The Economics of Resisting Censorship

Early peer-to-peer systems sought to resist censorship by distributing content randomly over the entire Internet. The most popular ones simply let nodes serve the resources they were most interested in. The authors offer the first model inspired by economics and conflict theory to analyze such systems' security.

George Danezis and Ross Anderson
University of Cambridge

Peer-to-peer designs have evolved in part as a response to the technical censorship of early remailer systems such as Penet (see http://wfmu.org/~davem/docs/penet.html) and early file distribution systems such as Napster.[1] Such centralized architectures make it possible to legally compel system owners to reveal user identities and suppress certain kinds of material. Peer-to-peer systems, however, distribute functionality across the network, thus avoiding single points of failure that could make them vulnerable to legal or technical attacks.

Two main paradigms in peer-to-peer systems have emerged during the past few years. The first is to scatter resources randomly across all nodes, hoping that doing so will increase the attacker's censorship costs (we call this the *random model*). The theory holds that censorship inconveniences everyone in the network—even nodes that aren't interested in the censored material—which increases the number of nodes that must be attacked for censorship to succeed. Eternity Service, Freenet, and Mojo Nation follow this strategy;[2–4] structured peer-to-peer systems, including distributed hash-table-based systems,[5] scatter files around in a deterministic way on random nodes, which achieves a similar effect.[6]

The second paradigm lets peer nodes serve any content that users have downloaded for personal use without burdening them with random files (we call this the *discretionary model*). Gnutella and Kazaa are popular real-world examples of such systems.[7] Newer designs incorporate distributed information retrieval techniques in an attempt to assist users in finding what they're looking for.[8]

Any comparison of these two paradigms necessarily concentrates on system and network engineering efficiency in terms of the cost of search, retrieval, communications, and storage. In this article, however, we'll compare the two paradigms' ability to resist censorship, which was the original intention of peer-to-peer systems. Our model of censorship and censorship resistance is inspired by conflict theory and economic analysis. It takes into account the peer nodes' heterogeneous interests and establishes the cost of the attack.[9]

A red–blue utility model

Before talking about censorship, we first need to define the preferences of the nodes in the network. Let's consider a network of N peer nodes. Each node n_i has a different set of interests from other nodes: it might prefer news articles to political philosophy essays, for example, or nuclear physics to cryptology. Nodes might even have different political views from each other. We model this by considering two types of resources: red and blue. (Despite real-life preferences' finer granularity, we follow the economics tradition of considering only two goods. Our results also apply to n goods.) We assign to each node n_i a preference for red $r_i \in [0, 1]$ and a preference for blue $b_i = 1 - r_i$ (note that $r_i + b_i = 1$).

Each node likes having and serving resources, but it prefers to have or serve a balance of resources according to its preferences r_i and b_i. For this reason, the utility function (which represents satisfaction) of a node holding T resources, out of which R are red and B are blue (where $T = R + B$), is

