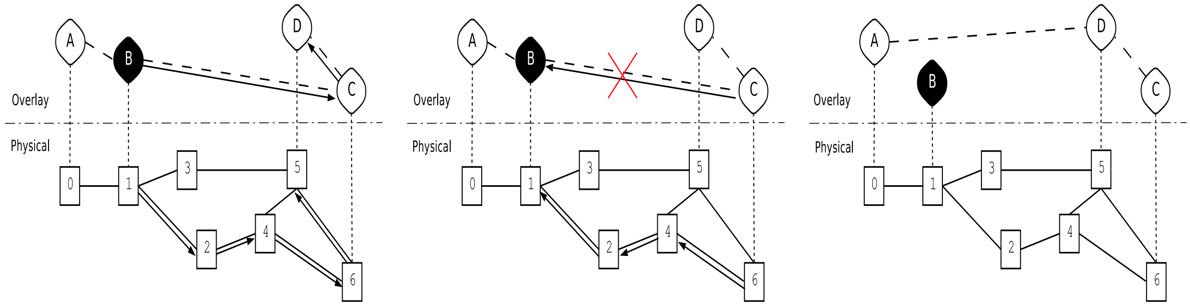
In this sense, selfish nodes have been modeled as rational players from a game theoretic perspective. They are viewed as players who wish to maximize their benefits when utilizing network services and to minimize their incurred costs when providing resources to other nodes. The benefits expected by a node may be conveniently translated to a utility function, that represents the anticipated payoff of each player, given a selected strategy. In [Chun et al. 2004] for instance, the utility function is given in terms of expected perceived latency, but other metrics such as bandwidth allocation, reliability or number of hops can be used as well. Although it is possible to use classical game theory [Qiu et al. 2003, Gollapudi et al. 2005] in this context, this approach only approximately describes a competitive game in overlay networks, since it requires that each player has the same game-relevant information as every other player. In this manner, all private strategies from the nodes should be public disclosed, which is clearly not a realistic assumption. The problem of creating a selfish overlay network through a non-cooperative incomplete-information game was introduced in [Fabrikant et al. 2003], and further scrutinized in [Chun et al. 2004]. They investigated the overlay topologies reached in the Nash equilibrium, and their properties, such as attack and failure resilience. Our work expands those studies by considering the free-rider problem and provides a fairness mechanism to promote reliability in overlay networks.

The question of fair resource allocation and reliability in competitive selfish environments has been the focus of attention of several studies. A common approach is to implement a pricing structure in the game, through which players can bid for resource units, such as QoS [Shu and Varaiya 2003] or bandwidth allocation [Wang and Li 2005]. The deployment of such mechanism is variable, either making use of a central entity to coordinate the process [Shu and Varaiya 2003] or distributing this responsibility to the players themselves [Wang and Li 2005] (which is more suitable in the overlay network case). But regardless of the chosen implementation, there is a common premise that fairness and reliability will be accomplished by explicitly putting a price on the network resources and creating a marketplace where players can pay and receive for services. This approach has the drawback of requiring a micro-payment system implementation that would allow nodes to carry out their transactions, which introduces a unnecessary complexity in the system and is by itself an ongoing research topic [Daras et al. 2003]. Furthermore, it is assumed that overlay nodes will be willing to operate based on a pricing scheme, with an explicit monetary charge. In contrast, many of the most successful cooperative networks nowadays use the same unit (e.g., bandwidth made available) both to reward and charge nodes, in conjunction with some kind of fairness policy [Cohen 2003]. This is the approach we take in this study.

3. Problem Formulation

The routing overlay problem consists of finding a path that satisfies some properties (e.g., shortest path, minimal latency, maximal bandwidth allocation, maximal reliability, etc.) to forward traffic over the overlay network. Figure 1 shows examples of overlay networks over the same physical topology, where the squares denote the routers and hosts of the substrate network and the circles denote the overlay nodes. We can see in Fig. 1(a) how node B uses the overlay to send a flow to node D, going through node C. This routing process is specially suited for inter-*autonomous systems*(AS's) routing, which carry the bulk of Internet traffic. The inter-domain links, fundamental for the operation of the network, are specially sensible to network failures, attacks and traffic congestion, and may take a potentially large time to converge after such events [Labovitz et al. 2000]. Routing overlay networks offer an alternative for the problem of inter-domain routing recovery. We assume that nodes in the overlay run their particular routing protocol which considers metrics of interest for the target application.

