Rebecca N. Wright (Ed.)

Financial Cryptography

7th International Conference, FC 2003
Guadeloupe, FrenchWest Indies, January 27-30, 2003
Revised Papers

On the Economics of Anonymity

Alessandro Acquisti[1], Roger Dingledine[2], and Paul Syverson[3]

[1] SIMS, UC Berkeley acquisti@sims.berkeley.edu
[2] The Free Haven Project arma@mit.edu
[3] Naval Research Lab syverson@itd.nrl.navy.mil

Abstract. Decentralized anonymity infrastructures are still not in wide use today. While there are technical barriers to a secure robust design, our lack of understanding of the incentives to participate in such systems remains a major roadblock. Here we explore some reasons why anonymity systems are particularly hard to deploy, enumerate the incentives to participate either as senders or also as nodes, and build a general model to describe the effects of these incentives. We then describe and justify some simplifying assumptions to make the model manageable, and compare optimal strategies for participants based on a variety of scenarios.

Keywords: Anonymity, economics, incentives, decentralized, reputation

1 Introduction

Individuals and organizations need anonymity on the Internet. People want to surf the Web, purchase online, and send email without exposing to others their identities, interests, and activities. Corporate and military organizations must communicate with other organizations without revealing the existence of such communications to competitors and enemies. Firewalls, VPNs, and encryption cannot provide this protection; indeed, Diffie and Landau have noted that traffic analysis is the backbone of communications intelligence, not cryptanalysis [9].

With so many potential users, it might seem that there is a ready market for anonymity services — that is, it should be possible to offer such services and develop a paying customer base. However, with one notable exception (the Anonymizer [2]) commercial offerings in this area have not met with sustained success. We could attribute these failures to market immaturity, and to the current economic climate in general. However, this is not the whole story.

In this paper we explore the incentives of participants to offer and use anonymity services. We set a foundation for understanding and clarifying our speculations about the influences and interactions of these incentives. Ultimately we aim to learn how to align incentives to create an economically workable system for users and infrastructure operators.

Section 2 gives an overview of the ideas behind our model. Section 3 goes on to describe the variety of (often conflicting) incentives and to build a general model that incorporates many of them. In Section 4 we give some simplifying

R.N. Wright (Ed.): FC 2003, LNCS 2742, pp. 84–102, 2003.

assumptions and draw conclusions about certain scenarios. Sections 5 and 6 describe some alternate approaches to incentives, and problems we encounter in designing and deploying strong anonymity systems.

2 The Economics of Anonymity

Single-hop web proxies like the Anonymizer protect end users from simple threats like profile-creating websites. On the other hand, users of such commercial proxies are forced to trust them to protect traffic information. Many users, particularly large organizations, are rightly hesitant to use an anonymity infrastructure they do not control. However, on an open network such as the Internet, running one's own system won't work: a system that carries traffic for only one organization will not hide the traffic entering and leaving that organization. Nodes must carry traffic from others to provide cover. The only viable solution is to distribute trust. That is, each party can choose to run a node in a shared infrastructure, if its incentives are large enough to support the associated costs. Users with more modest budgets or shorter-term interest in the system also benefit from this decentralized model, because they can be confident that a few colluding nodes are unlikely to uncover their anonymity.

Today, however, few people or organizations are willing to run these nodes. In addition to the complexities of configuring current anonymity software, running a node costs a significant amount of bandwidth and processing power, most of which is used by 'freeloading' users who do not themselves run nodes. Moreover, when administrators are faced with abuse complaints concerning illegal or antisocial use of their systems, the very anonymity that they're providing precludes the usual solution of suspending users or otherwise holding them accountable.

Unlike confidentiality (encryption), anonymity cannot be created by the sender or receiver. Alice cannot decide by herself to send anonymous messages — she must trust the infrastructure to provide protection, and others must use the same infrastructure. Anonymity systems use messages to hide messages: senders are consumers of anonymity and also providers of the cover traffic that creates anonymity for others. Thus users are better off on crowded systems because of the noise other users provide.

Because high traffic is necessary for strong anonymity, agents must balance their incentives to find a common equilibrium, rather than each using a system of their own. The high traffic they create together also enables better performance: a system that processes only light traffic must delay messages to achieve adequately large anonymity sets. But systems that process the most traffic do not necessarily provide the best hiding: if trust is not well distributed, a high volume system is vulnerable to insiders and attackers who target the trust bottlenecks.

Anonymity systems face a surprisingly wide variety of direct anonymity-breaking attacks [3,20]. Additionally, adversaries can also attack the efficiency or reliability of nodes, or try to increase the cost of running nodes. All of these factors combine to threaten the *anonymity* of the system. As Back et al. point out, "in anonymity systems usability, efficiency, reliability and cost become *secu-*

rity objectives because they affect the size of the user base which in turn affects the degree of anonymity it is possible to achieve." [3]

We must balance all of these tradeoffs while we examine the incentives for users and node operators to participate in the system.

3 Analytic Framework

In this section and those that follow, we formalize the economic analysis of why people might choose to send messages through mix-nets.[1] We discuss the incentives for agents to participate either as senders or also as nodes, and we propose a general framework to analyze these incentives. In the next section we consider various applications of our framework, and then in Section 5 we examine alternate incentive mechanisms.

We begin with two assumptions: the agents want to send messages to other parties, and the agents value their anonymity. How various agents might value their anonymity will be discussed below.

An agent i (where $i = (1, ..., n)$ and n is the number of potential participants in the mix-net) bases her strategy on the following possible actions a_i:

1. Act as a user of the system, specifically by sending (and receiving) her own traffic over the system, a_i^s, and/or agreeing to receive dummy traffic through the system, a_i^r. (*Dummy traffic* is traffic whose only purpose is to obscure actual traffic patterns.)
2. Act as an honest node, a_i^h, by receiving and forwarding traffic (and possibly acting as an exit node), keeping messages secret, and possibly creating dummy traffic.
3. Act as a dishonest node, a_i^d, by pretending to forward traffic but not doing so, by pretending to create dummy traffic but not doing so (or sending dummy traffic easily recognizable as such), or by eavesdropping traffic to compromise the anonymity of the system.
4. Send messages through conventional non-anonymous channels, a_i^n, or send no messages at all.

