DRAFT
COMMENTS WELCOME

The Economics of Peer-To-Peer Networks

Ramayya Krishnan, Michael D. Smith, Rahul Telang

{ rk2x, mds, rtelang}@andrew.cmu.edu

H. John Heinz III School of Public Policy and Management
Carnegie Mellon University, Pittsburgh, PA 15213

This Draft: August 2002

Acknowledgements: Financial support was provided by the National Science Foundation through grant IIS-0118767.

The Economics of Peer-To-Peer Networks

ABSTRACT

Peer-to-Peer (P2P) networks have emerged as a popular alternative to traditional client-server architectures for the distribution of information goods. Many new networks are being established and thousands of users are part of such P2P networks. In this chapter, we outline some important economic characteristics of these networks. We show that while the characteristics of services provided over P2P networks are similar to public goods and club goods, they have many important differences and hence there is a need for new theoretical models as well as empirical and experimental analysis to understand user behavior. In particular, we note the prevalence of free-riding on these networks may limit the scalability of such networks leading to rivalry in demand. In the same vein, we outline some incentive mechanisms that can improve functioning of these networks from the standpoint of both the individual and the collective. We also identify important research areas including privacy, copyright infringements, and liability on these networks.

1. Introduction

P2P networks allow a distributed community of users to share resources in the form of information, digital content, storage space, or processing capacity. The novel aspect of these networks is that, in contrast to client-server networks where all network content is located in a central location, P2P resources are located in and provided by computers at the edge of the network (a.k.a. "peers").

P2P networks hold out the promise of altering information sharing patterns for consumers and businesses. Consumer use of P2P networks has received the majority of press coverage in recent years — and with good reason. The Yankee group estimates that consumers swapped over 5 billion music files over P2P networks in 2001 (Dignan 2002). A recent study by Ipsos-Reid finds that 23% of the American population over age the age of 12 have downloaded MP3s over the Internet.[1] This proportion is likely to be much higher among the critical 12-21 year-old demographic. In 2000, a Gartner study found that Napster accounted for up to 75% of the traffic on some university LANs (Shuchman 2000). More recently, the March 2002 shutdown of Morpheus resulted in a 50% drop in the number of packets sent on Carnegie Mellon's wireless network.[2]

The high use of P2P networks for sharing of consumer products is reflected in the development of these networks. In the early 1990s Internet Relay Chat (IRC), which originally was restricted to passing messages between members, evolved to allow users to share files directly. Napster, founded in May 1999, popularized this form of information sharing by simplifying search and retrieval and improving scalability versus IRC implementations. Before being shutdown through

[1] Source: http://www.ipsos-reid.com/media/dsp_displaypr_us.cfm?id_to_view=1414
[2] Conversation with Marvin Sirbu, April 2002.

legal action by the Recording Industry Association of America (RIAA), Napster claimed to have 38 million registered users.

Napster is an example of a P2P network with a centralized database of content.[3] Users who logged into the Napster network upload a list of the content they are sharing to a mirrored set of computers owned by Napster. Users who wanted to access content on the network would issue a query to this central database that would then point them to a list of peers who had the content. The searching peer would then download the file directly from the peer who provided the content. This architecture gave users a high degree of visibility to content on the network and thus improved the ease of user search. However, it also introduced a vulnerability to the network: The network ceased to function if the central servers were shutdown, as a judge ordered Napster to do following a lawsuit issued by the RIAA.

Other P2P file sharing networks have adopted architectures that eliminate a central file catalog. For example, version 0.4 of the Gnutella protocol featured a distributed catalog of files. To connect to a network, a Gnutella peer would establish simultaneous connections to 3-5 other peers on the network. These peers would also maintain simultaneous connections to other peers. In this mesh architecture peers would maintain a list of their own files. A peer can issue a search to its neighbors with a time-to-live (TTL) definition (typically 7). This search will be forwarded through the network up to the number of times defined by the TTL field. If a peer that receives the search message has the requested content, they will send a reply back through the chain to the originating computer. The advantage of this architecture is that Gnutella 0.4 networks do not have a central point of vulnerability. However, this advantage comes at a cost of reduced visibility to network content.