A basic principle of a routing overlay network is that, when a node routes the flow of some other node, it can expect that will have the service retributed when asked. In Fig. 1(a) for instance, node B (in black) used C's resources to route traffic to D. In this way, C might expect that, since it made its resources available to B, it can also use B's resources to route traffic to node A. Figure 1(b) however, shows a situation where this assumption fails, and B decides to deny C's request. This behavior is possible when nodes are selfish and rational, as in the focus of our study, because such nodes seek to maximize their benefits while minimizing their costs. In fact, becoming a *free-rider* [Adar and Huberman 2000] may even be the best strategy in some scenarios, if there



(a) Node B is a free-rider(in black), and use the overlay to route traffic

(b) Node C has his routing request denied by node B

(c) Possible solution: free-rider excluded and network reconnected

Figure 1. A Selfish Overlay Network

are no incentives against it.

The *free-rider problem* resides in detecting and removing such nodes from the network, since they only consume resources and don't contribute to the community. In our example, fig. 1(c) shows a possible solution, where node B is excluded and the overlay is reorganized. In practice, it is really not necessary to remove all free-riders' connections to cut them off the network, since the same effect is reached if fair nodes stop using those connections and start denying free-riders' requests. Our approach set ground policies that explore the nodes' egoistic behavior to exclude free-riders which promotes fairness and increases reliability.

4. Reputation Strategy

The free-rider problem can be viewed under the perspective of trust. In that sense, free-riders are just nodes that cannot be trusted by others, since they will not pay back the services they used when asked to do so. In our model, we use the concept of *reputation* as a measure of how reliable a node is. A rational node quickly realizes that there is no benefit in giving away resources to unreliable nodes, and the reputation mechanism is used to determine whether it is worth or not to trust a certain node. Nodes with higher reputations are more likely to have their requests served, while low reputation ones have only a small probability of getting access to the overlay resources. Thus, nodes have a clear incentive to be reliable and keep the best reputation possible. The reputation of each node is built upon the *individual experiences* with that node and the *peer testimonials* about it.

4.1. Individual Experience

We define the *individual experience* as the personal view of a node i about the reliability of another node j, based only on their past interactions. Conceptually, every time node i tries to make use of node j's resources, it updates its individual experience according to j's behavior, increasing it if j acts fairly and decreasing it otherwise. In this way, in order to formally define the individual experience, we must first state the concept of *fairness* in the overlay network. As fairness can be very subjective, we propose the following policy.

Definition 1 *Let* $S^t_{i(j)}$ *indicates the amount of service node* i *has provided to node* j *up to time slot* t*. If* i *requests* j*'s resources, then* j *is unfair to* $i \iff j$ *denies the request and* $S^t_{i(j)} - S^t_{j(i)} > 0$.

Definition 1 states clearly when any player in the game can consider another one unfair. In this way, a node j is not required to serve i's requests if the *balance* of provided services (e.g. routed traffic) between them is positive to j. In this manner, a player can be considered unfair when it does not assume the costs associate with the use of overlay services. We define the individual experience function as follows.

Definition 2 *Let* $I^t_{i(j)}$ *be the individual experience of node* i *regarding node* j *in a time* t*,* $0 \leq I^t_{i(j)} \leq 1$*. Let* r *be the amount of service requested by* i*,* p *the amount of service provided by* j *and* n *be the number of times* j *failed to completely process* i*'s requests,* $n \geq 1$*. Finally, assume that* $S_{i(j)} - S_{j(i)} > 0$*, i.e,* j *is in debt to* i*. Then,* $I^t_{i(j)}$ *can be defined as:*

$$I^t_{i(j)} = \begin{cases} min(I^{t-1}_{i(j)} + \alpha, 1) & if\, p = r \\ \\ max(I^{t-1}_{i(j)} - (1 - \frac{p}{r})\alpha n^2, 0) & otherwise \end{cases} \tag{1}$$

As we can see in Def. 2, the individual experience of the players about each other evolves with time, reflecting the fact that a node's strategy (and its behavior) may vary during the game. So, bad reputation nodes can be reintegrated if they change their behavior, and nodes with high reputations can be excluded should they become free-riders. The α parameter defines the minimal experience units involved in each interaction between i and j, i.e., how much the individual experience of i about j can change during each interaction.

There are also more interesting aspects in the individual experience formulation. The first one regards how fast it rises and falls. As it can be observed in Eq. 1, while the function linearly increases, it decreases much more quickly, as a function of n^2. This express the fact that is much easier to lose credibility than to acquire it, and that incentives players to keep good reputation levels. It also prevents players from turning into free-riders after reaching high reputations, since this behavior would lead to a fast decrease in reputation, and the consequent lost of access to overlay services. At the same time, the formulation is forgiving to players that occasionally fail in carrying out requests, since n has small values in those cases and the penalty is light. It is also worth to point out that partially served requests also take a lesser penalty, proportional to the amount of resources denied, as we can see by the term $(1 - \frac{p}{r})$. The intention is also not to punish fair nodes excessively, since a partially served request is more likely to be a failure then a deliberate denial of service.

