Economic Modelling and Incentive Mechanisms for Efficient Resource Provision in Peer-to-Peer Systems

Doctoral dissertation

Panayotis Antoniadis

Dissertation committee

Costas Courcoubetis, Professor, Athens University of Economics and Business
George Doukidis, Professor, Athens University of Economics and Business
George C. Polyzos, Professor, Athens University of Economics and Business
Panos Constantopoulos, Professor, Athens University of Economics and Business
George D. Stamoulis, Professor, Athens University of Economics and Business
Ioannis Stavrakakis, Professor, National and Kapodistrian University of Athens
Vasilis Vassalos, Assistant Professor, Athens University of Economics and Business

March 2006

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
DEPARTMENT OF INFORMATICS

To all who give with nothing in return

Acknowledgements

There are many people with the absence of which I would have never completed this dissertation. First of all, is my professor and supervisor Costas Courcoubetis. I really feel very proud for receiving his constant and inspiring guidance over the last twelve years. After all this time, he is still teaching me, not only how to do research in computer science, but how to face the most important challenges in life and how to become a better person. Unfortunately, there are no words for expressing my gratitude. Then it is all our co-authors and collaborators and particularly Prof. Robin Mason (University of Southampton), Prof. Richard Weber (University of Cambridge), and Ben Strulo (BT Exact) who provided invaluable input and feedback.

Moreover, I would like to thank Professors George C. Polyzos and George D. Stamoulis for both their direct assistance and for creating (together with Costas Courcoubetis and all my colleagues) a very friendly and productive working environment in our research laboratory. I would also like to thank professors Panos Constantopoulos, Vassilios Vassalos, and George Doukidis, from Athens University of Economics and Business, and Ioannis Stavrakakis from National and Technical Capodistrian University of Athens, for the honour of accepting to be members of my dissertation committee, and also Martha Sideri from Athens University of Economics and Business for her important help and trust.

Many thanks to all the participants in the MMAPPS IST project consortium, in the context of which the large percentage of this dissertation was carried out (and financially supported), and my colleagues Elias Efstathiou, Athanassios Papaioannou, Manos Dramitinos, Sergios Soursos, and Marina Bitsaki who provided useful feedback and stimulating discussions.

Last but not least, as they say, I wouldn't have managed anything if it wasn't for my close friends and family that constantly supported me throughout this long lasting effort with their patience and unconditional love.

Abstract

Personal computers and mobile devices have today significant capabilities, they are connected through high-speed Internet connections or directly, forming ad-hoc networks, and thus provide the potential for the deployment of sophisticated distributed applications which seek to exploit the vast amount of aggregated underutilized resources available; the so-called peer-to-peer (p2p) applications. File sharing, which was the application responsible for the p2p hype started around the year 2000, is still the most popular and the only one widely deployed. But many more are being proposed and designed so as to exploit different types of resources such as computing power, storage, access bandwidth, and more. A fundamental difference of p2p applications compared to traditional distributed systems is the fact that decisions of the individual peers are based on their own self-interest and this in principle leads to inefficient system operation. In particular, a rational peer would participate in the system without contributing any resources following the so-called 'free riding' strategy, which could in many cases have a detrimental effect on system's efficiency.

So, suitable incentives should be given to peers in order to behave towards maximizing the overall efficiency. But there are very challenging research questions to be addressed in this context both in terms of economic modelling and implementation which have attracted lately many researchers from all the relevant research fields of computer science and economics. In this dissertation we identified and categorized the most important related concepts due to current practice and research activity, and we chose to focus on the issue of content provision in p2p file sharing systems as our main case study. Content availability is a non-rivalrous resource since files are not consumed by downloading them, and thus it has the main property of what is called in economics a 'public good'. Hence, we model content provision in p2p systems as a problem of private provision of a public good. Three characteristics of p2p systems that play important role in our analysis are the following: the very large number of participants, the high degree of heterogeneity, and the very challenging implementation issues due to the fully distributed and untrusted environment. Notably, the latter is the one that gives new interest to this classic, and very challenging, problem in the economics literature.

The large system size reduces, in general, the incentives to peers to contribute the efficient amount of resources for the provision of the public good because they feel less pivotal for its construction and thus would choose to minimize as much as possible their own contribution. However, when exclusions are possible, as suggested by recent important asymptotic results (see Courcoubetis and Weber (2005) and references therein), a fixed contribution scheme, where all peers contribute the same fee when they choose to participate, is within $O(1/n)$ from the maximum social welfare that could be achieved using Mechanism Design (and the optimal game definition) as the number of peers in the system becomes large. We demonstrate the very attractive properties of the fixed contribution scheme for the case of our economic model for file sharing compared to other alternative mechanisms and discuss certain theoretical issues of practical interest such as the importance of heterogeneous file popularity, the conditions that are required for system stability, and possible ways to estimate system parameters that were originally assumed known. However, the performance of the fixed contribution scheme decreases with the heterogeneity of the peers in terms of their preference parameters. Hence, categorizing peers into different groups could be greatly beneficial. So, we study under which assumptions one could exploit such additional information and the incentives that need to be given to peers to agree to be distinguished and declare their true group type if this is not observable.

Moreover, the fully distributed and untrusted p2p environment makes even such a simple mechanism (as is the fixed contribution scheme) very difficult to enforce in practice since some sort of accounting is required, which notably faces very challenging attacks in this context. To overcome this obstacle, we propose a memory-less enforcement mechanism which ensures that peers contribute to the system during the interval of time they are consuming resources for themselves by dictating a fixed (not too high) throughput for uploading files for all peers. We present the basic system requirements for supporting the necessary functionality for this approach and formulate a suitable economic model; in this model, the cost is directly related to the time a peer is forced to stay on-line and hence contribute to others. We then assess the performance of our mechanism in terms of economic efficiency, which depending on certain parameters could be comparable to the optimal one. Additionally, this model provides the means to compare the proposed mechanism to other popular in the literature system rules incentivizing contribution by imposing constraints on the consumption of resources, such as the rule equating downloads and uploads performed by each peer. Interestingly, we show that the fact that in our mechanism we can tune appropriately its critical parameter (in particular, we can control the amount of time peers should stay connected per download) leads to increased efficiency compared to the one achieved using the rule equating downloads and uploads. Additionally, its elasticity in terms of the ratio of the actual uploads and downloads required per peer ensures in general a better stability of the system.

We also formulate for completeness two interesting economic models focusing on different aspects of the activity in a p2p file sharing system. Firstly, we propose a more elaborate economic model capturing uploading costs in addition to file sharing costs. We then introduce a general class of reciprocity rules for enforcing a desirable relation between consumption and contribution of resources according to this model and demonstrate the difficulties that arise due to this more detailed modelling approach. More specifically, we show that a simple linear rule could maximize efficiency but for the computation of its coefficients we would require complete information of peers' types and most importantly, they would need to be in general personalized (different for different peers). Unfortunately, the optimal uniform policy under incomplete information in this context is unknown and we could not compare how our simple rule would perform in the limit based on information from the distribution of the peer types as we did for the proposed fixed contribution scheme. Secondly, we identify an interesting content distribution game concerning the introduction of costly items into a p2p community and formulate a suitable market model for content distribution which leads to useful analytical results. We use this model to study under which circumstances specific privileges, and of which kind, should be granted to peers who choose to purchase and share expensive content items in order to be beneficial for them to do so.

Finally, we have also made some initial steps in order to apply parts of our results to other contexts as well. More specifically, we formulate a public good model for the case of scientific grids and consider the possibility of implementing a memory-less 'contribute while consuming' mechanism for packet forwarding in ad-hoc networks.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

The considerable and often superfluous capabilities of ordinary PCs in terms of memory and processing power, and the recent expansion of broadband has provided the means for the proliferation of the so-called peer-to-peer (p2p) systems. That is, systems consisting of machines with comparable capabilities and identical roles (the peers), and typically intermittent connectivity, which form an overlay network in order to exploit their unused resources (storage, CPU cycles, access bandwidth) and/or share available content or human presence.

Notably, the Internet was fundamentally designed as a peer-to-peer system and USENET, a fully distributed news application that is running since 1979, is considered the grandfather of today's p2p applications. However, over time the client/server model has dominated, with millions of clients requesting services (mainly downloading content) from a restricted set of servers. The increase of PC capabilities and bandwidth, the enormous growth of the Internet, the encoding MPEG, which made feasible the exchange of multimedia content, and a killer application (namely, p2p file sharing) changed drastically this model and were probably the main reasons for the p2p hype that started with Napster[1] around 2000 and until today remains in the center of attention of millions of Internet users, the industry, and researchers.

Today, numerous different file sharing applications are being developed and many novel

[1]Napster was the first p2p file sharing application which put the spark for the p2p revolution in 2000 and, at the height of its popularity, had between 40 and 80 million subscribers with up to 5 million users logged on simultaneously.

p2p systems are being proposed and thoroughly studied in order to enable the sharing of all types of unused resources. The most important are grid computing, data preservation systems, and ad-hoc networks. Although they have all attracted significant and increasing attention in the research community they have not yet been widely deployed. The main reason is that many theoretical and practical underlying issues still remain open. More specifically, free riding and distributed implementation constitute the most challenging research issues in this context.

Regarding distributed implementation, although not a necessary requirement, complete autonomy of central servers is considered an important aspect of a p2p system and is often used as a criterion for the "p2p-ness" of a specific implementation. In the case of file sharing, this need has arisen due to the severe legal battles of the RIAA (Recording Industry Association of America) against copyright infringement that is common place in most applications. Naspter was closed exactly due to its dependence on a central server to serve as the directory service. And even though it was a perfectly peer-to-peer system at the application level (users acting both as clients and servers exchanging content) it is considered a 'hybrid' p2p application. On the other hand, Gnutella, which was among the first 'pure' p2p applications implementing all the core functionality for the exchange of files between peers in a fully distributed fashion, still survives, and provides a proof of concept for the feasibility and effectiveness of a pure p2p system.

Notice that complete decentralization reduces significantly the efficiency of certain system functionality (e.g. service discovery in Napster was much more efficient and user friendly than in Gnutella) and in practice many p2p systems rely on the existence of super-peers to overcome such shortcomings. However, autonomy from central servers is desirable because, in addition to the legal issues, it improves dramatically the scalability of the system and more importantly it provides the ability to p2p communities to self-organize (the protection of personal data and the independence from profit maximizing companies is most of the time a top priority issue in p2p communities) but also to be very easily deployed without the need for investments on infrastructure. This property of the fully distributed p2p systems, contributes towards the creation of an ideal environment for experimentation on new revolutionary p2p applications and provides a very fertile ground for the inspiration and creativity of the very large and active community of Internet application developers. Thus, a significant part of the research on p2p systems focuses on the technical issues related to the complete decentralization of all necessary functionality.

Service discovery [Lua et al. (2005)], security issues [Castro et al. (2002)], and accounting [Vishnumurthy et al. (2003);Feldman et al. (2004)] are some of the most important areas of research to that respect.

But the most important characteristic of a p2p system is the fact that a collective outcome is produced by small and dynamic in nature contributions of a large number of self-interested individuals. This is more obvious to see when peers participate by simply contributing resources to a common pool, as for example in the case of another pioneering hybrid p2p application, SETI@home, which exploits the unused computational power of personal computers by using distributed computation in Search for Extraterrestrial Intelligence (SETI)[2]. But this is also the case when sharing happens through bilateral or multilateral exchange of resources. For example, in the case of file sharing, the collective outcome produced by the exchange of multimedia files between end-users is the overall content availability achieved. That is, the probability with which a user participating in such a system will manage to find and download a certain (content) item —being indifferent as for whom will be the peer that will actually provide the specific file.

So, the success of a p2p system and thus the value acquired by each peer depends directly on the degree of contributions made by all other peers. This means that there are *positive externalities*[3] associated with the amount of resources each peer decides to contribute. But which is the incentive for a certain peer to actually contribute her own resources? In the absence of explicit incentives (and suitable mechanisms for their enforcement) and when the cost associated with contribution is non-zero, the rational strategy of all peers would be to choose to free ride on the efforts of others. The equilibrium in this extreme case would be the collapse of the system.

Since the first formal identification of the 'free riding' phenomenon in Gnutella by Adar and Huberman (2000), significant research focus has been given to the incentive mechanisms that could be used to eliminate or constrain opportunistic behaviour in p2p systems. Notice that this is a fundamental issue in human societies and it has been studied thoroughly by economists and sociologists in the last century. However, the unique characteristics of p2p systems and more specifically their very large size, the high heterogeneity in terms

[2]Note that the existence of a central server for allocating the jobs to different computers and the fact that all participating users act only as servers to this single client is the reason why SETI@home is not considered a pure p2p application.

[3]In economics, externalities exist in a system when, roughly speaking, the decisions of an agent affect directly the utility or cost of another agent. Pollution is a standard example of negative externalities and provision of public goods a standard example of positive externalities.

of capabilities and peer types, the dynamic nature of the system, several intangible value and cost generators, the difficulty or inappropriateness to use prices and the important implementation restrictions, have opened a new research area with the obvious name "p2p economics", which has attracted an increasing number of researchers from all the relevant fields (computer science, economics, sociology, and business).

Although the underlying theoretical issues are very interesting as such, a realistic economic approach for addressing free riding in p2p systems should not ignore the implementation limitations of accounting and enforcing any kind of incentive mechanisms, mainly due to the distributed and untrusted p2p environment. We believe that the answers to these research questions will turn out to be rather critical as more and more users join the Internet and as the millions of connected mobile devices (phones, PDAs, etc), sensors, public and private WLAN access points will start to form a very complicated communications environment where autonomic communication and self-organization will become a necessity.

1.2 Problem formulation

In this dissertation we address the issue of designing enforceable and economically efficient incentive mechanisms for resource provisioning in p2p systems. The majority of our work is performed in the context of file sharing and focuses on content availability as the main good provided.

Detailed modelling of the economic transactions carried out in a p2p file sharing system is in general a very complex task. The main reason is that participating peers should contribute different types of resources (bandwidth, storage, CPU cycles, content) with fundamentally different characteristics, the provision of which could generate additional costs such as legal risks and time spent using the system. Moreover, utility is often derived by intangible factors additionally to the consumption of resources, such as altruistic motives, community spirit, and more. So, it is very important to make the necessary abstractions in order to construct meaningful and tractable economic models.

As far as economic efficiency is concerned (i.e. maximization of social welfare), besides the complicated modelling, an additional challenging issue is the lack of information concerning the types and preferences of the individual peers, which is required for the computation of the optimal allocation of resources and cost in a p2p system. We will see that 'first-best', optimal, mechanisms require full information on peers' type and in many cases personalization (each peer facing a different price or contribution rule) and thus realistic approaches should compromise with sub-optimal solutions.

In addition to this really challenging economic resource provision and allocation problem, a p2p system designer has to deal with the inability to rely on trusted software[4] or on central entities that can monitor and account for peers' transactions to ensure that they contribute and consume the amount of resources dictated by an underlying economic model. A significant part of the research literature in p2p economics adhere to the principle that all peers should acquire from the system as much as (or proportional to what) they contribute in terms of actual downloads/uploads and studies the game theoretic and implementation issues related to the effective accounting of peers' transactions (by means of reputation, credits, etc.) and enforcement in a fully distributed and untrusted environment (e.g. Feldman et al. (2004); Hausheer et al. (2003); Kamvar et al. (2003b) among many others). The most challenging issues in this context are the whitewashing and false trading attacks (see Section 2.6.2).

BitTorrent [Cohen (2003)] is an example of a real world application which implements such a reciprocative incentive scheme but without relying on information about the past transactions of peers: peers downloading the same item exchange directly their upload bandwidth. Because the incentive scheme does not rely on tracking the long term behaviour of peers it is simple to implement and largely immune to problems of false trading and whitewashing. However, since it is designed to incentivise the provision of bandwidth, file sharing applications based on this technology do not necessarily tackle the issue of content availability, which we believe that is of great importance, especially as access bandwidth increases and flat based pricing is used by most operators; this means that download times get shorter and extra incentives will be probably required in the future for peers to stay online sharing their files. Notably, BitTorrent-based applications usually provide a relatively small number of very popular (and often very large) content items.

Recent articles in the popular press (initiated by Anderson (2004)) discuss the importance of the 'long tail' of content; that large part of the set of content in which individual files are not popular, but which together constitute the majority of the total requests, and

[4] Kazaa is a characteristic example of a real world application that tried to implement a reciprocative incentive mechanism (giving priority to peers that contribute more in terms of uploads/downloads ratio), which failed due to a hacked version of its software —see http://www.k-lite.tk/

thus often generating larger value than the popular ones. The provision of this part of the content in a p2p system requires different types of incentives than the ones usually discussed in the p2p economics literature and used by most popular p2p applications (like Bittorrent).

On the other hand, Direct Connect[5] is another p2p application which relies mainly on central control to enforce minimum contribution rules, expelling misbehaving peers based on their IP addresses. Most of these rules dictate a minimum total size of content shared by each peer without any restriction on the number of requests issued and thus gives priority to content availability. Notably, the participants in the Direct Connect network usually have high access bandwidth and very large storage spaces and even standard techniques for distributed content distribution of single items, such as swarming, are not used at all. Also note that BitTorrent and Direct Connect are examples of the two extremes of p2p file sharing software whose focus in general ranges from mainly content distribution (as BitTorrent does) to mainly content availability (as in the case of Direct Connect).

We define content availability as the total number of distinct files made available in a p2p file sharing network over time. So, the contribution of each peer, in terms of content availability, is the number of files stored in the shared folders of her PC and the time she stays connected in the p2p network. We believe that this time (peer availability) will become a critical parameter of the contribution of peers in this context, especially as people store more and more content in their PCs for their own use and access bandwidth becomes cheaper and excessive. So, in the majority of our work we assume that uploading cost is negligible. Moreover, being particularly interested in the long-tail of the content, we choose to consider systems where the probability of a certain file being requested is low but the overall value of satisfying such requests is much greater than for popular items; mainly because of their large number as explained above but also because it is in general more difficult to find unpopular or rare items, in many cases even if one wishes to pay for them. So, we don't account for congestion effects[6]. Under these assumptions, content availability is a *public good*: the copying of a file by one peer does not prevent another peer also from copying it; but contributing files to the common pool is costly.

The design of the incentive mechanisms required for such a system to address free riding

[5] At the time of writing of this thesis the official web page of Direct Connect was closed due to the legal battles of the RIAA —see http://en.wikipedia.org/wiki/NeoModus_Direct_Connect.

[6] Note that this assumption is still valid under the existence of popular items when the number of their copies is such that the corresponding request rate is similar to that of unpopular items.

and operate at acceptable levels of efficiency involves challenging research questions both in terms of economic modelling and system design. Our first goal is to formulate a simple but expressive economic model of a p2p file sharing system and define the notion of economic efficiency in this context. In order to achieve this we should first identify the most relevant concepts from economic theory (mainly from public good theory and mechanism design) and review the existing results evaluating their applicability in our context.

The overall objective of this theoretical part of our work is to devise suitable incentive mechanisms that would bring the system to an efficient equilibrium, as this is characterized by our economic model, under different assumptions on critical system parameters and informational constraints. Our focus is on content availability and the public good perspective of p2p resource provisioning but we also wish to explore the case where costs for uploading and/or due to congestion are not negligible, formulating and analyzing suitable economic models. Moreover, we will study under which circumstances the suitable incentives —and with what kind of mechanisms— could be provided in order for high value items to be introduced for the first time and shared in a p2p community.

Our ultimate goal is to design a system that implements the incentive mechanisms produced by our theoretical work without requiring the existence of any sort of user memory (means to gather and manage accounting information concerning users' past transactions) and/or the ability to permanently expel peers from the group, following the implementation principle of BitTorrent and thus avoiding the very challenging issues involved in the implementation of a reliable accounting functionality in a fully distributed and untrusted p2p environment. We also wish to identify which aspects of our solutions have general applicability and then explore to what extent they could be also suitable in other contexts such as grid computing and ad-hoc networking.

So, the main research questions that we address in this dissertation include the following:

- Which are the most dominant characteristics of p2p systems as far as economic modelling and incentives for resource provision are concerned?

- Which is the simplest economic model that captures the aspect of content availability in p2p file sharing systems? To what extent similar models could be applied in different p2p systems as well?

- Which are the trade-offs that arise due to the underlying informational and implementation restrictions of a realistic p2p application?
- What is the effect of negative externalities caused by content requests? To what extent can we address this additional modelling complexity?
- What kind of markets should be created in order for high value items to be introduced into a p2p community?
- Can we design enforceable incentive mechanisms that eliminate or at least limit free riding and bring the system to near optimal equilibria (according to theory)? Under which assumptions?

1.3 Contribution

We have tried to answer the above questions in the order presented. That is, we followed a top-down approach trying first to model and analyze p2p resource provisioning and design optimal incentive mechanisms for increasing the economic efficiency of a p2p file sharing system and then explore ways to enforce these incentive mechanisms or good approximations of them in a realistic context. The fact that we search for realistic approaches motivated us to avoid using prices and payments as an incentive mechanism but constrain ourselves to practical system rules controlling directly the consumption and/or contribution of resources[7].

As already mentioned, content availability is a public good. Thus, our problem resembles to a classic problem in the economics literature: the private provision of a public good. According to economic theory (see [Varian (1992)]), the existence of externalities in this problem rules out as inefficient the adoption of a free-market approach (either price or token based) for controlling a p2p system and thus some sort of regulation is required to lead the system to efficient equilibria (those that maximize social welfare). Then, the most challenging problem is how the system designer can acquire the necessary information from the participants (the peers) to compute an efficient provision and allocation of resources.

Fortunately, there is an important result due to recent asymptotic analysis of certain public good models (see Courcoubetis and Weber (2005) and references therein) that proves the following: under incomplete information (when only the distribution of the peers' types

[7] Prices are only used for comparison purposes.

are known to the system designer) and when the aim is to maximize social welfare, a fixed contribution scheme can be asymptotically optimal as the number of participants, n, grows to infinity. Most importantly, we can compute such contributions using a simple optimization problem. Such a simple scheme eliminates free riding, is incentive compatible and obtains a value of social welfare that is within $O(1/n)$ of that obtained by the second-best policy of the corresponding mechanism design formulation of the problem. We adapt such models for the case of p2p file sharing and show using numerical analysis the attractive properties of a fixed contribution scheme, one dictating the sharing of a specific number of files per unity of time by each participating peer. This result means that it is not necessary to collect large amounts of information, and/or to undertake complicated calculations, in order to implement the correct incentives in a large peer-to-peer system.

So, we focused on the fixed contribution mechanism and studied several interesting theoretical aspects concerning its application in a realistic environment. More specifically, we analyse in depth the case of group formation, when peers belong to different classes (such as DSL and dial-up users). This analysis is particularly interesting in the case of p2p systems which, according to measurements, usually consist indeed of a very heterogeneous mixture of peers in terms of capabilities and behaviour. In our work, we explore how one could take advantage of this extra information to increase the efficiency of the system. We identify under which assumptions peers will agree to participate disclosing their true type and when they won't have the incentive to do so or even prefer to form their own distinct groups, showing the corresponding trade-offs that arise under the different possible incentive mechanisms employed. We also extend our public good model to account for file popularity, and discuss properties of the resulting equilibria. Moreover, we consider the evolution of the system to equilibrium in its early life, when both peers and the system planner are still learning about system parameters. We finally discuss the game that occurs when peers know that a fixed fee will be used, but the distribution of their valuations is unknown to the system designer.

We also analyse the case of negative externalities. In order to do so we formulate a more detailed economic model capturing uploading costs, in addition to file sharing costs. In this model peers are characterized by both their preference parameter (how valuable is content availability to them) and their request rate. Note that in our original public model, where uploading is assumed to impose negligible costs on uploaders, peers are characterized only by their preference parameter. We introduce then a general class of reciprocity rules enforcing a desirable relation between consumption (controlling the request rate of a peer which imposes costs on others) and contribution (enforcing a certain contribution by each which increases the value of others) and compute the optimal rule coefficients (and the corresponding optimal prices) in the context of our model. We show that a linear such rule can maximize efficiency but for the computation of its coefficients we need complete information and they need to be in general personalized (different for each peer)[8]. Hence, one should compromise as far as economic efficiency is concerned if a uniform such rule is to be employed, the only practical option in almost all cases. Unfortunately, the optimal uniform policy under incomplete information in this context is unknown and we could not compare how our simple rule would perform in the limit based on information from the distribution of the peer types as we did for the proposed fixed contribution scheme. As a result, only simple heuristics could be considered in this case (such as averaging), which although they could be employed in practice to increase system's efficiency, they have limited theoretical interest. Finally, we extend our model to capture the notion of the quality of service offered by peers in the context of a novel p2p application for WLAN access sharing and compute for this case also the optimal prices and rules.

We conclude our theoretical work addressing a slightly different problem: the provision of incentives for the introduction of high value items in a p2p community. We formulate a model which treats peers as independent economic agents buying and selling content. We investigate the basic economic properties of such a market managed p2p content distribution network and the corresponding trade-offs. Initially, we assume that no peer has the content, and there is a substantial initial cost to bring it within the peer group. The bargaining position and hence the price that can be posted by an agent having the content depends on the cost to transport the content to the requesting peer, its value, and the number of other agents providing the same content. We discuss the influence of parameters such as the maximum number of the competitive offers allowed by the system, content popularity, its value to the agents, and the transport costs, taking into account the risk of the first agent incurring the initial content cost. We believe that complete lack of control of the available information may be detrimental in certain cases where substantial costs are involved. The goal of our model is to substantiate this and provide for some quantitative results.

[8] unlike prices for which, as the size of the network grows, we can compute a uniform one that would maximize social welfare.

Although the main motivation of this dissertation was the challenging theoretical issues concerning p2p resource provision, our ultimate goal is to propose incentive mechanisms that have both interesting economic properties but are also enforceable and thus constitute possible solutions for realistic p2p file sharing applications.

So, we also explore how to implement economic incentives in a memory-less p2p system, one that doesn't have the ability to account for peers' contributions over time but should rely only on the time peers are consuming resources to ensure that they contribute as they should according to the underlying economic model. More specifically, according to the fixed contribution scheme, we first require that they provide a certain number of content items of their choice, regardless of request rate (both their own and this of the other peers). Note that the independence from request rate is motivated by the fact that we focus on content availability rather than content distribution and that we consider uploading costs to have a second-order effect[9]. But we also need to give incentives to remain in the system and providing content, not simply to have provided at some time in the past. As already mentioned, we believe that this part of a peer's contribution will probably become dominant in future p2p systems.

We thus propose the following 'contribute while consuming' incentive mechanism enforced by uploading peers, who

1. check that the downloaders share a predefined number of valid files, and
2. use a certain (not too low) upload throughput in order to ensure that these files are made available for a significant amount of time.

Notice that this mechanism cannot enforce the same contribution for all peers as we would ideally wish. The reason is that the final contribution of a peer depends on her request rate, since it is only while consuming resources that she is forced to contribute (however, as we demonstrate in Section 5.4.6, this fact could under certain assumptions be beneficial for the system). Moreover, there are some important implementation and incentive issues that arise in this context (e.g. ensuring the validity of files shared, the need for super peers in order to avoid the requirement of cycles of requests to be formed, and more). However, we believe that our proposed mechanism constitutes a good compromise

[9]Notice that this is especially so because according to the requirements of our mechanism, uploading cost will be incurred only while peers are consuming resources for themselves. And Feldman et al. (2003) show that in many realistic cases including the ones we are considering, prioritizing TCP ACKs constitutes the cost of uploading negligible while downloading.

between economic efficiency and implementability and should lead to some interesting and practical solutions for providing incentives for content availability in p2p systems.

In this dissertation we describe the functionality that should be supported for enforcement and discuss the additional incentive issues that arise in this context, proposing some practical solutions to address them. We also formulate and analyse a suitable economic model with interesting properties, in order to build a sound theoretical framework for the study and evaluation of the qualitative characteristics of our memory-less incentive mechanism and provide insights for the appropriate tuning of its basic parameters: the number of files shared and the upload throughput used by all peers in the system. Our results show that the resulting efficiency depending on certain system parameters could be comparable to the optimal one. Moreover, this model provides the means to compare our proposed mechanism to other popular in the literature system rules incentivizing contribution by imposing constraints on the consumption of resources such as the rule equating downloads and uploads performed by each peer. Interestingly, we show that the fact that in our mechanism we can tune appropriately its critical parameter (the amount of time peers should stay connected per download) leads to increased efficiency compared to the one achieved using the rule equating downloads and uploads. Additionally, its elasticity in terms of the ratio of the actual uploads and downloads required per peer ensures in general a better stability of the system. We believe that these results are encouraging concerning the applicability and effectiveness of our proposed mechanism taking into account its limited implementation requirements.

Concluding our work, we have made some first steps towards applying parts of our results to other contexts as well. More specifically, we have formulated a public good model for the case of scientific computational grids and consider the possibility of implementing a 'contribute while consuming' mechanism for packet forwarding in ad-hoc networks. These are part of our on-going and future research.

1.4 Dissertation outline and publications

We present in this section the contents of this dissertation summarizing at the same time our contribution. In Chapter 2, we analyze the key concepts involved in the design of incentive mechanisms for p2p systems and present the state of the art in each area, placing into the picture our own contribution as well. In Chapter 3, we present some important results from public good theory and motivate their applicability for incentivizing resource provision in various p2p systems such as file sharing, grid computing, and WLAN peering. In Chapter 4, we formulate an economic model of p2p file sharing focusing on content availability and use a simple instance of our public good model in order to compare the complete and incomplete information cases using numerical analysis and demonstrate the attractiveness of the proposed fixed contribution scheme; see also [Antoniadis et al. (2004a)]. We then focus on the fixed fee contribution policies and discuss several theoretical issues associated with the employment of the fixed contribution scheme in a realistic p2p environment, such as the advantages and incentive issues of group formation, dealing with heterogeneous file popularity, ensuring the stability of the system, and discovering system parameters that are originally assumed known; see also [Antoniadis et al. (2004b)]. In Chapter 5, we present and analyze in depth, both in terms of implementation requirements and economic efficiency, a realistic version of the fixed contribution scheme, which does not require the use of system memory but relies only on the time peers are consuming resources to ensure that they contribute adequately; see also [Antoniadis et al. (2005)]. In Chapter 6, we propose a more detailed economic model capturing the cost for uploading a file. We introduce a class of reciprocity rules relating consumption with contribution which could be used instead of prices for maximizing social welfare and compute the optimal prices and rules under different assumptions concerning the availability of information. We also extend our model to capture the notion of the quality of service offered by each peer, in the context of a novel p2p application for WLAN access sharing and compute again the optimal prices and rules (see [Antoniadis et al. (2003a); Antoniadis et al. (2003b)]). This part of our work is included in its entirety in [Antoniadis et al. (2003)]. In Chapter 7, we propose a market model for the case where high value items are to be introduced in a p2p community and study the trade-offs that arise in relation to the information made available to potential content buyers; see also [Antoniadis and Courcoubetis (2002)]. Finally, Chapter 8 summarizes our results and discusses the future perspectives of our work.

Chapter 2

P2P Economics: basic concepts and state of the art

In this chapter we identify the most important aspects of p2p systems that constitute the problem of providing the suitable incentives for resource contribution so challenging, and novel, both in terms of theory and implementation. We introduce some fundamental principles and concepts from the theory of economics and system design that play a central role to this end, demonstrating their inter-dependencies. We also present a taxonomy of the different research questions that arise in this context, and the related work in each case, placing into the picture our own contribution as well.

2.1 Applications

In a peer-to-peer system there is always some sort of resource exchange between end-users. The type of the main resources involved and the corresponding application influence to a large extent the incentives of peers in terms of resource contribution and consumption, and their overall behaviour. Hence, the theoretical analysis and implementation issues can differ significantly from system to system. The main focus of this dissertation is on p2p file sharing. However, we will explore in this section the special characteristics of other applications as well in order to make an effort to distinguish between the general and application-specific aspects of p2p systems and investigate to what extent our solutions (or parts of them) are applicable in other contexts as well.

Clearly, we are interested in applications in which the incentives are not aligned. That is, when peers do not have any direct benefit from contributing resources but they would ideally prefer to consume resources provided by other peers, free riding on their efforts[1]. We describe below some important p2p applications and the different resources involved and which are susceptible to free riding.

2.1.1 File sharing

Since the beginning of the p2p hype in 1999, the killer p2p application in the Internet continues to be file sharing. That is, the direct exchange of files (mainly multimedia and software) between end users. Although research on p2p economics is not at all constrained to this type of applications, most researchers in this field have been constantly keeping an eye on the evolution of file sharing; as new file sharing applications attracting the vast majority of users are coming into the picture, new challenging research questions arise and a number of measurements and analysis follow to quantify the levels of cooperation achieved and understand the behaviour of the users and the characteristics of the traffic generated.

We describe below the evolution of file sharing applications as a result of the efforts of their designers to achieve two different and in many cases conflicting goals: to protect the software companies and the users from the legal battles of the RIAA and to reduce free riding.

Confronting the RIAA

Because the majority of files exchanged in p2p file sharing applications are copyrighted, their evolution has been much affected by the severe and on-going legal battles of the RIAA (Recording Industry Association of America) against companies which build software that "enables users to exchange copyrighted material" or even the users themselves. Napster was the first victim of these efforts due to its dependence on a central server for indexing all files shared and serving the content queries[2]. Gnutella, which was created in response to the valid attacks towards any centralized body like Napster, still survives due to its

[1]Notice that collaborative applications or on-line games do not belong to this category since peers acquire direct personal benefit from contributing their resources.

[2]This meant that Napster had the ability to filter copyrighted files and thus failed to convince the court that its service fell within the provisions of the Digital Millennium Copyright Act, 17 U.S.C.A. 512(a). Napster is now resurrected as a commercial online music service that competes with other commercial services like iTunes and Rhapsody.

complete independence from central servers.

However, complete decentralization of Gnutella came at a cost: the inefficiency of content search (both in terms of effectiveness and significant signalling traffic) due to the equal participation of all nodes in the search protocol. The next generation of p2p applications (including Kazaa/FastTrack[3], eDonkey, Gnutella2 and many more) acknowledged the high degree of heterogeneity between Internet users and lied somewhere between these two extremes: they employed an hierarchical architecture in which peers with more resources (access bandwidth, availability, storage space, etc.), the super-peers, form a core overlay network and each one is responsible for a subset of the whole index (the part provided by the normal peers attached to her). This solution makes content search much more efficient and scalable, and reduces signalling traffic without relying on a central index.

KaZaA was the most popular application of this second generation [4]. Since KaZaA did not have the means to stop its users from exchanging copyrighted material, was protected by the Digital Millennium Copyright Act. Thus the reaction of the RIAA was to start filing civil lawsuits against normal users[5] increasing the cost for sharing and thus harming content availability. Despite all these efforts p2p file sharing networks continue to flourish and the numerous applications such as Grokster, eDonkey, BitTorrent, OpenNap and WinMx, Gnutella, MP2P, Soulseek, host millions of users and petabytes of files shared but the battle has not still ended[6]. So, a third generation of p2p applications, which tries to preserve the anonymity of its users (and thus protect them from possible lawsuits) such as the so called Friend-to-Friend networks[7] and p2p anonymity systems[8] have started to gain the attention of the developers and become popular amongst users.

Another strategy of record companies against piracy was to inject a massive number of

[3] Kazaa is actually the most popular p2p client of a more generic p2p network called FastTrack.

[4] Saroiu et al. (2003) focused on the file-sharing workload of KaZaA from a traffic engineering perspective and provide some indications for the user behaviour and content popularity. Liang et al. (2004) analyze the architecture and overlay behaviour of KaZaA, providing evidence for the existence of a significant amount of peers that decide to act as super-peers without any particular incentives for doing so. The assumption for the existence for such peers is in many cases crucial for the design of realistic incentive mechanisms.

[5] http://www.microscope.co.uk/Article124720.htm

[6] At the time of writing of this thesis, the Supreme Court has offered movie studios and record labels an important victory against file swapping, ruling that peer-to-peer companies such as Grokster could be held responsible for the copyright piracy on their networks. However, besides the close down of Grokster, over 10 millions of users are still on-line at any time in several p2p networks and thus the RIAA is now extending its efforts towards European countries as well.

[7] In friend-to-friend networks, only your friends can know that your address is used to exchange files.

[8] p2p anonymity systems preserve the anonymity of a sender by relaying her packet through a number of intermediate nodes [Figueiredo et al. (2005)]

noisy or corrupted multimedia files (mainly popular ones) into the most popular p2p networks creating pollution and reducing significantly the value of the system. Christin et al. (2005) analyze the potential effect of this strategy (called poisoning) on content availability, while Liang et al. (2005) demonstrate its negative consequences in practice, performing measurements in the Kazaa/FastTrack network and propose some possible mechanisms to detect polluted files.

Overcoming free-riding

The provision of suitable incentives for contribution was the second most significant driving force of the evolution of p2p applications. Adar and Huberman (2000) were the first to verify that p2p file sharing systems are vulnerable to free riding. Analyzing user traffic on the Gnutella network they found that 70% of Gnutella users share no files and 90% of the users answer no queries.

MojoNation[9], was the first real life p2p system that made an effort to combat free riding. It followed a free market approach (using an internal currency called Mojo) and attracted the interest of the researchers on the emerging at that time field of p2p economics. In Mojo Nation, users could choose to contribute disk space, bandwidth, and processing cycles to the network in exchange for Mojo. They were enabled to set their own prices for these online resources. However, it very soon shut down possibly because the free market created failed to capture the positive externalities involved (see Section 2.4.1). Morever, peers earned Mojo by operating services for the network rather than offering content, as one might suppose. When they published content, it cost some Mojo because their software had to pay block servers to store their blocks (information was split into many pieces – blocks– and was distributed to several block servers). Probably this fact and the small value of Mojos (since it was rather easy to obtain free Mojo) played an additional role in its failure.

Some second generation p2p file sharing applications incorporated simple system rules for incentivizing contribution of resources. In most cases these rules were not very strict but just ‘compensating’ (i.e. good behaviour resulted in increased QoS). In Kazaa for example, the contribution of each peer is computed[10] and, according to its level, peers have the

[9]Although MojoNation is not functioning anymore, archived information is available at http://web.archive.org/web/20010207192406/mojonation.net/docs/faq.shtml.

[10]The client application calculated the difference between bytes downloaded and uploaded and provided a characterization expressing the peer contribution level.

corresponding priority in case of congestion. Other systems had even more simple rules targeting the number of files shared rather than the actual uploads offered. For example, in Audiogalaxy the number of concurrent downloads that a peer could make depended on the number of files that she shared. Direct Connect, as already mentioned, was the first real application to employ strict membership rules (e.g. many communities required from their members to share a minimum —often very large— total size of multimedia files) and non-conforming members were expelled from the specific community (based on IP addresses and using central control).

The next important evolution in p2p software concerning incentives came with BitTorrent [Cohen (2003)]. BitTorrent software implements a reciprocative mechanism for bandwidth sharing between downloaders of popular content items from central servers (the trackers). More specifically, peers choose to upload to those peers from whom they are currently receiving the maximum download throughput and as a result some sort of a direct exchange of resources is taking place. The success and interesting properties of BitTorrent's incentive mechanism have motivated an increasing number of research work on BitTorrent and similar approaches following the principle of memory-less enforcement, as does our approach, which we review in Section 2.6.4.

Note that BitTorrent is just a content distribution software, enabling peers to share their upload capacity for downloading quickly a single popular file, and could be actually part of any generic file sharing application in situations where more than one peers download the same file from another peer. In general, depending on the cost of uploading and congestion, the focus of p2p file sharing software ranges from solely content distribution (or p2p streaming in case of real-time content), as BitTorrent, to solely content availability, as Direct Connect, which did not provide even the standard 'swarming' technology.

2.1.2 Content sharing

Another interesting observation regarding p2p file sharing is that in general content shared by a peer could be either created by herself, or copied from an external source (a music file ripped from a CD, a photo scanned from a book, etc.) typically copyrighted, or just copied from another peer (again either self-made or copyrighted).

The majority of content in p2p file sharing systems today is copyrighted multimedia files and software and thus the corresponding incentive issues are highly affected by this fact (legal issues, altruism, focus on content distribution, etc.). We use the term content sharing to denote the sharing of self-made content. For example, recommendation systems and USENET are such applications, since peers contribute their personal reviews or expertise for items of interest. Again, the contribution of content is subject to free riding (e.g. users that take advantage from others' reviews but never contribute their own), but in these systems the cost of content distribution is negligible and actually, in most cases central servers are used for hosting all the posted information. On the other hand, if the information contributed has considerable size then both content creation and distribution would be subject to free riding and thus suitable incentives should be provided for both. This would be so for example in the case of self-made multimedia content (e.g. home videos).