$$U_i(R, B) = -T(R/T - r_i - 1)(R/T - r_i + 1). \quad (1)$$

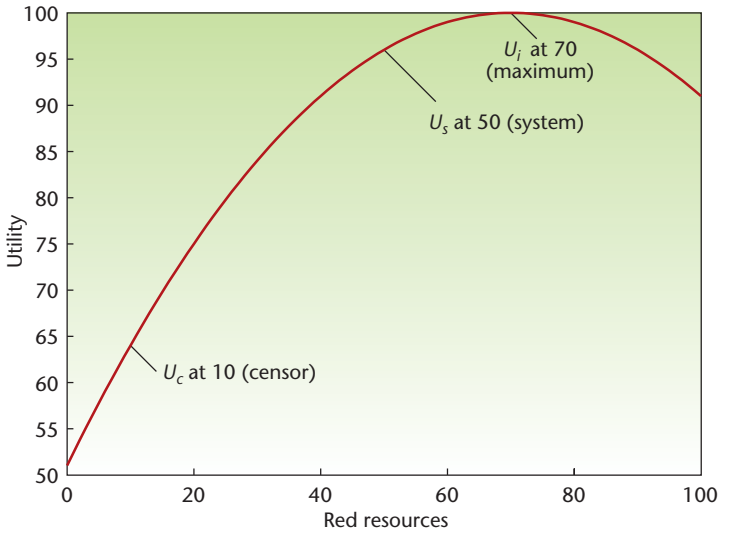

Figure 1. Utility model for discretionary, random, and censored distribution. The curve shows the utility of a node with $r_i = 0.7$ in a system following the discretionary model (U_i), the random model with $r_s = 0.5$ (U_s), and under censorship $r_c = 0.1$ (U_c).

As we see in Figure 1, this quadratic function has its maximum at $R = r_i T$, scaled by the overall number of resources T that node n_i holds. This utility function increases with the total number of resources, but it's also maximal when the balance between red and blue resources matches the node's preferences. Other unimodal functions with their maxima at $r_i T$, such as a normal distribution, give broadly similar results.

Our model diverges from traditional economic analysis of peer-to-peer networks, in which peers have no a priori incentives to share.[10] Although this assumption might be true for copied music, it doesn't hold for other resources such as news, political opinions, or scientific papers. A node with left-wing views, for example, might prefer to read and redistribute 80 percent of the articles from *The Guardian* and 20 percent from *The Telegraph* ($r_i = 0.8$, $b_i = 0.2$), one in the middle of the political spectrum might want to read and redistribute them equally ($r_i = 0.5$, $b_i = 0.5$), and one on the right might prefer 80 percent from *The Telegraph* and 20 percent from *The Guardian* ($r_i = 0.2$, $b_i = 0.8$). The nodes' respective utility will increase the more they can distribute this material in volumes that align with their political preferences.

Discretionary and random distribution utility

Let's examine the utility of network nodes when they can choose which files to store and help serve. Assuming that a node has the ability to serve T files in total, its utility is maximized for a distribution of red and blue resources that perfectly matches its preferences: $R = r_i T$ and $B = b_i T$. Our proposed utility function U_i is indeed maximal for $U_i(r_i T, b_i T)$ at each node n_i.

Distributed hash tables and architectures such as those in Eternity scatter the red and blue resources randomly across all nodes n_i, so what is each node's average or expected utility? If the system has a total of $\mathcal{R}$ red resources and $\mathcal{B}$ blue resources, we can define a system-wide distribution of resources (r_s, b_s) that each node in the system will hold on average:

$$r_s = \frac{\mathcal{R}}{\mathcal{R}+\mathcal{B}} \quad b_s = \frac{\mathcal{B}}{\mathcal{R}+\mathcal{B}}. \tag{2}$$

On average, each node n_i will have a utility equal to $U(r_s T, b_s T)$. The utility each node will attain in the random case is always lower than or equal to the utility a node has under the discretionary model:

$$U_i(r_i T, b_i T) \geq U_i(r_s T, b_s T). \tag{3}$$

For this reason, we prefer the discretionary peer-to-peer paradigm, given the choice and absence of other mechanisms.

Let's explore in more detail the implications of the lower utility provided by the random distribution model. The equality $U_i(r_i T, b_i T) = U_i(r_s T, b_s T)$ is true when $r_s = r_i$ and $b_s = b_i$—in other words, when the system's distribution of resources aligns with a particular node's preferences. However, this can't hold true for all nodes unless they share the same preferences. Moreover, it's in every node's self-interest to try to tip the balance of $\mathcal{R}$ and $\mathcal{B}$ toward its own preferences. With a utility function slightly more biased toward serving—sometimes called an "evangelism utility"—the network could be flooded with red or blue files according to preferences.

An alternative is subversion. Red-loving nodes that consider the network overly biased toward blue could just as easily try to deny service of blue files—in the extreme case, they could even try to deny general service. Distributed hash tables and systems such as Freenet are quite prone to flooding, of both the evangelism and denial-of-service varieties.