[3] Open Source Napster (a.k.a. OpenNap) uses the same architecture.

The Gnutella version 0.6 protocol, Kazaa networks, and Morpheus networks adopt a compromise architectural design. In each of these networks a small set of computers are selected to maintain local content databases. In Gnutella these are called ultrapeers and in Kazaa and Morpheus they are called supernodes. In both cases peers on the network indicate their willingness to serve as local content databases and are selected for this task by the network protocol. Other peers connect directly to these ultrapeers and upload their list of content as in the Napster protocol. The ultrapeers are then interconnected in much the same way as peers on a Gnutella network. Searches for content are first issued to the ultrapeer and then these queries propagate to other interconnected ultrapeers using the same TTL field discussed above. Thus, these networks have the advantage of increased content visibility and enable more efficient search. However, shutdown risks are minimized because new ultrapeers can be selected if some existing ultrapeers are disabled.

While P2P file sharing networks are among the most popular applications of P2P technology, P2P technology is also gaining adoption in a variety of other arenas. For example, P2P has gained adoption for distributed computing (e.g., SETI@Home), enterprise knowledge sharing (e.g., Bad Blue), and user collaboration (e.g., Groove Networks). Further, while most of the previous examples of consumer P2P networks have a high proportion of copyrighted content, there is no reason that Digital Rights Management (DRM) cannot be incorporated into such P2P networks. For example, the subscription-based Napster service launched in early 2002 used DRM technology within its P2P architecture to protect copyright holders.

While a great deal of research has been conducted on these network architectures in computer science departments, little is known about the economic characteristics of such networks. For example, how do user incentives impact the efficiency of such networks, what are the scalability

characteristics of the networks from an economic perspective, and what are the free-riding characteristics?

This chapter attempts to provide some background on the economic characteristics of these networks, provide a preliminary understanding of some of these questions, and identify areas for future research. This chapter proceeds as follows. In section 2, we analyze the economic characteristics of peer-to-peer networks focusing on comparing the services offered over peer-to-peer networks to traditional private, public, and club goods. In section 3, we extend this understanding of the characteristics of peer-to-peer networks to analyze the impact of economic incentives on the efficient operation of these networks. In section 4, we conclude and discuss fruitful areas for future economic research relating to peer-to-peer networks.

2. Economic Characteristics

P2P networks are emerging as a significant vehicle for providing distributed services (e.g. search, content integration and administration) both on the Internet and in enterprise. During this growth period of P2P networks, the main focus has been on the technology side of P2P. Many new improved networks and efficient algorithms have appeared in the market (see Ratnasamy et. al 2001). They have made it easier for users to participate and contribute on the network. Only recently, researchers have started to focus on the economic and social aspects of these networks. Understanding the amount of user participation, the volume of content sharing and distribution and evolving social and economics dynamic of these networks is critical for their efficient operations and survival in the long run.

While P2P networks vary in their architectural design, in all P2P networks files are transferred directly between the computers of users (a.k.a. peers) connected to the network. Further, once

these files have been delivered, the user accessing the file, by default, becomes a provider of that content. Thus, in an ideal case the provision of content on the network will scale to match the level of demand for the content. To the extent this holds, P2P networks can be modeled in the context of public goods and club goods. Public goods have the characteristics of non-excludability in supply (individuals can't be excluded from consuming the product) and non-rivalry in demand (one individual's consumption does not diminish another user's value of the product) (Hardin 1968). Clean air is a typical example of a public good. Club goods are goods that are excludable in supply but non-rival in demand (Buchanan 1965).

