Payment-based Incentives for Anonymous Peer-to-Peer Systems

Daniel R. Figueiredo[1] Jonathan K. Shapiro[2] Don Towsley[1] *

[1]Dept. of Computer Science
University of Massachusetts
Amherst, MA 01003
{ratton, towsley}@cs.umass.edu

[2]Dept. of Computer Science and Engineering
Michigan State University
East Lansing, MI 48824-1226
jshapiro@cse.msu.edu

Computer Science Technical Report 04-62 †

July 27, 2004

Abstract

Peer-to-peer anonymous communication systems are vulnerable to free-riders, peers that use the system while providing little or no service to others and whose presence limits the strength of anonymity provided by the system. To complicate matters, the identity of the free-rider is obscured by the very anonymity such systems are designed to provide, imposing challenging design constraints for incentive mechanisms that aim to discourage free-riding. In this paper, we address two aspects of cooperative behavior: participation and compliance. We propose a novel mechanism that creates financial incentives for peers to participate in the P2P system. This mechanism is based on the exchange of currency in return for service, and is implemented embedding small anonymous payments to each hop in the anonymous path. We also propose a complimentary retaliation mechanism to promote compliance with the anonymous protocol – peers that are not compliant are eventually identified and isolated. We present a simple analysis to quantify the effectiveness of the retaliation mechanism. Our mechanisms are particularly well suited for integration with anonymous protocols that use source routing and layered encryption, allowing the initiator of a message to safely embed payments for each peer along the path.

Keywords: anonymous communication, micropayment mechanism, incentive mechanism

1 Introduction

Despite widespread concern about the lack of privacy on the Internet, network applications such as the World Wide Web still fail to provide adequate privacy to their users. Some have recently argued [13, 2], that the

*This research has been supported in part by the NSF under grant awards ANI-0085848, ANI-0070067 and EIA-0080119, and by CAPES (Brazil). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

†Obsoletes Computer Science Technical Report 03-31

underlying economics of Internet-based services actually provide disincentives for protecting users' privacy. In this case, one option for privacy-concerned users is to rely on anonymous communication protocols to help guard their identity. However, the commercial availability of such an option is itself in doubt. Recent efforts to deploy a high quality commercial anonymity system based on the mix network architecture in [8] have failed, mainly due to high operational costs coupled with low user subscription rates [18]. An alternative to the commercial deployment of strong anonymity systems is the peer-to-peer (P2P) paradigm, where users that care about privacy collectively cooperate with each other in a decentralized manner.

Several P2P anonymous communication systems have been proposed in recent literature [26, 17, 27, 4]. In such systems, each user operates a peer that has functionality similar to a mix. The group of peers collectively provide anonymity by randomly forwarding each message among themselves an arbitrary number of times before sending it to its intended recipient, rendering the message initiator anonymous within the group of collaborating peers. The strength of anonymity provided by such systems can be characterized using any of a number of metrics that have been proposed to asses anonymous protocols—some related to the probability of correctly identifying a message initiator [26, 22, 20, 30], others related to the protocols resilience against specific types of attack [31, 32]. These metrics share the property of improving the strength of anonymity monotonically in the number of collaborating peers.

However, in P2P anonymity systems, collaboration is not guaranteed. Like many other P2P applications, anonymity systems are vulnerable to free-riders, peers that consume the service while providing little or no service to other peers in the system. Free-ridership has been observed in many P2P applications [1, 19, 6, 7]. In recognition of this problem, some researchers have recently suggested building incentive mechanisms into such applications in order to promote cooperative behavior. Within an anonymous P2P system, there are two aspects of cooperation that are of main concern. First, it is essential that peers are *compliant* with the protocol in the sense that they faithfully forward each others' traffic while joined to the system. Second, it is important that peers *participate* in the application by remaining joined to the group for extended periods of time. Mechanisms focused solely on enforcing compliance do not fully address the problem of free-riding in anonymity systems. In particular, free-riding may take the form of a fully compliant peer joining the system when it needs to establish anonymous communications, providing service while joined, but then leaving the system once its immediate needs are fulfilled. This behavior has the potential to undermine anonymity systems because of two undesirable consequences. First, the presence of free-riders tends to reduce the overall number of peers in the group at any particular point in time, which consequently reduces the strength of anonymity provided by the system. Second, the frequent turnover in group membership caused by free-riders joining and leaving the system imposes a high group maintenance overhead and can facilitate certain types of malicious attacks.

For an anonymous protocol to be scalable and effective, peers must be both compliant and participatory; yet there certainly are reasons not to do so. For example, the cost of dedicating local computer resources (e.g., network bandwidth, CPU cycles) to the system may limit a peers willingness to participate. Perhaps more insidiously, the perceived risk of scrutiny due to participation in an anonymous protocol may induce peers to willingly forward traffic within the group but not to its final destination.[1] Notwithstanding numerous examples of dedicated participants in such systems acting altruistically, many (if not most) users are likely to behave selfishly when presented with such costs and risks by becoming free-riders. In systems where the quality of the service provided depends on the number of participating nodes (as in anonymity), providing an explicit incentive to remain joined to the service is of great importance.

[1] In our terminology, such peers would be participatory, but not necessarily compliant.

In this work, we consider a novel technique that uses a payment mechanism to provide explicit incentives for peer participation along with a retaliation mechanism to promote compliance. This paper makes three specific contributions:

- We present a lightweight anonymous micropayment scheme with attractive properties, including payer-payee anonymity. Most previous anonymous payment schemes prevent the Bank from linking transactions to the payer. When used in conjunction with an anonymous communication protocol, our scheme also prevents the payee from identifying the payer.
- We incorporate our micropayment scheme into a class of mix-based anonymous communication protocols to provide positive financial incentives for participation in P2P anonymity systems.
- We introduce a mechanism for retaliation against peers that misbehave and do not comply with the protocol. The mechanism identifies and isolate malicious peers. We present a simple analysis to quantify its effectiveness. The retaliation mechanism is complementary to the payment mechanism we propose. Whereas the incentives provided by payments promote participation (i.e. joining and remaining joined to the group), the threat of retaliation assures that currently joined peers will operate in compliance with the proposed protocols.

The remainder of this paper is organized as follows. Section 2 provides a short overview of incentives mechanisms and payments schemes. In Section 3 we discuss why a payment mechanism is well suited for anonymity systems. Section 4 presents the proposed micropyament scheme followed by a discussion and a few optimizations. We discuss possible attacks on the enhanced P2P system in Section 5. In Section 6 we present the retaliation mechanism and an analysis of its effectiveness. Finally, Section 7 summarizes the paper.