Various benefits and costs are associated with each agent's action and the simultaneous actions of the other agents. The expected benefits include:

1. Expected benefits from sending messages anonymously. We model them as a function of the subjective value each agent i places on the information successfully arriving at its destination, v_{r_i}; the subjective value of keeping her identity anonymous, v_{a_i}; the perceived level of anonymity in the system, p_{a_i} (the subjective probability that the sender and message will remain anonymous); and the perceived level of reliability in the system, p_{r_i} (the subjective probability that the message will be delivered). The subjective value

[1] Mixes were introduced by David Chaum (see [6]). A mix takes in a batch of messages, changes their appearance, and sends them out in a new order, thus obscuring the relation of incoming to outgoing messages.

of maintaining anonymity could be related to the profits the agent expects to make by keeping that information anonymous, or the losses the agents expects to avoid by keeping that information anonymous. We represent the level of anonymity in the system as a function of the traffic (number of agents sending messages in the system, n_s), the number of nodes (number of agents acting as honest nodes, n_h, and as dishonest nodes, n_d), and the decisions of the agent. We assume the existence of a function that maps these factors into a probability measure $p \in [0, 1]$.[2] In particular:

- The level of anonymity of the system is positively correlated to the number of users of the system.
- Acting as an honest node improves anonymity. Senders who do not run a node may accidentally choose a dishonest node as their first hop, significantly decreasing their anonymity (especially in low-latency anonymity systems where end-to-end timing attacks are very hard to prevent [3]). Further, agents who run a node can undetectably blend their message into their node's traffic, so an observer cannot know when the message is sent.
- The relation between the number of nodes and the probability of remaining anonymous might not be monotonic. For a given amount of traffic, sensitive agents might want fewer nodes in order to maintain large anonymity sets. But if some nodes are dishonest, users may prefer more honest nodes (to increase the chance that messages go through honest nodes). Agents that act as nodes may prefer fewer nodes, to maintain larger anonymity sets at their particular node. Hence the probability of remaining anonymous is inversely related to the number of nodes but positively related to the ratio of honest/dishonest nodes. (On the other hand, improving anonymity by reducing the number of nodes can be taken too far — a system with only one node may be easier to monitor and attack. See Section 5 for more discussion.)

If we assume that honest nodes always deliver messages that go through them, the level of reliability in the system is then an inverse function of the share of dishonest nodes in the system, n_d/n_h.

2. Benefits of acting as a node (nodes might be rewarded for forwarding traffic or for creating dummy traffic), b_h.
3. Benefits of acting as a dishonest node (from disrupting service or by using the information that passes through them), b_d.

The possible expected costs include:

1. Costs of sending messages through the anonymous system, c_s, or through a non-anonymous system, c_n. These costs can include both direct financial

[2] Information theoretic anonymity metrics [8,22] probably provide better measures of anonymity: such work shows how the level of anonymity achieved by an agent in a mix-net system is associated to the particular structure of the system. But probabilities are more tractable in our analysis, as well as better than the common "anonymity set" representation.

costs such as usage fees, as well as implicit costs such as the time to build and deliver messages, learning curve to get familiar with the system, and delays incurred when using the system. At first these delays through the anonymous system seem positively correlated to the traffic n_s and negatively correlated to the number of nodes n_h. But counterintuitively, more messages per node might instead *decrease* latency because nodes can process batches more often; see Section 5. In addition, when message delivery is guaranteed, a node might always choose a longer route to reduce risk. We could assign a higher c_s to longer routes to reflect the cost of additional delay. We also include here the cost of receiving dummy traffic, c_r.

2. Costs of acting as an honest node, c_h, by receiving and forwarding traffic, creating dummy traffic, or being an exit node (which involves potential exposure to liability from abuses). These costs can be variable or fixed. The fixed costs, for example, are related to the investments necessary to setup the software. The variable costs are often more significant, and are dominated by the costs of traffic passing through the node.
3. Costs of acting as dishonest node, c_d (again carrying traffic; and being exposed as a dishonest node may carry a monetary penalty).

In addition to the above costs and benefits, there are also *reputation* costs and benefits from: being observed to send or receive anonymous messages, being perceived to act as a reliable node, and being thought to act as a dishonest node.

Some of these reputation costs and benefits could be modelled endogenously (e.g., being perceived as an honest node brings that node more traffic, and therefore more possibilities to hide that node's messages; similarly, being perceived as a dishonest node might bring traffic away from that node). In this case, they would enter the payoff functions only indirectly through other parameters (such as the probability of remaining anonymous) and the changes they provoke in the behavior of the agents. In other cases, reputation costs and benefits might be valued *per se*. While we do not consider either of these options in the simplified model below, Sections 5 and 6 discuss the impact of reputation on the model.

