Incentive-Driven P2P Anonymity System: A Game-Theoretic Approach

Souvik Ray[1], Giora Slutzki[2] and Zhao Zhang[1]

[1]Department of Electrical and Computer Engineering
Iowa State University
Ames, IA 50011

[2]Department of Computer Science
Iowa State University
Ames, IA 50011

Abstract–

Anonymous communication systems built on P2P infrastructures using anonymity forwarders are frequently affected by the churn problem, i.e. frequent joins and leaves of nodes. The problem unavoidably affects the quality of provided anonymity: The availability of anonymity forwarders will be decreased, which reduces the anonymity set; and the frequency of path reformation will increase, which increases the chance of successful intersection attacks. We propose an incentive-based P2P mechanism as an approach to providing reliable anonymity forwarding. It uses incentives to induce the peer nodes to provide anonymity forwarding as reliable service and to make *stable and distributed* forwarding decisions to minimize the frequency of path reformations. To support incentive, a payment system has been designed which meet the anonymity requirement and can handle typical scenarios of cheating and malicious attacks. To make sound forwarding decisions, we use game theory to carefully design the forwarding strategies used by the peer nodes. We have used event-driven simulations to evaluate the quality of anonymity provided by the mechanism under high churn and with the presence of malicious nodes. The results show that the quality of anonymity is maintained in those scenarios.

1. Introduction

Many anonymity protocols and systems have been developed in recent years to support anonymous communications (initiator anonymity, responder anonymity or unlinkability). Most of those systems assume some forwarding insfrastructure, e.g. trusted forwarding servers or a large set of P2P forwarding nodes. Nevertheless, there lacks research on the insfrastructure itself. It is not a trivial issue: The high operational cost of trusted infrastructure has resulted in the commercial failure of such systems, e.g. the Freedom Network [16]. P2P-based forwarding systems [21, 14] is more commercially viable because they use unreliable and untrusted forwarding nodes, and has the advantage of using non-centralized forwarding nodes. However, P2P-based forwarding systems are affected by the churn problem, i.e. the frequent leaves and joins of nodes. It is common in P2P systems because of free riding [11], a scenario that many nodes join a P2P system for a short time to enjoy its benefit but not to provide the expected service. There are two negative consequences for P2P-based forwarding systems: It affects the availability of forwarding nodes, which reduces the size of anonymity set; and it forces frequent reformations of forwarding paths, which make the system vulnerable to intersection attacks [27].

To address the churn problem, incentive mechanism has been introduced to induce P2P nodes to provide stable service [13]. Nevertheless, an incentive mechanism for anonymous forwarding must consider the quality of anonymity beyond the stability of service. In a simple incentive mechanism, a forwarder may align its routing decisions to maximize its local interests; for example, to minimize communication costs. Additionally, the churn problem may only be alleviated – even with incentive, new nodes may continue to join and old nodes may leave in typical P2P systems. The problem of frequent forwarding path reformation still exists and should be considered in the design of the incentive mechanism.

In this study, we propose an incentive mechanism that induces the forwarders to make forwarding decisions aligned with the quality of initiator anonymity. The objective is achieved by binding the incentive received by a local node with the quality of initiator anonymity at the system level – a local node may maximize its interests by using a routing strategy aligned with the goal of anonymity. To have a sound foundation, we use game theory to design and analyze the forwarding strategy. We also propose payment-based incentive mechanism that keeps the anonymity of involved parties.

The rest of the paper is organized as follows. The next section describes the incentive-based forwarding and routing model. Section 3 presents the experimental results and Section 4 discusses the related work. Finally Section 5 concludes this study.

2. Incentive-based Forwarding and Routing Model

2.1. Motivation and Design Objectives

P2P and forwarding-based anonymity systems are primarily distinguished by their routing and forwarding infrastructure. In Onion routing [20] and MIX-based systems [8], the routing is done before the forwarding and therefore a forwarder merely performs the forwarding. In Crowds [21], a forwarder makes the routing decision. There is a hidden-action problem [12], i.e. the actions of the forwarders are hidden from the initiator and the quality of the path is decided by the decisions of the forwarders, which is not necessarily aligned with the quality of anonymity. We believe that a properly designed incentive mechanism is required to ensure appropriate forwarding and routing of packets in such systems. Designing such a mechanism for an anonymity system, however, is challenging because the mechanism itself cannot leak the identity information.

Intersection attack is a serious concern for anonymity systems that are used by applications with recurring activities; for example, those using HTTP, FTP, NNTP or raw sockets [26]. In those applications, an initiator usually makes repeated connections to a set of specific responders. The reformations of the forwarding path, which can be caused by frequent node joins and leaves, will increase the chance of exposing the initiator and the responder to the intersection attacks. In an intersection attck, the attacker observes the intersection of the sets of active nodes at different times and may find out the initiator or responder by reducing the intersection set.

