Promoting contribution in peer-to-peer communities[♦]

Miranda Mowbray
Enterprise Systems and Software Laboratory
HP Laboratories Bristol

peer-to-peer, online
communities, user
contribution

This paper gives practical tips for administrators interested in promoting contribution within P2P communities used in the workplace.
 - Users contribute to P2P communities both to maximize their own personal utility, and for social/cultural reasons. Administrators should consider both types of factor when designing their communities to encourage contribution. I will describe some ways to enhance these two types of motivation.

 - Administrators should identify the needs and motivations of their prolific contributors and make it a priority to support them.

 - The easier it is for users to contribute, the more contribution there will be. I will describe some ways to make contribution easier for users.

 This paper is based on research with Matei Ripeanu (UBC, Canada), Nazareno Andrade and Aliandro Lima (UFCG, Brazil) on contribution in P2P communities that use BitTorrent software.

Approved for External Publication

Promoting contribution in peer-to-peer communities

Miranda Mowbray (HP Labs Bristol, UK)
Based on work with Matei Ripeanu (UBC, Canada), Nazareno Andrade
and Aliandro Lima (UFCG, Brazil) [10]

Keywords: Peer-to-peer, user contributions

1. Introduction

Peer-to-peer (P2P) is a disruptive technology [8] which allows digital content to be transmitted from one user to another using the computer power and bandwidth of the users' own computers, rather than relying on centralized servers. According to [3], by the end of 2004 P2P represented more than half of all core Internet traffic.

Initially the most common use of P2P technology was to distribute digital music and video in violation of copyright. However, it is now frequently used for perfectly legal and respectable applications, and there is increasing interest by media companies and other businesses in commercial applications of P2P.

P2P distribution of digital content relies on contributions from users. This contribution may be of the content itself, or of bandwidth for distributing content to other users.

In this paper I will give practical tips to administrators and information professionals interested in promoting contribution within P2P communities used in the workplace. These tips are based on studies of the relative amounts of contribution in different communities that use a particular P2P protocol called BitTorrent. The data and some of the discussion in this paper come from [10], and more detail about our studies of BitTorrent communities is available there.

The structure of the rest of the paper is as follows. I will briefly describe how BitTorrent works, and then I will present data from BitTorrent communities on factors that are associated with high levels of contribution. Next I will outline different reasons that exist for contributing in P2P communities. I will discuss a phenomenon which may come as a surprise: contribution in P2P communities which does not require any motivation on the part of the contributor. I will describe a general method of encouraging contribution in P2P communities, and end with three key learning points.

2. BitTorrent

BitTorrent is a P2P cooperative publishing tool, created by Bram Cohen [4]. Before the days of BitTorrent, if you had a file which a large number of people wanted to read, you would put it on a web site and allow them to download it from there. As the number of people simultaneously downloading the file increased, the bandwidth to the web site server would fill up, it would take longer and longer for readers to finish downloading the file, and eventually the site would crash and no-one could read the file at all. BitTorrent solves this problem. The users download the file using special

BitTorrent software clients. (In this paper I will use the word "peer" to mean a user's client). The file is chopped up into chunks, and the web server sends different chunks of the file to different peers. The peers swap the chunks between themselves until they have the full file. Thus the bandwidth used to distribute the file is not just the bandwidth between users and the server, but also the bandwidth between different users. This extra bandwidth allows readers to obtain the file quickly, the web site doesn't crash, and it all works nicely.

Solving this problem has made BitTorrent rather popular. According to data from CacheLogic, Inc. [3], in 2005 nearly 20% of all Internet traffic used the BitTorrent protocol. BitTorrent is used to distribute a wide variety of types of content, some in violation of copyright, others perfectly legal and respectable. There are BitTorrent distributions of open source software, pirate movies, University lectures, anime, and sermons; and several film and television companies have partnered with BitTorrent.com to use this method to publish some of their movies and TV shows [1].

Users in BitTorrent contribute bandwidth to other users by uploading chunks to them. BitTorrent is designed to provide an incentive for uploading chunks to other users currently downloading the same file. The incentive works on a tit-for-tat basis: peers preferentially upload chunks to other peers that have recently uploaded chunks to them. By uploading to other users, you increase priority to receive chunks from them, and thus theoretically increase the speed at which you receive the file. However, this is only an incentive to upload to other users while you are downloading a file. There is also a behaviour called *seeding*, which is to upload chunks of a file to other users after you have downloaded the whole file. This is another way to contribute bandwidth. There is no built-in incentive in the BitTorrent protocol to seed: but people do it nevertheless. The BitTorrent client is designed so that if you leave the client open for a while after you have finished downloading a file, the client will continue to seed until you close it.