We assume that agents want to maximize their expected payoff, which is a function of expected benefits minus expected costs. Let S_i denote the set of strategies available to agent i, and s_i a certain member of that set. Each strategy s_i is based on the the actions a_i discussed above. The combination of strategies $(s_1, ..., s_n)$, one for each agent who participates in the system, determines the outcome of a game as well as the associated payoff for each agent. Hence, for each complete strategy profile $s = (s_1, ..., s_n)$ each agent receives the expected payoff $u_i(s)$ through the payoff function $u(.)$. We represent the payoff function for each agent i in the following form:

$$u_i = u \begin{pmatrix} \theta\left[\gamma\left(v_{r_i}, p_{r_i}\left(n_h, n_d, a_h^s\right)\right), \partial\left(v_{a_i}, p_{a_i}\left(n_s, n_h, n_d, a_h^s\right)\right), a_i^s\right] + b_h a_i^h + b_d a_i^d \\ -c_s\left(n_s, n_h\right) a_i^s - c_h\left(n_s, n_h, n_d\right) a_i^h - c_d\left(..\right) a_i^d - c_r\left(..\right) a_i^r + \left(b_n - c_n\right) a_i^n \end{pmatrix}$$

where $\theta(.), \gamma(.)$, and $\partial(.)$ are unspecified functional forms. The payoff function $u(.)$ includes the costs and benefits for all the possible actions of the agents,

including *not* using the mix-net and instead sending the messages through a non-anonymous channel. We can represent the various strategies by using dummy variables for the various a_i.[3] We note that the probabilities of a message being delivered and a message remaining anonymous are weighted with the values v_{r_i}, v_{a_i}, respectively. This is because different agents might value anonymity and reliability differently, and because in different scenarios anonymity and reliability for the same agent might have different impacts on her payoff.

In Section 4, we will make a number of assumptions that will allow us to simplify this equation and model certain scenarios. We present here for the reader's convenience a table summarizing those variables that will appear in both the complete and simplified equations, as well as one that describes the variables used only in the more complete equation above.

Variables used in both full and simple payoff equations	
	u_i payoff for agent i
	v_{a_i} disutility i attaches to message exposure
	p_a simple case: $p_{a_i} = p_a$ for all i. See next table.
number of nodes (other than i) in mix-net	n_s sending agents (sending nodes)
	n_h honest nodes
	n_d dishonest nodes
dummy variables: 1 if true, 0 otherwise	a_i^h i is an honest node and sending agent
	a_i^s i sends through the mix-net
costs	c_h of running an honest node
	c_s of sending a message through the mix-net

Variables used only in full payoff equation	
	v_{r_i} value i attaches to sent message being received
	p_{a_i} prob. for i that a sent message loses anonymity
	p_r prob. that message sent through mix-net is received
benefits	b_h of running an honest node
	b_d of running a dishonest node
	b_n of sending a message around the mix-net
dummy variables	a_i^d i runs a dishonest node
	a_i^n i sends message around the mix-net
	a_i^r i receives dummy traffic
costs	c_d of running a dishonest node
	c_r of receiving dummy traffic
	c_n of sending a message around the mix-net

Note also that the costs and benefits from sending the message could be distinct from the costs and benefits from keeping the *information* anonymous. For example, when Alice anonymously purchases a book, she gains a profit equal

[3] For example, if the agent chooses not to send the message anonymously, the probability of remaining anonymous p_{a_i} will be equal to zero, $a^{s,d,r,h}$ will be zero too, and the only cost in the function will be c_n.

to the difference between her valuation of the book and its price. But if her anonymity is compromised during the process, she could incur losses (or miss profits) completely independent from the price of the book or her valuation of it. The payoff function $u(.)$ above allows us to represent the duality implicit in all privacy issues, as well as the distinction between the value of sending a message and the value of keeping it anonymous:

Anonymity	*Reliability*
Benefit from remaining anonymous / cost avoided by remaining anonymous, or	Benefit in sending message that will be received / cost avoided by sending such a message, or
Cost from losing anonymity / profits missed because of loss of anonymity	Cost from a message not being received / profits missed by message not being received

Henceforth, we will consider the direct benefits or losses rather than their dual opportunity costs or avoided costs. Nevertheless, the above representation allows us to formalize the various possible combinations. For example, if a certain message is sent to gain some benefit, but anonymity must be protected in order to avoid losses, then v_{r_i} will be positive while v_{a_i} will be negative and p_{a_i} will enter the payoff function as $(1 - p_{a_i})$. On the other side, if the agent must send a certain message to avoid some losses but anonymity ensures her some benefits, then v_{r_i} will be negative and p_{r_i} will enter the payoff function as $(1 - p_{r_i})$, while v_{a_i} will be positive.[4]

With this framework we can compare, for example, the losses due to compromised anonymity to the costs of protecting it. An agent will decide to protect herself by spending a certain amount if the amount spent in defense plus the expected losses for losing anonymity after the investment are less than the expected losses from not sending the message at all.

4 Applying the Model

In this section we apply the above framework to simple scenarios. We make a number of assumptions to let us model the behavior of mix-net participants as players in a repeated-game, simultaneous-move game-theoretic framework. Thus we can analyze the economic justifications for the various choices of the participants, and compare design approaches to mix-net systems.

Consider a set of n_s agents interested in sending anonymous communications. Imagine that there is only one system which can be used to send anonymous messages, and one other system to send non-anonymous messages. Each agent has three options: only send her own messages through the mix-net; send her messages but also act as a node forwarding messages from other users; or don't use the system at all (by sending a message without anonymity, or by not sending

[4] Being certain of staying anonymous would therefore eliminate the risk of v_{a_i}, while being certain of losing anonymity would impose on the agent the full cost v_{a_i}. Similarly, guaranteed delivery will eliminate the risk of losing v_{r_i}, while delivery failure will impose the full cost v_{r_i}.

the message). Thus initially we do not consider the strategy of choosing to be a bad node, or additional honest strategies like creating and receiving dummy traffic.

We represent the game as a simultaneous-move, repeated game because of the large number of participants and because of the impact of earlier actions on future strategies. A large group will have no discernable or agreeable order for the actions of all participants, so actions can be considered simultaneous. The limited commitment produced by earlier actions allows us to consider a repeated-game scenario.[5] These two considerations suggest against using a sequential approach of the Stackelberg type [14, Ch. 3]. For similar reasons we also avoid a "war of attrition/bargaining model" framework (see for example [21]) where the relative impatience of players plays an important role.

