11th April 2003

Mapping Peer-to-Peer Networks

Peer-to-peer (P2P) file-sharing networks are perhaps the most innovative new information space to emerge on the Internet in recent years. These new networks have attracted many users and much press attention, along with the ire of some powerful media corporations and trade organisations who feel threatened by this alternative, bottom-up mode of information distribution. Some commentators see P2P networks as offering a significantly more 'democratic', perhaps even 'anarchic', information topography and engendering new forms of online social structures completely lacking in central points of organisation. P2P breaks down the dominant information distribution paradigm on the Internet, the client-server model of the Web **[1]**.

P2P networks are, of course, most popular for sharing mp3 music files and much of this interaction involves the 'trading' of copyright materials or 'content'. As such these networks are a key component of the so called 'darknet', an informal cluster technologies and communities for users to share digital content **[2]**. The bastions of the media industries, most particularly the Recording Industry Association of America (RIAA), have attempted to choke-off the development and use of P2P networks, by various means, as they see them simply as conduits for wholesale digital theft **[3]**.

Emergent Networks

P2P network topologies emerge spontaneously from a multitude of individual actions - users simply connect to the P2P network when they like and leave when they like . The networks are typified by their lack of any centralised mechanisms for control, storage or exchange. Consequently, P2P network structures evolve, second by second, depending on which users are connected. The users provide all the necessary infrastructure to the network - the hard-disk storage, the processing cycles, the bandwidth and, of course, the vital information content. In the world of P2P the users are both the providers and the consumers. Unlike other information sharing networks, such as gopher or WWW, P2P networks are genuine users' networks: they have no external existence beyond those people who are on the network at a given time. Clay Shirky has argued that the key distinguishing feature of P2P networks is the shift in power to the 'edges' of the Internet, as they are "leveraging previously unused resources, by tolerating and even working with the variable connectivity of the hundreds of millions of devices that have been connected to the edges of the Internet in the last few years" **[4]**. Users form their own information space which can be immense in size, although they are only ever aware of, and concerned, with their immediate neighbourhood of peers.

Thedevelopment of P2P protocols, and the free software applications used to browse the resulting networks, has been an area of great creativity in the last couple of years, sparked in large measure by the runaway success of Napster, first released in the summer of 1999. Although Napster is no more, it has spawned a growing number of other P2P networks, such as *Kazaa*, *Gnutella* and *FreeNet* **[5]**. The particular topological characteristics of these different network spaces (especially the ways in which they propagate queries and responses across the network) vary depending on the exact P2P protocol design employed, but they are all ad-hoc and fluid, linking together multitudes of PCs on the edges of the Internet. Of course, many of the underlying concepts of fully distributed networks, where all nodes are peers and co-operatively share the overheads of running the network, can be traced back to well before the Napster- inspired, mp3 file-sharing boom of the last few years. For example, peer network structures have long been a popular means of structuring collaboration systems and instant messaging: so-called person-to-person networks. In many ways, contemporary P2P developments and growth are commensurate with the original Internet concept of co-operating networks without a centre. They are also a classic example of a 'disruptive technology', that is seen as threatening to many entrenched interests but also offer huge potential for innovation and overall market growth **[6]**.

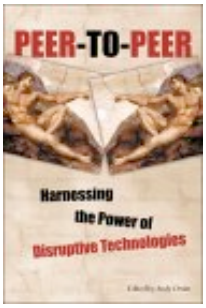
Revealing P2P Networks

Further Reading:

» For more information about this subject, the following resources are recommended.

[1] Dan Bricklin's paper "A Taxonomy of Computer Systems and Different Topologies: Standalone to P2P", June 2001.

Peer-to-Peer: Harnessing the Power of Disruptive Technologies, edited by Andy Oram (O'Reilly, 2001).



Buy This Book

[2] See "The Darknet and the Future of Content Distribution", by Peter Biddle, Paul England, Marcus Peinado and Bryan Willman, *2002 ACM Workshop on Digital Rights Management*, 18th Nov, 2002 Washington DC, USA. [word doc format]

[3] For a readable introduction to contentious and ongoing battle between P2P networks and the 'content' industries see Janelle Brown's article "The music revolution will not be digitized", *Salon*, 1st June 2001.