Preliminaries

We will use the term *anonymous path* or simply *path* refer to the sequence of peers traversed by a message before it is forwarded to its intended destination.

We make two key assumptions about the P2P anonymous protocols to which our mechanisms can readily be applied. First, a message header destined to some intermediary peer along a path is only visible in plain text to that peer. Second, the initiator has complete knowledge of the constructed path. Note that mix-based P2P applications, such as Tarzan [17], MorphMix [27] and GAP [4], generally satisfy these requirements. Our mechanisms would not trivially apply to systems like Crowds [26], where the path is not known to the initiator.

We also assume that the majority of the peers in the system are self-interested and will respond rationally to well-defined incentives, such as financial incentives or incentives designed to provide better quality of service. However, we allow for the existence of malicious peers, who, from the perspective of the incentives, may appear to behave irrationally.

As customary in micropayment mechanisms, utmost security is not required as payments have small monetary value. Therefore, losses of small amounts can be tolerated by a peer and a small steady-state rate of monetary loss can be viewed as the financial overhead of using the system. In Section 6, we adopt this assumption and provide a mechanism to effectively reduce the rate of monetary loss.

2 Related Work

In order to mitigate the impact of free-riders, researchers have recently introduced explicit incentive mechanisms into P2P applications in different domains, with ad hoc routing having received much attention. The mechanisms proposed for promoting cooperation have focused on two approaches: (i) reputation mechanisms, whereby peers—individually or collaboratively—identify free-riders and punish them by declining service to peers with bad reputations [6, 12, 11, 28]; (ii) payment mechanisms, whereby peers exchange tokens in return for service [3, 7, 34]. All of these recent efforts introduce incentive mechanisms to promote compliance among participating peers. We are not aware of any work that adopts the specific notion of free-riding introduced here, which encompasses both noncompliance and nonparticipation.

Micropayment schemes provide an efficient mechanism to transfer money between two entities and several such schemes have been proposed in the literature over the past years (e.g. [29, 23, 33]). These schemes typically achieve simplicity at the expense of full anonymity and untraceability. Since the ultimate goal of the systems considered in our work is to provide anonymity to their users, we require a payment mechanism that is similar to anonymous digital cash [9, 5], which guarantees that payments cannot be linked to any specific payer. In addition, the application also requires a feature beyond that provided by anonymous digital cash, namely that the payer is also anonymous to the payee.

In [16], the authors propose a micropayment mechanism for mix networks where users of an anonymous communication system can pay for service. The proposed scheme integrates payment with the data forwarding to provide full anonymity and untraceability and shares some features with our proposed mechanism. However, their design targets a model in which mix operators wish to be compensated for providing anonymity service while users are clients of the system and do not participate in providing anonymity. Moreover, their scheme relies on tamper-proof hardware at the users. As we are concerned with large-scale P2P anonymity systems with infrastructures composed of individual users' computers, we must avoid designs that require specialized hardware. Finally, whereas their mechanism implicitly provides an incentive for mix operators to offer an anonymity service, our approach uses payments to provide explicit incentives for high peer participation and low group turnover in a P2P system.

3 Payment-based Incentives for Cooperation

The idea of exchanging tokens for service to encourage cooperative behavior among peers arises naturally from an attempt to define a mechanism that can operate effectively under the primary design constraints imposed by anonymous P2P systems. Recall that we are concerned here with a particular type of non-cooperative behavior where a peer uses the system to deliver its anonymous messages but provides little or no forwarding service to others, by exiting the system as early as possible. By requiring the exchange of a token in return for service, a peer requesting service can show concrete evidence of its past cooperation.

For such a mechanism to be robust against fraud, the token must be irrefutable, which typically requires it to be issued by a trusted third party, and peers must be prevented from reusing previously spent tokens. Furthermore tokens must be transferable in the following sense: a token presented to its issuer can be exchanged for a new token that the presenter may then use to acquire service. At the same time, such mechanism should be efficient and lightweight if it is expected to scale to large number of users.

Efficient mechanisms for issuing and exchanging such tokens already exist in the form of several payment and micropayment schemes developed over the last decade. An inherent tradeoff between security and efficiency exists for all micropayment mechanisms. Accommodating very small payments in practice usually requires off-line payment protocols, where the Bank is not contacted at the time of payment and fraud can only be detected after it occurs. This after-the-fact detection together with retaliation measures taken outside the system (such as banning a particular user) is often considered sufficient to discourage abuse as long as payments are traceable. For an application like anonymous communication, however, the identity of a user must be concealed at all times. To support such applications, we require a low overhead mechanism for anonymous micropayments that still discourages abuse. We present such a mechanism in Section 4.

Although we may intend for tokens to be transferred in exchange for service, we cannot prevent tokens from being transferred in exchange for other things, like real money. Such unintended transfers make tokens fungible.[2] Fungibility might first appear to be undesirable because it allows peers to be free riders by acquiring tokens for money without providing service for the system. Thus a payment-based scheme would not strictly prevent free-ridership. Such a scheme would, however, attach a real monetary cost to free-riding which can be avoided by providing service to others, thereby providing a financial incentive for cooperative behavior.

In addition to its incentive properties, allowing money to be exchanged for tokens provides a natural way to bootstrap the system. Peers who are willing to pay would become a source of new tokens, which would propagate into the system to be collected by cooperative peers and later used to acquire service free of charge. However, tokens can also leave the system, as a peer might collect tokens by providing service and simply "cash out". Although this might appear undesirable, the ability to cash out tokens could improve the quality of anonymity by attracting peers who have little desire for anonymity services themselves, but join the system to collect money. Such peers would otherwise never join the system and their presence could increase the overall number of peers providing a better anonymity service. A thorough investigation of this aspect of payment-based incentives is beyond the scope of this paper, but is an interesting area for further research.

4 Micropayment Mechanism

We propose a debit based micropayment mechanism that allows peers to provide anonymous untraceable payments to each other in exchange for message forwarding service. The protocol is off-line, meaning that the Bank is not contacted at the time of a transaction. Our micropayment mechanism is based on the concept of hash-chains and was inspired by the PayWord scheme proposed in [29] and has has similar features to that used in [16]. In short, the mechanism works as follows:

A peer P purchases a certificate for a given amount from the bank B. The certificate is signed blindly by the bank and has no explicit binding information to the identity of P. Peer P sends this certificate anonymously to a peer, say Q, that will be used as an intermediary mix. Peer Q checks the validity of the certificate using bank B's public key. To pay Q for its forwarding service, P includes a token in each anonymous message it sends through Q. This token is verified against the certificate. At a later point in time,

[2] A real-world example of this phenomenon is the well-documented trade in the virtual entities of online multi-player games [10], which has established a basis of exchange among virtual and real currencies (see, e.g. http://www.gamingopenmarket.com/).