4.2. Reputation

In our model, the reputation of a node j states how reliable a player i considers it, and it is used by i to determine if it is worth to provide its resources to j in a given time slot.

Definition 3 *Let* N *be the set of nodes in the overlay network, and* $R^t_{i(j)}, 0 \leq R^t_{i(j)}, \leq 1$ *be the reputation assigned by a node* i *to a node* j *in time* t*, then* $R^t_{i(j)}$ *is given by:*

$$R^t_{i(j)} = R^{t-1}_{i(j)} + \beta_i(T^t_{i(j)} - T^{t-1}_{i(j)}) + (1 - \beta_i)(I^t_{i(j)} - I^{t-1}_{i(j)}) \tag{2}$$

where $T^t_{i(j)}$ is given by:

$$T^t_{i(j)} = \frac{\sum_{k \in N} I^t_{k(j)} R^t_{i(k)}}{\sum_{k \in N} R^t_{i(k)}} \tag{3}$$

In other to normalize $R^t_{i(j)}$, we set it to 1 when the value of Eq. 2 exceeds one and set it 0 when it assumes a negative value. We can note in (3) that the reputation of a player is a function of its prior reputation, the changes in individual experience with that player and the variation in the *peer testimonials* $T^t_{i(j)}$. Peer testimonials(Eq. 3) are simply statements from every node in the game about the reliability of a particular node, or in other words, the opinion of the community about a specific player. The testimonial of each node is weighted by its own final reputation, so that the opinion of high reputation nodes has more impact than the one of low reputation ones. That is important, because low reputation nodes may be free-riders that are trying to maliciously defame fair nodes, and since the contribution of each node to the final reputation is proportional to the reputation of the nodes themselves, the opinion of free-riders is diluted and defaming is avoided.

Thus, a player can count on both its own experience with a node and the opinion of its peers to infer the currently reputation of that node. The β_i parameter, $0 \leq \beta_i \leq 1$, is used to control how important a node i considers the network opinion and how important it considers its own previous experiences with a player j. So, higher values of β_i emphasize the peer testimonials and the opinion of the overlay neighbors, while lower values of β_i favor the individual experience. It is easy to realize by Eq. 2 that the reputation of a node is not unique defined in the entire overlay network, but rather built individually by each player, based on different sources of data. We show that, even though there is no global reputation in the network, our mechanism leads most nodes to attribute low reputation to free-riders and exclude them from the network, increasing service reliability.

5. A Game-Based Approach

We model the interaction between the n nodes as a non-cooperative game, where each node is a player whose strategy is to select which nodes to connect in order to maximize its benefits. In the game, each player starts with a reputation of 0.5, which represents a default reliability value, i.e, neither is the node considered extremely reliable nor unreliable. Each player also has a *minimal acceptable reputation*, as shown in Def. 4.

Definition 4 *Let P be the set of players. Then, $\forall i \in P, \exists$* minimal acceptable reputation *$R_{min(i)}$, $0 \leq R_{min(i)} < 0.5$, such that if $R^t_{i(j)} < R_{min(i)}$, then i does not interact with j.*

If a player's reputation is bellow the minimal acceptable reputation of the node it wants to interact with, then all of its requests are denied. Furthermore, a player does not even connect to a node whose reputation does not meet the minimal acceptable reputation criteria, and it also never tries to use that node's resources, since it is identified as a free-rider. Note that since every player starts the game with a reputation of 0.5 and $R_{min(i)} < 0.5$ by definition, every node is considered reliable in the beginning of the game. The minimal reputation $R_{min(i)}$ can be seen as an individual measure of how much risk that a player i is willing to take when sharing its resources. The fact that each player i has its individual $R_{min(i)}$ is crucial for the mechanism operation, since strict players are the first to deny service to free-riders, and loose players make possible to include back nodes that have acquired low reputations in the past, but changed their behavior. Thus, if a free-rider changes its strategy (i.e., behaving fairly), resource sharing becomes possible again as soon as its reputation rises above the minimal. This is guaranteed by the utility function, as we will show in the next section.

5.1. Utility Function

The utility function is an important part of the game, and it mathematically characterizes a node's self-interest by quantifying the empirical benefits a chosen strategy can yield. An initial utility function could consider only the relative latency between the nodes in order to rate each connection. The relative latency (Def. 5) represents the latency of a connection to j compared to the latency of the best connection that i can establish in the overlay. We could define the utility function as proportional to the inverse of the relative latency (i.e., $1/L_{i,j}$), so that the smaller the relative latency, the greater would be the value of the utility function, and therefore the benefits associated with this connection. That formulation fits the rational and selfish nature of the players, since it meets their desire to maximize their benefits with the best possible connections.

