SCIENTIFIC and TECHNOLOGICAL COOPERATION between RTD ORGANISATIONS in GREECE and RTD ORGANISATIONS in U.S.A, CANADA, AUSTRALIA, NEW ZEALAND, JAPAN, SOUTH KOREA, TAIWAN, MALAISIA and SINGAPORE

SecSPeer: Secure and Scalable peer-to-peer computing and communication systems

(Contract no: ΗΠΑ-021)

D5.1 "Commercial Viability Study"

Abstract: This document describes the exploitation and commercial viability of SecSPeer, a solution of scalable and secure unstructured Peer-to-Peer systems.

Contractual Date of Delivery	29 December 2005
Actual Date of Delivery	October 2006
Deliverable Security Class	Public
Editor	Antonis Misargopoulos
Contributors	Fotis K. Liotopoulos, Haris Papadakis

The SecSPeer Consortium consists of:

FORTH-ICS	Coordinator	Greece
University of Pittsburgh	Partner	USA
Virtual Trip Ltd.	Partner	Greece

Table of Contents

Table of Contents 2
List of Figures 3
List of Tables 3
Executive Summary 4
Chapter 1 Project Presentation and Positioning 5
1.1 Synopsis of the Project 5
1.1.1 P2P Systems 5
1.1.2 P2P Limitations 6
1.1.3 SecSPeer Objectives 7
1.2 Marketing Strategy 7
1.3 Business Model Options 8
1.4 Technical requirements for support business goals 9
1.5 Grid-based P2P Systems 9
Chapter 2 Overview of our Approach 10
2.1 Introduction 10
2.2 The Underlying Methodology 11
2.2.1 Duplicate Messages Reduction 11
2.2.2 Message Flooding Mechanism 13
2.2.3 System Security 16
Chapter 3 Exploitation of SecSPeer 18
3.1 P2P File-Sharing Systems 18
3.2 P2P Technologies in Corporations 21
3.3 Research Dissemination and Interest 22
3.4 National Security Interest 22
3.4.1 U.S. Government Cyberspace Security Strategy 23
Chapter 4 Presentation of the Foreseen Intangible Assets 25
4.1 Introduction 25
4.2 SecSPeer SWOT Analysis 26
4.3 SecSPeer PEST Analysis 27
Chapter 5 Discussion 30
5.1 Potential barriers in the exploitation of SecSPeer 30
5.1.1. Legal barriers 30
5.1.2. Technical barriers 30
5.2 Finding new sources of financial support 31
5.2.1. Promoting Acceptance 31
5.3 Profiling of the SecSPeer Results 32
Chapter 6 References 34

List of Figures

Figure 1: Example of duplicates' generation (left) – Red thick edges form the node A's shortest path tree (right) ... 12
Figure 2: Illustration of the Gnutella network and *Partition* method ... 15
Figure 3: Internet Protocol Trends 1993 to 2006 ... 18
Figure 4: P2P Traffic Monitoring June 2004 ... 19
Figure 5: BitTorrent vs. KaZaA Evaluation ... 20

List of Tables

Table 1: SWOT Analysis with respect to the SecSPeer consortium. ... 27
Table 2: PEST Analysis with respect to the SecSPeer consortium. ... 29
Table 3: Profiling of SecSPeer Results. ... 32

Executive Summary

This document provides the foundations of what shall result in the Exploitation Plan for the SecSPeer project by Month 24. As such, the Deliverable presents the overall approach and results of the activities developed inside Work Package 5 (WP5) for defining the exploitation potential of the SecSPeer project, taking into account the market, financial, managerial and technical aspects of it. The report has been issued by the R&TD dept. of Virtual Trip Ltd., responsible for WP5, with the collaboration of all project partners. As far as the contents are concerned, in the first section of the Deliverable the *Exploitation Plan* the overall positioning is presented in terms of marketing strategy and business model definition needs.

The outline of this document is as follows:

- Chapter 1 provides a brief review of the SecSPeer project, as presented in previous Deliverables.
- Chapter 2 provides a brief description of our proposed algorithms
- Chapter 3 describes the research and commercial interest of this project
- Chapter 4 gives a presentation of the foreseen intangible assets of SecSPeer, including SWOT and PEST analyses.
- Chapter 5 discusses potential barriers (legal and technical) in the exploitation of the SecSPeer project

Finally, the document concludes with a list of the used references.

Our main conclusion is that the SecSPeer project presents a significant potential for commercial exploitation, nationally and internationally, mainly due to the proliferation of peer-to-peer systems world-wide. For this reason, appropriate sources of funding and capital investors must be searched for, in order to expedite the in-time exploitation of the SecSPeer project's results.

Chapter 1 Project Presentation and Positioning

1.1 Synopsis of the Project

1.1.1 P2P Systems

During the last decades, Internet applications have been widely spread becoming a new way of efficient and direct communication between users. The development of World-Wide Web has completely transform the way users work, learn, study, shop and in general the way they interact with each other. However, the increasing number of requests, the need of multimedia traffic and the number of Internet users that continuously increases have stressed two significant important factors of effectiveness and confidentiality: scalability and security of the existing Internet computing and communication infrastructure. In addition, various malicious attacks, called Denial-of-Service attacks, seem to be a restrictive bottleneck to the effectiveness of the servers as sometimes their service cannot be commit to the customers at all.

The limitations in scalability of computing and communication infrastructures can hardly be traced in the "*client-server*" distributing model that most Internet services employ [1]. However, the scalability remains a problem as the increasingly number of users-clients implies a huge pool of requests that the servers have to address, reducing effectiveness and reliability. Also, a possible failure of servers causes the total failure of the service that is offered.

For the reasons above, last years another distributed computing model has been proposed [2], called *peer-to-peer* (*P2P*), and that is what we study in this project. Contrary to the traditional client-server model, the P2P model advocates that all computers are both servers and clients at the same time: they are *peers* or *servents* (**serv**ers + clie**nts**). Thus, in a P2P system, the service is not provided only by a limited set of servers: it is provided (potentially) by all peers, allowing the service to scale to large number of users. Indeed, since all computers in the system are both clients and servers at the same time, increasing the number of clients, implies that the number of servers is aloes increased. Therefore by default, such kind of systems is inherently scalable.

1.1.2 P2P Limitations

P2P systems seem to address scalability and reliability against client-server systems. However as demonstrated in Deliverable 1.1: *Requirement Analysis,* there is a significant important problem observed, that limits scalability and needs to be studied [4]. Most current P2P systems do not optimizes the traffic they generate, wasting their resource without reason. It is important that they seem to repeatedly generate the same requests and same traffic over and over, even if there is a significant amount of locality in P2P access patterns [3].

