Incentives for Combatting Freeriding on P2P Networks.

Sepandar D. Kamvar, Mario T. Schlosser, and Hector Garcia-Molina

Stanford University

Abstract. We address the freerider problem on P2P networks. We first propose a specific participation metric, which we call a peer's EigenTrust score. We show that EigenTrust scores accurately capture several different participation criteria. We then propose an incentive scheme that may be used in conjunction with any numerical participation metric. We show that, when these incentives are used in conjunction with EigenTrust scores, they reward participatory peers but don't exclude less active peers.

1 Introduction

A notable problem with many of today's P2P file-sharing networks is the abundance of *freeriders* on the network – peers who take advantage of the network without contributing to it. Up to 70% of Gnutella clients do not share any files, and nearly 50% of all responses are returned by 1% of the peers [1]. This abundance of freeriders, and the load imbalance it creates, punishes those peers who do actively contribute to the network by forcing them to overuse their resources (e.g. bandwidth).

We address this problem by providing incentives for peers to make active contributions to the network. For example, active participators may get preference when they are competing for another peer's resources, such as bandwidth. Our approach is simple: each peer gets a certain participation score, and it receives rewards based on its participation score. The challenges here lie in how to structure incentives that reward active participators without completely excluding peers that are less active.

Previous work in this area has focused primarily on currency-based systems wherein peers gain currency for uploading files, and use currency when downloading files [2]. We take a different approach, rewarding peers with high participation scores with advanced services, such as faster download times or an increased view of the network. Our approach may be used in conjunction with currency-based approaches.

In this work, we describe a scoring system that accurately quantifies participation, even in the presence of malicious peers trying to subvert the system, and we propose some incentives, and show empirically that these incentives benefit participatory peers in a fair manner.

2 EigenTrust

In the incentives that we propose in this paper, we assume the existence of some scoring system that measures the relative participation levels of peers in the system. One useful metric is a peer's EigenTrust score [4]. EigenTrust was developed as a reputation metric for P2P systems. In this work, we show that EigenTrust is also a good measure of a peer's relative participation level.

To test how well a peer's EigenTrust score reflects it's participation level, we simulate a P2P network in the manner described in [3] The simulator described in [3] is an event-driven simulator that proceeds by query cycles. At each query cycle, peers submit

and respond to queries according to certain distributions over peers' interests and the files peers share, and download files from peers who respond to their queries. Freeriders and malicious peers sharing inauthentic files are modeled as well as active participatory peers.

In Figure 1 we plot each peer in the network on a graph where the x-axis represents the EigenTrust score of the peer, and the y-axis represents the number of authentic uploads that the peer provides in a given timespan (15 query cycles). Notice that the EigenTrust score is correlated with the number of authentic uploads; those peers that provide many authentic uploads also have high EigenTrust score. The correlation coefficient is 0.97 indicating a close relationship between the number of authentic uploads and the EigenTrust score.
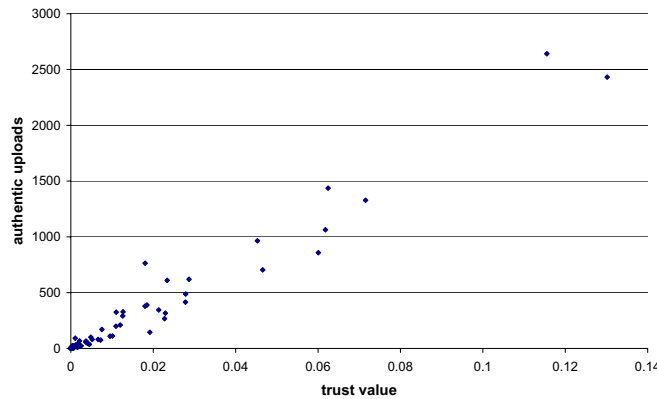

Fig. 1. Correlation between EigenTrust scores (y-axis) and the number of authentic uploads (x-axis). Each point represents a peer.

In order to give the reader further intuition on EigenTrust scores, we examine the following simple example users in our P2P simulation. These users, and their characteristics, are described below and summarized in Table 1.

Angela is an active participator and shares many popular files across many content categories. Bob is an average user who shares a moderate number of files, many of them popular, some of them unpopular, from a couple of content categories. Corcoran is an occasional user who shares a few files, most of them popular. David is an eccentric user, sharing many obscure files that few people want. Ebeniezer is a freerider, who doesn't share any files. And Forster is a malicious user, who shares many corrupt files.

Based on these descriptions, Angela is the most active participator, followed by Bob, Corcoran, David, Ebeniezer, and Forster, in that order. Table 1 shows that EigenTrust captures this ranking. Notice that the EigenTrust scores of Ebeniezer the freerider and Forster the malicious peer are both 0. One may wish for a scoring system in which the malicious peer gets a lower score than a freerider. However, it should be noted that, due to the ease of entry in P2P networks, giving a malicious peer a score below that of a freerider is not very effective. A malicious user with a poor score may simply create a new peer and enter the network as a new freerider.

