

Privacy In Peer-to-Peer Networks

Mika Suvanto
Helsinki University of Technology
mika.suvanto@hut.fi

Abstract

A big part of Internet traffic is nowadays peer-to-peer (P2P) traffic. Various peer-to-peer softwares have millions of users, sharing data by means of decentralized network. Well-known examples of such programs are Kazaa [1], Gnutella [2], and eDonkey [3], which allow their users to search and share their files within the P2P community. The Internet telephony system, VoIP, is another interesting and widely used P2P application.

This paper discusses the privacy aspect of today's P2P networks and the information that can be collected from the users of P2P software. The importance of the collected information and possible ways to use it are also discussed. The main focus is on currently available file-sharing applications and analysis of how they handle sensitive information.

KEYWORDS: peer-to-peer, privacy, Kazaa, Freenet, MUTE, VoIP

1 Introduction

File sharing within a P2P network is a relatively new application for end-users. The first widely used software was Napster, which appeared in 1999. In 2001, Napster was judged illegal because of copyright violations [4]. After this, the P2P networks have evolved into a completely decentralized model - there is no central server, instead the clients discover each other independently.

Privacy within P2P networks requires attention from the user. The user has to know how to use the software and what kind of information is being shared. It is quite possible to share the entire hard drive, including sensitive information such as mailbox and private documents [5]. The user has to make sure, that the shared documents do not contain personal information which could be misused. Chapter 2.4 gives more details about this.

Another aspect of privacy is the information that is being sent by the P2P software. To establish a connection, we need some essential information, such as the IP address. If this address is used openly in P2P networks, the identity of the user is easy to discover. This is a problem with many currently used P2P applications, such as Kazaa. These privacy threats of file-sharing and VoIP applications are discussed in chapter 2.

In chapter 3, we take a closer look at the technology of common P2P applications. Chapter 4 introduces some new applications, Freenet and MUTE, which address the privacy problem within P2P networks. Finally, chapter 5 discusses

the real threats and their effects from the user point of view.

2 Privacy Threats

2.1 Tracks that we are leaving

To use any program for Internet communication, we need to reveal some information about ourselves, such as an IP address and a port number. This information combined with time is enough to identify the peers, at least the ISP (Internet Service Provider) has the required knowledge. This easily collected information was used for example by the Recording Industry Association of America (RIAA) when they sued P2P users for copyright violations [6, 7].

A server software has access to much more information than just the IP and port number. This might be a severe security problem, because some of the P2P programs are closed-source code. It is impossible to verify that such a program sends only the information that is intended, and nothing else. These applications often advertise their version numbers, and the underlying operating system could be identified, at least if the application is available for some specific operating system only. This kind of information might be used for example for hacking purposes.

2.2 Information Collecting

Using computer networks always reveals some data to other parties. This data may be limited and useless alone, but it can be combined with data from other sources. This may give a valuable information about the user. The VoIP (Voice Over IP) is a good example and an interesting target for misuse.

2.2.1 VoIP applications

Basically, VoIP allows Internet users to call each other using a VoIP software. Speech and control data is transferred via Internet without the normal telephone network infrastructure. Recently, VoIP has gained plenty of new users. The most successful client software by far is Skype [8]. Skype makes also calls to the regular phones possible. This is done with PSTN gateways. Skype is a closed-source software, and thus its security cannot be easily verified.

A telephone network is much more controlled than Internet, and only a limited number of people have an access and required knowledge to tap the phone lines. Telephone calls are also protected by strict laws in many countries. All this protection becomes useless by VoIP calls. It is quite easy to eavesdrop Internet (and VoIP) traffic, at least in a local

area network. The VoIP protocols usually offer some kind of encryption to prevent this. It is worthwhile to note that encrypting the payload only is not enough. If this is the only protection, it is possible to collect data about who is talking to who, when, how long and similar information. This information is something that must be kept confidential.

The VoIP is vulnerable to many same threats as other Internet applications. Although many of these threats are currently theoretical, the caller ID spoofing is one realized threat that seriously affects the VoIP users [9]. The attacker may fake his ID and pretend to be someone he/she really is not, and in some VoIP applications, this is quite easy [9]. This enables many serious identity attacks - for example, an attacker could call the company's computer support and pretend to be a legitimate user who has forgotten his password. The support staff relies that the call comes from the users office phone since the number is correct, and sets the new password. Similar attacks are possible in many ways with only the creativity of the attacker as a limit.