4.1 Adversary

Although strategic agents cannot choose to be bad nodes in this simplified scenario, we still assume there is a percentage of bad nodes and that agents respond to this possibility. Specifically we assume a global passive adversary (GPA) that can observe all traffic on all links (between users and nodes, between nodes, and between nodes or users and recipients). Additionally, we also study the case when the adversary includes some percentage of mix nodes. In choosing strategies agents will attach a subjective probability to arbitrary nodes being compromised — all nodes not run by the agent are assigned the same probability of being compromised. This factor influences their assessment of the anonymity of messages they send. A purely passive adversary is unrealistic in most settings, e.g., it assumes that hostile users never selectively send messages at certain times or over certain routes, and nodes and links never selectively trickle or flood messages [23]. Nonetheless, a *global* passive adversary is still quite strong, and thus a typical starting point of anonymity analyses.

4.2 Honest Agents

If a user only sends messages, the cost of using the anonymous service is c_s. This cost might be higher than using the non-anonymous channel, c_n, because of usage fees, usage hassles, or delays. To keep things simple, we assume that all messages pass through the mix-net in fixed-length free routes, so that we can write c_s as a fixed value, the same for all agents. Users send messages at the same time, and only one message at a time. We also assume that routes are chosen randomly by users, so that traffic is uniformly distributed among the nodes.[6]

If a user decides to be a node, her costs increase with the volume of traffic (we focus here on the traffic-based variable costs). We also assume that all agents know the number of agents using the system and which of them are acting as

[5] In Section 3 we have highlighted that, for both nodes and simpler users, variable costs are more significant than fixed costs.

[6] Reputation considerations might alter this point; see Section 5.

nodes. We also assume that all agents perceive the same level of anonymity in the system based on traffic and number of nodes, hence $p_{a_i} = p_a$ for all i. Finally, we imagine that agents use the system because they want to avoid potential losses from not being anonymous. This subjective sensitivity to anonymity is represented by v_{a_i} (we can initially imagine v_{a_i} as a continuous variable with a certain distribution across all agents; see below). In other words, we initially focus on the goal of remaining anonymous given an adversary that can control some nodes and observe all communications. Other than anonymity, we do not consider any potential benefit or cost, e.g., possible greater reliability, from sending around the mix-net. We later comment on the additional reliability issues.

$$u_i = -v_{a_i}\left(1 - p_a\left(n_s, n_h, n_d, a_i^h\right)\right) - c_s a_i^s - c_h\left(n_s, n_h, n_d\right) a_i^h - v_{a_i} a_i^n$$

Thus each agent i tries to *minimize* the costs of sending messages and the risk of being tracked. The first component is the probability that anonymity will be lost given the number of agents sending messages, the number of them acting as honest and dishonest nodes, and the action a of agent i itself. This chance is weighted by v_{a_i}, the disutility agent i derives from its message being exposed. We also include the costs of sending a message through the mix-net, acting as a node when there are n_s agents sending messages over n_h and n_d nodes, and sending messages through a non-anonymous system, respectively. Each period, a rational agent can compare the payoff coming from each of these three one-period strategies.

Action	Payoff
a_s	$-v_{a_i}\left(1 - p_a\left(n_s, n_h, n_d\right)\right) - c_s$
a_h	$-v_{a_i}\left(1 - p_a\left(n_s, n_h, n_d, a_i^h\right)\right) - c_s - c_h\left(n_s, n_h, n_d\right)$
a_n	$-v_{a_i}$

We do not explicitly allow the agent to choose *not* to send a message at all, which would of course minimize the risk of anonymity compromise. Also, we do not explicitly report the value of sending a successful message. Both are simplifications that do not alter the rest of the analysis.[7]

While this model is simple, it allows us to highlight some of the dynamics that might take place in the decision process of agents willing to use a mix-net. We now consider various versions of this model.

[7] We could insert an action a^0 with a certain disutility or cost from not sending any message, and then solve the problem of minimizing the expected losses. Or, we could insert in the payoff function for actions $a^{s,h,n}$ also the payoff from successfully sending a message compared to not sending it (which could be interpreted also as an opportunity cost), and solve the dual problem of maximizing the expected payoff. Either way, the "exit" strategy for each agent will either be sending a message non-anonymously, or not sending it at all, depending on which option maximizes the expected benefits or minimizes the expected losses. Thereafter, we can simply compare the two other actions (being a user, or being also a node) to the optimal exit strategy.

Myopic Agents. Myopic agents do not consider the long-term consequences of their actions. They simply consider the status of the network and, depending on the payoffs of the one-period game, adopt a certain strategy. Suppose that a new agent with a privacy sensitivity v_{a_i} is considering using a mix-net with (*currently*) n_s users and n_h honest nodes.

Then if

$$\begin{gathered}-v_{a_i}\left(1-p_a\left(n_s+1,n_h+1,n_d,a_i^h\right)\right)-c_s-c_h\left(n_s+1,n_h+1,n_d\right)\\ <-v_{a_i}\left(1-p_a\left(n_s+1,n_h,n_d\right)\right)-c_s, \text{ and}\\ -v_{a_i}\left(1-p_a\left(n_s+1,n_h+1,n_d,a_i^h\right)\right)-c_s-c_h\left(n_s+1,n_h+1,n_d\right)\\ <-v_{a_i}\end{gathered}$$

agent i will choose to become a node in the mix-net. If

$$\begin{gathered}-v_{a_i}\left(1-p_a\left(n_s+1,n_h+1,n_d,a_i^h\right)\right)-c_s-c_h\left(n_s+1,n_h+1,n_d\right)\\ >-v_{a_i}\left(1-p_a\left(n_s+1,n_h,n_d\right)\right)-c_s, \text{ and}\\ -v_{a_i}\left(1-p_a\left(n_s+1,n_h,n_d\right)\right)-c_s<-v_{a_i}\end{gathered}$$

then agent i will choose to be a user of the mix-net. Otherwise, i will simply not use the mix-net.