P2P networks share many of the properties with the public or club goods, but they are different from both of them on many dimensions. In the ideal case, P2P networks will exhibit both non-excludability (information is made available to all members of the network) and non-rivalry (consumption by one user doesn't decrease network download possibilities in the absence of free-riding). However, in the presence of free-riding, P2P networks will exhibit levels of rivalry (Asvanund, Clay, Krishnan, and Smith 2001), which distinguishes them from either public goods or club goods. Thus, it may be possible to think of services provided by P2P networks as a new class of economic goods distinct from the existing private, public, and club good classes. For example, unlike other public goods, the provider of the good is also the consumer of the good. Hence on P2P networks, a user can be a provider, or consumer, or both. This has implications for the formation and sustainability of P2P networks as shown by Krishnan, Smith, Telang and Tang (2002).

An important observation from public goods literature that extrapolates to P2P networks is the prevalence of free-riding on such networks. The failure of each consumer to consider the benefits for others of her public good provision is known as the free-rider problem. Following Hardin

(1968), people refer to it as "the tragedy of commons." In P2P networks free-riding could occur if users consumer network resources without providing resources to the network in the form of shared files. Thus, while most P2P clients enable sharing by default, users also usually have the option of disabling sharing. Users may choose to disable sharing because of legal concerns,[4] or because of the scarcity of networking or computing resources. Free-riding in this way allows users to utilize their resources (disk space, band-width, content) privately while consuming other users' resources.

Given the similarities between the services provided over P2P networks and public and club goods it is not surprising that empirical studies have found high degrees of free-riding in P2P networks. For example, Adar and Huberman (2000) find that 70% of peers on the Gnutella network do not provide songs and that the top 1% of sharing hosts return 50% of all responses. Similarly Asvanund et al (2002) find evidence consistent with increasing free-riding in larger P2P networks. Traditional economic theory predicts that the free-rider problem causes inefficient private provision of public goods and calls for a central intervention to remedy this problem. In the context of P2P networks, Adar and Huberman (2000) observe that "free-riding leads to degradation of the system performance…if this trend continues copyright issues might become moot compared to the possible collapse of such systems."

In contrast, the second observed characteristic of P2P networks is that they appear to persist in spite of these high levels of free-riding. For example, according to a CNN report the number of P2P sites totals nearly 38,000, up 535 percent in the past year. Given these two empirical facts, an important question is: how can these networks survive in-spite of wide-spread free-riding? This also leads to the issue of scaling. What can we say about the ability of these networks to

[4] For example, the FBI recently announced they would target individual P2P users who shared copyrighted files.

scale as more users join them? Again, while the technological solutions, such as better algorithms, are being proposed, the social scientists have begun to pay attention only recently. Krishnan et al (2002), using the game theoretical approach, argue that as more users join the networks, it causes more congestion. This, in turn, might lead to some users turning on their sharing. This fact, to an extent, may allow the network to scale in the long run. Hence, unlike public goods that are usually static (i.e. either they are provided or not) P2P networks may be more dynamic and scalable.

The literature on user behavior and group size may also serve as a starting point for understanding the characteristics of P2P networks. Olson (1968) conjectured that public goods are less likely to be provided in larger groups as users have more incentive to free-ride on other's contributions. This prediction is supported by the analytic models of Palfrey and Rosenthal (1984), Hendriks and Pancs (2001), and Dixit and Skeath (1999). These results are based on pure self-interest. Including the possibility of altruism on the part of users increases the provision of public goods (Andreoni and Miller 2001). This is potentially a significant factor for P2P networks as many authors have observed that altruism plays a significant role in the provision of public goods (Ledyard 1993).

3. Economic Incentives and User Behavior

One obvious question raised by the previous section is how will user behavior respond to the economic characteristics of P2P networks and how can network designers influence this behavior through incentive mechanisms. This question is particularly important in light of recent observations of the importance of incentive alignment for Information Technology design (Ba,

Stallaert, and Whinston 2001). One obvious area where economic incentives find application in P2P networks is controlling free-riding behavior on the part of users.

