Free-Riding and Whitewashing in Peer-to-Peer Systems

Michal Feldman, Christos Papadimitriou, John Chuang, and Ion Stoica

***Abstract*—We devise a model to study the phenomenon of free-riding and free-identities in peer-to-peer systems. At the heart of our model is a user of a certain *type*, an intrinsic and private parameter that reflects the user's willingness to contribute resources to the system. A user decides whether to contribute or free-ride based on how the current contribution cost in the system compares to her type. We study the impact of mechanisms that exclude low type users or, more realistically, penalize free-riders with degraded service. We also consider dynamic scenarios with arrivals and departures of users, and with *whitewashers*—users who leave the system and rejoin with new identities to avoid reputational penalties. We find that imposing penalty on all users that join the system is effective under many scenarios. In particular, system performance degrades significantly only when the turnover rate among users is high. Finally, we show that the optimal exclusion or penalty level differs significantly from the level that optimizes the performance of contributors only for a limited range of societal generosity levels.**

***Index Terms*—Free-riding, incentives, peer-to-peer (P2P), whitewashing.**

I. Introduction

WHY is free-riding widespread among users of peer-to-peer (P2P) systems? How does free-riding affect system performance? What mechanisms discourage free-riding? How does whitewashing affect the performance of P2P systems?

These are the questions that motivate us.

P2P systems rely on voluntary contribution of resources from the individual participants. However, individual rationality results in free-riding behavior among peers, at the expense of collective welfare. Empirical studies have shown prevalent free-riding in P2P file sharing systems [1], [2]. Various incentive mechanisms have been proposed to encourage cooperation in P2P systems [3]–[7]. At the same time, it has been suggested that free-riding can be sustained in equilibrium and may even occur as part of the socially optimum outcome [8].

Manuscript received February 15, 2005; revised January 15, 2006. This work was supported in part by the National Science Foundation under ITR Award ANI-0085879, ITR Award ANI-0331659, ITR Award ANI-0225660, and Career Award ANI-0133811. The work of M. Feldman was done while with the School of Information Management and Systems, University of California, Berkeley, CA.

M. Feldman is with the School of Computer Science and Engineering, Hebrew University of Jerusalem, Jerusalem 91904, Israel (e-mail: mfeldman@cs.huji.ac.il).

C. Papadimitriou and I. Stoica are with the Computer Science Division, University of California, Berkeley, CA 94720-1776 USA (e-mail: christos@cs.berkeley.edu; istoica@cs.berkeley.edu).

J. Chuang is with the School of Information Management and Systems, University of California, Berkeley, CA 94720 USA (e-mail: chuang@sims.berkeley.edu).

Digital Object Identifier 10.1109/JSAC.2006.872882

Empirical research in behavioral economics has demonstrated that models of purely self-interested agents usually fail to explain observed behavior of people [9]. The *warm-glow* model [10], [11] attempts to account for observed user behavior by incorporating into one's utility function the utility he gains from the mere act of giving. Consistent with this approach, we develop a modeling framework that takes users' *generosity* into account. At the heart of our model is a user as a rational agent with a private and intrinsic characteristic called her *type*, a single parameter reflecting the willingness of the user to contribute resources. Type can be intuitively thought of as a quantitative measure of decency or generosity.

Each user decides whether to contribute or free-ride based on the relationship between the cost of contribution and her type. We assume that the cost of contributing is the inverse of the total percentage of contributors, because when many people free-ride, the load on contributors increases. Thus, if at present a fraction x of the users contribute, the decision of a rational user with type θ_i is:

$$\begin{array}{ll} \text{Contribute,} & \text{if } \frac{1}{x} < \theta_i \\ \text{Free-ride,} & \text{otherwise.} \end{array}$$

Even within this minimalistic framework, we can already see some interesting implications. In this "free market" (FM) environment, the percentage x of contributors in equilibrium is determined as the intersection of the type distribution, $x = Pr(\theta_i \geq t)$ with the curve $x = 1/t$.

Fig. 1 demonstrates the equilibria when the *type* is uniformly distributed between 0 and a maximal value, θ_m. In this case, there are three equilibria in the system. The first two are the two intersection points of the curves and the third equilibrium, which always exists, is $x = 0$. Obviously, when no user contributes, the contribution cost becomes too high for someone to contribute. Consider the natural fixpoint dynamics of such a system, i.e., starting at some initial x, users arrive at individual decisions, their aggregate decisions define a new x, which leads to a new aggregate decision, and so on. When the system is out of equilibrium, the direction in which the system moves depends on the relative heights of the type distribution curve and the cost curve. If the cost curve lies above the distribution curve, contribution cost is higher than the fraction of users who are willing to contribute at this cost, so the fraction of contributors decreases. For example, in Fig. 1, this happens for $x < x_2$ or $x > x_1$. In contrast, if the cost lies below the distribution curve, contribution cost is lower than the willingness to contribute, so contribution level increases. This is the case for $x_1 < x < x_2$ in the figure. Therefore, $x = x_1$ and $x = 0$ are the two attractors of the fixpoint dynamics, so as long as the initial x lies above the lower

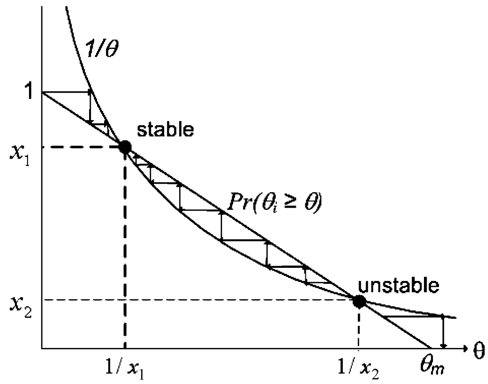


Fig. 1. The intersection points of the two curves represent two equilibria of the system. The curve $1/\theta$ represents the contribution cost, and $Pr(\theta_i >= \theta)$ represents the generosity CDF, assuming $\theta_i \sim U(0, \theta_m)$. The higher equilibrium (contribution level x_1) is stable. The point $x = 0$ is an additional equilibrium of the system.

intersection point (x_2), the process converges to the upper one (x_1). If the initial x is below the lower intersection point or if there is no intersection, i.e., when there are too many selfish rascals around, then x becomes 0 and the system collapses.