2.2.2 File-sharing applications

Not only the VoIP, but the common file-sharing applications are possible targets of information gathering. For example, the keywords that are used in searches might be interesting, as well as the files downloaded and served by a peer. These could be combined with the IP address and over time, form a comprehensive database about the user. The file-sharing application should protect this information, but commonly used second generation applications fail to do this, since they can not offer true anonymity. At least the peer from where the data is downloaded knows the downloader's IP address and the content of the downloaded data. This way it is possible to keep a track of the users who are interested of his/her content.

2.3 Spyware, malware and similar leeches

Another privacy risk with P2P application is spyware. Spyware is software that collects information about the user and sends it to a third party without user's knowledge. Many P2P applications contain software that can be considered as spyware. They are definitely a privacy risk, because they collect and transmit much sensitive information about the user. This information is valuable for companies who can gain a great amount of information about their potential customers. Many freely available applications are bundled with spyware, for example Kazaa [12]. The spyware programs are needed in order to use such applications. The end user has little ways of knowing what these programs are doing in his/her machine. Even the software firewalls are useless, because the spyware programs can use the same network connection as their host program, which needs the server rights.

Spyware, adware and malware risks are not limited to information collection. They always consume computing and network resources. This might be a small amount for single user, the costs of spyware are significant for companies. According to [10], it is suspected that half of PC software failures are caused by poorly coded spyware programs. This means increased need for user support and loss of productivity.

Peer-to-peer networks are a good source of spyware, viruses and similar malicious software. According to [11], in Bruce Hughes's investigation 45 % of executable files downloaded from Kazaa P2P network contained such code. As the sources in P2P network are not trustworthy, programs downloaded from P2P networks are a threat to the security of the system and this way to the privacy of the user. If the peer-to-peer networks are to be used in software distribution, there is much to be done to establish a reliable way to ensure that downloaded software works as expected, and does not contain any spyware/virus/malware leeches.

2.4 Sharing files without intention

Using P2P software turns a computer to a server. Running a secure and controlled server in open Internet is not a simple thing that anyone can do. Because of this, servers are usually administrated by professionals, not by end-users. Millions of users prove that configuring and using P2P software is easy, at least at first glance. But using file-sharing application in a controlled way is much more complicated.

Nathaniel S. Good and Aaron Krekelberg performed an interesting user study about Kazaa usability [5]. They showed that it is possible to share even the whole hard drive without knowing it. Many email inboxes, text documents and system files were found from Kazaa network. As probably no-one wants to share this kind of material with everyone, there must be some kind of misconfiguration in the application. As God's and Krekelberg's paper shows, the usability of P2P application relates closely to its security, and Kazaa's usability could be improved in many ways. The default values of the application should be so strict, that the user has to willingly configure the application to share his/her files. Unfortunately, this is something that the P2P developers do not want to do. There must be as much as possible attractive content in P2P network - otherwise the users would move to an another one. It may be feared that if the applications's default configuration is "no sharing", then only a small amount of users would turn the sharing on.

It is clear that a badly configured P2P application can be a serious threat to security and privacy. A possible scenario would be that a peer-to-peer software is installed on a home computer which is being used by several members of the family - father uses it for doing his work, and his son uses it to download new software and music from P2P networks. In this scenario, badly configured P2P application could easily share important and confidential documents which are stored on the computer. As more and more novice users are starting to use P2P software, this problem will surely exist.

3 Kazaa - A case study

Kazaa [1] has been the most successful successor of Napster, with millions of users. It uses a distributed network (Fast-track) which works over Internet. Kazaa is mostly used for file transfer. It has much content thanks to its wide user base and a working search functionality. Kazaa can be considered as a second-generation P2P network - it is distributed, but does not offer advanced options such as anonymity. Figure



Figure 1: Kazaa

1 shows the Kazaa user interface. Kazaa's technology and architecture is discussed in [14].

3.1 Spyware and adware content

There has been some conversation about Kazaa's privacy and its suspected spyware content [13]. Today, Kazaa's website promises "*No spyware*" on their frontpage. There is also some information about Kazaa's privacy and security [15, 16]. They admit that Kazaa contains adware programs from GAIN and Cydoor. Yet, they take no responsibility of those various third party programs that Kazaa contains: "Sharman Networks has no responsibility or control over the GAIN AdServer software, the GAIN Network or the GAIN Publishing's data practices and shall not be liable for any losses, damages or injuries arising therefrom." [15]

Kazaa's installation program establishes many different network connections. A packet capture shows that the third party software bundled with Kazaa is installed as a part of the process. Sites include for example *altmet.com*, *gator.com* and *joltid.com*. The installation procedure requires, that the end user accepts the license of Kazaa and installation of those third party programs. Otherwise the installation will not complete.