Several other systems, including Free Haven,[11] Mojo Nation, and Eternity, recognize that where the utility function places more value on consumption than service, nodes have an incentive to take a free ride by downloading as many resources as they can. Eternity proposed—and Mojo Nation tried to implement—a payment mechanism to align the incentives for storing and serving files. By performing these functions, nodes acquire *mojo*, a notional currency that lets them get service from other nodes. Due to implementation failures, poor modeling, and inflation, however, Mojo Nation provided substandard service and didn't take off.[4] Free Haven used a reputation system, in which peers could rate the quality of service other nodes provide and then

prioritize their service to good providers. Perhaps we could use such a system to rate nodes: some collaborative mechanism could establish r_s and b_s via voting and then rate peers in accordance with their closeness to this social norm. This isn't trivial; voting theory, also known as social choice theory, tells us that it's hard to create a voting system that is both efficient and equitable.[12,13] The additional constraints of peer-to-peer networks—nodes frequently joining and leaving, transient identities, and decentralization—make a "democratic" system even more complex to implement in practice.

Some systems attempt to hide their stored or served resources from nodes via encryption or dispersion. This is thought to protect the nodes by giving them plausible deniability in the face of censorship, but it also prevents the nodes from deleting any resources they don't like. In our framework, such techniques amount to hiding from the nodes the actual distribution of red and blue resources they hold, and can even go as far as hiding the system's overall distribution r_s and b_s. Unfortunately, hiding this information makes these systems very expensive. The effects of the participating nodes' state of uncertainty on their incentives to participate honestly in such a network should be the subject of further study.

Censorship

So far, we've compared the utility of nodes in the random versus discretionary models, and learned that the latter always provides as good or higher utility for all nodes in the absence of censorship. Let's now examine how the nodes react to censorship.

We assume the attackers are exogenous—that is, external to our system. We model censorship as an external entity's attempt to impose a particular distribution of files r_c, b_c on a set of nodes. The censor's effect isn't fixed; rather, it depends on the amount of resistance the affected nodes offer.

Assume a node that isn't receiving attention from the censor can store up to T resources. A node under censorship can choose to store fewer resources $(T - t)$ and invest an effort level t to resist censorship. We define the probability that a node will successfully fight censorship (and reestablish its previous distribution of resources) as $P(t)$. With probability $1 - P(t)$, the censor will prevail and impose the distribution r_c, b_c.

Let's first consider the discretionary case, in which nodes select the content they serve. Knowing the nodes' preferences r_i, b_i, the censor's distribution r_c, b_c, the total resource bound T, and the probability $P(t)$ that it defeats the censor, we can calculate the optimal amount of resources a node will invest in resisting censorship. The expected utility of a node under censorship is the probability of success, times the utility in that case, plus the probability of failure times the utility in that case:

$$U = P(t)U_i(r_i(T-t),\ b_i(T-t)) + (1-P(t))U_i(r_c(T-t),\ b_c(T-t)). \quad (4)$$

Our utility functions U_i are unimodal and smooth, so, assuming that the functions $P(t)$ are sufficiently well-behaved, we can find an optimal investment in resistance t in $[0, T]$ by setting $dU/dt = 0$.

We'll start with the simplest example—namely, where the probability $P(t)$ of resisting censorship is linear in the defense budget t. Assume that if a node invests all its resources in defense, it will definitely prevail but won't have anything left with which to actually serve files. At the other extreme, if the node spends nothing on lawyers (or any other relevant mode of combat), the censor definitely prevails. Therefore, we define $P(t)$ as

$$P(t) = \frac{t}{T}. \quad (5)$$

By maximizing Equation 4 with $P(t)$ as defined in Equation 5, we find that the optimal defense budget t_d will be

$$t_d = \frac{T}{2}\frac{2U_i(r_c,b_c) - U_i(r_i,b_i)}{U_i(r_c,b_c) - U_i(r_i,b_i)}. \quad (6)$$

The node diverts t_d resources from the file service to resist censorship. We also assume, for now, that the attack's cost to the censor is equal to the node's defense budget t.