The availability of forwarders has a strong impact on the success rate of intersection attack. In a P2P system, the availability of a peer node can be expressed as the ratio of the sum of its sessions times to its lifetime, where the lifetime is from the time of the initial entry of the peer node into the system to the time of its final departure, and a session time is the time between the arrival and the departure during a single session [22]. Consider the anonymous forwarding from an initiator I to a responder R. The higher the availability of a peer node, the higher the probability of it being selected as a forwarder. If a different set of forwarders are selected for each recurring connection between I and R, the probability of an successful intersection attack increases. In other words, if F_1,F_2,...., F_t are the set of all forwarders involved in the anonymous forwarding from I to R, then one should minimize this metric: $Q = |\bigcup_{i=1}^{t} F_i|$. Therefore, two conditions are desired in the systems we are concerned: (1) a relatively static set of intermediate nodes, and (2) stable selection of forwarders for all connections between I and R. The first condition is related to the availability of nodes, the second is concerned with the routing decision at each intermediate node.

Our goal is to design an incentive mechanism that not only induces peer nodes to provide stable forwarding service but also encourage them to make routing decisions aligned with the system objective of providing anonymity. We assume that the peer nodes are untrusted and unreliable in general. We quantify the problem as follows: Let $\pi = \{\pi^1, \pi^2 \cdots \pi^k\}$ be the set of k recurring connections between I and R and let L denote the average length of forwarding paths between I and R. We define the path quality of π, denoted by $Q(\pi)$, as $\frac{L}{||\pi||}$ where $||\pi||$ represents the size of the forwarder set[1]. The system objective is to maximize $Q(\pi)$ by minimizing $||\pi||$.

2.2. Outline of Incentive Mechanism

We model the system as a network of N nodes which participate in anonymous forwarding of data packets. Each node s maintains information about a fixed number d of neighbors which can be used as potential forwarders. This neighbor set is denoted by $D(s)$ (for a detailed description of the system, we refer the reader to the corresponding technical report [19]). When an initiator I decides to set up a connection to a responder R, it uses the following mechanism. It makes a commitment to pay an amount P_f to any intermediate forwarder, per forwarding instance (**forwarding benefit**). In addition it also decides to pay a total shared benefit (**routing benefit**) equal to P_r to all the forwarders. Thus if a forwarder participates in m forwarding instances, its benefit is $mP_f + \frac{P_r}{||\pi||}$. The idea of separating the total benefit into routing and forwarding components serves the following purpose. The forwarding benefit induces availability of nodes because even if a forwarder makes a random routing decision, it still stands to gain by just participating in the forwarding process. The routing benefit induces the nodes to make routing decisions which are aligned with the system objective of minimizing $||\pi||$.[2] This can be achieved by making non-random routing decisions. Note that the routing benefit induces an implicit cooperation among forwarders. For example, in Figure 1, node X is not available for forwarding in π_2; consequently its forwarding benefit is smaller than the scenario in Figure 2. Moreover, the routing benefit for each forwarder is reduced from $\frac{P_r}{8}$ to $\frac{P_r}{3}$. Thus the utility function must be designed in such a way that in trying to maximize their utilities, nodes take forwarding and routing decisions which are aligned with the system objective.

We next define the utility function which is used by a forwarder X to select the next hop on the path from I to R. Let

[1] L is used to normalize the forwarder set size for a given average path length.

[2] Note that we are basically concerned with the forwarder set for appropriate path lengths. The system objective is to ensure a minimum size forwarder set for path lengths which are appropriate for anonymity systems. For example in Crowds, tweaking the value of forwarding probability appropriately results in appropriate path lengths.

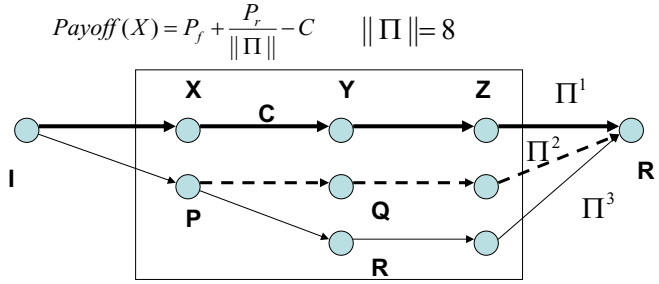

Figure 1: Random routing by P and unavailability of X leads to a large size of the forwarder set.