In any case, an important issue for sharing self-made content is indexing. Notice that in case of copyrighted multimedia content this is provided for free from record companies. Interestingly, collaborative tagging, the process by which many users add metadata in the form of keywords to shared content [Golder and Huberman (2005)] has made much more effective p2p content sharing applications such as photo sharing (e.g. Flickr). The reason is that it provides a flexible and distributed solution to the problem of indexing. Earth coordinates are also a predefined indexing scheme based on which very interesting content sharing applications could be built as for example sharing of high definition photos of different 'geotagged' places implementing a p2p version of the earth.google application with finer granularity.

In our approach, we do not distinguish between copyrighted files or not, but we do consider systems in which content creation and availability is in the responsibility of peers themselves, and their cost is more important than this of content distribution. An important characteristic of content sharing applications, which differentiates them from sharing of copyrighted files, is that there exist additional social incentives for contributing content such as personal recognition, community spirit, etc. (see [Lui et al. (2002)] and references therein), which should be taken into account by the underlying application. We do not include such social incentives in our modelling work but we believe that the formal study of the social perspective of p2p systems is a very interesting future direction.

2.1.3 Ad-hoc networks

Ad-hoc networks are, typically, wireless multi-hop networks formed by a set of mobile nodes without relying on a preexisting infrastructure. In order to make an ad-hoc network functional, the nodes are assumed to follow a self-organizing protocol, and the intermediate nodes are expected to relay messages between two distant nodes.

Hence, the basic service exchanged between peers in an ad-hoc network is routing and packet forwarding. Unlike most p2p networks, end-to-end communication in ad-hoc networks requires interaction of intermediate nodes to create an end-to-end path from the source to the destination; Indeed, in ad-hoc networks direct sharing of resources occurs among neighboring nodes, whereas in p2p networks any two peers can directly share their resources, independently of their location. So, a large body of the literature in this context addresses issues related to network formation and routing.

Regarding incentives, a very important characteristic of ad-hoc networks is that the main resources involved are congestible (radio spectrum) and consumable (battery power), whereas the main resources in most p2p systems are either non-rivalrous (e.g. content) or at least renewable (e.g. bandwidth). So, designing suitable incentive mechanisms for resource provisioning is crucial in this context and actually a large percentage of the p2p economics literature concerns the provision of incentives for packet forwarding in ad-hoc networks (see for example [Srinivasan et al. (2003); Buchegger and LeBoudec (2002); Zhong et al. (2003a)] among many others).

However, most of the related research questions remain open and there are not yet any real life applications of ad-hoc networks widely deployed despite the fact that a large percentage of people now carry mobile devices with significant capabilities. Another reason for this is possibly the fact that no killer application has appeared yet in this context. Interestingly, due to the scarcity of resources entailed, recent approaches focus more on less demanding applications such as delay-tolerant networks[11]. [Huang et al. (2004)] discuss the underlying trade-off between implementing incentives for contribution and encouraging participation in ad-hoc networks.

[11] http://www.dtnrg.org/wiki

2.1.4 Grid Computing

According to Foster et al. (2001), grid computing refers to "resource sharing that is not primarily file exchange but rather direct access to computers, software, data, and other resources, as is required by a range of collaborative problem solving and resource-brokering strategies emerging in industry, science, and engineering". SETI@home is considered as one of the first computational grid applications, where end-users offer their computing resources in search for extraterrestrial intelligence with mainly altruistic motives (the possibility to be included in the top contributor lists could provide additional incentives for some of them).

The allocation and provisioning of resources in a Grid system is a very challenging problem both in terms of economic efficiency and implementation. Although the existence of the appropriate incentives for participants to share their resources is a critical requirement for its success the majority of the current research efforts on grid computing has focused on important standardization and implementation issues including resource abstraction, scheduling, security, etc. Globus, NextGrid, and Akogrimo are the most active players in the development of grid middleware.

However, economic modelling and design of suitable incentive mechanisms will play a central role for the deployment of grid services and it should be given the necessary focus before Grid technology is fully settled. This of course depends on the envisioned application of the grid technology. Two are the main directions: 1) commercial exploitation and innovative business models and 2) resource sharing communities (such as Planetlab).

The resources exchanged in a Grid system are mainly commodities (e.g. CPU cycles or storage space) and rivalrous (when they are used by a user they are not available to others). This fact motivated many researchers to consider the implementation of market mechanisms for determining prices of resources, envisioning a global Grid market as in the work of Krauter et al. (2001). There is no grid-specific work, however, on how these should operate and is assumed that somehow any existing mechanism from the economic theory (auctions, commodity markets, bartering, etc.) could be used to supply prices for the resources traded.

Many companies such as IBM are developing business models to create very large pools of resources and exploit them commercially offering tremendous capabilities to possible customers (from large companies who wish to outsource certain functionalities to individual scientists who want to run very demanding experiments), moving from the "buy a computer" business model to "computing on demand". To this end, Kenyon follows a more focused business oriented approach assuming a fully assymetric Grid in terms of contribution and consumption of resources (that is, there are some participants that act mainly as providers, and others that act mainly as consumers) and propose some financial engineering paradigms for Grid computing [Kenyon (2004); Kenyon (2003)].

But grid computing could, in principle, be also operated in a peer-to-peer basis and there are those that envision a global universal grid (making possible the Sun's famous statement "The network is the computer") where all users could share their unused resources to built a single supercomputer for running computational intensive applications. However, such a grid application has not yet widely deployed in the Internet (and there is criticism as whether will ever happen[12]). There are though more constrained systems under development, limited in the exchange of resources among well defined and symmetric communities, such as the scientific grids, which have many similarities with file sharing concerning the design of incentives mechanisms (see Section 3.6.3 for a discussion on the application of our public good model to resource provisioning in scientific grids).

2.1.5 Data preservation systems

The main idea behind p2p storage is the fact that local storage space is often overprovisioned while there is often a strong need for redundant storing of important material to ensure its preservation in the future. There are several ambitious such systems under development the last years such as Oceanstore[13] and PAST[14], many more have been proposed in the literature such as pastiche [Cox et al. (2002)], samsara [Cox and Noble (2003)], and pStore [Batten et al. (2001)], and some, more focused ones, already implemented, such as LOCKSS[15]. In all of them, peers seek to exploit the unused storage space of their peers in order to create redundant copies of their data. They could be seen as long-term storage grids.

And this is exactly the unique characteristic of preservation systems: long-term service provision; peers allocate resources to one another for very long periods of time and they do not just provide a specific service with a certain duration. Notice that the value doesn't lie

[12] see for example http://www.shirky.com/writings/grids.html
[13] http://oceanstore.cs.berkeley.edu/
[14] http://research.microsoft.com/ antr/PAST/default.htm
[15] http://lockss.stanford.edu/

on the resource itself (storage space) but on replication since in most cases a peer should allocate a proportional amount of storage used for the data of her peers. So, in contrast with computational grids and file sharing systems, in preservation systems peers dedicate part of their local resources "forever" enjoying the use of exactly the same type of resources in a different place.

This fact strongly influences the design of mechanisms for incentivizing the contribution of resources. First of all, it makes its need even more compulsory since the amount of available resources is finite (in contrast with file sharing where resources are renewed over time) and generous peers could only to some extend support the system alone.

On the other hand, it facilitates the identification of peers for a reciprocative exchange of resources either directly between pairs, as assumed by Cooper and Garcia-Molina (2002) and Lillibridge et al. (2003), or enabling the formation of cycles of service provisions, for example, as proposed by Cox and Noble (2003), with the use of storage 'claims'. Other systems make this selection of peers deterministically based on DHT (Distributed Hash Table) overlays[16] (see Section 2.2.8 for a short summary of DHTs) and thus enable significant savings of storage space when identical items are stored by different peers. But then accounting, which is a very challenging problem (see Section 2.6), is required to provide the suitable incentives However, the long-term characteristic of service provision in p2p storage systems facilitates accounting as well since it provides the ability to peers to constantly audit whether another peer fulfills her obligations. Ngan et al. (2003) propose such an auditing mechanism relying on the existence of some trusted means of certification. Finally, in addition to the incentive issues, storage systems have also to address the trade-off between consesus and preservation, discussed by Bungale et al. (2005), security issues [Wallach (2002)], anonymity [Clarke (1999)], and more, which are out of the scope of this dissertation.

2.2 Important characteristics of p2p systems

As it is probably apparent, many interesting and challenging research questions are related to p2p resource sharing and they differ significantly from application to application. In this section we will try to identify and categorize some generic characteristics of p2p systems and highlight their importance concerning the provision of suitable incentives for resource

[16]PAST is based on Pastry, pStore [Batten et al. (2001)] on Chord, and Oceanstore on Tapestry.

contribution.

2.2.1 Public good aspect

Besides the bilateral or multilateral exchanges of resources in a p2p system (e.g. file uploads between two or more peers), which in general generate utility to the clients and cost to the providers of the specific transactions, peers making available their resources in a peer system, being the only providers, participate in the provision of a *public good*. For example, in the case of p2p file sharing, the public good constructed is content availability. Similarly, peers at SETI@home and similar applications (e.g. cancer research), contribute to the advance of science. Even in the case of ad-hoc networks or the Grid, where all resources involved are rivalrous, it is the participating peers themselves that provide the overall service (e.g. the ability to route a packet through an ad-hoc network) and thus all together contribute to the provision of a "public good", in this case the system itself.

Thus, in p2p systems, the number of peers in the system affects significantly the value that each one of them receives by participating, by influencing the amount of resources available. In other words, there exist positive externalities[17] associated with the amount of resources a peer will decide to contribute, a fundamental characteristic of public goods.

The provision of public goods has received significant attention in the economics literature. Formally, a pure public good has two characteristics. First, it is *non-rivalrous in consumption*—that is, one agent's consumption of the good does not diminish the amount of the good left for other agents to consume. Second, the good is *non-excludable* —there is no (feasible) technology whereby agents can be prevented from consuming the good. The classic examples of public goods are street lighting and national defence. When one individual benefits from national defence, it does not reduce the amount of national defence left for other individuals. Similarly, content in p2p systems is not consumed when a file is exchanged between two peers. And short of extradition, individuals cannot be excluded from the benefits of national defence. Nor can anyone be excluded from benefiting from a discovery made by a cancer research p2p system. There are also goods that have aspects of a public good, but do not have these extreme features. For example, some goods are non-rivalrous in consumption, but are potentially excludable. Such goods are known as

[17] In economics, externalities exist in a system when, roughly speaking, the decisions of an agent affect indirectly the utility or cost of another agent. Pollution is a standard example of negative externalities and provision of public goods a standard example of positive externalities.

club goods. Other goods are non-excludable, but are rivalrous in consumption; these are *impure*, or congestible public goods.

The standard problem with public goods is ensuring that agents have sufficient incentives to contribute to their provision. When self-interested peers are allowed unrestricted use of the system (as is roughly the case with e.g., Gnutella or Kazaa), they will not consider the benefits that accrue to other peers. The result is *free riding*: each peer has an incentive to decrease its contribution below the efficient level, while still enjoying utility from the contributions of others. Although free riding is a fundamental and crucial issue in resource provisioning problems, studied from the very beginning of the formation human societies (and there is a vast related literature), there exist some special characteristics of p2p systems that provide novel dimensions to this ancient research question, discussed in the following.

2.2.2 Complicated cost modelling

Before designing mechanisms to discourage free riding, one should first identify the dominant cost factors associated with the participation of a peer into a p2p community. The fact that there are many different types of resources involved, and in different levels of system functionality, constitutes our cost modelling task a very challenging problem. So, we start by decomposing resource contribution into its different dimensions and analyze in detail the corresponding costs. From this detailed picture we will then make the necessary abstractions in order to manage to formulate a simple and tractable economic model.

Initial contribution

First of all, each peer “brings” an initial amount of resources joining into a p2p community, her initial contribution. This comprises of the capabilities of her PC/device (storage space, CPU power, and battery for mobile devices), her access bandwidth and, the amount of content stored in her hard drive. The costs associated with each one have different characteristics and not all of them are of significant value for all types of applications.

The cost for contributing an initial amount of content has to do with the effort of creating or acquiring the content outside the system (e.g. ripping CDs, taking and/or scanning photos, making a review etc.) and the corresponding storage space if required. These are ’pure’ costs in that peers do not have a direct benefit from sharing their content (unless social motives exist). But notice that in the case of file sharing, a peer could be useful to the community even if she doesn't offer any initial content at all since she can contribute to the efficient distribution (and further availability) of the existing one.

As far as hardware and ISP (access) costs are concerned, in most cases these are paid by a user for her own use and not for contributing them in a p2p community; unless both the value of participating is significant and suitable incentive mechanisms are in place so as to force a peer to make such an investment (when it is not justified by her own needs). However, the maximum capabilities of a user are often correlated with her personal cost (and utility) from participating in a p2p system, what we call her 'type'. For example, a DSL user has probably more utility (and less cost) from content sharing than a dial-up user. In any case, it is the unused amount of these resources that in principle p2p systems seek to exploit. When this is not possible, their contribution can incur additional costs during service provision (see below).

Peer Availability

Peer availability, is also a very important contribution of a peer in a p2p community. That is, how much time she stays on-line making available her resources (satisfying the service requests made throughout that time). Of course the cost for staying on-line is related to the cost of service provision (see below) and the average number of service requests received per unity of time. But ignoring this cost for the moment, there are also costs associated solely with the time a peer decides (or forced by an incentive mechanism) to stay on-line.

There are two different cases: If a peer's primary task is the p2p application, this cost corresponds to the time required for her service requests to be satisfied (mainly because of wasting time but also due to per time fees paid by dial-up and ISDN users in some cases). If the p2p application is run at the background, then in addition to the above cost (whose "weight" would be probably less in this case) there will be an additional (probably small) cost due to the amount of CPU power, memory and access bandwidth consumed by the application to run and thus deprived from the main task of the peer. Additionally, in case of illegal content sharing, peers run legal risks proportional to the amount of time they stay on-line (plus the amount of content they choose to share). Finally, some users feel insecure to leave their PC vulnerable to possible attacks and even if they have no cost to leave their PC switched-on (even when they don't use it at all) some would be reluctant to stay logged in a p2p system during all this time.

We believe that, especially for the case of file sharing applications, as available bandwidth increases and download times reduce significantly, peer availability will become a very important aspect for their success, and especially for those providing items of medium size, like music files, high definition photographs, etc. To see this, consider the limiting case where downloads take almost zero time to complete. Then if peers act rationally in terms of their availability, in equilibrium no files will be actually shared.

Service provision

During service provision the actual exchange of resources between peers is taking place. It entails mainly rivalrous but renewable resources. For example, in the case of file sharing the most important resource consumed for uploading one file is bandwidth. Bandwidth is finite and its consumption reduces the available amount to others, but it cannot be conserved. So, unless a peer is paying per volume charges to her ISP, the cost for offering upload capacity to other peers depends on the extent to which this capacity is not deprived from her own applications and service requests (i.e. downloads).

But most p2p systems seek to exploit exactly the under-utilized resources of Internet users. So, the development of the necessary technology, which will ensure that only the truly unneeded resources of peers will be offered for service provision, is very important for their success. This was the intention of the developers of SETI@home, which was meant to run in the background only when the screen-saver of a PC was activated (indicating that the user is not using it) and it was possibly only the installation and psychological costs (combined with the imperfect implementation) that kept many people from participating.

But this is not always possible. For example, in the case of ad-hoc networks service provision is always costly since it invlolves the consumption of the valuable battery. On the other hand, in the case of file sharing not only the cost of service provision could be minimized (Feldman et al. (2003) show that in many realistic cases including the ones we are considering, prioritizing TCP ACKs constitutes the cost of uploading negligible) but further increases the availability of resources through duplication (to some extent of course because of limited storage space)[18]. This was the main reason Clark Shirky criticized[19] the efforts of systems like MojoNation, which actually created a distributed market for service

[18]This is a unique characteristic of p2p file sharing, where consumption has positive externalities

[19]http://www.openp2p.com/pub/a/p2p/2000/12/01/shirky_freeloading.html

provision, proposing the employment of upfront incentives for the provisioning of the real valuable resources of peers such as access bandwidth, storage space and availability[20]. In our work we follow the same principle but proposing different types of incentives.

Core functionality

In pure p2p systems peers are also responsible for implementing all the core functionality for the system to operate (including the enforcement of the incentive mechanism, if available). Depending on the underlying requirements on computational and bandwidth resources this cost could become prohibitive, as for example in the early versions of Gnutella, which generated a large amount of signalling traffic in order to perform content search in a p2p fashion. In 2nd generation p2p systems this problem is addressed by exploiting the existence of peers with significant capabilities and resources and the suitable incentives to implement this functionality, the super-peers (see also Section 5.2.1.

Negative externalities

Finally, the amount of service requests satisfied by a peer (and the corresponding cost) is not only related to her initial contribution and availability but also on the request rate and size of the network. So, when service provision costs are significant, the overall service request rate in a p2p system, could be considered a public *bad*: the request rate of an agent increases the load on all other agents equally (if we assume that it is uniformly distributed among them). Moreover, when the average request rate is very high, we could also have congestion, which reduces the value of the service itself. So, there are situations where the externalities besides being positive, could be also negative. When this is true modelling and implementation of incentive mechanisms are further complicated (see Chapter 6).

2.2.3 Heterogeneity

Saroiu et al. (2002) performed a detailed measurement study of both Napster and Gnutella and, in addition to verifying the results of Adar and Huberman (2000) concerning free riding, observed that bandwidth, latency, availability and the degree of sharing vary across peers by between three and five orders of magnitude.

[20]http://www.openp2p.com/pub/a/p2p/2001/01/18/shirky_umbrellas.html

This is a natural consequence from the fact that Internet users have in general very diverse capabilities in terms of access bandwidth, cpu power, etc and different degree of familiarity with their PC (they range from university students and hackers to simple home users). For example, a user that has anyway a large amount of content stored on her hard drive for her own use, has a broadband connection with fixed monthly charges and leaves anyway her PC always on, has almost zero cost for uploading a file to a peer (the same holds for university students[21]). On the other hand, for a dial-up user with a noisy PC that pays per time and rarely uses the Internet or even her PC, the corresponding costs could be much higher.

So, an incentive mechanism dictating that all users should contribute the same amount of resources would generate disproportional costs to peers of different 'type' and if, additionally, consumption is related to contribution, peers with small capabilities would probably fail to acquire from the system the minimum value required to participate. Second generation p2p systems recognized the inability of all users to contribute equally and exploit the increased capabilities of certain peers, the super peers, which have enough motivation to voluntarily support the core functionality of the system and in many cases serve the majority of the requests.

This observation is consistent with the game theoretic analysis of Robin Mason included in [Antoniadis et al. (2003)], which shows (applying recent results by Mason and Valentinyi (2003)), that in a heterogeneous network, there is a unique equilibrium in which the those peers who value the shared resource the most contribute the most. This is the case when the degree of heterogeneity is above a certain level. Of course, when the heterogeneity is large it might be beneficial to apply different types of incentive to different groups of peers. Antoniadis et al. (2004b) address such issues (see also Section 4.6) and study when this is the case and how these should be designed so as the overall efficiency to be increased.

2.2.4 Size

P2p applications host already millions of users and since they seem to be very scalable, they could potentially reach a universal participation (according to the popularity of the services offered). It is not an exaggeration to say that they will probably become (if they are not already) the largest unregulated economies ever created by humans.

[21] and notably, such users are identified by measurements as the main contributors in such systems and thus the main target of the filings of the RIAA.

Interestingly, the role of size has conflicting effects on resource provision. One one hand, due to the positive externalities involved, the success of p2p systems is closely related to the number of their participants. And since each one of them contributes a relatively small amount of resources, it is only when a very large number of peers participate in a p2p network that it provides enough value. On the other hand, it is a standard result on economics that the power of the incentives for contributing to the provisioning of a public good diminishes as the size of the participants becomes larger. But when exclusions are possible, size could actually help. However, when negative externalities are present, the participation of a large number of peers should be accompanied with the sufficient individual resource provision and regulation of resource consumption. We will analyze these trade-offs concerning size in the next chapter.

2.2.5 Highly dynamic environment

In a client/server system, the resources made available by the central server are relatively stable, adjusting slowly over time. In contrast, the resources of a p2p system are spread across the network of peers, and hence are inherently more dynamic, potentially changing rapidly and unpredictably as peers join and leave the system.

So, a unique characteristic of p2p systems compared to classic public good theory is that in most cases the good is constructed *while (and because) peers are using it.* Thus, the provision and availability of resources depends strongly on the time peers decide to stay on-line. In most cases this time is used by peers to consume resources, unless they have no cost staying on-line and sharing their resources or an effective incentive mechanism is in place. So, it is this time, the one needed for consuming resources, which should be exploited in order to ensure that peers are contributing adequately. Otherwise, and because peers rarely interact with the same peers (due to the dynamic environment and very large size) accounting would be required, which is in many cases very difficult to be implemented (see Section 2.6).

Additionally, in a highly dynamic system, no peer can acquire a global view of the system. Many of its characteristics change constantly and the computation of the optimal configuration parameters of basic functionality (e.g. an incentive mechanism) cannot be computed off-line but should be dynamically adapted based on the current state of the system. This becomes even more challenging when no central entity is in place to gather all this information and perform the tuning of the protocols parameters but this should be implemented in a fully distributed fashion. For example, size, an often critical parameter could be only approximately estimated as most approaches for the aggregation of network information in pure p2p systems are based on random walk techniques (see [Kempe et al. (2003); Bawa et al. (2003)]).

2.2.6 Cheap pseudonyms

It is obvious that there is an inherent tradeoff between anonymity and accountability. Marx (1999) and Kling et al. (1999) present the benefits of anonymity and the costs due to reduced accountability. Interestingly, the consensus of a AAAS (American Association for the Advancement of Science) sponsored conference, as reported by Teich et al. (1999), was that the tradeoffs should sometimes be resolved in favor of anonymity in the Internet and reality validates this conclusion: most on-line games, auction sites, and interactive forums allow users to choose a pseudonym when they register. Even services that identify users based on their email addresses do not prevent identity changes, since users can easily acquire new e-mail addresses by the numerous free e-mail providers.

P2p file sharing applications also follow this regime and allow peers to create new pseudonyms at no (or very limited) cost. But although this is very important for encouraging participation, and protects peers from legal threats (to name a few of the benefits of anonymity), it poses additional challenges to any candidate accounting mechanism that would need to relate an amount of information to a peer's identity (see Section 2.6.2).

2.2.7 Unpredictable quality of service and hidden action

In a very large, rapidly changing, and anonymous system, it can be very difficult to assess the quality, reliability and security of resources and services offered by peers. This unpredictability of quality of service could result in an inefficient level of system usage by peers. A typical way to do this is through reputation mechanisms providing information for certain services (e.g. content items) or peers themselves. Of course, reputation mechanisms, being some sort of an accounting mechanism, suffer from all the aforementioned (and those discussed in the following) challenges.

But in some cases, not only the quality of service is unpredictable but additionally it is not possible (or easy) to exert the corresponding effort put of the providing peer. In other words, the specific action taken by a peer upon a service request, cannot be fully determined by the successful or not outcome of her action. For example, in ad-hoc networks, a forward request is not always successful (the packet does not reach its final destination) even if the next hop peer forwarded it to her next hop as promised. Similarly, in computation grids the outcome of a computation does not always depend deterministically on the corresponding computational power used.

This fact poses further challenges for the design of incentive mechanisms and effectiveness of a possible reputation system. The principal-agent (PA) model has been widely used in the economics literature to study such situations. Feldman et al. (2005) study the issue of hidden action in the case of ad-hoc networks, interestingly showing that in certain situations per-hop monitoring (often proposed in this literature for discovering malicious or uncooperative nodes) does not necessarily improves social welfare.

2.2.8 Distributed implementation

A trusted central server could support a p2p application in many different ways; it could serve as a bank to issue currencies, enforce strong identities, exclude misbehaving peers, collect entry fees, store and update reputation values, gather useful system statistics (e.g. the size of the network), and more[22]. Hence, complete decentralization, which as we discussed in the previous chapter is desirable for many reasons such as increased scalability, independence and ability to self-organize, easy deployment, legal threats, and more, comes at cost.

The first problem that needs to be tackled in a pure p2p system is service discovery, on which a significant part of the research work on p2p systems has focused the last years. Two are the basic approaches on forming overlay networks (in order to perform searching): structured and unstructured[23]. Structured overlays, also called DHTs (Distributed Hash Tables), assign to each node a unique identifier (hashing her IP address for example) and nodes create links between them deterministically, based on this identifier (either according to just the arithmetic relation of the identifiers, as in Chord [Stoica et al. (2003)] or enhanced with information from the physical layer to avoid inefficient paths, as in

[22]However, accounting, if required, has still to be performed by peers themselves since it is not realistic to assume a central entity that monitors all transactions; notice that one cannot rely on the software.

[23]In both cases, for bootstrapping reasons it is always required that some ‘known’ nodes are made somehow available (e.g. through a public web page) so as for a new member to be able to connect to the network and learn the addresses of some of its nodes.

Tapestry [Zhao et al. (2004)] and Kademlia [Maymounkov and Mazieres (2002)]. Similarly, unique identifiers are assigned to content items and each node is responsible for the content items whose identifier match her own (belong to a certain range computed based on the node's identifier). The most popular such approaches are Chord [Stoica et al. (2003)], Pastry [Rowstron and Druschel (2001)], Tapestry [Zhao et al. (2004)], CAN [Ratnasamy et al. (2000)], and Kademlia [Maymounkov and Mazieres (2002)]. On the other hand, in unstructured overlays nodes are free to choose their neighbours. And they do so, either probabilistically or trying to maximize their own benefit. Gnutella is such a system. In these systems the problem is which algorithm (e.g. breadth first search, depth first search, etc.) to use to forward the search queries in order to increase effectiveness while minimizing communication costs (see [Lua et al. (2005); D. Zeinalipour-Yazti and Gunopulos (2004)] for a survey on p2p overlay formation and information retrieval).

In unstructured overlays an interesting network formation game is to be played among the nodes since the overlay network is formed by self-interest neighbour selection (based on cost-benefit estimation of every possible link being created). The trade-off is that an agreement is useful (since it creates a direct connection with a node and thus avoids intermediaries) but incurs a cost to one (or both) of the parties. Moreover, full mesh networks are not necessary because two nodes can communicate indirectly. It has been observed [Jackson and Wolinsky (1996); Fabrikant et al. (2003); Corbo and Parkes (2005)] that in such scenarios the networks to be formed (if an equilibrium is reached) are not always efficient in terms of cost. The reason is that system-wide efficiency does not always coincide with personal objectives (for a recent survey of network creation see [DIgnazio and Giovannetti (2004)]). Moreover, in both structured and unstructured overlays, additional incentives should be provided to peers to run obediently the service discovery protocol, forward queries, maintain the routing information, etc. (see [Sun and Garcia-Molina (2004); Cooper and Garcia-Molina (2005)] and references therein).

The second critical functionality that should be implemented in a distributed way is the enforcement of the possible incentive mechanisms. Notably, enforcement always poses a certain cost on peers themselves (independently of the "p2p-ness" of the application): at least to check and report the outcomes of the transactions. But, of course, this task is even more challenging in case of a distributed implementation (see also Section 2.5.1).

2.2.9 Rationality vs. altruism

An interesting controversy appears in p2p file sharing systems: While measurements (and theory) indicate that free riding is the dominant strategy, file sharing applications are apparently very successful and are responsible for millions of files exchanged every day. And this is actually the reason why the RIAA makes tremendous efforts and spends a lot of money to close them down.

Many researchers have tried to give their own explanations for this controversy. For example, Krishnan et al. (2003) argue that file contribution is a rational motive for some peers so as to attract part of the demand and increase the download speed of the files they wish. Strahilevitz (2003) argues that the reason of the success of p2p systems is the fact that the p2p software hides free riders and gives the impression that everybody else shares, what is called "charismatic code". Finally, many believe that intangible factors encouraging altruistic behaviour play a decisive role towards this end.

An important characteristic that promotes altruistic behaviour in p2p systems is the fact that the contribution of a peer, although an irrational action, benefits all other peers. Interestingly, Andreoni (1995) showed through experiments that there is a behavioural assymetry between the 'warm-glow'[24] of doing something good (even if it is not rational) and "cold-prickle" of doing something bad. That is, people tend to act less rationally (being altruistic) when this has a positive effect on the whole community, but when their rational decision has a negative effect on others (e.g. congestion), they choose to maximize their own net benefit ignoring the negative externalities of their action[25].

But the most important reason lying behind this controversy is probably the high degree of heterogeneity between peers in existing systems (as indicated by measurements and experience). The role of heterogeneity is two-fold. First, peers with very high capabilities and low costs (e.g. college students) become seeds of content and are responsible for the majority of the transactions. On the other hand, it is probably exactly the existence of such powerful peers that motivates free riding by the rest of the peers, and thus explaining

[24] Warm-glow is the action of contributing to the common good (yours as well) even if it is not rational to do so. Note that, formally, altruism is slightly different than warm-glow. An altruistic peer acquires benefit from the action of contributing as such. So, even 'warm-glow' could be actually due to altruism is used as a term to describe the 'irrational' contribution of resources.

[25] To that respect, the "tragedy of the commons" [Hardin (1968)] often used to describe the social dilemmas associated with the provision of resources in a p2p system is not always the right example, since in the case of the commons the "tragedy" has to do with the consumption of a fixed amount of rivalrous resources, which means that the externalities in this case are negative while in many p2p systems they are positive.

the controversy. This is so, because, as suggested by Hindriks and Pancs (2001), altruism induces more free riding, since when there is high degree of altruism peers feel much less 'pivotal' for the provision of the public good and thus prefer to free ride.

However, despite the success of p2p file sharing systems, notice that content exchanged is constrained to be just the one initially contributed by the altruistic nodes and possibly a large degree of efficiency is sacrificed due to the lack of the appropriate incentive mechanisms. Moreover, the altruism would be much less in cases where costs are not negligible and the value is not so significant as in the case of copyright infringement! So, altruism can go some way to correcting the inefficiencies in a p2p system. Hence, additional incentives, carefully chosen and respecting the heterogeneity of peers should be devised. But policing the system by forcing service provision using strict reciprocity rules could be in many cases harmful and generate rational reactions otherwise not present. Thus, one should be careful not to harm the inherent community spirit in many p2p systems.

Summarizing, the issue of rationality is an important dimension when designing incentive mechanisms in this context. Varian (1995) argues that "the hyper-rational view of game theory may actually be an appropriate model for software agents, which have much better computational power than human beings" (a point also recognized by Rosenschein and Zlotkin (1994)). On the other hand, Christin et al. (2004) study the effect of using an alternative notion of rationality called *near-rationality* [Akerlof and Yellen (1985); Radner (1980)] in different network games such as network formation, TCP congestion control, and protection against security threats. In near rational equilibria, a player is satisfied to get close to (but does not necessarily achieve) her best response to the other players strategies. While the personal losses for a player are potentially very small, the equilibria derived often represent substantial departures from a prediction based on perfect Nash optimizing behavior. This is a very challenging avenue for future research. In our case, even small contributions may be enough due to the high degree of externalities involved.

Another interesting non-rational dimension is the social motivations inherent in many p2p applications, which can become dominant in many cases, as for example in content sharing [Lui et al. (2002); Gu and Jarvenpaa (2004)]. Feelings of membership and influence, identification of members (creation of their "image"), fulfilment of needs (means to provide support), and shared emotional connection (close relationships, community spirit) could play an important role in community based future p2p applications, in which users wouldn't have the incentive to be anonymous. For example, existing p2p applications have not managed to build long-lived and true communities around the sharing of resources (e.g. files), possibly due to the legal threats their users face, which we believe has probably had repercussions on efficiency. For instance, file sharing incentives would be stronger if a true community of members is built, which besides only sharing files, shared also their experiences and benefited from the opinions of experts, who, in turn, have proved their expertise by some group-specific internal process. For the formalization of such incentive mechanisms the concept of *bounded rationality* [Gigerenzer and Selten (2001)] could prove very useful, incorporating social norms, imitation, and other cultural tools as parts of rational strategies of agents.

Finally, social rationality is a very convenient notion of rationality, according to which agents make their decisions based on the overall system's net benefit [Hogg and Jennings (1997)]. But especially in the case of p2p systems it is not trivial in general to estimate the socially rational choices. Additionally, a system based on this notion of rationality would be susceptible to rational behaviour, which is always a viable strategy and it is very often unrealistic to assume that there are no peers who will choose to maximize their own net benefit, exploiting those that do not.

2.3 Defining efficiency

Resource allocation and provision in p2p systems is a very complicated problem. It includes the majority of economic concepts developed for networking over the last years. More specifically, service provision reduces the available bandwidth of providing peers and thus all the relevant work for bandwidth allocation like congestion pricing (see Courcoubetis and Weber (2003)) could be applied in this subproblem (Crowcroft et al. (2003) follow such an approach in the context of ad-hoc networks). Moreover, all the game theoretic approaches on network formation and routing are also highly relevant for the implementation of the core functionality of a p2p network in forming the overlay and implementing service discovery. Finally, our public good problem of content availability resembles, theoretically, to multicast cost sharing [Feigenbaum et al. (2001); Herzog et al. (1997)], with the differecnce that the amount of resources provided is not predefined[26], there is a single level (no tree structure), and the enforcement issues are much more complex (due to the lack of a central entity controlling membership).

[26]or among a finite set of possible alternatives as in the case of layered multicast.

As already explained, we have chosen to focus on the public good aspect of p2p systems, both theoretically and practically. Using the economics terminology, we have to solve a problem of private provision of a public good. Some of the main characteristics of p2p systems that one should take into account are the inappropriateness of prices as an incentive mechanism, the very large number of participants, the dynamic environment and symmetric roles of players, and the requirement for a distributed enforcement mechanism relying on the peers themselves for its implementation. Two are the most popular objectives for judging the success in the provision of a public good and determining the individual contribution of the participating users: 1) social welfare maximization and 2) fairness.

Social welfare is defined in economics as the sum of the utilities of the participants in a system consuming (or using) a certain amount of resources minus the cost for producing them. The *utility function* is an abstract construction that translates resources or services consumed to a satisfaction metric, different in principle for different people. One could think of the utility of an agent for a specific amount of resources made available, as the amount of money he is willing to pay for it or just the overall satisfaction received (comparatively).

The fact that this information is private and in many cases people may lose by revealing it truthfully is one of the motivations of the fairness (or equity) approach. For example, maximizing social welfare typically means that agents with higher utility should acquire larger proportions of a finite resource in a resource allocation problem. But in this case, people with low valuations would declare a higher value. The equity approach treats all agents the same in terms of utility and cost and applies a resource allocation (or defines contribution rules) subject to a specific fairness criterion, based on observable only characteristics of agents. In the simplest case, each agents acquires (or contributes) exactly the same amount of resources (or if demand differs amongst agents, an agent should get more than others only if the demand of the latter is fully satisfied —see also the concept of max-min fairness in data networks [Bertsekas and Gallager (1992)]).

The equity approach treats all agents the same in terms of utility and cost and applies a resource allocation subject to a specific fairness criterion, based on observable only characteristics of agents. In the simplest case, each agents acquires (or contributes) exactly the same amount of resources (or if demand differs amongst agents, an agent should get more than others only if the demand of the latter is fully satisfied —see also the concept of max-min fairness in data networks [Bertsekas and Gallager (1992)]). When there are observable characteristics of agents that differentiate them, the equity approach requires that agents with the same characteristics are allocated the same amount of resources (e.g. tax liability as a function of observable wealth). Interestingly, in the context of p2p streaming, Chu et al. (2004) propose the distribution of ‘wealth’ (upload capacity) amongst peers using taxation assuming that the capabilities of each peer are observable by the source.

But, in general, it is not at all obvious which of the observable characteristics are important and with which weight. So, following the equity approach one should first answer the question “equality of what?”[27] which is central in the philosophical debates concerning resource allocation and provision. Most importantly, it is easy to see the inefficiencies that may arise when the ‘equity’ approach is followed. Taking the example of multicast cost sharing, sharing equally the cost of a multicast transmission (e.g. 10 units) equally between two agents, with utilities 9 and 3 units respectively, would lead to none of the two agents receiving the transmission although their overall utility is greater than the corresponding cost (see also chapter 11 of [Courcoubetis and Weber (2003)]). We will thus try to design our system so as to maximize the social welfare of the system providing the suitable incentives to peers to reveal their private information truthfully (see next section) and when we refer to an efficient equilibrium or an efficient system we often mean a system which has achieved a maximum level of social welfare.

2.4 Maximizing social welfare

Social welfare maximization, by definition, requires the appropriate modelling of utility and cost. And this modelling is critical for the final outcome. But even if we manage to build a realistic economic model, there is one more important obstacle in order to achieve our objective function: information.

The reason is that the required information (the utility and cost functions of the agents) for maximizing social welfare is private and since it is often critical concerning the position of an agent the incentives to report it truthfully are not always there. In the following we present the tools provided from economics for implementing efficient incentive mechanisms under different assumptions concerning the available information. We discuss several existing modelling approaches and introduce our own in the next section.

[27] http://plato.stanford.edu/entries/equality/

2.4.1 The need for regulation

The great advantage of competitive markets for private goods is that no firm needs to know anything about the tastes of individual consumers. (And consumers do not have to know anything about the production technology of firms.) The only thing a firm has to know—other than its own production technology—is the price at which it can sell the good. The only things a consumer has to know—other than her own tastes—are the prices of the goods and services available. This is an enormous advantage for private markets. The first and second fundamental theorems of welfare economics (see e.g., Mas-Colell et al. (1995)) imply that the allocation resulting from perfect competition maximizes social welfare and no agent has an incentive to change her behaviour. That means that perfect competition is just as good as could be achieved by a benign social planner. But the big advantage of private markets over central planning is that a benign social planner would have to know everyone's tastes, and all firms' technologies, in order to implement an efficient allocation. Under perfect competition, there is no need for anyone to learn this information, and therefore no problem in getting people to reveal their preferences or firms to reveal their technologies.

This is however not the case when externalities (positive or negative) are present as in the case of p2p resource provisioning. For example, a rational peer will not take into account the benefit acquired from all other peers when deciding on the amount of resources made available (typically contributing less than the optimal amount) nor will tune her consumption level based on the negative externalities created due to congestion or uploading costs. The key difference between markets for private goods, and the pseudo-markets for public goods, is that regulation is required in the case of public goods because more information is needed to implement an efficient allocation and externalities should be internalized through suitable incentive mechanisms.

In our case, the regulator is the p2p software designer who should encode the proposed incentive mechanism in a specific protocol enforcing the rules dictated by the mechanism and which will be part of the software that peers should run in order to join the community. The parameters of this protocol should be also automatically configured according to certain system characteristics (e.g. its size) since it will not be easy for the software designer to intervene and correct the values of these parameters. But since p2p software cannot be considered trusted, these rules should be seen as "proposed" strategies, which should be designed in such a way that at least rational peers don't have a strong incentive to alter them and malicious ones cannot harm the system (see also Section 2.5.1 below).

2.4.2 Complete Information

In the ideal case, full information about the entire system is available to calculate the optimal form of regulation to correct peers' behaviour. Then the standard economic approach is to apply *Lindahl prices*—a tax or subsidy that is specific to each peer to internalize the externalities caused by the peer. Alternatively, one could apply specific system rules enforcing the corresponding levels of consumption and contribution that would be achieved through the prices (see Section 4.5.1).

The standard criticism of Lindahl prices is that they require a large amount of information to implement. It is very optimistic to suppose that any real-life regulator or planner would have this amount of information. For some p2p systems (likely to be small-scale with a stable set of similar peers), this could be possible. For other systems, with a large number of rapidly changing, heterogeneous peers, only limited and local information can be used to regulate their peers. An additional disadvantage of this approach is the fact that the optimal prices (or rules) are in general personalized (their values are different for each different peer)[28]. In addition to the informational obstacles, personalization poses serious implementation ones, which constitute such an approach impractical. We use it, however, as a benchmark for assessing the efficiency achieved by different sub-optimal solutions that could be devised.