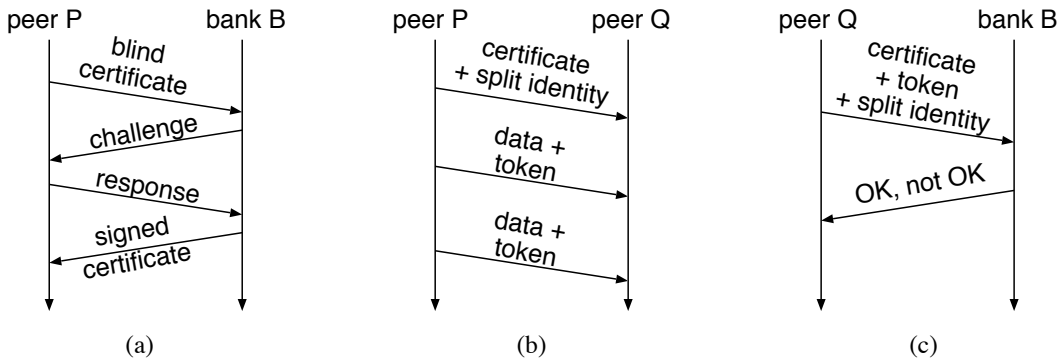

Figure 1: Messages exchanged during each interaction of the protocol: (a) purchasing a certificate, (b) making payments; (c) redeeming payments

peer Q presents the certificate as well as the last token received to bank B to redeem its payments. Bank B verifies the validity of the certificate and the tokens and credits the proper amount into P's account. The three types of interaction are illustrated in Figure 1.

There are two types of abuse that must be prevented in this scheme. First, the Bank clearly must avoid multiple redemptions of identical tokens, or *double payment*. This form of abuse can be prevented if Bank B maintains a record of the certificate and the last token redeemed. However, a malicious payer could still exploit the off-line nature of the payment mechanism to spend the same unit of currency with more than one payee, a practice known as *double spending*. To discourage double spending, we may borrow standard techniques used in anonymous digital cash schemes that use cryptographic payment protocols that reveal the identity of P (with high probability) if a certificate P purchases is used with more than one payee. This can be accomplished by binding the split identity information of P onto the certificate in a verifiable way. Therefore, requiring P to reveal half of its identity binding information when giving its certificate to Q, will also prevent Q from reusing the certificate, as Q does not know P's secrets.

The signing of the certificate described above must be done blindly, so that the bank cannot link the payment with the peer that purchased the certificate. An existing mechanism for obtaining blind signatures and binding the split identity information of a peer to the certificate, such as cut-and-choose [9] or single-term [5, 14], can be used.

The details of each interaction in the micropayment mechanism are described below.

4.1 Peer-Bank Interaction: Purchasing Certificates

A peer that desires to send messages through the system must first purchase signed certificates from a bank. To reduce the load on the bank, the peer is responsible for preparing the certificates, which will only be verified by the bank. The certificate generated by the peer contains a globally unique identification number (chosen randomly by P), its monetary value v, the committed split identity information of the peer and the final value of a hash chain of length v. In order to obtain the final value of a hash chain, the peer will have to randomly select a seed and recursively apply a well-known hash function to this value v times. This certificate is then sent to the bank over a secure and authenticated channel.

The bank verifies the account balance of P for enough funds and checks for the committed identity information as well as the uniqueness of the certificate identification. If the certificate is successfully verified, the bank withdraws the respective amount from P's account, signs the certificate *blindly* and sends it back to peer P. Note that the bank does not need to verify the hash chain, as generating an adequate chain of length v is in the peers' own interest. Finally, the peer receives the signed certificate and verifies the bank's signature. A diagram of this interaction is illustrated if Figure 1.a.

4.2 Peer-Peer Interaction: Making Payments

A peer will have to pay other peers in order to have its messages relayed through them. The first time peer P selects another peer, say peer Q, to be part of its anonymous path, it will send that peer one of its signed certificates. This message is sent through the anonymous path itself, and not directly, such that peer Q does not know the identity of P. The establishment of this bi-directional anonymous path will be discussed in Section 4.4.

Upon receiving the signed certificate, peer Q checks the bank signature and issues a challenge to peer P requesting half of its committed split identity. This challenge request traverses the reverse anonymous path. Peer P replies with the response to the challenge posed by Q. Peer Q verifies that the challenge is adequate and accepts to forward packets on behalf of P.

When P sends an anonymous message that goes through Q, it embeds a token (the next hash chain value) together with the certificate identification number in the message. This token is the payment for forwarding this message. Node Q verifies that the payment is valid by applying the well-known hash function to this token and comparing it with the previously received token (or the token in the certificate).

Peer Q must keep track of the last token and the number of tokens received. Note that the certificate only ensures payment of v tokens and, thus, Q should accept at most v tokens from P. Of course, node P also keeps track of its balance with node Q and must send a new new signed certificate when its balance reaches zero if it intends to continue to use Q to forward its messages.

For improved efficiency, the challenge-response interaction to obtain half of the split identity of P can be removed. Node P can define the challenge by computing the hash of a message that contains the certificate concatenated with the identity of Q. This challenge value cannot be easily predetermined and can be easily verified by Q. Peer P sends the response to this challenge value together with the certificate to peer Q, which can then verify the validity of the challenge/response pair. This interaction is illustrated if Figure 1.b.

4.3 Peer-Bank Interaction: Redeeming Payments

A peer that provides forwarding service to others will accumulate certificates and corresponding tokens. These tokens should be redeemed at the bank that signed the certificate periodically, but preferably when the total number of tokens guaranteed by a certificate has been received. To redeem its payments, peer Q sends the corresponding bank the certificate, the committed split identity information of peer P, the last token received for that certificate and the total number of tokens received since the last time it redeemed this certificate.

The bank checks its own signature in the certificate. It then verifies if the token being redeemed is valid, by recursively applying the well-known hash function to the token until the final hash value printed in the certificate is obtained (or until it matches the last token redeemed associated with this certificate). The number of hash function calls should match the value claimed by Q. The bank then credits the proper the amount of money into Q's account. A diagram showing this interaction is presented in Figure 1.c.