On the other hand, security becomes a bottleneck for P2P systems comparing with traditional client-server. In P2P systems, if attackers manage to compromise either the “server” or the “client” part of a peer, they will be able to compromise all the computers that participate in the P2P network. In their attempt to find ways to overcome traffic limitations imposed by strict firewall administrators, P2P systems have developed a rich set of techniques that bypass the traditional security rules imposed by firewalls, installing potential security threats in all computers of an organization. Additionally, P2P systems are usually installed in home and office computers that are not always administered by experienced system administrators, becoming more susceptible to security intrusions.

1.1.3 SecSPeer Objectives

In this project, we studied P2P systems with emphasis on improving their *scalability* and *security*, extending *expressiveness* and *quality of service* issues as well. We explored the locality aspects of these systems and proposed methods that capitalize on this locality. In addition we introduced mechanisms that monitor the security of such systems and identify security breaches early enough, as presented in previous deliverables [5], [6].

1.2 Marketing Strategy

The marketing strategy for the SecSPeer project will be based on a phased approach; the initial phase will concentrate on the direct exploitation of the project and the demonstrator cases in United States of America and Greece. It will focus on the national partners in these two countries and their existing professional contacts and co-researchers.

Since the national partners already operate successfully in this market, no further market studies seem necessary. The primary mechanism for marketing will be:

- exploitation of existing business contacts of the partners;
- presentations of the SecSPeer demonstrator at exhibitions, conferences, and technology fares;
- as accompanying measures, publications of articles in scientific and professional journals, features, and editorials describing the system in appropriate technical journals;
- contacting to potential users and corporations in order to integrate SecSPeer proposals into existing P2P systems;
- use of the Internet as an advertising medium.

The important is that the exploitation of SecSPeer holds both research and commercial interest. The latter is regarded as the most important intangible asset that can be created within the project lifetime.

1.3 Business Model Options

As Deliverable 1.1: *Requirement Analysis* demonstrates, getting peer-to-peer systems in the value chain of ICT and ICT-related industries implies that the business/commercial exploitation of the SecSPeer project shall be taken into account. Therefore, the business models that have to be supported are the following four:

- **Pure-license-based**

In this case, one can just purchase an instance of a client and through this will connect to a number of peer-to-peer networks that are “associated” to this package.

- **Pure-subscription-based**

In this case, one gets for free the client(s) software and pays subscription fees proportionally to which peer-to-peer network(s) is connected and other qualitative and/or quantitative criteria.

- **Hybrid (license & subscription)**

The Hybrid model is just the combination of the above two models.

- **Consulting**

The Consulting business model is probably the more interesting among the four proposed models. It is considered that a *“Free Peer-to-Peer Service Infrastructure”* exists, where a number of services are provided through various peer-to-peer networks. However, consulting services are required in order to make the offered services useful for each participating organization. For example, Skype-like P2P telephony could be incorporated in a company having offices at several sites. This would require major integration and consulting services.

1.4 Technical requirements for support business goals

The SecSPeer project needs to address the issues of *service monitoring*. Since appropriate tools for monitoring, as well as for authentication and authorization are available, the *billing* issue must be addressed, especially for subscription-based business model.

Distributed and trustworthy management of billing and authentication/authorization services present a technical challenge to the SecSPeer project, with major business implications.

Finally, interoperation with trust-management systems, including distributed reputation and recommendation systems, would be highly desirable from a business perspective, since content- and service-trust are required for business exploitations.

1.5 Grid-based P2P Systems

Peer-to-peer systems suffer from low Quality-of-Service, which is an obstacle for a number of commercial applications. This happens mainly because of the Internet infrastructure, which is, in a large extent, not reliable.

The Grid paradigm, which is widely accepted by ICT industries, will provide a solid infrastructure for the deployment of services. However, a well-know disadvantage of Grid -based systems is low flexibility. A P2P approach for the deployment of e-Services on Grids would probably overcome this serious problem. Moreover, this will lead to the creation of business synergies, since major ICT vendors, including IBM, Oracle, SUN, etc, have invested in Grid Computing and would be very interested in getting a channel for deploying services targeted to the market of peer-to-peer systems.

Chapter 2 Overview of our Approach

2.1 Introduction

The goal of the SecSPeer project is to handle the two main drawbacks observed in well-known P2P distributed systems; scalability and security. At this time, traditional World Wide Web distributed content-delivery systems treat each participant in the system as an individual stakeholder that plays a distinct role. Specifically, web servers play the role of the content providers, while clients (web browsers) request content from the servers. On the other hand, in P2P system each participant (servent) offers content to the rest of the participants and at the same time can also request content from them. To do this, each servent has to be aware of the existence and the probably the location of each other servent participating in the topology. For that reason, the prevailing method used is *broadcasting* (i.e.: trying to ask everyone in the system). This technique that is adopted mostly by the current P2P systems suffers since the number of peers continuously increase and the size of data transferred get larger dramatically, causing many scalability optimization issues. In addition, this large and unstructured – at most cases – topology is extremely vulnerable to malicious cyber-attacks over the web, such as DDoS and DoS attacks.

In next section, we make a brief overview of the problems that SecSPeer deals with and the algorithm solution we propose; i.e., reducing duplicate messages and reinforcing security against DDoS and DoS attacks [13].

2.2 The Underlying Methodology

In section 2.2.1 we describe the proposed algorithm for the elimination of *redundant messages*, in order to reduce the number of worthless messages during a broadcast. *Worthless* messages are messages that do not increase our chance of locating the required piece of data. The system tries to avoid forwarding a message to a participant that may have already received it by learning from traffic history, through the use of explicit duplicate notification from the receiving participant. In section 2.2.2, we present another scalability improvement algorithm, which tackles the issue of *blind broadcast (message flooding)*, by adding *semantic information* to the network, so as to be able to broadcast to only a subset of participants of the network. This way we try to avoid generating another type of worthless messages, namely messages sent to participants that do not have the data being looked up. Finally, section 2.2.3 briefly describes SecSPeer consortium's proposals of dealing with security issues, such as *spam generation* and *DDoS attacks*. Broadcasting creates significant amount of traffic which is shared by all the nodes of the system, on behalf of the broadcasting node. To avoid exploitation, it is essential that amount of traffic load a node injects in the system be relative to other nodes' amount of traffic load it serves. Spam generation is created from malicious nodes that reply to queries for content they do not have. Since the amount of results they send is arbitrary, this reduces the amount of "good" results the requesting node receives.