[4] "What is P2P…. And What Isn't", by Clay Shirky, O'Reilly Network, November 2000.

[5] See the O'Reilly P2P Directory, maintained by Lisa Rein. This gives you an idea of the large number of the different p2p networks.

[6] For insightful discussion of file sharing networks, see:

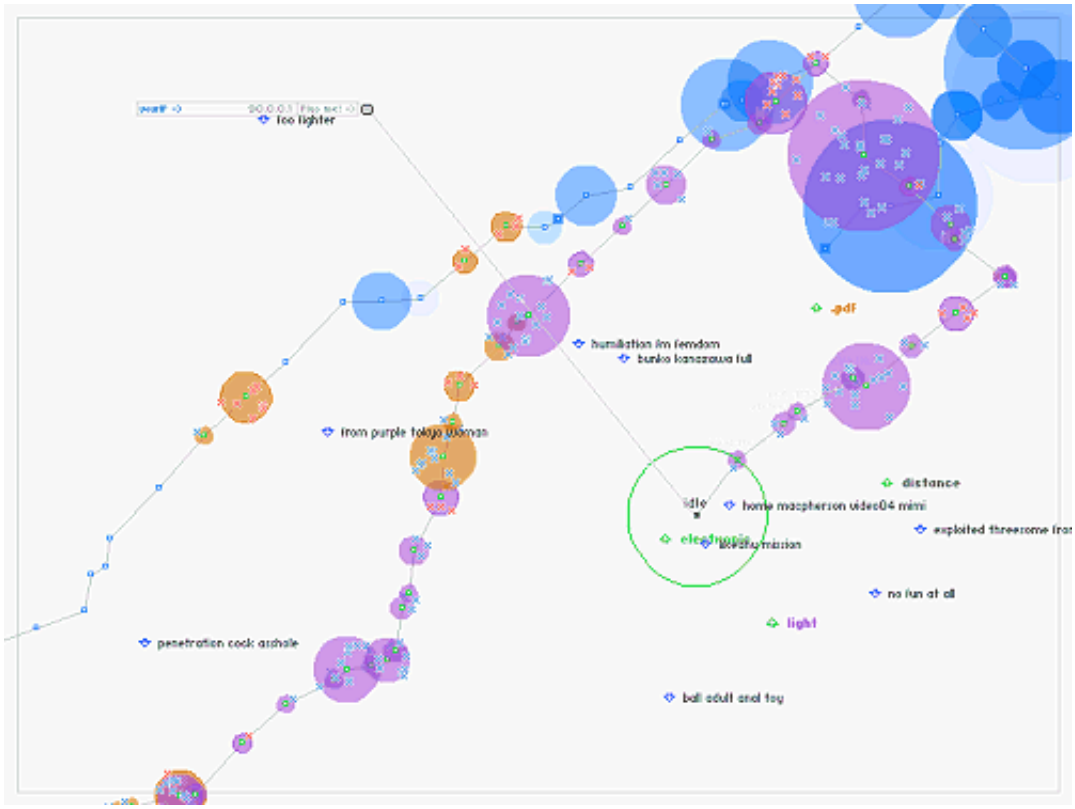
"Piracy is progressive Taxation, and Other Thoughts on the Evolution of Online Distribution", by Tim O'Reilly, December 2002.

John Snyder and Ben Snyder, *Salon*, 1st February 2003.

[7] Details and free download of Minitasking is at www.minitasking.com

The social networks that form the fabric of our lives are structured by space and time. P2P can fundamentally be seen as social space and as such it structured by topological relations in space and time of its users. Yet, the complex time-space topologies of P2P networks, like Gnutella, are never normally apparent to users, being hidden from view behind simple interfaces. After all, users do not need to be able to see the structure of the network to use it. But what would the time-space structures look like if the interface became somehow transparent, thereby revealing the fluctuating patterns of nodes and links as the network evolves? How might one map the time-space structures of P2P networks? A new application, called *Minitasking*, provides one interesting answer **[7]**.