As discussed above, the bank maintains a record of the certificate and of the last token redeemed with it to guard against paying peer Q more than once. The bank should also keep the split identity information of associated with the certificate to detect possible double spending by the initiator.

For efficient use of storage space at the banks, the certificates should be redeemed by a given date. This expiration date can be stamped in the certificate during its creation by peer P. The bank would only redeem tokens to certificates that have not been expired, and would only need to maintain state until that date.

4.4 Communicating Anonymously

The proposed micropayment scheme readily applies to P2P anonymous systems that are based on Chaumian mixes [8], such as the systems described in [17, 27, 4]. Two fundamental characteristics of Chaumian mixes that are required by our scheme are the recursive encryption and source routing. We first review the basic functionality of mix networks.

In a simplified mix network, a peer wishing to send an anonymous message M (known as the *initiator*) to some destination D, constructs a path through a set of collaborating peers in the system. The last peer on this path is responsible for forwarding the message to its ultimate destination D. The initiator selects the intermediary peers of the path, possibly at random, and using their respective public keys constructs a message that is recursively encrypted and has the form:

$$O = S_1, \{S_2, \{S_3, \{\ldots, \{S_L, \{M, D\}_{K_L^+}\}_{K_{L-1}^+}\}\ldots\}_{K_2^+}\}_{K_1^+} \quad (1)$$

where S_i is the address of the i-th peer in the path, L is the path length, and $\{X\}_{K_i^+}$ denotes message X encrypted with public key K_i^+.[3] In onion routing [25], this recursively encrypted message is known as an *onion* and we adopt this nomenclature in the subsequent text.

After constructing, the onion, the initiator then forwards it to the first peer of the path, S_1. Each intermediary peer i in the path will have access to a payload after decrypting the message with its private key. The payload contains the address of the next hop S_{i+1} and an encrypted payload to be passed to this next hop. At each hop, state is installed at the peers to allow for future messages and/or messages traversing the reverse path. Eventually, the message reaches the last node in the path, which then forwards message M to destination D. The peers forward any response from D along the reverse path towards the initiator.

Our proposed scheme will embed either a certificate or a token to each hop of the anonymous path. If an initiator is requesting service from a peer for the first time or if all tokens associated with the certificate currently held by the peer have been spent, then a new certificate is embedded. Otherwise, a single token

[3] For clarity of presentation, we describe a simplified encryption mechanism for the onion. Real protocols carry more information on their onions and rely on public keys only initially to distribute symmetric keys, which are used thereafter.

will be embedded. Thus, the payload received by node i is

$$P_i = \{S_{i+1}, P_{i+1}, C_i\}_{K_i^+} \tag{2}$$

where C_i is either a token or a certificate to the i-th hop of the path.[4]

Note that the inherent source-routing mechanism (initiator determining the anonymous path) provided by the above protocol is particularly well suited for integration with a payment mechanism, as the initiator can safely embed a payment for each hop in the path. Moreover, the structure of the onion ensures that the certificate or token destined to a given peer is visible only to that peer and no other peer.

4.5 Discussions and Optimizations

Once assigned to a given peer, a certificate cannot be redeemed by the initiator. To take full benefit of its value, the certificate must be used until its balance reaches zero. Although one can propose a mechanism for redeeming the value remaining in a certificate, we opt for simplicity and discard this possibility. Note, however, that peer P has control of the value of the certificate it generates. This can be used to gage the quality of service provided by a given intermediary peer Q. A certificate of low value can initially be issued and given to peer Q and as its credibility increases, P can start to assign certificates of higher values. Of course, there is a risk that peer Q will suddenly refuse service (or simply leave the system) and peer P will lose its money. However, there is little incentive for Q to refuse service, as Q cannot obtain the remaining balance from the certificate. Section 6 describes a mechanism to counteract against malicious peers that consistently refuse to forward messages in exchange for tokens.

If an unreliable communication medium is used for exchanging messages, it is possible that either a token or certificate is lost during transmission. However, the micropayment mechanism proposed is resilient to losses of certificates and tokens. If a message containing a token to intermediate node Q is lost, then peer Q can recover that token in the subsequent message it receives, by simply using the subsequent token and applying the hash function. If a certificate is lost, then the initiator can retransmit the same certificate to the same intermediate peer without sacrificing anonymity. Such properties are useful when dealing with a scenario where messages can be lost.

The micropayment scheme above provides payment from the initiator to all intermediate peers along the forward path—that is, the path taken by the message as it travels to the destination. Thus far, we have not provided payments for response messages that originate at the destination and follow the reverse path to the initiator. Although we could assume that payments in the forward path provide sufficient incentives for intermediary peers to relay back response messages, it is also possible that explicit payments will be required for forwarding these responses.

There are several ways in which payment for response messages can be incorporated. One particular mechanism, would be for the initiator to provide payments for reverse messages in a separate recursively encrypted message after it has received the response. This delayed payment can perhaps be piggybacked

[4] In earlier versions of this work, the payment for the i-th hop was encrypted with a symmetric key to be returned in an acknowledgment by the subsequent hop, thereby requiring hop i to forward the message in order to be paid. We now rely on the retaliation mechanism introduced in Section 6 to promote compliant behavior, enabling us to remove the acknowledgment scheme and its associated overhead.

in a subsequent request. This method has the advantage of allowing the initiator to pay in proportion to the size of the response. However, this mechanism requires intermediate peers to trust the anonymous initiator in paying for the response messages after they have been forwarded back. If an initiator fails to pay, intermediate peers can punish it by dropping subsequent requests and forcing the initiator to create a new path. The creation of a new path is known to reveal information about the initiator's identity [32]. Thus, initiators who presumably value anonymity, therefore have an incentive to behave honestly with respect to issuing payments.

The bank is likely to impose a fee for its service of distributing and redeeming certificates to peers. It would be adequate then, to have a transaction fee charged to peer Q whenever it redeems tokens for money. This would also instimulate peer Q from going to the bank too frequently and encourage it to accumulate enough tokens before redeeming them.