Note that the description of the algorithms proposed is out of the scope of this document. Though, our goal is to present the description of the issues we studied and some metrics we retrieved based on our algorithm solutions. More technical description can be found in previous Deliverables.

2.2.1 Duplicate Messages Reduction

- **The problem**

As Deliverable 2.1: *System Design* describes, the duplicate problem stays when two or more peers send the same message to another peer and consequently a peer receives the same message from two or more other peers. *Figure 1* (left) depicts such a possible scenario of duplicate message transferring, since B and C nodes both sent the same message to D. Additionally, D forward the first message it receives, just when it receives it. Assuming that B's message reached D before C's message, this will mean that D will forward the message to all of its neighbors, except the one it received it from (that is B), which includes C. This causes the generation of two duplicates, one from C to D and one from D to C.

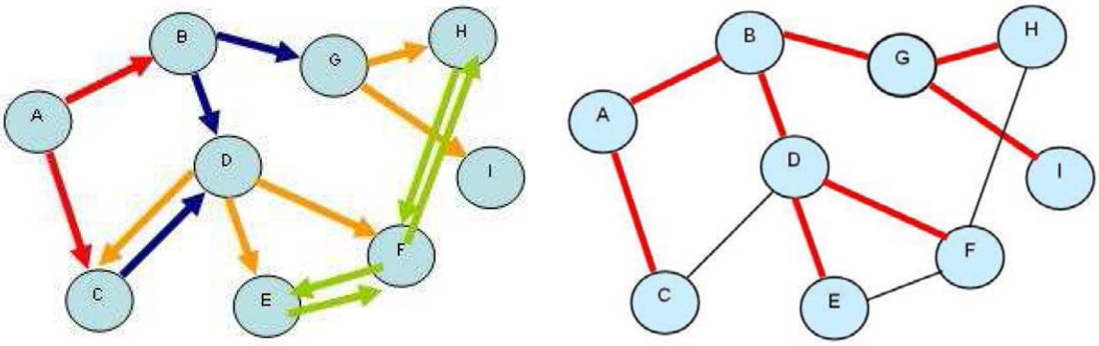

Figure 1: Example of duplicates' generation (left) – Red thick edges form the node A's shortest path tree (right)

During a single flooding process (originating from any node X), if a message traveling over and edge, reaches a node Y, which has yet to receive a message of this flood, this edge is part of X's shortest path tree in that graph of the network. On the other hand, if the node that receives the message has already received another message of the same flood, it means that the edge traveled by the duplicate message to reach the node is not part of X's shortest path tree. *Figure 1* (right) depicts such a potential scenario. Even if there are two shortest paths from a node X to a node Y, Y will process the messages that arrive at the same time, sequentially, which means that the path used by the first message to be processed, will be deemed shortest. Any message sent over an edge which is not part of the shortest path tree of the node that initiated the query, will be a duplicate. Each node has a different shortest path tree and this spanning tree does not change, if the network structure does not change.

- **Our approach**

An initial approach to eliminate the duplicate messages during flooding is that each node need not be aware of the shortest path tree of each distinct node X it may receive a message from, but rather, which of its edges are part of the shortest path tree of X. However this design is also not very scalable because of the requirement that each node hold information equal to the size of the P2P network *N* multiplied by the degree of the nodes *n*; i.e., *O(N*n)* in the worst case. According to this algorithm, each message was distinguished based on the originating node, thus leading to *N* categories, one for each node in the network. To make the algorithm scalable, we found some other *criterion* which defines a small number of categories, regardless of the network size, to distinguish between messages, rather than the node that initiated the flood.

It has been observed that many messages from several originating nodes exhibit the same behaviors as far as duplicate generation is concerned for any single node, depending on the distance traveled and the “direction” from which they arrived. Thus, messages are instead grouped by and on any single node based on the distance already traveled when they arrive, and the immediate neighbor through which they arrived. These group definitions instead define a much smaller number of categories of messages. As mentioned, the definition of these categories is based on the need to group together messages that have similar possibility of creating duplicates (either low or high). After a warm-up period, each node discards messages which belong to categories with high duplicate generation probability.

2.2.2 Message Flooding Mechanism

In this section, we describe the proposals for adding semantic information to the network, in order to avoid sending messages altogether to nodes which most surely will not contain the piece of data we are looking for.

- **The problem**

Besides duplicate messages problem, message flooding issue comes upon the scalability of P2P systems. The flooding mechanism becomes scalable by using the TTL field, at the expense of greatly reducing coverage and thus making locating less popular items very difficult. The only way to increase coverage by using the same amount of messages, is to ensure that every message reaches a new servent (i.e., no duplicates) and thus, that we do not waste messages. However, the goal of reaching as many nodes as possible is a consequence of the fact that every node has the same chance of containing the piece of data we are looking for. If the information we are looking for is popular, (i.e. is replicated to many nodes in the network), flooding will locate it quickly, even with a small TTL. If however, the information resides at just one node, flooding would have to reach almost every node in the network to locate it. This means that in unpopular searches, a lot of bandwidth is wasted contacting nodes that do not have the information we need. If there was some way of knowing which servents are less likely to contain the information, we could use some way to avoid wasting messages by sending them to those servents.

- **Our approach**

So far, a number of techniques that tries to handle flooding problem has been come up and already used in existing systems; i.e., *Directed Breadth First Search (DBFS)*, Ultrapeers in the *Gnutella 2* network and Supernodes in the *FastTrack* network, and lastly the approach of thematic partitioning of the network proposed by Crespo and Garcia-Molina.

As demonstrated in Deliverable 2.1: *System Design*, SecSPeer adopts the basic idea of the SON proposal and defines i) how the overlay subnetworks are formed and ii) how search is performed through those networks. The idea in the core of our design is the partitioning of the Gnutella network in independent subnetworks. This partitioning will be based on some categorization of the content in the network.