Definition 5 *For the time slot* t*, let* $l^t_{i,j}$ *be the latency in a connection between node* i *and node* j *and let* $l^t_{min(i)}$ *be the minimal latency of any link between node* i *and every other overlay node. Then, the* relative latency $L_{i,j}$ *of* i *to* j *is given by:*

$$L_{i,j} = \frac{l^t_{i,j}}{l^t_{min(i)}} \tag{4}$$

However, in many cases the optimal theoretical network performance would only be possible considering that all nodes are evenly fair. Since that is not the case, connections to free-riders will bring no real benefits to the system and the real network performance is in practice much worse than the theoretical bound. In fact, not only do free-riders take their toll on performance, but on service reliability as well, since a node can not predict if there is a free-rider in the route path. The reputation concept is included in the formulation of the utility function to account for those situations, which is then given by:

$$u^t_{i(j)} = \begin{cases} \frac{R^t_{i(j)}}{L_{i,j}} & \text{if } R^t_{i(j)} > R_{min(i)} \\ \\ -\infty & \text{otherwise} \end{cases} \tag{5}$$

It is clear when we look at Eq. 5 why a node i does not interact with a player whose reputation is bellow $R_{min(i)}$: the expected payoff is $-\infty$. The basic goal of the players is to set up low latency links to *trustworthy* nodes. The utility function is used by player i to determine its strategy, which is the set of nodes chosen to be connected to. The total benefit of a graph G given by the overlay topology, with a set of chosen connections C in time t is given by Eq. 6.

$$B^t_i(G) = \sum_{j \in C} u^t_{i,j} \tag{6}$$

That is the function players try to maximize.

5.2. Game Model

The game takes place as an iterative procedure, in which players make their decisions based on their present perception of the network in order to select which actions to take next, such as adding or removing links and assigning reputation values to their peers. The game starts with a random connected graph; in each turn, each player modifies its strategy (the set of established connections) to maximize its benefits. The Nash equilibrium is reached when every player maintains its strategy unchanged, and the final overlay topology is obtained. Our mechanism demands that each player should have at least two links to other distinct overlay nodes, so they have sufficient conditions to provide routing services. This policy avoids that free-riders duck their responsibilities by simply becoming leaves in the topology graph.

The game starts as follows. In the beginning of every round (denoted by time t), each player updates the reputation from all nodes whose reputation has changed since the past round. The changes in the reputation of the peers are sent by each player in the end of the previous round, piggybacked in the regular overlay routing protocol packages. With the reputation information updated, the player starts computing the new payoff for each connection it participates, dropping those whose peer reputation falls bellow R_{min}.

After executing the initial actions, a player starts to effectively search the strategy space for better configurations, i.e, the addition of better links and the removal of the worse ones. This procedure is shown in Alg. 1

Algorithm 1 Strategy search for player i

```
 1: for all node $j \in$ overlay network do
 2:   if $R^t_{i(j)} > R^t_{min(i)} \wedge R^t_{j(i)} > R^t_{min(j)}$ then
 3:     Compute $B^t_i(G)$ with a link to $j$
 4:     if $B^t_i(G) - B^{t-1}_i(G) > \gamma$ then
 5:       Add link to $j$
 6:       $B^{t-1}_i(G) \leftarrow B^t_i(G)$
 7:     end if
 8:   end if
 9: end for
10: for all link $l$ from $i$ do
11:   if number of links $> 2$ then
12:     Compute $B^t_i(G)$ without $l$
13:     if $B^{t-1}_i(G) - B^t_i(G) < \gamma$ then
14:       Drop $l$
15:       $B^{t-1}_i(G) \leftarrow B^t_i(G)$
16:     end if
17:   end if
18: end for
```

To choose the best strategy in the time slot t, each node starts evaluating which new connections could be advantageously established, and which current ones give little return and could be safely removed. The procedure displayed in Alg. 1 can be divided into two parts. In the first one (lines 1-8), player i calculates the payoff of each possible connection not yet established to the other overlay nodes. First, i checks if the reputation of the node j, to whom it wants to connect, is above its minimal acceptable reputation. Since the information about its peers was updated in the beginning of the round, i has a very precise view of the current state of j's reputation, and it refuses to interact with this node if the minimum reputation criteria is not met. But player j has also to believe that it will not establish a connection with a free-rider that will only consume its resources. If both these conditions hold, it is possible to create the link, and then i calculates the benefits of the graph with the new link added (Eq. 6) and compares to the benefits without the link. The link is definitely added if the gain it brings is bigger than a given threshold γ. In the second part of Alg. 1 (lines 9-16), the player evaluates if any connection set up can be removed, provided that at least two other connections are still operational. A link can be removed if that imply in losses smaller then the threshold, or in other words, if the benefits of keeping the link are smaller then γ, a threshold that depends on the size of the network. Similar overlay network creation protocols [Chu et al. 2000, Chawathe 2000] also use a link periodical evaluation and a threshold dependent on the number of nodes to add or drop links. When the strategy is consolidated, the player can start using overlay resources, and adjust the personal reputation of the nodes with whom it interacts according to the behavior of those nodes, using the the mechanism described in Def. 2. After the player has used the overlay services, it broadcasts the recomputed reputation of the nodes involved in the past transactions, piggybacked in the regular overlay routing protocol packages, adding very little extra message overhead to the system.