	# Content Cat.	# Files	Popularity	ET Score	ET Rank
ANGELA	10	1,000	Typical	.02875	1
BOB	3	100	Typical	.00462	2
CORCORAN	3	50	Typical	.00188	3
DAVID	3	200	Unpopular	.00115	4
EBENIEZER	0	0	-	0	5 (tie)
FORSTER	10	2,000	Malicious	0	5 (tie)

Table 1. EigenTrust vs. other rankings for each of our seven sample users. Each of these users has the same uptime (20%), and the same number of connections (10).

We also examine the following pairs of peers, where the first member of each pair is more participatory in a different way. Again, we use these examples to give the reader intuition on the behavior of EigenTrust scores.

1. The first scenario involves Tim and Jim. Tim and Jim share the exact same files, except Tim is always online while Jim turns his computer off at night.
2. The second scenario involves Stan and Jan. Stan and Jan share files in the same content categories. However, Stan shares twice as many files as Jan.
3. The third scenario involves Stephanie and Bethany. Stephanie and Bethany share five files each from the same content category. However, Stephanie shares five very popular files, and Bethany shares five unpopular files.
4. The final scenario involves Mario and Luigi. Mario and Luigi share the same number of files, but Mario has a diverse collection, sharing files in many content categories, while Luigi has narrow interests, and shares files from only one content category.

Table 2 compares the EigenTrust scores of each of these pairs of peers, and shows that each of these characteristics we tested (uptime, number of shared files, popularity of shared files, and diversity of shared files) are reflected in the EigenTrust score of a peer. For example, Tim (who shares the same files as Jim, but has a greater uptime) has a greater number of authentic uploads, and a greater EigenTrust score, than Jim.

	# Content Cat.	# Files	Popularity	# Connections	Uptime	# Auth. Uploads	ET Score
TIM	3	1000	Typical	10	92%	965	0.045
JIM	3	1000	Typical	10	35%	415	0.028
STAN	3	500	Typical	10	33%	325	0.011
JAN	3	1000	Typical	10	33%	379	0.018
STEPHANIE	1	5	5 Most Popular	10	35%	15	0.00038
BETHANY	1	5	5 Least Popular	10	35%	0	0
MARIO	5	5000	Typical	10	25%	267	0.023
LUIGI	1	5000	Typical	10	25%	81	0.0052

Table 2. Comparing Pairs of Peers

3 Incentives

Our goal is to provide incentives that reward participatory peers and punish freeriders and malicious peers. However, we do not want to punish freeriders so much that they are not able to download files and share them to become participators if they so choose.

Two ways to reward participatory peers is to award them faster download times, and grant them a wider view of the network. In this section, we propose two score-based incentive schemes: a bandwidth incentive scheme and a TTL incentive scheme. Notice that these score-based schemes that we propose are completely general, and may be used with any scoring scheme that gives scores for which peers should be rewarded.

Bandwidth. The first incentive that we propose is to give active participators preference when there is competition for bandwidth. More specifically, we propose that, if peer i and peer j are simultaneously downloading from another peer k, then the bandwidth of peer k is divided between peer i and peer j according to their participation scores. So if Tim and Jim are simultaneously downloading from some peer k, then Tim will get $\frac{S_T}{S_T+S_J} * 100\%$ of peer k's bandwidth (where S_T represents Tim's participation score, and S_J is Jim's participation score), and Jim will get $\frac{S_J}{S_T+S_J} * 100\%$ of peer k's bandwidth.

This also works when more than 2 peers are simultaneously downloading from the same peer. The protocol for this incentive is given in Algorithm 1.

Peer i with available bandwidth b, assigns bandwidth to each peer j downloading from it as follows:
foreach *peer j downloading from peer i* **do**
$\qquad bandwidth(\text{peer } j) = \frac{\text{score(peer } j)}{\sum_j \text{score(peer } j)} b;$
end

Algorithm 1: Bandwidth incentive

Notice that if a peer has a participation score of 0, it will get none of peer k's bandwidth if it is competing against other peers for peer k's bandwidth. However, this doesn't exclude that peer from the network, since it is able to download from peers that are not servicing other peers. We show in Section **??** that freeriders are not excluded from the network when this incentive is implemented.

TTL. Currently, Gnutella assigns each peer a time-to-live of 7 for each query. Our second incentive is to assign each peer a TTL based on its participation score, giving active participators a wider view of the network. There are many ways to do this, but one simple way would be to give each peer who has an above-average participation score a high TTL (for example, 10), and to give each peer who has a below-average participation score a low TTL (for example, 5). The protocol for this incentive is given in Algorithm 2.