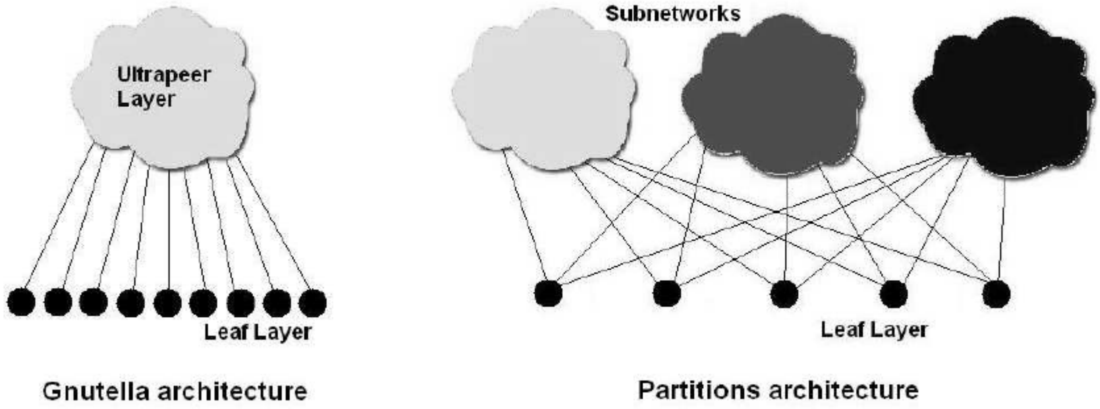

Figure 2: Illustration of the Gnutella network and *Partition* method

We proposed the formation of categories based on easily applicable rules. Such a simple rule is to apply a uniform hash function on each keyword describing the files. This hash function maps each keyword to an integer, from a small set of integers and each integer defines a different category. As it is obvious, we categorize the keywords instead of the contents itself, since lookup is keyword-based. We also define one subnetwork (*partition*) for each defined category. Because of the uniformity of the hash function, the subnetworks will be roughly equal in population. Given a small set of integers, it is very likely that each peer will contain at least one keyword from each possible category. Thus, Leaf nodes connect to all subnetworks. However, they only publish to each one of them, only the part of their content appropriate to that subnetwork. *Figure 2* shows Gnutella architecture with *Partition* architecture.

Connecting to a subnetwork can be done the same way servents connect to the P2P networks today; i.e., a *webcache* can be maintained for each of the subnetworks defined by the categories. Each query is performed based on the rarest of its keywords and is flooded only to the subnetwork corresponding to the same category as that keyword. Thus, the search space of each individual flooding is restricted to a single partition, considerably limiting it, and thus reducing the overload volume of traffic produced by flooding.

The evaluation of the scalability of SecSPeer comprises some measurements on both maintenance cost of the system an operational costs (i.e. query load), as described in Deliverable 2.1: *System* Design and Deliverable 4.1: *Deployment – Evaluation*.

2.2.3 System Security

- **The problem**

A known malicious behavior, observed in the Gnutella system, is the generation of Query Replies for each Query received by the malicious peer. That is, a malicious peer can monitor every Query packet which is routed to it, parse its Search Criteria and produce a Query Reply packet with imaginary embedded responses; i.e. *spam generation*. The responses are created by adding a known file extension to the original Search Criteria and by performing a type of frequently used capitalization. Although, the responses have imaginary filenames, the files have a valid content, which is an advertisement message.

Another potential security threat in unstructured peer-to-peer systems is *Denial of Service* (*DoS*) attacks. A malicious peer can generate random Query traffic; that is Queries with random Search Criteria, HOPS and TTL fields. Since, each query floods the system via the traditional flooding mechanism, or portions of the system via dynamic querying, a peer can emit to the system unnecessary message traffic, which will eventually grow following exponential rates. A second type of DoS attack, which nature is completely distributed, can be achieved by emitting Query Responses instead of Queries. This method may target any machine, which listens to a known port and it is connected to the Internet. There is no need for the target machine to be a Gnutella participant. A malicious peer can monitor the Queries it receives and generate responses for every Query message. Each Query Reply packet will carry the IP address and the Port of the target machine. Since, there is no mechanism to indicate if the IP address in the Query Reply message matches the IP address of the machine, which generates the Query Reply message; all generated responses will be routed to the original queriers. Thus, there is a chance that a vast amount of download attempts to a single computer may be performed in a short time of period.

- **Our approach**

Abstractly, in order to eliminate spam generation, we propose a strategy in which a legal peer queries with a random Query of TTL=1, at a random time period upon handshaking with a new node, the new node it handshaked with. If it receives a reply for the random Query then it should drop the connection. Malicious peers though may try to detect the strategy by (a) not responding to Queries with TTL=1 and HOPS=1; that is Queries originated by one of their neighbors, (b) not responding to Queries for the first few minutes (c) hide over a legal peer, part of the malicious infrastructure (that is, the legal peer does not follows our strategy), which serves as a gateway to the Gnutella system. Our strategy can be enhanced in order to overcome (a) and (b) by making the HOPS field of the random Query message also random and by re-querying peers in random time intervals, respectively. As far as (c) is concerned, we may issue also Queries with TTL=2. If a response of a random Query with TTL=2 is received, then we can safely judge the neighbor as a legal client that hides a malicious one and the connection should be dropped.

This spam preventing strategy with random querying could be also used to prevent DoS attacks of the second type, where peers respond to every Query with results that contain the IP address of the target machine. The peer that originates the attack will also answer the random Query and hence will be judged as malicious. Furthermore, we introduce a complete load-balancing solution based on *"coupon exchanging"* between peers, which attempts to prevent Query flooding based attacks. Our load-balancing algorithm prevents peers from generating enormous Query traffic and thus DoS attacks, as presented in more details in previous Deliverables. The base architecture of SecSPeer, as far as the security issues are concerned, is build in the *SEALING* (*Short Term Safe Listing)* algorithm, as demonstrated in the previous Deliverables.

Chapter 3 Exploitation of SecSPeer

3.1 P2P File-Sharing Systems

Over the last five years, P2P has become one of the most popular user applications on the Internet and is acknowledged as one of the key drivers for consumer broadband uptake. This popularity has positioned P2P as the dominant protocol on the Internet, representing between 60% and 80% of total traffic on the networks operated by Internet Service Providers (ISPs), as *CacheLogic™* announced [14]. *Figure 3* presents the Internet Protocol trends since 1993.