2.4.3 Incomplete information

Thus, a major challenge for the design of a suitable incentive mechanism is to acquire the necessary information for achieving an efficient outcome. This may require peers to reveal information that may raise their costs. Hence, any approach should be *incentive compatible*: for those aspects that cannot be directly observed, peers must be given incentives to behave correctly, or to report information truthfully. In the ideal case we should be able to design mechanisms that satisfy the requirement in the economics of mechanism design to be detail-free—known popularly as the 'Wilson doctrine'. The idea is that mechanisms, ideally,

[28] For certain economic models, however, and any given degree of heterogeneity among peers, as we show in Section4.3, the Lindahl prices, but not the corresponding rules, can be approximated by a uniform price in a sufficiently large network.

should be universal (i.e., independent of the details of the distribution of peers' types of valuations) and anonymous (i.e., the identity of a peer is irrelevant). The motivation for this requirement is that any mechanism that depends on fine detail will be difficult to implement in practice, when the mechanism designer (i.e., system manager) has little or no relevant information.

Mechanism design

So, the informational burden on the planner with Lindahl prices or rules suggests an alternative approach in which peers are given incentives to reveal truthfully their net utilities. This a fundamental problem in resource and cost allocation problems since in most cases agents have a strong incentive to misreport their value (higher than the real one for resource allocation and lower in cost allocation) so as to increase their personal net benefit as computed by an efficiency maximizing social choice function.

In the context of auction design for single items, Vickrey (1961) devised the celebrated second-price sealed-bid auction or simply the Vickrey auction: Bidders simultaneously submit sealed bids for the item on sale. The highest bidder wins the item, but (unlike standard sealed-bid auctions) the winner pays the second-highest bid. For example, if the winning bid is 10 and the second-highest bid is 7, the winner pays 7. Under this mechanism, a bidder can never affect the price it pays (if he wins) and thus no bidder has the incentive to misrepresent his value. The bids determine only whether the corresponding bidder wins or not, and only by bidding his true value can he be sure to win exactly when he is willing to pay the price. So, under this mechanism, which was also extended for the case of multiple identical items in [Vickrey (1961)], it is a dominant strategy for bidders to report their values truthfully and the resulting outcomes are efficient (the item is bought by the bidder that values it the most). Interestingly, Green and Laffont (1979) and Holmstrom (1979) show that, under weak assumptions, the VCG mechanism is the unique direct reporting mechanism with dominant strategies, efficient outcomes, and zero payments by losing bidders. Since Vickrey's original contribution, his auction design has been melded with the Clarke (1971) and Groves (1973) design for public good problems. Various names are used to describe the extended Vickrey mechanism. We will refer to it as the Vickrey-Clarke-Groves or VCG mechanism.

As with Vickrey's original design, the VCG mechanism computes efficient outcomes and charges agents the social opportunity cost because of their participation. That is, each agent pays the benefit that is deprived from the other agents because of her existence, the social opportunity cost. Formally, given the complete report of agents' utility functions, $u_i(x)$, $i = 1, \ldots, n$, where x being an allocation of the good to the agents, the price agent i has to pay is given by

$$p_i = \sum_{j \neq i} u_j(x^*_{-i}) - \sum_{j \neq i} u_j(x^*) \tag{2.1}$$

where $x^* = \arg\max_x \sum_{i=1}^n u_i(x)$ is the optimal allocation and x^*_{-i} is the optimal allocation when peer i doesn't participate[29]. It is straightforward to see that under this scheme, each agent has a strict incentive to report its net utility truthfully, since the final outcome is independent from her own declaration; and with this information, the planner is able to implement the efficient allocation. But despite its very attractive and unique characteristics the VCG mechanism has some very serious weaknesses [Ausubel and Milgrom (2005)], which we discuss in the next section.

The VCG mechanism is a special case of a "mechanism", which as defined by the theory of mechanism design (MD) provides a mapping of agents valuations to an outcome (an allocation). More specifically,

- the agents, knowing the mechanism, report their valuations wishing to maximize their own net benefit (strategic players)

- the social planner decides on the allocation of resources and payments made by or paid to the agents

In other words, mechanism design studies the definition of the suitable game so as to achieve a desired outcome, while game theory typically studies the outcome of specific games. This is the reason it is often called "inverse game theory". The major issues in this context are information and complexity. A mechanism is called *strategyproof* when it is always the optimal (dominant) strategy of agents to truthfully report their valuations. Classical mechanism design (MD) assumes a central entity which is responsible for all the communication with the agents, and computations of the payments, ignoring the corresponding computational complexity.

[29]Notice that in a cost allocation problem this payment equals to the amount of cost not covered from the declarations of the rest of the agents (see Section 3.3.2) —which is again independent from an agent's own declaration. This is often called the *marginal cost mechanism*.

Algorithmic mechanism design (AMD), introduced by Nisan and Ronen (1999), addresses the issue of complexity, proposing a formal model of centralized computation that combines incentive compatibility and computational tractability, while distributed algorithmic mechanism design (DAMD), introduced in Feigenbaum et al. (2001) (see also Feigenbaum and Shenker (2002) for an overview), seeks solutions that are both fully distributed and computationally tractable.

A rich body of literature has been dedicated to the application of MD to software agents and Internet applications. Many of them are implementations of VCG mechanisms in a distributed environment. For example, Feigenbaum et al. (2002) propose a strategyproof mechanism (of the VCG family) for shortest paths in BGP routing that induces truthful revelation of transit costs. And Feigenbaum et al. (2001) proposes a distributed implementation of the VCG mechanism for cost allocation in the context of multicast cost sharing. In the context of p2p systems, Zhong et al. (2003b) designs a payment mechanism for packet forwarding in mobile ad-hoc networks (ressembling to mechanism design), in which the payment to intermediate nodes who carry the packets is based on reports sent by the nodes to a central center. The payments are designed in a way that encourages truthful revelation about the service that has been provided to and consumed by nodes.

The second-best allocation

The VCG mechanism is universal, anonymous, and incentive compatible. It has, however, a very important weakness: the transfers made to peers (to ensure that they reveal truthfully their utilities) can result in a very large deficit for the network. That is, it is not *budget-balanced*. For example, in the Vickrey auction, if the second-highest bid is zero the winner will pay nothing and thus the auctioneer will not manage to cover the cost of the item sold.

Myerson and Satterthwaite (1983) show that this is a generic problem (under any possible strategyproof mechanism) and even in the average case. So, it is impossible to achieve full efficiency when agents have private information, participation is voluntary, and budget balance must be achieved. Of course, there is always a maximimum level of efficiency that could be achieved by an appropriate mechanism (see next chapter for a formal characterization of second-best in the context of our proposed economic model). This level of efficiency is called “second-best” since it is the best a system designer can do under incomplete information (when the exact types of the participants are not known).

However, what is considered known in most cases for the calculation of the second best is the distribution(s) from which the values of peer types are drawn. Note that the dependence on the distributions is very difficult to be removed and this could be achieved only in situations where there is some sort of additional information available. For example, in the case of auctions a system designer could acquire some ex-post knowledge of the distribution of user types from the bids of the participants, as is the case in the work of Segal (2003).

But even with this assumption (the knowledge of the distributions), the computation of second-best mechanisms is very complicated. Fortunately, for the problem of private provisioning of public goods (which we believe captures nicely the problem of content availability in p2p file sharing) there are recent results that suggest that when the number of agents is large (which is consistent with the situation in p2p networks) we can achieve an asymptotically optimal equilibrium (within $O(1/n)$ from this achieved by second best) using a very simple mechanism. We make this more concrete in the next chapter.

2.5 Modelling

Before designing and evaluating an incentive mechanism in terms of economic efficiency one should first formulate the corresponding economic model, define the utility and cost functions, describe the agents behaviour in time, etc. Obviously, the decisions taken at this stage affect significantly the final outcome.

However, even if it was possible to capture all aspects of a system, the corresponding detailed model would be impossible to analyze analytically or even numerically. So, the goal of an economic model is to capture the most important utility and cost generators, compromising practical details for the sake of simplicity and tractability. Besides its critical role for devising efficient incentives mechanisms, economic modelling could also be used to assess the efficiency of existing systems and applications and quantify the related trade-offs. In the following we summarize the related work in this area.

2.5.1 Prices vs. rules

A standard approach for co-ordinating and regulating behaviour is to use prices. The price approach is often advocated by economists, and in many situations has a number of advantages over alternative methods. Prices could be used to incentivize contribution (e.g. a payment to peers for resources offered) or control consumption (e.g. a charge for resources consumed). Alternatively, one could create a market for resources whose price per unit could be either fixed (equating this way consumption with contribution —a barter economy) or freely determined by providing peers themselves (a free market), as in the case of MojoNation.

However, the use of real money is impractical in many p2p systems due to e.g., the anonymity of peers and the sheer volume of low value transactions but is also against the community spirit that is dominant in many existing applications. Moreover, as Odlyzko (2001) suggests[30], the mental burden involved for making numerous decisions for purchasing or not very cheap resources is in most cases unacceptable by users who are in favor of flat pricing schemes even if the total cost is greater. This holds also to some extend even when virtual currencies are used, which are often proposed as a possible incentive mechanism in the context of p2p systems.

Not surprisingly, price-based approaches have been proposed mainly in the context of ad-hoc networks where the involved resources are truly scarce (e.g. [Crowcroft et al. (2003); Jakobsson et al. (2003); Zhong et al. (2003b)] among many others). In p2p file sharing systems *rules*, non-price-based methods of regulating peers' behaviour may be needed. Rules will constrain what peers are able to do (e.g., restricting the number of downloads to be no greater than the number of uploads) or set participation constraints (e.g. requiring a certain level of sharing for a peer to be able to participate); nevertheless, peers are free to choose aspects of their behaviour (e.g., the number of uploads that they perform). In a price-based system, peers choose their resource contribution and consumption subject to a budget set: the set of feasible contribution and consumption choices, given their budget and the prevailing prices. In a rule-based system, peers choose their resource contribution and consumption subject to the rules: the set of feasible contribution and consumption choices. In other words, system rules define the constraints imposed on peers by the software. Then peers will choose actions within those constraints according to their local policy which encapsulates their approach to maximising their utility. Notice that the use of virtual currencies to implement a barter economy actually coincides with a system rule dictating that contribution of peers should be always greater or equal to their consumption of resources. So, there are cases where we can have exactly the same outcome in terms of resource allocation and provisioning with both approaches.

[30]see also http://www.openp2p.com/pub/a/p2p/2000/12/19/micropayments.html

We shall examine criteria to assess the performance of rules; and assess the extent to which they can achieve efficient system usage and can substitute for a price-based approach. We anticipate that rules are more appropriate for simple symmetric p2p systems of relatively low cost resources, as the majority of existing p2p applications are. Prices, or some other form of monetary compensation, may be necessary in situations where both high and low value services are exchanged and peers do not have balanced contributions (some peers may be more on the provider side while other peers may be more on the consumer side).

The choice of these system rules is based on two separate considerations. Firstly, the rules must cause the co-operative behaviour and achieve efficiency that is our objective (In general, rules could either exactly specify efficient actions or could be such that the utility-maximising action of each peer is efficient). But, secondly, the rules must be enforceable given the control that is possible within the system being described. Certainly not all formal rules can necessarily be implemented in any particular system. An all-powerful system manager could, perhaps, impose any system rule on peers and ensure that it was obeyed. But, in any practical system, there will be state that is private to peers and not visible to any external observer. Then a manager can only base a system rule on variables that it can observe, and enforce rules only on the basis of its power to control or punish behaviour. In a system with no central manager, this observation and enforcement *must be carried out by the other peers in the system.*

Thus, many system rules will require additional rules to be enforced. For example, we may need peers to aggregate history about each other and then to punish (e.g. by exclusion) the constraint-breaking peers. In other words, a social planner does not have actually the power to 'regulate' a p2p system. What she can do is encode in the default client software certain rules which should be designed that rational (or at least 'obedient') peers do not have an incentive to alter them (that is, they should be incentive compatible). Moreover, given that a number of peers uses the "proposed" strategies, anyone who uses a strategy outside this set should be worse off (typically because they will be discovered and punished). But note that in any case enforcement has a cost. And although there is experimental evidence (see Fehr and Gachter (2002)) that people are willing to pay some cost in order to punish free riders, an enforcement mechanism shouldn't rely on this and minimize such costs but most importantly, it shouldn't allow peers to take advantage of it (through any feasible means) for their own benefit.

2.5.2 Economic Modelling (aiming at efficiency)

Golle et al. (2001) made a first effort to model the utilities and costs associated with the participation in a p2p file sharing system. Using game theoretic analysis, they propose the use of micro-payments for achieving the desirable equilibria. More specifically, in the model proposed, each peer has two orthogonal actions available in each time period with 3 different levels: 1) Sharing: s_0 (no sharing), s_1 (moderate sharing) or s_2 (sharing all files); 2) Downloading: d_0 (no downloads), d_1 (moderate downloads) or d_2 (heavy downloads). The utility of each peer depends on the amount of downloads, network variety, disk space used, bandwidth used, altruism, and financial transfers. According to the proposed mechanism, at the end of each period, each user is charged an amount $C = f(d - u)$. The function f maps the difference between downloads d and uploads u to a financial transfer from the agents to the system, which could correspond or not to real money. The model is extended to include rewards for sharing files rather than actually uploading them. In each different case and under different assumptions the equilibrium strategy (s_i, d_k) is intuitively deduced. The main weakness of this approach is that it assumes that there is a central entity that has full information of all the transactions between peers and which handles all the financial transfers, which is impractical in a p2p context.

Krishnan et al. (2003) give focus on congestion effects and model the net benefit of peers (they assume it is the same for everyone) as the value they derive from each unit of content that is potentially downloadable $v(k)$ (where the amount of content depends on the number of sharers k) and the likelihood that a user can download content from the network $f(n, k)$, which is a function of congestion in the network, which in turn is increasing in the number of network users n and decreasing in the number of sharers k. The strategy of peers is binary: either to share or not. Sharing incurs a lump-sum cost c, again the same for all peers. The most interesting case is the intermediate range of cost c where despite free-riding, there are always k^* users willing to share their content, in spite of the existence of free-riders. By sharing they are able to reduce the congestion on other nodes that share, making it easier for them to download their desired content[31]. However, this model is rather simple (since it assumes all peers have the same preferences and they have a binary sharing strategy) and it focuses on congestion effects which we believe that

[31]It is assumed that a node can deduce whether a potential downloader is a file sharer, and contributors are always ranked ahead of non-contributors in a queue (as an additional incentive for peers to become contributors since there will be a probability p that they will not obtain their desired content).

have a second-order effect in the problem of resource provision in p2p file sharing systems.

Feldman et al. (2004) assume that each peer i is characterized by a generosity level, her type t_i and decides whether to contribute or free-ride based on how her generosity compares to the current contribution cost in the system, which is the inverse of the total percentage of contributors, x. They show how the levels of societal generosity affect the resulting efficiency, when social welfare is a concave function of x. Under the assumption that free-riding is observable and could be punished through the deterioration of a free rider's benefit by a fraction of $(1-p)$ (e.g. service is refused with probability p), they model a dynamic system with arrivals, departures and whitewashers[32] in order to compute the levels of identity cost and punishments that should be imposed under different system parameters in order to reach an efficient equilibrium. Notice that this work (as this of Krishnan et al. (2003)) doesn't address the issue of efficient resource provision since it just distinguishes between cooperators and defectors assuming that all cooperators share an adequate amount of resources (social welfare depends only on the number of cooperators).

Buragohain et al. (2003) incorporate the issue of resource provision in their model. They define the cost of peer i to be the amount of space (used for storing shared files) or/and bandwidth allocated, denoted by d_i, and her utility the sum of the contributions of all the other peers multiplied by a factor b_{ij} expressing how valuable the content made by a certain peer j is for peer i. That is, $u_i = \sum_j b_{ij} d_j$. The authors assume that the system can enforce a certain level of reciprocity achieving a probability $p(d_i)$ with which a content request from peer i is accepted, similarly with Feldman et al. (2004). So the larger the contribution d_i the larger the percentage of the maximum possible value of u_i is acquired. Their main goal is to compute the nash equilbrium achieved given the above incentive mechanism for a given function $p(.)$. For the case of two players they show that in equilibrium efficiency is maximized. For the general case it is just shown (analytically for the homogeneous case and using simulations in the heterogeneous case) that the system is lead to a desirable nash equilibrium (different from the one where $d_i = 0 \forall i$) but not necessarily to the most efficient one.

Our basic public good model for content availability, presented in detail in Chapter 4, is similar to the model of Buragohain et al. (2003). We, however, follow a regulation based approach based on the theory of private provision of public goods, in order to design system rules dictating an efficient level of contribution and/or consumption of resources.

[32]We discuss the whitewashing attack in Section 2.6.1.

We believe that there are certain system parameters (such as the number of files shared per peer), which if tuned appropriately would lead the system to a better equilibrium. Of course there are in this case also challenging issues concerning enforcement which we address separately (see Chapter 5).

2.5.3 Evolutionary games (aiming at cooperation)

A large body of the p2p economics literature follows a different approach and models resource provision and allocation in p2p systems as a classical prisoner's dilemma game (e.g. [Ranganathan et al. (2003); Feldman et al. (2004)]), without taking into account the possibly different utilities and costs from resource consumption and contribution of different peers. The goal of these approaches is to encourage cooperation in a more abstract sense.

In the simple two-player prisoner's dilemma each player has two possible strategies: either to Cooperate or to Defect. The fact that in a (Cooperate, Defect) situation the cooperator looses more than if both had defected and the defector has larger payoff than if both had cooperated, makes the nash equilibrium of the one shot game to be that both players defect even though the benefit of both cooperating is greater.

For the repeated version of the game, Axelrod (1984) has shown that the generous tit for tat strategy, which always cooperates on the first move (it is optimistic) and reciprocates what the other player did on the previous move thereafter, outperforms all other strategies. More formally, tit for tat is an 'evolutionary stable' strategy. Evolutionary game theory studies the equilibria of games where agents change their strategies over time (as they interact with other agents with possibly other strategies) trying to maximize their own net benefit (e.g. mutating to strategies that achieve higher benefits). Evolutionary stable are the strategies that prevail throughout this process.

However, the tit for tat strategy requires the players to meet often. But in the types of p2p systems that we are considering the probability to encounter the same peers is small (see Section 2.2.5). Hence, in this case, reciprocity should be enforced *indirectly* in order to be effective: a peer should reward or punish another peer according to her past behaviour towards *other* peers. The notion of indirect reciprocity in human societies was first introduced by Alexander (1987). Nowak and Sigmund (1998) introduced the 'image score' of an agent (which expresses the number of her total cooperations minus her total defections) so as to enable reward and punishment of agents according to their past behaviour. The 'Image strategy' (which cooperates with peers that have an image score above a threshold) can defeat defectors even in games with large populations and few repeated transactions.

But the ability of peers to easily change their identities in most p2p systems constitutes the notion of the image score useless, since peers with a low image score can whitewash their past creating a new identity. Hence, either an effective distributed accounting mechanism, which can provide reliable information for the past behaviour of a peer (addressing the whitewashing attack) should be in place or the nature of the service should allow for a memory-less implementation of (variants of) the tit-for-tat strategy. We discuss such issues in detail in the next section.

Moreover, and even more importantly as far as modelling is concerned, the notion of cooperation and defection should be defined. How the 'image score' of a peer should be computed? When she must characterized as cooperative and when as defector? In the work of Nowak and Sigmund (1998) agents are assumed to meet regularly and decide whether to cooperate or to defect. But this is not exactly the case in p2p systems. Treating a successful upload as a cooperative action and an unsuccessful one as a defection, fails to capture the notion of content availability. For example, a peer that cooperates (uploads) very rarely sharing a very small number of files but always successfully would have the same score with someone that uploads very often and shares a much larger number of files. Feldman et al. (2004) have used the term 'untraceable' defects to describe this situation and proposed the ratio of downloads/uploads as a more suitable image score. Then setting the image score threshold to one would result in equating consumption and contribution in terms of actual uploads/downloads, which is also the goal of many token-based approaches (see Section 2.6.3). Ignoring the challenging enforcement issues, they are discussed in the next section, such a mechanism would achieve a certain level of content availability which will in general depend on demand and on the equilibrium of the sharing strategies of the peers[33]. But in our work we wish to exploit possibly available information (i.e. the distribution of the peers' preferences and their expected number) and regulate peers' behaviour in order to increase the efficiency of the system in terms of content availability,

[33]There is actually a 'congestion game' [Rosenthal (1973)] to be played amongst peers to that respect because when their sharing decisions are guided by strategies related to the corresponding enforcement mechanism (e.g. trying to temporarily maximize their upload rate so as to minimize the time required to stay on-line sharing files for satisfying their demand —see Section 5.4.7), popular files will be 'congested' since this would lead some peers to change their sharing strategies, and so on and it is not clear what the equilibrium of this game will be.

according to an underlying economic model. In Chapter 5, we propose a practical memoryless approach for tuning certain important system parameters towards this goal, and discuss in more detail how our approach is compared with the 'downloads=uploads' mechanism (see Section 5.4.7).

Finally, the notion of reward and punishment should be also defined. A standard punishment strategy is to deny access to defectors (e.g. peers with low image score) as assumed in the work of Feldman et al. (2004) discussed above. However, less strict incentive mechanisms have also been proposed, which also seek to achieve a certain level of *cooperation* rather than an economically efficient equilibrium. They propose the use of qualitative rewards and punishments based on one's past contribution (what Buragohain et al. (2003) expressed as a probability to receive service in their model). For example, the view of the network is filtered (the number and/or type peers that a peers is allowed to interact with) as a function of a peer's contribution (e.g. [Papaioannou and Stamoulis (2004); Ranganathan et al. (2003)]) or priority is given in case of contention to the most generous peers (as for example the reputation mechanism employed in Kazaa).

2.5.4 Free-market models

Some researchers see p2p systems as distributed markets for buying and selling resources, where prices (either real money in the form of micro-payments or tokens) are freely determined by providing peers (according to demand, congestion, and other possible criteria). But the failure of MojoNation and the theory of externalities suggest that such approaches could be suitable only when the costs for service provision are significant and the corresponding resources rivalrous and/or when the public good aspect is limited.

Hence, free markets could be suitable for applications which involve rivarlous or consumable resources such as ad-hoc networks [Crowcroft et al. (2003); Jakobsson et al. (2003)] and grid computing [Krauter et al. (2001)]. See also the work of Figueiredo et al. (2005) for providing incentives for relaying packets in a p2p anonymity system (which has some similarities with the case of packet forwarding in ad-hoc networks). In the context of file sharing, there are some approaches that focus on service provision costs and propose distributed implementations of a free market [Vishnumurthy et al. (2003); Hausheer and Stiller (2005)]. We also propose a market model for p2p content distribution in the case where costly items are to be introduced in a p2p community.

2.6 Enforcement mechanisms

2.6.1 Accounting

We first discuss the notion of accounting and under which assumptions it is feasible to account for peers past transactions in order to provide the required information to the underlying incentive mechanism. The most easy to account transaction is service provision (e.g. a file download). Other contribution parameters, such as availability, require in general regular auditing and are almost infeasible to be measured in a distributed system[34]. Existing accounting mechanisms either compute a balance between contribution and consumption of resources in terms of service provision or provide a qualitative characterization of a peer (i.e. a reputation value) expressing her 'type' in terms of generosity, truthfulness, etc. A very challenging problem in this context is ensuring the reliability of this information (see next section).

Local accounts

The simplest accounting mechanism is one that trusts each peer (actually her client software running on her machine) to keep a record of her own transactions, possibly aggregated (e.g. a counter of uploads minus downloads). In the best case, just this information could be used to enforce the correct behaviour of the corresponding peer. This could be possible under the existence of tamper-proof hardware at each node. As assumed for example in the approach of Buttyan and Hubaux (2003) for providing incentives for forwarding in ad-hoc networks, where for each packet transmitted by the source node, a credit counter is decreased by an integer amount equal to the number of intermediate hops from the source to the destination. But, of course, in most cases this mechanism is insecure since peers have the incentive to alter their records for their own benefit by altering their application for example (as happened in the case of Kazaa). In certain cases, this attack could be addressed by requiring receipts signed by a member of the community in order for a record presented to be trustworthy. But then the issue that arises is which are the valid members of the community?

Nevertheless, in the worst case, these records could be used from the peer herself to

[34] It is one of the contributions of this dissertation the idea that one could directly control the availability of peers in a system (instead of accounting for it) by controlling the time they need for consuming the required resources (see Chapter 5).

account for the contribution and consumption of the rest of the peers. But the very small probability to interact with the same peer in the enormous and dynamic p2p systems constitutes this accounting method inadequate.

Public accounts

An approach to address the limitations of local accounts proposed in IST project MMAPPS [MMAPPS-Consortium (2004)], is to require peers to keep complete public accounts of their transactions. Peers could regularly audit the accounts of other peers to discover any false accounts.

Such an approach is relatively easy to implement in the case of preservation systems, where peers should constantly provide service [Ngan et al. (2003)]. But in the case of file sharing, where millions of service provisions take place in short time scales there should be some way to reduce the information stored. For example, a mechanism by which history can eventually be rolled up into a carried-forward balance. This could be achieved by asking a group of peers to sign any proposed balance. Once a balance is adequately attested (say, by a minimum number of other peers) it can be added to the account and the history deleted.

A different way to implement a public account is to have a third party hold the account. For example, we can designate one (or a redundant group of) other peer(s) via the peer addressing mechanism, if a DHT is used. Each signed receipt is sent by the provider to his account holders and then from them on to the account holders for the consumer. The account holders need to hold these records only for a short time at most and then they can roll them up into the aggregated balance. A similar mechanism is proposed in Vishnumurthy et al. (2003), where the account of a peer corresponds to an amount of "karma" acquired through service provisions whose value (in terms of karma) is defined through auction mechanisms. But even if a central entity existed for storing the accounting records it would still rely on peers to report the outcomes of the transactions, which as we discuss in the following makes the system susceptible to serious attacks.

Reputation

Reputation could be defined [Wilson (1985)] as a 'summary' of an agent's past behaviour, expressing a probability that future transactions with this agent will have a certain outcome. In other words, it is meant to provide an indication of her ‘type’: Is she honest or not? Is she altruistic or egotistic? What is the quality of service offered? Is she reliable? Or even more subjective characteristics such as her taste on different aspects.

In real life, businesses, organizations, humans depend on their reputation in order to attract customers, members, friends. In the Internet, reputation mechanisms have recently attracted significant attention as a mechanism for building trust and encouraging cooperation in on-line trading communities, such as eBay, or as a powerful tool to guide consumer’s choices through recommendation systems, such as eOpinions (see [Resnick et al. (2000)] for a short introduction and [Dellarocas (2005)] for the state of the art). The positive role of reputation in these systems, have motivated many researchers to study the application of reputation mechanisms for the increase of the performance and quality of service offered by p2p systems. In the following we present a short overview of the concept of reputation and discuss to what extend it could be helpful in our context mentioning some important results.

The most fundamental attribute of a reputation system is the action of *rating*. Rating is the feedback that agents provide to the system after a transaction has taken place, evaluating its outcome in order for the reputation values of all agents to be computed and information concerning their type to be elicited. Depending on the application, feedback could be objective (a successful service provision or not in a p2p system) or subjective (how useful a review of a certain peer was in a recommendation system) or both, as in the case of eBay, where buyers do not rate sellers only for sending an item or not, but also for the quality of the packaging, the speed of delivery, etc. In general, objective ratings are binary (success/failure) while subjective usually take more than two values depending on the degree of subjectivity.

The next step is to *aggregate* all ratings into a single reputation value. Taking an average of all ratings is the most common aggregation method. An additional parameter of the aggregation algorithm could be time, since reducing the reputation value as time elapses provides the incentive to agents to keep providing efforts in order to sustain their reputation. For example, Josang et al. (2003) propose the *Beta* aggregation, where each peer’s reputation equals the fraction of the weighted number of successful service provisions over the weighted total number of her service provisions, with the weight being a negative exponential function of the elapsed time. Dellarocas (2002) proposes the use of Buyes’ rule which under certain assumptions about the possible different types of peers, expresses the a posteriori belief that she belongs to a certain type given the history of her transactions.

Note that such mechanisms value equally all ratings. But, in general, users could have different types as far as their credibility or quality as raters is concerned. In other words, their ratings could be untruthful (in the case they are objective) or, in case of subjective feedback, of low quality for certain peers. And indeed, in Amazon users trust more the raters which have given them good advice in the past (or have a good reputation as raters[35]) and as a result the final reputation of each item is different for different agents according to these valuations.

So, the standard way to address the issue of credibility is to put different weights on ratings according to the personal or overall evaluation of a peer as a rater. For example, in the context of p2p file sharing, Kamvar et al. (2003a) propose the eigentrust algorithm, which uses the personal experience of peers (concerning the relative quality of a peer as a service provider) as a weight for the calculation the page rank [Ng et al. (2001)] on the recommendation graph. As a result, the final global reputation value is given by a weighted sum of ratings providing a different set of reputation values for each agent according to her personal ratings. But this approach assumes that the quality of a rating is correlated with the reputation of the rater as a service provider.

There are researchers that have identified the need to distinguish between these two characteristics of a peer and have recently proposed the introduction of a separate reputation metric to account for the credibility of the peers independently from their performance in terms of service provision. More specifically, Papaioannou and Stamoulis (2005) propose to punish peers that disagree for the outcome of a transaction (since with high probability one of them is lying) proportionally to their reputation. Buchegger and LeBoudec (2004) propose a distributed system where each individual peer computes a reputation value and put less value on conflicting reports. However, the problems due to the existence of untrustworthy peers is multiplied because of cheap pseudonyms and the so-called 'sybil attack' which we discuss in the following.

In general, reputation could be employed to reveal different types of information according to the specific application. In the case of p2p systems, many such types of information are important: the reliability of peers, their credibility as raters, and their level of contribution. But these are in principle independent variables of a peer's behaviour: a high reputation value for reliability does not give enough evidence of whether a peer is a free

[35] A large percentage of users "has found their review useful".

rider or not in terms of her overall contribution of resources compared to consumption. So it is important for a system designer to understand the exact information revealed by a reputation mechanism. For example, reputation for reliability is often confused with reputation for contribution or altruism in p2p file sharing systems.

In this dissertation we are particularly interested in addressing the free riding issue and providing the suitable incentives to peers for sharing their content and resources. In this context, feedback is objective. This means that the outcome of a transaction (e.g. a file having been downloaded or not) does not depend on the perception of the participating peers. The reason a peer would wish to misreport his valuation of a service provision is either because she is malicious or acting rationally. Rational behaviour arises due to the necessary incentive mechanisms that should be built on top of the reputation system in order for peers to wish to increase their reputation (such as the strategies discussed in Section 2.5.3).

So, reputation mechanisms are always useful for supporting any accounting mechanism that depends on peer's reports for the outcome and number of transactions (as for example the public accounts mechanism described above) and could provide useful information concerning the credibility of peers as providers of accounting information. But can they help for assessing the levels of contribution of a peer? When the reputation value equals the balance (e.g. uploads vs. downloads) as computed by an accounting mechanism then the two approaches are very similar[36]. Another way to implicitly use reputation (e.g. for quality of service) to incentivize contribution is to apply to the reputation values computed an appropriate decreasing function over time so as to force peers to contribute constantly in order to keep their reputation at high levels. But note that all these mechanisms could be mainly used to enforce cooperation rather than achieving an economically efficient equilibrium which is our objective.

2.6.2 The role of identity

The notion of identity plays a central role in implementing any type of accounting mechanism. Most importantly, in the distributed and untrusted p2p environment there are some very serious attacks that need to be addressed.

[36]Except that reputation is a more generic concept and in practice is used for determining a peer's type, as for example the characterizations used in Kazaa, rather than an exact balance and the corresponding incentives are less strict (e.g. priority in case of congestion).

Whitewashing

First of all, the ability of peers to create a new identity with low or no cost, enables them to *whitewash* their accounting records if it is for their benefit (depending on the corresponding incentive mechanism). A straightforward solution, implementable only when a central trusted authority is in place, is to assign to nodes free but irreplaceable pseudonyms, possibly based on a verifiable real-world identity (see Castro et al. (2002)). Otherwise, it may be necessary (in both centralized and distributed implementations) to impose a certain cost on newcomers (without being able to distinguish between a new member and a whitewasher); what Friedman and Resnick (2001) have described as the social cost of cheap pseudonyms.

Such costs include monetary entry fees, imposing delays through cryptographic puzzles, or the assignment of the lowest reputation value to newcomers. However, none of these alone could fully solve the problem. Money is very effective but could discourage participation as well. Delays could be surmounted by creating the necessary identities early enough. And in the case of not trusting any newcomer, there would (and should) be always a positive probability that a newcomer will be finally provided service. So, if the cost of creating a new identity is smaller than the cost of the minimum contribution, then a peer has the option to be constantly creating new identities, receiving a small amount of resources for each one.

Feldman and Chuang (2005) examine the effect of whitewashing on the evolutionary stability of reciprocity strategies. Not surprisingly, they find that stability degrades due to the whitewashing attack. Only by imposing a sufficiently large cost of creating a new identity can result to a stable equilibrium in the presence of whitewashers. A way to avoid unnecessary incurrences of the social cost is, as Feldman et al. (2004) propose, for peers to employ a 'stranger policy' which adapts to the previously observed behavior of strangers.

False Trading

The fact that p2p accounting mechanisms always rely on peers' reports to gather the necessary information, makes them susceptible to the false trading attack. That is, peers falsely claiming to have positively or negatively transacted with others. In reputation systems this is translated to raising one's reputation through positive ratings (false praise) or harm the reputation of legitimate peers (false accusation). So, it is critical for all types of accounting mechanisms to address this issue and provide the means to peers to identify questionable reports based on their own experience (we have already discussed some possible solutions). Systems that unconditionally trust peers' ratings or reports (as for example Vishnumurthy et al. (2003)) would be vulnerable to false trading, which could be disasterous when peers can create multiple identities as we discuss next.

The 'sybil attack'

When a peer can easily create multiple identities, it is possible to create artificial large colluding groups with her own identities and thus constitute the aforementioned attacks even more disasterous in a system which depends on accounting information to sustain an efficient level of operation. Douceur (2002) termed the forging of multiple identities a 'sybil attack'.

Cheng and Friedman (2005) introduced the notion of sybilproofness in reputation schemes. The term is used to characterize a system that is immune to false trading (and thus constituting the ability to create multiple identities useless) as 'sybilproof'. The main result is that symmetric reputation functions (ones which depend only on the edge structure of the reputation graph and they are thus invariant under a renaming of the nodes) cannot be sybilproof. The reason is that symmetric reputation functions (as the one proposed by Kamvar et al. (2003a)) cannot distinguish between groups of sybils and real nodes. In contrast, assymetric reputation functions, by assuming that some specified nodes are trusted or having users compute the reputation of other users with respect to themselves (for example trust only users that have directly or indirectly have transacted with them in the past), are more resistant to such attacks. The work of Feldman et al. (2004), Richarson et al. (2003), and Efstathiou et al. (2006) propose approaches that are based on such assymetries to address sybil attacks.

2.6.3 Virtual Currency - Trading

In real life, currency is used in order to avoid complicated accounting and enable trading of resources and services. Instead of aggregating accounting records one could use a mechanism like digital cash [Wayner (1997)] (see also [Rivest and Shamir (1996); Chaum et al. (1990); Jakobsson et al. (2003)] for the implementation of real currency schemes in p2p systems) according to which peers hold their own account (no user memory is required) represented by a number of cryptographic tokens which they cannot forge and which are exchanged according to the rules of an underlying market. However, such tokens need to be issued by someone and strong identities are required in most cases. Furthermore, to control the possibility of a peer double-spending a single token they must eventually return to the issuer (in the absence of trusted hardware at the peer). Three are the possible options.

First, as in the usual digital cash case, the issuer could be a trusted bank external to the peer group. In the context of ad-hoc networks, Zhong et al. (2003a) proposed the use of a central credit clearance server through which nodes can re-deem credit from forwarding packets. Second, it is also possible for a group of peers to create an internal distributed bank which can mint its own currency, using threshold cryptography as proposed in Hausheer et al. (2003). Note that the sub-group is required only for the verification of the validity of the digital currency and not for storing and updating the actual users' accounts. But even if there is a trusted way to issue tokens, the inability of the issuer to observe transactions can still allow peers to cheat on each other. Liebau et al. (2005) address this problem by proposing the use of tokens that can be sent in two steps. Before the service is provided, a token is sent by the client of the specific transaction without her signature (and thus it cannot be used anymore). When the service worth one token is delivered, the signed token is sent to the service provider (the client has the incentive not to sign it only if the service promised is not delivered) and thus both peers loose their incentive to cheat. Alternatively, Vishnumurthy et al. (2003) propose the issueing of receipts for every one of many small fragments of a certain file in order to discourage cheating.

However, the fact that in both the above schemes there should be always some initial endowment offered to new members for bootstrapping reasons constitute necessary the implementation of strong identities, which is an important factor that could discourage participation. Additionally, standard problems of real-world currencies such as inflation and deflation should also be taken care, possibly posing additional administrative overhead.

A particularly appealing approach is to allow each peer to mint their own currency (a form of IOU[37] receipts). This means that providing peers will rapidly collect coins from a variety of banks and to control double-spending, will need to regularly redeem coins at the issuer. But then the credibility of the issuer would be crucial. If he is trustworthy his currency could be further reused and thus the notion of reputation is important. Actually, Moreton and Twigg (2003) show that self-minted currencies is a generalization of tokens

[37] I Owe You

and reputation but this means that in this case some sort of user memory is required in order for reputation to be effective. Interestingly, Efstathiou et al. (2006) propose the use of an assymetric reputation function (see Section 2.6.2) to ensure that the IOU receipts are issued by a trusted peer (one that has directly or indirectly has offered her service in the past) and thus uses only local information to ensure their validity.

2.6.4 Memory-less mechanisms

If no trustworthy accounting information is available, then other “direct” (memory-less) mechanisms should be employed. Such mechanisms are very powerful in terms of enforcing reciprocity but they lack the flexibility provided by the use of accounting information or trading, which allow peers to dissociate consumption from contribution.

Direct exchange

Direct exchange of resources is the simplest to implement incentive mechanism. It is enforced by definition and is totally memory-less and anonymous. BitTorrent [Cohen (2003)], already introduced in Section 2.1, employs such a mechanism and is actually one of the few real p2p applications that manage to enforce a specific incentive mechanism in a p2p fashion[38]. Also notice that direct exchange is a ‘natural’ mechanism employed in preservation systems [Cooper and Garcia-Molina (2002)] (see also Section 2.1.5) and p2p multicast streaming [Chu et al. (2004)].

The success of BitTorrent has recently attracted the interest of many researchers who try to both analyze and improve its performance. Qiu and Srikant (2004) use a simple fluid model to study steady state performance for a BitTorrent-like system. They prove that, under certain conditions, a Nash equilibrium exists, and derive analytical upper bounds for free-riding, which is possible due to the optimistic nature of the optimistic nature of the tit-for-tat strategy implemented in BitTorrent. Interestingly, despite the ability to free ride in BitTorrent (see also [Jun and Ahamad (2005)]), and in contrast with networks such as gnutella, users actually do not [Andrade et al. (2005)]. One of the reasons might be that uploading is not the dominant cost factor in a p2p network especially while downloading [Feldman et al. (2003)], one of our main assumptions. Hales and Patarin (2005) argue that it is the inherent ‘tagging’ taking place in BitTorrent communities that is responsible for

[38]Note that the role of the central server of BitTorrent is not to regulate the contributions of peers but to act as a seed for the content provided.

this fact; see Section 2.6.4. However, the nature of the BitTorrent protocol doesn't address the issue of content availability.

Anagnostakis and Greenwald (2004) propose exchange based mechanisms for providing incentives for cooperation. In the proposed system peers give higher service priority to requests from peers that can provide a simultaneous and symmetric service in return. This approach is generalized to n-wise exchanges among rings of peers (which improves the incentives for content availability —at the cost of the requirement for forming 'cycles' of content requests) and a search algorithm for locating such rings is presented. Cox and Noble (2003) propose a similar approach in the context of p2p storage.

Tag-based cooperation

An interesting approach for improving cooperation in memory-less systems is the use of 'tags' [Holland (1993)]. Tags are labels freely selected by peers playing a repeated multi-person prisoner's dilemma (as discussed in Section 2.5.3). These labels do not have any direct behavioural implication for the agents that carry them but are used as a means for self-organized group selection mechanism. Peers prefer to interact with peers that have the same label.

The result is that cooperating peers, who face low performance due to the existence of a large percentage of free riders in a group (which will always happen after some time since free riding outperforms all other strategies), will have to change often their tags and thus achieve acceptable levels of performance by always "running away" from free riders [Hales and Edmonds (2003)].