In terms of efficiency, the payment mechanism proposed above will clearly introduce additional communication and computation overheads to the anonymity system. This overhead will materialize in the form of delays when communicating anonymously. In a system where peers' computational resources are plentiful, such delays will be dominated by the communication costs. It is possible to hide part of this overhead by executing tasks in parallel or ahead of time, such as purchasing certificates in batches well prior to establishing anonymous communication. A more serious issue is the need of a trusted Bank in the payment mechanism. However, we argue that the Bank should be a publicly accessible authority outside the P2P system that issues digital cash and certificates for other purposes besides the anonymity system. Users that intend to join the system would then be required to open an account with the bank.

5 Attacks on Anonymity

When augmenting an anonymous protocol with a payment mechanism, as we have proposed, an important concern is whether new attacks on anonymity and on the system itself can be enabled by the payment mechanism. Clearly, the Bank is a new potential attacker and may even collude with peers in the system.

We argue that the Bank alone will not trivially be able to reveal the identity of an initiator. The anonymity properties of digital cash prevent the Bank from linking a signed certificate to the payment made to some peer. However, a more sophisticated Bank might use traffic analysis to correlate the issue of certificates to a given peer with later redemptions by other peers. To counteract traffic analysis, an initiator might buy large quantities of certificates divided into a few face values, as opposed to requesting the Bank to issue certificates on a per session time-scale. However, such traffic analysis seems no easier than launching a predecessor attack against a node in the absence of a payment mechanism.

Furthermore, even if the Bank colludes with one or more nodes the initiator is not trivially revealed. Since the Bank cannot add any definite information to what colluding nodes might already know, at best they can join efforts in performing traffic analysis. In this case, traffic analysis can be stronger as more events might be correlated. If the Bank is, as we advocate, a publicly accessible authority providing digital cash services, performing coordinating payments with other system events is potentially more difficult, as transactions from the anonymous protocol peers will be interleaved with unrelated transactions.

We argue that a payment-based incentive mechanism improves the quality of the anonymity service

offered by the system. Intuitively, this mechanism will reduce the number of free-riders that respond to financial incentives, by increasing the participation level of such peers. This increase in participation will lead to a higher average number of peers in the system and reduce the average turnover rate. These effects have direct positive impact in the quality of anonymity the system offers. For example, attacks on anonymity, such as the predecessor attack and the intersection attack, will require more effort from the attackers in order to be successful. These attacks are known to depend directly on the average group size and turnover rate [32]. Last, as we have more rigorously argued in [15], adding payment-based incentives for participation has the effect of increasing the average number of peers in the system.

Finally, in a payment-based system it is possible that malicious users would wish to exploit the system for financial gain, thus new attacks designed to obtain digital cash from the system can emerge. Although we are not aware of any trivial and effective attack, it is possible that by colluding with each other, a set of malicious users could pose such a threat to the system. The effectiveness of such attacks is inherently limited if the anonymous path is defined solely by the initiator, reducing the chances of colluding nodes appearing in the same path. Systems where the anonymous path is defined cooperatively are more threatened by such collusion. Due to the threat that malicious paths pose to anonymity, however, such systems typically include mechanisms for detecting [27] or limiting [17] such collusive behavior.

5.1 Malicious Behavior

The problem of exchanging digital products (e.g., certificates for message forwarding) between two parties that do not necessarily trust each other has been widely studied over the years. Protocols that guarantee the success of such transactions are known as fair exchange protocols and a number of them have been proposed in the literature [21, 24]. Efficient protocols that avoid after-the-fact disputes rely on a common trusted third party. Protocols that do not use a trusted third party can have a high communication overhead or unreasonable assumptions (such as requiring the parties to have identical computing power), limiting their practicality [24]. Moreover, fair exchange protocols are even more complicated if anonymity is required among the parties.

Although the system proposed here could potentially make use of an existing fair exchange protocol that provides anonymity among the parties, we choose instead, for the sake of system simplicity, to accept a certain amount of inherent lack of trust among the peers involved in a transaction. Lack of trust can be tolerated because a peer that has participated in an unfair transaction can retaliate against the misbehaving peer. The possibility of retaliation will motivate rational peers to behave fairly when engaging in a transaction. However, if a peer is irrational and acts maliciously, then the retaliation mechanism should identify and isolate this peer. Such a mechanism is described in Section 6.

6 Identifying and Isolating Malicious Peers

The micropayment mechanism previously described assumes that intermediary peers will correctly forward messages in exchange for tokens. Therefore, once a certificate is given to a peer, the initiator hopes to make full use of its certificate by selecting that peer as an intermediary hop for future messages. However, as noted earlier, it is possible that some peers might not be willing to cooperate under any financial incentives

or simply be malicious. Although a peer might gain a few tokens without forwarding any messages, a careful initiator would attempt to minimize its exposure to such malicious peers by monitoring the quality of service received and avoiding the use of high-value certificates on untrusted peers. Thus, we propose a simple mechanism where an initiator can identify and isolate malicious peers. By identifying malicious peers, an initiator will make use only of peers that are willing to cooperate by forwarding messages in exchange for tokens. Peers that are identified as malicious will not be requested for service. Thus, malicious peers will not receive any tokens from other peers and will have to pay the full price in case they use the anonymity service. Note that peers that are rational and respond to financial incentives will be cooperative to reduce their cost when using the system. Moreover, an isolated peer will not attract any traffic to itself thereby placing its own anonymity at risk, as observed by Bennett and Grothoff [4]. A peer that cooperates will attract traffic from others, which can serve as cover traffic and help disguise the messages itself generates.

The proposed scheme does not completely avoid the loss of certificates, but merely reduces the consistent loss of certificates due to uncooperative peers. Failures of network links or nodes as well as occasional encounters with uncooperative nodes may lead to broken paths and unrecoverable certificates. However, the decision of how much risk assume in this regard remains in the hands of the initiator, who may decide on the value of the certificates that are outstanding in the network at any time. In particular, a risk averse initiator will choose to create low value certificates at the expense of increased interaction with the Bank. It may be possible to deploy a mechanism to completely avoid loss of certificates (e.g., a fair exchange protocol), but these tend to be rather complex and require trusted parties. Since certificates can have a very small value, we believe that a simple lightweight mechanism that prevents consistent losses is more suitable for a P2P system.

In our proposed scheme, each initiator individually establishes the reputation of other peers in the system and stops requesting service from them when reputation is low. Note that the initiator can observe the aggregate behavior of the intermediary peers it selects to form its anonymous path. If a message fails to reach the destination (or an expected response fails to return) the initiator decreases the reputation of all peers in that path. Note that it is also possible to have an initiator learn the reputation of other peers indirectly through a third party it interacts with, as discussed in [6]. A mechanism similar to the one in [6], that uses second-hand information to improve the convergence of peers' reputation, could be coupled to our mechanism.