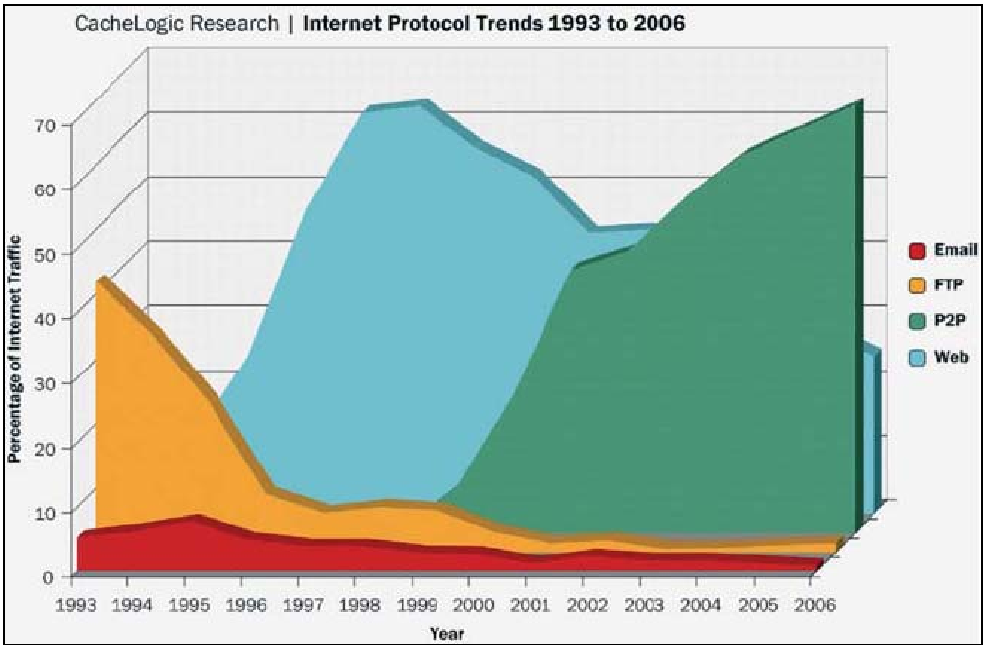

Figure 3: Internet Protocol Trends 1993 to 2006

In recent years, aspects as legacy and privacy have been come up. P2P systems, such as Napster, Gnutella, KaZaA, Grokster and others have become synonymous with illegal file-sharing platforms, enabling the users to download songs, films and other content for free. Although all the legal problems have not yet been resolved, P2P has the potential to become an efficient, legally conformant and non-expensive means for delivering content to the general public. P2P systems can be used for the distribution of either "*linear*" (i.e. real-time video or audio streaming) services and/or "*non-linear*" (on-demand) services (i.e. file downloading) over the Internet. P2P can also be used as a channel for networked Personal Video Recorders (PVRs). *Figure 4* shows what was the market share of the different P2P approaches was in June 2004, according to CacheLogic™ research study.

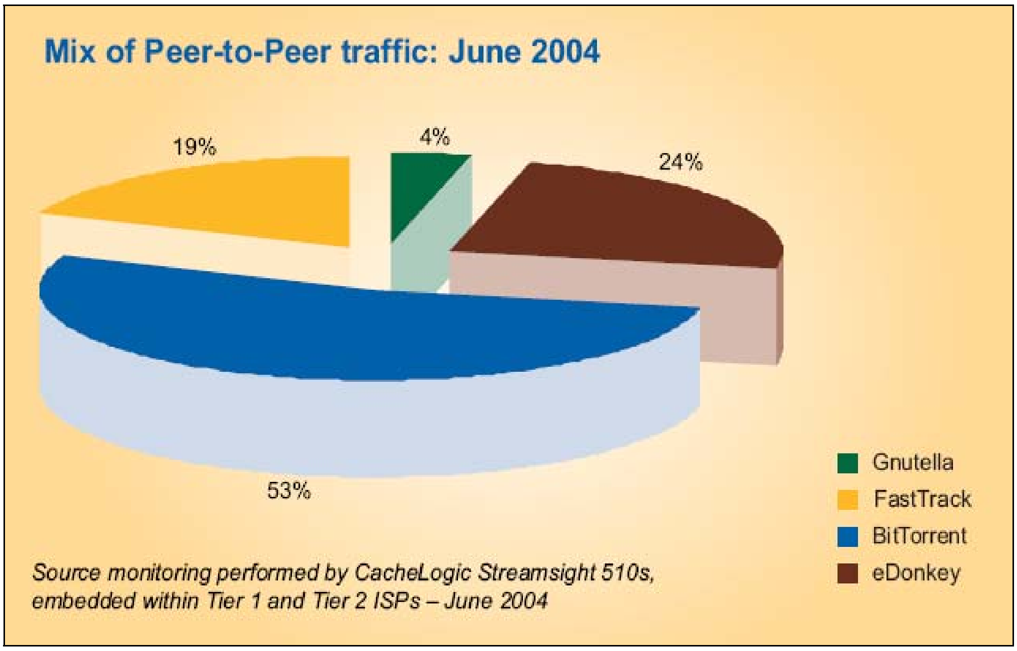

Figure 4: P2P Traffic Monitoring June 2004

BitTorrent [15] is increasingly being used for distributing legitimate content and takes more than half of all P2P traffic. P2P carries a mixture of audio (11%), video (61%) and other data traffic (about 28%). Almost half of all the traffic uses Microsoft Windows Media formats, while most of the audio files use the mp3 format (65%). This study has been compiled from data gathered from CacheLogic's *Streamsight Analysis Network* over a 48 hour period between 16 and 17 July 2005. By the end of 2004, BitTorrent was accounting for as much as 30% for all Internet Traffic. But December 2004 saw a crackdown on the major BitTorrent sites (Subrnova and others). In 2005, over the 25% of the entire number of broadband subscribers in United States and 20% in China exploit BitTorrent services.

In many geographies though, the majority of traffic has shifted towards an alternative network – *eDonkey*. Notice that eDonkey has been localized to a wide range of languages and is fully decentralized, in the manner that there are no “tracker” sites to shut down. BitTorrent levels have been dramatically affected by the closure of the key tracker sites; however there was not an immediate occurrence. For that reason, a new fully decentralized version of BitTorrent known as *eXeem* released to rocket in popularity. In 2005, eXeem accounts for less than 1% of BiTorrent traffic, probably due to the authors decision to include spy ware technology at launch; one of the key factors in the decline of KaZaA along with the legal pressures. While Asia is predominantly BitTorrent with notable exception of South Korea, Unite States has seen growth in eDonkey and Gnutella.