As noted above, free-riding behavior occurs when a user consumes network resources without providing any resources — typically in the form of sharing content files — and this situation may deteriorate in larger P2P networks where social norms are likely to be weakened (Olson 1968). A variety of solutions have been proposed to reduce the problem of free-riding on P2P networks. The most common proposal follows a pricing model, where incentive compatibility is achieved by pricing a scarce network resource (e.g., MacKie-Moson and Varian 1995 and Wang, Peha, and Sirbu 1996). In the spirit of pricing network resources, Golle Leyton-Brown, and Moronov (2001) propose to charge for the use of P2P network capacity through a system of micro-payments. Similarly, Chandan and Hogenborn (2001) analyze the use of micro-payments in the context of wireless P2P networks and find that micro-payments may be able to provide an incentive compatible solution to the free-riding problem.

However, it is also interesting to note that direct payments between peers may be impractical in many common P2P implementations. For example, it is difficult to imagine transfer payments between users of a knowledge sharing P2P network with an enterprise. Likewise, in many consumer P2P networks direct micro-payments will be difficult to implement because of the anonymous nature of network usage. In these settings it will be particularly important to develop non-priced incentives to encourage efficient behavior on the part of P2P users. Some examples of non-priced incentives could include delay times (e.g., providing priority queuing to users to share more content with the network), network membership (e.g., threatening to remove non-sharing members from the network), or peer ratings of content providers. Non-priced incentives are discussed in more detail by Krishnan, Smith, Tang, Telang (2002).

Another interesting question is how can copyright holders use economic theory to reduce the attractiveness of P2P networks for sharing of unlicensed files. In a typical scenario, a user's net utility from consuming a product is $U - p - sc$ where U is the utility of the product, p is the price of the product, and sc is the search cost associated with obtaining the product. In the case of MP3 file sharing for a record company's file sharing site to yield higher utility to a potential customer than using a copyright-infringing site the record company would need $U_r - p_r - sc_r > U_p - sc_p$ where the price to use the copyright-infringing site is assumed to be 0. What is interesting for the record companies is that, unlike in a typical situation where they can only control their own utility, price, and search cost, in the case of MP3 file sharing networks they can also influence the utility and search cost of the competing network. Record companies can influence the utility of content downloaded over copyright-infringing sites by threatening users with fines or lawsuits for illegally downloading copyrighted music. Search costs can be influenced by making it difficult to users to locate the content they are looking for.

Interestingly, both of these tactics are being used by the record companies, through the RIAA, to counter the threat of copyright-infringing file sharing. Increasing the search costs associated with finding content is a particularly interesting, and recent, tactic employed by the record companies (see Segal 2002 for example). In such a scenario, the record company would register numerous peers on the copyright-infringing site each with a set of MP3s that use the same naming structure as popular music content, but which contain no usable content.[5] The end result should be to increase the number of files a user must download before they find the "real" content they were looking for. By lowering the utility and increasing the search costs of their competitors' file

[5] The companies MediaDefender and Vidius both advertise software product to make this process easier for the record companies.

sharing networks, the record company's web sites will have more flexibility in setting the utility level of their content (e.g., by restricting the legal uses of the files) and their prices.

Another interesting policy question is the liability and responsibility in case of infringement of copyright laws? Who is responsible and who is liable? Is it the users who put copyrighted content on the networks? Is it the users who are downloading it or in some cases, is it the infrastructure provider? Recently, Sony Corp sued AT&T Broadband, Sprint, and UUNET for providing the infra-structure to users so that they can download songs from (www.listen4ever.com), a site in China (USA Today, August 20, 2002). Clearly, as P2P networks become more popular, the intellectual property rights, content distribution, and issues of liabilities have to be resolved from economic as well as a policy point of view.

Conclusions and Future Research

It is quite evident from the growth of P2P networks in last 2-3 years that they are becoming popular tool for content sharing and distribution. While the popularity of these networks has been mainly from consumer P2P networks, many enterprises have been using these networks for as knowledge management tools to share information across the enterprise.

We identify in the chapter that while the technical developments have kept pace with the growth of these networks, economic and social analysis of these networks is still in a nascent stage. Understanding P2P network operation from an economic and social perspective will be critical to developing protocols and systems to ensure the efficient operation of these networks.