Let's look at what happens when we scatter resources randomly around the network, with each node expecting to hold a mixture of files r_s, b_s. As in the previous example, the utility of a node under censorship depends on its defense budget t, the censor's choice of r_c, b_c, and the system's distribution of files r_s, b_s:

$$U = P(t)U_i(r_s(T-t),\ b_s(T-t)) + (1-P(t))U_i(r_c(T-t),\ b_c(T-t)) \quad (7)$$

We follow an approach similar to the one in Equation 6 to derive each node's optimal defense budget t:

$$t_s = \frac{T}{2}\frac{2U_i(r_c,b_c) - U_i(r_s,b_s)}{U_i(r_c,b_c) - U_i(r_s,b_s)}. \quad (8)$$

Not all nodes are motivated to resist the censor—some will find that $U_i(r_sT,\ b_sT) \leq U_i(r_cT,\ b_cT)$, which means their utility under censorship increases. This isn't an improbable situation: in a network in which half the resources are red and half are blue ($r_s = 0.5$, $b_s = 0.5$), a censor that shifts the balance to $r_c = 0$ will benefit the blue-loving nodes; if they're free to set their own defense budgets, they'll select $t = 0$, which means no resistance.

Who fights censorship harder?

We derive the defense budgets of a node in a discretionary and in a random network as t_d and t_s, respectively (see Figure 2). They also equal the censor's costs in the

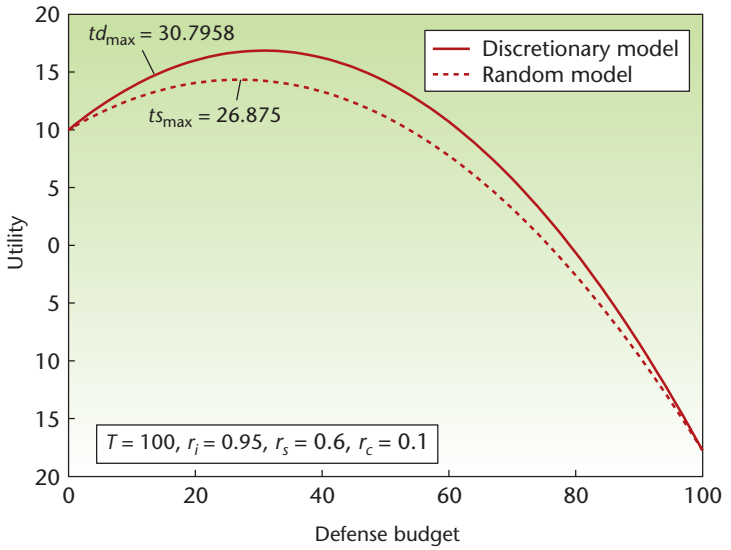

Figure 2. Defense budgets. The defense budget for the discretionary model (t_d) is larger than random (t_s), where r_i represents the node's preferences, r_s the random model parameter, and r_c the censor's imposed distribution.

two types of network. The aggregate defense budget, and thus the cost of censorship, is greater in the discretionary model than in the random one, except in the case in which all nodes have the same preferences (case equality holds).

For the maximum value of the defense budget t to be positive in the interval $[0, T]$, the following condition must be true:

$$0 < \frac{T}{2}\frac{2U_i(r_c,b_c) - U_i(r_s,b_s)}{U_i(r_c,b_c) - U_i(r_s,b_s)}. \quad (9)$$

In other words,

$$2U_i(r_c, b_c) < U_i(r_s, b_s). \quad (10)$$

When this is not true, a node maximizes its utility by not fighting at all and choosing $t = 0$ (see Figure 3).

Given these observations, it follows that

$$\forall i \in S,\ t_d \geq t_s \Rightarrow \sum_{i \in S} t_d \geq \sum_{i \in S} t_s\ . \quad (11)$$

Whatever the attacker's strategy, it is at least as costly or more so, to attack a network's architecture via the discretionary rather than the random model. Equality holds when $t_d = t_s$ for each node, which in turn means that $r_i = r_s$. This is the case for homogeneous preferences, but in all other cases, the cost to censor a set of nodes is maximized when resources are distributed according to preferences rather than randomly.