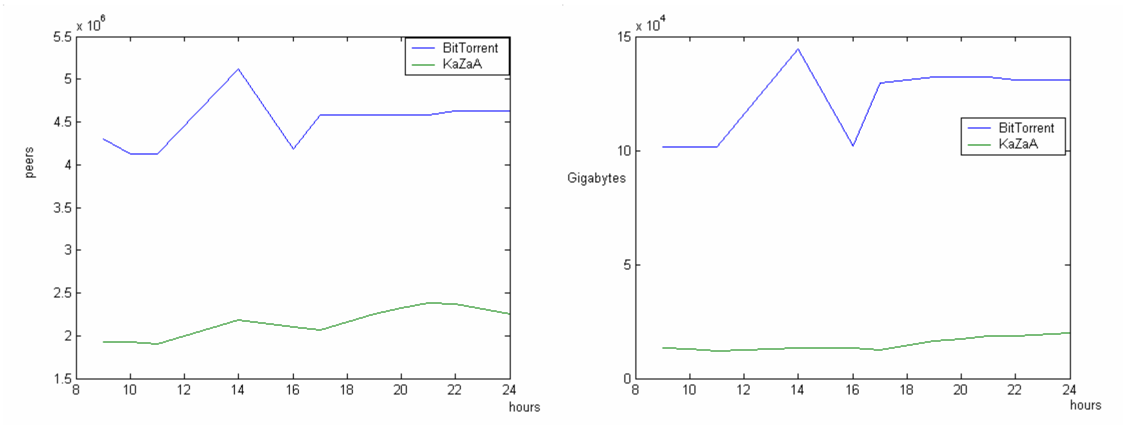

Figure 5: BitTorrent vs. KaZaA Evaluation

According to U.S. department of Commerce [8] “*corporate users of P2P technologies will rise from 60,000 currently (2002) up to 6.2 million by 2007, and that the enterprise P2P market will increase from $42.8 million today to $4.53 billion in 2007*”. Recent researches have proved the enormous number of peers and the size of the data transferred. For instance, *Figure 5* presents such a study in respect to BitTorrent and KaZaA.

3.2 P2P Technologies in Corporations

As the technology continues to mature, an increasing number of large companies are beginning to adopt P2P-based solutions to conduct business more cost-effectively. For instance, corporations such as *Intel*, *GlaxoSmithKline*, *Raytheon*, *Ernst & Young*, and *First Union Bank* are turning towards P2P in order to manage and share information across distances or as a collaborative tool. First Union is using a distributed computing system from the New York start-up *DataSynapse* to harness unused processing power from existing computer systems and to improve their existing distribution architecture. In addition Raytheon and GlaxoSmithKline are implementing file-sharing software from *Groove Networks*, a P2P company headed by Lotus Notes creator Ray Ozzie. Governments in countries like the United Kingdom and the Netherlands are using P2P systems to facilitate and enhance scientific research efforts.

Additionally, the market is also favorable for security-related systems. For example, *IDC*, the world's leading provider of technology intelligence industry analysis and market data, predicts that the security market will increase from $66 billion in 2001 to $155 billion in 2006 [11]. Given the world’s concern about security, the Internet security-related market will continue to increase for the years to come and thus, provide a fertile ground for our ideas and our results. In July, IDC found that security remained the number one concern of IT professionals, as 40% of about 1,000 IT managers surveyed rated security as their highest priority.

3.3 Research Dissemination and Interest

Besides the interest on commercial organizations, a large number of researcher pay great attention to matter like scalability and security in P2P systems. Many universities and research institutes form research teams obtaining measurements and experiments on existing P2P system like Gnutella, KaZaA, eDonkey, and Napster. As the number of users that use modern P2P applications is estimated to become 30 million per day [12], the need of such optimization research is paramount important.

Especially, as far the SecSPeer project partners are concerned, the study of scalability and security of P2P systems create new perspectives to knowledge and scientific experience. *FORTH* brings its expertise in web caching, P2P systems, scalable applications and Internet security. By having expertise in all the above areas, FORTH becomes particularly more suitable to conduct research in the scalability and security of P2P systems. The researches from the *University of Pittsburgh*, on the other hand bring a long history of expertise with databases and benefit on the challenges of an increasing number of web users. Finally, *Virtual Trip Ltd.* has significant expertise in protecting and penetrating Internet-connected systems and increases its scientific knowledge to P2P systems.

All SecSPeer results and proposals to message flooding, information duplicates and security problems on P2P systems can become a fertile ground for further research as well as an optimization prototype for existing systems.

3.4 National Security Interest

As we saw, during last years a great evolution of (semi-) automatic information exchange and service supply P2P systems over the Internet is remarked all the around the world. Security and scalability are a significant important perspective not only for companies and organizations, but for government as well. Therefore, both Greece and United States pay great attentions to national, political, economical, commercial, educational, and research areas against possible malicious attacks over cyberspace. For instance, during the Olympic Games 2004 in Athens, Greek Government spent enormous amount of money in order to face any possible terrorism and cyber-attack threats. Next section refers to "*National Security Strategy to Secure Cyberspace*", a framework announced by U.S. Government about information technology infrastructures protection policy over cyberspace.

On the other hand, research on scalability of distributed services is of significant importance for both countries. For example, in both countries a significant amount of services has been moved to the Internet. Such services include e-commerce, e-government, and e-science. According to the "*GREECE in the Information Society: Strategy and Actions 2002*" document, one of the four milestone goals of the Information Society for Greece as is to "*carry out the greatest part of transactions with the state in an electronic manner*". Such goals imply that millions of people will interact with the state electronically; stressing the limits of the aging client-server model of building Internet services [10].

3.4.1 U.S. Government Cyberspace Security Strategy

The United States that participate to this project, have already taken several steps towards securing their cyberspace, and are particularly active in reducing cyber-attacks. To underline the importance of such an activity, the "*President's Critical Infrastructure Protection Board*" in the "*National Security Strategy to Secure Cyberspace*" [7] strongly encourages research and development activities against any kind of malicious intrusions into servers, clients, databases and warehouses concerning secret national information. In particular, the National Strategy to Secure Cyberspace outlines an initial framework for both organizing and prioritizing efforts. It provides direction to the federal government departments and agencies that have roles in cyberspace security. It also identifies steps that state and local governments, private companies and organizations, and individual Americans can take to improve our collective cyber-security.

The strategic objectives of this National Strategy to Secure Cyberspace are to:

- Prevent cyber attacks against America’s critical infrastructures;
- Reduce national vulnerability to cyber attacks; and
- Minimize damage and recovery time from cyber attacks that do occur.

Information infrastructures over Internet originally designed to share unclassified research among scientists who were assumed to be uninterested in abusing the network. It is that same Internet that today connects millions of other computer networks making most of the nation’s essential services and infrastructures work. These computer networks also control physical objects such as electrical transformers, trains, pipeline pumps, chemical vats, radars, and stock markets, all of which exist beyond cyberspace.

A spectrum of malicious actors can and do conduct attacks against these critical information infrastructures. Of primary concern is the threat of organized cyber attacks capable of causing debilitating disruption to U.S. Nation’s critical infrastructures, economy, or national security. The required technical sophistication to carry out such an attack is high and partially explains the lack of a debilitating attack to date. Cyber attacks on U.S. information networks can have serious consequences such as disrupting critical operations, causing loss of revenue and intellectual property, or loss of life. Countering such attacks requires the development of robust capabilities, reducing vulnerabilities and detecting those with the capabilities and intent to harm these critical infrastructures. Underlining the important of computer security, on November 27 2002, President Bush signed $900 million “*Cybersecurity Act*”, legislation dedicating more than $900 million over five years to security and education to protect the US infrastructure against hackers and terrorists.