"Contribute while consuming"

In [Antoniadis et al. (2005)] we propose a different variation of the direct exchange mechanism. The 'exchange' is to be performed between a peer and the rest of the system as a whole. A way for the system to enforce this exchange in the case for content availability, for example, is to regulate the time needed for a peer to consume the desired resources (i.e. to download a file), so as not to be too small[39] ensuring that a minimum amount of files are shared throughout that time.

[39] Adar (2005) also proposes (in the same issue of sigecom exchanges journal —available at http://www.acm.org/sigs/sigecom/exchanges/volume_5/5.4-index.html) the reduction of upload bandwidth for the cost incurred by seeds in the performance.

We believe that this a very interesting class of memory-less mechanisms, which exploits the public good aspect of p2p systems and requires for a peer to contribute her resources to *any* other peer while consuming, and not necessarily to the peer from whom she is receiving service —and could be probably applicable in other similar contexts as well. Then the enforcement mechanism, run by the serving peer, should ensure the validity of the contribution, which is the most challenging problem in this context. We discuss all the practical and theoretical issues that arise in Chapter 5.

Chapter 3

Private provision of public goods

In this chapter we formalize the discussion on the public good perspective of p2p resource provisioning. We first introduce some standard notation and recent results from economics. Based on these results, we demonstrate that a simple fixed contribution scheme is asymptotically optimal and show the applicability of this public good model to different types of p2p systems. Then, in the next chapter, we will focus on content availability.

3.1 Introduction

We have already motivated our approach to model p2p resource provision as a public good problem and explained why private provision of public goods is not incentive compatible. To see the incentive issues that arise consider the following example. Assume that two players want to decide on the provision of a common facility of size = 1,2. The utility and cost of both players is as follows: $u_i(1) = 2, u_i(2) = 4, c_i(1) = 3$. Each player decides whether to contribute 1 unit of the facility to be provisioned (see Figure 3.1). If no incentive mechanism is in place, player i prefers the other player to contribute one unit in order to acquire 2 units of net benefit (compared to 1 when he also contributes 1 unit). This corresponds to the 'free riding' strategy that we have already argued that is the equilibrium strategy in private provision of public goods. However, they both play this strategy, they will both have zero net benefit. So, we have a 'prisoner dilemma' situation. This fact leads us to the major questions that this dissertation tries to answer:

- Which is the amount of aggregated p2p resources that maximizes social welfare?

Player A \ Player B	provision 1	provision 2
provision 1	1,1	-1,2
provision 2	2,-1	0,0

Figure 3.1: Players' payoffs in the private provision of a public good game

- Which proportion of this amount each individual peer should contribute?
- How can we encourage peers to reveal their private information so as to achieve an efficient allocation and provisioning of resources?

As we have already stated, no mechanism can be efficient, budget balanced (even in the average case) and individually rational [Myerson and Satterthwaite (1983)]. Furthermore, the mechanisms that are constrained-efficient i.e., achieve the best outcome subject to informational and budget constraints (called second-best) are, typically, very complex. In most mechanisms, a large amount of information has to be passed from the peers to the mechanism designer; and the subsequent calculation is complicated. Nevertheless, application of very recent analysis of a public good model by M. F. Hellwig (2003) and Norman (2004), who build on the earlier work of Mailath and Postlewaite (1990), suggests that in large systems, there is not significant loss of efficiency from a scheme that simply demands that each peer contributes a uniform minimum number of files to join the network. This scheme has two advantages: its simplicity; and the fact that prices are not involved. This is very attractive in a large, decentralized system in which implementing a currency can be difficult since peers can contribute 'in kind' (e.g. by making available files). Importantly, Courcoubetis and Weber (2005) have shown that for a large class of utility functions a simple optimization problem suffices to compute a fixed contribution fee that is asymptotically optimal, for a given distribution of peers' types. They also compute a tight bound on the implied efficiency loss. But let us start from the beginning.

3.2 A non-excludable public good model

Consider a situation where n agents bargain for the provision of a public good. The cost of Q units of public good equals to $c(Q)$, where $c(.)$ is a continuous and increasing function with $c(0) = 0$, and the cost is to be covered by the agents themselves (private provision).

If the quantity Q is provided, all agents are automatically in a position to enjoy it; there is no means for individual exclusion and no problem of crowding (congestion).

If Q units of the public good are provided and agent i contributes p_i units towards the cost $c(Q)$, she obtains net benefit

$$\theta_i u(Q) - p_i \, , \tag{3.1}$$

where $u(.)$ is a continuous and nondecreasing function, the same for all agents, with $u(0) = 0$.

We denote $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ the vector of the 'preference parameters' (the types) of the agents, which are assumed to be independent and identically distributed samples from a distribution on $[0, 1]$ with distribution function H and density function h [1]. This distribution, H, is known to all agents, but the value of θ_i is known to agent i alone.

The question is what quantity Q of the public good should be provided and what contributions, $p_1, \ldots, p_n$, the different agents are to make towards covering the cost $c(Q)$. This means that the solution should be *feasible*. Moreover, all agents should acquire a positive net benefit from participating, otherwise they would refuse to make the requested payment. That is, the solution should also be *individually rational*. Given these constraints, our overall goal is to maximize the social welfare: the total utility of the agents minus the total cost. If the true values of the vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ are known to the social planner, his task will be then to solve the following program:

$$\max_{Q \geq 0} \sum_i \theta_i u(Q) - c(Q) \, . \tag{3.2}$$

The solution of (3.2) is trivially feasible and individually rational since for any value of Q^* for which $\sum_i \theta_i u(Q^*) - c(Q^*) \geq 0$, we can find p_is such that $\sum_i p_i = c(Q^*)$ and $\theta_i u(Q^*) \leq p_i \forall i$. Since this is the maximum value of social welfare that could be achieved, given the true vector $\boldsymbol{\theta}$, this solution is called 'first-best'.

Hence, the only obstacle for solving this problem is the availability of the necessary information: the vector $\boldsymbol{\theta}$. And it is an obstacle because this is private information of the agents and a 'mechanism' is required for them to agree to reveal their preference parameters. More specifically, the system designer should first decide on an *allocation* function $Q(.), p_1(.), \ldots, p_n(.)$ which shows how the level of public-good provision and the

[1] Considering other and/or personalized distributions, H_i, does not affect the qualitative results obtained throughout our analysis.

different agents' payments depend on the vector $\boldsymbol{\theta}$. Formally, a mechanism is defined as follows:

1. The designer posts the functions $Q(\cdot)$ and $p_i(\cdot)$
2. The agents, after observing $Q(\cdot)$ and $p_i(\cdot)$, declare some value for their preference parameter θ_i (not necessarily the true one).
3. The designer builds the system to size $Q(\boldsymbol{\theta})$ and asks $p_1(\boldsymbol{\theta}), \ldots, p_n(\boldsymbol{\theta})$
4. Agents pay their contributions and the good is provided. Then all agents are free to enjoy it.

Because of the prior uncertainty about the agents' preference parameters the computation of the optimal allocation function will depend on expectations of their value, according to the assumed distribution H. So, the system designer has to solve the problem

$$\max_{Q(.)\geq 0, p_1(.), \ldots, p_n(.)} \mathbb{E}[\sum_i \theta_i u(Q) - c(Q)]\,, \tag{3.3}$$

subject to the following constraints.

First, the allocation $Q(\cdot)$ should be *feasible*. That is,

$$c(Q(\boldsymbol{\theta})) \leq \sum_{i=1}^{n} p_i(\boldsymbol{\theta}) \tag{3.4}$$

for all $\boldsymbol{\theta} \in [0,1]^n$. This will ensure that the public good is to be fully covered by the private contributions of the participants.

Second, the allocation should be *individually rational*. That is,

$$\mathbb{E}_{\theta_{-i}}[\theta_i u(Q(\theta_i, \boldsymbol{\theta}_{-i})) - p_i(\theta_i, \boldsymbol{\theta}_{-i})] \geq 0,\ \forall \theta_i \tag{3.5}$$

for all i and all $\theta_i \in [0,1]$. This will ensure that all peers gain a non-negative expected utility from participating; otherwise they would refuse to do so.

Finally, the allocation should be also *incentive compatible*. That is,

$$\mathbb{E}_{\theta_{-i}}[\theta_i u(Q(\theta_i, \boldsymbol{\theta}_{-i})) - p_i(\theta_i, \boldsymbol{\theta}_{-i})] \geq \mathbb{E}_{\theta_{-i}}[\theta_i u(Q(\hat{\theta}_i, \boldsymbol{\theta}_{-i})) - p_i(\hat{\theta}_i, \boldsymbol{\theta}_{-i})] \tag{3.6}$$

for all i and all $\hat{\theta}_i$ and θ_i in $[0,1]$, in order for no agent to have an incentive to lie and declare $\hat{\theta}_i$ as her preference parameter when its true value is θ_i. In other words, all peers must gain an expected utility from reporting their type honestly that is at least as large as the expected utility received when they are dishonest.

It is natural to ask whether imposing (3.6) means that the social welfare cannot be as great as if we optimized over all possible mechanisms, including those that are not incentive compatible. However, the well known 'revelation principle' in the theory of mechanism design states that in designing mechanisms, we only need to consider direct-revelation games, where each agent reports his type directly (completely and in a single step up front) and games in which every agent reveals his type truthfully in equilibrium[2]. So, we can safely restrict ourselves to this class of allocations.

But when we have incomplete information about the preference parameters of the agents (and thus the solution should be also constrained by equation (3.6)), unfortunately, due to the impossibility theorem of Myerson and Satterthwaite (1983), the maximum value of (3.3) in the expected sense cannot be achieved since the constraints (3.6) and (3.4) will be binding. In such cases, if the free maximum of (3.3) is to be achieved, there should be the ability of transfers to the community from an external source (as in the case of the VCG mechanism —see Section 2.4.3) in order for the allocation to be feasible, which however is not possible when provision is private (as in our problem).

This means that if we want to build the public good with private provision under incomplete information about the agents' valuations we will have to compromise to a suboptimal solution having to pay an 'information rent' to agents in order to reveal their preference parameters, called 'second-best'. Rephrasing the impossibility theorem of Myerson and Satterhwaite (1983) we have that always Second Best $<$ First Best in terms of efficiency. An interesting question is how close can second best be from first best. But before studying in more depth the relation of second and first best allocation we give a characterization of the second best.

We wish to maximize the average social welfare of the system (equation (3.3)) but with the feasibility, individual rationality and incentive compatibility as these are expressed by

[2]The intuition behind this principle is that given any mechanism, one can always construct a truthful direct-revelation mechanism whose performance is identical. This could be done by building an interface layer between the agents and the incentive compatible mechanism. This layer just inputs the actions of the agents into the original mechanism, and the resulting outcome is the outcome of the new mechanism. Since the interface layer acts "strategically on each agents best behalf", there is never an incentive to report falsely to the interface layer and the new mechanism, which is incentive compatible, has the same equilibrium with the original one. Treatments of the revelation principle can be found in [Fundenberg and Tirole (1992)] and [Mas-Collel et al. (1995)].

equations (3.4), (3.5), and (3.6) respectively. In fact, it can be shown that the following *ex post* requirement (called 'weak feasibility') is equivalent in all senses to the *ex ante* budget balance requirement and has the additional benefit of being easier to manipulate in the analysis:

$$\mathbb{E}[c(Q(\boldsymbol{\theta})) - \sum_{i=1}^{n} p_i(\boldsymbol{\theta})] \geq 0. \tag{3.7}$$

More specifically, an argument of Cramton et al. (1987) shows that for any incentive compatible and weakly feasible $(Q(.), p_1(.), \ldots, p_n(.))$, there exist payment functions $\hat{p_1}(.), \ldots, \hat{p_n}(.)$ such that $(Q(.), \hat{p_1}(.), \ldots, \hat{p_n}(.))$ is incentive compatible and feasible and, moreover, for each agent, both allocations generate the same relation between expected payments and values of the preference parameter θ_i.

It can be shown using standard manipulations, see M. F. Hellwig (2003), that the above reduces to a problem of maximizing (3.13) subject to the constraint

$$E\Big[\textstyle\sum_i g_i(\theta_i)u(Q(\boldsymbol{\theta})) - c(Q(\boldsymbol{\theta}))\Big] \geq 0\,, \tag{3.8}$$

where, $h(\cdot)$ being the probability density function of H, and

$$g(\theta_i) = \theta_i - \frac{(1 - H_i(\theta_i))}{h_i(\theta_i)}. \tag{3.9}$$

According to (3.8), the sum $\sum_i g_i(\theta_i)u(Q(\boldsymbol{\theta}))$ is the maximum possible payment that can be extracted from peers and hence the informational rent that needs to be given to them in order to reveal their private information truthfully is $(\theta_i - g_i(\theta_i))u(Q(\boldsymbol{\theta}))$.

Then the solution of

$$Q(\boldsymbol{\theta}) \in \arg\max_Q [\sum_{i=1}^{n} \theta_i u(Q) - c(Q) + \lambda(\sum_{i=1}^{n} g(\theta_i)u(Q) - c(Q))]\,, \tag{3.10}$$

for some value of the Lagrange multiplier, λ, is not only necessary, but also sufficient for an allocation to be second-best; this also guarantees that a second-best allocation exists (see M. F. Hellwig (2003)).

However, the incentive payments that need to be applied so as to achieve the second-best allocation are really complex. They are functions of the complete vector $\boldsymbol{\theta}$, involve money transfers between agents and there are no known simple approximations. Moreover, it can be shown that as the size of the network grows the ratio of second-best/first-best goes to zero. Before analyzing under which assumptions these two problems could be overcome, we present a simple example to make more concrete the above discussion.

3.3 A simple example

Assume $Q \in 0, 1$. That is, the public good is binary; it is either constructed or not, as for example a bridge. Also assume that $u(Q) = Q$ and $c(Q) = \alpha Q$, with $a \leq 1$. In this case, efficient is an allocation that constructs the good if and only if the sum of the utilities of the agents exceeds the cost.

3.3.1 First-best

We will compute the first-best allocation for the simple case where $n = 2$. Obviously, if $\theta_1 + \theta_2 < \alpha$, $Q^* = 0$ in which case the good wouldn't be provided and thus $p_1 = p_2 = 0$. On the other hand, if $\sum_i \theta_i u(Q^*) - c(Q^*) > 0$, $Q^* = 1$ and we could use any $p_1 \leq \theta_1$, $p_2 \leq \theta_2$ such that $p_1 + p_2 = \alpha$. For example we could define the contribution rule to be $p_i(\boldsymbol{\theta}) = \frac{\theta_i}{\theta_i + \theta_j}\alpha$ But why should agents declare their actual θs? For example, if $p_i(\boldsymbol{\theta}) = \frac{\theta_i}{\theta_i + \theta_j}\alpha$,

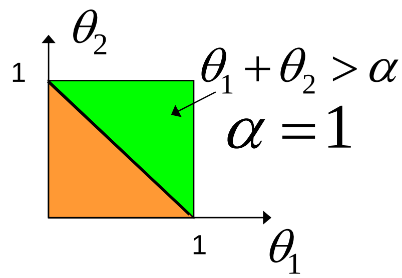


Figure 3.2: Example: First-best allocation

then the agent with the highest type would gain by declaring a $\hat{\theta}_i < \theta_i$. So, it is obvious that given the first-best allocation mechanism, agents don't have the incentive to report their types truthfully and this is the reason we need complete information to maximize social welfare. But what is the best we can do when our mechanism relies on the agents to truthfully report their types?

3.3.2 The VCG mechanism

As already explained in Section 2.4.3, the VCG mechanism has the very attractive property of incentive compatibility but fails to achieve a feasible allocation (it is not budget-balanced). To see this, let us compute the payments of the two agents based on their declarations under the VCG mechanism. Remember that under VCG the price of agent i is the social opportunity cost due to her participation. But which is the opportunity cost from the participation of an agent in our binary public good problem? There are two cases. If $\theta_1 + \theta_2 < \alpha$ the good is not provided at all and so $p_1 = p_2 = 0$. If on the other hand, $\theta_1 + \theta_2 \geq \alpha$ then $p_i = \alpha - \theta_j$.

This is derived by simple calculations from equation (2.1)[3], modified accordingly to include the cost for building the public good. That is, in our problem, according to the VCG mechanism we have that

$$p_i = [\sum_{j \neq i} \theta_j u(Q^*_{-i}) - c(Q^*_{-i}] - [\sum_{j \neq i} \theta_j u(Q^*) - c(Q^*)] \,. \tag{3.11}$$

This means that in the most interesting case where $\sum_{j \neq i} \theta_j < a$ ($Q^*_{-i} = 0$) and $\sum_j \theta_j > a$ ($Q^* = 1$), the first term of the above equation equals to zero since $Q^*_{-i} = 0$ and the second term equals to $\sum_{j \neq i} \theta_j - a$, and thus $p_i = a - \sum_{j \neq i} \theta_j$. In all other cases, where either $Q^*_{-i} = 1$ or $Q^* = 0$, both terms are equal and thus $p_i = 0$.

It is very easy to see that the sum of the payments would not necessarily cover the cost since always $p_1 + p_2 \leq \alpha$. In particular, if $\theta_1 > \alpha$ and $\theta_2 > \alpha$, $Q^* = 1$ and $p_1 = p_2 = 0$, and as a result the social planner would have to pay the full cost for the construction of the good. The incentive compatibility characteristic of the VCG mechanism just provided correct evidence of whether this would be an efficient decision. But note that the VCG mechanism in addition to not achieving budget balance is also vulnerable to manipulations. In our example, both agents could agree to declare a $\theta_i > \alpha$.

[3]Given the complete report of agents' utility functions, $u_i(x)$, $i = 1, \ldots, n$, where x being an allocation of the good to the agents, the price agent i has to pay is given by $p_i = \sum_{j \neq i} u_j(x^*_{-i}) - \sum_{j \neq i} u_j(x^*)$, where $x^* = \arg\max_x \sum_{i=1}^{n} u_i(x)$ is the optimal allocation and x^*_{-i} is the optimal allocation when peer i doesn't participate.

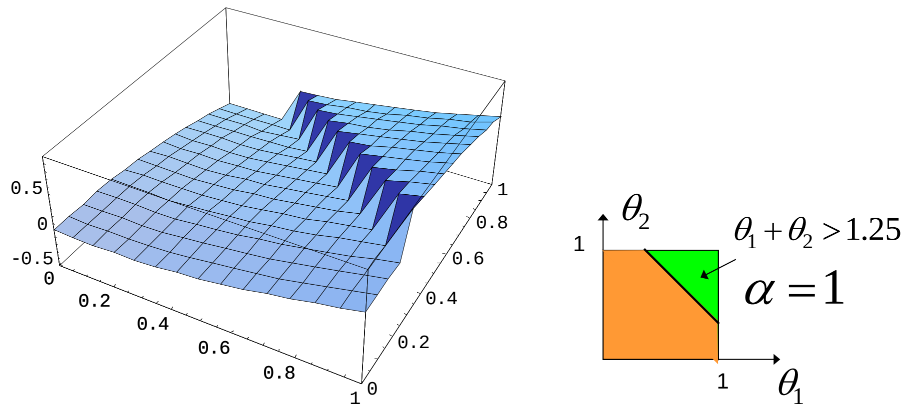


Figure 3.3: The payments for $n = 2$ and the corresponding informational rent required.

3.3.3 Second-best

So, what is the best we can do when require both the feasibility and incentive compatibility constraints to be satisfied? Let us compute the second-best allocation for our simple example. From (3.10), with standard calculations, we have that

$$Q(\boldsymbol{\theta}) = 1 \iff \sum_{i=1}^{n} \theta_i - \frac{\lambda}{1+\lambda} \sum_{i=1}^{n} \frac{(1 - H_i(\theta_i))}{h_i(\theta_i)} \geq c \tag{3.12}$$

Solving (3.12) for $n = 2$, we get that $\theta_1 + \theta_2 > 1.25$. That is, the aggregated informational rent in this case equals 0.25 (see Figure 3.3). To demonstrate how complicated the incentive payments can be, we plot in Figure 3.4 the payments for $n = 2$.

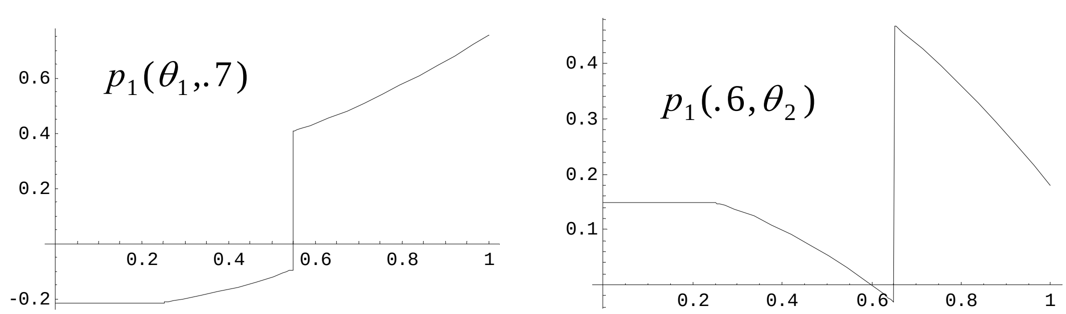


Figure 3.4: The payments for $n = 2$

3.4 Exclusions and the role of size

We now turn to the main question of interest: how close to full efficiency can the second-best mechanism get? By the Myerson-Satterthwaite theorem (and the form of the Lagrangian), we know already that any mechanism delivers total social welfare that is strictly less than the efficient level. The analysis here will focus on this issue when the number of peers n becomes large.

The results of Mailath and Postlewaite (1990), M. F. Hellwig (2003) and Norman (2004), along with the more general analysis of Al-Najjar and Smorodinsky (2000), suggest that a peer can be induced to make contributions only to the extent that the peer makes a significant difference to the totalquantity of the public good. The latter is of order $1/\sqrt{n}$. It is intuitive that it should be decreasing in n: the greater the number of peers, the less influence any individual peer will have on the total amount of good provided. Hence each agent is willing to contribute to the cost in the order of $1/\sqrt{n}$ times the benefit he expects to gain from the network. The total contributions that can be elicited from agents is therefore of the order of $n/\sqrt{n} = \sqrt{n}$ times some measure of the overall benefit per agent. In contrast, efficient contributions are of the order of n times a measure of the overall benefit per peer from shared files. Hence, in the limit as $n \to \infty$, the total utility level achieved by the second-best mechanism goes to zero as a fraction of the efficient level. This result is stated in the next proposition:

Proposition 1 *Let M_n denote the total utility achieved by the mechanism characterized in* (3.10)*; and let E_n denote the total utility achieved in the full efficiency case (in which the constraints* (3.6)–(3.5) *can be ignored). Then*

$$\lim_{n\to\infty} \frac{M_n}{E_n} = 0.$$

The problem highlighted in Proposition 1 is that peers are willing to contribute towards covering the cost of the good only if they believe that they have a significant effect on the total good that will be provided—i.e., that they are 'pivotal'. But as the number of agents grows, so does the chance that any particular agent is actually pivotal. If an agent does not expect herself to be pivotal, then she would report her type as being the lowest possible. Note that the mechanism does manage to obtain some contributions from agents;

and as the number of agents grows, so does the total amount of good provided. But the gap between full efficiency and the total utility achieved by the mechanism grows ever larger as the network grows in size.

A possible response to this problem is to reduce the incentive of agents with high types to mimic agents with low types. A mechanism to achieve this is by excluding agents with low types based on their reports. This will have two effects in the limit as the network becomes large. The first is to close the efficiency gap that is identified in Proposition 1. This fact is stated formally in the following proposition, proved in M. F. Hellwig (2003).

Proposition 2 *Let M'_n denote the total utility achieved by the mechanism that allows for peers to be excluded from the network; and let E_n denote the total utility achieved in the full efficiency case (in which the constraints* (3.6)–(3.5) *can be ignored). Then*

$$\lim_{N\to\infty} \frac{M'_n}{E_n} \in (0,1).$$

Hence excluding peers allows the mechanism designer to escape from the asymptotic full loss of surplus. Full efficiency cannot be achieved; but the limit of the ratio of second-best to first-best surplus is bounded away from zero.

To understand this result note that under an exclusion mechanism a minimum contribution would be required by all included agents; the ones whose reported type is high enough. So, the threshold type would pay her exact valuation, and higher types are willing to spend more only to the extend that expected consumption increases in announcement. The mechanism designer, except from the amount of the public good Q and the payments p_i decides on the value of the inclusion function $\pi_i(\boldsymbol{\theta})$, which equals to 1 if peer i is included in the system and 0 if not. Then a minimum contribution mechanism could be defined as follows.

Definition 1 *The mechanism $(Q, \boldsymbol{\pi}, p)$ is a minimum contribution mechanism if there exists a $\bar{\theta}$ such that $\pi_i(\boldsymbol{\theta}) = 1$ if and only if $\theta_i \geq \bar{\theta}$ and $p_i(\boldsymbol{\theta}) \geq \pi_i(\boldsymbol{\theta})\,\bar{\theta}\,f(\boldsymbol{\theta})$ for each i and $\boldsymbol{\theta} \in H^n$, where $f(\cdot)$ expresses a measure of the overall benefit per peer from shared files.*

Hence, when such a mechanism is applied the overall efficiency achieved would now depend on the number of agents n (instead of $\sqrt{n}$) and thus the efficiency loss is bounded.

Moreover, note that in a large economy, the average agent has little influence on the quantity provided. Transfers from participating agents are thus almost independent of type and the further efficiency loss when a *fixed* fee is used for all agents is negligible with many agents. So, asymptotically, all payments come from "entry fees" supported by the possibility of individual exclusion.

In the context of p2p file sharing such a fixed fee mechanism is very attractive since as we discuss in the next chapter it could be translated as a very simple scheme that requires from peers to share a fixed number of files, computed off-line based on the distribution of peer types and their number. This scheme has two advantages: its simplicity in terms of the form of the policy, being a simple contribution rule; and the fact that prices are not involved. The latter is attractive in a large, decentralised system in which implementing a currency can be difficult. To that respect is also very important the fact that no communication is required between the peers and the system designer. Peers report their type by simply deciding to participate or not based on the declaration of the system designer on the expected amount of files shared (see Section 4.6.3 for some interesting stability issues that arise).

So let us reformulate the mechanism design problem under the assumption that a peer may be excluded from using the good. Given knowledge of this mechanism, each peer declares his θ_i. The mechanism then sets $Q(\boldsymbol{\theta})$ and also decides which peers may use the good. If peer i is excluded from using the good then $\pi_i(\boldsymbol{\theta}) = 0$. If he is allowed to use it, then $\pi_i(\boldsymbol{\theta}) = 1$ and he must pay a fee $p_i(\boldsymbol{\theta})$. So we just rewrite equations (3.3), (3.7), (3.5), (3.6), including the function $\pi_i()$. The mechanism design problem now is to maximize expected social welfare:

$$\underset{\pi_1(\cdot),\dots,\pi_n(\cdot),\,Q(\cdot)}{\text{maximize}}\ E\left[\textstyle\sum_i \pi_i(\boldsymbol{\theta})\theta_i u(Q(\boldsymbol{\theta})) - c(Q(\boldsymbol{\theta}))\right] \tag{3.13}$$

subject to a 'weak feasibility constraint', which says that the expected payments must at least cover the expected cost:

$$E\left[\textstyle\sum_i \pi_i(\boldsymbol{\theta})p_i(\boldsymbol{\theta}) - c(Q(\boldsymbol{\theta}))\right] \geq 0\,, \tag{3.14}$$

'individual rationality' constraints, which say each peer can expect positive net benefit by participating:

$$E\left[\pi_i(\theta_i, \boldsymbol{\theta}_{-i})\left\{\theta_i u(Q(\theta_i, \boldsymbol{\theta}_{-i}) - p_i(\theta_i, \boldsymbol{\theta}_{-i})\right\}\right] \geq 0\,, \tag{3.15}$$

for all i, and 'incentive compatibility' constraints, such that each peer i does best by declaring his true θ_i rather than 'free-riding' by declaring some other θ_i':

$$\begin{aligned} &E\left[\pi_i(\theta_i, \boldsymbol{\theta}_{-i})\left\{\theta_i u(Q(\theta_i, \boldsymbol{\theta}_{-i}) - p_i(\theta_i, \boldsymbol{\theta}_{-i})\right\}\right] \\ &\geq E\left[\pi_i(\hat{\theta}_i, \boldsymbol{\theta}_{-i})\left\{\theta_i u(Q(\hat{\theta}_i, \boldsymbol{\theta}_{-i}) - p_i(\hat{\theta}_i, \boldsymbol{\theta}_{-i})\right\}\right]\,, \end{aligned} \tag{3.16}$$

for all i and θ_i'.

Note that the argument of Cramton et al. (1987) concerning the equivalence of 'feasibility' and 'weak feasibility' holds when no exclusions are allowed. However, as Courcoubetis and Weber (2005) show, a similar result holds also for the case of exclusions.

Similarly to the case of a non-excludable public good, it can be shown using standard manipulations, see M. F. Hellwig (2003), that the above reduces to a problem of maximizing (3.13) subject to the constraint

$$E\Big[\textstyle\sum_i \pi_i(\boldsymbol{\theta})g_i(\theta_i)u(Q(\boldsymbol{\theta})) - c(Q(\boldsymbol{\theta}))\Big] \geq 0\,, \tag{3.17}$$

where, $h(\cdot)$ being the probability density function of H, and $g(\cdot)$ as in (3.9).

This problem can also be solved using Lagrangian methods (see Courcoubetis and Weber (2005)). That is, there is a nonnegative λ such that it is equivalent to solve the unconstrainted problem

$$\underset{\substack{\pi_1(\cdot),\ldots,\pi_n(\cdot) \\ Q(\cdot)}}{\text{maximize}}\; E\Big[\textstyle\sum_i \pi_i(\boldsymbol{\theta})(\theta_i + \lambda g(\theta_i))u(Q(\boldsymbol{\theta})) - (1+\lambda)c(Q(\boldsymbol{\theta}))\Big]\,. \tag{3.18}$$

In the following we present the main results of Courcoubetis and Weber (2005) which show that for a large class of public good models including the one presented in the next chapter, a simple fixed fee policy which is independent of the declarations of the peers, is enough to get us within $O(1/n)$ of the second best. This is a very important result because it states that even such a very simple and intuitive scheme is the best one can do in terms of social welfare maximization in p2p systems that match our modelling approach (one that focus on their public good aspect).

3.5 Asymptotically optimal mechanism design

As already mentioned, the full solution of our problem is, in general, very complex. However, Courcoubetis and Weber (2005) show that we can use (3.19) below to approximate the solution of (3.13) subject to (3.17). They then show that when n is large, there is a simple mechanism designed by solving a simple optimization problem, which obtains a value of the objective function that is within $O(1/n)$ of the maximum achievable (the second-best).

More specifically, assuming $u(Q) = AQ^{\alpha}$ and $c(Q) = BQ^{\beta}$, where $A, B > 0$, $\delta > 0$, $0 < \alpha \leq 1$,$\beta \geq 1$, and $\alpha < \beta$, Courcoubetis and Weber (2005) prove the following Theorem[4].

Theorem 1 *Let P be the problem of maximizing* (3.13) *subject to* (3.17)*, with optimal value Φ_n. Let Q^* and θ^* be the optimizing decision variables in the problem P^*, defined as*

$$\underset{\theta\in[0,1],Q\geq 0}{\textit{maximize}} \left\{ nu(Q) \int_{\theta}^{1} \eta h(\eta) d\eta - c(Q) \right\} \tag{3.19}$$

subject to

$$n[1 - H(\theta)]\theta u(Q) - c(Q) \geq 0\,. \tag{3.20}$$

Let the optimal value be Φ_n^*. Suppose we take as a solution[5] for P the functions $\pi_i(\theta) = 1\{\theta \geq \theta^*\}$ and $Q(\boldsymbol{\theta}) = Q^*$. Then the expected social welfare under this (suboptimal) mechanism is Φ_n^*, and this is asymptotically optimal in the sense that $\Phi_n/\Phi_n^* \leq 1 + O(n^{-1})$

Notice that all peers who are allowed to use the good pay the same fee of $f = \theta^* u(Q^*)$. Peer i may use the good if and only if $\theta_i \geq \theta^*$. This is the same as the condition that his net benefit should be nonnegative, i.e., $\theta_i u(Q) - f \geq 0$. The expected number of peers for which this holds is $n(1 - H(\theta^*))$ and $Q = n(1 - H(\theta^*))f$. This means that there is actually no need for the planner to intervene in an active manner. Once f has been set, the optimum $\pi(\cdot)$ and Q arise simply by peers making their own self-interested decisions.

[4] In [Courcoubetis and Weber (2005)] this result is proved for a more general cost function, $c(n, Q)$. Moreover, under a weaker assumption, that $A_1 Q^{\alpha} \leq u(Q) \leq A_2 Q^{\alpha}$, and $B_1 h(n) Q^{\beta} \leq c(Q) = B_2 h(n) Q^{\beta}$, where $A_1, A_2, B_1, B_2 > 0$, $0 < \alpha \leq 1$,$\beta \geq 1$, and $\alpha < \beta$, the bound becomes $\Phi_n/\Phi_n^* \leq 1 + O(\sqrt{n})$.

[5] It may not be feasible since we compute the cost on the average case.

The intuition underlying this result is as follows. For each $\theta \in [0,1]$, let $S(\theta)$ be the set of all agents who have preference parameters in the interval $[\theta, 1]$. Denote the size of this set by $|S(\theta)|$. Now $E|S(\theta)| = n(1-H(\theta))$, for all θ. Suppose we had the stronger fact that $|S(\theta)| = n(1-F(\theta))$, for all θ. Since, by the remarks above, the optimal mechanism includes the set of agents with preference parameters greater than some θ, we find (using integration by parts) that P simplifies to P^*. The planner includes all agents in $S(\theta)$, for some θ, and then charges each of these agents the same fixed fee f. The mechanism is individually rational for all agents in $S(\theta)$ provided $f \leq \theta u(Q)$. So using the greatest charge consistent with this, namely $f = \theta u(Q)$, the total payment is $n(1-H(\theta))\,\theta u(Q)$ and by (3.20) this covers the cost of $c(Q)$.

Now return to the original problem P. The weak law of large numbers guarantees that $|S(\theta)|$ is close to $n(1-H(\theta))$ with high probability when n is large. So we can expect it to be very nearly optimal to adopt the mechanism above, i.e., to take $Q(\theta) = Q^*$ and set a fixed fee of $\theta^* u(Q^*)$, thus including those peers for which $\theta_i \geq \theta^*$.

As a simple illustrative example, suppose that $u(Q) = 0.6Q^{1/2}$, $c(Q) = Q$, and θ_i is uniformly distributed on $[0,1]$, so $H(x) = x$. The solution of (3.19)–(3.20) is $\theta = 1/4$, $Q = 0.0126563n^2$ and the social welfare is $0.006328125n^2$. The fee is $0.016875n$. We can compare this to the maximized social welfare that could be achieved if we were unrestricted by constraints of individual rationality and incentive compatibility. This social welfare would be $0.01125n^2$, which is achieved by $Q = 0.01125n^2$. The need to satisfy the constraints leads to a reduction in social welfare of 43.75%, for this specific numerical example.

Segal (2003) has considered a related setup in which a monopolist is trying to maximize his profit by selling units of a single good to n buyers (agents). He designs a mechanism such that the price each buyer pays is a function of all n buyers' declared valuations (our $\boldsymbol{\theta}$). He supposes that the underlying distribution H of these valuations is unknown (but is derived from the bids received) and shows how to define the mechanism so as n tends to infinity, it achieves the optimal monopoly profit that could be obtained had H been known. Essentially, the formulae are the same, except that θ_i is taken to be a sample from $\hat{H}_i(\cdot|\boldsymbol{\theta}_{-i})$, which is H conditional on knowing $\boldsymbol{\theta}_{-i}$. Like Courcoubetis and Weber (2005), he finds that the limiting policy is to offer each buyer the same fixed price, but he does not derive the $O(1/n)$ error, deal with the additional optimization over Q that our public good model involves, nor explain any application to a problem of p2p resource allocation.

Finally, note that even if we don't know the distribution of peers' types such a distribution always exist and as a result there is a simple scheme as the one presented that is asymptotically optimal.

3.6 Applications

We will now motivate the application of these very interesting theoretical results in the context of specific p2p systems, whose service provided has a public good flavor, such as file sharing, peering WLANs, and scientific grids. File sharing, and more specifically content availability, is covered in detail in the next chapter and WLAN peering in [Courcoubetis and Weber (2005)] and [Efstathiou and Polyzos (2003)]. Modelling of grid computing as a public good provision problem is a novel approach introduced in this dissertation and since in-depth analysis of this model has been left for future work, we will elaborate a more than the other two applications during its description in this section. We believe that our approach could be further applied in more p2p systems which share the same basic characteristic, the public good aspect. However, in case where congestion effects or service provision costs are not negligible (as assumed in the majority of our work for the case of file sharing) our model, in all of its variants, should be extended to account for this extra cost incurred by peers due to the consumption of others (negative externalities), either in terms of costs for resource provisions or in terms of congestion (see Chapter 6).

3.6.1 Content availability

In the case of p2p file sharing systems, when uploading does not incur significant costs, we have argued that content availability is the main good provided and could be treated as a public good: all peers benefit from the total number of available files while content is not consumed by downloading. So, Q equals to the amount of files shared per unity of time and the 'payment' of a peer i is the amount of files, f_i, stored —and shared— in her PC multiplied by the percentage of time she stays on line (her availability).

One of the main contributions of this dissertation is the adaptation of the public good model introduced in this chapter for the case of p2p file sharing and the in depth analysis of various theoretical and implementation issues that arise in this context. This part of our work is presented in the next two chapters and we won't elaborate more on this here.

3.6.2 WLAN peering

WLAN peering [Efstathiou and Polyzos (2005); Courcoubetis and Weber (2004)] is a novel application, which is motivated by the fact that WLANs provide large amounts of bandwidth that is mostly underutilized by its local users and moreover the pipe that connects the local WLAN users to the Internet is usually of a broadband nature (DSL) and may also be underutilized over large periods of time. So, WLAN owners can greatly benefit by sharing their WLANs[6] in order for all to benefit from the aggregated coverage offered.

So, in this case, our public good Q is the total coverage achieved by people sharing their WLANs, which could be modeled as the probability of a successful service request when roaming. Notice that service provision (allowing visitors to connect to the Internet through one's WLAN) doesn't reduce the overall ability of others to connect to the same WLAN, when there is no congestion. That is, coverage is non-rivalrous.

A way to model the contribution of peer i (see Courcoubetis and Weber (2005)), is to assume that a WLAN covers an area A from the total area B and peer i accepts service requests from roaming users with probability p_i (either due to her local policy or due to her availability —the amount of time she switches on her WLAN). Then we have $Q = \sum_i^n p_i A/B$ and the analysis is similar to this of file sharing. Courcoubetis and Weber (2005) analyze the more interesting case where the cost for contributing coverage A with probability p_i is proportional to the rate of service requests that a peer accepts (see also the discussion in Section 6.1.1).

3.6.3 Grid computing

Although grid computing (see Section 2.1.4) is mostly envisioned as a market place of computing resources (either for efficiency purposes [Krauter et al. (2001)] or for financial exploitation [Kenyon (2004)]) there are certain applications for which a public good model like the one presented in this chapter would be more appropriate. More specifically, scientific grids are a special case of a grid p2p community where participants are large organizations, such as universities, research laboratories, etc. which share their computing resources in order to be able to perform tasks with enormous requirements in computational power and/or storage space (for temporal data) such as experiments in the fields of biology,

[6] Existing technology allows WLAN administrators to control access to their networks and to limit the consumption of network resources by remote (roaming) users.

physics, astrophysics, and more. In such a system, the participating organizations have in general symmetric roles, since they all act as both providers and consumers of computing resources, and have in general similar, or at least comparable (potential) capabilities. That is, they are peers.

The benefit of the formation of a scientific grid is mainly due to the high degree of multiplexing that can be achieved, since demand for computing resources is in general very bursty. That is, jobs are submitted rarely but require an enormous amount of resources. This means that if an organization relied only on its own resources to satisfy its demand for executing computational (and/or storage) intensive tasks, it should make an enormous investment for resources, which would be needed only for the peak periods and thus would be underutilized most of the time. But clearly, if no explicit incentive mechanisms are in place, the rational strategy of all peer organizations would be to free ride on others' efforts. That is, share no resources and consume as much as possible. Then no organization would contribute any resources and the system would collapse.