Discussion

The model of heterogeneous preferences and censorship that we present here is very simple, but it still gives some important initial insights into the economics of censorship resistance. Censorship is an economic activity; the censor faces costs, regardless of whether a particular kind of material is repressed via criminal or civil law, or military force. Furthermore, the target's defense expenditures (for lawyers, lobbying, technical protection measures, or even armed conflict) can diminish the censor's prospect of success.

Most research on censorship resistance views censorship as a binary matter: a document is either proscribed by a court or it isn't, a technical system is either vulnerable to attack or it isn't. Such models are as unrealistic as the global adversary sometimes posited in cryptography (an opponent can record or modify all the messages on network links). All-powerful opponents make censorship uninteresting as a technical issue because resistance would be impossible. Similarly, assuming that nobody can censor any nodes provides little intuition into real systems.

Technology changes greatly affect parameters. The introduction of moveable type printing, for example, made it much harder to suppress books thought to be heretical or seditious—a change in the underlying economics that helped usher in the modern age. However, developments such as online publishing and trusted computing might make censorship easier again. We can only guess the effects such changes will bring, which is why trying to analyze the cost of both censorship and resistance to it is important.

Preferences and utility

Modeling a node's preferences provides important insights. Simply assuming that all nodes will fight censorship for an abstract notion of "freedom of speech" restricts the model to a fraction of potential real-world users. Think of the online tussles between Scientologists and people critical of their organization compared to the sexual material that is legal under California law but illegal in Tennessee. The average Scientology critic might not care that much about sexual freedom, while a collector of erotic literature might be indifferent to religious disputes. Although some individuals would take a stand for freedom of speech on a broad range of issues, many more are prepared to defend it on specific issues.

On the other hand, assuming that nodes will meekly surrender any disputed documents or photographs is also unfaithful to real-world experience. Allowing nodes to express heterogeneous interests in, and preferences for, material they want to promote and protect helps us enhance the system's stability and security. It also enables us to defend against certain types of denial-of-service attacks. When Eternity was initially implemented and opened for public use, one of the first documents placed

in it (by an anonymous poster) advocated sex between men and underage boys. Although some people defended such speech, many felt reluctant to use a system that expressed it. A discretionary peer-to-peer system can deal with such issues, much as ISPs currently decide which Usenet newsgroups to support depending on local laws and client preferences. Objectionable content need not provide a universal attack tool.

Our model provides a framework for thinking about such issues. In particular, we've found that in the presence of heterogeneous preferences, systems that distribute material randomly across all nodes are less efficient at resisting censorship than those that allow storage according to node preferences. As this inefficiency increases with heterogeneity, we expect random distribution to be more successful in groups with roughly homogeneous interests. When interests diverge, systems should let users choose their resources, or the networks will tend to be unstable. Nodes will prefer to form alternative networks that better match their preferences.

Our model can be extended in several ways. We use red and blue resources as a simple example, but nodes must be able to express arbitrary and much finer-grained preferences; accordingly, the results we present can be generalized to unimodal multidimensional utility functions. We chose to model node utility locally and didn't take into account a resource's "global" availability throughout the network. Forming a global view of availability is hard in many peer-to-peer systems because nobody has a total view of network membership and the state of all its nodes. We also ignore the costs associated with search and retrieval. Some systems, such as distributed hash tables, allow very efficient retrieval but at a high search cost; other systems are more balanced or less efficient overall. Our model is simple and flexible enough to be extended to describe various specific attacks on peer-to-peer systems.

Random distribution can introduce social choice problems that discretionary distribution avoids. It describes explicit mechanisms to determine r_s and b_s, the relative number of red versus blue files that a typical node is asked to serve, on average. Nodes have incentives to shift these system parameters toward their own preferences, so they might be tempted to manipulate the voting or reputation systems in use. Making these systems robust is a separate topic of research.