Public and club goods provide a useful starting point for the economic analysis of P2P networks. P2P networks share many characteristics with public and club goods, but differ from these goods

in important ways. Since the literature on public goods is quite extensive, it allows the researchers to extrapolate some of the results from this literature. But we note that not all results are equally applicable. Since these goods are essentially a different class of product, we need to understand the mechanism of these networks carefully before applying these results. In many cases, we need different models to understand user behavior.

Another important observation we have noted is the existence of extensive free-riding on these networks and its social and economic implication. Free-riding may significantly reduce the performance of P2P networks. Besides the economic literature, group behavior and social network literature sheds some light on free-riding on these networks. It is believed that larger groups lead to more free-riding. But at the same time, altruism could mitigate these effects quite significantly. As these networks continue to grow, we need new theoretical models as well as experimental and empirical data to understand user behavior on these networks.

Another important issue related to these networks is issue of incentive mechanism. What does the economic and social engineering theory tell us about creating incentive that makes these networks socially optimal? We note some of the mechanism like pricing, membership rules, micro-payments, peer rating etc. Again, empirical and experimental analysis which can measure the impact of these mechanisms on welfare of these networks is a very promising area of research.

Finally, we note the policy and economic implications of sharing content that is copy-righted or user behavior that infringe on such rights. Firms can and do resort to legal recourse. But we mention that the issue of liability in this case is quite nontrivial. The issue of liability, privacy and intellectual property rights on these networks is another promising area of research which

has academic, policy as well as commercial implications. Many firms have also been using these networks to influence the search costs of users. Undoubtedly, as these networks become even more popular, many profit maximizing firms would exploit the opportunity of a huger member base through different means (like advertising) and create another set of research questions.

Reference

Adar, E, B. A. Huberman. 2000. Free-riding on Gnutella. *First Monday* **5**(10).

Andreoni, J., J. H. Miller. 2002. Giving According to GARP: An Experimental test of the consistency of preferences for Altruism. *Econometrica* **70**(2) 737-753.

Asvanund, A., K. Clay, R. Krishnan, M. Smith. 2001. Preliminary Evidence on Peer-to-Peer Sharing of Music Files. Telecommunications Policy Research Conference, Alexandria, VA, October 27-29.

Ba, S., J. Stallaert, A.B. Whinston. 2001. Optimal Investment in Knowledge Within a Firm Using a Market Mechanism. *Management Science* **47**(9) 1203-1219.

Buchanan, J. M. 1965. An Economic Theory of Clubs. *Economica* **32**(125) 1-14.

Dignan, Larry 2002 "Report: File Swapping Sites To Thrive." *ZDNet News*. Accessed on August 16, 2002 from http://zdnet.com.com/2100-1106-949724.html.

Dixit, A., S. Skeath. 1999. *Games of Strategy*, New York, Norton.

Golle, P., K. Leyton-Brown, I. Mironov. 2001. Incentives for Sharing in Peer-to-Peer Networks. Working Paper, Stanford University, Palo Alto, CA.

Hardin, G. 1968. The Tragedy of the Commons. *Science*. 162 1243-48.

Krishnan R., M. Smith, Z. Tang, R. Telang. 2002. The Virtual Commons: Why Free-riding Can Be Tolerated in File Sharing Networks. Working Paper, Carnegie Mellon University.

Ledyard, J. O. 1993. Public Goods: A survey of Experimental Research. *Handbook of Experimental Economics*, Princeton University Press, Princeton, NJ.

Olson, M. 1965. *The Logic of Collective Action*, Harvard University Press (Cambridge, MA).

Palfrey, T., H. Rosentral.1984. Participation and the Provision of Discrete Public Goods: A Strategic Analysis. *Journal of Public Economics* **24** 171-193.

Ratnasamy, S., P. Francis, M. Handley, R. Karp, S. Shenker. 2001. A Scalable Content Addressable Network. Proceedings of the ACM SIGCOMM, San Diego, CA.

Segal, David. 2002. A New Tactic in the Download War: Online 'Spoofing' Turns the Tables on Music Pirates. Washington Post. Page A1, August 21.

Shuchman, L. 2000. Universities Get Schooled on Napster. The Industry Standard. September 8.