3. Results from BitTorrent communities

There are different communities of BitTorrent users, who share files of a particular genre. Typically, a web page for the community will indicate which files are available to community members, and where to upload them from. Our study [10] used data from several different BitTorrent communities, comparing the amount of contribution between different communities and between different files shared within the same community, in order to determine factors associated with a high contribution levels.

We were interested in whether the BitTorrent protocol did succeed in encouraging contribution in the form of uploading. We got data from three different BitTorrent communities on how much peers had uploaded and downloaded. We found that in each of the communities, less than 6% of the peers had uploaded nothing. These 6% were not necessarily unwilling to contribute; they might have been willing to upload but not have yet downloaded any chunks requested by other peers, or they might have been prevented from uploading by firewalls. In contrast, the percentage of peers in Gnutella who don't contribute anything has been measured as between 13% and 85% [5], depending on the details of the measurement method. Gnutella is a P2P system with roughly similar demographics and content type to the three BitTorrent

communities, but has a different protocol. So it looks like the design of the BitTorrent protocol results in more contribution.

One possible reason why most peers upload something is that most BitTorrent clients don't have a setting in which you upload nothing. To do make it do that, you have to hack the client. We think that the default setting is probably a major reason why most peers upload. In addition, BitTorrent's tit-for-tat incentive means that if you upload a reasonable amount in general you should be able to download the file faster, although we've discovered that that doesn't always happen. So it looks like the incentive is working.

Next we looked at the amount of seeding in BitTorrent communities. Remember seeding is uploading after you've finished downloading the file. It's a sharing behaviour that is not affected by the tit-for-tat incentive.

We investigated the length of time that peers in one BitTorrent community continued seeding. Most of the seeders we identified stopped seeding (as a result of the user closing their client and disconnecting from the community) after at most a few hours. But there were a small number of seeders who continued seeding for the full ten day period for which we sampled the community. It is possible that these were seeding files that they had originally published, deliberately contributing bandwidth to help with the file's distribution. It is also possible that some of the users just forgot to close their clients.

We found an intriguing result when comparing the amount of seeding in the four communities for which we had data and which used the standard BitTorrent protocol. Two distributed only legal content, and two distributed copyright-violating content. We found that the legal communities had more seeding than the illegal ones. The difference was statistically significant, and not explained by anything else we could measure. We think this is because some people downloading illegal content are reluctant to stay connected for long to dodgy sites, so they close their clients quickly after the downloading is finished. (I'd like to make it clear here that we didn't download any illegal data ourselves. We used information provided by the community administrators about the volume of data uploaded and downloaded, not the actual data.)

A BitTorrent community is basically a decentralized system, but it does have a small centralized component which tells peers where to find files and which other peers to upload chunks to. Some communities have extra mechanisms which make use of the centralized component and which have the effect of promoting contribution. One of the communities we looked at, easytree, originally had a high level of seeding which the admins attributed to a sharing culture among their users. Then there was a big influx of new users without this culture, and the amount of seeding went down. In order to raise it again, they introduced a mechanism called sharing ratio enforcement. This mechanism is also used by some other BitTorrent communities. Under sharing ratio enforcement, if you haven't uploaded at least a quarter of what you've downloaded, measured in megabytes, then you can't access any new files. This gives an incentive to seed.

Broadcatching is a way of automatically downloading new content when it appears. It's used in BTefnet, which distributes television programmes, and some other BitTorrent communities. The new content is announced using RSS, really simple syndication, and you can get your client to automatically download, for example, any new Dr Who episode when it's announced. After it has finished downloading, your client remains open and seeding until you notice and close it. So as a side effect of this mechanism, extra seeding may take place, carried out by clients that the user doesn't yet notice are open.

We found that there was more seeding in easytree and in BTefnet than in any of the four BitTorrent communities that used the standard BitTorrent protocol. So the extra mechanisms appear to be working.

Another way that users can contribute to a P2P community is by publishing files. We looked at the users who had published at least one file during a 21-day period in two BitTorrent communities. In both communities, just over half of these users had published just one file, and about 95% contributed 10 or fewer. A small number of users published a *lot* of files (over 50). There was a similarly skewed distribution of the total volume of data published by users. We think that one factor determining this distribution is users' technical capabilities: the distribution of the amount of bandwidth available to a user is also heavily skewed, and users with very low bandwidth are unlikely to be prolific publishers. However, just having the appropriate technical capability is not enough to ensure contribution; there is also the question of motivation. It's a good idea to pay attention to the prolific contributors in your P2P community. You should identify your frequent contributors, try to find out what motivates them, and aim to address their needs, because these few users will have a large effect on the system.