GAIN's privacy statement [17] describes the information that it collects. Some examples of such information includes:

- Non-personally identifiable information, such as the first 4 digits of a credit card number
- Operating system type, version, browser type, etc.
- Web pages viewed; their URL's, view time
- Online purchases
- Information about software installed on the computer
- "Occasionally, we may, automatically, or through other means, update, upgrade, or patch the Licensed Materials".

By installing Kazaa, the user agrees that the information mentioned above may be sent over Internet and to be shared with parties that may not even be known beforehand [17].

This is just one example of suspicious third party software that comes along with the Kazaa. Kazaa's "*No spyware*" - promise has little use, if the users have to install this kind of third-party software. It should be noted, that the last point allows installing software to the computer without users knowledge.

3.2 The protocol

Kazaa uses proprietary protocol for its network traffic. Some work has been done to reverse-engineer this protocol, for example the giFT project [18]. Through this access to the Kazaa network is possible also with other clients.

4 Third generation P2P networks - A possible solution?

As the concern about privacy is growing, new solutions are being developed. Anonymous P2P applications are one approach. They hide the users true identity, which is identified by IP address, and use some other mechanism instead.

Third generation networks offer improvements over the currently widely used second generation networks. Anonymity has become a major concern, and many new networks try to offer more protection for privacy. Anonymous networks are for example Freenet [19], GNUnet [22] and I2P [23].

Another subclass of third generation networks is friend-to-friend -networks [24]. They allow connections only to pre-selected nodes (friends), and other connections are routed via those friends. MUTE is one example of this kind of network architecture.

4.1 Freenet

Freenet [19, 20] is an another approach to distributed information sharing. It is an anonymous, completely distributed and self-organizing network implemented with Java. The common user interface is a normal web browser, though also other tools exists. As its same suggests, freedom of publishing has been main motivation on its development.

Freenet consists of nodes. Nodes offer disk storage and bandwidth for Freenet's content, a so-called virtual file system. The application selects which files are stored in a particular node. The most popular content of Freenet is found from many different nodes, while seldomly accessed content may be found from only one specific node. This way, the end-user has little possibilities to affect the material which is stored in his/her computer. This may be troublesome, since some of the content is probably illegal. The Freenet FAQ [21] states "*The true test of someone who claims to believe in Freedom of Speech is whether they tolerate speech which they disagree with, or even find disgusting. If this is not acceptable to you, you should not run a Freenet node.*"

Freenet's design objectives were [20]

- privacy for information producers, consumers, and holders
- resistance to information censorship

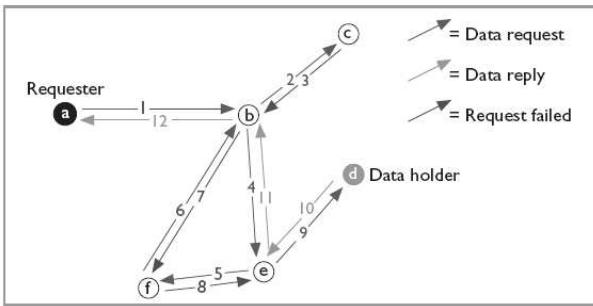


Figure 2: Routing in Freenet [20]

- high availability and reliability through decentralization
- efficient, scalable, and adaptive storage and routing

The privacy in Freenet is not perfect, but other design objectives seem to have been met better. The FAQ admits that “Freenet does not offer true anonymity in the way that the Mixmaster and cypherpunk remailers do”. The first Freenet node which receives the connection from a new peer knows its identity (IP address), and dedicated attacks could probably break the anonymity [21]. In [25], it is suspected that the Freenet’s anonymity has been compromised as the Japanese police arrested some P2P users who had been using an anonymous network Wimpy over Freenet. Yet it remains unknown if the full protections were used, and was the police really breaking the anonymity protections or using some other way to get to the tracks of those users.