Minitasking is a visual client for browsing the Gnutella network, released in April 2002. It provides a very different interface to P2P network spaces compared to the norm. According to the readme file that comes with the Minitasking, its aim is to give users "*…a visual manifestation of the properties of dynamic and temporarily created networks and introduces transparency to the exchange of data and network instability.*" The Gnutella network is mapped by *Minitasking* using the notion of a visual timeline with new nodes being added as they are encountered. The end result does not look much like a conventional network maps, being more like a growing necklace of beads. Yet Minitasking's visualisations are strangely compelling, showing us the time-space 'bubbles' of the local Gnutella network neighbourhood around one's machine. It is fascinating to watch the ebbs and flows of queries and replies that pop up on the map.

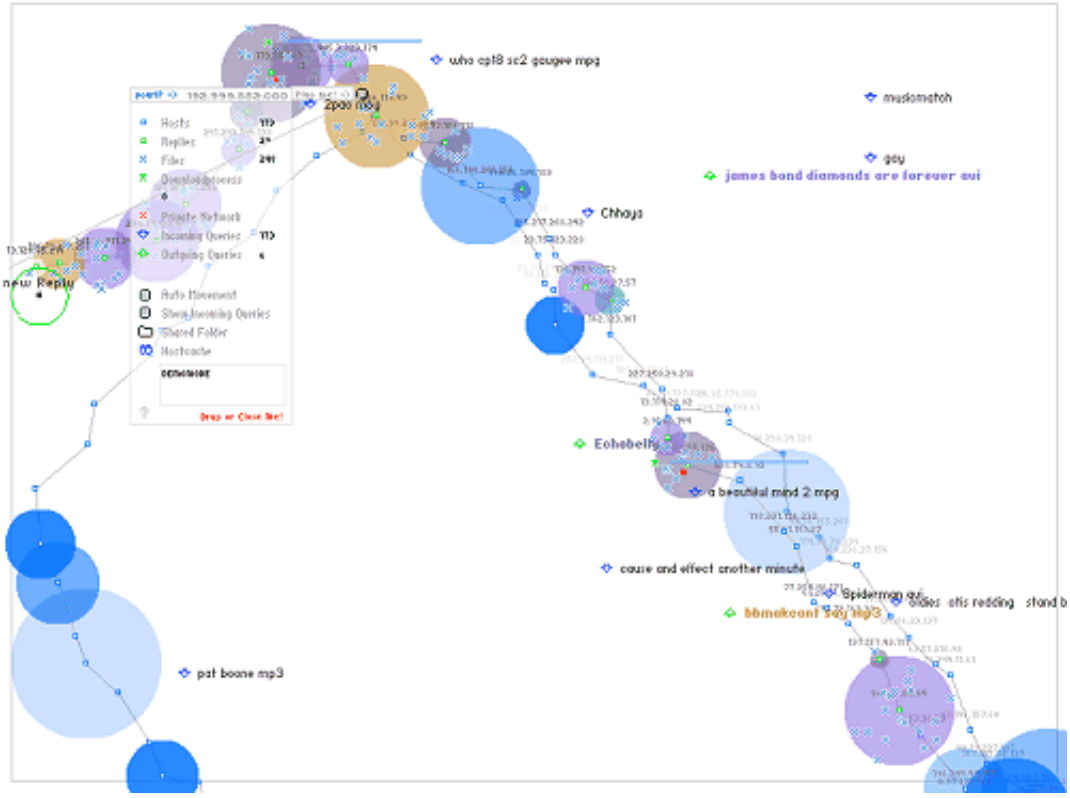


A screenshot of Minitasking showing a fragmentary view of a small part of the Gnutella network.

Minitasking was created by *Schoenerwissen*, an artistic research and design group of Anne Pascual and Marcus Hauer, in Berlin. It comes very much from the tradition of innovative net.art type software, and should be interpreted in terms of experimentation rather than efficiency; as the README says "*… usage is intended to depend more on chance than on routines*". Hauer and Pascual both studied media design and sciences at the Academy of Media Arts in Cologne going onto form *Schoenerwissen* in 1998. A literal translation of the German word *schoenerwissen* means 'more beautiful knowledge' and ***Map of the Month*** asked them a few questions, via email, to find out more about the way Minitasking maps Gnutella.

