Measuring Effectiveness of Mechanisms for Promoting Cooperation in BitTorrent

Miranda Mowbray, HP Labs Bristol
Nazareno Andrade, Gustavo Wagner, and Aliandro Lima, UFCG Brazil

Introduction: mechanisms for cooperation in BitTorrent

This paper investigates two mechanisms for promoting cooperation in BitTorrent (the standard tit-for-tat mechanism, and sharing-ratio enforcement), using measurements taken in popular BitTorrent sites. We also suggest some ideas for improving BitTorrent clients inspired by our measurements and by some tit-for-tat-like strategies in animals. We take a biologically-inspired approach to the security and effectiveness of the mechanisms, in the sense that we do not care if there are a few uncooperative peers as long as the system keeps on working with a good quality of service.

All collaborative computing systems potentially face the problem of *freeriders*: that is, users that consume resources of the system without contributing anything in return. This problem is acute for traditional peer-to-peer (P2P) file-sharing systems; Adar et al. [1] found in 2000 that 66% of Gnutella users did not make any files available for download to other users, and in 2005, Hughes et al. [2] classified 85% of Gnutella users as freeriders. Saroiu et al. [3] found that 20% to 40% of Napster users share few or no files.

BitTorrent [4] is a P2P file-distribution tool that moves the burden of bandwidth consumption from the content provider to the peers that download the file. The phenomenon where users of a resource contribute to its distribution to other users also appears in some biological systems, for instance in cooperative foraging by social insects [5]. BitTorrent has proved extremely popular: a CacheLogic report estimates that BitTorrent generated about 30% of all US Internet traffic in June 2004 [6].

BitTorrent works as follows. The original file distributor publishes details of the file on a web server, and creates a tracker that allows peers interested in the file to find each other. To download the file, peers access the tracker and join the torrent (a *torrent* is a group of peers connected to the same tracker). The file is divided into chunks, and, as a peer downloads chunks of the file, it also uploads, to other peers in the torrent, chunks that it has previously downloaded. The burden of bandwidth consumption is thus moved from the original content distributor to all peers in the torrent.

A distinguishing feature of BitTorrent is its tit-for-tat incentive to discourage freeriding and increase peer cooperation. A peer is most likely to upload (i.e. serve content) to those peers that have recently uploaded to him [7]. This gives an incentive for peers to cooperate by uploading to other peers while they download content. Similar behaviour has been observed in a wide variety of animal species: see [8] for an overview. The environmental harshness that appears to drive non-kin cooperation by (for instance) vampire bats [9] suggests that a harsher environment for uncooperative BitTorrent peers might be beneficial: as we will show, under some circumstances the standard BitTorrent protocol can be quite forgiving to uncooperative peers.

Optimum operation of the standard BitTorrent protocol, including the tit-for-tat incentive, should result in peers having a download rate close to their upload rate. This is fine for peers with a symmetric broadband link, but not so good for peers with asymmetric links. In order for these peers to receive a good download rate, it is necessary for some peers in the torrent to *seed*, that is, to continue uploading for some time after they have finished downloading the file, a cooperative behaviour that is not rewarded by the tit-for-tat incentive. Seeding may also increase the length of time that the file remains available for download to new peers.

BitTorrent files are usually published through websites that consist of listings of torrents with HTML links to the trackers used for joining them. This architecture allows sites to enforce additional regulation of BitTorrent peers, increasing environmental harshness for non-cooperative peers. Such institutional regulation cannot be implemented in P2P systems such as Gnutella or Napster, because these systems lack subnetworks with centralized components. One regulatory mechanism implemented in several BitTorrent sites is called *sharing-ratio enforcement*. These sites keep a long-term history of user cooperation. If a user's sharing ratio (the volume of data he has uploaded divided by the volume he has downloaded) falls below a certain threshold, he is prevented from gaining access to new content. The "age" of the user may also be taken into account. In practice, sharing ratio enforcement provides an incentive to seed as well as not to freeride.