6. Experimental Evaluation

In order to evaluate the proposed mechanism, we carried out extensive simulations to appraise the properties of reliability and fairness of the resulting topologies given by the game. Simulation results were averaged over at least 300 distinct simulation runs using the same initial set of parameters. The topology used, comprising 3477 hosts and 5049 links, was obtained by mapping the routes between several different nodes of the PlanetLab network[1] on the underlying IP infrastructure.

The metrics related to latency time and bandwidth consumption were also actual values gathered using that network. For the sake of simplicity, our simulations consider a snapshot of the network where latencies are considered constant, so that the variations in the parameters of our model can be isolated and better understood. For the simulation setup, we initially chose 100 nodes randomly from the substrate network to form the overlay, and characterized the ones with fastest links as free-riders. In this way, free-riders became preferential points of connections, and we could stress the worst possible scenarios, where opportunistic nodes are highly connected and likely to be chosen to serve requests. Fair nodes also have a 3% probability of failing to carry out a request, simulating periods of unavailability (hardware failures, congestion, etc.), when fair nodes apparently behave as free-riders. We expect that our mechanism can correctly differentiate those cases from free-riders. The minimal acceptable reputation of the players are uniformly distributed on the interval $0.0 \leq R_{min} < 0.5$ and the confidence interval is 95%.

6.1. Metrics and Strategies

To assess the effects of the mechanism, the following metrics were evaluated.

---

[1] `http://www.planet-lab.org/`

- *Service Provided to Free-Riders*: This metric shows how many players are willing to provide network service to free-riders as the game progresses. Ideally, free-riders should be denied access to all the overlay, which in practice would increase reliability by making more resources available to fair nodes. This metric shows how efficient a strategy is in detecting and removing free-riders.
- *Service Provided to Fair Players*: Similar to the metric above, it shows the average percentage of fair-players that are willing to provide service to other fair players. Thus, it demonstrates how susceptible a strategy is to false-positives, i.e., to the incorrectly assignment of fair nodes as free-riders.
- *Reliability*: In order to quantify the overlay reliability, or the probability of a node obtaining the service it requested, we use the ratio between the volume of traffic a node tried to send over the network and the volume of that flow which actually got delivered. That metric provides information about the quality and reliability of the overlay service. Values close to 1 indicate highly reliable topologies, while values close to 0 show a high free-rider incidence.
- *Reachable Nodes*: Defined as the average percent of fair players that every other fair player can reach in the graph. Clearly, if a node relies on a free-rider to route information to other peer, that peer is probably unreachable, since the free-rider will not fulfill the forwarding request.

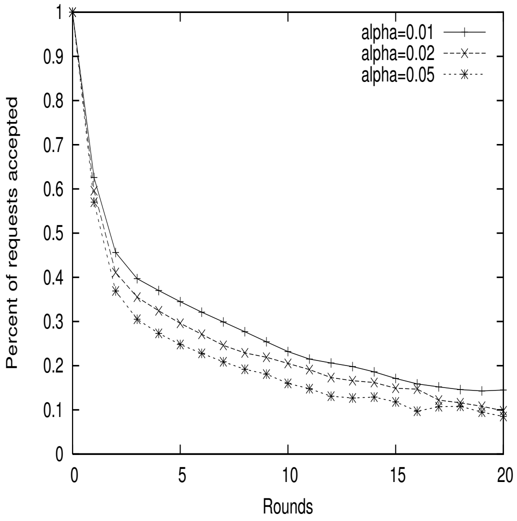


Figure 2. Requests by free-riders accepted varying α

First, we studied the impact of α parameter in our model to the service provided to free-riders. Fig. 2 shows the curves resulting from the variation of that parameter. We can observe that as the value of α is increased, the reputations of the players raise and decrease more quickly. This is related to how long it takes to the game to reach the Nash equilibrium. Since, in average, our simulations required the same number of rounds to converge to an overlay topology, higher values of α imposed higher penalties over free-riders on the same period of time. In other words, free-riders ended up with lower reputation in the Nash equilibrium, which in turn caused them to have more service denied by their peers in the network.