"Our objective", Hauer and Pascual explain, *"was to show the 'invisible' elements of data processing, to reintroduce technological elements like a protocol in the use of software as a way to adopt the dynamics and structure of networks."* The limits and boundaries of the Gnutella protocol forced them to work with the most undetermined parts of the network, the

temporal flow of space. This led them to conceive and build *"… a new form of the software and we reshaped our vision to produce not only a visual interface, but"*, following the ideas of media theorist Pit Schultz, *"...to get rid of usability in favour of visibility."* They focused on the functionality of the Gnutella P2P protocol and designed a very novel, and very minimalist, visual front-end to it.



Another screenshot of Minitasking in action.

The interface of *Minitasking* is a single large, dynamic map, which displays the evolving relationships of Gnutella nodes on the network, along a single growing timeline. Because of the nature of the Gnutella network protocols, one can never get a complete map, but can only see the structure viewed from one's own particular vantage-point in virtual time and space. *Minitasking* represents the active nodes it encounters in the local Gnutella neighbourhood as bubbles that vary in size and colour depending on the number of files they are sharing. The transparency of the colour of the circle gives an indication of distance away from one's own vantage-point - that is, a topological distance measured in network hops. So, the lighter the circular bubble, the further 'away' it is. When a mouse passes over a bubble, a pop-up displays useful data on the node like IP address, number of files and hop distance. When you submit a query to Gnutella, via *Minitasking*, the colour-coded matches are shown in another bubble. Double-clicking on a file of interest will start them downloading. The social nature of Gnutella is also apparent in *Minitasking* as you can see the 'voices' of other users in terms of queries they are making. These incoming queries are themselves revealing in the oftentimes specific and highly specialised content being sought after.

A close up view of part of a Minitasking map.

[8] See for example, "Modeling Large-scale Peer-to-Peer Networks and a Case Study of Gnutella", by Mihajlo A. Jovanovic, Laboratory for Networks and Applied Graph Theory, University of Cincinnati, 2001.

"Mapping the Gnutella Network: Properties of Large-Scale Peer-to-Peer Systems and Implications for System Design", by Matei Ripeanu, Ian Foster and Adriana Iamnitchi, *IEEE Internet Computing* (Vol. 6, No. 1, 2002). [pdf format file]

"A Quantitative Analysis of the Gnutella Network Traffic", by Demetrios Zeinalipour-Yazti and Theodoros Folias, Department of Computer Science, University of California Riverside, 2002.

[9] "Free Riding on Gnutella", by Eytan Adar and Bernardo A. Huberman, First Monday (Vol. 5, No. 10, October 2000).

[10] See the Gnutellavision homepage and their paper "Animated Exploration of Graphs with Radial Layout" by Ka-Ping Yee, Danyel Fisher, Rachna Dhamija, Marti Hearst Proceedings of the IEEE Symposium on Information Visualization 2001 (Infovis 2001), October 2001.

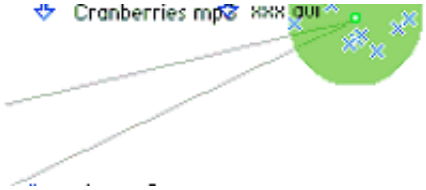
[11] The Notre Dame networks research group homepage.

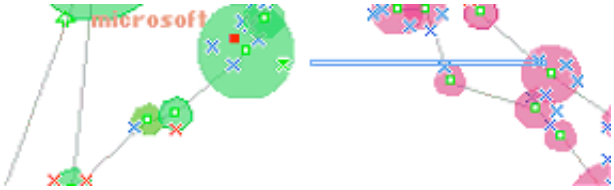
A-L Barabási, *Linked: The New Science of Networks* (Perseus Publishing, 2002).

Buy This Book

[12] *The Rise of the Network Society*, by Manuel Castells (Blackwell, 1996).