To map misbehavior to a reputation value, each initiator j will maintain a set of counters for the number of times each peer has failed to cooperate. Let $N_i^j(T), i = 1, \ldots, n$ be j's counter value for peer i after T paths have been constructed, where n is the total number of peers in the system. When a path is created and the initiator suspects that a peer in the path is failing to cooperate, then the counter for all peers in that path is incremented by one.

A simple rule to identify that a given peer is misbehaving, is to define a threshold for the counters and claim that a peer misbehaves if its counter is above that threshold. Let H be a threshold value. Then peer j classifies peer i as misbehaving if $N_i^j(T) \geq H$. To isolate misbehaving peers from new anonymous paths, an initiator should only select intermediary peers that are classified as well behaved peers. Note that by selecting from this reduced set of peers, an initiator will perceive better service, as it will be less likely to loose certificates or to have its messages discarded.

It is possible that an initiator will wrongly classify a peer that is willing to cooperate as a malicious peer. This can occur if a well-behaving peer is selected by an initiator to be on a path together with malicious peers

too many times. In this case the counter for the well-behaving peer will exceed the threshold value H and the peer will be classified as being malicious. To compensate for this potential misclassification a redemption mechanism should be deployed. For example, the reputation value of a given peer could improve with time (by reducing the counters N_i^j that are greater than zero). The improvement rate can be peer dependent and adaptive to reflect the past behavior of this peer. For example, a peer that has shown to be consistently misbehaving should have a much lower redemption rate than a peer who has only just been classified as misbehaving. A periodic redemption mechanism was successfully used in [6] to reintegrate nodes that were misclassified, and we believe a similar mechanism can be readily adopted in our system.

6.1 Analysis of Identification Mechanism

We now perform a simple analysis on the effectiveness of the identification mechanism presented above. We assume that system membership is static and that paths are constructed randomly and independent from each other. Moreover, we will not consider any redemption mechanism. In support for this last assumption note that any redemption mechanism should operate in a time-scale that is much larger than the rate at which paths are created.

Let the anonymous communication system contain n peers, out of which m are misbehaving. Without loss of generality, let peers $1,\ldots,m$ be the misbehaving peers. Let $p_i, i = 1,\ldots,n$ be the probability that an initiator selects peer i to be on a given path. We assume that every time a misbehaving peer is selected to be on a path it misbehaves (e.g., refuses to forward a message) and that such behavior is perceived by the initiator (i.e., will have its counter value increased). Moreover, assuming that each path is composed of L intermediary peers and that each peer is equally likely to be selected for a path, the probability that a given peer is selected to be on a path is $p_i = 1-(1-1/n)^L, i = 1,\ldots,n$.

Recall that peer k will be classified as misbehaving by peer i if its counter $N_k^i(T)$ reaches the value of H. The probability a misbehaving peer k is indeed classified as a misbehavior is given by the probability peer k is selected at least H times. The probability peer k is selected to be on any given path is simply p_k. Thus, the probability it is classified as a misbehaving is given by the binomial distribution:

$$\Pr[N_k^i(T) \geq H] = \sum_{t=H}^{T} \binom{T}{t} p_k^t (1-p_k)^{T-t} \tag{3}$$

Where peer $k = 1,\ldots,m$ is a misbehaving peer. Note that this probability approaches one as T becomes large, independent of the value for the threshold H. Thus, given enough path creations, it is always possible to positively identify a misbehaving peer.

Figure 2 illustrates the probability of identifying a malicious peer as a function of T for different values of the threshold[5]. As noted, this probability converges to one as T becomes large. We also observe that using a lower threshold value will more quickly identifies the malicious peer.

It is also instructive to understand how initiators lose certificates due to malicious peer. Initially, an initiator loses a certificate whenever it chooses a malicious peer for its path. However, as malicious peers are identified and no longer selected to form paths, this probability goes to zero. This is illustrated in Figure

[5]The numerical analysis presented has the following parameters: $n = 100$, $c = 10$, $L = 5$. Results using other parameters showed similar trends.

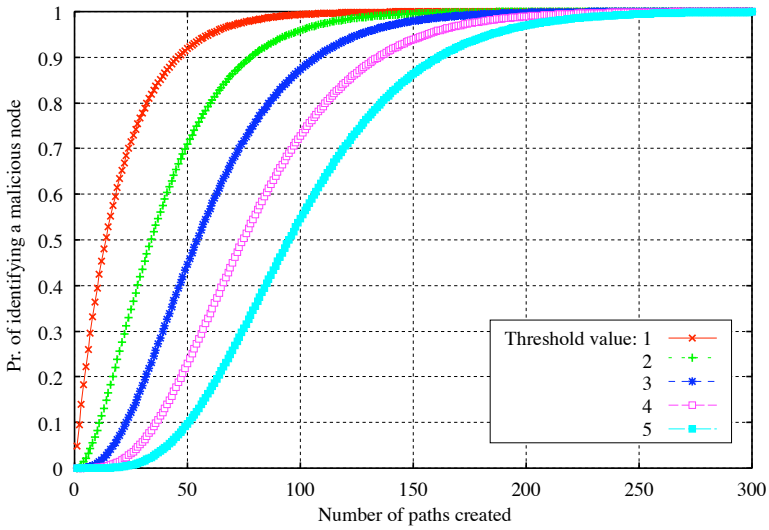

Figure 2: Probability of identifying a malicious peer as a function of the number of paths created for different values of the threshold.

3, which shows the probability of losing a certificate as a function of T for different values of the threshold. Note that with 10% of the peers being malicious and path length of 5, the probability of losing a certificate is less than 0.05 after 50 paths are constructed by the initiator.

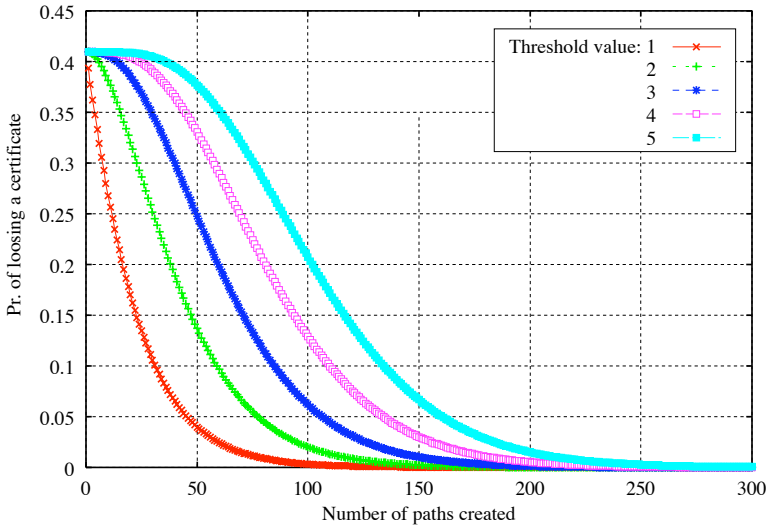

Figure 3: Probability of losing a certificate to a malicious peer as a function of the number of paths created for different values of the threshold.