Strategy	Utility Function
Greedy	$\frac{1}{L_{i,j}}$ (Eq. 4)
Probabilistic	$\frac{1}{L_{i,j}}$ with $P(j) > \frac{1}{2^n}$
Reputation-based	$u_{i(j)}$ (Eq. 5)

Table 1. Evaluated strategies with their respective utility functions

To evaluate the effectiveness of our approach, we compared our mechanism with two other strategies. In the simulations, we keep $\alpha = 0.05$ but vary the β parameter to study its impact on the behavior of the system. We describe the other strategies evaluated bellow, and summarize the utility functions used by each one in table 1. The two following strategies are also evaluated.

Greedy: This is probably a typical strategy used in a selfish overlay network. Players do not try to identify free-riders and assume that all other players are evenly fair. The utility function is given only in terms of the relative latency.

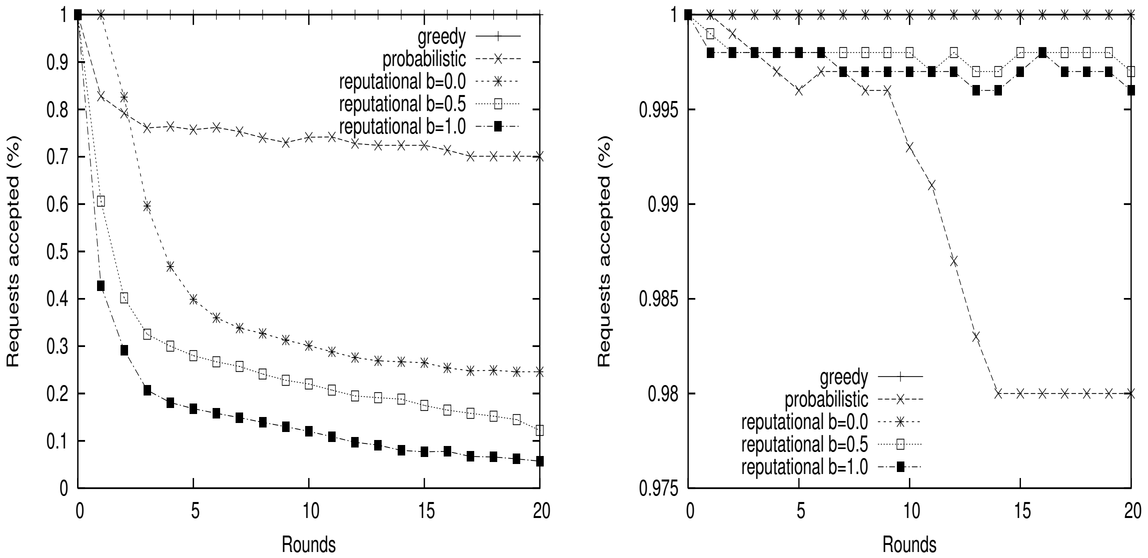
Probabilistic: In this strategy, not only do players consider the relative latency but also probabilistically decide how to establish connections and provide resources to other players, based on past interactions. A player A has a probability of $1/2^n$ to cooperate with a player B, where $n \geq 0$ is the number of times A's requests were denied by B.

We also wanted to evaluate the impact of using *peer testimonials* in the computation of reputations. For that we considered three values for β: 0, which meant reputations were computed solely based on individual experiences (no use of peer testimonials), 1, which meant the opposite, and 0.5, which represented a balanced combination of individual experiences and peer testimonials.

6.2. Results

Figure 3(a) shows the service provided to free-riders throughout the simulations using each strategy. When using the greedy utility function, players always serve free-rider's requests. That is expected, since by no means the overlay nodes try to isolate unfair peers in the network. The probabilistic approach shows to be more effective than the greedy one, and it is able to deny resource access to a significant percentage of free-riders. However, it is the reputation-based utility function which more efficiently identifies and prevents free-riders from having access to network services, with less than 5% of fair nodes still supplying free-riders' requests.

A free-rider detection mechanism should not only be able to identify unfair nodes, but also correctly separate them from the fair ones, and assures that the later will not be improperly penalized. Figure 3(b) shows that all strategies behave quite well regarding the willingness to provided service to fair players. Basically, these players are not mistaken as free-riders by their peers, and seldom have their requests denied. In spite of that, it is not sufficient that a routing request is accepted by fair players, it is also necessary that the route path does not contain any free-riders, since they will not honor the request and the destination host will be unreachable. We can observe the percentage of reachable nodes in Fig. 4(b). The reputation-based strategy clearly outperforms the other strategies evaluated, assuring that links are not established with free-riders and consequently, forming routes composed only by fair nodes. It is interesting to note that, although the re-