In general, grid resources could have different values for an organization during different time periods (exactly due to the bursty nature of demand). So, if we divide time into k slots, we could decide to built an exchange economy (see Varian (1992), Chapter 17) of k commodities. In such an economy each organization i is characterized by its initial endowment $w^i = (w_1, \ldots, w^k)$ of each of the k resources —in our case it could be possibly $w_1 = w_2 = \ldots = w_k$— and its utility function $u_i(x_1^i, \ldots, x_k^i)$, where x_j^i is denotes the amount of resources consumed in slot j. In a pure exchange economy, the agents trade the goods among themselves attempting to make themselves better off. According to demand, different (relative) prices for the resources at the different slots will be formed based on which the exchange of resources will be performed. The first and second theorems of welfare economics suggest that when agents behave competitively (that is, they take prices as given, independent of their action), in equilibrium, demand will be equal to supply and social welfare will be maximized.

So, why not solve the resource provision and allocation problem based on the exchange economy model? There are two weaknesses of such a system. First, it requires the implementation of market mechanisms for the computation of prices and hence additional core market functionality, such as accounting and charging. Second, it assumes organizations have a well defined utility function in the above model, in the sense that they know in advance their exact needs in the various slots so that they can do the exchange for all the slots of a given time period at some earlier enough point in time. Thus, in cases where demand of specific slots is not predictable in advance and contention is rare, we believe that an alternative economic system, which would treat the total amount of resource contributed as a public good could be more appropriate.

But notice that unlike classic public goods, the resources involved in a computational grid are rivalrous: a peer consuming storage space or CPU cycles at a node prevents another peer for consuming the same resources of this node during the same time period. However, although rivalrous, grid resources are not consumable: when one organization finishes her task the storage space will be fully available for the next one. So, when there are not frequent resource contention overlaps, it could be assumed that grid resource are non-rivalrous, over time. In other words, the sum of all the resources contributed by peers in a Grid system, which typically consist from the initial investment on infrastructure from each organization, the total 'grid capacity', exhibits an aspect of a public good, which under a fair scheduling scheme it will be potentially available to all participants for executing computationally intense tasks. And the more amount of resources is available the less completion time for the submitted tasks could be potentially achieved for everyone. The only difference from standard public good models will be that the participants cannot use this public good all together at the same time which may be not be an issue if resource usage could be scheduled in some non-overlapping way.

So, an interesting mathematical abstraction of a scientific grid is a server with capacity C, which is to be provisioned by the clients themselves. The delay experienced by clients certainly depends on the overall rate request but the higher the total capacity the less delay they will all experience. If they were themselves responsible for building the capacity C how much each one should be incentivized to contribute? This is a challenging resource provision problem, which again has many similarities with the classical mechanism design problem the 'private provision of a public good' and so all the results presented in this chapter are relevant in this context as well. We formulate below a simple public good model for resource provisioning in scientific grids in order to formalize the above discussion and then discuss some grid-specific aspects of this problem, some of which have also been addressed in the context of file sharing.

A simple public good model for scientific grids

Assume that there are n organizations each one contributing s_i amount of storage space and c_i CPU power in a scientific grid and that they dedicate these resources only for 'grid usage'. We could imagine that these resources are managed by a central scheduler and peers cannot have access to them to run external to the Grid tasks. For simplicity, assume that we have one variable to express the amount of resources made available by peer i, q_i, and that the total amount of resources available by the Grid, Q, is given by $\sum_i q_i$.

Also assume that the cost for providing these unused resources is zero (no delivery/bandwidth costs), but there is a cost, $c(q_i)$, for purchasing them. For example, $c(q_i) = \alpha q_i$. Finally, assume that all tasks have the same storage and CPU time requirements and each task uses the full infrastructure while it runs. Nodes are again differentiated only by their 'preference parameter' θ_i, their type, which measures how much valuable is the service to them.

Our overall goal is to maximize the social welfare (the total net benefit of the agents minus the total cost):

$$\max_{Q \geq 0} \sum_i \theta_i u(Q) - c(Q)\,, \tag{3.21}$$

where $u(,)$ is a concave function expressing the increasing, but with diminishing returns, value of peers (for the delay faced per task) as the size of the infrastructure, Q, grows. In the simplest case we assume that nodes submit their jobs in scheduled times during the week for example, and the cost is the completion time of a job after it is submitted.

In this system a peer acquires utility proportional to how fast her tasks finish. Obviously, the completion time of a task depends both on the total amount of resources Q and on the sum of the request rates of peers (due to congestion). However, when there is enough capacity for all tasks requested in one day to finish during that day (no congestion effects), we can assume that completion time depends only on Q. In this case, a round robin scheduling of tasks would suffice for all tasks to finish on time. We believe that although this assumption seems simplistic, it is actually realistic in many existing scientific grids where demand is sporadic but with enormous requirements for storage and CPU power. We discuss in the next section how one could address the case of congestion.

Note that under this model, it is the initial amount of resources "invested" by peers that counts. And it is exactly to this respect that they have the incentive to free ride: purchasing a very small amount of resources and relying on the investments of others to be able to run their tasks fast. This is actually the same model we analyzed above and the main result, that a fixed contribution mechanism in terms of initial amount of resources made available by each peer is asymptotically optimal as far as economic efficiency is concerned, holds for this case also. This is a nice result, because it says that we don't need to have much information about the valuation of peers (only its distribution) nor implement complicated incentive mechanisms to be able to maximize the economic efficiency of the system.

Dealing with heterogeneity and exclusions

An important characteristic of p2p systems in general is that they comprise a highly heterogeneous mixture of peers in terms of utilities and capabilities. It is easy to see that the ability to categorize peers into different groups with different fixed contributions and possibly granting differentiated levels of participation (or quality of service) to each group can increase significantly the overall efficiency achieved. Most importantly, as we show in our analysis for group forming in file sharing systems (see Section 4.6.1), even a very small number of such groups could be enough to achieve the maximum possible efficiency on average. The exploitation of this extra information is a particularly interesting topic in the context of resource provision in scientific grids for two main reasons.

First, the fact that participation in scientific grids is not anonymous and there are observable characteristics of the participating organizations (e.g. their size), constitutes such a categorization more realistic than it is for other p2p systems such as file sharing and thus no incentives for self-selection are required, which as shown in Section 4.6.1, would harm efficiency.

Second, the ability to create different levels of quality of service could also eliminate (or constrain) the need for exclusions, which are in many cases undesirable in this context. More specifically, there are situations where exclusions are not desirable due to political reasons. For example, certain organizations cannot be excluded from a national-wide grid even if they have not enough utility or resources to contribute the required minimum contribution. Notably, the administrative costs and the overhead for aggregating resources constitute very small contributions more costly than useful and it would be actually beneficial if organizations with limited capabilities participated with zero contribution. Then a suitable mechanism should be built which would ensure that the incentives remain aligned and no organisation can exploit this fact for its own benefit. That is, organizations who will choose to participate as free riders should receive an adequately reduced quality of service or have access to a limited percentage of the overall resources provisioned. We analyze in detail this problem as well in Section 4.6.1.

Dealing with congestion

Another important aspect of a grid system is that, unlike content, resources contributed are actually rivalrous: when demand is high, there could be contention for computing power and/or storage. Thus, in this case there will be congestion effects, since the utility of organizations will be affected by the consumption of others (there will be negative externalities). The possible negative effects of contention are magnified by the fact that in reality grid capacity is not fully available to everybody (as assumed by our abstraction of a grid system) due to the restrictions that arise by the fragmentation of tasks to different sites. So, scheduling becomes a very important dimension of a candidate economic model and the corresponding incentive mechanisms proposed and it should not be considered an independent system functionality.

When contention is not very rare, we could assume that more sophisticated incentive mechanisms are provided by the scheduler (e.g. in case of contention peers that have contributed a larger amount of resources may have priority to use the public good) and it is a very interesting avenue for future research to study the corresponding Nash equilibria under different incentive mechanisms of this kind and give some insights for the selection of the best ones.

The problem in this case is that organizations should be now characterized by an additional parameter, their service request rate, in order for one to model the negative externalities due to contention. This complicates significantly the analysis since our results do not hold when agents are characterized by more than one parameter. The extension of our simple model for capturing the possible congestion effects is left for future research (see also Section 6.1.2).

Chapter 4

The case of content availability

In this chapter we formulate a specific public good model for content availability in p2p file sharing systems. We describe the possible incentive mechanisms under complete and incomplete information and, using a simple example of our model, compare their efficiency as the number of peers grows demonstrating the attractiveness of the fixed contribution scheme. We also analyze some interesting theoretical issues that would arise in a realistic p2p system employing the fixed contribution scheme.

The majority of the results presented in this chapter are joint work with Robin Mason and Richard Weber (see Antoniadis et al. (004a) and Antoniadis et al. (004b) respectively).

4.1 A public good model for p2p file sharing

We apply the ideas presented in the previous chapter to a problem of p2p file sharing by defining the appropriate functions $u(\cdot)$ and $c(\cdot)$. Suppose that n peers make available files to share with one another. Suppose also that the utility obtained by peer i when the expected number of distinct supplied files per unity of time is Q, is $\theta_i u(Q)$, where $u(.)$ is concave in Q. A social planner should decide on the optimal quantity of content Q^* that should be built and what percentage of the cost $c(Q^*)$ each peer should contribute (depending on her type θ_i). Now it is the number of files shared per unity of time, denoted by f_i, instead of a payment p_i that defines the contribution of peer i.

The payoff function of peer i is given by

$$b_i(Q(F)) = \theta_i u\left(Q(F)\right) - f_i, \tag{4.1}$$

where $\sum_i f_i = F$.

Q is the total amount of distinct files and peer i contributes a percentage of these files f_i. Since it is the number of distinctly different files which are shared that matters, we should account for the possibility that more than one peer will make available the same file. So, if F is the total amount of files contributed by peers, $F = \sum_i f_i$, there is in general a concave function $Q(F)$ that gives us the number of distinct files Q as a function of F. We discuss the shape and characteristics of this function in the following.

Here, the utility function $u(\cdot) \geq 0$ is assumed to be continuously differentiable, increasing and strictly concave in its argument. It is the same for all peers. As in the public good model of Chapter 3, peers differ in the payoff parameter θ_i (these are drawn from a distribution H and density function $h(\cdot)$) and they all have the same cost function for sharing files: linear with no fixed cost. This assumption is not crucial for the qualitative results we obtain. It allows payments to be done in 'kind' (i.e., files) instead of actual money which would be then converted to files according to a more complicated cost function.

Notice that we attempt to give incentives to peers to stay in the system providing any content content item, and not simply to have provided at some time in the past, regardless of request rate[1]. We describe in Chapter 6 some alternative models for situations where request rate is costly and analyze in depth one of them, the one that takes into account the uploading costs.

Notice also that peers exhibit some degree of altruism since they obtain benefit from their own contributions. This modelling assumption is consistent with observed peer behaviour in existing content sharing systems, where most peers contribute at least some (low) amount of files with no explicit external incentives. However, we don't claim to have modelled altruism appropriately. We just note that this simplification is not unrealistic.

Finally, for any n-vector $\boldsymbol{\theta}$ of payoff parameters, without loss of generality we order the peers so that $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_n$ (we will use this for the computation of several of our proposed incentive schemes in Section 4.5).

4.1.1 The function $Q(F)$

We might imagine that each peer provides the same number of files, say f, choosing these randomly from amongst a set of N maximum content items possibly available (we relax

[1] see Section 2.2.2 for the motivation of this modelling decision.

this equal contribution assumption later). Then the expected number of distinct files that will be provided is

$$Q = N(1 - (1 - f/N)^n)\,,$$

so

$$f(Q) = N\left(1 - (1 - Q/N)^{1/n}\right)\,.$$

Suppose that each peer incurs a cost in providing files that is proportional to the number she contributes. For simplicity we let the constant of proportionality be one (noting that we could always re-scale the utility function). Thus the total cost is $c(F) = F$, where $F = nf(Q)$, and this is a convex increasing function of Q. Also, for any fixed Q, the cost nf rapidly increases with n to the asymptote of $-N\log(1-Q/N)$. This is greater than Q, the total cost if there were no duplication in the files peers supply. Note (see Figure 4.1) that for a large range of values of Q the cost is almost linear in Q, but then increases rapidly as Q approaches N. For example, for $n = 100$, we find $nf(Q) = Q + 0.00005Q^2 + 3.32\times 10^{-9}Q^3 + \cdots$. This justifies an approximation $c(F) = Q$ when Q/N is of moderate size.

A slightly more sophisticated model might imagine that the peers share different numbers of files. Suppose $n\rho_i$ of peers each share i files, each of them choosing his i files randomly from amongst a set of $N = na$ files, $a > 0$. Let m be an upper bound on the number of files that any one peer can share, and $\sum_k \rho_k = 1$. The expected number of distinct files supplied will be

$$\begin{aligned} Q &= na\left[1 - \prod_{k=1}^{m}\left(1 - \frac{k}{na}\right)^{n\rho_k}\right] \\ &= na\left[1 - e^{-\sum_k k\rho_k/a}\right] + o(1)\,. \end{aligned} \tag{4.2}$$

Now $F = n\sum_k k\rho_k$ is the total number of files provided by the peers. So we can again use the same approximation as above:

$$Q(F) = N\left(1 - e^{-F/N}\right)\,. \tag{4.3}$$

Note that as $Q(F)$ is concave in F, we have that $\bar{u}(F) = u(Q(F))$ is also concave in F.

For a given Q we will require

$$F(Q) = -N \log\left(1 - Q/N\right) \; . \tag{4.4}$$

Of course when Q/N is not close to 1, $F(Q) = Q(1+\frac{1}{2}(Q/N)+o(Q/N))$, so again $c(F) \approx Q$.

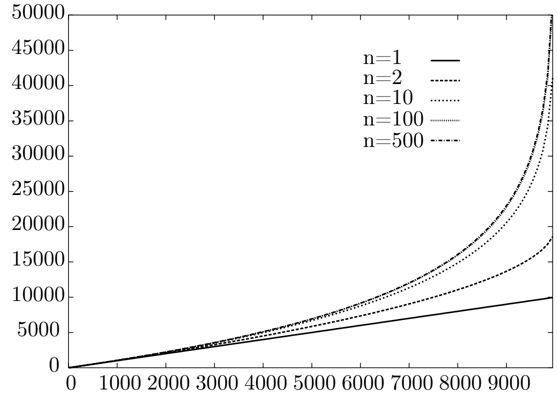


Figure 4.1: $nf(Q)$ for $N = 10000$; $n = 1,\ 2,\ 10,\ 100,\ 500$

We make this simplifying assumption when necessary, which is not crucial for the qualitative results we obtain. Otherwise we can compose the functions $u(Q)$ and $Q(F)$ into another utility function where our public good is F (as we do in Section 4.5 below).

Both of the above lead to models that are covered by the results of the previous chapter. Before analyzing the corresponding incomplete information model for the case of file sharing we first show the inefficiency of equilibrium and compute the lindahl prices for the complete information case.

4.2 Inefficiency of Equilibrium

It is a standard observation that when there are externalities present, and without any explicit incentive mechanism in place, the equilibrium of the public good model proposed will be inefficient (see e.g., Mas-Colell et al. (1995)). This observation holds irrespective of the information structure of the game (i.e., whether peers' valuations are private information or common knowledge).

The reason is the number of files f_i that each peer will choose to share will be the one that maximizes his own net benefit ignoring the positive externalities imposed to other peers (f_i would increase the quantity Q and thus the net benefit of all other peers). In the extreme case where peer i acquires no benefit from his own files f_i, each peer would set $f_i = 0$.

It is simplest to illustrate inefficiency in the complete information game. Each peer chooses the number of her shared files to maximize her own net benefit:

$$\max_{f_i \in \mathbb{R}} b_i(\theta_i, f_i, f_{-i}). \tag{4.5}$$

In contrast, efficiency requires that each agent chooses the number of her shared files to maximize the total net benefit:

$$\max_{f_i \in \mathbb{R}} \sum_{j=1}^{n} b_j(\theta_j, f_j, f_{-j}). \tag{4.6}$$

Let the number of files shared by peer i in equilibrium in this case be denoted f_i^*. Then notice that the resulting amount of distinct files Q^*, where $Q^* = Q(\sum_{i=1}^{n} f_i^*)$, will be the same with the result of the program (3.2). That is, it would result to the first-best provision as defined in the previous chapter.

The difference between the two choices for peer i is represented by the externality term

$$E_i \equiv \sum_{j \neq i} b_j(\theta_j, f_j, f_{-j}). \tag{4.7}$$

Because E_i does not appear in peer i's maximization problem (and this holds for any i), equilibrium will not be efficient. This leads naturally to the question: what policy can be used to correct the inefficiency of equilibrium?

4.3 Lindahl Pricing

The standard economic approach to the problem of externalities is to establish prices that close the gap between private and social incentives. It is simplest to show what is necessary when each peer i's utility is continuously differentiable in its own shared files, f_i and the aggregator function of other peers' shared files, f_{-i}. From the previous section, it is clear that the relevant price for peer i is

$$p_i \equiv \sum_{j \neq i} \frac{\partial b_j(\theta_j, f_j, f_{-j})}{\partial f_i} \tag{4.8}$$

so that peer i is charged p_i for each file that it shares. This price may be positive or negative, depending on the form of externalities involved. In our case it would be positive (i.e. peers are paid for the files they share). So, the payoff function of peer i will be now $b_i(\theta_i, f_i, f_{-i}) + p_i f_i$. The price p_i is exactly the total externality imposed by peer i on all other peers and it is easy to see that using these prices in equilibrium each peer will decide to share f_i^* files as computed by equation (4.6).

Notice that the resulting prices are *non-uniform*: in general, $p_i \neq p_j$ for $i \neq j$. This is an instance of *Lindahl* pricing. The Lindahl prices are personalized prices: potentially there is a different price for each peer in the system. With ordinary private goods, agents face the same price for the good; they then choose the quantity they wish to consume, and these quantities will differ between agents. With a pure public good, such as here, the opposite situation arises: everyone consumes the same quantity of the good, but agents' prices are different.

So, in our case of a p2p system, the Lindahl prices require that all peers' valuations for content availability are known and different prices being computed for different peers. Instead of using prices, one could also consider the use of system rules enforcing directly the resulting equilibrium choices of peers. That is, rules requiring from each peer i to share f_i^* files in order to participate. Notice that there rules are again in general personalized and require complete information in order to be computed.

4.4 Incomplete Information

In the previous chapter we explained the difficulties that arise when the private information of peers (in our case their preference parameter θ_i) is not known to the system designer. We showed however that, when the distribution of the preference parameters is known and exclusions are possible, employing a fixed contribution scheme dictating a minimum number of files each peer should share to participate would asymptotically lead to the second-best efficiency (the best one can do under incomplete information).

Using the payoff function (4.1) instead of (3.1), we can directly use the results presented in Section 3.5, replacing price p_i with the number of files f_i each peer will be required to share and defining the amount of public good Q as the total effective number of files shared (see also Section 4.1.1). So, using Theorem 1, in order to attain to within $O(1/n)$ of the optimum (the second-best) we have to choose θ and F so as to

$$\begin{aligned} &\underset{F,\theta}{\text{maximize}}\;\; nu(Q(F)) \int_{\theta}^{1} \eta h(\eta)\, d\eta - c(F) \qquad (4.9) \\ &\text{s.t. } n[1 - H(\theta)]\theta u(Q(F)) - c(F) \geq 0\,. \end{aligned}$$

where F is the total number of files that must be provided for achieving an effective number of shared files Q.

This program has a very simple interpretation. It sets two variables: the total effective number of shared files $Q^* = Q(F^*)$ (i.e., the argument of the utility function $u(\cdot)$); and the type of the marginal peer, θ^*, who is just indifferent between joining the network and not. All peers who join the network are required to contribute $\theta^* u(Q)$ files (i.e., the benefit of the marginal peer). Peers with a payoff parameter greater than θ^* are willing to join; peers with $\theta < \theta^*$ are not. The total contribution is covered by the contributions of the peers which will participate, which are $n(1 - H(\theta^*))$ on the average. The average value of file sharing per peer is $u(Q) \int_{\theta^*}^{1} x dH$.

Hence (4.9) maximizes the expected social welfare over the choice of fixed fee policies. The optimal policy will correspond to the optimal values of the two variables F and θ^*. Moreover, this fee can be paid 'in kind', i.e., without monetary payments since each included peer pays her fee by contributing the same number of files: $F/n(1 - H(\theta))$.

Even though many challenging implementation issues are related to the enforcement of such an incentive mechanism, this result constitutes a significant progress in defining economic efficiency and designing a simple and practical mechanism to provide incentives in a p2p system. Notice for example that there is no need for peers to report their types, as they do in the direct mechanism. Hence, from a potentially complicated scheme (requiring complicated computations, exchange of large amounts of information and payments between peers) to achieve the second-best allocation, as described in Section 3.2, the problem has been simplified to just requiring each peer to submit a minimum number of files to join the p2p network. In Section 4.6 we discuss the most important theoretical and implementation issues that arise in this context, and in Chapter 5 we make one step further proposing a memory-less mechanism which achieves comparable levels efficiency but without requiring complicated accounting mechanisms to be enforced.

Note that there is also the alternative to use averaging over the optimal f_is computed under complete information (i.e. solving equation (4.6)) for a large number of realizations of the distribution H. In the following we formulate a simple example in order to be able to compare numerically all the possible incentive mechanisms under the proposed model and demonstrate the attractiveness of the fixed contribution scheme.

4.5 Comparing economic incentives in p2p systems

We will now formulate a specific simple model of file sharing in order to make more concrete the above discussion and be able to compare the complete and incomplete information cases. Each peer i decides on the number of files f_i it shares. We remind that the distribution function H from which θ_i takes values, has support normalised to the unit interval $[0, 1]$. The payoff of peer i from the p2p system with n peers and an n-vector $\boldsymbol{f} = (f_1, \ldots, f_n)$ of shared files is

$$b_i(Q) = \theta_i u \left(\sum_{j=1}^{n} \sqrt{f_j} \right) - f_i. \tag{4.10}$$

In our simple example the argument of the utility function is the sum of the square roots of the files shared by each peer. That is, $Q(F) = \sqrt{F}$. As explained in Section 4.1.1, this models file duplication: a contribution of f_i files by peer i results in a 'useful' amount of content which is less than f_i, since some of these files are already made available by other peers. Again, this assumption simplifies calculations, but is otherwise not crucial.

Peer i seeks to maximize its payoff: $\max_{f_i} b_i(\boldsymbol{f})$, taking as given the files shared by all other peers. A Nash equilibrium is comprised of a vector of files simultaneously maximizing their payoffs, for this specific vector. The first-order condition for peer i's maximisation problem is

$$\frac{\theta_i}{2\sqrt{f_i}} u' \left(\sum_{j=1}^{n} \sqrt{f_j} \right) - 1 \leq 0 \qquad f_i \quad \geq 0 \tag{4.11}$$

where $u'(\cdot)$ denotes the derivative of the utility function.

Let the equilibrium choice of peer i be denoted $\hat{f}_i$. In equilibrium, only peers with high enough values of θ_i will make strictly positive contributions. Let the payoff parameter of the marginal peer i.e., the peer with the lowest value of θ_i out of those that make positive contributions, be denoted θ_e where $e \in \{1, \ldots, n\}$; and let $E \equiv \{e, \ldots, n\}$. For those peers who make positive contributions i.e., with $i \in E$, the equilibrium contribution is

$$\hat{f}_i = \left(\frac{\theta_i}{2} u' \left(\sum_{j=e}^{n} \sqrt{\hat{f}_j} \right) \right)^2 . \tag{4.12}$$

Hence

$$\sum_{j=e}^{n} \sqrt{\hat{f}_j} = \frac{\Theta_e}{2} u' \left(\sum_{j=e}^{n} \sqrt{\hat{f}_j} \right) \tag{4.13}$$

where $\Theta_e \equiv \sum_{j=e}^{n} \theta_j$. Equation (4.13) is an implicit equation for the variable $\hat{Q} \equiv \sum_{j=e}^{n} \sqrt{\hat{f}_j}$; with the concavity assumption on $u(\cdot)$, there is a unique solution for $\hat{Q}$ (which, it should be noted, depends on the value of Θ_e—we return to this point below), and hence a unique value $\hat{\boldsymbol{f}}$ for the n-vector of equilibrium contributions.

Finally, the identity of the marginal peer is determined by the indifference condition

$$\theta_e u(\hat{Q}) - \hat{f}_e = 0. \tag{4.14}$$

Due to integer constraints, this indifference condition may not hold exactly; if it does not, the marginal peer is identified as the 'last' peer with a non-negative payoff.

In summary: in the Nash equilibrium, $n - e$ out of the n peers—those with the highest payoff parameters—make strictly positive contributions, given by equation (4.12); the others share no files. The total number of files shared in equilibrium is given by the solution to equation (4.13).

Contrast this characterisation of the Nash equilibrium with the solution that arises when the system is run by a benevolent and fully-informed manager who chooses the contributions of peers to maximize the total payoff of all peers as follows:

$$\max_{\{f_1, \ldots, f_n\}} \sum_{i=1}^{n} \left(\theta_i u \left(\sum_{j=1}^{n} \sqrt{f_j} \right) - f_i \right) .$$

The first-order condition for peer i is

$$\frac{\theta_i}{2\sqrt{f_i}} u' \left(\sum_{j=1}^{n} \sqrt{f_j} \right) - 1 + \frac{\sum_{j \neq i} \theta_j}{2\sqrt{f_i}} u' \left(\sum_{j=1}^{n} \sqrt{f_j} \right) \leq 0 \qquad f_i \quad \geq 0. \tag{4.15}$$

Denote the resulting number of shared files f_i^* for $i \in \{1, \ldots, n\}$. Note that in this solution, if it is optimal for any peer to contribute files, then it is optimal for all peers to share files (since the marginal costs and benefits of file sharing are the same for all peers).

An important difference between equations (4.11) and (4.15) is the presence in the latter of externalities, which can be measured by

$$\frac{\sum_{j \neq i} \theta_j}{2\sqrt{f_i}} u' \left(\sum_{j=1}^{n} \sqrt{f_j} \right)$$

for peer i. In the next section, we analyse specific approaches to ensure that peers consider these externalities when deciding how many files to share.

The contribution of each peer is

$$f^* = \left(\frac{\Theta}{2} u'(Q^*) \right)^2 \tag{4.16}$$

where $\Theta \equiv \sum_{j=1}^{n} \theta_j$ and $Q^* \equiv n\sqrt{f^*}$. This gives the implicit equation for the total number of unique shared files in this case:

$$Q^* = \frac{n\Theta}{2} u'(Q^*). \tag{4.17}$$

(There is a unique solution to this equation, due to the strict concavity of $u(\cdot)$.) Comparison of equation (4.17) with equation (4.13) shows that $Q^* \geq \hat{Q}$, since $\Theta \geq \Theta_e$ and $n \geq 1$.

In summary: in the 'social optimum', in which every peer is concerned about the payoffs of all peers in the system, each peer shares f^* files, where $f^* \geq \max_i \hat{f}_i$; consequently the total number of files shared in the social optimum is greater than in equilibrium. This means that the total payoff attained in the social optimum, denoted S_{SO}:

$$S_{SO} \equiv \sum_{i=1}^{n} \left(\theta_i u \left(\sum_{j=1}^{n} \sqrt{f_j^*} \right) - f_i^* \right)$$

is greater than the total payoff attained in equilibrium, denoted S_{NE}:

$$S_{NE} \equiv \sum_{i=1}^{n} \left(\theta_i u \left(\sum_{j=1}^{n} \sqrt{\hat{f}_j} \right) - \hat{f}_i \right);$$

i.e., $S_{SO} \geq S_{NE}$. In short, equilibrium is inefficient.

4.5.1 Incentive Schemes

In this section, we derive analytical expressions for a number of incentives schemes that might be used to correct the externalities and resulting inefficiencies identified in the previous section. We start by considering the 'first-best', when the scheme designer[2] has complete and perfect information about the payoff parameters of all peers and is able to set personalised incentives for each peer. We then move on to less ideal situations, in which participation incentives must be given; personalisation is not possible and information is incomplete.

The first-best

The simplest method of ensuring full efficiency of the peer-to-peer system is to require each peer to share f^* files. In summary:

Scheme 1 (First-best rule) *Each peer shares f^* files, where f^* is given by equation* (4.16)*.*

An equivalent approach uses prices to give peers the correct incentives (see Section 4.3 above). In the simple case of file sharing that we are considering, the prices are *subsidies* paid[3] to peers to encourage them to share files. If negative externalities (e.g., congestion) were present, then payments might occur in the opposite direction: the peers would be charged to discourage the activity generating negative externalities (see Chapter 6). By comparing equations (4.11) and (4.15), it is apparent that the appropriate price for peer i

[2]In our view, as already discussed in Section 2.4.1, the scheme designer coincides with the designer of the software of a p2p application and does not need to be implemented by a physical entity participating in the actual operation of the system.

[3]Such subsidies may be paid by a third party who indirectly benefits from the efficient operation of the system. Although this is a standard procedure in the provision of public goods, it may be impractical in many realistic p2p contexts where such a payment system is hard to implement. In any case we use it for comparison purposes.

is

$$p_i \equiv \frac{\sum_{j\neq i}\theta_j}{2\sqrt{f^*}}u'(Q^*);$$

substituting in the expressions for f^* and Q^*, this reduces to the very simple form

$$p_i = \frac{\Theta_{-i}}{\Theta} \tag{4.18}$$

where $\Theta_{-i} \equiv \sum_{j\neq i}\theta_j$. This leads to

Scheme 2 (First-best prices) *Peer i is paid a price p_i per file shared, where p_i is given by equation* (4.18).

There are three facts to note about the first-best schemes. First, both use complete information: the rule and prices depend on the vector $\boldsymbol{\theta}$ of payoff parameters.[4] Secondly, the rule may involve some peers (those with low values of θ) receiving a negative payoff from being part of the system. This raises the issue of whether those peers can be forced to be part of the system, or whether they must be given incentives to join. Thirdly, as already mentioned, the first-best prices are personalised: each peer faces a different subsidy for each file that it shares. The first-best rule, on the other hand, is not personalised—each peer shares the same number of files (but note that this is not the case for rules with participation incentives as we discuss next).

Complete information rules with participation incentives

In this section, we consider rules that use information about peers' payoffs and provide incentives for peers to join the system. The problem faced by the manager is to choose contribution levels to maximise total payoffs:

$$\max_{\{f_1,\ldots,f_n\}} \sum_i \left(\theta_i u\left(\sum_{j=1}^n \sqrt{f_j}\right) - f_i\right)$$

[4] It is an interesting feature of the solution that information about the utility function is not always required. The first-best prices do not depend on the form of the utility function $u(\cdot)$. This feature arises because of the way in which peer heterogeneity appears in the problem: the payoff parameter multiplies a function which is the same for all peers. The first-best rule does, however, depend on the form of $u(\cdot)$.

subject to the constraints that each peer must receive a non-negative payoff:

$$\theta_i u\left(\sum_{j=1}^{n} \sqrt{f_j}\right) \geq f_i,\ \forall i. \tag{4.19}$$

The Lagrangian for this constrained maximisation problem is

$$L \equiv \sum_i \theta_i (1+\lambda_i) u\left(\sum_{j=1}^{n} \sqrt{f_j}\right) - \sum_i (1+\lambda_i) f_i. \tag{4.20}$$

The first-order conditions are

$$\sqrt{f}_i = \frac{\sum_j \theta_j \rho_j u'\left(\sum_{l=1}^{n} \sqrt{f_l}\right)}{2\rho_i} \tag{4.21}$$

where $\rho_i \equiv 1 + \lambda_i$. In these equations the λ_i are non-negative. They are also decreasing for the following reason. For the is where the participation constraint (4.19) binds, we have that $\lambda_i > 0$ and also that $f_i = \theta_i u(\cdot)$. Combining this with equation (4.21), we obtain that $\theta_i(1+\lambda_i)^2$ must be constant. Since we assumed that $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_n$, we must also have that the λ_is are decreasing in i.

The above suggests that rules take the following form: for some threshold k,

$$\tilde{f}_i = \begin{cases} \theta_i u(\tilde{Q}) & i < k, \\ \theta_k u(\tilde{Q}) & i \geq k \end{cases} \tag{4.22}$$

where $\tilde{F} \equiv \sum_{j=1}^{n} \sqrt{\tilde{f}_j}$. Given these levels of contributions, the threshold k can then be chosen to maximise total payoffs.

Hence we have the following incentive scheme:

Scheme 3 (Rules with participation incentives) *Peer i shares $\tilde{f}_i$ files, where $\tilde{f}_i$ is given by equation* (4.22)*.*

Non-personalised, complete information prices

Now suppose that we have complete information about the payoff parameters of the peers, but we are unable to establish their identities—so that personalisation of an incentive scheme is not possible. Another reason is that implementing a personalized incentive scheme may be too costly or infeasible. In this case, the system designer could not implement first-best prices or rules (with participation incentives). However, in the case of prices there are several ways to calculate a uniform price. We shall assess the following:

$$p = 1 - \frac{\max\{\theta_1, \ldots, \theta_n\}}{\Theta} \tag{4.23}$$

i.e., the personalised, complete information (first-best) price received by the peer with the highest payoff parameter. The numerical analysis in Section 4.5.2 confirms that all peers wish to participate given this price. [5]

Scheme 4 (Non-personalised price) *Each peer is paid a price p per file shared, where p is given by equation* (4.23).

Incomplete information schemes

We now turn to the case in which there is incomplete information about the peers' payoff parameters. To be specific: the number of peers n is known, as is the distribution from which the n peers' payoff parameters are drawn (they are i.i.d. random variables with distribution H) and the form of the utility function $u(\cdot)$. But only peer i knows the realisation of its payoff parameter θ_i; peer i does not know the payoff parameter of peer j, and we assume that the agent who designs the incentive scheme observes none of the payoff parameters.

This information structure rules out immediately all of the schemes considered above. We will consider adaptations of these schemes which will be compared to a scheme that is derived from full consideration of the incomplete information problem (i.e. the fixed contribution mechanism introduced in Section 4.4 above).

The first-best rule "share f^* files" depends on the realised $\boldsymbol{\theta}$, since f^* depends on Θ (see equation (4.16)). Making this dependence explicit by writing $f^*(\Theta)$, an average rule can be computed as

$$\bar{f} \equiv \int \cdots \int f^*(\Theta) dH^n(\theta)\,, \tag{4.24}$$

where H^n is the probability distribution of the random vector $\boldsymbol{\theta}$. An example helps to

[5] The choice of another definition for a uniform price, e.g., using instead of the maximum, the average of the θ_is, or their minimum, does not affect the qualitative results presented.

make this clear. Suppose that the utility function $u(\cdot)$ is iso-elastic: $u(x) = x^\alpha$ where $\alpha \in (0,1)$. Then

$$f^*(\theta) = \left(\frac{\alpha\Theta}{2} n^{\alpha-1}\right)^{\frac{2}{2-\alpha}} ,$$

and

$$\bar{f} \equiv \int \left(\frac{\alpha\Theta}{2} n^{\alpha-1}\right)^{\frac{2}{2-\alpha}} dH^n(\theta) ,$$

which can be calculated explicitly by making an assumption about the distribution of payoff parameters (e.g., independently drawn from the uniform distribution on $[0,1]$), or by using simulation.

Scheme 5 (Average rule) *Each peer shares $\bar{f}$ files, where $\bar{f}$ is given by equation* (4.24)*.*

Of course, it may be that some peers would receive a negative payoff if they contributed $\bar{f}$ files; hence under this rule, they will not take part in the network. This will affect any calculation of the efficiency achieved under this scheme in the simulation assessment of Section 4.5.2 below.

A similar procedure can be applied to prices to compute an average price:

$$\bar{p} \equiv \int \cdots \int \frac{\Theta_{-i}}{\Theta} dH^n(\theta). \tag{4.25}$$

Scheme 6 (Average price) *Each peer is paid a price $\bar{p}$ per file shared, where $\bar{p}$ is given by equation* (4.25)*.*

Notice that the average rule depends on both the functional form of $u(\cdot)$ and the distribution from which peers' payoff parameters are drawn. The average price depends only on the latter.

These adaptations of the previous schemes do not take into account explicitly additional factors that arise in the presence of incomplete information. In this case, it is likely that the scheme can be improved (in terms of the system efficiency achieved) by eliciting information from peers. A very simple and attractive mechanism to do so is presented in Section 3.5 and adapted in the case of file sharing in Section 4.4. Based on (4.10), we can reduce (4.9) into a simpler form. First observe that since all participating peers contribute equally some amount f, $F = m\sqrt{f}$, where m is the number of final participants. But also $c(F) = mf$, and hence $c(F) = F^2/m$. By doing integration by parts and using the cost constraint we obtain the equivalent program

$$\underset{F,\bar{\theta}}{\text{maximize}}\, u(\Phi)\int_{\bar{\theta}}^{1}(1-H(x))dx \tag{4.26}$$
$$\text{s.t. } n(1-H(\bar{\theta}))\bar{\theta}u(F)=\frac{F^2}{n(1-H(\bar{\theta}))}$$

where we substituted m by its average $n(1-H(\bar{\theta}))$.

Denote the expected total payoff from this program S_{CW}. According to Courcoubetis and Weber (2005), the difference between S_{CW} and the second-best is $O(1/n)$, and so becomes negligible as n becomes very large. This result motivates us to consider the following, 'fixed fee' scheme:

Scheme 7 (Fixed fee) *All peers who join the network are required to share $\bar{\theta}u(F)$ files, where $\bar{\theta}$ and F are given by the solution to the program* (4.26)*.*

4.5.2 Simulation Assessment of Incentive Schemes

In previous sections, we have defined seven incentive schemes:

1. first-best rule;
2. first-best prices;
3. rules with participation incentives;
4. non-personalised price;
5. average rule;
6. average price;
7. fixed fee.

In this section, we calculate the efficiency of these schemes for a specific functional form for the utility function: $u(x)=x^{\alpha}$, where $\alpha=0.5$; and a specific distribution function for the peers' payoff parameters (uniform on the unit interval, with iid draws). The objective is to assess the performance of the different schemes, in terms of the total payoff that they yield, as the number of peers increases. Then, in order to show the robustness of our results, we use a lognormal distribution, and we vary its variance to investigate how peer heterogeneity may affect the results.

The procedure we use is as follows:

1. Fix the number of peers at n.
2. Calculate the average rule and price in schemes 5 and 6[6].
3. Draw n values $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$ randomly from the specified distribution.
4. For this realisation of peers' valuations, calculate rules/prices in schemes 1–4.
5. For this realisation of peers' valuations, calculate total payoffs for all schemes.
6. Return to step 3 and repeat 100 times.
7. Average the total payoffs achieved over the draws of $\boldsymbol{\theta}$.
8. Increase the number of peers by 1 and return to step 2.

The outcome of the simulations is summarised in Figures 4.2 and 4.3. In these figures, the total payoff achieved by each scheme, averaged over realisations of peers' payoff parameters, is expressed as a proportion of the first-best total payoff, as the number of peers varies. So, the first-best rule or price would be represented by a flat line at 1 for all network sizes. As expected, all other schemes return a total payoff strictly less than the first-best level; hence the lines for these schemes lie strictly below 1. Figure 4.2 confirms the observation of Section 4.2: without an incentive scheme, the externalities that exist between peers leads to inefficiency. The figure shows the total payoff in the Nash equilibrium, as a proportion of the first-best level. In a small network (10 peers), the degree of inefficiency is marked: the Nash equilibrium achieves only 60% of the first-best total payoff. As the network grows, the inefficiency becomes worse and worse; by the time the network includes 100 peers, the Nash equilibrium total payoff is just 30% of the first-best. This illustrates the intuitive property that free-riding is worse when there are many agents. Each peer anticipates that she has little effect on the total number of files shared when the network is large. Consequently the incentive to contribute nothing and rely only on the contributions of others grows with network size.

[6] since these do not depend on the realization of the particular θs.