Censorship model

We carefully chose our model to avoid introducing additional social choice issues. The censor targets a set of individual nodes, with the success of the attack depending only on the targeted nodes' defense budgets. Of course, where nodes are subject to legal action, a victory against one might create a precedent that makes enforcement against other nodes cheaper in the future. Defense could thus take on some of the aspects of a public good. This approach's ultimate success will depend on whether the level of defense depends on the least effort, greatest effort, or the sum of all nodes' efforts.[14]

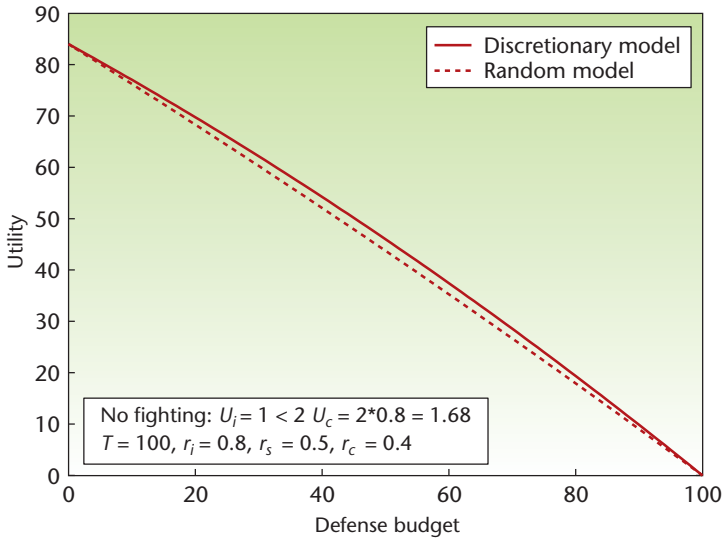

Figure 3. Zero defense budget. Nodes won't fight censorship unless their utility is halved. Instead, they will invest in different resources.

Our model also assumes that a censor wishes to impose a certain selection of resources on nodes, which might be appropriate when modeling censorship of the press, but maybe not for online music. The music industry's strategy in such cases is to increase the cost of censorship resistance to match the music's retail price. In this context, our model suggests that it would be much harder for the industry to take on a diffuse constellation of autonomous fan clubs than it would be to take on a monolithic file-sharing system. Some performers might be unwilling to alienate their fans by overly aggressive enforcement.

Our particular censorship model provides some further insights. For both random and discretionary distribution, the censor meets resistance from a node once its activities halve its utility because a node reacts to mild censorship by investing in other resources instead of engaging in combat. Mild censorship might provoke a small reaction, but at some point, nodes will start to fight back, starting with those nodes whose preferences differ the most from the censor's.

Under our model's assumptions, discretionary distribution is better. The more heterogeneous the nodes' preferences, the more it outperforms random distribution. In a discretionary model, nodes don't have to collectively manage the network's overall content, which

gives them fewer incentives to subvert control mechanisms, which in turn allows for simpler network designs that don't require election schemes, reputation systems, or electronic cash. The discretionary model also leads to a more stable network because each node can better maximize its utility and is less likely to leave the network to seek a better deal elsewhere.

Current debate also centers on whether increasing social diversity will necessarily undermine social solidarity.

In our work, we haven't modeled the censor's incentives or tried to find its optimal strategy in attacking a network. Better attack models will require more detail about network architecture and operation. Ultimately, we feel our model might have wider implications. Rather than fighting against government regulation and for market freedom in the abstract, firms are more likely to invest effort (through trade associations) in fighting for the freedoms most relevant to their own particular trades. Current debate also centers on whether increasing social diversity will necessarily undermine social solidarity.[15,16] Our model's relevance to such issues is a matter for discussion elsewhere. □

Acknowledgments

We thank Rupert Gatti and Thierry Rayna for their valuable input on the appropriateness of different utility functions to the problem studied and the anonymous reviewers for their valuable comments. This work was done with the financial assistance of the Cambridge MIT Institute (CMI) as part of a project on the design and implementation of third-generation peer-to-peer systems.