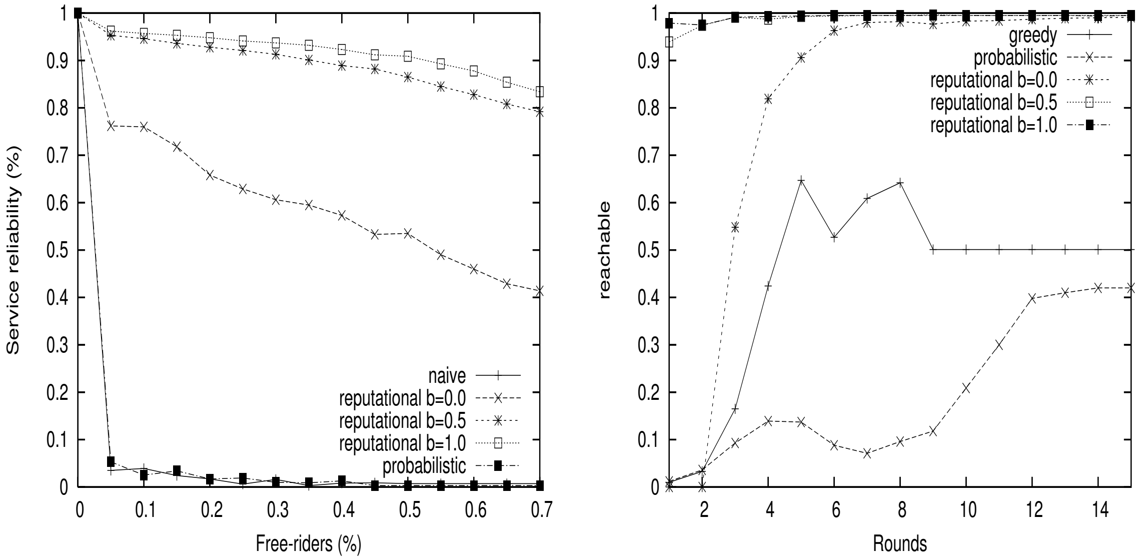


(a) Service to free-riders

(b) Service to fair nodes

Figure 3. Service ratios in a routing overlay network with 10% of free-riders

sources possibly available to fair nodes are quite substantial in the probabilistic approach (Fig. 3(b)), this strategy fails to keep routes from using free-riders. Therefore, the number of nodes effectively reachable is significantly lower than expected.

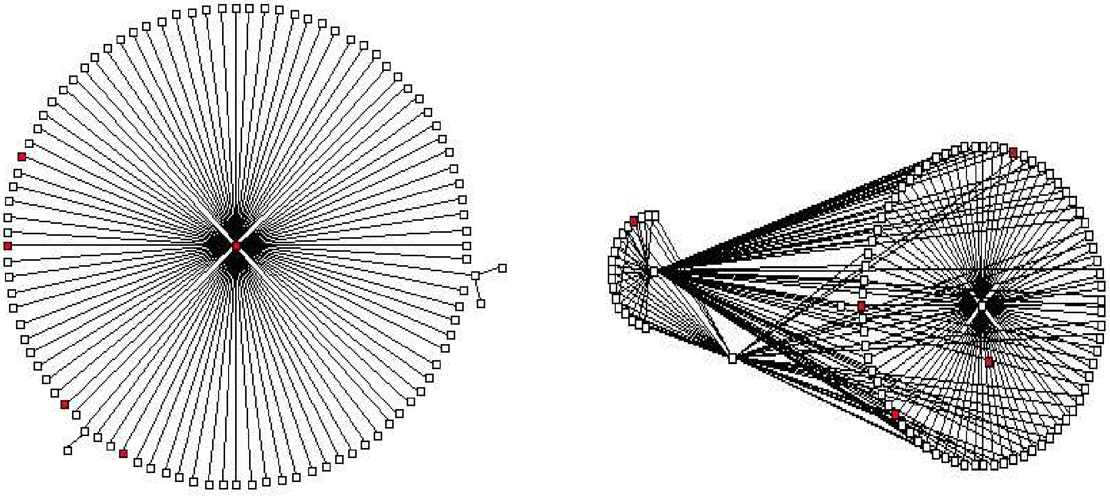


(a) Service reliability

(b) Reachable nodes

Figure 4. Reliability and reachable nodes

Fig. 4(a) shows how the presence of free-riders severely impairs the network capacity to meet the demands of its members. The *reliability* of the resulting topologies is



(a) Sample topology generated by the greedy strategy (free-riders in black)

(b) Sample topology generated by the reputational-based strategy (free-riders in black)

Figure 5. Sample topologies

significantly affected by the strategy used. Both the probabilistic and the greedy strategies fail to get around free-riders, and yield very low reliability ratios. The problem gets worse when the percentage of free-riders in the network grows, and with 10% of free-riders, almost no requests are completed anymore. In comparison, our mechanism is able to generate topologies free of opportunistic nodes, that could successfully increase reliability to about 83% of the submitted requests, even in the worst scenarios, where 70% of the network members are free-riders.