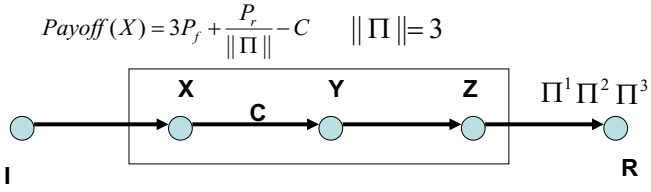

Figure 2: Stable set of forwarders.

q_e be the quality[3] of the forwarding edge $e = (X, Y)$ from X to Y on some path $\in \pi$ and C be the sum of participation cost and forwarding cost (to Y) incurred by X. From a system perspective, we would want that a forwarders utility be aligned with the global objective, i.e. its utility should increase if it selects a high quality edge. We therefore define the utility for a forwarder X as

$$U_X(Y) \quad = \quad P_f + q_e P_r - C \tag{1}$$

Note that in trying to maximize its utility, X selects high quality forwarding edges which in turn increases its payoff due to a decrease in $||\pi||$. Thus this utility model captures the effect of local decision making on the final payoff to a forwarder and aligns its interest with the system objective. An intermediate forwarder X decides to participate or not participate in forwarding and routing of the payload on the basis of its utility. It calculates its utility corresponding to each neighbor $g \in D(X)$ and selects the neighbor which gives it the maximum utility as the next hop. Ties are broken by selecting a neighbor with a higher quality. Note that since the identity of the intermediate nodes (except first hop) is not known to the initiator, the establishment of the forwarding path is based on propagation of contract information (P_f and P_r) through the intermediate nodes (Note that both Crowds like probabilistic forwarding and hop-distance based forwarding are applicable to our model). Finally after R receives the payload, it sends back a confirmation through the reverse path. Each intermediate forwarder also includes path informtion which is then used by I to recreate the path and validate it. After evaluating the path quality, the initiator uses a central entity (bank) to make payments to the forwarders. Note that although the identity of R is known to the intermediate nodes, the identity of I is not leaked and therefore the system achieves initiator anonymity. The payment is made by I only after all the connections in π are completed. Details about path quality evaluation and payment mechanism can be found in the technical report [19]. For the initiator, a high benefit corresponds to a low value of $||\pi||$ (such that $A(||\pi||)$[4] increases with decrease in $||\pi||$) and the cost it incurs is equal to the payment it makes to the forwarders. Therefore

$$U_I = A(||\pi||) - ||\pi|| P_f - P_r \tag{2}$$

Relationship between forwarding and routing benefits: A high value of P_f increases the probability of peer participation in the forwarding process. A high value of P_r gives a higher weightage to the benefit and this results in a higher profit for a forwarder. Since a high benefit corresponds to a high quality path, a high value of P_f results in the formation of high quality paths between I and R. Thus depending on its anonymity requirements, the initiator can select appropriate values for P_f and P_r. Note that the ratio of P_f and P_r also affects the decisions taken by the forwarders. If $P_r = \tau P_f$, then a small value of τ would induce nodes to forward traffic; however, it may not align their routing decisions to the system objective. On the other hand a high value of τ will induce nodes to make effective forwarding and routing decisions.

2.3. Edge Quality

Connection history. Each node stores history information about connections passing through it. Thus if a node s lies on a path π^i with connection identifier cid, it stores the corresponding predecessor and successor hops as shown in Figure 1. The history information at s for the k^{th} connection, represented as $H^{k-1}(s)$, consists of all outgoing edges from s which lie on $\pi^1, \pi^2, \cdots, \pi^i, \cdots, \pi^{k-1}$. Note that by using the predecessor information, a node can differentiate between outgoing edges for two different positions on the

Connection id	Predecessor	Successor
cid	X	Y

Table 1: History profile at node s.

[3] We introduce the notion of edge quality in the context of path reformations in section 2.3.

[4] We use A(.) as a function for quantifying the anonymity.

same path (e.g. if node s occupies two different positions on π^k).

Availability of neighbors. Each node also stores availability information about its neighbors. In the absence of a centralized entity for collecting availability information, each peer calculates availability of its neighbors using its own observations. A peer uses active probing [25] to monitor its neighbors. Mechanisms based on active probing have been used to estimate churn in peer-to-peer systems. We use a methodology similar to [4] to estimate availability. When a peer first joins the system, it initializes the session time of each of its neighbors to 0. At the start of each probing period a peer s checks the liveness of each neighbor. If the neighbor is alive, its session time t_s is updated as $t_s^{new} = t_s^{old} + T$, where T is the probing time period. If a new neighbor is found, its session time is updated as $t_s^{new} = rand(0,T)$ where $rand(0,T)$ is a uniformly distributed random value in the range $(0,T)$. Finally availability of a neighbor u, $u \in D(s)$ is calculated as $\alpha(u) = \frac{t_s(u)}{\sum_{v \in D(s)} t_s(v)}$. Thus a neighbor with a higher observed session time has a higher availability. This is in accordance with observed session times of peers in peer-to-peer file sharing systems, which is modeled using a pareto distribution [18]. We represent the availability of a node v as observed by s as $\alpha_s(v)$ s.t. $0 \leq \alpha \leq 1$