As noted earlier, a well-behaving peer k will be wrongly classified as a misbehaving peer if its counter reaches H. This will only happen if k appears together with any misbehaving peer in at least H paths. Moreover, k will only appear together on a path with a misbehaving peer j if j has not yet been identified as misbehaving (otherwise the initiator will no longer select this peer). Note that once all malicious peers are identified, a well-behaving peer can no longer be misclassified.

Thus, we are interested in computing the false positive probability, that is $\Pr[N_k(T) \geq H], k = m+1, \ldots, n$. Different from the previous probabilities, this expression is not trivial to compute for the general case, since misbehaving peers that are classified as such are eliminated from the path selection process. Instead, we obtain an upper bound for the probability of this event by assuming that at most one malicious peer can be on a given path. Under this new model, we can write a recursion on T for the probability $\Pr[N_k(T) = H]$ by keeping track of the number of malicious peers that have been identified. This recursion can be solved numerically to give the upper bound for $\Pr[N_k(T) \geq H]$. Intuitively, the false positive probability should converge to a value less than one, as all malicious peers are eventually identified.

Figure 4 illustrates the false positive probability as a function of T for different threshold values. We observe that the upper bound for this probability converges to a value away from one, as expected. We also note that increasing the threshold value has a dramatic impact on the false positive probability, becoming very small (less than 0.01) with a value of $H = 3$.

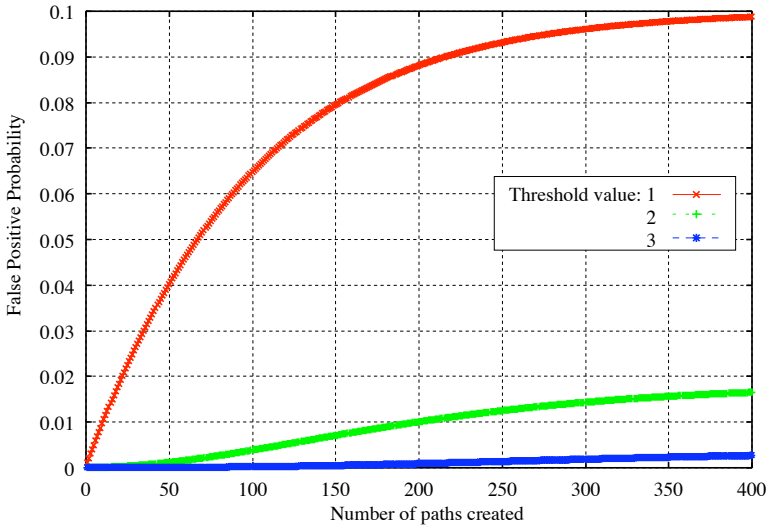

Figure 4: False positive probability as a function of the number of paths created for different values of the threshold.

The results above indicate the existence of a tradeoff on the threshold value. By using a low threshold the initiator can identify malicious peers faster and, thus, obtain better service (e.g., loose less certificates). However, a low threshold also increases the chances of wrongly classifying a well behaving peer as malicious. Wrongly classifying a good peer as malicious will unnecessarily reduce the number of peers an initiator can use as intermediaries, which can potentially impact the quality of its own anonymity. Moreover, it also penalizes the good peer as it will receive less payments and attract less traffic.

However, although one initiator might wrongly classify a good peer with a non-negligible probability, it is unlikely that all other peers wrongly classify that *same* good peer as also being malicious. In fact, the expected fraction of good peers that wrongly classify a given good peer as malicious is also given in Figure 4. As can be observed, at most 10% of the good peers will wrongly classify any given good peer as malicious when the threshold value is 1. Therefore a well-behaving peer will loose at most 10% of incoming payments and traffic (if all peers are identical). It seems the use of a low threshold value (e.g., $H = 1$), is more effective overall. Moreover, the misclassification will also be remedied by the redemption mechanism

that should run in parallel with the classification mechanism.

7 Summary

Providing explicit incentives to foster cooperation in P2P applications is emerging as an effective technique to combat free-riders. In P2P anonymous systems the notion of free-riding takes a new dimension as peers can fail to participate or fail to comply with the protocol. Free-riding can have a significant negative effect on the strength of anonymity provided by the system, as most users are likely to act selfishly.

An incentive mechanism where the initiator must pay for service and peers receive payments for forwarding messages will provide clear incentives for peers that, despite being self-interested, value money. To reduce the cost of using the system rational peers will be more cooperative, as free-riding has now a real monetary cost.

In this paper we have addressed the two aspects of free-riding enumerated above by presenting a payment-based incentive mechanism that coupled with a retaliation mechanism provides incentives for rational peers to cooperate. Our proposed mechanisms builds on primitives from micropayment systems and reputation systems that have been proposed in the literature, and can be readily applied to several anonymous P2P systems proposed, such as [17, 27, 4].

Although explicit incentives may be needed to drive rational users to take desired actions, it is worth mentioning that some anonymous protocols appear to have built-in incentives for cooperation that derive from specific protocol behavior. Such incentives are triggered solely from each individual peer's desire for stronger anonymity service. For example, in GAP [4], each peer should attract traffic from others in order to obscure its own communications. Similarly, in MorphMix [27], the collusion detection mechanism requires each peer to continuously construct anonymous paths. Such protocol-specific incentives and its relationship to explicit general purpose incentives are intriguing and worth of further study.

Acknowledgments

We thank Sonja Buchegger for fruitful comments in early drafts of this work and discussions on schemes to detect misbehaving peers in anonymity systems.

References

[1] Eytan Adar and Bernardo A. Huberman. Free riding on gnutella. *First Monday*, 5(10), 2000.

[2] Ross Anderson. Why information security is hard - an economic perspective. In *Proc. 17th Annual Computer Security Applications Conference*, New Orleans, LA, December 2001.