Our goal is to highlight the economic rationale implicit in the above inequalities. In the first case agent i is comparing the benefits of the contribution to her own anonymity of acting as a node to the costs. Acting as a node dramatically increases anonymity, but it will also bring more traffic-related costs to the agent. Agents with high privacy sensitivity (high v_{a_i}) will be more likely to accept the trade-off and become nodes because they risk a lot by losing their anonymity, and because acting as nodes significantly increases their probabilities of remaining anonymous. On the other side, agents with a lower sensitivity to anonymity might decide that the costs or hassle of using the system are too high, and would not send the message (or would use non-anonymous channels).

Strategic Agents: Simple Case. Strategic agents take into consideration the fact that their actions will trigger responses from the other agents.

We start by considering only one-on-one interactions. First we present the case where each agent knows the other agent's type, but we then discuss what happens when there is uncertainty about the other agent's type.

Suppose that each of agent i and agent j considers the other agent's reaction function in her decision process. Then we can summarize the payoff matrix in the following way:[8]

[8] We use parameters to succinctly represent the following expected payoffs:

$$\begin{gathered}A_w=-v_w\left(1-p_a\left(n_s+2,n_h+2,n_d,a_w^h\right)\right)-c_s-c_h\left(n_s+2,n_h+2,n_d\right)\\ B_w=-v_w\left(1-p_a\left(n_s+2,n_h+1,n_d\right)\right)-c_s\\ C_w=-v_w\\ D_w=-v_w\left(1-p_a\left(n_s+2,n_h+1,n_d,a_w^h\right)\right)-c_s-c_h\left(n_s+2,n_h+1,n_d\right)\\ E_w=-v_w\left(1-p_a\left(n_s+1,n_h+1,n_d,a_w^h\right)\right)-c_s-c_h\left(n_s+1,n_h+1,n_d\right)\\ F_w=-v_w\left(1-p_a\left(n_s+2,n_h,n_d\right)\right)-c_s\\ G_w=-v_w\left(1-p_a\left(n_s+1,n_h,n_d\right)\right)-c_s\end{gathered}$$

Agent i / Agent j	a_j^h	a_j^s	a_j^n
a_i^h	A_i, A_j	D_i, B_j	E_i, C_j
a_i^s	B_i, D_j	F_i, F_j	G_i, C_j
a_i^n	C_i, E_j	C_i, G_j	C_i, C_j

As before, each agent has a trade-off between the cost of traffic and the benefit of traffic when being a node, and a trade-off between having more nodes and fewer nodes. In addition to the previous analysis, now the final outcome also depends on how much each agent knows about whether the other agent is honest, and how much she knows about the other agent's sensitivity to privacy.

Of course, for an explicit solution we need a specific functional form for the probability function.[9] Nevertheless, even at this abstract level of description this framework can be mapped into the model analyzed in [19] where two players decide simultaneously whether to contribute to a public good.

In our model, when for example $v_{a_i} \gg v_{a_j}$ and v_{a_i} is large, the disutility to player i from not using the system or not being a node will be so high that she will decide to be a node even if j might free ride on her. Hence if j values her anonymity, but not that much, the strategies a_i^h,a_j^s can be an equilibrium of the repeated game.

In fact, this model might have equilibria with free-riding even when the other agent's type is unknown. Imagine both agents know that the valuations v_{a_i}, v_{a_j} are drawn independently from a continuous, monotonic probability distribution. Again, when one agent cares about her privacy enough, and/or believes that there is a high probability that the opponent would act as a dishonest node, then the agent will be better off protecting her own interests by becoming a node (again see [19]). Of course the more interesting cases are those when these clear-cut scenarios do not arise, which we consider next.

Strategic Agents: Multi-player Case. Each player now considers the strategic decisions of a vast number of other players. Fudenberg and Levine [13] propose a model where each player plays a large set of identical players, each of which is "infinitesimal", i.e. its actions cannot affect the payoff of the first player. We define the payoff of each player as the average of his payoffs against the distribution of strategies played by the continuum of the other players. In other words, for each agent, we will have: $u_i = \sum_{n_s} u_i\,(a_i, a_{-i})$ where the notation represents the comparison between one specific agent i and all the others. Cooperative solutions with a finite horizon are often not sustainable when the actions of other agents are not observable because, by backward induction, each agent will have an incentive to deviate from the cooperative strategy. As compared to the analysis above with only two agents, now a defection of one agent might

[9] We have seen above, however, that privacy metrics like [8,22] do not directly translate into monotonic probability functions of the type traditionally used in game theory. Furthermore, the actual level of anonymity will depend on the mix-net protocol and topology (synchronous networks will provide larger anonymity sets than asynchronous networks for the same traffic divided among the nodes).

affect only infinitesimally the payoff of the other agents, so the agents might tend not to punish the defector. But then, more agents will tend to deviate and the cooperative equilibrium might collapse. "Defection", in fact, could be acting only as a user and refusing to be a node when the agent starts realizing that there is enough anonymity in the system and she no longer needs to be a node. But if too many agents act this way, the system might break down for lack of nodes, after which everybody would have to resort to non-anonymous channels.

We can consider this to be a "public good with free-riding" type of problem [7]. The novel point from a game-theoretic perspective is that the highly sensitive agents actually *want* some level of free-riding, to provide noise. On the other side, they do not want too much free-riding — for example from highly sensitive types pretending to be agents with low sensitivity — if it involves high traffic costs.

So, under which conditions will a system with many players not implode?

First, a trigger strategy might be agreed upon among the many agents, so that the deviation of one single player might be met by the reaction of all the others (as described in [13]). Of course the only punishment available here is making the system unavailable, which has a cost for all agents. In addition, coordination costs might be prohibitive. This is not a viable strategy.