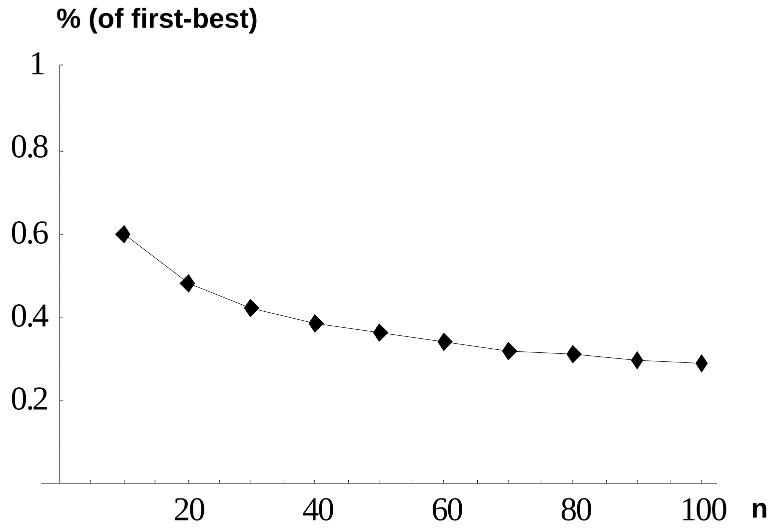


Figure 4.2: Total payoffs in the Nash equilibrium

There are several features that should be noted from Figure 4.3. First, the rules with participation incentives achieve close to full efficiency even for moderate network sizes (20 peers or more). This indicates that (within this set of calculations, at least) there is little efficiency loss arising from the need to give participation incentives to peers. The intuition for this result is that as the system gets larger, the participation constraint becomes easier to satisfy. This is because the value of the shared content to peers increases and peers with the same θ are willing to contribute more in order to participate. For very large n, a very small fraction of peers will be reluctant to pay the fixed fee contributed by the rest of the peers as defined by the optimal policy. Hence the participation constraint of the optimization problem becomes irrelevant, and the solution converges to the first-best. The fixed fee scheme yields a strictly lower level of utility. This demonstrates the Myerson-Satterthwaite result: incomplete information leads to efficiency loss. (A major difference between schemes 3 and 7 is, of course, that the former uses full information about peers' payoff parameters, which is not available for the latter.) The efficiency loss is relatively small, however: around 6–7%. An intuition for this can be gained by considering the sources of efficiency loss. One is that participating peers do not contribute the efficient amount of files (f^*); another is that some peers are excluded. The size of the second inefficiency is limited, however, by the fact that it is peers with low payoff parameters that are excluded; and they make little difference to the overall level of efficiency. Further,

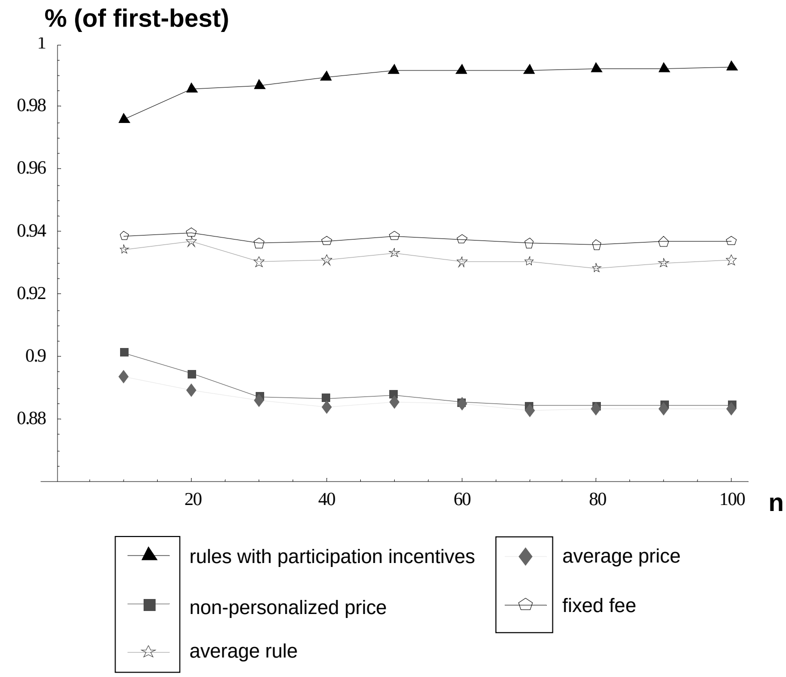
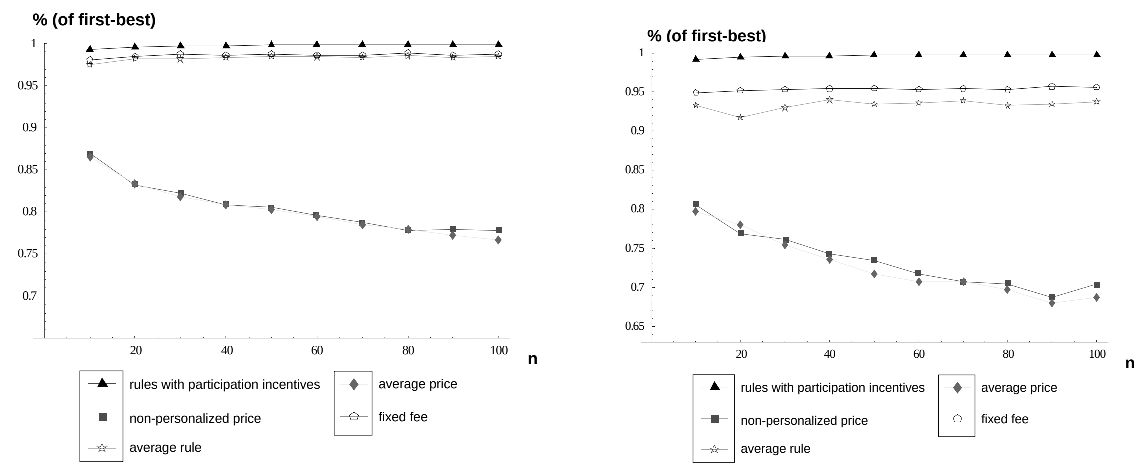


Figure 4.3: Comparison of incentive schemes

from the results of Norman (2004) and M. F. Hellwig (2003), as the number of peers becomes large, the fixed fee schemes approaches the full second-best mechanism. A major reason for this is that calculations based on averages or expectations—like the calculation of the fixed fee—become more and more accurate as the number of peers becomes large, for standard statistical reasons. There is a second, more subtle argument as we discussed in the previous section, which explains why, although in the full second-best problem, incentive compatibility constraints must be included—see equation (3.6), these incentive constraints are effectively ignored in the calculation of the fixed fee. Hence the numerical results indicate that these constraints become less important (in efficiency terms) as the number of peers grows. Since our fixed fee policy becomes asymptotically second-best optimal it is also optimal among the set of all fixed fee policies.

Moreover, the average rule scheme (number 5) does well for efficiency, but appears systematically to yield a strictly lower total payoff than schemes 3 and 7. This indicates that the approximation used in this scheme (taking a straight expectation of the first-best rule) is inferior to the fixed fee approximation. This is to be expected: the fixed fee method derives an instrument to maximise average payoff; the approximation in scheme 5 averages over an instrument that maximises total payoff under complete information. The former is better suited to the situation in which there is incomplete information—a fact reflected in the figure.

Finally, prices of any flavour (other than first-best) yield lower levels of efficiency than the other schemes. The degree to which they under-perform appears to be driven by the specification of our model: in particular, the linear "cost" term in equation (4.10). With this specification, small errors between the first-best price and an approximate price are translated into very large differences in the number of files contributed by peers. Hence the resulting efficiency level of approximate (non-personalised or average) prices are low. With alternative cost specifications (especially cost functions that are strictly convex), the performance of approximate price schemes improves.[7] In fact, in a variety of such models with convex cost functions, as n increases, uniform prices perform close to the first-best. In the next results we use a lognormal distribution of the peer payoffs. The reason is to



(a) $(\mu, \sigma) = (0.5, 0.6)$

(b) $(\mu, \sigma) = (0.5, 0.8)$

Figure 4.4: Comparison of incentive schemes using the lognormal distribution

investigate the effects of the variance of the θs to the performance of the various schemes. The results are in Figure 4.4, for two different values of the variance of the distribution. These results can be easily explained. Smaller variance makes uniform rules more effective

[7]This observation raises the question: why choose a quasi-linear payoff function? The reason is that this allows us to match the model directly to the public good provision models used by M. F. Hellwig (2003) and Norman (2004), and hence to apply their results in our context.

since nformation loss is less important while larger variance has a negative effect on the resulting efficiency. This suggests that if peers can be better classified in terms of their payoffs (by using some objective characteristic like the access speed of their modem) in groups having a higher degree of homogeneity, then simple fixed fee schemes for each group will perform better than having all peers in a single group (see Section 4.6.1).

4.6 Fixed contribution scheme

In this section we discuss several practical issues associated with the employment of the fixed contribution scheme in a realistic p2p environment, such as the advantages and incentive issues of group formation, dealing with heterogeneous file popularity, ensuring the stability of the system, discovering system parameters that are assumed known (such as the size and distribution H), and more. We don't provide complete solutions to all these challenging issues. However, the simplicity of our model allows us to study several interesting dimensions and gain some useful insights.

For our analysis in this section we will use the approximations presented in Section 4.1.1 and we will assume that $Q(F) \approx F$. That is the social planner will now have to solve the following simplification of (4.9) in order to compute the fixed contribution required from participating peers

$$\begin{aligned} \underset{Q,\,\theta}{\text{maximize}} \;\; & nu(F)\int_{\theta}^{1} \eta h(\eta)\, d\eta - F \\ \text{s.t.} \;\; & n[1-H(\theta)]\theta u(F) - F \geq 0\,, \end{aligned} \tag{4.27}$$

where $u(F) = F^a$, with $0 < a < 1$, and $F = \sum_{i=1}^{n} f_i$. Again, this assumption is not crucial for the qualitative results we obtain.

4.6.1 Group formation

The simplicity of the fixed contribution scheme comes at a cost. As we have seen in the previous section, as heterogeneity increases, the distance between the efficiency achieved by the fixed contribution scheme and the first-best also increases. However, it may sometimes be possible for the global planner to distinguish between types of peers and use this information to model the distributions of their preference parameters more accurately. Suppose, for example, that the population of peers consists of both ISDN dial-up users (group A)

and DSL users (group B) with uniform distributions H_A and H_B of their preference parameters on $[0, \beta]$ and $[\beta, 1]$ respectively, with $\beta < 1$. This reflects the fact that DSL users value more and benefit more from the shared content than do the dial-up users. It is clear that when this is possible, it would be for the benefit of the system to differentiate the fixed contributions of the users belonging to different groups.

As we have seen, the asymptotic problem that must be solved to determine the optimal fixed fees is, fortunately, quite simple. We can thus gain, through numerical analysis, some important insights into the formation of groups. There are three possible scenarios that the planner could pursue:

1. dial-up users and DSL users form *distinct groups* A and B and do not share content between them.

2. they form a single group, in which they share content, but their type distributions (H_A and H_B) are *indistinguishable* to the global planner. He only knows the initial proportions of users of each type.

3. they form again a single group, in which they share content, except that their type distributions are now *distinguishable*, i.e., peers disclose their type distribution to the planner.

Which of these is preferable for each user type? Does the answer depend on the relative number of the users of each type? In order to gain some interesting insights towards answering these questions we computed the following levels of efficiency. We first solved (4.6) for $H = H_A, n = n_A$ and $H = H_B, n = n_B$ to estimate the efficiency achieved by each group when they choose not to share files between them, where n_A and n_B are the number of peers belonging to each group respectively. Then we solved (4.6) assuming that the two groups form a larger group but without distinguishing between peers belonging to different groups. In this case the distribution H will be a weighted combination of two uniform distributions, with different weights on the intervals $[0, \beta]$ and $[\beta, 1]$ ($n_A/(n_A + n_B)$ and $n_B/(n_A + n_B)$ respectively). We finally computed the solution of the following program which takes into account the fact that the types of peers can be distinguished.

$$\underset{F,\theta_A\,\theta_B}{\text{maximize}}\ n_A u(F)\int_{\theta_A}^{1}\eta h_A(\eta)\,d\eta + n_B u(F)\int_{\theta_B}^{1}\eta h_B(\eta)\,d\eta - F \qquad (4.28)$$

$$\text{s.t. } n_A[1-H_A(\theta_A)]\theta_A u(F) + n_B[1-H_B(\theta_B)]\theta_B u(F) - F \geq 0\,, \qquad (4.29)$$

where $u(F) = F^a$, with $0 < a < 1$, and θ_A, θ_B denote the type of the marginal peer for each group.

In Figure 4.5 the values of social welfare for each group separately and in total and for each scenario are depicted, for the simplest case where $\alpha = 0.5$, $\beta = 0.5$, and $n_A = n_B$. We see that the users have always the incentive to form a larger group but the overall utility is maximized when the type distribution of different peers is distinguishable. However, in this case the DSL users (group B) would prefer to form a distinct group since they incur a welfare loss of 11% by agreeing to participate in the larger group declaring their type. In Figure 4.6 we plot the ratio of the welfare of dial-up users, group A (Figure 4.6(a)),

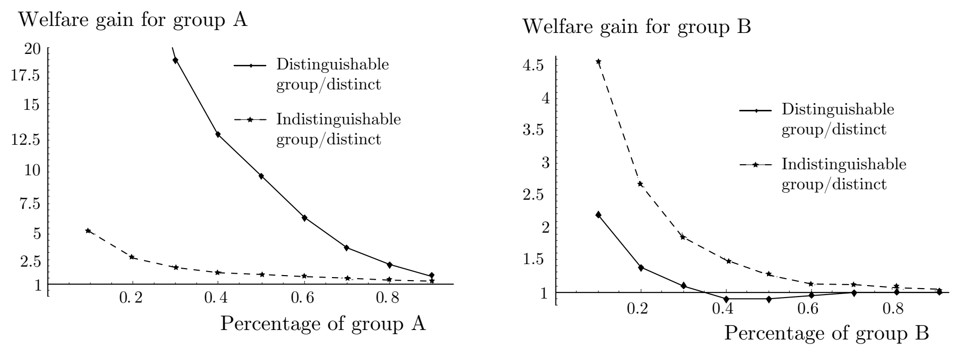
Welfare	Group A	Group B	Total
Distinct groups	3296	35156	38452
One group, indistinguishable	6976 (+ 111%)	44792 (+ 27%)	51768
One group, distinguishable	31249 (+ 848%)	31250 (-11%)	62500

Figure 4.5: Social welfare achieved in the different group forming scenarios for $\alpha = 0.5$, $\beta = 0.5$, and $n_A = n_B = 500$

and DSL users, group B (Figure 4.6(b)), with the welfare they obtain when they form distinct groups, when groups are distinguishable and indistinguishable, as a function of their percentage in the group ($n_A/(n_A + n_B)$ and $n_B/(n_A + n_B)$ respectively). There are two interesting observations to be made:

- As the proportion of DSL users decreases, the DSL users prefer more the second scenario: a large indistinguishable group; their next preferred option is the first scenario, when they form their own distinct group. However, this is not so when their percentage in the group is small (below 30%). Then they still prefer the large indistinguishable group but their second option is the large distinguishable group.

- As the proportion of dial-up users decreases, they favour the third scenario since at the social welfare optimum, the DSL users offer the majority of the content. The second scenario, of a single indistinguishable group, is not so attractive as they are forced to pay a substantial fixed fee. When their number is small they gain by the large amount of content made available by the DSL users, compared to the content they would obtain in their own group. This difference becomes negligible when they are the dominant type in the mixture.



(a) Group A (b) Group B

Figure 4.6: Welfare gain of dial-up (group A) and DSL (group B) users (%) when forming groups in which they are distinguishable and indistinguishable to the planner compared to the welfare they obtain when they form distinct groups.

So, unfortunately, the distinguishable group scenario is never the desired option for both groups even if it achieves the maximum efficiency. The underlying reason is rather simple: when forming a larger indistinguishable group, DSL users benefit from the larger content available and the fact that the fixed fee, which is the same for all, is less than the one they would pay if they could be distinguished. Furthermore, dial-up users who can afford it contribute a substantial amount to the common cost, whereas if they can be distinguished, they free-ride on the contribution of the DSL users. In other words, when more information is available, the total welfare increases, but the surplus of the DSL users decreases because they are forced to contribute a larger share of the cost (indeed all the cost!): when the percentage of DSL users is above a threshold, assigning a zero fee to the dial-up users maximizes their participation and hence the social welfare —see also the analysis for best group splitting below. In this case they rather prefer to form their own group and would refuse to join such a larger group if they are to be distinguished. The situation is different when the percentage of the DSL users is small. Then the dial-up users should also contribute a significant amount of content in order for social welfare to be maximized and thus DSL users benefit from forming a larger group; but note that again they prefer not to be distinguishable. We discuss in the following possible ways to create the necessary incentives to users to participate in the large group when their type is observable or declare it truthfully when it is not and the corresponding trade-offs.

From our numerical analysis is also deduced that the preference for the first scenario (forming a distinct group) becomes even more marked for DSL users as the distribution of the preferences becomes less spread, when β is greater (e.g., uniform on $[0.8, 1]$, rather than on $[0.5, 1]$); conversely, as it becomes more spread they tend to prefer the second scenario (forming a large indistinguishable group).

Finally, we so far have assumed that it is equally costly for DSL and dial-up users to share a file. This may be a reasonable approximation when the up-link speed of DSL users is small. But other times this may not be the case. Suppose that our two user types both have θs distributed uniformly in $[0, 1]$, but their costs for sharing f files are af and f respectively, $a > 1$. Assuming that the types can be distinguished by the global planner, suppose F_1, F_2, θ_1, θ_2 and solve

$$\max_{F_1, F_2, \theta_1, \theta_2} u(F_1 + F_2) \left\{ n_1 \int_{\theta_1}^{1} y dy + n_2 \int_{\theta_2}^{1} y dy \right\} - aF_1 - F_2 \,, \tag{4.30}$$

subject to the two conditions

$$n_1[1 - H(\theta_1)]\theta_1 u(F_1 + F_2) - aF_1 \geq 0 \,, \tag{4.31}$$

$$n_2[1 - H(\theta_2)]\theta_2 u(F_1 + F_2) - F_2 \geq 0 \,. \tag{4.32}$$

In this case, the fixed fee of $\theta_1 u(F_1 + F_2)$ that is paid by type 1 users should be divided by a to convert it to a number of files that each one of them has to provide. To say who gains by such a distinction we compare the net benefits obtained by peers of both types with the net benefits they obtain in cases that there is a single indistinguishable group or that each type forms its own group. The results of our experiments for this model are similar to those when peers have different distributions of preference parameters. The type 1 peers, for whom file sharing is more costly, benefit when all peers disclose their types, whereas the type 2 peers may not. Type 1 peers are better off both because of the larger content selection and because they contribute less.

Best splitting with average information

We have explored the benefits of group forming when we can categorize peers into two different groups. A natural question is whether it would be still beneficial to further divide peers into more than two groups, defining a different minimum contribution for each one of them. In other words, supposing that the global planner could obtain information that would allow him to partition peers into smaller subgroups, such that the preference parameters in each subgroup are distributed on non-overlapping subintervals of $[0, 1]$, can he always gain by doing this? The answer is that this is not always the case.

Before demonstrating this fact, we should first define the maximum efficiency that could be achieved assuming the maximum possible splitting capabilities. By 'splitting capability' we mean the definition of a specific number of non-overlapping subgroups for which the system designer has the suitable information so as to decide which peers belong to which subgroup based on a specific realization of their preference parameters. It is obvious that if the designer could divide the initial interval into a large number of subintervals and use the actual information from the realizations of the θ_is, i.e. how many peers n_j fall into the jth subinterval, to construct the optimal policy, he could eventually obtain the first best efficiency. An interesting and more practical case is when the designer first decides on the splitting and specifies the fixed contribution for each subinterval, and then the peers join the system and decide on whether to participate. In this case, the actual number of peers that falls in a particular subinterval is not known to the designer when he decides on the optimal policy and thus he must compute the optimal policy for the average case, i.e. using as only knowledge the initial distribution H[8]. In this case we show that splitting after some point does not produce any improvement.

To see that, observe first that since the information available to the designer when computing the fixed fees is the distribution H, the best he can hope for is the maximum

[8]Alternatively, he could define the optimal policy parameterized by the actual number of peers in each subinterval. In this case, first the peers would declare in which interval they fit and then the designer would disclose the fixed fee. As we mentioned, this policy can approximate the first best.

unconstrained average efficiency. That is, the solution of the following program

$$\max_{F \geq 0} \mathbb{E}[\sum_i \theta_i u(F) - c(F)] . \tag{4.33}$$

So, assume the distribution H of peers' preference parameters is uniform on $[0,1]$. This is to be divided in k non overlapping intervals where the end-points of each interval j are denoted by β_{j-1} and β_j. Thus the vector $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)$, where $b_0 = 0$, $b_k = 1$, and $0 < b_1 < \ldots < b_{k-1} < 1$, defines the k sub-intervals of $[0,1]$. Then the average number of peers that will eventually belong to each subgroup j will be $n(\beta_j - \beta_{j-1})$ and their types will be uniform on $[\beta_{j-1}, \beta_j]$.

Based on this information, the system designer has to decide on the fixed fees required by peers belonging to each subgroup (let $\boldsymbol{f} = (f_1, \ldots, f_k)$ be the vector of the required minimum contributions for each subgroup) so as to maximize the average social welfare achieved with the constraint to cover the cost. We will show that there is a minimum number k for which the social welfare of the constrained problem, denoted by $SW(k)$, will be the maximum possible (the solution of the unconstrained problem with complete information (4.33)), denoted by SW^*.

But notice that in order to achieve this level of efficiency no peer should be excluded from the system on the average, as assumed for the computation of SW^*. So, for the computation of $SW(k)$ we could restrict ourselves to the set of contribution rules that allow all peers to participate. That is, we consider a vector $\boldsymbol{f}$ for which

$$f_j = \beta_{j-1} u(F^*) \ , \forall j \in [1, k] , \tag{4.34}$$

where F^* is the value of F for which (4.33) is maximized. Notice that we compute the vector $\boldsymbol{f}$ as if the optimal amount of content F^* is actually provided. Then in order to check whether our assumption was correct, and our solution is feasible, what we would need to check is the condition for cost coverage

$$F^* \leq \sum_{j=1}^{k} n(\beta_{j-1} - \beta_j) f_j , \tag{4.35}$$

where f_j is as defined by (4.34), $n(\beta_{j-1} - \beta_j)$ is the average percentage of peers in the interval $[\beta_{j-1} - \beta_j]$, and $c(F) = F$. So, if (4.35) holds, it suffices to say that the average social welfare can be maximized when peers can be divided in k subgroups.

Before looking into the general case, let us consider a simple example. If $u(F) = F^a$, and the initial distribution is uniform on $[0,1]$, (4.33) is maximized for

$$F^* = \left(n\,a \int_0^1 \eta h(\eta)\,d\eta\right)^{\frac{1}{1-a}} = (0.5\,n\,a)^{\frac{1}{1-a}}. \tag{4.36}$$

We will now need to examine increasing values of k in order to find the minimum for which we can satisfy equation (4.35), choosing appropriately the vector $\boldsymbol{\beta}$. For $k = 2$, and recalling that $\beta_0 = 0$, (4.35) becomes

$$n(1-\beta_1)\beta_1 \geq (F^*)^{1-a} = 0.5\,n\,a. \tag{4.37}$$

So, if there is a β_1 for which $(1-\beta_1)\beta_1 = 0.5\,a$ then $k = 2$ is enough. Obviously, this depends on the value of a. It is easy to see that there is such a β_1 for $a \leq 0.5$ since the maximum value of $(1-\beta_1)\beta_1$, for $0 < \beta_1 < 1$, is 0.25. This means that for $a > 0.5$ we would require more than two groups to be formed. Let us examine whether $k = 3$ is enough. Equation (4.35) now becomes

$$n(\beta_2 - \beta_1)\beta_1 + n(1-\beta_2)\beta_2 \geq 0.5\,n\,a. \tag{4.38}$$

This means that we have to find the values of β_1 and β_2 that maximize the sum $(\beta_2 - \beta_1)\beta_1 + (1-\beta_2)\beta_2$. It is easy to see that this sum is maximized when $\beta_1 = 1/3$ and $\beta_2 = 2/3$; when we divide the interval $[0,1]$ into three equal subintervals with value 1/3. Hence, $k = 3$ is enough when $a \leq 2/3$. Interestingly, for $k = 4$ it turns out that we can find a vector $\boldsymbol{\beta}$ that achieves the average first-best efficiency for any value of a $(0 < a < 1)$.

What this analysis shows is that the ability to divide peers into groups is a very powerful tool and even a very small number of groups is in many cases enough to reach the efficiency of (4.34). And notice that the uniform distribution is a highly heterogeneous distribution. If we assume a less heterogeneous distribution —e.g. a discrete triangular distribution[9] in $[0,1]$, whose mean value is again 0.5— and following the above procedure (with slightly more complicated computations), one can see that for $k = 2$ we can maximize welfare when $a < 0.55$ (for $b_1 = 0.4$) and $k = 3$ is enough when $a < 0.7$ (for $b_1 = 0.3$ and $b_2 = 0.53$). On

[9] $P(X \leq x) = \frac{g(x)(g(x)-1)}{2\,n}$, for $x \leq 0.5$ and $P(X \leq x) = \frac{2\,n - g(1-x)(g(1-x)-1)}{2\,n}$, for $x \geq 0.5$, where $g(x) = [x(1+\sqrt{1+4\,n})] + 1$.

the other hand, when the distribution is more heterogeneous than the uniform one, one would expect that we need more groups (greater values of k) in order to reach the optimum efficiency.

In general, we can show that there is always such a finite k for which efficiency is maximized. To see why this is true, notice first that in order to compute the minimum value of k for different probability distributions we should just need to replace the term $(\beta_{j-1}-\beta_j)$ of our condition (4.35) with the percentage of peers that belong to the corresponding range of preference parameters, and the value of $\int_0^1 \eta h(\eta)\, d\eta$ with the corresponding mean of the new distribution. Now we choose to split peers into groups in a way that each subgroup has the same number of peers, n/k, on the average. Then, rewriting equations (4.34) and (4.36), assuming a generic distribution Z with density function z and support $[x, y]$, (4.35) becomes

$$\frac{n}{k}\sum_{j=1}^{k}\beta_{j-1}u(F^*) \geq F^* \Leftrightarrow \frac{n}{k}\sum_{j=1}^{k}\beta_{j-1} \geq (F^*)^{1-a} = n\, a \int_x^y \eta z(\eta)\, d\eta. \qquad (4.39)$$

Now observe that as k grows, $\frac{1}{k}\sum_{j=1}^{k}\beta_{j-1}$ converges to $\int_x^y \eta z(\eta)\, d\eta$. Hence, for $0 < a < 1$ there is always a k for which (4.39) holds.

Incentive issues

The analysis of best splitting above is encouraging concerning the number of groups required in order to maximize efficiency. However, there are two important obstacles a system designer has to overcome in order to take advantage of the additional information provided by splitting peers into different groups. The first is the ability to identify which peers actually belong to which group. This is not at all trivial and there is in practice a 'physical' limit concerning the splitting capability according to the inherent characteristics of the peers population, as for example the fact that there are some peers with DSL and other with dial-up connections, and the extent to which these are observable by the system designer (e.g. it is technologically feasible to check whether a peer has a DSL or a dial-up connection).

But even in the most convenient case (two well defined groups with observable characteristics), if the high type users have the option to form their own distinct group the uitable incentives should be provided to them in order to agree to participate in the larger group. As already shown the incentives are not aligned when their percentage in the group is significant. For example, when $n_B/n = 50\%$, the optimal allocation results in them losing 11% of the social welfare achieved when they form their own group (see Figure 4.5). Given that their type is observable, a solution would be to change the fee of DSL users to make it attractive for them to join. For example, we could allow the fee of the DSL peers to be lower than $\theta_B^* u(F^*)$, say $xu(F^*)$, where $\theta_A^* < x < \beta$ (recall that $\theta_A \in [0, \beta]$ and $\theta_B \in [\beta, 1]$), thus granting a positive net benefit for all members of this group. So, we have to solve the following program

$$\begin{aligned} &\underset{F,\theta_A,x}{\text{maximize}}\ n_A u(F) \int_{\theta_A}^{\beta} \eta h_A(\eta)\, d\eta + n_B u(F) \int_{\beta}^{1} \eta h_B(\eta)\, d\eta - F \qquad (4.40) \\ &\text{s.t.}\ \ n_A[1 - H_A(\theta_A)]\theta_A u(F) + n_B\, x\, u(F) - F \geq 0\,, \\ &\text{and}\ \ \theta_A < x < \beta\,, \end{aligned}$$

with the additional constraint that the total welfare of group B should be greater than this achieved when they form their own distinct group. In the numerical example of Figure 4.5, this value is 35156 units (note that in this example $\theta_A^* = 0$ and $\theta_B^* = 0.5$ as indicated by the best splitting analysis above) and solving the above program using this threshold for the total value that group B need to acquire, we get $x^* = 0.46$. Interestingly, we now also require from group A a minimum contribution ($\theta_A^* = 0.12$) and thus some peers will have to be excluded. So, total social welfare is reduced to 62359 units (from 62500). But now group B receives a greater share of this amount (35200 units), still less than the indistinguishable scenario, but larger than forming their own group. On the other hand, group A receives less welfare (27156 units instead of 31249), which is still larger from all the other options. Reducing more the value of x (which results in a higher fixed contribution from the low type group), one could make it even more advantageous to the high type group to participate at the cost of reduced total social welfare.

In Figure 4.7 we depict the total and individual efficiency achieved for the two groups for different values of x. Hence, when it is possible to objectively categorize peers into groups we can always compute suitable minimum contribution fees for the different groups is order to make it beneficial for all peers to participate in such a system (given that they will be distinguished) and achieve increased overall efficiency.

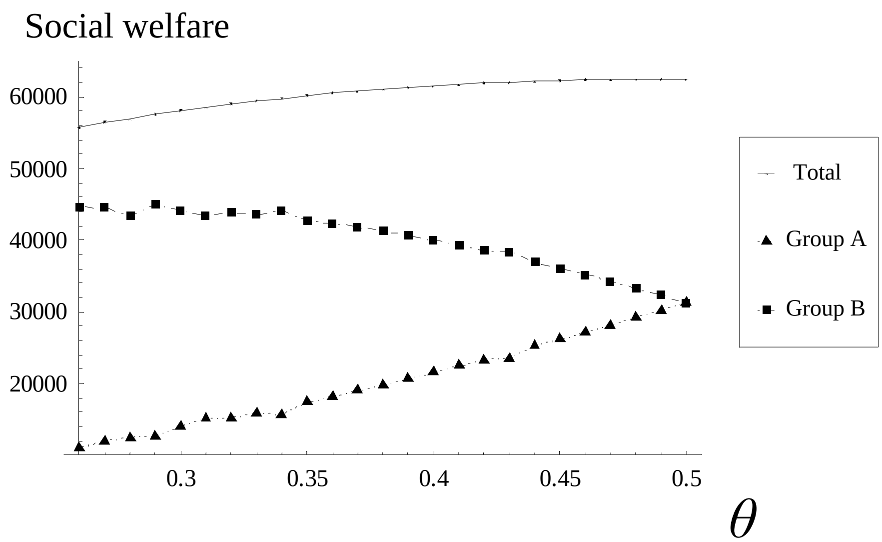


Figure 4.7: Total and individual efficiency achieved for the two groups for different values of x for $\alpha = 0.5$, $\beta = 0.5$, and $n_A = n_B = 500$

Versioning

But what if the group of a peer is not observable to the system designer? Then, clearly, all members of the group B in our example will have the incentive to claim that belong to group A in order to pay a smaller fee and thus acquire a larger net benefit. In this case one should provide to peers the appropriate incentives to declare voluntarily their group. In order to achieve this very attractive characteristic, we would again have to compromise. A standard approach is to permit more than one levels of participation creating different versions of the service (quality). Ideally, the allowed options would target the different peer types and provide the incentives for each peer to self-select the option that is targeted at her type. For instance, imagine that each peer has the choice between an effective utility $u(F)$, for a fee of f_2, or $u(\rho F)$, for a fee of f_1, where $\rho < 1$ and $f_1 < f_2$. Notice that in this case, we don't have to use the "physical limit" (if any) of the peer groups. We just have to make sure that the self-selection condition holds for our choice of f_1, f_2, and ρ and thus divide peers into groups arbitrarily (as we assumed we could do in our analysis for best splitting above). According to the value of these parameters, there will be a 'marginal' peer from the low type group with type θ_l (peers with $\theta_i < \theta_l$ will not participate) and a 'marginal' peer from the high type group with type θ_h who will be the first to prefer to pay the larger fee f_2 in order to acquire the full value of the system. That is, $\theta_h u(F) - f_2 \geq \theta_h u(\rho F) - f_1$.

So, the general program we will have to solve in order to maximize social welfare, given the ability of provide differentiated service to peers, will be the following

$$\begin{aligned} \underset{\theta_l, \theta_h, f_1, f_2, \rho}{\text{maximize}} \quad & nu(\rho F) \int_{\theta_l}^{\theta_h} \eta h(\eta)\, d\eta + nu(F) \int_{\theta_h}^{1} \eta h(\eta)\, d\eta - F \qquad (4.41) \\ \text{s.t.} \quad & \theta_l u(\rho F) - f_1 = 0\,, \\ \text{and} \quad & \theta_h u(F) - f_2 \geq \theta_h u(\rho F) - f_1\,, \\ \text{and} \quad & F = n[H(\theta_h) - H(\theta_l)] f_1 + n[1 - H(\theta_h)] f_2\,. \end{aligned}$$

Although this program is well-defined and there is always a solution, it is not easy to solve, not even numerically. Clearly, the fact that the system designer doesn't have now the required information so as to objectively categorize peers into groups offers an "informational rent" to the peers and thus the total social welfare achieved by (4.41) will be reduced compared to the one achieved under the distinguishable group scenario analyzed above. In particular, this is due to the need to offer worse value to a certain percentage of the population in order to incentivize truthful group type declaration and additionally require less contributions from high-value peers (note that in the original solution of our problem the marginal peers acquires zero net benefit) in order for them to have the incentive not to take the offer intended for the low valuation peers. Numerical experiments prove our assertion. We must also stress that in practice this approach imposes an additional cost to the system designer: the implementation of an enforcement mechanism which will ensure that different levels of service are granted to peers belonging to different groups. This is in many cases more difficult to implement than exclusions.

However, there is an interesting special case of (4.41), in which a system designer would wish to incur the cost incurred due to the self-selection incentives: when exclusions are not desirable. Using our notation, exclusions are avoided when $f_1 = 0$ (and thus $\theta_l = 0$). Then no peer will be excluded, but those that cannot afford f_2 will have to compromise to having access to just a percentage ρ of the total content made available. This approach could be additionally justified due to the fact that in many systems the cost for exploiting small contributions exceeds the overall value acquired, as for example, the cost of downloading files shared behind a very slow dial-up line[10]. In such cases, the system designer faces a

[10]This is especially so in the case of scientific grids (see Section 3.6.3), where the addition of an organization to participate has significant administrative costs and sharing very small small amount of resources could incur high scheduling costs and communication overhead during service provision. As a result, orga-

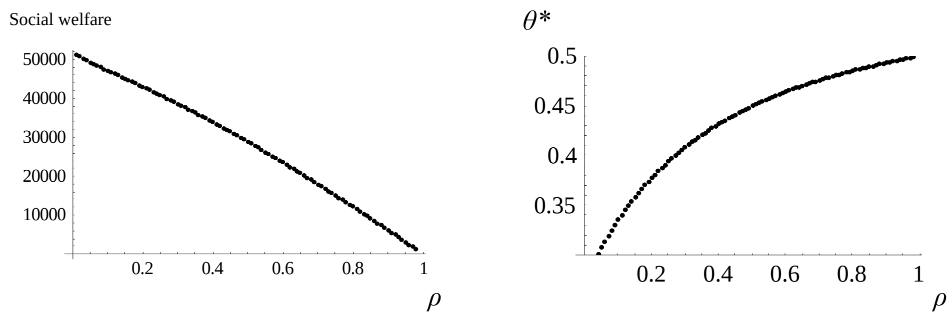
trade-off. To totally exclude peers that really cannot offer much could raise ethical issues depending on the application (e.g when several users don't have enough income to buy a broadband connection). On the other hand, to unconditionally allow them access to the pool of resources, as already explained, would ruin the incentive mechanism and at the extreme case no peer will contribute to the system. He should thus compromise if no peer is to be excluded from the system.

This is also expressed by our model since when no contribution is required by low type peers (when $f_1 = 0$), it is easy to see that the solution of (4.41) is the same with one of the indistinguishable group scenario. That is, $\rho^* = 0$. That is, low type peers should be excluded from using the good. To see why, notice that a marginal increase of ρ would make the marginal peer of the contributing group (the one with $\theta_i = \theta_h$) to move to the non-contributing group. Then F will be reduced by f_2 which means that the net benefit of $(n-1)$ peers will be reduced accordingly while the net benefit of the marginal peer will be increased only by ϵ. So, it is always beneficial, in terms of social welfare, to decrease ρ. This means that in this scenario the value of ρ is a 'political' decision rather than an economical one. One should decide on the desired value of ρ. That is, how much value peers that contribute nothing to the system is acceptable to acquire. Then what we will need to compute is the optimal value of f_2 and θ_h in order to maximize social welfare under the above system requirements. In Figure 4.8 we plot the total social welfare achieved and the corresponding θ_h of the marginal peer for different values of ρ.

4.6.2 Heterogeneous file popularity

In our proposed economic model all files shared are assumed to offer the same value to peers and incur the same costs. One could argue that this is not the case in realistic systems because of the heterogeneous file popularity. However, it is not straightforward how popularity actually affects value and cost. For example, a rare file could offer much higher value than a popular one which can be found more easily in general. Moreover, even though popular files attract much more demand, there are much more copies available and hence this increased demand is divided amongst more providers. So, in our analysis so far we have made the implicit assumption that all files have a similar (low) rate of requests

nizations with a very small amount to offer are actually incurring cost to the community. However, due to political reasons, exclusions are not always desirable in this context (e.g. in a European scientific grid such as the one built by the EGEE project —http://www.egee.org).



(a) Social Welfare as a function of ρ (b) θ_h as a function of ρ

Figure 4.8: Social welfare and θ_h according to (4.41) as a function of ρ when $f_1 = 0$ ($\alpha = 0.5$, $n = 1000$)

(either because they are unpopular or because they are distributed in proportion to their popularity) and are equally valuable to a certain peer that downloads them.

We could, however, extend our model to circumstances in which file popularity affects utility and cost. For example, suppose the following scenario. There are N_1 popular files and N_2 less popular files. A popular one is requested at twice the rate of an unpopular one (and so generates twice the cost to a peer who provides it), but is also twice as valuable. The total cost corresponding to the total upload rate is $c(F_1, F_2) = 2F_1 + F_2$ and the utility is now, say, $u(2F_1 + F_2)$. The analysis is much as above. Each peer is asked to contribute f_1 type 1 files and f_2 type 2 files. Notice that because the ratio of value to cost is the same for both file types, it is only the value of $f = 2f_1 + f_2$ that actually matters. So the planner has no preference for the precise combination of f_1 and f_2 by which a peer makes his contribution. If the value/cost ratios had been different for the two file types, then it would have been optimal to share only one type of file (the one with greater value/cost ratio). Hence, as in our original model, the planner can check that a peer is making his required contribution simply by verifying that the total number of files shared is f. This fact, and notions of equilibrium economics, suggest that all file types that are actually worth sharing will effectively have the same value/cost ratio.

4.6.3 Stability

Suppose that the social planner designs a mechanism on the basis that there are n peers. He expects that $(1 - H(\theta))n$ of them will pay a fee of $f = \theta u(F)$. Since the fee is paid 'in kind' and equates to providing f files, the total number of files that are provided will be $F = (1 - H(\theta))nf$.