Determining edge quality. We now outline a local mechanism for determining the quality of an edge at a node. Consider a node s which lies on π^k. Let $D(s)$ be the neighbor set of s. s calculates the quality of each outgoing edge, $q(s,v)$ using a procedure which takes as inputs v, $H^{k-1}(s)$ and $\alpha_s(v)$. Given an edge (s,v), s looks up its history information $H^{k-1}(s)$ (path information corresponding to $\pi^1, \pi^2 \cdots \pi^{k-1}$) for any entry corresponding to (s,v). The ratio of the number of entries corresponding to (s,v) and the maximum possible entries $(k-1)$ is called its selectivity and represented as $\sigma(s,v)$. Weights w_s and w_a are assigned to selectivity $\sigma(s,v)$ and availability $\alpha(v)$ respectively such that $w_s + w_a$=1. Finally, the edge quality is calculated as $q(s,v) = w_s\sigma(s,v) + w_a\alpha(v)$. The weights w_s and w_a signify the relative importance of selectivity and availability. A high value of w_a signifies a higher importance to the availability of the forwarders, with the objective that these forwarders would be available for forwarding for future connections. A high value of w_s on the other hand signifies higher importance for past history. Note that the edge quality of the last edge in the path π^k is always 1 because it ends in R. Note that w_s and w_a are system parameters which are set depending on the anonymity requirements of the system. The amount of history information stored at a node also influences the quality of the edge. The quality of a path π^k is then given by the sum of the qualities of the individual edges. We next show how incentive based non-random routing leads to reduction in path reformations.

Proposition 1. *Incentive based non-random routing by intermediate nodes leads to reduction in path reformations when compared with random routing.*

Proof. Consider an edge $e = (a,b)$ on path π^k. Let X be a random variable such that

$$X = \begin{cases} 0 & : \ if \ \ e \in \bigcup_{i=1}^{i=k-1} \pi^i \\ 1 & : \ otherwise \end{cases}$$

We need to show that $E[X]$ for random forwarding is greater than $E[X]$ for utility based non-random forwarding. If random forwarding is used, then $E[X] \geq \frac{N-(k-1+1)}{N} = 1 - \frac{k}{N}$. Since $k << N, E[X] \rightarrow 1$. In the case of utility based forwarding, a new edge is added only if there is no existing edge in $\bigcup_{i=1}^{i=k-1} \pi^i$. Let the probability that an edge $\in \ \pi^i$ is available in π^k be p_i. Then $E[X] = (1-p_1)(1-p_2)\cdots(1-p_i)\cdots(1-p_{k-1}) \ \forall \ 0 < w_s, w_a < 1$. Note that since $w_a > 0$, $p_i \rightarrow 1$ as $i \rightarrow k$. Consequently $E[X] \approx 0$. From equation 1, $U_a(b)$ increases with an increase in $q(a,b)$ and therefore a rational forwarder would always try to select high quality edges which in turn leads to low path reformations. □

2.4. Forwarding and Routing Strategy

We model the forwarding and routing mechanism as a finite multi-stage game [15] where the peers are the players. Consider a system containing N peers represented by the set $V = \{1, 2 \cdots N\}$. A nodes forwarding and routing strategy space SS is the set of all nodes in the system (except itself) along with the NULL entity, which corresponds to the case when a node does not participate in the forwarding path. Therefore $SS_i = \{1, 2, \cdots, i-1, i+1, \cdots, N, NULL\}$. The strategy profile of node i is represented as $\{S_i, S_{-i}\}$ where S_i and S_{i-1} represent the strategies of i and other nodes respectively. For a path π, the strategy profile of the set of nodes is therefore given by $SP = \bigcup_{i=1}^{N} \{S_i\}$. At each stage a node has three choices; a) not participate in forwarding, b) forward and route randomly, c) forward and route non-randomly. Note that the primary objective of an adversary in an anonymous forwarding system is to identify the end points of a communication and therefore its routing decision is not aligned with any economic incentive. We model an adversary's routing strategy as random routing. Note that the main objective of an adversary is to break initiator anonymity; therefore it is not concerned about the incentive. Our main objective here is to ensure that there is an equilibrium in the strategies for the selfish nodes (those who want to obtain maximum income) and in doing so the availability of nodes in the system increases. This in turn affects the anonymity of the system.

We first outline some game-theory basics. A *Dominant Strategy* [15] for a player i is a strategy which gives it an

optimal utility irrespective of the strategies taken by other players. A *Nash Equilibrium* [15] represented as $\{S_i^*, S_{-i}^*\}$ is a strategy profile in which each players utility is optimal given that the other players have also played their optimal strategies. An equilibrium is a weaker property than a dominant strategy. Finally, a *Subgame Perfect Nash Equilibrium (SPNE)* [15] is a special case where irrespective of the past, playing the assigned strategies from the current stage is still an equilibrium.