Our study

A number of previous studies [10,11,12,13] have analysed existing BitTorrent deployments or used analytical modelling and simulations to characterize the properties of the BitTorrent protocol and to improve its performance. However, these studies did not provide empirical evidence on the effectiveness of BitTorrent incentive mechanisms in promoting cooperative user behaviour in real deployments. To address this question, we analysed data sampled from several BitTorrent sites with thousands (or hundreds of thousands) of peers, *bt.etree.org*, *easytree.org*, *piratebay.org*, *torrentportal.com* and *btefnet.net*.

As we will show, we found that the BitTorrent tit-for-tat incentive did not work for torrents when a large proportion of the peers in the torrent were seeding, but despite this weakness the amount of freeriding in the BitTorrent sites was low. Judging by our comparative measurements between different sites, sharing-ratio enforcement can boost cooperation still further. Our study of seeding patterns in the sites suggests a way of changing the client to increase seeding effectiveness.

We also considered security of the incentive mechanisms: we investigated ways in which uncooperative users might try to fool or bypass the mechanisms that do not require alterations to the code of the BitTorrent clients, and found evidence these attacks are not currently causing problems.

Freeriding

We obtained data from the BitTorrent sites *easytree* and *etree* on the amount of data downloaded and uploaded by individual peers. (This data was not available from the other sites.) We found that only 5% to 6% of live peers in these two sites had not uploaded any data, a much lower rate of freeriding than the 20-40%, 66% or 85% reported by various studies of Gnutella and Napster [1,2,3]. It appears that the design of the BitTorrent protocol is successful at reducing the amount of freeriding.

Sharing-ratio enforcement

One of the five BitTorrent sites we studied, *easytree*, uses sharing-ratio enforcement. We found that it had significantly higher levels of seeding than all of the other four sites. In *easytree* torrents in which some sharing was going on, 59% of the peers were seeding at the instant of data capture, as opposed to at most 55% for the other sites. We verified using partial regressions that this significant difference in the amount of seeding between the sites was not explained by differences in the distribution of torrent ages, file sizes, or numbers of peers per torrent.

Effectiveness of the tit-for-tat incentive

Figure 1 gives the results of an experiment showing that, although BitTorrent is successful in penalizing freeriding in torrents when a small proportion of the peers in the torrent are seeding, in torrents with many seeders freeriders are better off than collaborating peers. This might be a direct manifestation of the cost of cooperation [14]: TCP acknowledgement packets compete with (and slow down) incoming data streams. The graph plots the download time experienced by a freerider divided by that experienced by a collaborating peer. The X axis is the ratio of seeders to non-seeders (leechers) in the torrent. The experiment was conducted in four BitTorrent sites. We used two clients, one of which had been modified to make it freeride. The raw data, and more details about the experiment, are available at http://www.lsd.ufcg.edu.br/~gustavo/bittorrent/raw_data.sxc

Despite this, in *easytree* and *etree*, the two sites for which we had site-wide information about the volume of data uploaded and downloaded by peers, we found that there was a significant positive correlation between upload rates and download rates for peers. So on the whole the tit-for-tat mechanism is working when viewed over these entire sites, although it may give a perverse incentive in individual torrents.