Suppose that there are indeed n peers, but initially some of them are dubious that F will be as large as the planner claims. Consequently, some do not participate and the number of files that is initially provided is $F_1 < F$. If $f > u(F_1)$ then no peer finds the system to be of sufficient size to be beneficial and so no one wishes to continue to participate. So let us suppose $f < u(F_1)$. Once the peers have observed F_1, those peers with $\theta_i > f/u(F_1)$ will realise that it is to their advantage to participate. Their fees will provide F_2 files where

$$F_2 = \left(1 - H\left(\frac{f}{u(F_1)}\right)\right) nf\,. \tag{4.42}$$

Write this as $F_2 = \phi(F_1)$ and imagine iterating $F_{k+1} = \phi(F_k)$, $k = 1, 2, \ldots$. In general, there can be more than one root to $F = \phi(F)$. For example, suppose $u(F) = 0.6F^{1/2}$, $f = 5$, $n = 120$, and θ_i is uniformly distributed on $[0, 1]$. Then $\phi(F) = \left(1 - 5/0.6F^{1/2}\right)(120)(5)$. In this example there are two roots, $F = 100.00$ and $F = 320.87$. One can easily prove

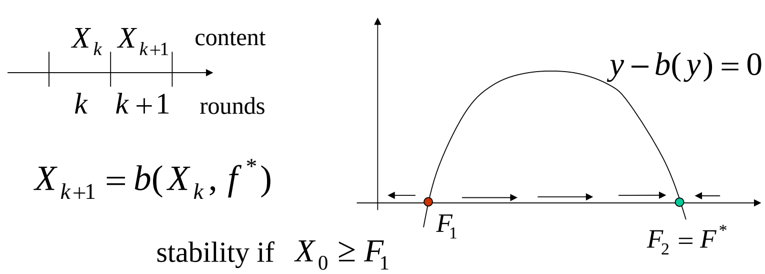


Figure 4.9: Stability

that if F_1 exceeds the smaller root then F_k tends to the larger root as k tends to infinity. Otherwise $F_k \to 0$. For $F = 100$ the social welfare is 10, whereas for $F = 320.87$ it is 184.4. Thus the greater F, to which the system converges, is also the root for which a greater number of peers participate and the greater social welfare is achieved.

4.6.4 Parameter discovery

We have assumed that n, $u(\cdot)$ and $H(\cdot)$ are known to the social planner when he computes the optimal fee f. It would be an interesting issue for further research to see how, in absence of this knowledge the planner might take advantage of what we have already learned about the form of the optimal policy to design an adaptive policy that learns the parameters. Here, we make only some preliminary remarks. If only n is not known, the planner could set an entrance fee and then observe the size F at which the system stabilizes (after iterations of (4.42)). Then n can be estimated by $F/(1 - H(f/u(F))$.

If only H is not known, and peers actually declare their preference parameters, then the planner might estimate H from the empirical distribution of the declared preference parameters, say $\hat{H}$, and then implement the solution that is optimal for $\hat{H}$. For n large he should have $\hat{H} \approx H$, provided the peers are truth-telling. This will be so if we assume that peers have different preference parameters in different instances of the game (e.g. upon different times they join to the system). The reason is that in a repeated game formulation, in which the preference parameter of a typical peer is repeatedly sampled from $H(\cdot)$, then her average net benefit will be the $1/n$th of the total social welfare and it will indeed be optimal for him to be truth-telling.

However, if a peer's preference parameters are chosen once for all or remain relatively static for a long time then there may be an incentive for a group of peers to lie about their preference parameters, hoping to fool the social planner into mis-estimating H and doing better for themselves thereby. Consider the following example (for which calculations were done with Mathematica). Suppose that $N = 100000$, $n = 100$, $u(Q(F)) = 2\sqrt{N(1 - \exp(-F/N))}$. Suppose that θ_i is one of $\{0.25, 0.5, 0.75, 1.0\}$ with frequencies 0.4, 0.4, 0.1 and 0.1 respectively. The mechanism which maximizes social welfare takes $\theta = 0$ and $F = 2183.5$. Each of the peers who has $\theta = 0.5$ makes net benefit of 24.6390. However, suppose that the peers who have $\theta = 0.5$ act in concert and arrange for one-quarter of them to untruthfully declare it as 0.75. The central planner will conclude that the frequencies of the four possible parameter values are 0.4, 0.3, 0.2 and 0.1, and for this he maximizes social welfare by taking $\theta = 0$, $F = 2411.14$. Under this mechanism, the peers who have $\theta = 0.5$ will now make a greater net benefit of 24.6975. They profit from being untruthful. (However, if all the peers who have $\theta = 0.5$ untruthfully declare it as 0.75 then they do not do better than if they are all truthful.) Another way to view this example is that it shows there can be more than one Nash equilibrium in the n-person game being played by the peers.

4.6.5 Other possible fixed contribution schemes

A feature of our limiting mechanism is that all participating peers pay the same fee. One can devise other mechanisms for which this is true. Two obvious such mechanisms suggest themselves. In Mechanism 1 the planner announces that he shall provide the good in quantity Q and then share the cost $c(Q)$ amongst all those who volunteer to participate. If m out of n choose to participate then each pays $c(Q)/m$. Those who wish to participate must make a commitment to do so, before knowing how many others will participate.

In Mechanism 2, we charge a fee of ϕ and then build the largest facility whose cost can be met by the number m who choose to participate, namely Q such that $c(Q) = m\phi$. As before, peers must make a commitment to pay ϕ without knowing how many others will also participate. The strongly feasible policy would do a bit better than a scheme in which Q is fixed a priori, because although it provides the same Q on average, it automatically saves bigger values of Q for the times that more users participate. Note that $mu(Q)$ is convex in m. A better strongly feasible equal contribution scheme would be one in which, having learnt that m wish to participate, the planner builds a facility of size $Q(m)$ and charges each participant $c(Q(m))/m$. Potential participants know the function $Q(m)$. Courcoubetis and Weber (2005) show that none of the above mechanisms can be more than a factor $1 + O(1/n)$ of second best better than the simple one we propose.

4.6.6 Implementation Issues

Theoretically, for a system to force peers to share a minimum number of files per unit of time, it needs to calculate the total time a peer has stayed in the system (during an enforcement period) and make sure that the required number of valid files were shared —and served upon request— throughout that time. Based on this information, it can then appropriately penalise misbehaving peers during the next enforcement period.

A more realistic approach could treat the actual uploads provided by a peer as a proxy for the number of files shared and use as discussed in Section 4.6.2. For example, one could consider issueing expiring receipts signed by peers when downloading content, which will be checked by future uploaders before serving a serving request. Peers will be able to download a file only if they can show a valid number of such receipts. Tuning the expiring time appropriately, the system designer would be able to enforce the required level of sharing at each enforcement period. Notice that such a mechanism differs from mechanisms equating consumption with contribution (either using token-base accounting or reputation) in that, as suggested by our model, request rate is not regulated. But such an enforcement mechanism (and any other that is based on accounting) would be susceptible to false trading (peers signing fake receipts for their friends or even for themselves creating multiple identities) and whitewashing as discussed in Section 2.6.

In the next chapter we propose a memory-less incentive mechanism motivated by the proposed fixed contribution scheme. Its most important difference is that the time peers stay in the system sharing a fixed, and common for everyone, amount of files, depends on their request rate. This is actually the price one has to pay if he wants to implement any incentive mechanism in the untrusted environment of p2p systems, where accounting of past transactions is problematic or even not feasible. That is, the need to rely on the time peers are consuming resources in order to force them to contribute.

Another important practical issue that needs to be addressed by a system designer who wishes to regulate peer behaviour in order to increase system's efficiency, is the dynamic tuning of the most important parameters (such as the required number of files shared per unity of time) as a function of its size and overall activity, which could change drastically over time in a realistic p2p system. This procedure would in general depend on experience and on many system-specific characteristics. We have left the modelling of such a dynamic model for future research. But note that again the ability to actually implement the adjustment of certain parameters during the system life time is also a very challenging task both in terms of measuring and update, which we have taken into account in the design the proposed memory-less mechanism (see next chapter).

4.7 Summary

In this chapter we made the connection of file sharing with public goods. We have thus defined the notion of efficiency of a p2p file sharing system and demonstrated the nice characteristics of a fixed contribution scheme. Most importantly, that such a simple mechanism is enough to lead the system, asymptotically, to the maximum level of efficiency that could be possibly achieved by any mechanism under incomplete information.

However, as we showed, the performance of the fixed contribution scheme decreases with the heterogeneity of the peers in terms of their preference parameters. Hence, categorizing peers into different groups could be greatly beneficial. We demonstrated this fact and actually showed that in many cases only a very small number of such groups is enough to maximize efficiency on the average. However, in most cases suitable incentives should be provided in order for peers to agree to participate declaring their true group type or agreeing to be distinguished accordingly if this is observable. We analyzed some possible solutions in order to manage to exploit the additional information due to group forming and discussed the underlying trade-offs.

We also provided some insights towards addressing several challenging problems that arise when such a mechanism is to be employed in a realistic p2p system such as the effect of heterogeneous file popularity, the conditions required for stability, the difficult problem of discovering system parameters that are assumed known (such as the size and distribution H), and the very challenging implementation issues involved concerning the enforcement of the proposed mechanism, which we discuss in detail in the next chapter.

Note that although our proposed mechanism is not always easily enforceable and many practical issues are not fully addressed, it provides us with a very useful theoretical benchmark for evaluating and comparing different practical incentive mechanisms such as the one proposed in the next chapter. Moreover, a large part of the analysis presented is also applicable in other p2p systems that have a strong public good aspect as well, such as scientific grids and WLAN peering, and which could have different implementation requirements. This is the reason we have treated economic modelling and implementation as orthogonal aspects of the design of incentive mechanisms in p2p systems since the technology has not yet settled and it is critical to make the right abstractions in this complicated modelling environment in order to understand what is important and what is not. In our view, the public good aspect of p2p systems plays in many cases a dominant role regarding resource allocation and provision and thus we believe that the importance and applicability of our results in this chapter are not constrained to p2p file sharing applications.

Chapter 5

A Memory-less Enforcement Mechanism

In this chapter we present a candidate memory-less mechanism enforcing a (close) to our theoretical fixed contribution scheme, which was analyzed in detail in the previous chapter. Under certain assumptions, this mechanism could lead a realistic file sharing system to satisfactory efficiency levels, a very attractive characteristic given the fact that no accounting is required. We discuss the system requirements that arise and formulate a new economic model that captures the basic economic transactions that take place in order to assess the efficiency of our proposed mechanism and compare it with other alternatives.

Part of this chapter is joint work with Ben Strulo (see also [Antoniadis et al. (2005)]).

5.1 Introduction

The ultimate goal of this dissertation is to design a system that focuses on content availability (as Direct Connect) without requiring the existence of any sort of user memory or the ability to permanently expel peers from the group, following the implementation principle of BitTorrent. Thus, we hope to improve the economic efficiency of the system without suffering from whitewashing and false trading and without the implementation requirements of longer term tracking of peer behaviour.

A robust memory-less enforcement mechanism needs to rely only on the time a peer is consuming resources in the p2p system (i.e. downloading a file) and thus use the serving peer (i.e. the *uploader*) as the 'enforcing entity'. Direct exchange mechanisms, discussed in Section 2.6.4, rely exactly on this principle to enforce reciprocity between peers. We follow a less strict approach, which does not require from peers to reciprocate in terms of actual resource provision (e.g. forming pairwise or n-way exchanges) but to contribute to the system as a whole while consuming, giving emphasis on its public good aspect. This approach has of course its own weaknesses (e.g. it depends on the existence of super peers) but we believe that its tolerant nature (and explicit focus on content availability) makes it an attractive mechanism for realistic p2p systems where there are actually many peers that wish to take the role of a super-peer, and strict approaches equating consumption with contribution could harm content availability and the community spirit, which seems a very fundamental characteristic for their success. We elaborate more on the several related trade-offs in the following.

We first summarize our main assumptions. First of all, we consider uploading cost to be of limited importance, especially while downloading. This is especially so since, anticipating one of the main characteristics of our mechanism, we design the system to aim to provide a particular upload throughput for all requests at all times. Hence, since all peers use the same upload throughput, we have a homogeneous system in which the prioritization of TCP ACKs largely eliminates the cost of uploading while downloading (see Feldman et al. (2003)).

Moreover, our focus on content availability (and especially on the 'long tail' of the content) rather than bandwidth offered for uploads motivates us to also assume that: 1) congestion on the upload link of an individual peer is rare and 2) all files have a similar (low) rate of requests, either because they are unpopular or because they are distributed in proportion to their popularity. Finally, we assume that peers act rationally in their own self-interest and not maliciously.

5.2 Enforcement mechanism

Two are the main attributes of a peer's contribution towards content availability:

- number of file shared
- her own availability (the time she stays on-line sharing and serving these files)

So, we propose a system where the uploader first checks whether a candidate *downloader* shares the amount of files required before providing a requested file. Moreover, she should ensure that these files are accessible to the rest of the group (e.g. by checking with the search mechanism for their availability). If the downloader refuses to actually upload one of his files to a potential *requestor*, the latter will inform the uploader to stop the transfer[1]. The same would happen if the file sent by the downloader is not valid as discussed in Section 5.3.1.

The second attribute of a peer's contribution, and probably even more important, is the time that she spends in the system sharing the required amount of files. Since we have assumed that peers can be forced to contribute only during the time they download files, this time is directly related to the upload throughput offered by the uploader. So, in order for the system to enforce a certain contribution by peers, it also needs to constrain the minimum average download time that they face. This could be easily implemented using a simple and lightweight protocol that broadcasts the fixed throughput that all peers will be using for uploading files thus allowing it to be dynamically tuned.

Increased average download time would have a positive impact both on the content availability achieved and on the probability that a file of a downloader is requested (in which case he would be forced to also upload a file to continue downloading). But it also reduces the utility of peers by requiring them to wait longer for their downloads to complete. We study such trade-offs concerning this crucial parameter of our mechanism in Section 5.4.3.

In the following, we describe the main requirements of a p2p system implementing the proposed memory-less mechanism.

5.2.1 Super-peers

The current generation of p2p file sharing systems exploits the heterogeneity previously observed among peers [Saroiu et al. (2002); Yang and Garcia-Molina (2003)]. Unlike pure p2p systems such as the original Gnutella network, where all peers share the same responsibilities, current hybrid p2p file sharing systems, such as KaZaA and the new Gnutella protocol (Gnutella 2.0), assign more responsibilities to peers with more resources (i.e., higher access bandwidth, larger processing power and content to share) and/or stable connectivity. Such peers, the *super-peers*, form a backbone which is responsible mainly for answering queries

[1]Remember that we do not consider the case of possible malicious behaviour from peers. Otherwise, malicious peers could falsely report that some other peers are misbehaving just to ruin the system. A rational peer doesn't have an incentive to do so.

for content. The rest of the peers, the *leaves*, are connected to this backbone through one or more super-peers. Super-peers forward queries only between them and thus avoid network flooding, which was responsible for the many inefficiencies that arose in the operation of the original Gnutella network. For instance, in Gnutella 2.0[2], a node searching for a given file, sends a query message to the super-peer to which it is connected. This super-peer first searches for the file in its local repository and in those of the leaves directly connected to it. At the same time, it forwards the query to its neighbors, which, in turn, search for the file in their local network but do not forward the message any further. One could also imagine of different optimizations for increasing the probability of a successful query forwarding between super-peers based on history, interests, etc.

The benefits of such an hierarchical architecture are apparent. Actually, current practice and the tremendous sucess of p2p applications constitute the proof of concept for the scalability and efficiency of hybrid p2p systems (see also [Benevenuto et al. (2004)] for a quantitative evaluation). A natural question arises however. Which are the incentives for peers to operate as super-peers in a p2p system? The answer is that there are many users that have very small operational costs (such as university students and expert users) with broadband connections and constant connectivity to the Internet, very large content repositories, and altruistic motives, who voluntarily wish to contribute their excessive resources for the efficient operation of a p2p community. Liang et al. (2004) present measurements in the Kazaa network verifying that indeed a large number of users choose to play this role even though it is not rational to do so.

Hence, it is not at all unrealistic to assume the existence of some peers who have the resources and the suitable incentives to play this role. Note that in any case, their existence is necessary for even more fundamental system functionality such as service discovery and thus we believe that practical incentive mechanisms that do not make this assumption could unnecessarily restrict themselves to worse levels of efficiency than these that could be achieved otherwise. So, even if additional incentives are needed in order for users to decide to play the role of a super-peer (and this is actually an additional interesting direction for research in p2p economics) we assume that these are in place.

In our system, we need super-peers for two important reasons. First, we rely on them to act as seeds for the content by providing a certain initial amount of files and by becoming the roots of the envisioned trees of downloaders which would further increase the availability

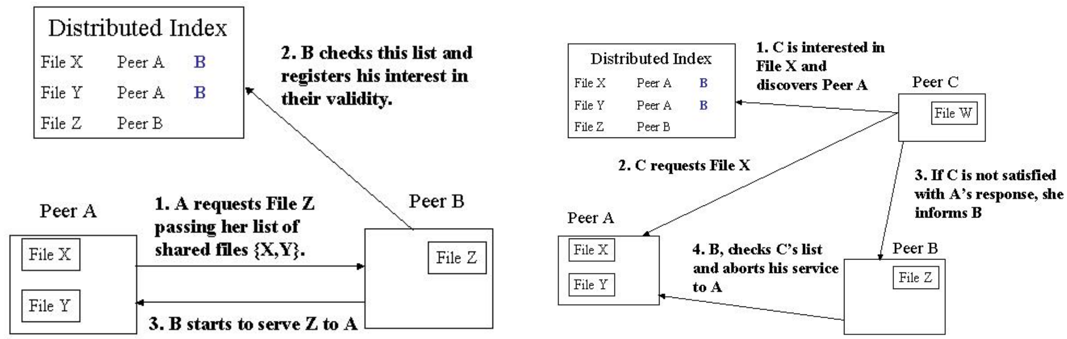
[2]http://www.gnutella2.com

of content in the system. Otherwise, our mechanism would require "cycles" of content requests to be formed resembling to a direct exchange mechanism (see [Anagnostakis and Greenwald (2004)]). Second, our super-peers are responsible for computing useful system information (such as the size of the system, the number of files shared, etc.) and tuning important system parameters, such as the fixed upload throughput and the minimum number of files shared per peer. Note that this is the main aspect of our mechanism that differentiates our approach from existing p2p systems and related research work. That is, our effort to incorporate some sort of regulation towards improving the economic efficiency of the system as this is defined by our modelling work. Moreover, the existence of super-peers makes the implementation of this functionality more realistic and helps avoid specific incentive issues that arise, as we discuss next.

It is out of the scope of this dissertation to give a detailed system design of the proposed mechanism. We believe that its main functionality described in the following could be easily incorporated in the specification of a hybrid p2p architecture such as Gnutella 2.

5.2.2 Checking file validity

We first describe an abstract protocol which passes the necessary information between peers in order to achieve the most challenging functionality of our enforcement mechanism: ensuring that downloaders make available the predefined number of valid files to the rest of the group. Initially, as depicted in Figure 5.1(a), Peer A discovers files of interest through the system index and sends a request to an appropriate uploader (who serves the specific files), Peer B. This request includes a list of files currently being served by A as evidence of her contribution. Peer B needs to check the validity of this list (or in order to reduce the corresponding cost, a random set of files from this list). He does this by querying the index himself. If he fails to find them in the index he will refuse to serve A. If he does find them, he adds a note to these index entries indicating his current interest in the reliability of these services. He then begins to serve to A. This procedure is necessary because the fact that a certain number of files are made available by a peer does not mean that she will actually serve these files when requested. So, during the time in which B is serving A, it is possible for another peer C to wish to request a file from A. C will discover via the index that B has an interest in the validity of these services. Then if the service to C is not performed adequately, (the files are not actually served, or are corrupt or mis-labeled,



(a) Checking file validity (b) Punishing misbehaving peers

Figure 5.1: Basic functionality of the enforcement protocol

etc.) then C can tell B (see Figure 5.1(b)). At this point, B will punish A by aborting the download of file Z, after checking the validity of files shared by C (to ensure that this was not the reason she wasn't served).

In order to minimize the validity checks, since they could be costly (see also Section 5.3.1), this extra check would be performed only if there is really a conflict. That is, when Peer B denies the statement of C (saying that B is defecting). Note that B does not have the incentive to do so if C is a lawful peer and he has not served him adequately. Similarly, C (if rational) doesn't have the incentive to falsely accuse B for not serving him.

5.2.3 Parameter tuning

The second attribute of our enforcement mechanism, the fixed upload throughput, is much more easier to be controlled by peers. In relatively stable systems (with a constant average number of peers with similar behaviour) this functionality would be implemented in the client software which will be configured to upload files using a hard coded predefined throughput.

However, in most cases, p2p systems can change drastically over time both in terms of participation and behaviour. Hence, an appropriate protocol should be designed in order for the value of upload throughput used by the client software to be updated by super-peers based on measurements of the overall activity in the system. This is particularly important since, as we will see in Section 5.4, when the throughput's value is below a certain threshold the system becomes unstable. Thus, there should be ways for the adaptation of this crucial parameter as the system evolves.

The updating part is relatively easy, since by the time super-peers agree on the value of the throughput that should be used, each one of them can broadcast to each own subnetwork (the normal peers that are directly attached to them) the new value that should be used by their client software while uploading files. Performing the computation of this value in a distributed fashion while the system evolves is a more challenging task, but again is feasible due to the existence of the super-peers. The first task of the super peers would be to collect the system statistics based on which they will check whether a new value of upload throughput needs to be broadcasted. These statistics include the number of the participating peers, the probability of a successful request, the percentage of aborted downloads, and many more system specific parameters. When a consesus is reached about the value of these parameters an algorithm hardcoded in the software of super-peers (possibly based on results similar to those presented in Section 5.4 below) would decide whether an update should take place and which should be the new value. Again, some sort of consesus should be established and the new value would be broadcasted to all normal peers.

5.3 Incentive Issues

Does the proposed mechanism actually provide the required incentives to peers under our system assumptions? If peers choose not to advertise shared files then they will be unable to download. However, can peers pretend to be sharing but actually fail to do so? We discuss this and other incentive issues that arise under our proposed mechanism below.

5.3.1 Sharing invalid files

Since we don't require a peer to actually upload a file while downloading, we need to do our best to check the validity of the files shared as early as possible. The first check is that the files are actually shared in the distributed index. Thus peers cannot claim to be sharing without running the risk of receiving a request.

A cheating peer could advertise files with meaningless or deliberately obscure names thereby reducing the chances of receiving request. They need not actually have the claimed content available. We can imagine processes, perhaps run by super-peers, which flush invalid titles from the index, perhaps by using something similar to the freedb service[3] (either centralized or implemented in a distributed way). This will make it difficult for peers to advertise such names. The uploading peer could also perform such checks.

A cheating peer may also advertise files with legal names but corrupted or invalid content. This could also be checked by the uploading peer (or more likely by future requestors of the file) using techniques like the ones proposed by Liang et al. (2005). However, if the probability of receiving a request is low, this may still be a profitable attack. However, there are a number of mechanisms we can use to reduce this risk. For example

- super-peers could perform random checks on shared files, perhaps focusing on new peers, or newly shared content. If a trusted subset of super-peers is available then passing this check could be rewarded with a signed certificate that could be used in future transactions.

- peers could be incentivized to contribute their own (possibly rare) tastes into the community and thus become members of social networks of interest. Or there could be social incentives such as improved social reputation for the provision of high quality content.

- a more demanding scenario in terms of effort exerted by super peers could be the implementation of "wish list" functionality. That is, unsatisfied content requests by simple peers could be stored in a distributed "wish list". Whenever a peer with such a file enters the network super peers will download the file on behalf of the peer that requested it in the first place. She will then be able to download it from a super-peer when she will again enter the system. This operation will increase the probability that files are requested and provide an additional metric for the validity of files shared.

In this chapter our main focus is on how to incentivize peers to provide content for a sufficient time to improve its availability. Thus we do not provide a complete solution to this question of content validity and more research needs to be done in this area. We discuss this further in Section 5.5.

[3]http://www.freedb.org/

5.3.2 Failing to serve

But still, even if we ensure that peers share valid files, they could always fail to serve them to potential requestors. Nevertheless, this form of misbehaviour should be detectable as soon as one of the shared files is requested by a third peer. The defecting peer can choose at this point to refuse (or abort) the corresponding download. But then her own download will be also aborted. Of course, the misbehaving peer could always enter into the system at a later time to continue any aborted downloads. But this means that her download will take more time than would have taken if she had cooperated. So, since we have assumed that uploading a file does not incur any additional cost, a downloader, having to pay anyway the cost for sharing the required amount of files and for waiting for his download to complete, doesn't have the incentive not to serve an incoming request. As a result, we don't consider this attack a rational decision in our context.

Notice that this attack resembles the free riding issue in BitTorrent, where potential free riders can receive a —limited— service due to 'optimistic unchoking' (see Qiu and Srikant (2004)). Notably, in the case of BitTorrent this attack is not actually performed by the majority of participating peers [Andrade et al. (2005);Hales and Patarin (2005)] and this supports our belief for its limited appearence in our context as well. Note also that, as in the case of content validity discussed above, the failing to serve strategy becomes more profitable as the probability of receiving a request for a file decreases. So, some of the mechanisms proposed for addressing the sharing of invalid files, would discourage the failing to serve strategy as well.

5.3.3 Incentives for enforcing

As already discussed all incentive mechanisms in the context of p2p systems, both fully distributed and hybrid, pose a certain cost (much more significant the more distributed is the implementation of the system's core functionality) for their enforcement to participating peers. Our enforcement protocol has two parts: a) checking file validity and aborting uploads to misbehaving peers and b) constraining the upload throughput offered.

The first one incurs a certain fixed cost to the peers executing it. However, we expect this cost to be relatively small at least in cases where bandwidth is not scarce as assumed in our system. But of course, a system designer should ensure that it is minimized by careful design and by moving the maximum possible part of its functionality to super peers. Then, at least for some peers, this cost may be outweighed by the direct benefit of having an audit done; i.e. that other peers are prevented from abusing the system. Notably, it seems that human nature includes a bias towards punishing cheats even at some cost to the punisher, as described by Fehr and Gachter (2002).

The second one, fixing the upload rate to a prescribed value, carries no cost and thus peers will have no incentive to increase it, especially since its enforcement will be hardwired into the application. Decreasing the upload rate is also not a rational strategy under our assumptions. However, if it is considered, for some reason, a valid attack, the system has the following options (amongst others) to address it:

- A peer receiving lower than expected throughput could treat the corresponding peer as "failing to serve" with some probability (depending perhaps on the probability that a peer's upload capacity is congested —in which case the lower throughput should be excused)

- Alternatively, a peer could be limited to receiving the same —reduced— download throughtput with the one offered to her own downloaders. In this case, the report from such a downloader would include the measured throughtput received from the possibly defecting peer. Using the notation in Figure 5.1, Peer C could inform B in order to adjust his upload throughput to the value A offers to C.

Finally, colluding peers that agree to upload to each other with higher than the prescribed upload rate will not gain much when they also interact with normal or super peers, since they will still have to wait for these downloads to finish. If they are not interacting with any other peers then they effectively form their own isolated system which is always an option in any case.

5.4 Economic Modelling

We will now reformulate our public model of Chapter 4 to focus on peer availability. That is, consider only the time a peer decides to stay on-line as her main contribution towards content availability rather than both the number of files shared and availability. We will then extend it to capture the special characteristic of the proposed enforcement mechanism, namely the fact that peers are assumed to contribute their resources only while consuming themselves.

Suppose that peers $1, \ldots, n$ are to share the use of a public good: the expected number of distinct files made available in the system at some arbitrary time. The good can be provided at quantity Q for a rate of cost $c(Q)$. If N is the maximum number of distinct valid files, we now use Q/N, the probability that a random request is satisfied, to express the content availability achieved in a system of size Q.

As in Chapters 3 and 4 peer i has a utility for the good of $\theta_i u(Q/N)$, where θ_i is a 'preference parameter' which is known only to peer i, but which is again a random sample from a distribution on $[0, 1]$, with distribution function $H(\cdot)$ and density function $h(\cdot)$. The function $u(\cdot) \geq 0$ is again assumed to be continuously differentiable, increasing and strictly concave in its argument (i.e. $u(x) = x^\beta$, where $\beta < 1$ is a positive constant). In order to build the public good Q, each peer has to share f_i files for a fraction t_i of time ($0 \leq t_i \leq 1$). Then, at some arbitrary time the total average number of not necessarily distinct files shared F, will be $F = \sum_{i=1}^{n}(f_i t_i)$. Due to duplication, the number of available distinct files Q will be in general a concave function of F, which when Q/N is not very close to 1 and all files are equally popular, could be approximated by F, as shown in Section 4.1.1. That is, in our range of parameters, $Q(F) \approx F$, and we can use Q instead of F. In the following we make this simplifying assumption, which is not crucial for the qualitative results we obtain.

Since we do not account for uploading costs, no limitation is posed on the rate with which peers request and download files. Moreover, the number of files shared is considered a 'sunk' cost, incurred by a peer before entering the system and includes mainly the costs for acquiring (e.g. ripping a CD) and storing the content. In our analysis we have assumed that the system designer has determined beforehand a fixed and common number of files f required to be stored on each peer, based on system parameters like the expected number of participating peers, the maximum number of distinct files N, the average size and value of files shared, etc.

So, during system operation, the rate of cost peers have to contribute for building the public good Q is only due to the fraction of time t_i they have to stay on-line, sharing the fixed amount of f files. We assume that it is linear in t_i and the same for all peers. So, $c(Q) = \alpha \sum_{i=1}^{n} t_i$, where α converts time units to monetary units, and Q will be equal to $\sum_{i=1}^{n} t_i f$, and thus $c(Q) = \alpha Q/f$.

In the following we remind the first-best and fixed contribution schemes with the above slight modifications, and present and analyze in depth the new economic model which we formulated in order to capture the fact that peers can be forced to contribute their files only while downloading.

5.4.1 First-Best

Under complete information (when the payoff parameter θ_i of every peer i is known) and unlimited enforcement capabilities, the system designer should decide on the optimal amount of Q built and the fraction of time t_i that each peer i should stay on-line, solving the following optimization problem

$$\underset{\{t_1,\ldots,t_n\},Q}{\text{maximize}} \sum_{i=1}^{n} \theta_i u(Q/N) - c(Q) \tag{5.1}$$
$$\text{s.t. } 0 \leq t_i \leq 1 \;\&\&\; \theta_i u(Q/N) \geq \alpha t_i, \; \forall i.$$

When the optimal size of the system Q^* is computed using (5.1), it is then trivial to compute a set of feasible t_i^*s (i.e. such that $\theta_i u(Q^*/N) \geq \alpha t_i^*, \; \forall i$).

The solution of this problem (the first best provision) will form our 'benchmark' for assessing the efficiency of our proposed mechanism, since the efficiency achieved solving (5.1) is the maximum possible.

5.4.2 Fixed contribution mechanism

According to the fixed contribution scheme analyzed in depth in the previous chapter, the system designer should choose $\bar{\theta}$ and Q according to the following

$$\underset{Q,\bar{\theta}}{\text{maximize}} \; nu(Q/N) \int_{\bar{\theta}}^{1} x dH - c(Q) \tag{5.2}$$
$$\text{s.t. } n(1 - H(\bar{\theta}))\bar{\theta} u(Q/N) = c(Q)$$

As already demonstrated (see Chapter 4), (5.2) above maximizes the expected social welfare over the choice of fixed fee policies. The optimal policy will correspond to the optimal values of the two variables Q and $\bar{\theta}$. Solving (5.2) with $c(Q) = \alpha Q/f$ and the additional constraint that $t^* \leq 1$, we can compute Q^* and then the minimum contribution of each participating peer would be $t^* = \bar{\theta} u(Q^*/N)/\alpha$.

Although the fixed contribution mechanism is in theory a very simple and attractive incentive scheme for content availability, it cannot be enforced by a realistic p2p system since it requires constant auditing of a peer's contribution and user memory (see Section 4.6.6). For instance, in the limiting situation where the access lines of peers have infinite capacity, and hence the average download time is zero, a rational peer participating in such a system would set $t_i = 0$, without affecting his own use of the system. Thus the inability to incentivize peers to stay in the system longer than needed by their downloads to complete, would result to 'market failure'.

5.4.3 Fixed upload throughput mechanism

In the context of our memory-less mechanism, we propose the control of the throughput b with which peers upload files to each other, in order to force peers to stay in the system sharing their files for at least the amount of time they consume resources for themselves. We denote by d the average download time of one file, which is in general a decreasing function of b and we will thus use d as our control paremeter in our analysis below and for simplicity assume that $d(b) = s/b$, where s is the average file size. In Section 5.4.4 we discuss in more detail the relation of these two variables and the significance of this assumption.

Under the 'fixed upload throughput' mechanism, the contribution of peers will depend on their request rate multiplied by the probability that their requests are successful (and thus it will not be the same for everybody as in the fixed contribution scheme). We define $r(\theta)$ to be a function that maps a peer's type to its request rate. For our analysis we have chosen $r(\theta) = \theta^2$, assuming a convex relation between the type and the request rate of peers. Let x_i $(0 \leq x_i \leq \theta_i)$ be the value to which a peer is willing to reduce her type, and hence her request rate, when facing an average download time d in a system of size Q. Then the fraction of time that peer i will be downloading in the system, say t_i, will equal the fraction of time that a $M/D/\infty$ queue has at least one customer present. Assuming that t_i is small, we have $t_i = r(x_i)(Q/N) \times d$, that is, the rate of successful requests, $r(x_i)(Q/N)$, times the average download time per successful request, d.[4]

We describe now the iterative procedure that will converge to the limiting value of Q. Suppose that Q_0 files are made available by super-peers and Q_1 additional files are made available by the peers themselves while they are present and downloading in the system.

[4] We see this by Little's formula, $L = \lambda W$, with $L = t_i$, $\lambda = r(x_i)(Q/N)$ and $W = d_i$. Note that we are assuming that a peer may be making more than one download at the same time. But since t_i is assummed to be small, the number of downloads taking place is almost always 0 or 1, and hence $t_i \approx L$.

Q denotes the total content in the system, i.e., $Q = Q_0 + Q_1$. Each peer i will choose an optimal x_i by solving the following local optimization problem.

$$\underset{x_i}{\text{maximize}}\Big\{x_i u((Q_0+Q_1)/N) - \alpha r(x_i)((Q_0+Q_1)/N)d\Big\} \tag{5.3}$$
$$\text{such that } 0 \le x_i \le \theta_i \text{ and } r(x_i)((Q_0+Q_1)/N)d \le 1.$$

The cost is taken to be proportional to the fraction of time that the peer is downloading (and hence to the fraction of time that he is making files available for upload). Recalling that $r(x_i) = x_i^2$, the solution will be where

$$x_i(Q) = \min\left\{\theta_i, \bar{\theta}(Q)\right\}\ , \text{ where } \bar{\theta}(Q) = \frac{u(Q/N)}{2\alpha d(Q)/N}\,.$$

Note that $x_i(Q)$ is a decreasing function of Q. This fact remains true under the assumptions only that u is concave and g is convex, as we must have

$$g'\left(\bar{\theta}(Q)\right) = \frac{u(Q/N)}{\alpha d(Q)/N}\,.$$

and the right hand side is decreasing in Q.

Assuming that while he is downloading, a peer makes f files available for uploading, her choice of x_i will have an effect on the number of distinct files, $Q_0 + Q_1$, that are available for others to upload. Suppose peer 1 measures the number of files in the system as $Q_0 + Q_1$ and decides to change her type from x_1 to $x_1(Q_0 + Q_1)$. Then Q_1 will change from the solution to

$$Q_1 = \sum_i r(x_i)((Q_0+Q_1)/N)df$$

to the solution of

$$Q_1' = \left[r(x_1(Q_0+Q_1)) + \sum_{i\neq 1} r(x_i)\right]((Q_0+Q_1')/N)df$$

Now peer 2 measures the number of files in the system as $Q_0 + Q_1'$ and decides to change his type to $x_2(Q_0 + Q_1')$. Now

$$Q_1'' = \left[r(x_1(Q_0 + Q_1)) + r(x_2(Q_0 + Q_1')) + \sum_{i \neq 1,2} r(x_i) \right] ((Q_0 + Q_1'')/N) df$$

And so repeating this iteratively for $k = 1, 2, \ldots$ we have

$$Q_1^{k+1} = \sum_{i=1}^{n} r(x_i(Q_0 + Q_1^k))((Q_0 + Q_1^k)/N)\, d\, f\,.$$

Our first observation is that this procedure will always lead to a fixed point

$$Q_1 = \frac{\sum_{i=1}^n r(x_i) Q_0\, d\, f}{N - \sum_{i=1}^n r(x_i)\, d\, f}\,, \tag{5.4}$$

unlike the case where $Q_0 = 0$, where we require that $\frac{\sum_{i=1}^n r(x_i)\, d\, f}{N} = 1$ and hence

$$\rho := \frac{\sum_{i=1}^n r(\theta_i)\, d\, f}{N} \geq 1\,, \tag{5.5}$$

since if $\rho < 1$ then we would have

$$Q_1^{k+1} = r(x_i(Q_1^k))(Q_1^k/N)\, d\, f \leq r(\theta_i)(Q_1^k/N)\, d\, f < \rho Q_1^k \leq \rho^k Q_1^{(0)}\,,$$

and so $Q_k \to 0$ as $k \to \infty$.

We provide now a complete stability analysis for a particular case, which we expect should hold in a more general setting. Consider first the case $Q_0 = 0$. Suppose that n is large and we are at an equlibrium. Then we must have

$$\frac{\sum_{i=1}^n r(x_i) df}{N} = 1 \tag{5.6}$$

where

$$x_i = \min\left\{\theta_i, \frac{u(Q/N)}{2\alpha(Q/N)d}\right\} = \min\left\{\theta_i, \bar{\theta}\right\}\,,$$

for some Q and $\bar{\theta}$, where

$$\bar{\theta} = \frac{u(Q/N)}{2\alpha(Q/N)d}. \tag{5.7}$$

Assuming now that $\theta_1, \ldots, \theta_n$ are i.i.d. $U[0,1]$ random variables, and hence so are the

$r(x_i)$, $i = 1, \ldots, n$, and using simple calculations we get that

$$\frac{1}{n} E\left[\sum_{i=1}^{n} r(x_i)\right] = (1/3)\bar{\theta}^3 + (1-\bar{\theta})\bar{\theta}^2 \,. \tag{5.8}$$

Rewriting our stability condition as

$$\frac{1}{n} \sum_{i=1}^{n} r(x_i) = \frac{N}{d\,f\,n} \tag{5.9}$$

and observing that $\frac{1}{n}\left[\sum_{i=1}^{n} r(x_i)\right] - E\left[\sum_{i=1}^{n} r(x_i)\right]$ is an $O(1/\sqrt{n})$ quantity (by the Central Limit Theorem) our stability condition (5.9) becomes

$$\frac{1}{n} E\left[\sum_{i=1}^{n} r(x_i)\right] + \frac{A}{\sqrt{n}} = \frac{N}{d\,f\,n}. \tag{5.10}$$

where the first term of the left hand side is a $O(1)$ quantity given by (5.8).

Assuming that n is large, we can neglect the $A/\sqrt{n}$ term and obtain the approximate stability condition

$$\frac{1}{n} E\left[\sum_{i=1}^{n} r(x_i)\right] = \frac{N}{d\,f\,n} \Leftrightarrow (1/3)\bar{\theta}^3 + (1-\bar{\theta})\bar{\theta}^2 = \frac{N}{d\,f\,n}. \tag{5.11}$$

The expected social welfare, using (5.7) is

$$SW = n\Big[(1/2)\bar{\theta}^2 + (1-\bar{\theta})\bar{\theta}\Big] u(Q/N) - n\alpha\Big[(1/3)\bar{\theta}^3 + (1-\bar{\theta})\bar{\theta}^2\Big](Q/N)d \tag{5.12}$$

$$= n\Big[(1/2)\bar{\theta}^2 + (1-\bar{\theta})\bar{\theta}\Big] 2\alpha(Q/N)d\bar{\theta} - n\alpha\Big[(1/3)\bar{\theta}^3 + (1-\bar{\theta})\bar{\theta}^2\Big](Q/N)d \tag{5.13}$$

$$= n(Q/N)d\alpha\Big[(2/3)\bar{\theta}^3 + (1-\bar{\theta})\bar{\theta}^2\Big] \tag{5.14}$$

Suppose we take $u(Q) = Q^{1/2}$. Then $Q/N = 1/(2\alpha d\bar{\theta})^2$ using (5.7). So we get by substituting in (5.14) that

$$SW = \frac{n^2(3-\bar{\theta})(3-2\bar{\theta})\bar{\theta}^2 f}{36\alpha N} \,. \tag{5.15}$$

Observe that this is an equation involving only $\bar{\theta}$, which is a function of d. This is maximized at $\bar{\theta} = (3/16)(9-\sqrt{17}) \approx 0.9144$. Hence we should take

$$d^* = \frac{3.06\,N}{nf} \tag{5.16}$$

using (5.7) and substituting $\bar{\theta}$. It is interesting that this optimal decision does not depend on α.