2.4.1 Cost Model

The costs incurred by peers includes the following: a) Participation cost and b) Transmission cost. Any participating peer incurs a certain cost which depends on the nature of the application. A number of internet protocols including HTTP, FTP, NNTP and raw sockets etc. are characterized by a recurring traffic pattern [26] and therefore any client application which uses these protocols is vulnerable to path reformations. Consequently our cost model should be generic and must not be limited to any particular application. The cost of participation therefore includes the cost of running a software associated with a particular application for a peer session. This cost is represented as C^p. For the initiator, the participation cost is equal to the sum of payments that it makes to the intermediate forwarders. The transmission cost for a peer is associated with forwarding the payload to the next hop and is represented as C^t. If the payload size is b and per unit transmission cost to the next hop is l, then $C^t = bl$. We ignore the cost of transmitting control packets which is negligible. Note that the participation cost is incurred by a peer for participating in the anonymity system and is a one time cost, while the forwarding cost is incurred per forwarding instance. A peer tries to maximize its own access bandwidth for sending its own traffic and therefore during data forwarding for other peers, its rational (selfish) nature will make it forward traffic on low bandwidth links. This type of selfish behavior by peers has been modeled for peer-to-peer streaming [24].

2.4.2 Utility Model I

The forwarding and routing mechanism can be treated as a game where each forwarder can be modeled as a player. The path formation can be modeled as a sequential reasoning process at each forwarder such that the decision is influenced by the utility to the forwarders. Having outlined a possible mechanism for evaluating edge quality in section 2.3, we can now formalize the utility model for the i^{th} peer (from equation 1).

$$U_i(j) = P_f + q(i,j)P_r - (C_i^p + C^t(i,j))$$

In this model the benefit to a forwarder i is proportional to the quality of the edge from i to its successor j. The rationale behind using this model is that each forwarder can make a local decision based on the quality of the edges to each of its neighbors. Since a path is composed of edges, ensuring a high quality for each individual edge can also lead to formation of a path with high quality. In this case the determination of the optimal next hop requires sorting the utilities corresponding to each edge and has a complexity of $O(logd)$. From a system perspective, we would like to derive the conditions under which there is a dominant forwarding and routing strategy for each forwarder. Ideally, we would like peers to participate in the forwarding process and route the payload to the best quality neighbor. We next show the conditions under which forwarding can be induced and also show the existence of a dominant routing strategy for good nodes.

Proposition 2. *If we assume a constant participation cost C^p and a constant fowarding cost C^t for all forwarders, then the condition $P_f > \frac{C^p N}{Lk} + C^t$ can induce peers to participate in forwarding.*

For a detailed proof of the proposition, we refer the reader to the technical report [19].

Proposition 3. *If $P_f > (C_i^p + C_i^t)$, forwarding is a dominant strategy for the forwarding stage.*

Due to space limitations, we again refer the reader to the corresponding technical report [19].

2.4.3 Utility Model II

We next consider an utility model whereby the utility is proportional to the quality of the path from i to the responder R. The intuition is that i can select a neighbor corresponding to a high quality path from i to R. Here $q(\pi(i,j,R))$ represents the quality of the path from i to R which goes through j.

$$U_i(j) = P_f + q(\pi(i,j,R))P_r - (C_i^p + C^t(i,j))$$

The path formation can be modeled as a L stage game for a path of length L such that at each stage only one player makes a move. We also define a history information at stage l. The history information corresponds to the position of the forwarder on the path and the identity of the predecessor. A subgame perfect nash equilibria is then a strategy profile SP such that each subgame $G_l \ \forall l = 1, 2 \cdots, L$ is a nash equilibrium. The equilibrium strategy profile (S_i^*, S_{-i}^*) can be derived using backward induction. We again refer the reader to the technical report [19] for a detailed discussion of this utility model.

3. Experiments

We analyze the effect of forwarding and routing benefits on the path quality and study the effect of malicious nodes on the path equilibrium and how it affects the payoff of good nodes. We use a discrete event simulator to perform

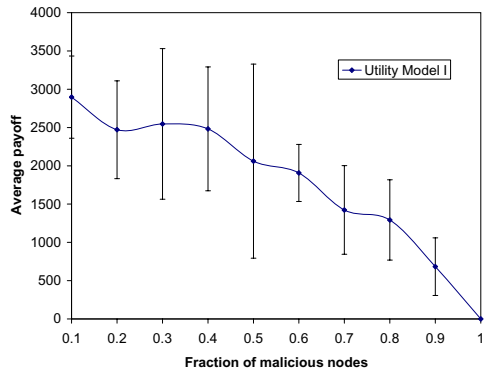

Figure 3: Average payoff for a non-malicious node.