---

if $score(peer\,j) > mean\_score$ **then**
$\qquad TTL(\text{peer j}) = high\text{-}ttl;$
else
$\qquad TTL(\text{peer j}) = low\text{-}ttl;$
end

Algorithm 2: TTL incentive

One problem here is that each peer must know the average participation score in the network, and explicitly computing this can be prohibitively costly in terms of message complexity, since it would require each peer to know the participation scores of every other peer in the network. However, if EigenTrust scores are used, this isn't a problem. Since the EigenTrust scores in the network sum to 1, the average EigenTrust score for any network is $1/n$, where n is the number of peers in the network (We assume that a peer either knows or can approximate the number of peers in the network. If this is impractical for the given network, a peer can simply substitute its own EigenTrust score for $1/n$.) Therefore, each peer can compare its own EigenTrust score to $1/n$, and if its EigenTrust score is greater, the peer may issue a query with a TTL of 5. Otherwise, the peer may issue a query with a TTL of 3. A peer never needs to explicitly compute the average participation score in the network.

4   Experiments

We have shown in Section 2 that EigenTrust scores are a good measure of participation. Our task in this section is to show that the proposed incentives achieve their stated goals: to reward participatory peers with faster download times and a wider view of the network without completely excluding less active peers from the network. Again, we use our sample peers to give the reader intuition on how these incentives reward peers in the network.

Bandwidth. The bandwidth incentive aims to reward peers with faster downloads. To test this, we again examine the sample peers Angela et. al., and measure their average download speeds in our simulations with and without the bandwidth incentive implemented.

In Table 3, we show the average download speed for each of our sample users when the bandwidth incentive is implemented in conjunction with the EigenTrust scoring scheme, compared to the average download speed for each of our sample users when no incentive is implemented. To simulate a congested network, we employ the following congestion model: for each download that peer i begins to download from peer j, he competes for peer j's bandwidth resources against somewhere between 0 to 4 other peers (chosen from a discrete uniform random distribution) who are also downloading files from peer j. Notice that, when the bandwidth incentive is implemented in this simulation in conjunction with EigenTrust scores, active participators are compensated for their participation, but not at too great an expense to the less active peers.

		Bandwidth Incentive	No Incentive
	EigenTrust Score	% Bandwidth Download Speed	% Bandwidth Download Speed
ANGELA	.02875	.6549	.4566
BOB	.00462	.5307	.4566
CORCORAN	.00188	.5978	.4566
DAVID	.00115	.4737	.4566
EBENIEZER	0	.2	.4566
FORSTER	0	.2	.4566

Table 3. The average download speed (and percent bandwidth) for each peer using the bandwidth incentive in conjunction with the EigenTrust Score (left), and with no incentive (right)

TTL. Again, the important issue to investigate is whether this incentive compensates participators enough to be a useful incentive while giving nonparticipators enough resources so that they have the option of becoming participators if they so choose.

To do this, we simulate 15 query cycles with the TTL incentive scheme activated as in Algorithm 2. Since this is a small network (100 peers), we define the default TTL to be 4 (peers can reach 75 other peers in this network on average with a TTL of 4). In this case, we define *high-ttl* to be 5, and *low-ttl* to be 3.

For reporting our results, we split the peers into two groups: premium users (those peers with EigenTrust scores larger than the average EigenTrust score $\frac{1}{N}$), and moderate users (those peers with EigenTrust scores less than the average EigenTrust score $\frac{1}{N}$). Table 4 shows the number of premium users and moderate users and the average number of peers within their respective TTL range. Note that activating the TTL incentive scheme with these settings decreases the query load on the network while beefing up the service levels for premium users, which is a very desirable result: With the TTL scheme switched on, $77 * 27 + 23 * 98 = 4367$ query messages will be generated throughout the network when all peers issue a query. With the TTL scheme switched off, $100 * 75 = 7500$ messages will be generated in the same process. Also, notice that even the freeriders will not be excluded from the network, as they receive a TTL of 3 for their queries.

		TTL Incentive		No Incentive	
	EigenTrust Score	#Peers	#Peers in TTL Range	#Peers	#Peers in TTL Range
MODERATE USERS	$< \frac{1}{N}$	77	27	74	75
PREMIUM USERS	$> \frac{1}{N}$	23	98	26	75

Table 4. The TTL, average number of peers reached (PR), and average number of responses per query (RPQ) for each peer using the TTL incentive in conjunction with the EigenTrust Score (left) and with no incentive (right).

5   Conclusion

Two main results are presented in this paper: First, we show that a peer's EigenTrust score is a good participation metric according to three natural participation criteria. Second, we present two incentive protocols based on these scores, and show that they reward participatory peers without completely excluding non-participatory peers.

References

1. E. Adar and B. Huberman. Free riding on gnutella. *First Monday*, 5(10), October 2000.
2. P. Golle, K. Leyton-Brown, I. Mironov, and M. Lillibridge. Incentives for sharing in peer-to-peer networks. In *WELCOM01*, 2001.
3. S. Kamvar and M. Schlosser. Simulating a File-Sharing P2P Network. In *SemPGRID03*, 2003.
4. S. Kamvar, M. Schlosser, and H. Garcia-Molina. The EigenTrust Algorithm for Reputation Management in P2P Networks. In *WWW 2003*, 2003.