In the next section I will describe different motivations for contributing to P2P communities.

4. Motivations for contribution to P2P communities

There are two strains of research into contribution to P2P communities. The first strain comes from economics, and assumes that users aim to maximize their own utility. Several incentives for contribution in P2P communities have been developed which are designed to ensure that users who contribute will receive something in exchange which increases their own personal utility. This strain of research is exemplified by papers in the NetEcon conferences (see eg. http://netecon-ibc.si.umich.edu/). The second strain comes from social science (and from observations of user behaviour) and observes that users contribute to P2P communities for social, ideological, identity-related or altruistic reasons. An example of this second strain is Kevin McGee and Jorgen Skågeby's work on gifting technologies [6]. Designers of P2P communities have tended to provide either economic-type or social-type incentives for contribution, according to their assumptions about what will motivate users. So which of the two strains of research is correct? Do users seek to maximize their own utility, or to increase the utility of others?

The answer appears to be that both are correct. We found a wide variety of different motivations for contribution in P2P communities, ranging from the purely egoistic to the purely altruistic. Both egotistic and altruistic motivations for cooperation were simultaneously in action for the same community, and sometimes for the same user. The implication is that when if you want to encourage contributions in your P2P community, you should aim to harness *both* economic-type *and* social-type motivations for contribution. I will now describe some of the different motivations that we observed, and some ways that they can be harnessed.

4.1 Economic motivations

The most straightforward example of an economic type motivation in an online community is when contributions to a community are paid for: either with actual currency, or with points which can be exchanged for resources within the community (see eg. http://www.kongregate.com, www.mojonation.net). A slightly less direct example of a personal reward in return for a contribution is the incentive for uploading built into BitTorrent, where users who contribute bandwidth for others to use obtain bandwidth for their own use in return. Similarly, the P2P community OurGrid (www.ourgrid.org), in which users share computing power, encourages contributions through its resource allocation method, in which peers give priority to satisfying requests for computation power by peers that have donated a large amount to them in the past. Several P2P communities use reputation schemes to mediate this type of reciprocation; by contributing to the system a user will increase their reputation, and users with high reputation obtain greater utility from the system. A controlled experiment by Paul Resnick et al. [9] found that eBay purchasers were prepared to pay 8% more for a good if its seller had a high reputation than if its seller was a newcomer.

Another form of economic motivation is the wish to avoid punishment. Sharing ratio enforcement in BitTorrent works by punishing users who do not contribute enough.

Economic motivation does not necessarily require a guaranteed link between contribution and personal benefit: it may be enough that contributing increases the probability of receiving a personal benefit. One example of this is resource pooling, in which a group of users share a set of resources to decrease the likely expense or inconvenience of individually obtaining all the resources to meet their own needs. Another example is the promotional distribution of music tracks to promote album sales and concert tickets.

4.2 Social motivations

There is no economic encouragement for seeding in BitTorrent, however there is social encouragement: notices on BitTorrent clients and sites inform users that it is considered polite to leave your client open for a while after you have finished downloading a file. It seems likely that this setting of a social norm of contribution does have the desired effect of encouraging contribution. In easytree, the sharing culture of the original users led to high contribution levels - and the loss of this culture decreased the amount of contribution. Users who contribute for social reasons may not necessarily be seeking the good of the whole community: they may just be aiming

to impress (or assist, or entertain) their friends. This appears to be the motivation for a large number of blog writers, for example. The community-wide distribution of their contributions means that these users can potentially benefit a large number of people outside their circle of friends. Some contributions in P2P communities are motivated by ideological reasons, for example to promote a user's favourite genre of music or type of software: see [6] for some examples. Finally, purely altruistic contributions do exist: for example, in a survey of SETI@home contributors, over half said that they had contributed computer power to the project "for the good of humanity" [11].

Some practical tips for ways to design your community so as to harness these types of motivations include making visible a measure (or several different measures) of the amount that each user has contributed; providing extra publicity for particularly generous or prolific contributors; linking digital content to a profile page for the user who contributed it; allowing users to set up teams or buddy lists which form subsets of the community, with tools for communicating with each other; setting goals for the community or for subgroups, with information how these goals may be beneficial for a wider constituency and information on progress; induction for new members to assist their use and help build social norms; and integrating the online community with offline activities and organizations [7].