The common way to share content in Freenet is by adding a link containing a content’s key to a commonly known Freenet forum. The key is a string which identifies the content within the network. It is calculated with SHA-1 hash function [20]. From the user point of view, Freenet has some problems - different than for example Kazaa. It works quite slowly. There is no search mechanism which is common in P2P file-sharing applications, and due to design, it is hard to implement. This is a way to protect the data from dedicated attacks, but a search function is something that the end user is eagerly looking for.

The slowness of Freenet can be understood by looking at Freenet’s routing system in figure 2. While common P2P applications use direct connections for file transfer, in Freenet this is not possible. Every node knows only the route to their neighbours, so they cannot establish direct connections. Instead, every message is routed through the chain of nodes. Freenet’s routing system is clearly explained in [20].

4.2 MUTE

MUTE [26] is a GPL licensed file-sharing application, which hides the IP address of its users. The software is new, and still under active development. Current version, 0.4, is available for Unix, Windows and Mac platforms.

MUTE works by replacing IP addresses with hashes as seen in figure 3. A peer knows only the hashes of the other peers, not the IP addresses. Since TCP/IP traffic always needs an IP address, a MUTE peer has to know some IP addresses to establish any connection. Those peers, whose

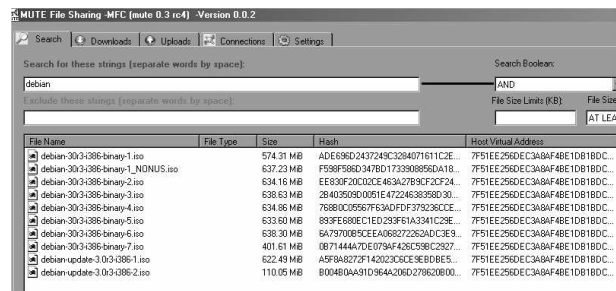


Figure 3: MUTE file-sharing

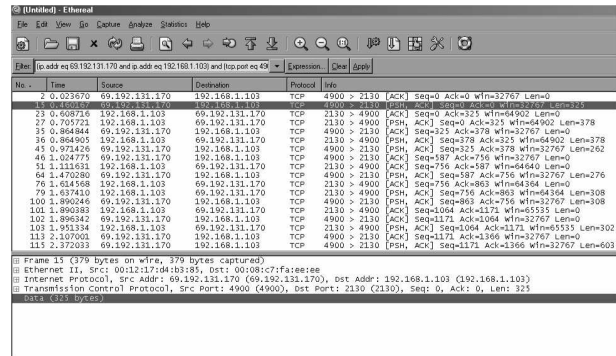


Figure 4: MUTE packets captured

address is known, are known as neighbours. All traffic to the MUTE network goes through these computers. MUTE calls this an “ant-inspired” routing algorithm. The addresses of neighbours can be configured manually, and this way it is possible to generate a true friend-to-friend-network.

Figure 4 shows some network traffic captured while searching the MUTE network. At least some packet analyzer attack fails to find any useful information. The connection is made to a neighbour peer, and all traffic goes through this peer as the documentation defines. This IP address is all that an attacker would get, and it will have little use. Even if the attacker would eavesdrop the neighbour node, he/she will not know from where the packets originally came from - they could be routed from the other side of the network or from the very next neighbour.

Compared to Kazaa, MUTE is light-weight and pretty easy to use. Kazaa has dozens of options, while MUTE has only the basic ones. MUTE’s installation program asks the user to select his/her shared folder, which is far better than silently defaulting to some application specific folder - this will make sure, that the user realizes he/she is actually sharing something.

Experiences of MUTE were not fully positive. The MUTE network seems very quiet - there are not many peers connected. Thus, the amount of interesting files to download is low. Much of the peers seem to be located far away, like USA or Italy. This, in addition to MUTE’s network architecture where everything is transported via many different nodes, causes quite low download speeds. The problem is similar of Freenet’s routing as discussed earlier in section 4.1.

5 The effects of privacy threats

There certainly are some risks involved when using peer-to-peer applications. But are they important and worthwhile taking? This chapter discusses the risks of P2P from the point of individual users, who use their own computers, and from the point of business world. The topic area is large and continually growing as new applications and threats emerge, so it is possible to only briefly introduce some of the most important threats.

5.1 Individual users

A great part of the peer-to-peer traffic is caused by home computers, and many home users have a little experience with computers and networks. Security awareness is still quite low. It should be realized that the P2P applications are also a security risk. Probably most trouble comes from viruses and malware, but depending on the computer usage, also unintentional file sharing may cause severe problems.