So far, we have been interested only in costs. To understand system performance, we need to consider system benefits as well. What is a user's benefit when the level of contribution is x? We assume that the benefit a user receives from participation in the system (whether or not she contributes) is proportional to the contribution level in the system, and thus a function of the form αx for some constant $\alpha > 0$.

The performance of the system, denoted by W_S, is defined as the difference between: 1) the total benefit received by all users (including both contributors and free-riders) and 2) the total contribution cost experienced by all users, which effectively include only the contributors because free-riders incur no costs. Note that the network size is normalized to 1. We get

$$W_S = \alpha x - \frac{1}{x} x = \alpha x - 1.$$

With this, we are ready to tackle more questions.

1) Would excluding low-type users from the system improve performance? The answer seems to be true only if the societal generosity level is low (see Section III).
2) The exclusion scenario is unrealistic because users' types are private. What if free-riding *behavior* brings some form of penalty, that is, deterioration of benefits by a fraction of p? We find that the penalty mechanism is effective in discouraging free-riding behavior when the threat is sufficiently high relative to the contribution cost (see Section IV). Moreover, for a sufficiently high threat, no social cost is incurred because no user is effectively penalized, so the optimal performance is achieved.
3) Imposing penalties on free-riders require a way to identify free-riders and distinguish them from contributors. Reputation systems [12], [13] may help, but these systems are vulnerable to the *whitewashing* attack, where a free-rider repeatedly rejoins the network under new identities to avoid the penalty imposed on free-riders [14]. The whitewashing attack is made feasible by the availability of low cost identities or *cheap pseudonyms*. There are two ways to counter whitewashing attacks. The first is to require the use of free but irreplaceable pseudonyms, e.g., through the assignment of strong identities by a central trusted authority [15]. In the absence of such mechanisms, it may be necessary to impose a penalty on all newcomers, including both legitimate newcomers and whitewashers. This results in a social cost due to cheap pseudonyms, as suggested by Friedman and Resnick [14]. In Section V, we quantify the social loss due to cheap pseudonyms and find that performance is significantly affected only if the turnover rate is high.
4) *System* performance measures the total performance realized by *all* users in the population, including the free-riders. One may wish to treat the performance realized by contributors differently than that of free-riders on grounds of fairness. How would the results change if we consider only the performance of contributors? How is the exclusion or penalty levels that optimize system performance different than the ones optimizing contributor performance? We address these questions in Section VI and find that the values that optimize system performance and contributor performance coincide when the societal generosity level is either high or low.

II. Model

At the heart of our model is a user as a rational agent with a private and intrinsic characteristic, called her type—a single parameter reflecting the willingness of the user to contribute resources to the system. Each user contributes resources if her type, θ_i, is greater than the contribution cost in the system, which is assumed to be inversely proportional to the contribution level. We get that the contribution level, x, is the fraction of users whose generosity (type) exceeds $1/x$. Thus, the fraction of users who contribute is derived by solving the fixpoint equation

$$x = \text{Prob}\left(\theta_i \geq \frac{1}{x}\right). \quad (1)$$

A user's type is a random variable with unknown distribution. To solve this equation, we need to make assumptions about the type distribution. We consider a distriubtion in which a fraction of the users are uniformly distributed between 0 and θ_m, and the remaining users are equally split between having type 0 and type θ_m. Formally:

- fraction ϕ of the users are uniformly distributed, with $\theta_i \sim U(0, \theta_m)$;
- fraction $(1-\phi)/2$ of the users are of type $\theta_i = 0$;
- fraction $(1-\phi)/2$ of the users are of type $\theta_i = \theta_m$.

The parameter $\phi \in [0, 1]$ determines the degree of bimodality of the distribution, with $\phi = 0$ corresponding to an extreme bimodal distribution and $\phi = 1$ corresponding to a uniform distribution. θ_m is the maximum willingness to contribute resources, and the expected type is always $\theta_m/2$, independent of the value of ϕ. θ_m is thus an important parameter of the system, as it reflects the societal "generosity."

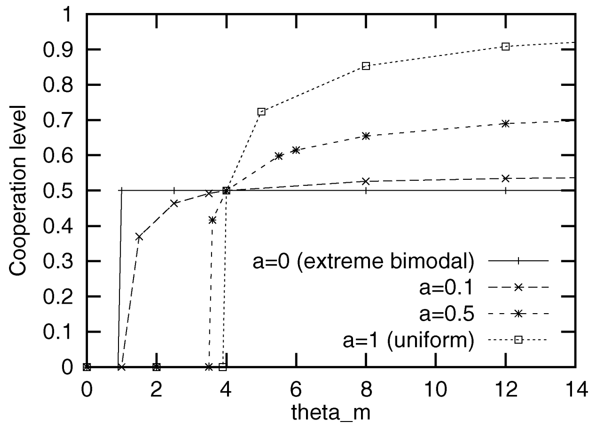


Fig. 2. Contribution level as a function of the societal generosity level for different generosity distributions.

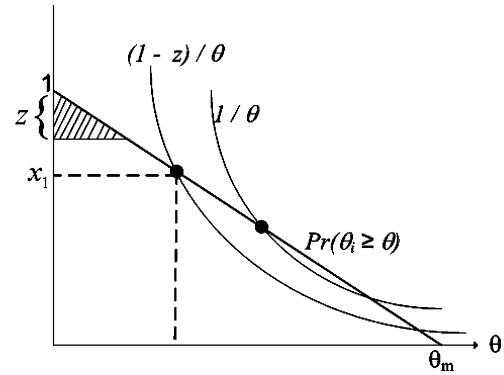


Fig. 3. This figure demonstrates the effect of the exclusion mechanism. Under the exclusion mechanism, the cost curve shifts from $1/t$ to $(1-z)/t$, consequently, the attractor (x_1) becomes higher. The shaded area represents the excluded users.

Based on (1), the contribution level, x, is the solution to the fixpoint equation

$$x = \phi \frac{\theta_m - \frac{1}{x}}{\theta_m} + \frac{1-\phi}{2}$$

which yields

$$x_{1,2} = \frac{\theta_m \phi + \theta_m \pm \sqrt{\theta_m^2 \phi^2 + 2\theta_m^2 \phi + \theta_m^2 - 16\theta_m \phi}}{4\theta_m}. \quad (2)$$

The larger root, x_1, is the attractor of the system (see Fig. 1). As demonstrated in Fig. 2, the contribution level varies depending on the type distribution and range (reflected by ϕ and θ_m).