Second, we must remember that highly sensitive agents, for a given amount of traffic, prefer to be nodes (because anonymity will increase) and prefer to work in systems with fewer nodes (else traffic gets too dispersed and the anonymity sets get too small). So, if v_{a_i} is particularly high, i.e. if the cost of not having anonymity is very high for the most sensitive agents, then they will decide to act as nodes regardless of what the others do. Also, if there are enough agents with lower v_{a_i}, again a "high" type might have an interest in acting alone if its costs of not having anonymity would be too high compared to the costs of handling the traffic of the less sensitive types.

In fact, when the valuations are continuously distributed, this *might* generate equilibria where the agents with the highest valuations v_{a_i} become nodes, and the others, starting with the "marginal" type (the agent indifferent between the benefits she would get from acting as node and the added costs of doing so) provide traffic.[10] This problem can be mapped to the solutions in [4] or [17]. At that point an equilibrium level of free-riding might be reached. This condition can be also compared to [15], where the paradox of informationally efficient markets is described.[11]

The problems start if we consider now a different situation. Rather than having a continuous distribution of valuations v_{a_i}, we consider two types of agents: the agent with a high valuation, $v_{a_i} = v_H$, and the agent with a low valuation, $v_{a_i} = v_L$. We assume that the v_L agents will simply participate sending traffic if the system is cheap enough for them to use (but see Section 6.3), and we also assume this will not pose any problem to the v_H type, which in fact has an

[10] Writing down specific equilibria, again, will first involve choosing appropriate anonymity metrics, which might be system-dependent.

[11] The equilibrium in [15] relies on the "marginal" agent who is indifferent between getting more information about the market and not getting it.

interest in having more traffic. Thus we can focus on the interaction between a subset of users: the identical high-types.

Here the "marginal" argument discussed above might not work, and coordination might be costly. In order to have a scenario where the system is self-sustaining and free, and the agents are of high and low types, the actions of the agents must be visible and the agents themselves must agree to react together to any deviation of a marginal player. In realistic scenarios, however, this will involve very high transaction/coordination costs, and will require an extreme (and possibly unlikely) level of rationality for the agents. This equilibrium will also tend to collapse when the benefits from being a node are not very high compared to the costs. Paradoxically, it also breaks down when an agent trusts another so much that she prefers to delegate away the task of being a node. The above considerations however also hint at other possible solutions to reduce coordination costs. We now consider some other mechanisms that can make these systems economically viable.

5 Alternate Incentive Mechanisms

As the self-organized system might collapse under some of the conditions examined above, we discuss now what economic incentives we can get from alternative mechanisms.

1. *Usage fee.* If participants pay to use the system, the "public good with free-riding" problem turns into a "clubs" scenario. The pricing mechanism must be related to how much the participants expect to use the system or how sensitive they are. Sensitive agents might support the others by offering them limited services for free, because they need their traffic as noise. The Anonymizer offers basic service at low costs to low-sensitivity agents (there is a cost in the delay, the limitation on destination addresses, and the hassle of using the free service), and offers better service for money. With usage fees, the cost of being a node is externalized. A hybrid solution involves distributed trusted nodes, supported through entry fees paid to a central authority and redistributed to the nodes. This was the approach of the Freedom Network from Zero-Knowledge Systems. The network was shut down because they were unable to sell enough clients to cover their costs.
2. *"Special" agents.* Such agents have a payoff function that considers the social value of having an anonymous system or are otherwise paid or supported to provide such service. If these agents are paid, the mechanism becomes similar to the hybrid solution discussed above, except anonymity-sensitive agents, rather than act as nodes, pass the money to a central authority. The central authority redistributes the funding among trusted entities acting as nodes.
3. *Public rankings and reputation.* A higher reputation not only attracts more cover traffic but is also a reward in itself. Just as the statistics pages for seti@home [5] encourage participation, publicly ranking generosity creates an incentive to participate. Although the incentives of public recognition and

public good don't fit in our model very well, we emphasize them because they explain most actual current node operators. As discussed above, reputation can enter the payoff function indirectly or directly (when agents value their reputation as a good itself).

If we publish a list of nodes ordered by safety (based on number of messages passing through the node), the high-sensitivity agents will gravitate to safe nodes, causing more traffic and improving their safety further (and lowering the safety of other nodes). In our model the system will stabilize with one or a few mix nodes. In reality, though, p_a is influenced not just by n_h but also by jurisdictional diversity — a given high-sensitivity sender is happier with a diverse set of mostly busy nodes than with a set of very busy nodes run in the same zone. Also, after some threshold of users, latency will begin to suffer, and the low sensitivity users will go elsewhere, taking away the nice anonymity sets.

More generally, a low-latency node may attract many low-sensitivity agents, and thus counterintuitively provide *better* anonymity than one that waits to batch many messages for greater security.

6 A Few More Roadblocks

6.1 Authentication in a Volunteer Economy

Our discussions so far indicate that it may in fact be plausible to build a strong anonymity infrastructure from a wide-spread group of independent nodes that each want good anonymity for their own purposes. In fact, the more jurisdictionally diverse this group of nodes, the more robust the overall system.

However, volunteers are problems: users don't know the node operators, and don't know whether they can trust them. We can structure system protocols to create better incentives for honest principals and to catch bad performance by others, e.g. by incorporating receipts and trusted witnesses [10], or using a self-regulating topology based on verifying reliability [11]. But even when this is feasible, identifying individuals is a problem. Classic authentication considers whether it's the right entity, but not whether the authenticated parties are distinct from one another. One person may create and control several distinct online identities. This *pseudospoofing* problem [12] is a nightmare when an anonymity infrastructure is scaled to a large, diffuse, peer-to-peer design; it remains one of the main open problems in the design of any decentralized anonymity service. The Advogato trust metric [16] and similar techniques rely on humans to make initial trust decisions, and then bound trust flow over a certification graph. However, so far none of these trust flow approaches have provided a clear solution to the problem. Another potential solution, a global PKI to ensure unique identities [24], is unlikely to emerge any time soon.