Chapter 4 Presentation of the Foreseen Intangible Assets

4.1 Introduction

In this Chapter, we present a SWOT and PEST analysis for the SecSPeer project.

The *SWOT* analysis is an extremely useful tool for understanding and decision-making for a variety of situations in business and organizations. SWOT is an acronym for *Strengths, Weaknesses, Opportunities,* and *Threats*. The SWOT analysis headings provide a good framework for reviewing the strategy, position and direction of a company or business proposition, or any other business concept. A SWOT analysis is a subjective assessment of data, which is organized by the SWOT format into a logical order that helps understanding, presentation, discussion and decision-making. The four dimensions are a useful extension of a traditional two heading list of pro's and con's.

On the other hand, the *PEST* analysis is a useful tool for understanding market growth or decline, and as such the position, potential and direction for a business. A PEST analysis is a business measurement tool. PEST is an acronym for *Political, Economic, Social* and *Technological* factors, which are used to assess the market for a business or organizational unit. It can be used for business and strategic planning, marketing planning, business and product development and research reports and thus A PEST analysis measures a market; while a SWOT analysis measures a business unit, a proposition or idea. Note that the PEST model is sometimes extended (some would say unnecessarily) to seven factors, by adding *Ecological* (or *Environmental*), *Legislative* (or

Legal), and *Industry* analysis (the model is then known as *PESTELI*). Arguably if completed properly, the basic PEST analysis should naturally cover these 'additional' factors: Ecological factors are found under the four main PEST headings; Legislative factors would normally be covered under the Political heading; Industry analysis is effectively covered under the Economic heading.

4.2 SecSPeer SWOT Analysis

In this Section, we recognize and examine the strengths, weaknesses, opportunities and threats in respect with specific criteria. Actually, this classification can mostly be of use either at the early stages of preprocessing input to achieve the *formation* of the set of each organization's core competences or to provide a *stress test* for each partner. In this respect, it can be treated as a "rough guide" to help people like corporate decision-makers and research directors to improve their processes and increase the value they shall be getting from the project. However, even at this stage the SWOT analysis of SecSPeer could be useful for commercial managers and market-analysts for PEST construction.

Strengths	Factor	Weaknesses	Factor
P2P penetration study	1		
Security issues knowledge	1		
Existing P2P system architecture knowledge	2		
Significant research and commercial results	2		
Expected publications and papers	2		
Potential enhancement into well-known P2P file-sharing technologies and governmental infrastructures	2	This poses a challenge to copyright law. Tens of millions of people regularly violate copyright using P2P networks, a tide of infringement copyright holders have been unable to stop. Numerous commentators have proposed legal responses to the P2P phenomenon, but either they take a myopic view of the problem, proposing a quick fix that does not address the larger issues raised by P2P, or they propose a radical copyright overhaul, likely to take years, without any provision for current harms.	1

Opportunities	Factor	Threats	Factor
SecSPeer could establish a standard protocol where mostly proprietary infrastructures can use today.	1		
SecSPeer proposals could be adopted by well-known existing P2P systems	1		
SecSPeer proposals could be the basic architectural idea for new P2P systems	1		
Extensive reputation of SecSPeer in conferences, meetings and other activities	2	Other similar hybrid architectures appearing before the end of the project	1
		Low rate of adaptation of the developed technology	1
Increasing Government budgets for scalability and security issues	1		
Establish/improve expertise in web services, semantics, P2P networking	1		
Project uses innovative approach newly proposed in P2P research community.	2		
Closer cooperation in research and development of Greek and U.S. partners	2		

Table 1: SWOT Analysis with respect to the SecSPeer consortium.

Table 1 presents the SWOT analysis in respect to SecSPeer final product. The first column is the *criteria* list and the right the *factors* according to them. In specific, for factor column (**0**) indicates a neutral position, (**1**) a potential agreement, and (**2**) a strong agreement of the consortium. Finally, (**-**) indicates a strongly negative position of the specific information presented in the corresponding category.

4.3 SecSPeer PEST Analysis

In this Section, we present the PEST analysis of SecSPeer project; i.e., the evaluation of the project's potential from a *Political, Economic, Social* and *Technological* point of view. *Table 2* summarizes the results of this analysis.

Political criteria	Description
Greek Government interest	Security of governmental information systems is a critical concern for Greece especially during important period like the Olympic Games 2004. The government is funding corporations and research institutes for security and stability matters study.
U.S. Government interest	Security is one of the most critical problems that the U.S. Government is dealing with. Numerous cyber-attacks carry out a real threat for the safety of national and scientific data of U.S. nation. U.S. president Bush signed more that $900 million to protect the US infrastructure against hackers and terrorists.
Worldwide interest	Other nation all around the world pay great attention and interest for national security issues; like China, Russia, France, etc.
Reduce national vulnerability to cyber-attacks	SecSPeer provides a stable shield against malicious cyber-attacks. This high-level protection increases the sense of safety and prestige of a nation and people.
Minimize damage and recovery time from cyber attacks that do occur	Provides the capability for quick damage estimation and easy recovery in case of successful malicious cyber-attack.
Prevent cyber attacks against other critical governmental infrastructures	Eliminates the spread of cyber-attack effects from main P2P systems to smaller low-level infrastructures.
Stability and scalability of governmental infrastructures	Besides security, SecSPeer handles scalability problem against the enormous number of users and the large size or required transferred data over P2P system.
Capability of secure and stable data transmission between nations	SecSPeer could be the framework for a secure and stable information transmission over systems of nations all around the world.

Economic criteria	Description
Low cost	Low cost for research and development. Low cost for support and maintenance.
Job opportunities	Further research and improvement implies new job opportunities for researchers and scientific personnel.
Possible SecSPeer proposals adoption by existing P2P systems	SecSPeer proposals can be adopted by well-known existing P2P systems with considerable financial benefits.
Public market interest in SecSPeer model innovation	Besides the existing trade P2P systems, many companies and organizations preserving P2P infrastructures may adopt SecSPeer proposals as well.

Social criteria	Description
SecSPeer consortium gains theoretical and technical knowledge over P2P architecture	All partners learn how well-known existing P2P systems are organized; like Gnutella and KaZaA. They study the vulnerability potentials on malicious cyber-attacks and go deep into security issues.
Public positive reaction	Users take advantage of the security and the stability of the P2P infrastructures. SecSPeer reinforce the sense of safety and insurance for data transmission between remote users and trade transaction with respect to e-commerce P2P systems.