Claim 1: The contribution level in equilibrium increases in θ_m and converges to $(1+\phi)/2$ as θ_m goes to ∞. In addition, contribution level falls to zero when $\theta_m < \max\{1, (16\phi/(1+\phi)^2)\}$.

Proof: It is easy to see that when θ_m goes to ∞, the only users who do not contribute are the users whose type equals to 0. Therefore, the contribution level is $1-(1-\phi/2) = 1+\phi/2$. To derive the threshold value, note that the contribution level is greater than 0 only when there is an intersection between the two curves, which happens when the expression within the square root in (2) is greater or equal to zero. We get

$$\theta_m^2 \phi^2 + 2\theta_m^2 \phi + \theta_m^2 - 16\theta_m \phi \geq 0 \Leftrightarrow \theta_m \geq \frac{16\phi}{(1+\phi)^2}.$$

However, if $\theta_m < 1$, then for all i, $\theta_i < 1$, and since $(1/x) \geq 1$, based on (1), we get that $x = 0$ is the only solution. Therefore, the threshold is $\max\{1, (16\phi/(1+\phi)^2)\}$. ∎

A uniform distribution and an extreme bimodal distribution are two special cases of the above distribution with $\phi = 1$ and $\phi = 0$, respectively. Under a uniform distribution, the system is sustained for $t_m \geq 4$, compared with $t_m \geq 1$ for a bimodal distribution. Thus, a bimodal type distribution can better sustain the system when the societal generosity level is low. Conversely, when the societal generosity is high, a uniform type distribution can realize a higher contribution level and system performance.

In the remainder of this paper, we restrict attention to the uniform type distribution (i.e., $\phi = 1$) and leave a more thorough analysis of additional type distributions for future work.

III. Exclusion Mechanism

The analysis presented above suggests that when the societal generosity is low, the system cannot be sustained without intervention. In this section, we analyze the effect of intervention in the form of exclusion.

If we had perfect information about the type of each individual user, we could *exclude* the users of the lowest type in order to increase the contribution level. This shifts the cost curve downward (see Fig. 3), resulting in a higher contribution level. However, exclusion also decreases performance by limiting the number of users who enjoy the system's benefits. The trade-off is optimized at a particular exclusion level.

If a fraction z of users are excluded, the cost of contribution becomes $(1-z)/x$, and the fixpoint equation describing contribution level is

$$x = \text{Prob}\left(t_i \geq \frac{1-z}{x}\right) \quad (3)$$

which yields the attractor

$$x_1 = \frac{\theta_m + \sqrt{\theta_m^2 - 4\theta_m + 4\theta_m z}}{2\theta_m}.$$

Based on this expression, x is defined only for $z \geq 1 - \theta_m/4$. Also, notice that x represents the contribution level in the entire system rather than that in the post-exclusion system; therefore, the effective contribution level cannot exceed $1-z$, and we get

$$x = \min(x_1, 1-z). \quad (4)$$

With the exclusion in effect, the system performance becomes

$$W_S(z) = (\alpha x - 1)(1-z).$$

The optimal exclusion level is

$$z_s^* = \arg\max_z W_S(z).$$

Claim 2: The optimal exclusion level (z_s^*) and the corresponding contribution level (x) are the following.

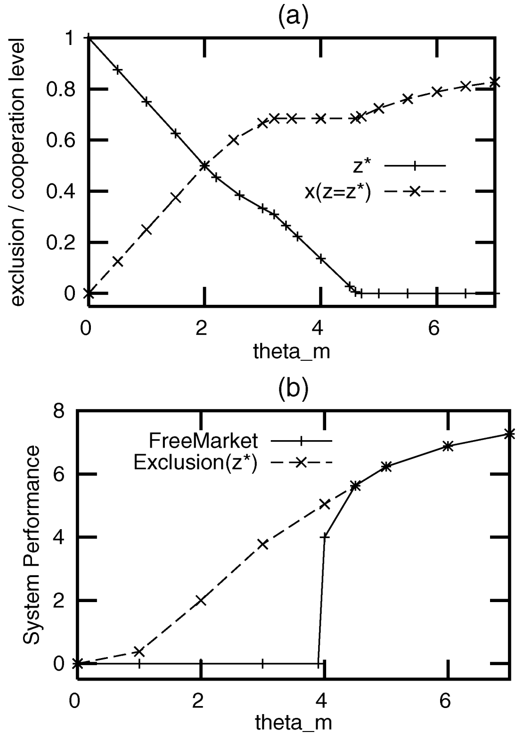


Fig. 4. (a) Optimal exclusion level (z_s^*) and corresponding contribution level ($x(z = z_s^*)$) under the exclusion mechanism. (b) System performance (W_S) under exclusion and FM (no exclusion). $\alpha = 10$.

- For $\theta_m \in [0, 2)$: $z_s^* = 1 - \theta_m/4$, and $x = 1 - z$.
- For $\theta_m \in [2, \theta_m')$: $z_s^* = 1/\theta_m$, and $x = 1 - z$.
- For $\theta_m \in [\theta_m', \theta_m'')$

$$z_s^* = \Big(9\alpha^2 - \theta_m\alpha^2 - \alpha\theta_m(2 + \sqrt{\alpha^2 - \alpha + 1}) + 2\theta_m(1 + \sqrt{\alpha^2 - \alpha + 1})\Big)/(9\alpha^2)$$

and

$$x = \frac{\theta_m + \sqrt{\theta_m^2 - 4\theta_m + 4\theta_m z}}{2\theta_m}.$$

An interesting property of this region is that x is independent of θ_m (by substituting $z = z_s^*$ in x).
- For $\theta_m \in [\theta_m'', \infty)$: $z_s^* = 0$, and $x = (\theta_m + \sqrt{\theta_m^2 - 4\theta_m})/2\theta_m$.

The proof can be found in the appendix.

Fig. 4 presents the behavior of the system under the exclusion mechanism as a function of θ_m. Fig. 4(a) presents the optimal exclusion level and its corresponding contribution level, and Fig. 4(b) presents the optimal system performance with and without exclusion.

Based on these results, if the societal generosity is sufficiently large ($\theta_m >\sim 4.6$), $z_s^* = 0$; that is, intervention is not necessary since any performance improvements realized by the high-type users are offset by the loss in benefits for the excluded low-type users. In contrast, for low θ_m values ($\theta_m <\sim 3.2$), exclusion of low-type users is effective in preventing a total collapse in cooperation. Indeed, the optimal exclusion level is such that all of the users that remain in the system contribute. For intermediate θ_m values, ($\sim 3.2 < \theta_m <\sim 4.6$), optimal system performance occurs when a few low-type users are excluded, yet the remaining population is a mixture of contributors and free-riders. In this region, the optimal cooperation level is independent of θ_m.