6.2 Dishonest Nodes vs. Lazy Nodes

We have primarily focused on the strategic motivations of honest agents, but the motivations of dishonest agents are at least as important. An anonymity-breaking adversary with an adequate budget would do best to provide very good service, possibly also attempting DoS against other high-quality providers. None of the usual metrics of performance and efficiency can identify dishonest nodes. Further, who calculates those metrics and how? If they depend on a centralized trusted authority, the advantages of diffusion are lost. Another approach to breaking anonymity is to simply attack the reliability or perceived reliability of the system — this attack flushes users to a weaker system just as military strikes against underground cables force the enemy to communicate over less secure channels.

On the other hand, when we consider strategic dishonest nodes we must also analyze their motivations as rational agents. A flat-out dishonest agent participates only to compromise anonymity or reliability. In doing so, however, a dishonest agent will have to consider the costs of reaching and maintaining a position from which those attacks are effective — which will probably involve gaining reputation and acting as a node for an extended period of time, a cost if the goal is to generally break reliability. Such adversaries will be in an arms race with protocol developers to stay undetected despite their attacks [11]. The benefits from successful attacks might be financial, as in the case of discovering and using sensitive information or a competitor's service being disrupted; or they could be purely related to personal satisfaction. The costs of being discovered as a dishonest node include rebuilding a new node's worth of reputation; but being noticed and exposed as the adversary may have very serious negative consequences for the attacker itself. (Imagine the public response if an Internet provider were found running dishonest nodes.) Thus, all things considered, it might be that the laws of economics work against the attacker as well.

A "lazy" node, on the other hand, wants to protect her own anonymity, but keeps her costs lower by not forwarding or accepting all of her incoming traffic. By doing so this node decreases the reliability of the system. While this strategy might be sounder than the one of the flat-out dishonest node, it also exposes again the lazy node to the risk of being recognized as a disruptor of the system. In addition, this tactic, by altering the flow of the traffic through her own node, might actually reduce the anonymity of that agent.

Surveys and analysis on actual attacks on actual systems (e.g., [18]) can help determine which forms of attacks are frequent, how dangerous they are, and whether economic incentives or technical answers are the best countermeasures.

6.3 Bootstrapping the System and Perceived Costs

Our models so far have considered the strategic choices of agents facing an already existing mix-net. We might even imagine that the system does not yet exist but that, before the first period of the repeated-game, all the players can

somehow know each other and coordinate to start with one of the cooperative equilibria discussed above.

But this does not sound like a realistic scenario. Hence we must discuss how a mix-net system with distributed trust can come to be. We face a paradox here: agents with high privacy sensitivity want lots of traffic in order to feel secure using the system. They need many participants with lower privacy sensitivities using the system first. The problem lies in the fact that there is no reason to believe the lower sensitivity types are more likely to be early adopters. In addition, their *perceived* costs of using the system might be higher than the real costs[12] — especially when the system is new and not well known — so in the strategic decision process they will decide against using the mix-net at all. Correct marketing seems critical to gaining critical mass in an anonymity system: in hindsight, perhaps Zero-Knowledge Systems would have gotten farther had it placed initial emphasis on usability rather than security.

Note that here again reliability becomes an issue, since we must consider both the benefits from sending a message *and* keeping it anonymous. If the benefits of sending a message are not that high to begin with, then a low sensitivity agent will have fewer incentives to spend anything on the message's anonymity. We can also extend the analysis from our model that considers the costs and benefits of a single system to the comparison of different systems with different costs/benefit characteristics. We comment more on this in the conclusion.

Difficulties in bootstrapping the system and the myopic behavior [1] of some users might make the additional incentive mechanisms discussed in Section 5 preferable to a market-only solution.

6.4 Customization and Preferential Service Are Risky Too

Leaving security decisions up to the user is traditionally a way to transfer cost or liability from the vendor to the customer; but in strong anonymity systems it may be unavoidable. For example, the sender might choose how many nodes to use, whether to use mostly nodes run by her friends, whether to send in the morning or evening, etc. After all, only she knows the value of her anonymity. But this choice also threatens anonymity — different usage patterns can help distinguish and track users.

Limiting choice of system-wide security parameters can protect users by keeping the noise fairly uniform, but introduces inefficiencies; users that don't need as much protection may feel they're wasting resources. Yet we risk anonymity if we let users optimize their behavior. We can't even let users pay for better service or preferential treatment — the hordes in the coach seats are more anonymous than the few in first class.

[12] Many individuals tend to be myopic in their attitude to privacy. They claim they want it but they are not willing to pay for it. While this might reflect a rational assessment of the trade-offs (that is, quite simply, the agents do not value their anonymity highly enough to justify the cost to protect it), it might also reflect "myopic" behavior such as the hyperbolic discounting of future costs associated to the loss of anonymity. See also [1].

This need to pigeonhole users into a few behavior classes conflicts with the fact that real-world users have a continuum of interests and approaches. Reducing options can lead to reduced usability, scaring away the users and leaving a useless anonymity system.

7 Future Work

There are a number of directions for future research:

- Dummy traffic. Dummy traffic increases costs but it also increases anonymity. In this extension we should study bilateral or multilateral contracts between agents, contractually forcing each agent to send to another agent(s) a certain number of messages in each period. With these contracts, if the sending agent does not have enough real messages going through its node, it will have to generate them as dummy traffic in order not to pay a penalty.
- Reliability. As noted above, we should add reliability issues to the model.
- Strategic dishonest nodes. As we discussed, it is probably more economically sound for an agent to be a lazy node than an anonymity-attacking node. Assuming that strategic bad nodes can exist, we should study the incentives to act honestly or dishonestly and the effect on reliability and anonymity.
- Unknown agent types. We should extend the above scenarios further to consider a probability distribution for an agent's guess about another agent's privacy sensitivity.
- Comparison between systems. We should compare mix-net systems to other systems, as well as use the above framework to compare the adoption of systems with different characteristics.
- Exit nodes. We should extend the above analysis to consider specific costs such as the potential costs associated with acting as an exit node.
- Reputation. Reputation can have a powerful impact on the framework above in that it changes the assumption that traffic will distribute uniformly across nodes. We should extend our analysis to study this more formally.
- Information theoretic metric. We should extend the analysis of information theoretic metrics in order to formalize the functional forms in the agent payoff function.