4.3 No motivation

We found that a surprising amount of contributions were made with *no* motivation by the contributor. This type of contribution was a default, or a by-product of an action that the contributor was carrying out for some other reason. For example, there are users of Skype (www.skype.com) who have no idea that are contributing bandwidth to the Skype P2P community: they just know that they are getting free international telephone calls. As Dan Bricklin [2] has pointed out, Napster managed to create a huge index of music because Napster users automatically contributed to the index whenever they carried out an action such as burning a CD. In BTefnet, additional seeding is a by-product of the feature of automatically downloading a track, and this plausibly explains why BTefnet has high levels of seeding. The fact that uploading is the default in BitTorrent has the result that most peers upload.

5. A general method of encouraging contribution

A general principle for encouraging contribution in your P2P community is that people will contribute more if it's easy for them to do so.

There are several ways that you can make it easy to contribute. Look for ways of increasing the effectiveness of contribution: for example, effectiveness can be increased by tools that match users who require a particular resource with users who are able to supply it. Reduce technical barriers to contribution. It is definitely worth spending time to make your system as simple to understand and to use as possible. Make it safe to contribute: as the difference in contribution levels in legal and illegal BitTorrent communities illustrates, you are likely to get more contribution if you reduce risks associated with contribution, whether those are legal, ethical or security risks. If you can, make contribution possible with no effort at all by the user, by making it default behaviour of the system, as is the case for uploading while downloading in BitTorrent, or by making it a byproduct of actions carried out by

users for their own benefit. If it is not possible to enable contributions without some effort by the contributing user, then it is a good idea to enable different kinds of contribution requiring different levels of effort. For example in BitTorrent uploading while downloading requires no effort, seeding only requires the minor effort of keeping the client open after downloading, and publishing new content requires more. Users who do not wish to go to the trouble of publishing new content can (and do) still contribute in BitTorrent in other ways.

The search for ways to make the system easier to use can be a spur to additional creative improvements. As the Brazilian writer Mário Quintana said: *Laziness is the mother of progress. If Man hadn't been too lazy to walk, he wouldn't have invented the wheel.*

6. Learning points

Three key learning points of this paper are:
1. Users contribute to P2P systems both to maximize their own personal utility, and for social/cultural reasons. Administrators should consider both types of motivation when designing their systems to encourage contribution.
2. P2P communities typically include a small number of prolific contributors who have a large effect on the system. Administrators of P2P systems should identify the needs and motivations of their prolific contributors and make it a priority to support these users.
3. The easier it is for users to contribute, the more contribution there will be. So to encourage contribution, make it easy.

REFERENCES

[1] BitTorrent, Inc. BitTorrent strikes digital download deals with 20th Century Fox, G4, Kadokawa, Lionsgate, MTV Networks, Palm Pictures, Paramount and Starz media. Press release, 26 November 2006.
[2] Dan Bricklin. The Cornucopia of the Commons. *O'Reilly P2P conference*, 14 Feb 2001. http://www.bricklin.com/speeches/c-of-c/
[3] CacheLogic, Inc. P2P in 2005. Research report. http://www.cachelogic.com/home/pages/research/p2p2005.php
[4] Bram Cohen. Incentives Build Robustness in BitTorrent. *First Workshop on the Economics of P2P Systems (P2PEcon),* Berkeley CA USA, 2003. http://www2.sims.berkeley.edu/research/conferences/p2pecon/papers/s4-cohen.pdf
[5] Daniel Hughes, Geoff Coulson and James Walkerdine. Freeriding on Gnutella revisited: The bell tolls?" *IEEE Distributed Systems Online* 6 (6), pp. 90–97, 2005.
[6] Kevin McGee and Jorgen Skågeby. Gifting technologies. *First Monday* 9 (12), Dec 2005. http://firstmonday.org/issues/issue9_12/mcgee/
[7] Miranda Mowbray. Designing online learning communities to encourage cooperation. In: Niki Lambropoulos and Panayiotis Zaphiris (eds), User-centered design of online learning communities. Hershey, PA, USA: Information Science Publishing, Chapter V, pp. 102-120, 2007.
[8] Andy Oram (ed), Peer–to–peer: Harnessing the power of disruptive technologies. Sebastopol, CA, USA.: O'Reilly & Associates, 2001.
[9] Paul Resnick, Richard Zeckhauser, John Swanson and Kate Lockwood. The value of reputation on eBay: a controlled experiment. *Experimental Economics* 9(2), pp.79-101, 2005.

[10] Matei Ripeanu, Miranda Mowbray, Nazareno Andrade and Aliandro Lima. Gifting Technologies: a BitTorrent case study. *First Monday* 11(11), Nov 2006.
[11] SETI@home. SETI@home poll results. 2002. http://seticlassic.ssl.berkeley.edu/polls.html