References

1. B. Carlsson and R. Gustavsson, "The Rise and Fall of Napster: An Evolutionary Approach," *Active Media Technology*, LNCS 2252, J. Liu et al., eds., Springer-Verlag, 2001, pp. 347–354.
2. R. Anderson, "The Eternity Service," *Proc. 1st Int'l Conf. Theory and Applications of Cryptology* (Pragocrypt 96), Czech Tech. Univ. Publishing House, 1996, pp. 242–252.
3. I. Clarke et al., "Freenet: A Distributed Anonymous Information Storage and Retrieval System," *Int'l Workshop on Design Issues in Anonymity and Unobservability*, LNCS 2009, H. Federrath, ed., Springer-Verlag, 2001, pp. 46–66.
4. B. Wilcox-O'Hearn, "Experiences Deploying a Large-Scale Emergent Network," *First Int'l Workshop* (IPTPS 02), LNCS 2429, Springer, 2002, pp. 104–110.
5. I. Stoica et al., "Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications," *ACM SIGCOMM Conf. Applications, Technologies, Architectures, and Protocols for Computer Comm.*, ACM Press, 2001, pp. 149–160.
6. W.W. Weatherspoon, C. Wells, and B.Y. Zhao, "Oceanstore: An Architecture for Global-Scale Persistent Storage," *Proc. 9th Int'l Conf. Architectural Support for Programming Languages and Operating Systems* (SIGPLAN), 2000, pp. 190–201.
7. Q. Lu, S. Ratnasamy, and S. Shenker, "Can Heterogeneity Make Gnutella Scalable?" *Proc. 1st Int'l Workshop* (IPTPS 02), LNCS 2429, Springer, 2002, pp. 94–103.
8. C. Tand, Z. Xu, and S. Dwarkadas, "Peer-to-Peer Information Retrieval Using Self-Organizing Semantic Overlay Networks," *Proc. SIGCOMM*, ACM Press, 2003, pp. 175–186.
9. J. Hirshleifer, *The Dark Side of the Force*, Cambridge Univ. Press, 2001.
10. P. Golle, K. Leyton-Brown, and I. Mironov, "Incentives for Sharing in Peer-to-Peer Networks," *ACM Conf. Electronic Commerce*, ACM Press, 2001, pp. 264–267.
11. R. Dingledine, M.J. Freedman, and D. Molnar, "The Free Haven Project: Distributed Anonymous Storage Service," *Int'l Workshop on Design Issues in Anonymity and Unobservability*, LNCS 2009, H. Federrath, ed., Springer-Verlag, 2001, pp. 67–95.
12. A. Sen, "Social Choice Theory," *Handbook of Mathematical Economics*, vol. 3, K. Arrow and M.D. Intriligator, eds., Elsevier, 1986, pp. 1073–1181.
13. A. Serjantov and R. Anderson, "On Dealing with Adversaries Fairly," *3rd Ann. Workshop on Economics and Information Security* (WEIS 04), 2004.
14. H. Varian, "System Reliability and Free Riding," *Workshop on Economics and Information Security*, Univ. California, Berkeley Press, 2002; www.sims.berkeley.edu/resources/affiliates/workshops/econsecurity/econws/49.pdf.
15. D. Goodhart, "Discomfort of Strangers," *The Guardian*, 2 Feb. 2004, pp. 24–25.
16. "The Kindness of Strangers?" *The Economist*, 26 Feb. 2004.

George Danezis *is a research associate at Cambridge University's computer laboratory. His research interests include anonymous communications, traffic analysis, censorship-resistant publishing, and information hiding. Danezis has a PhD in computer security from the University of Cambridge. Contact him at George.Danezis@cl.cam.ac.uk.*

Ross Anderson *is a professor of security engineering at the University of Cambridge and one of the founders of the study of security economics. His research interests also include cryptography, protocols, hardware tamper-resistance and peer-to-peer systems. He is the author of the textbook* Security Engineering—A Guide to Building Dependable Distributed Systems. *Contact him via www.ross-anderson.com.*