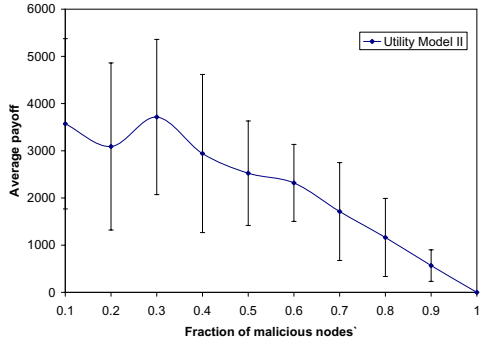

Figure 4: Average payoff for a non-malicious node.

the various experiments. For simulation simplicity, we use a small network size of $N = 40$ to study the effect of the utility models. A poisson process is used to simulate the joining of nodes and each node randomly selects d nodes as its neigbors (unless otherwise specified, d is selected as 5 in our experiments). A set of nodes are randomly selected as Initiators and Responders. A (Initiator,Responder) pair is then randomly selected as the end points of an anonymous message transmission. The number of maximum transmissions for the same (I, R) pair is controlled using a parameter $max - connections$ in our simulations. A typical simulation setup involves 100 (I, R) pairs and a total of 2000 message transmissions, for an average of 20 communications rounds for a single (I, R) pair. The forwarding benefit, P_f for a (I, R) pair is randomly selected from the range $[50, 100]$ (since we did not have any particular application in mind, these values are arbitrary, however we believe they are reasonable to study the effect of benefits through simulations) and τ is selected from the set $(0.5, 1, 2, 4)$. Unless otherwise specified, the weights w_s and w_a are chosen as 0.5 and 0.5 respectively. We model the transmission cost between two peers as being proportional to the communication bandwidth between them. The session time of peers is modeled using a Pareto distribution and the median session

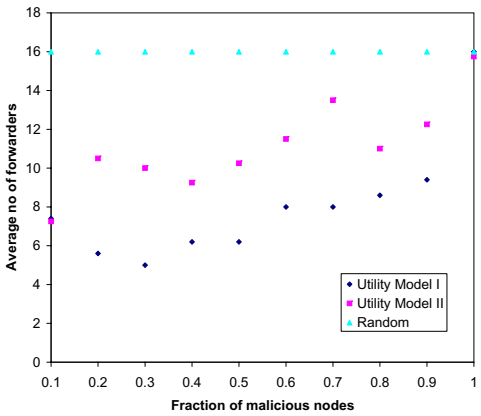

Figure 5: Comparison of the quality of a set of recurring connections between I and R using different routing strategies.

	$\tau = 0.5$	$\tau = 1$	$\tau = 2$	$\tau = 4$
f=0.1	409	390	391	456
f=0.5	299	298	332	306
f=0.9	85	91	72	122
Mean	296	303	301	360

Table 2: Routing efficiency for utility model I.

time is set as 60 mins in accordance with the analysis done in [23]. A certain fraction f of nodes are selected as adversaries and an adversary's routing strategy is modeled as random routing.

Impact of malicious nodes on the payoff for good nodes Adversarial or malicious nodes randomly select the next hops for forwarding the packets which increases the size of the forwarder set for a set of connections between I and R. Consequently, the expected payoff for good nodes can decrease because the routing benefit gets shared by a large number of nodes and this can weaken their incentive to cooperate. We study this effect through simulations. Figures 3 and 4 show the decrease in average payoff for good nodes for varying fractions of adversarial nodes (error bars show 95 % confidence interval). Both utility models exhibit similar nature. Note that at low values of f, the average payoff is appreciably high. We also analyze the effect of the size of forwarding and routing benefits and τ (ratio between routing and forwarding benefits) on the payoff to a good node. We use *routing efficiency* (ratio of average payoff and average number of forwarders) as a metric to quantify the effectiveness of the routing strategy of forwarders for a given value of τ. Note that a high value of *routing efficiency* is aligned with the system objective of inducing forwarders to make routing decisions so as to minimize the size of the forwarder set. Figure 2 shows that a high value of τ tends to increase the routing efficiency.

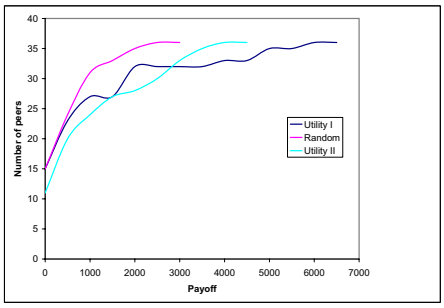

Figure 6: CDF of payoff for good nodes when f=0.1.