IV. Penalty Mechanism

The exclusion mechanism suffers from two problems. First, user types may not be observable in practice. Second, even if user types were observable, excluding users based on their innate type rather than their actual behavior precludes the possibility of rational decision-making by the users in face of incentives.

In this section we introduce the *penalty* mechanism. The penalty mechanism assumes that free-riding behavior is observable, even though innate user types may not be; that is, users are labeled as either contributors or free-riders, and being a free-rider entails a penalty—deterioration of a user's benefits by a fraction of p. While this assumption may still be too strong, various mechanisms that have been proposed and analyzed (see, e.g., [3], [12], and [16]) support this approach. An example penalty would be exclusion with probability p. Another example penalty, which is mathematically equivalent to the first, is service differentiation, under which free-riders' system benefits are reduced, while contributor benefits are not.

Downgrading the performance of the free-riders has two effects, both of which lead to a higher contribution level. First, since free-riders get only a fraction $1-p$ of the benefits, the load placed on the system decreases to $x + (1-x)(1-p)$; therefore, contribution cost becomes $(x + (1-x)(1-p))/x$. Second, the penalty introduces a threat, since users who free-ride know that they will get reduced service.

Let Q denote the individual benefits, R denote the reduced contribution cost and T denote the threat. Under the penalty mechanism, the realized performance of contributors and free-riders is

$$\begin{aligned} W_C &= Q - R = \alpha x - \frac{x + (1-x)(1-p)}{x} \\ W_{\text{FR}} &= Q - T = \alpha x - p\alpha x \end{aligned}.$$

Consequently, the contribution level x is derived according to the following expression:

$$x = \text{Prob}(t_i \geq R - T).$$

That is

$$x = \text{Prob}\left(t_i \geq \frac{x + (1-x)(1-p)}{x} - p\alpha x\right). \quad (5)$$

The attractor of this fixpoint equation is

$$x = \frac{p - \theta_m + \sqrt{p^2 + 2\theta_m p + \theta_m^2 - 4\theta_m + 4p\alpha - 4p^2\alpha}}{2(-\theta_m + p\alpha)}.$$

System performance now becomes

$$W_S(p) = (\alpha x - 1)(x + (1-x)(1-p))$$

and the optimal penalty level is

$$p_s^* = \arg\max_p \quad W_S(p).$$

Imposing a penalty on free-riders, while increasing the contribution level, entails some social loss. This may lead one to conclude that we should not exceed a particular penalty level. However, if p is sufficiently high, the threat exceeds contribution cost, thus all users cooperate and no penalty is actually imposed. Based on (5), this is achieved when $p \geq (1/\alpha)$. In this case, $x = 1$ and the maximal system benefit (which equals to α) is attained. For example, if $\alpha = 10$, we only need a mechanism that can catch and exclude a free-rider with 10% probability, but if $\alpha = 1.1$, we will need to increase the probability to over 90%.

These results suggest that if we impose a sufficiently high penalty, or are able to identify and exclude free-riders with high probability, we can achieve optimal system performance. However, in many cases it may be difficult or costly to identify free-riders with certainty, and p will be restricted by a maximal feasible value, denoted by p_m by a maximal feasible value. As long as $p \geq 1/\alpha$, optimal system performance can be achieved, regardless of the value of θ_m. On the other hand, if the penalty is set too low, the resulting performance is not significantly better than the FM outcome.

V. Social Cost of Free Identities

In Section IV, we show that a penalty mechanism can discourage free-riding behavior. However, the effectiveness of penalties can be undermined by the availability of cheap pseudonyms. In particular, a free-rider might choose to *whitewash*, i.e., leave and rejoin the network with a new identity on a repeated basis, to avoid the penalty imposed on a free-rider. The lower the cost of acquiring new identities, the more likely a free-rider will engage in whitewashing. Since whitewashers are indistinguishable from legitimate newcomers, it is not possible to single them out for the imposition of a penalty. Of course, it is possible to counter the whitewashing strategy by imposing the penalty on all newcomers. However, this results in a social cost, as shown by Friedman and Resnick [14].

In this section, we are interested in quantifying the social cost of cheap pseudonyms in terms of reduced system performance. We do so by extending our model from Section IV into a dynamic model in which users join and leave the system. To quantify the performance reduction due to cheap pseudonyms, we consider two dynamic scenarios: *permanent identities* (PI) and *free identities* (FI).

Under PI, identity costs are taken to be infinity (i.e., they are irreplaceable), while under FI, they are free.[1] As we will see below, these extreme cases provide important insights while preserving simplicity.

We model a system where some users leave and newcomers join, with a turnover rate of d (Fig. 5). We assume that arrivals and departures are type-neutral and therefore do not alter the type distribution. In addition, we assume that the time scale of the service policies is relatively large with respect to user behavior. The model can be extended in future work by considering more sophisticated dynamics, as discussed in Section VII.

In this model, users can be distinguished along two dimensions: existing members versus newcomers, and contributors

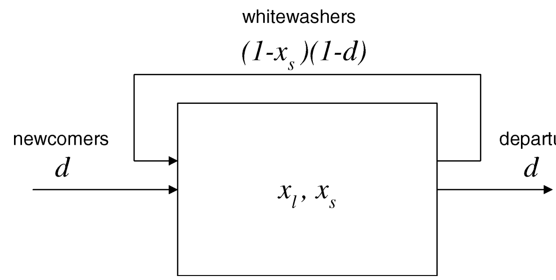


Fig. 5. Dynamic system with arrivals, departures, and whitewashers. A fraction d of users depart and are replaced by the same number of newcomers. At the same time, a fraction $(1-d)(1-x_s)$ of users choose to whitewash when new identities are costless.

versus free-riders. This means that the population is comprised of four groups of users: existing contributors (EC), existing free-riders (EF), new contributors (NC), and new free-riders (NF). In a system with FI, the EF may choose to adopt the whitewashing strategy (WW) if the penalty imposed on newcomers is smaller than that imposed on EF.