Buy This Book





Other analysis and views of Gnutella

While there have been several research projects interested in the quantitative measurement and analysis of the structural properties of Gnutella network **[8]**, it would be fair to say no one knows much about what is going on inside the P2P network as a whole, and over time. However, it seems quite likely that P2P network spaces are being intensively monitored, and perhaps visually mapped, by corporations and trade bodies like RIAA. Some activists suspect that the media industries, and people working directly and indirectly for them, are forcibly trying to change the formulation of P2P space to make it less useful through, for example by deliberate 'pollution' with multiple fake files and floods of false queries. Indeed, controversially RIAA has been seeking new laws from U.S. Congress to give content copyright holders the right to P2P networks and users machines.

A One of the major concerns with P2P networks like Gnutella has been their ability to scale capacity, in an efficient manner, to cope with many millions of users. This 'scaling' issue is further compounded by concerns that P2P is especially prone to 'selfishness' as users of the network are not carrying equal overhead loads, with many wanting to 'take' without 'giving' in return. As researchers, Eytan Adar and Bernardo A. Huberman found in their empirical study of a day in the life of Gnutella in 2000, "… we established that almost 70% of Gnutella users share no files, and nearly 50% of all responses are returned by the top 1% of sharing hosts" **[9]**. Clearly, the network is not as evenly distributed as the ideals of P2P would suggest.

There are many possible ways to map P2P network structures. Beside *Minitasking*, the most notable visualisation work was the Gnutellavision project undertaken in 2001 by Ka-Ping Yee, Danyel Fisher, and Rachna Dhamija, at University of California Berkeley, producing interesting radial radar maps of Gnutella **[10]**. There is still scope to try to undertake a more comprehensive mapping of the whole Gnutella network space.

Understanding Networks

Being able to see and understanding networks, is increasingly important. Many aspects of the human and physical world can be modelled as networks. Mapping the topological structure of a network, along with understanding the behavioural rules of how individual nodes become linked, provides a powerful explanation of the underlying processes at work. Going way beyond P2P networks on Internet, a number of scholars have advanced claims that networks are keys to explaining much of the complexity of the world that has not been easily understood with more traditional formulations.

One of the leading researchers developing new theories of network structure is the physicist, Albert-László Barabási, who runs a research group at the University of Notre Dame studying them. In his recently published book, ***Linked***, he convincingly argues for the explanatory power of a "new science of networks" in a wide range of social, economic, technological and biological realms **[11]**. As Barabási says, *"Today we increasingly recognize that nothing happens in isolation. Most events and phenomena are connected, caused by, and interacting with a huge number of other pieces of a complex universal puzzle. …. We have come to grasp the importance of networks."* Most significantly, Barabási goes onto argue that research must begin to move beyond just looking at network structure to consider the actual dynamic processes that the network supports, to begin to model what actually flows over the links; *"Networks are only the skeleton of complexity, the highways for the various processes that make our world hum."* I would argue that part of the innovative significance of Minitasking is that it tries to go beyond topology to map out spatio-temporal patterns of traffic.

In the realm of the 'big theories' for understanding society and the economy, the noted social theorist Manuel Castells has used the concept of networks to explain the changing dynamics of the world through globalisation and the information revolution. In ***The Rise of the Network Society*** [**12**], Castells argues powerfully that "*… as a historical trend, dominant functions and processes in the information age are increasingly organized around networks. Networks constitute the new social morphology of our societies, and the diffusion of networking logic substantially modifies the operation and outcomes in processes of production, experience, power and culture*" (page 469). Clearly Gnutella is but a minuscule part of the whole Network Society.

So what can we discern of Gnutella's time-space structure from Minitasking? "*What you see is not a network!*" according to when asked about what can one learn from observing Minitasking time-space maps of Gnutella. "*The potential of a Minitasking maps is not to provide the 'perfect' view of the Gnutella network*", say Pascual and Hauer, "*but to point the attention to obscure details by observing them for a while.*" Although one suspects that others, probably with less benign motives will try to survey and map out P2P space with much great precision and clarity.