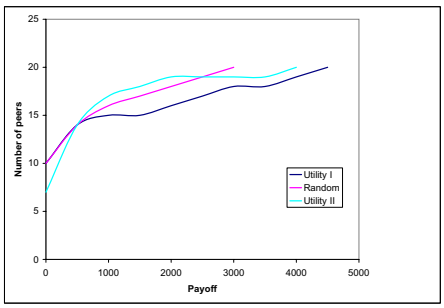

Figure 7: CDF of payoff for good nodes when f=0.5.

Comparison of the effect of routing strategies on path quality We use the average size of the forwarder set as a metric to compare the path quality under different routing strategies. Figure 5 shows the average size of the forwarder set for different routing strategies under varying fractions of malicious nodes. Both utility models I and II appreciably outperform random routing. Utility model I shows the minimum size of the forwarder set. We also analyze the distribution of payoffs among the good nodes for the different routing strategies to gain insight into the effectiveness of the utility models. Figures 6 and 7 show the cumulative distribution function of the payoffs for the good nodes. This includes both the forwarding and routing benefits. We observe that the maximum payoff is highest in the case of Utility I and both models I and II show similar values for average payoff for good nodes. However, the payoff distribution has the maximum variance in the case of model I. In comparison random routing shows a much smaller variance. The reason is that in case of utility model I, if a peer is selected for forwarding traffic during a particular connection (e.g. based on its availability), then it is very likely that it will be selected again for future connections. This results in a skewed distribution of the payoffs. A similar reasoning applies for utility model II. On the other hand, when random routing is used, there is an uniform probability (1-f) that any peer will be selected for forwarding traffic. Note that utility model I results in a high average and maximum payoff and is therefore effective in inducing nodes to participate in forwarding and also route non-randomly.

4. Related Work

A quantitative analysis of anonymous communications was presented in [17] and the effect of path length on anonymity was studied. Mutual anonymity protocols for hybrid peer-to-peer systems was proposed in [28]. The economic aspects of anonymity was first addressed in [1]. It outlined the reasons why anonymity systems are hard to deploy and enumerated the incentives to participate for both initiators and intermediate forwarders. Previous work in this area has treated protocol compliance [10, 9] and availability separately [13]. We address both the issues within the same framework. Reputation mechanisms were used to address the issue of compliance in MIX networks [9] and remailers [10] respectively. Reputation-based schemes are based on feedback about nodes in a system which are made through observations. As mentioned in [13], schemes based on system wide monitoring are not ideally suited for anonymity systems. Moreover, an inherent problem with a scoring or reputation mechanism is that nodes can collude with each other to increase their score or reputation and therefore increase their probability of being selected in the forwarding path. The work presented in [13] proposed the use of an incentive mechanism to ensure the availability of forwarders in a fixed length forwarding system. Although it addresses the issue of availability in forwarding-based anonymity systems, the proposed mechanism is limited to systems in which the identity of the intermediate nodes is known to the initiator. Incentive mechanisms for protocol compliant forwarding and routing have been proposed for both wired [2] and wireless [7, 3, 5] networks. The hidden-action problem in routing was addressed in [12] and an incentive mechanism was proposed to overcome the problem through the use of direct and recursive contracts. Micro-payment based schemes were proposed in [29] and [6] to stimulate cooperation in adhoc networks. However, these mechanisms are not ideally suited for anonymity systems in which the identity of the initator must be hidden from other peers.

5. Discussions and Conclusion

Our incentive mechanism is vulnerable to the following anonymity attacks: (1) availability attacks; where malicious nodes become highly available and wait for paths to be reformed through them, (2) traffic Analysis Attacks and (3) attacks through the use of connection identifier in the history information stored at a malicious node. However, our system implementation can address these attacks. We refer the reader to the technical report [19] for a detailed description of the various system implementation issues and how the system handles various attacks on anonymity. We also describe the payment infrastructure and the various cryptographic operations involved in route formation and verifica-

tion.

We have proposed an incentive mechanism to increase peer availability and reduce path reformations for forwarding based anonymity systems. We propose two utility models and use game theory to evaluate forwarding and routing strategies of intermediate peers. We also show the appropriateness of these utility models for forwarding based anonymity systems and evaluate their effectiveness in aligning the forwarding and routing strategies of peers. We compare the effectiveness of routing under these models with random routing. Our simulation results show that the incentive mechanism is quite effective both under churn and presence of malicious adversaries. We also outline a payment mechanism and show that in trying to increase the system anonymity, the payment mechanism does not actually decrease it. In designing the incentive mechanism, we have tried to address the two issues of compliance and availability in an anonymity system [13] within the same framework.

References

[1] A. Acquisti, R. Dingledine, and P. Syverson. On the economics of anonymity. In *Proceedings of the 7th International Financial Cryptography Conference*, 2003.