An important property of the dynamic scenarios is that not all users care about the threat. The users who leave the system at the end of each period are not affected by the penalty they would have paid had they stayed in the system. Consequently, we get two separate contribution levels: x_l: the contribution level of users who leave, and x_s: the contribution level of users who stay. The values of x_s and x_l in equilibrium are determined by the following equations:

$$x_l = \text{Prob}(t_i \geq R) \quad (6)$$
$$x_s = \text{Prob}(t_i \geq R - T) \quad (7)$$

(recall that R and T denote the contribution cost and the threat, respectively). Consequently, the average contribution level in the system, denoted by x_a, is: $x_a = dx_l + (1-d)x_s$.

The contribution level of users who stay is always greater than or equal to that of users who leave. Unlike the static system, where $x = 1$ can be achieved for a sufficiently high p, full cooperation cannot be achieved in dynamic scenarios because the threat does not affect users who intend to leave the system at the end of the current period.

The user's contribution cost in each period is determined by the ratio between the fraction of users who get the full benefit of the system and those who get the reduced benefit. In what follows, we compare the system performance in two cases.

- Case 1) Uunder PI, it is feasible to penalize only the free-riders. Therefore, EF receive reduced service, but all other users receive full service.
- Case 2) Under free-identities, in order to penalize the free-riders, it is necessary to penalize all the newcomers.

The table shown at the bottom of the next page presents the fraction of users who get the full and reduced benefit under the two scenarios.

Based on this table, the contribution cost under PI, when newcomers are not penalized is

$$R_{\text{PI}} = \frac{(1-d)x + d + (1-d)(1-x)(1-p)}{x}$$

[1]The initial identity cost is considered to be a sunk cost. Identity cost refers to the cost of acquiring any additional identity after the first.

TABLE I
SIZE AND REALIZED PERFORMANCE LEVEL OF THE DIFFERENT GROUPS UNDER THE PI AND THE FI SCENARIOS

Group (j)	Group Size (f_j)	Realized Performance (W_j)	
		Permanent identities	Free identities
EC	$(1-d)x$	$Q - R_{PI}$	$Q - R_{FI}$
EF / WW	$(1-d)(1-x)$	$Q(1-p)$	$Q(1-p)$
NC	dx	$Q - R_{PI}$	$Q(1-p) - R_{FI}$
NF	$d(1-x)$	Q	$Q(1-p)$

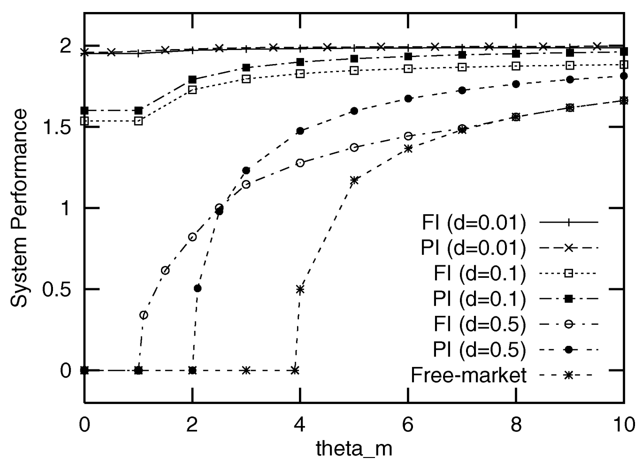


Fig. 6. System performance W when the system is subject to penalty p^*, as a function of θ_m, for different turnover rates d, when identities are either permanent (PI) or free (FI). The FM baseline has a penalty $p = 0$. The value of α is 3.

and the contribution cost under FI, when newcomers are penalized is

$$R_{\rm FI} = \frac{(1-d)x + d(1-p) + (1-d)(1-x)(1-p)}{x}$$

under FI, imposing a penalty always results in reduced performance, because all newcomers are penalized independent of their behavior. In contrast, setting p sufficiently high under PI can lead to a situation in which no user is penalized in effect (similar to the static system scenario, see Section IV).

Table I presents the fraction in the population and the realized performance level of each group under the two scenarios. System performance is given by:

$$W_s(p) = \sum_j (f_j * W_j(p)).$$

Fig. 6 presents the system performance subject to a penalty $p = \mathrm{argmax}_p\, W_s(p)$ (which was obtained numerically) under PI (with no penalty to newcomers) and FI (with penalty to newcomers) as a function of θ_m for different turnover rates (d). We make the following observations.

- If the turnover rate is low $(d = 0.01)$, both the fraction of newcomers and the fraction of users who leave the system are small. This means that only a few newcomers exist, thus penalizing newcomers does not significantly affect system performance. In addition, because the population is fairly permanent, a low p imposes a sufficient threat to obtain many contributions. In this case, the system performs close to its optimal level and no notable performance gap exists between the PI and FI scenarios.
- If the turnover rate is high $(d = 0.5)$ and the societal generosity is low $(\theta_m <\sim 2)$, system collapse can only be avoided by reducing the demand placed on the system. Assessing a penalty on all newcomers is one method to limit the demand. In these situations, penalizing newcomers actually helps to sustain the system by reducing system overload. Therefore, no social loss is incurred due to penalty to newcomers.
- If the turnover rate is high $(d = 0.5)$ and the societal generosity is intermediate $(\sim 2 < \theta_m <\sim 10)$, the imposition of penalty on all newcomers in order to discourage whitewashing incurs a social loss. If penalty can be imposed only on free-riders, as in the PI scenario, a higher system performance can be obtained.
- If the societal generosity is high, a high contribution level is obtained even in the absence of intervention. Therefore, the best policy under both scenarios is to impose a small penalty or no penalty at all. Hence, no notable social loss is incurred due to FI.
- If the benefits of the system (α) are high, even a small p results in a high threat to free-riders. Once again, the optimal p is small, and so no notable gap occurs.

In summary, a notable social loss due to FI is incurred only when a penalty on all newcomers is unnecessarily imposed. That is, in scenarios where the system can otherwise (under PI) tolerate the newcomers. In particular, the cost is incurred only under high turnover rates (d) and only in conjunction with intermediate contribution levels (θ_m) and low system benefits (α).

While we considered only the two extreme cases in which the identity cost is either zero or infinity, identity cost can take any positive finite value, and users decide whether to whitewash depending on how the identity cost compares to the penalty imposed on free-riders and newcomers. In these cases, it is possible to set the penalty of newcomers lower than that of free-riders, while still preventing whitewashing behavior. In particular, setting $p_{\rm NC}$ and $p_{\rm FR}$ (the penalty levels of newcomers and free-riders, respectively) such that $W_{\rm NC} = W_{\rm FR} - c$, where c is the identity cost, is sufficient to prevent whitewashing. This implies that the social loss due to positive-cost identities can actually be lower than that with zero-cost identities.