What are the assets that should be protected, what is so important and interesting for others that an individual user should do something about it? This of course depends on the usage of the computer, but some examples are listed below.

- Personal documents
- E-mails
- ID information - phone number, address, social security number, etc. information, which may be stored in personal documents
- Credit card number, bank account information
- Passwords, user names

Identity thefts have increased recently, and they are one of the most serious risks for both home and corporate users. One concrete example of this kind of threats is Korgo virus [31], which installs a key logger on the victim's computer. The key logger gathers passwords and form data which the user fills while surfing the WWW, and sends this information to the crackers. P2P file-transfer programs are one possible source of viruses like this.

Home users generally have no user support to clear and fix their computers, if something goes wrong. The adware and spyware programs may cause so much network and processing overhead that the computer's usability decreases dramatically, especially at home where processor and connection speeds are not always top-level. Adware/spyware is also often hard to detect and remove, at least for inexperienced end user.

Some easy precautions should be taken, if P2P software is to be used. A firewall and up-to-date anti-virus software are essential. There are plenty of P2P applications available, so there is a possibility to favour an application that does not contain malware. Most importantly, the end user should learn to use the software and configure it properly. The usability of the software is important because of these security and privacy threats.

5.2 Corporate users

Many companies nowadays have forbidden the P2P applications, since they have realized their problems [28, 29]. For companies, the possibility of unintended file sharing is very annoying, as they have much to hide - business secrets, plans, contact information, e-mails - the list is endless. Advanced network monitor systems/software can be used to detect the P2P programs, and commonly used ports can be blocked from the firewall. Still, if the end user has a possibility to install software to his/her computer, there is always a risk that somebody succeeds to install and use some P2P program. Port filtering may be used, but many applications have found ways to avoid this, so it should not be relied on. A common way to bypass the firewall is to use a commonly used port, which is left open. For example, Skype has an option to use port 80 for its traffic. This port is used in normal HTTP connections, so it is usually left open.

The VoIP is no doubt an interesting application for companies as well as individuals, as it offers low-cost replacement for the traditional phone network. But this comes with a prize - the security of the VoIP is not convincing. The called ID spoofing, which was discussed in 2.2.1, is a real privacy threat to the company. The possibility of eavesdropping and call logging should be remembered as well - the companies generally have much sensitive information which requires a trustworthy protection and the low costs of VoIP systems might not be enough to compensate this.

Unfortunately, much of the content in file-sharing P2P networks is somewhat illegal, and this means another risk for the companies where P2P is used. Music, videos and computer programs are the most common copyrighted material which is illegally distributed with these networks. Since the P2P applications (usually) share the files that are being downloaded, the user becomes a host for illegal material at the same time. His/her company could then be sued for piracy.

Growing number of employees work at home, either part-time or full-time. They generally need an Internet connection, and a connection to the company's private network. These users are a big risk for the companies - they have a full access to their computers, and can install any software and use it as they like. Companies security experts have limited ways to control them.

6 Conclusion

The peer-to-peer networks have become increasingly popular in a short time. Nowadays a fast processor, plenty of hard disk space and a fast Internet connection is available to millions of people, and this keeps feeding the growth of P2P networks.

File-sharing has been the most popular service of peer-to-peer networks. Recently, the VoIP and particularly Skype have showed that new applications can emerge and take the advantage of the power of P2P. It should be noted, that those new applications might also introduce new security threats, and this should be considered when designing and applying new technology.

First versions of computer software quite often have leaking security - it has been enough, that the software works.

Today, this is no longer the case. Security and privacy should be considered from the very beginning of development process since adding it later is often difficult or even impossible. The peer-to-peer applications are facing the same problem. First- and second generation networks can not offer the required level of privacy. A reaction to this is the emerge of new third-generation networks.

As discussed in this paper, third-generation networks are still at development state, and lacking some functionality and ease of use that second-generation networks can offer. Eventually, they will get more popular as they evolve and public awareness of risks in P2P rises. There are plenty of promising alternatives where to choose from, and within time probably one or two of those will gain superiority, just like its successors Napster and Kazaa once did.

Though new solutions are being developed, it is questionable will they help the situation. Many of the problems we are currently facing can not be completely solved with pure technology solutions. Anonymity and trust are one hard combination - the new anonymous P2P networks may offer anonymity at least to some degree, but they can not guarantee the safety of the content. Thus, spyware, malware and viruses are likely to remain our problem in future. The anonymity, while it is an important improvement, might even make it harder to trace malicious users and to protect from them.