[2] M. Afergan and J. Wroclawski. On the benefits and feasibility of incentive based routing infrastructure. In *Proceedings of the ACM SIGCOMM Workshop on Practice and theory of incentives in networked systems*, 2004.

[3] L. Anderegg and S. EidenBenz. Ad hoc-vcg: a truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents. In *Proceedings of the Ninth International Conference on Mobile Computing and Networking*, 2003.

[4] F. E. Bustamante and Y. Qiao. Friendships that last: Peer lifespan and its role in p2p protocols. In *Proceedings of the International Workshop on Web Content Caching and Distribution*, 2003.

[5] L. Buttyan and J. P. Hubaux. Enforcing service availability in mobile ad-hoc networks. In *Proceedings of the First ACM Workshop on Mobile Ad Hoc Networking and Computing*, 2000.

[6] L. Buttyan and J.-P. Hubaux. Stimulating cooperation in self-organizing mobile ad hoc networks. In *Mobile Networks and Applications*, volume 8, 2003.

[7] M. Cagalj, S. Ganeriwal, I. Aad, and J. P. Hubaux. On selfish behavior in csma/ca networks. In *Proceedings of the IEEE Infocom*, 2005.

[8] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–88, Feb. 1981.

[9] R. Dingledine, N. Mathewson, and P. Syverson. Reliable mix cascade networks through reputation. In *Proceedings of the Sixth International Financial Cryptography Conference*, 2002.

[10] R. Dingledine, N. Mathewson, and P. Syverson. Reputation in p2p anonymity systems. In *Proceedings of the Workshop on Economics of Peer-to-Peer Systems*, 2003.

[11] M. Feldman and J. Chuang. Overcoming free-riding behavior in peer-to-peer systems. In *ACM Sigecom Exchanges*, volume 6.1, 2005.

[12] M. Feldman, J. Chuang, I. Stoica, and S. Shenker. Hidden-action in multi-hop routing. In *Proceedings of the ACM E-Commerce Conference*, 2005.

[13] D. Figueiredo, J. Shapiro, and D. Towsley. Incentives to promote availability in peer-to-peer anonymity systems. In *Proceedings of the IEEE International Conference on Network Protocols*, 2005.

[14] M. J. Freedman and R. Morris. Tarzan: A peer-to-peer anonymizing network layer. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, 2002.

[15] D. Fudenberg and J. Tirole. Game theory. 1991.

[16] I. Goldberg. Zeroknowledge to discontinue anonymity service. 2001.

[17] Y. Guan, X. Fu, R. Bettati, and W. Zhao. A quantitative analysis of anonymous communications. In *IEEE Transactions on Reliability*, 2004.

[18] D. Liben-Nowell, H. Balakrishnan, and D. Karger. Analysis of the evolution of peer-to-peer systems. In *Proceedings of the 21st Annual Symposium on Principles of Distributed Computing*, 2002.

[19] S. Ray, G. Slutzki, and Z. Zhang. Incentive-driven p2p anonymity system: A game-theoretic approach. Technical report, Department of Computer Engineering, Iowa State University, Ames, 2007.

[20] M. G. Reed, P. F. Syverson, and D. M. Goldschlag. Anonymous connections and onion routing. In *IEEE Journal on Selected Areas in Communications Special Issue on Copyright and Privacy Protection*, 1998.

[21] M. K. Reiter and A. D. Rubin. Crowds: Anonymity for web transactions. In *ACM Transactions on Information and System Security*, 1998.

[22] S. Rhea, D. Geels, T. Roscoe, and J. Kubiatowicz. Handling churn in a dht. In *Proceedings of the USENIX Annual Technical Conference*, 2004.

[23] S. Saroiu, P. K. Gummadi, and S. D. Gribble. A measurement study of peer-to-peer file sharing systems. 2002.

[24] V. Shrivastava and S. Banerjee. Natural selection in peer-to-peer streaming: from the cathedral to the bazaar. In *Proceedings of the International Workshop on Network and Operating Systems Support for Digital Audio and Video*, 2005.

[25] D. Stutzbach and R. Rejaie. Understanding churn in peer-to-peer networks. In *Proceedings of the Internet Measurement Conference*, 2006.

[26] M. Wright, M. Adler, B. N. Levine, and C. Shields. An analysis of the degradation of anonymous protocols. In *Proceedings of the Symposium on Network and Distributed System Security*, 2002.

[27] M. Wright, M. Adler, B. N. Levine, and C. Shields. Defending anonymous communications against passive logging attacks. In *Proceedings of the IEEE Symposium on Security and Privacy*, 2003.

[28] L. Xiao, Z. Xu, and X. Zhang. Mutual anonymity protocols for hybrid peer-to-peer systems. In *Proceedings of the 23rd International Conference on Distributed Computing Systems*, 2003.

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