	PI : NC not penalized	FI : NC penalized
penalized	$(1-d)(1-x)$	$d + (1-d)(1-x)$
not penalized	$(1-d)x + d$	$(1-d)x$

VI. Fairness

A. Motivation

Throughout this paper, we have used the *system performance* (W_S) metric to quantify the effect of free-riding and evaluate the performance of the incentive mechanisms. This metric measures the total performance in the population, assigning equal weight to the realized performance of each individual in the population, regardless of his behavior.

However, one may have reservations about this metric on grounds of fairness, claiming that we should not care about the performance realized by free-riders as much as we care about the performance realized by contributors. For example, under the exclusion mechanism (see Section III), we might leave some free-riders in the system for the sake of maximizing *system* performance, while such a decision improves the benefits of the free-riders at the expense of contributors. We might wish to maximize the total (or average) performance realized by contributors rather than that realized by the society as a whole.

While W_S denotes the total system performance, W_C denotes the total performance of contributors in the system. In the FM

$$W_C = x\left(\alpha x - \frac{1}{x}\right) = \alpha x^2 - 1.$$

B. Contributor Performance in the Exclusion Mechanism

Under the exclusion mechanism, the overall system performance is $W_S = (\alpha x - 1)(1 - z)$, and the contributor performance is given by $W_C = x(\alpha x - (1 - z/x))$.

Claim 3: The exclusion levels that optimize the total system performance and the contributor performance for different values of θ_m is given in the following table (expressions of z', θ'_m and θ''_m as in Section III).

	W_S	W_C
$\theta_m < 2$	$z^* = 1 - \frac{\theta_m}{4}$	$z^* = 1 - \frac{\theta_m}{4}$
$2 < \theta_m < \theta'_m$	$z^* = \frac{1}{\theta_m}$	$z^* = \frac{1}{\theta_m}$
$\theta'_m < \theta_m < \theta''_m$	$z^* = z'$	$z^* = \frac{1}{\theta_m}$
$\theta_m > \theta''_m$	$z^* = 0$	$z^* = \frac{1}{\theta_m}$

The proof can be found in the appendix.

Looking at the table and Fig. 7, we can divide the results into three regions (expressions of θ'_m and θ''_m as in Section III).

- If $\theta_m < \theta'_m$, $z^*_c = z^*_s$. In this case, $x = 1 - z$, meaning that *all* of the remaining users contribute. When the societal generosity is low, an exclusion level that results in a system with no free-riders is optimal from the perspectives of both system and contributor performance.
- If $\theta'_m < \theta_m < \theta''_m$, $z^*_s < z^*_c$. System performance is maximized for an exclusion level that results in a mixture of contributors and free-riders. If we impose $z = z^*_s$, contributor performance *decreases* in θ_m, and may incur a significant loss [compared with $W_C(z^*_c)$].
- If $\theta''_m < \theta_m$, $0 = z^*_{W_S} < z^*_{W_S}$. From the perspective of system performance, it is optimal to exclude no users, which entails some loss to contributors. However, as θ_m gets higher, $z^*_{W_C}$ approaches 0 as well, and the gap between $W_C(z^*_{W_C})$ and $W_C(z^*_{W_S})$ shrinks.

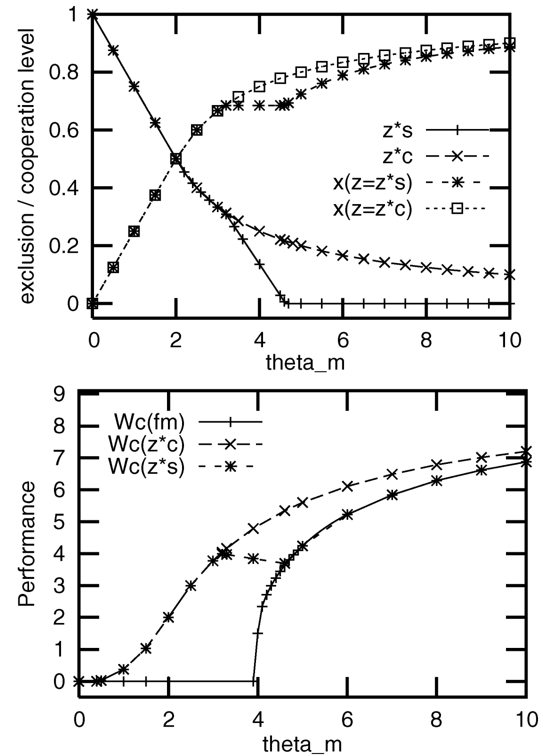


Fig. 7. (a) Optimal exclusion level z^* and corresponding contribution level x when maximizing system performance W_S or contributor performance W_C. (b) Contributor performance W_C realized for different exclusion level (FM) baseline where $z = 0$. The value of α is 10.

C. Contributor Performance in the Penalty Mechanism

We found that under the exclusion mechanism, $z^*_c \geq z^*_s$. That is, if we care more about the performance realized by contributors, we may wish to exclude a higher fraction of users. What about the penalty mechanism? Recall from Section IV that $p = 1$ maximizes system performance. When $p = 1$, all users contribute, and therefore it also maximizes W_C. We get: $p^*_c = p^*_s = 1$ (where p^*_c denote the penalty level that maximizes contributor performance). This statement is true in the static model (see Section IV), in which a high enough penalty leads to 100% cooperation. Conversely, in the dynamic model, where users join and leave and newcomers are penalized to deter whitewashing (see Section V), there is a tradeoff. While penalty increases contribution level, it also decreases the benefit of *new* contributors. Therefore, especially in cases of high turnover, a high p may result in lower contributor performance. Yet, similar to our findings under the exclusion mechanism, our numerical results confirm that $p^*_c \geq p^*_s$.

Fig. 8 presents the results of a dynamic system with turnover rate of 0.5. It shows $W_C(p^*_s)$ and $W_C(p^*_c)$ under FI and PI. Under PI, the penalty is not imposed on newcomers nor departing users. It is only imposed on free-riders who stay in the system. A sufficiently high penalty level achieves 100% cooperation in that group, and therefore, there is no gap between $W_C(p^*_s)$ and $W_C(p^*_c)$. However, with FI, we observe a similar pattern to the results obtained under the exclusion mechanism.