This situation is better understood when we examine the sample topologies generated in our simulations (Fig. 5(a) and 5(b)). When there is no restriction on the maximal degree of a node, most players choose to establish connections with only one node, possibly the one with the smallest latency, forming star topologies, as shown in Fig 5(a). When a free-rider happens to be the center of the star, all routes must include that free-rider as a hop. Since the free-rider will deny any routing request, all overlay nodes become in fact unreachable. This is not the case in the reputation-based strategy, as seen in the sample topology in Fig. 5(b). Although the topology is still star-like shaped, multiple centers are formed, which guarantee a significant increase in the amount of available routes. Moreover, since the free-riders have very low reputation throughout the game, they never become preferential connection points, and consequently, are never one of the topology centers. We can clearly see that, the more a player uses information provided by others to compute the reputation values, the better are the exclusion of free-riders and the reliability ratios. That is due to the fact that when a node relies on the peer testimonials (β closer to one), it gets much more data about other nodes than when it only considers its own interactions with them, and thus can take more precise and informed decisions in the game.

7. Conclusions

In this work we present a novel reputation-based mechanism to increase the reliability and fairness of selfish overlay networks which combine both direct experiences between nodes and experiences reported by others.We use a non-cooperative game model to evaluate the effectiveness of our approach, and test it against different linking cost functions. We verify that our mechanism is effectively capable of preventing free-riders from draining network resources, without incorrectly penalizing fair-nodes. We also find that our mechanism can increase service reliability to 83% of fair nodes' requests, even in the worst case scenarios. There are several interesting directions for future work. We intend to study dynamic networks and solutions to refrain players who acquire bad reputations to leave and then re-join the network with a new identity. Previous works [Friedman and Resnick 1998] have indicated that imposing a penalty on newcomers (e.g., a minimal number of served requests) can prevent this kind of behavior, and we plan to examine how such strategies can be included into our model. We also want to examine the behavior of our mechanism in a dynamic latency environment, adding congestion and traffic fluctuations to our simulations and to conduct a more comprehensive intestigation about the impact of topologiy sizes in our approach.

References

Adar, E. and Huberman, B. (2000). Free riding on gnutella. Technical report, Xerox Parc.

Andersen, D. G., Balakrishnan, H., Kaashoek, M. F., and Morris, R. (2001). Resilient overlay networks. In *Symposium on Operating Systems Principles*, pages 131–145.

Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books, New York.

Chawathe, Y. (2000). Scattercast: An architecture for internet broadcast distribution as an infrastructure service.

Chu, Y.-H., Rao, S. G., and Zhang, H. (2000). A case for end system multicast. In *Measurement and Modeling of Computer Systems*.

Chun, B., Fonseca, R., Stoica, I., and Kubiatowicz, J. (2004). Characterizing selfishly constructed overlay networks. In *Proceedings of IEEE INFOCOM'04*, Hong Kong.

Cohen, B. (2003). Incentives build robustness in bittorrent. In *Proceedings of the First Workshop on the Economics of Peer-to-Peer Systems*, Berkeley, CA.

Daras, P., Palaka, D., Giagourta, V., Bechtsis, D., Petridis, K., and Strintzis, M. G. (2003). A novel peer-to-peer payment protocol. In *Proceedings of IEEE EUROCON*.

Fabrikant, A., Luthra, A., Maneva, E., Papadimitriou, C., and Shenker, S. (2003). On a netowrk creation game. In *Proceedings of ACM PODC*. ACM Press, New York.

Friedman, E. and Resnick, P. (1998). The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, 10:173–199.

Gollapudi, S., Sivakumar, D., and Zhang, A. (2005). Exploiting anarchy in networks: A game-theoretic approach to combine fairness and throughput. In *Proceedings of IEEE INFOCOM'05*, Miami.

Labovitz, C., Ahuja, A., Bose, A., and Jahanian, F. (2000). Delayed internet routing convergence. In *SIGCOMM*.

Qiu, L., Yang, Y. R., Zhang, Y., and Shenker, S. (2003). On selfish routing in internet-like environments. In *Proceedings of the ACM SIGCOMM*, Germany.

Shu, J. and Varaiya, P. (2003). Pricing network services. In *Proceedings of IEEE INFOCOM 2003*. IEEE.

Wang, W. and Li, B. (2005). Market-driven bandwidth allocation in selfish overlay networks. In *Proceedings of IEEE INFOCOM 2005*. IEEE.