Technological criteria	Description
Technological and research achievements on P2P systems	• Scalability on large number of users/peers • Stability on the large number of user requests and size of data transferred • Flooding messages process handling • Duplicate-message elimination • Protection on spam-message generation • Security against DDoS and DoS attacks
Papers and publications	SecSPeer offers the possibility of publishing papers in conferences and journals regarding (but not limited to) studies of P2P systems limitations, the proposed algorithms, etc.

Table 2: PEST Analysis with respect to the SecSPeer consortium.

Chapter 5 Discussion

5.1 Potential barriers in the exploitation of SecSPeer

The potential barriers to be encountered in the foreseen exploitation/transfer of the project's results can be identified as of legal and technical nature.

5.1.1. Legal barriers

- The exploitation of the project results requires a corresponding legal framework harmonized with the existing P2P systems policy.
- In the scope of security, national barrier may come up in respect to our cyber-attacks security approach
- Till now the standards on data exchange are not fully specified and legalized and that could be regarded as a barrier.

5.1.2. Technical barriers

- SecSPeer algorithms can easily be installed and modified into existing P2P systems for exploitation and evaluation perspectives.
- In this context, the different authorities involved, perform the necessary actions to achieve the interoperability of their systems (both software and 'soft' business process systems) after having compared their respective system and information access policies.

5.2 Finding new sources of financial support

Although the initial steps for establishing the ‘exploitation network’ have already started by the SecSPeer partners, collectively on the basis of responsibilities taken up by the individual participants, available budgets and therefore the intensity of these commercialization activities are limited. Without substantial external financial support it will be virtually impossible to come anywhere near the target of a well running SecSPeer project. The more financial support becomes available, the more it will be possible to speed up the process.

5.2.1. Promoting Acceptance

Promoting acceptance is to be seen as the crucial activity in the initial phase of the strategy to establish SecSPeer results from a commercial point of view. In this respect, the four activities below should be seen in connection with the dissemination plan which gives further details.

1. **Publish papers:** All partners will collectively or individually publish papers on the SecSPeer system, its components and the approach taken, in well-established scientific and professional journals as well as other relevant publications. *For the acceptance of our results by the scientific community, publication in major scientific journals is considered to be of utmost importance*.

2. **Presentations in relevant conferences:** The SecSPeer partners will collectively and individually present papers on the project and its components in relevant international scientific and professional conferences. Also presentations in such conferences about projects undertaken by the partners, in whom the SecSPeer approach is used are also welcome, and can be utilized as an additional carrier for dissemination.

3. **Organize international workshops:** Besides presentations in the scope of international scientific meetings, the SecSPeer consortium will organize workshops within the framework of established conferences and other events, in order to inform relevant parties about the project results.

4. **Inform the Market:** Besides presentations in scientific workshops and conferences to inform the scientific community, targeted actions will be undertaken to inform commercial organizations about SecSPeer; like Gnutella, KaZaA, etc.

Result	Type	Deadlines	Comments
Requirement Analysis	- "know-how"	- Month 6	- related work - P2P penetration study - gathering and analysis of project
SecSPeer System Design	- Concept - Proposal - Contribution	- Month 12	- SecSPeer objectives - SecSPeer architecture - Traffic pattern generation – probe development - Traffic analysis – discovery of locality patterns - Development and analysis of cashing methods - P2P limitations: - Flooding problem - Message Duplicate problem - Spam Generation problem - DoS and DDoS attacks
SecSPeer Implementation	- Software - Simulation	- Month 18	- Development and analysis of network re-organization mechanisms - Development and analysis of semantic proximity mechanisms - Flood Drive algorithms – Divide and Conquer - Preventing spam generation algorithm - Preventing DoS attacks algorithm
Deployment and Evaluation	- Study - Demonstration	- Month 24	- Experimental results

Table 3: Profiling of SecSPeer Results.

5.3 Profiling of the SecSPeer Results

The consortium exploitation activities are strongly influenced by the project approach and by the individual exploitation plans of the partners.

In general terms, the SecSPeer development concepts can be considered as value-added communication applications for existing P2P systems, by providing improved algorithms related to the exchange of large size of data between many users with a secure and scalable way.

In Table 3 we present a preliminary profiling of the project results.

Chapter 6 References

[1] A. Sinha, *Client-server computing*. Communications of the ACM, 35(7):77--98, June 1992

[2] D. Clark, *Face-to-Face with Peer-to-Peer Networking*, Computer, 34(1), 2001

[3] Evangelos P. Markatos, *Tracing a large-scale Peer to Peer System: an hour in the life of Gnutella*, in the Proceedings of CCGrid 2002: the second IEEE International Symposium on Cluster Computing and the Grid. May 2002, pp. 65-74.

[4] Spyros Antonatos, *Secspeer deliverable 1.1: Requirement Analysis*, 2004.

[5] Elias Athanasopoulos and Harris Papadakis, *Secspeer deliverable 2.1: System design*, 2005.

[6] Elias Athanasopoulos and Harris Papadakis, *Secspeer deliverable 3.1: System implementation*, 2005

[7] U.S. Government, *The National Strategy To Secure Cyberspace*, February 2003.

[8] U.S. Department of Commerce, *Exportit Market Brief: Peer-to-peer Tecnologies*, by Raymond Cho, February 2002, http://exportit.ita.doc.gov.

[9] Sandvine Incorporated: Intelligent Broadband Networks, http://www.sandvine.com

[10] Information Society: The Official Greek Portal for I.S., http://www.infosoc.gr

[11] IDC, http://www.newsfactor.com/perl/story/19809.html

[12] Sandvine Incorporated: Intelligent Broadband Networks, *Peer-to-Peer File Sharing: The impact of file sharing on service provider networks. An Industry White Paper*, 2002, available from www.sandvine.com

[13] Daswani, N., Garcia-Molina, H. “Query-flood DoS Attacks in Gnutella” Proceeding of Ninth ACM Conference on Computer and Communications Security, Washington, DC (2002)

[14] *CacheLogic™:* Advanced Solutions for P2P Networks, http://www.cachelogic.com/

[15] *BitTorrent*: Peer-assisted data delivery platform, http://www.bittorrent.com/

[16] *eXeem*: The Ultimate BitTorrent Program, http://www.exeem.com/

[17] EATS Project, D6.1, “*EATS Provisional Exploitation Plan*”, Jan. 2000, http://www.siva.it/ftp/eats_deliverable6.pdf

[18] ARTEMIS Project, D2.1.1: “*Exploitation plan*”, http://www.srdc.metu.edu.tr/webpage/projects/artemis/deliverables/ExploitationPlan_FirstPeriod.doc