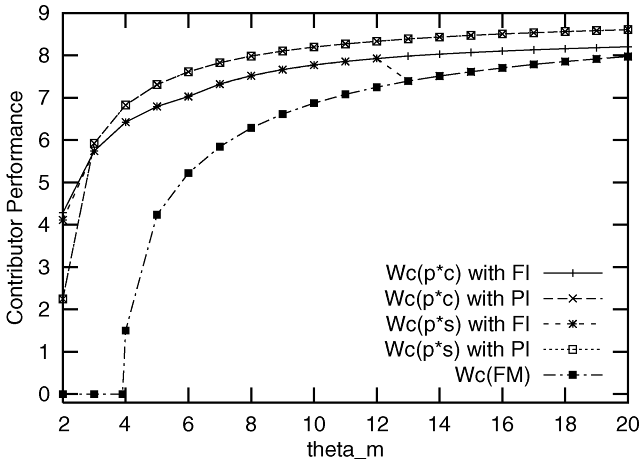


Fig. 8. Contributor performance W_C subject to $p = p_c^*$ and $p = p_s^*$ under FI, PI, and FM, where $p = 0$. Turnover rate $d = 0.5$, $\alpha = 3$.

More specifically, the gap between $W_C(p_s^*)$ and $W_C(p_c^*)$ is significant for a limited range of θ_m values. When considering the performance of the system as a whole, we may impose a lower penalty level at the expense of contributors.

VII. Conclusion

By constructing a simple economic model of user behavior, we gain some insight into free-riding and whitewashing in P2P systems. First, we find that when societal generosity level is low, a mechanism that penalizes free-riders can improve system performance by simultaneously imposing a threat to potential free-riders and reducing the cost of contribution. Second, if identities are freely available, penalizing all newcomers can be highly effective in discouraging whitewashing behavior, and incurs a social loss only for a limited range of societal generosity levels, and only in conjunction with a high turnover rate.

The focus of this work is not to propose a new incentive scheme, but rather to understand the effect of incentives on user behavior and system performance. Therefore, we have strived to keep the model as simple as possible. However, the model retains a flexibility that allows us to account for a diverse set of system characteristics, and can be easily extended in several directions.

First, the model can be extended to analyze scenarios of resource heterogeneity. A preliminary analysis shows that the effect of heterogeneity depends to a large extent on the correlation between a user's resource capacity and her generosity level. To model resource heterogeneity, each user can be split into a number of virtual users, proportional to the amount of resources she has. In this context, one can experiment with different contribution cost functions that may be more reflective of the opportunity cost of contribution. For example, resource-rich users may experience a lower opportunity cost per unit of contribution, even when they contribute more resources [17]. Users can also choose different levels of contribution, rather than the binary choice of being a contributor or a free-rider.

Second, in modeling dynamic behavior, the model can be extended to allow: 1) departure rates that depend on system performance; 2) arrival rates that are affected by the imposed penalty level; and 3) dynamics affecting societal generosity by postulating type-dependent arrivals and departures. In particular, one could imagine that the choice of imposed penalty p would affect arrival rate in different directions. On the one hand, imposing penalties on newcomers may discourage them from joining the system, thereby reducing the arrival rate. On the other hand, contributors may find such a system attractive to join because whitewashers are discouraged, and therefore the system performance level may actually be higher. Depending on how the penalty level p affects the arrival rate, the performance impact of FI may increase or decrease relative to our results. The challenge here is to model these effects, while preserving the simplicity of the model.

We believe our abstract modeling of exclusion and penalty mechanisms captures the essence of many real-world P2P incentive mechanisms such as reputation-based service differentiation and entrance fees, and therefore our findings may find some applicability in the understanding of these systems, particularly when they are threatened by whitewashers. At the same time, our model also suggests other possible forms of incentive engineering in P2P systems that are yet to be explored, e.g., explicit system partitioning, that may also turn out to be effective against free-riders and whitewashers.

Appendix I
Proofs of Claims 2 and 3

Proof of Claim 2: If $\theta_m < 2$, $1 - z < x_1$ for all $z \in [1 - \theta_m/4, 1]$, which means that $x = 1 - z$, which decreases with z. In this case, it is optimal to impose the minimum possible z, which is $z_s^* = 1 - \theta_m/4$. $x = 1 - z$ indicates that all of the users that are not excluded contribute.

if $\theta_m \geq 2$:

- For $z \geq 1/\theta_m$, $x = 1 - z$, which yields
$$W_S = (\alpha(1-z) - 1)(1-z).$$
It is straightforward to verify that this function has no maximum in $[1/\theta_m, 1]$. Therefore, the optimal z in this region is
$$z_s^* = \frac{1}{\theta_m}.$$
- For $z \leq 1/\theta_m$, $x = x_1$, which yields
$$W_S = \left(\alpha \frac{\theta_m + \sqrt{\theta_m^2 - 4\theta_m + 4\theta_m z}}{2\theta_m} - 1\right)(1-z).$$
By equating the first derivative to 0, we get the optimal z in this region shown in the first equation at the top of the next page.

However, there are two exceptions to the above result.

If it is greater than $1/\theta_m$, then $z_s^* = 1/\theta_m$. This happens for $\theta_m < \theta_m'$, shown in the second equation at the top of the next page.

For example, if $\alpha = 10$, $\theta_m' \sim= 3.2$. In addition, if it is smaller than or equal to 0, then $z_s^* = 0$. This happens for $\theta_m > \theta_m''$, where
$$\theta_m'' = \frac{9\alpha^2}{\alpha^2 + 2\alpha + \alpha\sqrt{\alpha^2 - \alpha + 1} - 2 - 2\sqrt{\alpha^2 - \alpha + 1}}.$$
For example, if $\alpha = 10$, $\theta_m'' \sim= 4.6$.

$$z_s^* = \frac{9\alpha^2 - \theta_m\alpha^2 - \alpha\theta_m(2+\sqrt{\alpha^2-\alpha+1}) + 2\theta_m(1+\sqrt{\alpha^2-\alpha+1})}{9\alpha^2}$$

$$\theta'_m = \frac{3\alpha(3\alpha + \sqrt{5\alpha^2 - 8\alpha - 4\alpha\sqrt{\alpha^2-\alpha+1} + 8 + 8\sqrt{\alpha^2-\alpha+1}})}{2(\alpha^2 + 2\alpha + \alpha\sqrt{\alpha^2-\alpha+1} - 2 - 2\sqrt{\alpha^2-\alpha+1})}$$

Proof of Claim 3: Let z_c^* denote the exclusion level that maximizes total contributor performance

$$z_c^* = \mathrm{argmax}_z(W_c).$$

If $\theta_m < 2$, $z_c^* = z_s^* = 1 - \theta_m/4$, as explained above. The following analysis holds for $\theta_m \geq 2$.