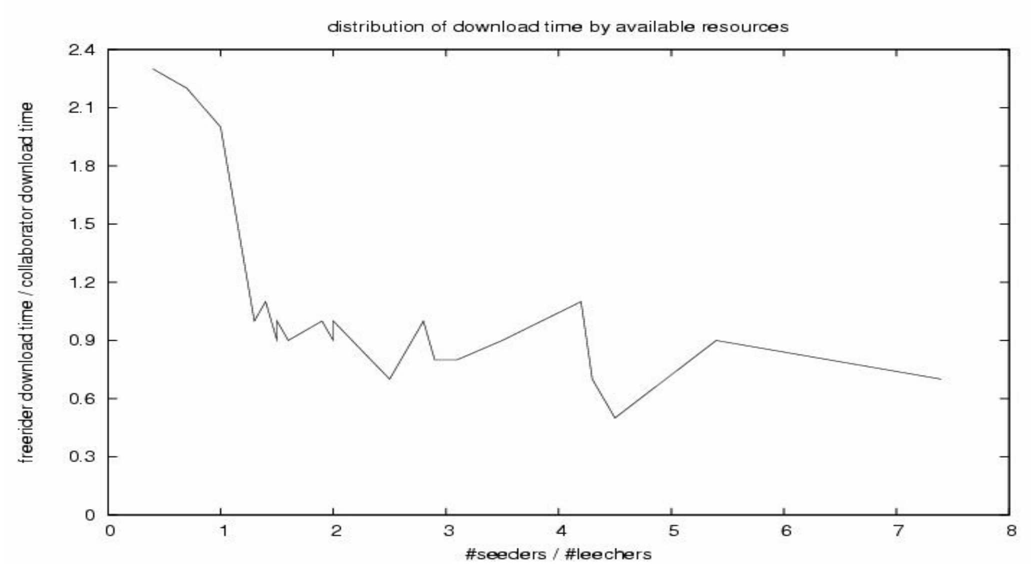

Figure 1: Experiment measuring effectiveness of the tit-for-tat mechanism

Security of the incentives: checking for uncooperative users (with unhacked clients)

Changing the code of the BitTorrent client is probably beyond the capability of most BitTorrent users, but there are several ways that an uncooperative user might try to fool or bypass the incentive mechanisms that do not require him to do this. We have checked that these are not currently a significant problem in *easytree*.

One BitTorrent client apparently can be made to report its downloads as uploads [15]; contrary to the belief of some BitTorrent users this would not fool the tit-for-tat mechanism, because clients actually decide who to upload to based strictly on the transfer rates they experience directly [16], but it might circumvent sharing-ratio enforcement. However, if we ignore the peers reported as having zero download, there is still more seeding in *easytree* than in *etree*, so if the attack on sharing ratio enforcement using this client is happening in *easytree*, it does not appear to be common enough to cause degraded functionality of the site.

An uncooperative user might try to improve his sharing ratio by leaving the system and re-entering as a new user. The administrators of *easytree* report that they have successfully limited this problem by imposing restrictions on the creation of new accounts, so that a user who tries to re-enter the system with a new identity may have to wait for a long time.

Finally, he might try to maintain his sharing ratio just above the threshold level, reducing the amount of upload to the bare minimum. The histogram of sharing ratios of peers in *easytree* does not show a peak at or just above the threshold, confirming that such behaviour is rare in *easytree*.

Ideas for improving the client

The experimental result shown in Figure 1 suggests that BitTorrent clients are not seeding in the most effective way. If clients could spend more of their seeding time in torrents with less seeding (and less in torrents with more seeding) then this could potentially reduce every peer's download time without requiring any increase in the amount of time peers spend seeding. In all the sites we measured, we found a negative correlation between the amount of seeding in a torrent and the size of the file shared by the torrent. (All the correlations were significant at the 0.01 significance level.) We suggest therefore that clients could be adapted to spend more of their seeding time in torrents with larger file sizes: clients already know the size of the file, and this value is static, whereas the amount of seeding in the torrent is not.

Finally, we suggest two possible improvements to BitTorrent clients inspired by tit-for-tat-like animal behaviour. First, the aquatic worm *Ophryotrocha diadema* appears to choose tit-for-tat-like partners carefully, but not to punish them if they cheat later [17]; analogously, it might be worthwhile for a BitTorrent client to keep on uploading for a time to a peer that previously had uploaded to it with high bandwidth but whose upload bandwidth was recently low, in case this was just a temporary problem. Second, the strategy of attitudinal reciprocity adopted by brown capuchin monkeys [18] suggests that it might be advantageous for a client to upload not to a set number of peers with the highest recent upload rates, but rather to peers who recently uploaded at above a particular rate.