References

- [1] Kazaa <http://www.kazaa.com>
- [2] Gnutella <http://rfc-gnutella.sourceforge.net/>
- [3] eDonkey <http://www.edonkey2000.com/>
- [4] Napster History <http://napster.music.us/history.htm>
- [5] Nathaniel S. Good, Aaron Krekelberg. *Usability and privacy: a study of Kazaa P2P file-sharing* June 2002 Available at <http://www.hpl.hp.com/research/idl/papers/kazaa/index.html>
- [6] Internetnews.com. March 23, 2004. *RIAA Keeps Pressure on P2P Users* <http://www.internetnews.com/xSP/article.php/3330071>
- [7] CNET News. July 8, 2003. *P2P's little secret* http://news.com.com/2100-1029_3-1023735.html
- [8] Skype - Internet telephony system <http://www.skype.com/>
- [9] CNN News. March 18, 2005. *Internet phones a hacking risk?* http://money.cnn.com/2005/03/18/technology/personaltech/scam_phones.reut/index.htm?cnn=yes
- [10] Computer Reseller News. June 28, 2004. *Spyware support costs run into millions* <http://www.crn.vnunet.com/news/1156261>
- [11] Wired News. News article. January 9, 2004. *Kazaa Delivers More Than Tunes* <http://www.wired.com/news/business/0,1367,61852,00.html>
- [12] Tech Law Advisor. News article. January 3, 2005. *RIAA or Overpeer Seeding P2P Files with Spyware?* <http://techlawadvisor.com/2005/01/riaa-or-overpeer-seeding-p2p-files.html>
- [13] Cnet News. News article. November 26, 2004. *CA slaps spyware label on Kazaa* http://news.com.com/CA+slaps+spyware+label+on+Kazaa/2100-1025_3-5467539.html
- [14] Jian Liang, Rakesh Kumar, Keith W. Ross. September 15, 2004. *The KaZaA Overlay: A Measurement Study* Available at <http://cis.poly.edu/~ross/papers/KazaaOverlay.pdf>
- [15] Kazaa Ad Support <http://www.kazaa.com/us/privacy/ad-support.htm>
- [16] Sharman's No Spyware Commitment http://www.kazaa.com/us/help/new_nospay.htm
- [17] GAIN Privacy Statement http://www.gainpublishing.com/help/privacy_statement.html
- [18] The giFT Project <http://gift.sourceforge.net/>
- [19] Ian Clarke, Oskar Sandberg, Brandon Wiley, Theodore W. Hong. *Freenet: A Distributed Anonymous Information Storage and Retrieval System* University of Edinburgh, 1999. <http://freenetproject.org>
- [20] Ian Clarke, Scott G. Miller, Theodore W.Hong, Oskar Sandberg, Brandon Wiley *Protecting Free Expression Online with Freenet* Available at <http://freenet.sourceforge.net/papers/freenet-ieee.pdf>
- [21] *Freenet Frequently Asked Questions* <http://freenet.sourceforge.net/index.php?page=faq>
- [22] GNUnet <http://gnunet.org/>
- [23] I2P <http://www.i2p.net/>
- [24] Friend-to-friend networks explained <http://www.answers.com/topic/friend-to-friend>
- [25] Cnet News. News article. December 3, 2003. *Covert P2P network fails to hide users* <http://news.zdnet.co.uk/internet/security/0,39020375,39118255,00.htm>
- [26] Jason Rochrer. *MUTE File Sharing* <http://mutednet.sourceforge.net/>
- [27] Alan Davidson. May 15, 2003. *Peer-to-Peer File Sharing Privacy and Security* <http://www.cdt.org/testimony/030515davidson.shtml>

-
- [28] Osterman Research, Inc. 2004. *Managing IM and P2P Threats in the Enterprise* Available at http://wp.bitpipe.com/resource/org_971197299_840/Osterman.pdf
- [29] Websense, Inc. June 1, 2004. *Emerging Threats: Peer-to-Peer File Sharing* http://www.websense.com/products/resources/wp/EmergingThreats_P2P.pdf
- [30] Martin Boldt, Johan Wieslander *Investigating Spyware in Peer-to-Peer Tools* Available at <http://psi.bth.se/mbo/masters.thesis.pdf>
- [31] BBC News. June 4, 2004. *Worm eyes up credit card details* <http://news.bbc.co.uk/1/hi/technology/3776247.stm>