- For $z \in [0, (1/\theta_m))$: $x = \theta_m + \sqrt{\theta_m^2 - 4\theta_m + 4\theta_m z}/2\theta_m$, which increases in z. Since W_C increases in both z and x

$$z_c^* = \frac{1}{\theta_m}. \tag{8}$$

- For $z \in [(1/\theta_m), 1]$: $x = 1 - z$, and by simple calculus, we get

$$z_c^* = \frac{1}{\theta_m}. \tag{9}$$

From (8) and (9), we get that for all $\theta_m \geq 2$

$$z_c^* = \frac{1}{\theta_m}.$$

That is, to maximize W_C, it is always optimal to reach an exclusion level that results in a system with no free-riders. Decreasing z below this point may improve system performance, but necessarily degrades W_C.

Acknowledgment

The authors would like to thank the anonymous reviewers for their insightful comments and suggestions.

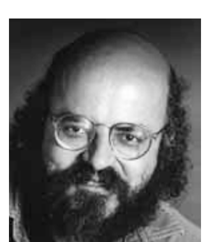
References

[1] E. Adar and B. A. Huberman, "Free riding on Gnutella," *First Monday*, vol. 5, no. 10, Oct. 2000. [Online]. Available: citeseer.nj.nec.com/adar00free.html.
[2] S. Saroiu, P. K. Gummadi, and S. D. Gribble, "A measurement study of peer-to-peer file sharing systems," in *Proc. Multimedia Computing and Netw.*, 2002. [Online]. Available: http://citeseer.nj.nec.com/article/saroiu01measurement.html.
[3] C. Buragohain, D. Agrawal, and S. Suri, "A game-theoretic framework for incentives in P2P systems," in *Proc. Int. Conf. Peer-to-Peer Comput.*, Sep. 2003, pp. 48–56.
[4] J. Crowcroft, R. Gibbens, F. Kelly, and S. Ostring, "Modeling incentives for collaboration in mobile ad hoc networks," in *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2003.
[5] P. Golle, K. Leyton-Brown, I. Mironov, and M. Lillibridge, "Incentives for sharing in peer-to-peer networks," in *Proc. 3rd ACM Conf. Electron. Commerce*, Oct. 2001. [Online]. Available: http://crypto.stanford.edu:/pgolle/papers/peer.pdf.
[6] P. Michiardi and R. Molva, "A game theoretical approach to evaluate cooperation enforcement mechanisms in mobile ad hoc networks," in *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2003.
[7] A. Urpi, M. Bonuccelli, and S. Giordano, "Modeling cooperation in mobile ad hoc networks: A formal description of selfishness," in *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2003.
[8] R. Krishnan, M. Smith, Z. Tang, and R. Telang, "The virtual commons: why free-riding can be tolerated in peer-to-peer networks," in *Proc. Workshop on Inf. Syst. Econ.*, Dec. 2003.
[9] C. F. Camerer, *Behavioral Game Theory*. Princeton, NJ: Princeton Univ. Press, 2003.
[10] J. Andreoni, "Giving with impure altruism: applications to charity and Ricardian equivalence," *J. Pol. Economy*, vol. 97, no. 6, pp. 1447–58, 1989.
[11] ——, "Impure altruism and donations to public goods: A theory of warm-glow," *Economic. J.*, vol. 100, no. 401, pp. 467–477, 1990.
[12] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The eigentrust algorithm for reputation management in P2P networks," in *Proc. 12th Int. World Wide Web Conf.*, May 2003.
[13] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity by image scoring," *Nature*, vol. 393, pp. 573–577, 1998. [Online]. Available: citeseer.nj.nec.com/nowak98evolution.html.
[14] E. Friedman and P. Resnick, "The social cost of cheap pseudonyms," *J. Econ. Manage. Strategy*, vol. 10, no. 2, pp. 173–199, 1998. [Online]. Available: citeseer.nj.nec.com/friedman98social.html.
[15] M. Castro, P. Druschel, A. Ganesh, A. Rowstron, and D. S. Wallach, "Security for structured peer-to-peer overlay networks," in *Proc. Multimedia Comput. Netw.*, 2002.
[16] M. Feldman, K. Lai, I. Stoica, and J. Chuang, "Robust incentive techniques for peer-to-peer networks," in *Proc. ACM Conf. Electron. Commerce*, 2004.

Michal Feldman received the Ph.D. degree from the School of Information Management and Systems, University of California at Berkeley in 2005.

She is a Postdoctoral Fellow at the School of Computer Science and Engineering, Hebrew University of Jerusalem. Her research interests lie in the intersection of game theory, microeconomic theory and communication networks.

Dr. Feldman was the recipient of the Lady Davis Fellowship (2005–2007).



Christos Papadimitriou received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ.

He is Professor in the Computer Science Division, University of California at Berkeley. He has taught at Harvard University, MIT, Athens Polytechnic, Stanford University, and University of California at San Diego prior to Berkeley. He is interested in the theory of algorithms and complexity, and its applications to databases, optimization, AI, networks, and game theory. He has written four textbooks in these areas, as well as a novel about computation.

John Chuang received the B.S. and M.S. degrees in electrical engineering from the University at Southern California, Los Angeles, and Stanford University, Stanford, CA, respectively, and the Ph.D. degree in engineering and public policy from Carnegie Mellon University, Pittsburgh, PA.

He is a Associate Professor of Information Management and Systems at the University of California, Berkeley. His research focus is on economics-informed design of computer networks and distributed systems.

Ion Stoica received the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, in 2000.

He is an Assistant Professor in the Department of Electrical Engineering and Computer Science, University of California at Berkeley, where he does research on resource management, scalable solutions for end-to-end quality-of-service, and peer-to-peer network technologies in the Internet.

Dr. Stoica is the recipient of a Presidential Early Career Award for Scientists and Engineering (PECASE) 2002, the Sloan Foundation Fellowship (2002), and the ACM Doctoral Dissertation Award (2001).