[3] B. Wilcox-O'Hearn. Experiences deploying a large-scale emergent network. In *Proceedings of the First International Workshop on Peer-to-Peer Systems (IPTPS '02)*, Cambridge, MA, March 2002.

[4] Krista Bennett and Christian Grothoff. Gap – practical anonymous networking. *Proc. Workshop on Privacy Enhancing Technologies (PET)*, 2003.

[5] Stefan Brands. Untraceable off-line cash in wallets with observers. In *Proc. on Advances in cryptology (CRYPTO'93)*, volume 773, pages 302 – 318. Springer-Verlag, 1995.

[6] Sonja Buchegger and Jean-Yves Le Boudec. The effect of rumor spreading in reputation systems for mobile ad-hoc networks. In *Proc. WiOpt'03 (Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks)*, 2003.

[7] Levente Buttyan and Jean-Pierre Hubaux. Stimulating cooperation in self-organizing mobile ad hoc networks. *ACM/Kluwer Mobile Networks and Applications (MONET)*, 8(5), October 2003.

[8] David Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–90, 1981.

[9] David Chaum, Amos Fiat, and Moni Naor. Untraceable electronic cash. In *Proc. on Advances in cryptology (CRYPTO'88)*, volume 403, pages 319–327. Springer-Verlag, 1990.

[10] Julian Dibbell. Unreal estate boom, or, the 79th richest nation on earth doesn't exist. *Wired Magazine*, December 2002.

[11] Roger Dingledine, Nick Mathewson, and Paul Syverson. Reputation in P2P anonymity systems. In *Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, 2003.

[12] Roger Dingledine and Paul Syverson. Reliable mix cascade networks through reputation. In *Proc. Sixth International Financial Cryptography Conference - FC02*, Mar 2002.

[13] Joan Feigenbaum, Michael Freedman, Tomas Sander, and Adam Shostack. Economic barriers to the deployment of existing privacy technology. In *Proc. Workshop on Economics and Information Security*, Berkeley, CA, May 2002.

[14] Niels Ferguson. Single term off-line coins. *Advances in Cryptology—EUROCRYPT '93, Lecture Notes in Computer Science*, 765:318–328, 1994.

[15] Daniel R. Figueiredo, Jonathan K. Shapiro, and Don Towsley. Incentives for cooperation in anonymity systems. Technical Report UM-CS-2003-031, University of Massachusets at Amherst, Dept. of Computer Science, 2003.

[16] Elke Franz, Anja Jerichow, and Guntram Wicke. A payment scheme for mixes providing anonymity. In *Proc. Trends in Distributed Systems for Electronic Commerce (TREC'98)*, volume 1402 of *Lecture Notes in Computer Science*, pages 94 – 108. Springer-Verlag, 1998.

[17] Michael J. Freedman and Robert Morris. Tarzan: A peer-to-peer anonymizing network layer. In *Proc. of the 9th ACM Conference on Computer and Communications Security (CCS 2002)*, Washington, DC, November 2002.

[18] Ian Goldberg. Zeroknowledge to discontinue anonymity service. Slashdot. URL: http://slashdot.org/comments.pl?sid=22261&cid=2388977, October 2001.

[19] Philippe Golle, Kevin Leyton-Brown, Ilya Mironov, and Mark Lillibridge. Incentives for sharing in peer-to-peer networks. *Lecture Notes in Computer Science*, 2232, 2001.

[20] Yong Guan, Xinwen Fu, Riccardo Bettati, and Wei Zhao. An optimal strategy for anonymous communication protocols. In *Proc. 22nd IEEE International Conference on Distributed Computing Systems (ICDCS 2002)*, Jul 2002.

[21] Steve Kremer, Olivier Markowitch, and Jianying Zhou. An intensive survey of non-repudiation protocols. *Computer Communications Journal*, 25(17):1606–1621, 2002.

[22] Brian N. Levine and Clay Shields. Hordes: A protocol for anonymous communication over the internet. *ACM Journal of Computer Security*, 10(3), 2002.

[23] Tomi Poutanen, Heather Hinton, and Michael Stumm. NetCents: A lightweight protocol for secure micropayments. In *Proc. of the Third USENIX Workshop on Electronic Commerce*, pages 25–36, 1998.

[24] Indrajit Ray and Indrakshi Ray. Fair-exchange in e-commerce. *SIGecom Exchanges*, 3.2:9–17, 2002.

[25] Michael G. Reed, Paul F. Syverson, and David M. Goldschlag. Anonymous connections and onion routing. *IEEE Journal on Selected Areas in Communication Special Issue on Copyright and Privacy Protection*, 1998.

[26] Michael K. Reiter and Aviel D. Rubin. Crowds: anonymity for Web transactions. *ACM Transactions on Information and System Security*, 1(1):66–92, 1998.

[27] Marc Rennhard and Bernhard Plattner. Introducing morphmix: Peer-to-peer based anonymous internet usage with collusion detection. In *Proc. of the Workshop on Privacy in the Electronic Society (WPES)*, Washington, DC, November 2002.

[28] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. Reputation systems. *Communications of the ACM*, 43(12):45–48, 2000.

[29] Ronald L. Rivest and Adi Shamir. Payword and micromint–two simple micropayment schemes. In *Proc. International Workshop on Security Protocols*, volume 1189 of *Lecture Notes in Computer Science*, pages 69 – 87. Springer-Verlag, 1996.

[30] Andrei Serjantov and George Danezis. Towards an information theoretic metric for anonymity. *Proc. Workshop on Privacy Enhancing Technologies (PET)*, 2482, 2002.

[31] Matt Wright, Micah Adler, Brian N. Levine, and Clay Shields. An analysis of the degradation of anonymous protocols. In *Proc. ISOC Network and Distributed System Security Symposium (NDSS 2002)*, Feb 2002.

[32] Matt Wright, Micah Adler, Brian N. Levine, and Clay Shields. Defending anonymous communication against passive logging attacks. In *Proc. of the IEEE Symposium on Security and Privacy*, Oakland, CA, May 2003.

[33] Beverly Yang and Hector Garcia-Molina. Ppay: micropayments for peer-to-peer systems. In *Proc. of the 10th ACM Conference on Computer and Communication Security*, pages 300–310. ACM Press, 2003.

[34] S. Zhong, J. Chen, and Y. Yang. Sprite: A simple, cheat-proof, credit-based system for mobile ad hoc networks. In *Proc. of Infocom 2003*, 2003.