Acknowledgements

Most, but not all, of the content of this paper was first reported in "Influences on Cooperation in BitTorrent Communities", by this paper's authors and Matei Ripeanu, published at P2PEcon '05. Thanks to Jupiter and Erwe of Easytree for their useful discussions and data.
This work was partially supported by HP Brazil R&D.

References

[1] E. Adar and B. Huberman. Free Riding on Gnutella. *First Monday* **5**(10), October 2000.
[2] D. Hughes, CG. Coulson and J. Walkerdine. Freeriding on Gnutella Revisited: The Bell Tolls? Submitted to *IEEE Distributed Systems Online*, February 2005.
[3] S. Saroiu, P. K. Gummadi and S. D. Gribble. A Measurement Study of Peer-to-Peer File Sharing Systems. In *Proc. Multimedia Computing and Networking 2003 (MMCN'02)*, San Jose CA USA, January 2002. http://www.cs.washington.edu/homes/gribble/papers/mmcn.pdf
[4] BitTorrent, Inc. BitTorrent web site, 2001-2005. http://www.bittorrent.com
[5] C. Anderson and J. J.Bartholdi, III. Centralized versus decentralized control in manufacturing: lessons from social insects. *Complexity and Complex Systems in Industry*, Warwick UK, September 2000, pp.92-105.
[6] A. Parker. The true picture of peer-to-peer sharing. CacheLogic, July 2004. http://www.cachelogic.com/research
[7] B. Cohen. Incentives Build Robustness in BitTorrent. In *Proc. Workshop on Economics of Peer-to-Peer Systems*, Berkeley CA USA, June 2003.
[8] Lee Alan Dugatkin. *Cooperation among animals: an evolutionary perspective*. Oxford University Press, 1997.
[9] G. Wilkinson. Reciprocal food sharing in vampire bats. *Nature* 308: 181-184, 1984.
[10] J. A. Pouwelse, P. Garbacki, D. H. J. Epema and H. J. Sips. The BitTorrent P2P file-sharing system: Measurements and analysis. In *4th International Workshop on Peer-to- Peer Systems (IPTPS'05),* February 2005.
[11] M. Izal, G. Urvoy-Keller, E. W. Biersack, P. Felber, A. A. Hamra and L. Garcés-Erice. Dissecting BitTorrent: Five Months in a Torrent's Lifetime. In *Proc. Passive and Active Measurements (PAM 2004)*, Antibes, France, April 2004.
[12] D. Qiu and R. Srikand. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. *ACM SIGCOMM Computer Communications Review* 34(4), October 2004.
[13] A. Bharambe, V. Padmanabhan and C. Herely. Understanding and Deconstructing BitTorrent Performance. Technical report MSR-TR-2005-03, Microsoft Reseach, Febrary 2005. http://www.research.microsoft.com/~padmanab/papers/msr-tr-2005-03.ps
[14] M. Feldman, K. Lai, J. Chuang and I. Stoica. Quantifying disincentives in peer-to-peer networks. *Workshop on Economics of Peer-to-Peer Systems*, June 2003.
[15] "Mr Pink", "RE: Fake Stats", *G3 torrent project message board*, 4 July 2004. http://sourceforge.net/forum/message.php?msg_id=264
[16] B. Cohen, *Slashdot* interview, 3 June 2002. http://interviews.slashdot.org/article.pl?sid=03/06/02/1216202&mode=thread&tid=126&tid=185&tid=9

[17] G. Sella. Reciprocation, reproductive success and safeguards against cheating in a hermaphroditic polychaete worm Ophryotrocha diadema. *Biol. Bull* 175: 212-217, 1988.
[18] Frans B. M. de Waal. Attitudinal reciprocity in food sharing among brown capuchin monkeys. *Animal Behaviour* 60(2) 254-261, August 2000.