8 Conclusions

We have described the foundations for an economic approach to the study of strong anonymity infrastructures. We focused on the incentives for participants to act as senders and nodes. Our model does not solve the problem of building a more successful system — but it does provide some guidelines for how to think about solving that problem. Much research remains for a more realistic model, but we can already draw some conclusions:

- Systems must attract cover traffic (many low-sensitivity users) before they can attract the high-sensitivity users. Weak security parameters (e.g. smaller batches) may produce *stronger* anonymity by bringing more users. But to attract this cover traffic, they may well have to address the fact that most users do not want (or do not realize they want) anonymity protection.
- High-sensitivity agents have incentive to run nodes, so they can be certain their first hop is honest. There can be an optimal level of free-riding: in some conditions these agents will opt to accept the cost of offering service to others in order to gain cover traffic.
- While there are economic reasons for distributed trust, the deployment of a completely decentralized system might involve coordination costs which make it unfeasible. A central coordination authority to redistribute payments may be more practical, but could provide a trust bottleneck for an adversary to exploit.

Acknowledgments. Work on this paper was supported by ONR. Thanks to John Bashinski, Nick Mathewson, Adam Shostack, Hal Varian, and the anonymous referees for helpful comments.

References

1. Alessandro Acquisti and Hal R. Varian. Conditioning prices on purchase history. mimeo, University of California, Berkeley, 2002. `http://www.sims.berkeley.edu/~acquisti/papers/`.
2. The Anonymizer. `http://www.anonymizer.com/`.
3. Adam Back, Ulf Möller, and Anton Stiglic. Traffic analysis attacks and trade-offs in anonymity providing systems. In Ira S. Moskowitz, editor, *Information Hiding (IH 2001)*, pages 245–257. Springer-Verlag, LNCS 2137, 2001.
4. Theodore Bergstrom, Lawrence Blume, and Hal R. Varian. On the private provision of public goods. *Journal of Public Economics*, 29:25–49, 1986.
5. UC Berkeley. SETI@home: Search for Extraterrestrial Intelligence at Home. `http://setiathome.ssl.berkeley.edu/`.
6. David Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–88, 1981.
7. Richard Cornes and Todd Sandler. *The Theory of Externalities, Public Goods and Club Goods*. Cambridge University Press, 1986.
8. Claudia Díaz, Stefaan Seys, Joris Claessens, and Bart Preneel. Towards measuring anonymity. In Roger Dingledine and Paul Syverson, editors, *Privacy Enhancing Technologies (PET 2002)*. Springer-Verlag, LNCS 2482, 2002.
9. Whitfield Diffie and Susan Landau. *Privacy On the Line: The Politics of Wiretapping and Encryption*. MIT Press, 1998.
10. Roger Dingledine, Michael J. Freedman, David Hopwood, and David Molnar. A Reputation System to Increase MIX-net Reliability. In Ira S. Moskowitz, editor, *Information Hiding (IH 2001)*, pages 126–141. Springer-Verlag, LNCS 2137, 2001. `http://www.freehaven.net/papers.html`.
11. Roger Dingledine and Paul Syverson. Reliable MIX Cascade Networks through Reputation. In Matt Blaze, editor, *Financial Cryptography (FC '02)*. Springer-Verlag, LNCS 2357, 2002.

12. John Douceur. The Sybil Attack. In *1st International Peer To Peer Systems Workshop (IPTPS 2002)*, March 2002.
13. Drew Fudenberg and David K. Levine. Open-loop and closed-loop equilibria in dynamic games with many players. *Journal of Economic Theory*, 44(1):1–18, February 1988.
14. Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, 1991.
15. Sanford J. Grossman and Joseph E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, June 1980.
16. Raph Levien. Advogato's trust metric. `http://www.advogato.org/trust-metric.html`.
17. Jeffrey K. MacKie-Mason and Hal R. Varian. Pricing congestible network resources. *IEEE Journal of Selected Areas in Communications*, 13(7):1141–1149, September 1995.
18. David Mazières and M. Frans Kaashoek. The Design, Implementation and Operation of an Email Pseudonym Server. In 5^{th} *ACM Conference on Computer and Communications Security (CCS'98)*. ACM Press, 1998.
19. Thomas R. Palfrey and Howard Rosenthal. Underestimated probabilities that others free ride: An experimental test. mimeo, California Institute of Technology and Carnegie-Mellon University, 1989.
20. J. F. Raymond. Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems. In H. Federrath, editor, *Designing Privacy Enhancing Technologies: Workshop on Design Issue in Anonymity and Unobservability*, pages 10–29. Springer-Verlag, LNCS 2009, July 2000.
21. Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50:97–110, 1982.
22. Andrei Serjantov and George Danezis. Towards an information theoretic metric for anonymity. In Roger Dingledine and Paul Syverson, editors, *Privacy Enhancing Technologies (PET 2002)*. Springer-Verlag, LNCS 2482, 2002.
23. Andrei Serjantov, Roger Dingledine, and Paul Syverson. From a trickle to a flood: Active attacks on several mix types. In Fabien Petitcolas, editor, *Information Hiding (IH 2002)*. Springer-Verlag, LNCS 2578, 2002.
24. Stuart G. Stubblebine and Paul F. Syverson. Authentic attributes with fine-grained anonymity protection. In Yair Frankel, editor, *Financial Cryptography (FC 2000)*, pages 276–294. Springer-Verlag, LNCS 1962, 2001.