The case of $Q_0 > 0$ is as follows. Recall that Q denotes the total content in the system and Q_1 the content provided by the peers at the equilibrium, i.e., $Q = Q_0 + Q_1$. Then

$$Q_1 = \frac{\sum_{i=1}^{n} r(x_i)(Q_0 + Q_1)\, d\, f}{N} ,$$

or equivalently the stability condition becomes

$$\frac{\sum_{i=1}^{n} r(x_i)\, d\, f}{N} \approx \frac{n[(1/3)\bar{\theta} + (1-\bar{\theta})]\bar{\theta}^2\, d\, f}{N} = \frac{Q - Q_0}{Q} . \tag{5.17}$$

Assuming again that $u(Q) = Q^{1/2}$ we will also have, according to 5.7, that

$$Q = \frac{N}{(2\alpha d\bar{\theta})^2} . \tag{5.18}$$

Now (5.17) and (5.18) define a second degree polynomial in d

$$n[(1/3)\bar{\theta} + (1-\bar{\theta})]\bar{\theta}^2\, d\, f + Q_0(2\alpha\bar{\theta})^2\, d^2 - N = 0 , \tag{5.19}$$

which allows us to compute its positive root. Substituting this for d in (5.14) allows us to have an expression of the social welfare as a function of $\bar{\theta}$ as before. Again, we could compute the maximizing value of $\bar{\theta}$ as a function of the rest of the parameters, and use (5.19) to find the dependence of the optimal d on the rest of the parameters.

5.4.4 The parameter d

Notice that during our analysis above we have made an implicit assumption: that once a download starts, then it will be completed, and thus for each successful request a peer will always stay connected for time d. In order for this to happen, if a request for a file arrives during time d, the peer that serves this file must provide the upload to the requesting peer without aborting it, even if her own download (the reason that she was on) is completed before this upload finishes. However, she is not required to answer other new requests for files after her downlad has finished and until any pending uploads are completed.

Otherwise, if peers closed their connection immediately after their own downloads finish, no upload would ever complete under the assumption of constant d. To see this, consider a peer A who starts downloading a file at time 0 from a super-peer and assume for simplicity that all files have exactly the same size (and thus take exactly the same time d to be downloaded). Then if another peer B requests a file from A at any time greater than 0 and less than d, will not manage to complete her download because in the mean time A will have completed her own and thus won't have the incentive to remain connected for B to complete her own as well. This will be also the case even if files have different sizes with a growing probability for a download to be aborted for peers that lie further down the 'downloading trees' that will be formed.

But this is not a realistic behavioural model in the first place. Peers in general make parallel requests, they don't disconnect immediately after their downloads complete, they keep searching for files and thus a new download can start while another is in progress, etc. What our model really requires is the ability of a system designer to force peers to stay more time in the system per download and we believe that setting the upload throughput to a certain (not low) value would have the desirable result. So, in order to avoid complicated behavioural models we have made the aforementioned assumption (that peers cannot abort uploads that started before their own download finishes[5]) and ignored this extra (but uniform) cost imposed on peers. Under this assumption our model is valid and we believe that it captures the most important aspects of the proposed enforcement mechanism.

Of course, the analysis of a more detailed behavioural model in the light of the fixed upload throughput mechanism and the more accurate estimation of the relation between d and b as a function of the rest system parameters is a very interesting avenue for future work (and part of our on-going research) in this context.

5.4.5 Stability

We now continue discussing some interesting properties of our model and more sperifically the role of the parameter d. We consider the case where $Q_0 = 0$ and $u(Q) = Q^{1/2}$ as before, which leads to simpler expressions that are easier to analyze but also focuses on the amount of content built by peers themselves. For the computation of the social welfare achieved under different choices of our control parameter d we will assume that there

[5] It would be interesting to explore under which assumptions such a rule could be actually enforced. For example, one could imagine that files are transmitted encrypted using a certain key that is to be delivered to the receiver by a responsible super-peer when all its pending uploads finish.

is some initial content for bootstrapping purposes, which is withdrawn after peers start downloading content from each other.

Observe that the optimal value of d when $Q_0 = 0$ is very close to the minimum possible according to the stability condition (equation (5.5)). This is so because at the optimum $\bar{\theta} = 0.9144$ (close to 1) and $x_i = \min\left\{\theta_i, \bar{\theta}\right\}$ while the minimum feasible value of d, d_{min}, is given by

$$d_{min} = \frac{N}{(1/n)\sum_{i=1}^{n} r(\theta_i)\, f\, n}\,, \tag{5.20}$$

So since $\frac{1}{n}\left[\sum_{i=1}^{n} r(\theta_i)\right] \approx \frac{3\,N}{nf}$ for large n, d_{min} is close to d^* (recall that $d^* = \frac{3.06\,N}{nf}$).

Hence, using d^* as the optimal choice of d, if for a particular realization of the θ_is, $\frac{1}{n}\sum_{i=1}^{n} r(\theta_i) < 1/3.06$ (A), then the stability condition will not hold because in this case $d_{min}(\boldsymbol{\theta}) > d^*$. But since the mean of $r(\theta_i)$ is 1/3, the probability of this event (A) goes to zero very fast as a function of n (probably exponential fast, since (A) consists a large deviation). On the other hand, if $\frac{1}{n}\sum_{i=1}^{n} r(\theta_i) > 1/3.06$, there is a value $d^*(\boldsymbol{\theta})$ for which the social welfare would be greater than this achieved for $d = d^*$. But note that both $d_{min}(\boldsymbol{\theta})$ and $d^*(\boldsymbol{\theta})$ are quantities that need full information to compute. This is the reason we use $d^* = \frac{3.06\,N}{nf}$, which is an approximation of $d^*(\boldsymbol{\theta})$ for large n that uses the information of the distribution of the θ_is.

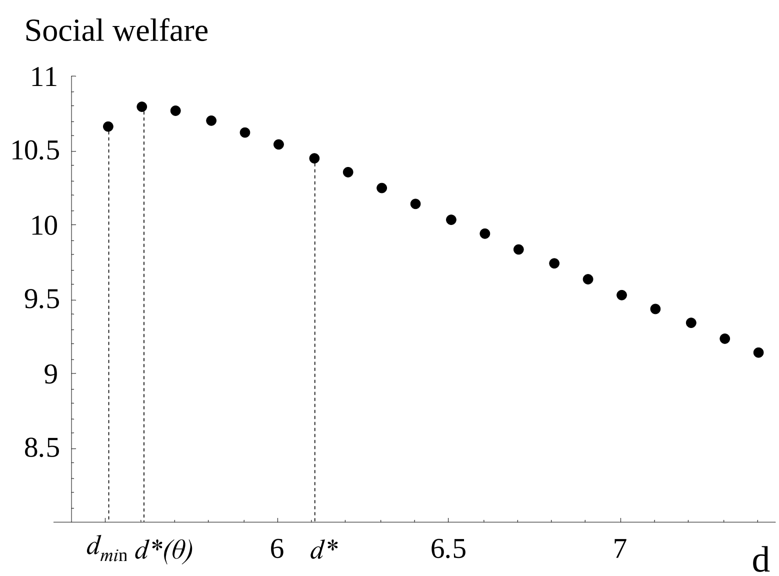


Figure 5.2: Social welfare as a function of d ($n = 100$, $\beta = 0.5$, $\alpha = 0.3$, $N = 10^4$, $f = 50$)

To make this more clear we depict in Figure 5.2 above how social welfare is affected by our choice of d by plotting the value of social welfare at equilibrium for different values of d above d_{min}. This is done for a specific realization of $\boldsymbol{\theta}$ for which $\frac{1}{n}\sum_{i=1}^{n} r(\theta_i) > 1/3.06$ and this is why d^* is not in this case the optimal selection for d, as it is shown in the figure. Moreover, as expected, the complete information optimal value of $d^*(\boldsymbol{\theta})$ is again close to the threshold value, d_{min}.

So, the fact that d^* is close to threshold value (at least for our choice of $u(.)$ and $r(.)$), suggests that under incomplete information and for moderate values of n, one could think of a more conservative tuning of d, larger than d^*, in order to avoid the unfortunate situation where $\frac{1}{n}\sum_{i=1}^{n} r(\theta_i) < 1/3.06$ and thus losing a little in efficiency for some cases but gaining in robustness of the equilibrium point.

5.4.6 Evaluation

Having characterized the levels of efficiency achieved by first-best allocation and the fixed contribution mechanism when the main contribution of peers is the time they spent in the system sharing their files (see Sections 5.4.1 and 5.4.2), a natural question is how the level of efficiency achieved by the fixed upload throughput mechanism is compared with them. But note that there are two differences of the model analyzed above with the fixed contribution scheme (and first-best).

First, the fact that the latter doesn't consider the possibility for peers to decide on their request rate and most importantly on their type, which was originally assumed fixed. However, under the fixed contribution scheme peers wouldn't have the incentive to reduce their type (and request rate) since the amount of time they will stay on-line (their contribution) is predefined and enforced by some external enforcement mechanism. This is especially so if we assume that, normally, a file download takes almost zero time and thus the time peers spent in the system is only the one dictated by the fixed contribution scheme. So, if we assume that this relation of type and request rate holds in both cases and that peers have the option to decide on their type in the case of the fixed contribution scheme as well, participating peers would always choose $x_i = \theta_i$, since choosing a x_i such that $0 \leq x_i < \theta_i$, wouldn't reduce their cost; only their utility. Obviously, such an option wouldn't change the decision of excluded peers (not to participate) as well. The same holds for the case of first-best provision as well.

Notably, the second difference plays a more important role. More specifically, in the 'fixed upload throughput' mechanism we have assumed that peers request content with a rate that depends on their true type. As we discuss in the following the existence of this relation and the shape of the corresponding function $r(.)$ have a great impact on the efficiency achieved by our mechanism. Moreover, in this new game setup created we don't know which is the optimal mechanism under incomplete information (this is a very interesting topic for future work in this context) and now the fixed contribution scheme is to be treated as just an alternative incentive mechanism (and not the one that asymptotically achieves the second-best efficiency). Actually, as it turns out the fixed contribution scheme is no longer, always, the best choice. We demonstrate this fact through numerical experiments.

In Figure 5.3 we plot the efficiency achieved on average by the first-best allocation, the fixed contribution scheme, and the fixed upload throughput mechanism as a function of a variable δ expressing the shape of the function $r(.)$ (i.e. $r(\theta) = \theta^{\delta}$). For the computation of the social welfare, for given n, f, N, and δ we did the following: Draw n values $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$ randomly from the specified distribution H and for this realisation of peers' valuations

- Compute the first-best efficiency solving the program (5.1).
- Compute the efficiency achieved using the fixed fee computed by (5.2).
- Compute the efficiency achieved by the fixed upload throughput mechanism[6].

Repeat this procedure 100 times and average the total payoffs achieved over the draws of $\boldsymbol{\theta}$ for all three mechanisms.

As one can see, the proposed mechanism achieves better efficiency levels than the fixed contribution scheme for $\delta < 2.3$. Notably, for a small range of values of δ (for $1.5 \leq \delta \leq 1.7$) the efficiency of the fixed upload throughput mechanism gets very close to first-best. The reason is that because of our assumption that the request rate of peers

[6] We again assumed a certain number of files initially available (but withdrawn after the first step of our simulation) and used for the local maximization of each peer $x_i = \min\left\{\theta_i, \left(\frac{u(Q/N)}{\delta\alpha(Q/N)d}\right\}\right)^{\frac{1}{\delta-1}}$. To avoid complicated calculations, for each δ and realization of the vector $\boldsymbol{\theta}$, we used $d = d^*(\boldsymbol{\theta})$ (computed numerically), which based on the discussion in Section 5.4.5 for large n is very close to the value d^* that we would have computed for the specific value of δ. That is, we have used a close approximation of the efficiency that would have been achieved under incomplete information. Notably, as it turns out from our experiments, $d^*(\theta)$ is always close to d_{min} but their distance increases as δ grows.

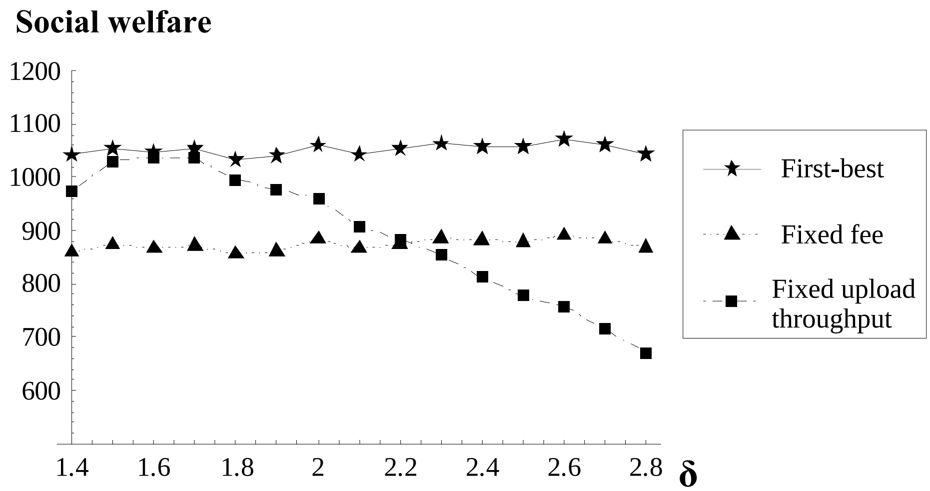


Figure 5.3: Evaluation of the 'fixed upload throughput' mechanism ($n = 10^4$, $\beta = 0.5$, $\alpha = 0.3$, $N = 10^6$, $f = 50$)

depends on their type, choosing d appropriately gets them to reveal important information concerning their actual preference parameters (by choosing the corresponding request rate assumed for the computation of d), which would be otherwise not available. As already demonstrated, for large n, d^* converges to the optimal $d^*(\boldsymbol{\theta})$ and thus by using $d = d^*$ is like having complete information of peers' types. Additionally, our mechanism allows the system designer to impose a different contribution on peers depending on their actual type and avoid exclusions without loosing the ability to acquire significant contributions from high-value peers.

However, the performance of our mechanism highly depends on the shape of this function $r(.)$, and as is shown in Figure 5.3 there are cases where it is worse than this of the fixed contribution scheme. So, the shape of the function $r(.)$, which we should stress that it is an external parameter not controlled by the system designer, strongly affects the behaviour of participating peers and thus the total amount of public good constructed and the overall efficiency achieved.

In any case, we believe that our first results are encouraging concerning the positive effect of the proposed mechanism on system's economic efficiency as this is defined by our public good model. More work is required however in order to show its incentive compatibility in a formal mechanism design formulation and assess its performance under different assumptions concerning the properties of the request rate of the peers.

5.4.7 Upload/download ratio

In our proposed mechanism we have defined the contribution of a peer to be the time required to stay on-line sharing a fixed amount of files independently from incoming requests. It would be however interesting to assess how well our mechanism would perform in terms of actual service provisions performed (i.e. uploads) on average under this scheme; we can do this now because we have incorporated into our model the rate with which peers request files.

There is an interesting observation to be made towards this end. First notice that the average number of uploads a peer offers during a unity of time being connected, equals to $U = \sum_{j=1}^{n} r(x_j)(f/N)$ (the total request rate multiplied with the probability that one of her f files is requested). So, the total expected number of uploads a peer i will offer under our scheme will be equal to $r(x_i)(Q/N)dU$, i.e. the fraction of time she stays on-line due to sucessful content requests multiplied by the average number of uploads per unity of time. But at equilibrium, when $Q_0 = 0$, $dU = 1$ (from (5.6)). Hence, in this case peers will actually upload as many files as downloaded on average ($r(x_i)(Q/N)$), independently from the value of d. Similarly, when $Q_0 > 0$, from (5.17) we have that the upload/download ratio at equilibrium will be $\frac{Q-Q_0}{Q}$.

But more interestingly, having estimated the average number of uploads per download peers offer in our system, we can now compare our scheme with the very popular mechanism enforcing the equation of downloads with uploads using virtual currencies or other means (we will refer to this scheme as the '1-1 scheme' since it dictates a 1-1 upload/download ratio for each peer). In order to do this we will assume that the number of files all peers share under this scheme is again f although in reality this decision is part of peers' strategy in such a system and as already mentioned in Chapter 2, there is actually a 'congestion game' [Rosenthal (1973)] to be played amongst peers to that respect: each peer would try to share the files that generate the more incoming requests in order to reduce the time they will have to stay on-line for acquiring the necessary credit to satisfy their demand. This would lead to some files becoming 'congested' and this would lead to different file sharing strategies, and so on.

We also assume that the cost of a peer participating in a system with a rule dictating a 1-1 uploads/downloads ratio will be again the time she will be required to stay on-line sharing her files. Notice that this time now depends directly on the actual behaviour (i.e.

request rates) of the rest of the peers, since a peer should stay on-line as long as it is needed for her to acquire the necessary credit (e.g. tokens) for satisfying her own demand. This is the most fundamental differentiation between the two mechanisms: in our memory-less mechanism we impose a certain fixed contribution cost on each download in order to maximize a certain objective (i.e. economic efficiency) while in a virtual market of uploads the only task of the system designer is to impose a certain constraint on peers' behaviour independent from their preference parameters or the size of the system.

Returning to the notation of our model, there is again a certain amount of time d, let it be denoted by d_{1-1}, that a peer will have to stay on-line sharing her files for each download, which now depends on the actual total request rate in the system. That is, $d_{1-1} = \frac{N}{\sum_{i=1}^{n} r(x_i)\,f}$. To see why, notice that d_{1-1} is the time required by a peer to stay on-line on average to offer one upload. And this is exactly the time that should be 'contributed' by each peer for each download she performs. (We have also assumed for simplicity that without any restriction on the upload throughput the average download time is zero).

It is not easy to compute analytically the corresponding equilibrium. However, we expect that the level of efficiency achieved under the 1-1 scheme will be in general less than the one achieved by our proposed mechanism. The reason is that in our case we compute the optimal value of d so as to maximize efficiency, while in the case of the 1-1 rule there is no reason to expect that in equilibrium the value of d_{1-1} will be the globally optimal. The numerical experiments we have conducted indicate that when an equilibrium is reached, its efficiency is on the average 40% less than this of the fixed upload throughput scheme. Additionally, as also indicated by our numerical experiments, the existence of equilibrium is not guaranteed under the 1-1 scheme but depends in general on the initial conditions of the experiment (i.e. the amount of content assumed initially available for bootstrapping purposes).

Another weakness of the 1-1 scheme in the context of our model is that it doesn't take advantage of the number of files Q_0 possibly contributed by super-peers as does our mechanism under which the more is Q_0 related to the overall provision Q the less uploads/downloads ratio is required from normal peers. In particular, according to (5.17), the ratio of uploads/downloads performed by each peer on average will be $\frac{Q-Q_0}{Q}$ at equilibrium. But even more importantly, the 1-1 scheme would always fail to reach an equilibrium when $Q_0 > 0$ under our model. And the reason is exactly the fact that it doesn't allow for a different ratio of uploads/downloads than 1-1 as the stability condition (5.17), for

$Q_0 > 0$, dictates. Hence, in the case of the 1-1 scheme, the upload/download ratio should be adapted according to Q_0 in order for the system to be stable, but this is not practical in general.

On the other hand, the 1-1 scheme has the advantage of controlling the actual service provision while in our scheme this is done implicitly and as we discuss in the following there could be cases where certain attacks which would be successful in bypassing its actual enforcement. But this is also the case for the 1-1 scheme which requires accounting mechanisms for which there are also serious attacks to be addressed (such as whitewashing and false trading), unless there is the possibility for direct exchange, either bilateral or multilateral (see also Section 2.6.1).

Finally, there is an additional qualitative characteristic that differentiates the two approaches. The 1-1 scheme requires an ex ante contribution from peers in order to acquire service while the requirement in our scheme is somehow ex post since everyone is free to consume with no restrictions and contribution comes as a consequence from consumption. Moreover, in our scheme contribution is fixed (in terms of availability per download) and thus requires in general less mental burden for the decisions taken and poses less uncertainty on the corresponding costs compared to the 1-1 scheme. So, we believe that the fixed upload throughput scheme is more friendly in terms of simplicity and elasticity, and it wouldn’t harm significantly the ‘community spirit’ inherent in many p2p application, which actually seems to play a very important (often decisive) role for their success.

5.5 Effectiveness of the ‘contribute while consuming’ mechanisms

Our memory-less mechanism belongs to a more general class of enforcement mechanisms which we could term ‘contribute while consuming’ mechanisms. These mechanisms besides being memory-less (they don’t require any long-term accounting functionality to be implemented) they additionally avoid the synchronization in terms of service provision required by other memory-less mechanisms such as the direct exchange of resources. However, the key challenge of any enforcement mechanism of this kind (one that ensures a peer is contributing to the community while consuming) is to what extend a peer providing service to another can check if the consuming peer actually contributes throughout that time her own resources.

For example, the most serious weakness of applications of the proposed mechanism is likely to be attacks based on the provision of invalid content, particularly where the cost of provision of an individual piece of content is high and its probability of request is low. Newcomers may then often be able to get away with advertising invalid content for long enough to successfully complete their downloads. However, in other scenarios, where the cost of provision of new content is less high, then this could be a much less attractive attack. We believe that in the future many groups of people will have much content on their PCs, perhaps legally available, and of wide interest within their social network. Then the cost for providing this content will be small (not much larger than the cost of creating a realistic faked file) with any advantage of cheating outweighed by the probability of having your downloads stopped.

Other incentive mechanisms could be also built on top of our basic incentive scheme, such as the construction of social networks based on taste within which peers will have an incentive to make available the files that best represent their personality. Other psychological incentives, such as encouraging peers to identify with the well-being of the community as a whole, also appear to have an effect in some file-sharing communities and provide interesting avenues for exploration. We believe that the formal discussion of social incentives in the context of p2p applications is an interesting direction for further work.

Finally, we should note that besides the direct exchange and our much less strict mechanism in terms of actual service provision required by peers, one could think of 'contribute while consuming' mechanisms that lie between these two extremes. We sketch below such a mechanism which could provide the suitable incentives to nodes to forward packets belonging to bilateral communications between distant nodes of an ad-hoc network (see Section 2.1.3) in order to demonstrate the possible applicability of this class of incentive mechanisms in other contexts as well.

5.5.1 A 'contribute while consuming' mechanism for packet forwarding in ad-hoc networks

In ad-hoc networks, nodes are required to forward packets indifferent for them and without any direct benefit. Instead of implementing trading (token-based) or indirect reciprocity (reputation-based) schemes, we wish to explore the possibility to force the nodes while consuming the resources of a peer node (they forward one of their packets through it) to contribute their resources to *any* other node in the system. As we describe in the following, in this case, unlike our memory-less mechanism for content availability, this contribution could be defined to be an actual service provision towards some other nodes of the network. Again the challenge in terms of enforcement would be to verify the validity of this contribution.

A memory-less mechanism in this context should exploit the fact that participating nodes would be in general interested themselves to forward their own packets and thus it could require from peers to forward a certain number of additional packets piggybacked with their own. In addition to acting as a ‘proof’ for contribution, piggybacking will have a positive effect on the overall cost for packet forwarding, since “the number of packets has greater impact on energy consumption than packet size does” [Feeney and Nilsson (2001); Feeney (2001)].

But the requirement for a peer to first receive a number of requests for forwarding before sending her own packets could be very restrictive in general. However, when the corresponding application is delay-tolerant and all peers have symmetric needs and thus all want to send packets towards some destination (e.g. in the context of a p2p implementation of a social mobile application such as serendipity [Eagle and Pentland (2005)]) all nodes will have in general packets to send and thus the probability to receive requests would be high enough,[7]. Hence, in this context, such a scheme could be used in order to provide the appropriate incentives to peers to contribute their resources.

Then a major threat the system designer may face is the fact that a node can produce fake packets to deceive the proposed incentive mechanism avoiding the cost of receiving packets from the network. In general, however, peers are willing to receive packets in case they are the destination and thus the cost for performing such an attack may not justify the corresponding gains. But in cases where such an attack is beneficial a system designer should devise ways for peers to detect such fake packets. As already mentioned, this is the major challenge one should address when implementing a ‘contribute while consuming’ mechanism. But we believe that its attractive properties constitute the research of practical

[7] In any case, there could be nodes (e.g. those at the edges of the network) that do not have enough packets to piggyback. However, such nodes could rely on the need of others to piggyback packets and thus have their packets —with much smaller probability— reach their destination. But what if all nodes rely on this fact and never piggyback others’ packets? In certain cases (with low traffic) this could be a profitable attack but it is the nodes at the center of the network who will discourage this behaviour since they will receive more packets than required for their needs and as a result they would give priority to piggybacked packets. Thus moving to the center of the network only such packets will manage to get through.

ways to verify contribution to the system as a whole while consuming, for different types of p2p applications, an interesting avenue for future work.

5.6 Summary and future work

In this chapter we have proposed a memory-less incentive mechanism addressing an important dimension of a peer's contribution in a p2p file sharing system towards content availability: the amount of time she stays on-line sharing her files. We studied the implementation issues for its enforcement and formulated a suitable economic model that captures the main aspects of such a system.

This is an important contribution of this dissertation because it provides a practical solution for addressing the free riding problem in p2p file sharing systems. Moreover, the corresponding theoretical framework enables us to estimate the performance of our mechanism in terms of economic efficiency and compare it with the optimal one and this achieved by the fixed contribution mechanism, but also with other alternative practical incentive mechanisms such as the popular rule dictating equal consumption and contribution of resources in terms of actual uploads/downloads for each peer. We do believe that this class of 'contribute while consuming' mechanisms could have applications in other contexts as well and we hope that our analysis in this chapter provides the means for understanding the underlying trade-offs and thus constitutes a first step towards providing incentives for resource provision in realistic p2p applications.

Our on-going work includes the formulation of alternative, more detailed, models incorporating the notion of request rate and the regulation of service time required per download in order to validate the results of our initial modelling approach. For example, one could consider models in which the request rate of peers is in general independent from their type and/or their utility is a function of actual download rate (peers then would wish control their request rate in order to achieve a target download rate).

Moreover, we wish to model in more detail the transactions that take place, the trees of downloaders formed, their interdependencies, and how the total average download time and thus content availability is affected by the tuning of the upload throughput used by all peers.

Finally, a very interesting topic for future research is the design of a real system implementing such an enforcement mechanism and providing the means for peers to set and tune its critical parameters based on dynamic measurements of the main system's attributes such as its size, its content availability, and overall activity (e.g. average request rate and peer availability, percentage of aborted downloads, etc.).

Chapter 6

More general cost functions

In this chapter we propose an alternative economic model capturing the costs incurred due to content requests (mainly due to uploading) and explore the use of simple system rules enforcing a certain relation between consumption and contribution (generalizing the popular uploads=downloads rule often discussed in the p2p economics literature) for improving the economic efficiency of the system.

Parts of this chapter are joint work with Robin Mason (see [Antoniadis et al. (2003)]).

6.1 Negative externalities

Since our main focus in this dissertation is on content availability, we have made the important assumption that serving content (i.e uploading) does not impose significant costs on peers. Below we discuss how our basic economic model could be extended in order to account for such costs (due to uploading or congestion).

6.1.1 Uploading costs

The first possible extension to our model is to assume that the cost of sharing files depends also on the corresponding request rate that these files generate due to uploading costs: a peer's cost is proportional to the rate at which she serves upload requests (e.g. because she pays per volume charges to her ISP or her upload capacity is deprived from her own use).

Assuming that all peers have the same request rate and that files are equally popular means that the total cost incurred by all peers will be proportional to the product of the number of participating peers (that generate the requests) and the number of unique files, i.e., $c(Q) = (\sum_i \pi_i)Q$. If peers can only access files held within a certain neighbourhood of their location, this might be better modelled as $c(Q) = (\sum_i \pi_i)^\beta Q$, where $0 < \beta < 1$. This modelling decision would be suitable for pure p2p systems like Gnutella where peers in general don't have access to the full content made available due to the limitations posed on the search algorithm (e.g. the value of the 'time to live' field for the propagation of content queries) in order for the corresponding communication overhead to be constrained. Courcoubetis and Weber (2005) show that the main optimality results holds for this case also.

Another approach could be to assume that the request rate is somehow related with the type of a peer, θ_i (e.g. by a function $r(\theta_i)$, the same for all peers). This approach we followed in the previous chapter as well (assuming that $r(.)$ is a convex function) for capturing the effect the proposed incentive mechanism would have on the request rate of peers. What we gain with this assumption is that peers would still be characterized by a single variable, their preference parameter θ_i. And this fact makes it probably feasible to extend our public good model to account for uploading costs since again this cost would depend on the number of participating peers (since request rates are strictly increasing on peers' types).

But what if request rate is an independent variable of a peer's behaviour and thus, in addition to her type θ_i, a peer i is also characterized by her average request rate r_i? Then the provision cost would depend not only on the number of participating peers but on who exactly they are and this means that the results of Chapter 4 would not hold anymore. We will now need to regulate both the contributions of peers in terms of files shared per unity of time but also their request rate since this has negative externalities on the net benefit of all the other peers. So, in this chapter we propose a more detailed economic model for p2p file sharing, which treats peers' request rate as an independent variable. We again seek for system rules which regulate peers' behaviour by relating the allowed request rates and with the number of files shared and show that a linear rule is adequate to maximize efficiency. However, we also show that such a rule, unlike prices, should be personalized in order for efficiency to be maximized. Moreover, the optimal uniform rule under incomplete information in this context is unknown and we can use only simple heuristics based on averaging of the optimal personalized rule coefficients in order to asses how a uniform rule would perform in this context.

An even more detailed modelling approach could assume that not all peers offer equally valuable services to the rest of the community. This could be due to taste similarities between certain peers and the opposite. Buragohain et al. (2003) model this situation by assuming that peer i acquires different value from different peers j, denoted as b_{ij}. However, this assumption further complicates the analysis of the model. Actually in the work of Buragohain et al. (2003) all analytical results presented concern the homogeneous case and this more detailed scenario is studied only through simulations. In Section 6.4.6 we present a similar model, originally designed for the WLAN peering application (see Antoniadis et al. (2003b)), which assumes that each peer i has different request rates for the service provided by different peers j, r_{ij} and additionally each peer offers a different level of quality of service (e.g. a different blocking probability) Q_i. We compute for this case also the optimal system rules, which again, as expected, must be personalized.

6.1.2 Congestion

Congestion costs as a function of the request rate of a peer are mainly incurred by other peers requesting files from the same site due to the fact that upload capacity is limited. In this case the total request rate does not increase the cost of peers due to file sharing (as in the case of uploading costs) but reduces their utility due to worse quality of service received (e.g. low throughput, blocking probability, etc.).

So, extending our public model to account for service quality degradation due to congestion, the net benefit of peer i would become

$$b_i = \theta_i u(Q, \sum_j r_j) - c(f_i)\,, \tag{6.1}$$

where $u(\cdot)$ is increasing and concave in Q and decreasing on $\sum_j r_j$. But adding an additional parameter for characterizing peers makes again the problem very difficult to solve. We could try to solve this problem making the simplifying assumption of the previous chapter, that there is a function $r(.)$ such that the request rate of a peer is given by $r_i = r(\theta_i)$. This is left for future research.

Finally, notice that congestion despite the cost that imposes on participating peers, it constitutes an opportunity for differentiated treatment and thus provides some means incentivizing contribution. For example, in case of congestion peers that are cooperative could be given priority. As already discussed in Section 4.6.1, such a differentiated service could be used only when peers can be categorized to different groups based on an objective metric or if the suitable incentives are given to self-select. Similar approaches have also been proposed in the context of accounting mechanisms (such as reputation) which provide the necessary information that could be used to differentiate between peers (e.g. [Papaioannou and Stamoulis (2004); Ranganathan et al. (2003)]). We believe that the theoretical study of such approaches in order to compare them in terms of Nash equilibria and efficiency achieved is a very interesting avenue for future research.

6.2 An alternative economic model of peering

We formulate below an alternative economic model that captures uploading costs assuming request rate is independent from the preference parameter. So, each peer i decides on the rate r_i at which it requests content from other peers, and on the number of files f_i it shares. Her net benefit with n pariticipating peers and n-vectors $\mathbf{r} = (r_1, \ldots, r_n)$ and $\mathbf{f} = (f_1, \ldots, f_n)$ of request rates and shared files is

$$b_i(\mathbf{r}, \mathbf{f}) = u_i\left(r_i, \sum_{j=1}^{n} f_j\right) - c_i\left(\sum_{j=1}^{n} r_j, f_i\right). \tag{6.2}$$

Here, the utility function $u_i(\cdot, \cdot) \geq 0$ is assumed to be continuously differentiable, increasing and strictly concave in both of its arguments, and the cost function $c_i(\cdot, \cdot) \geq 0$ to be continuously differentiable, increasing and strictly convex in both its arguments, for all i. Further, the functions are assumed to have properties such that the maximization problems considered in this chapter (corresponding to equilibrium and the efficient solution) have unique solutions. Note that for simplicity we assume that the peer gains a benefit (as well as incurring a cost) from the files that it shares; and incurs a cost (as well as gaining a benefit) from its own requests. Notice that the above model is a generalization of our simple public good model analyzed in the previous chapters[1].

File sharing is again a public good. But content request is now a public bad: a request rate by agent i increases the load on all other agents equally (in an appropriate statistical sense). The standard problem with public goods is that the (Nash) equilibrium in which

[1] where $u_i\left(r_i, \sum_{j=1}^{n} f_j\right) = \theta_i u(Q(\sum_{j=1}^{n} f_j))$ and $c_i\left(\sum_{j=1}^{n} r_j, f_i\right) = f_i$ (recall that content request do not incur cost to peers in our public good model analyzed in Chapter 4).

agents determine their contribution levels to maximize their own utility is, typically, inefficient relative to the social optimum, in which contributions are set to maximize the sum of all utilities, due to free-riding. But the same holds also for the negative externalities imposed due to content requests in our generalized model. That is, a peer maximizing her own net benefit wouldn't limit her request rate in order to reduce the cost imposed to others. In other words, each peer i would prefer to step up her request rate r_i and reduce (to zero) her contribution f_i and, in equilibrium, request rates will be too high and the number of shared files too low, relative to the efficient levels.

Peer i seeks to maximize her net benefit: $\max_{\{r_i,f_i\}} b_i(\mathbf{r},\mathbf{f})$, taking as given the request rates and files shared of all other peers. A Nash equilibrium is comprised of vectors of request rates and files shared such that all peers are simultaneously maximizing their net benefits. The first-order conditions for peer i's maximization problem are

$$\frac{\partial u_i\left(r_i, \sum_{j=1}^n f_j\right)}{\partial r_i} - \frac{\partial c_i\left(\sum_{j=1}^n r_j, f_i\right)}{\partial r_i} \leq 0 \qquad r_i \geq 0, \tag{6.3}$$

$$\frac{\partial u_i\left(r_i, \sum_{j=1}^n f_j\right)}{\partial f_i} - \frac{\partial c_i\left(\sum_{j=1}^n r_j, f_i\right)}{\partial f_i} \leq 0 \qquad f_i \geq 0. \tag{6.4}$$

Let the equilibrium choices of peer i be denoted $\hat{r}_i$ and $\hat{f}_i$.

A benevolent social planner (the system designer) would choose rates $\mathbf{r}$ and files $\mathbf{f}$ so as to maximize the sum of peers' net benefits:

$$S \equiv \max_{\mathbf{r},\mathbf{f}} \sum_{i=1}^{n} \left[u_i\left(r_i, \sum_j f_j\right) - c_i\left(\sum_j r_j, f_i\right)\right].$$

The efficient request rates and number of shared files are given by the first-order conditions

$$\frac{\partial u_i(r_i, \sum_j f_j)}{\partial r_i} - \sum_j \frac{\partial c_j(\sum_k r_k, f_j)}{\partial r_i} \leq 0 \qquad r_i \geq 0, \tag{6.5}$$

$$\sum_j \frac{\partial u_j(r_j, \sum_k f_k)}{\partial f_i} - \frac{\partial c_i(\sum_j r_j, f_i)}{\partial f_i} \leq 0 \qquad f_i \geq 0 \tag{6.6}$$

for $i = 1, \ldots, n$. Denote these rates and number of files r_i^* and f_i^*.

An important difference between equations (6.3)/(6.4) and (6.5)/(6.6) is the presence in the latter of externalities, which can be measured by the following sums

$$-\sum_{j\neq i}\frac{\partial c_j(\sum_k r_k, f_j)}{\partial r_i}, \qquad \sum_{j\neq i}\frac{\partial u_j(r_j, \sum_k f_k)}{\partial f_i}.$$

That is, when peer i increases r_i she imposes a cost of $\sum_{j\neq i}\frac{\partial c_j(\sum_k r_k, f_j)}{\partial r_i}$ on all other peers (the externality is negative), while by increasing f_i all other peers acquire benefit $\sum_{j\neq i}\frac{\partial u_j(r_j,\sum_k f_k)}{\partial f_i}$ and thus the externality is positive. Then, the assumptions that we make about $u_i(\cdot,\cdot)$ and $c_i(\cdot,\cdot)$ lead naturally to our previous remark: without any explicit incentive mechanism, in the context of our model, the social efficiency at equilibrium will be almost zero since f will tend to zero. In the next section, we analyse different alternative approaches for ensuring that peers consider these externalities when deciding on their request rate and files shared per unity of time.

6.3 Lindahl prices

6.3.1 Fully efficient prices

As already mentioned, the standard economic approach to the problem of externalities is to establish prices (i.e. lindahl prices) that close the gap between private and social incentives. Comparison of equations (6.3)–(6.6) show that the relevant prices are

$$p_i^r = \sum_{j\neq i}\frac{\partial c_j(\sum_k r_k^*, f_j^*)}{\partial r_i}, \tag{6.7}$$

$$p_i^f = -\sum_{j\neq i}\frac{\partial u_j(r_j^*, \sum_k f_k^*)}{\partial f_i}, \tag{6.8}$$

so that peer i is charged p_i^r on her rate of requests and receives p_i^f for each file that she shares. These prices are exactly the total externalities (negative and positive) imposed by peer i on all other peers. So, now each peer will decide on her rate and contribution based on the following optimization

$$\underset{r_i,f_i}{\text{maximize}}\left[u_i\left(r_i,\sum_{j=1}^n f_j\right) - p_i^r r_i + p_i^f f_i - c_i\left(\sum_{j=1}^n r_j, f_i\right)\right] \tag{6.9}$$

Notice that the prices are *non-uniform*: in general, $p_i^r \neq p_j^r$ for $i\neq j$ (and similarly for

p_i^f).

6.3.2 Approximately efficient uniform prices

In this section, we show that when the p2p system is sufficiently large, the non-uniform, fully efficient prices can be approximated by uniform prices. In order for the argument to work, we make the following assumption.

Assumption 1 (Bounded Heterogeneity) *There exist* $\bar{b}, \underline{b}, \bar{c}$ *and* $\underline{c}$*, where* $0 < \underline{x} \leq \bar{x} < +\infty,\ x \in \{b, c\}$ *such that*

$$\begin{aligned} \sup_{f \in \mathbb{R}_+, r \in \mathbb{R}_+} \frac{\partial u_i(r,f)}{\partial f} &\in [\underline{b}, \bar{b}], \\ \sup_{f \in \mathbb{R}_+, r \in \mathbb{R}_+} \frac{\partial c_i(r,f)}{\partial r} &\in [\underline{c}, \bar{c}] \end{aligned}$$

for all i.

Definition 2 *Let*

$$\begin{aligned} p^r &\equiv \sum_{j=1}^{n} \frac{\partial c_j(\sum_k r_k^*, f_j^*)}{\partial r}, \\ p^f &= -\sum_{j=1}^{n} \frac{\partial u_j(r_j^*, \sum_k f_k^*)}{\partial f}. \end{aligned}$$

Note that the prices p^r and p^f are uniform i.e., do not depend on the identity of any peer.

Proposition 3 *For given values of* $\bar{b}, \underline{b}, \bar{c}$ *and* $\underline{c}$*, and for a constant* $\epsilon > 0$*, there exists a critical number of peers* $\bar{n}$ *such that for a number of peers* $n > \bar{n}$*,*

$$\frac{|p_i^x - p^x|}{p^x} < \epsilon, \quad x \in \{r, f\}, \quad i = 1, \ldots, n.$$

Proof. The proposition will be proved for p^r and peer i: an identical argument holds for other peers and p^f. Note that $|p_i^r - p^r| = \frac{\partial c_i(\sum_k r_k, f_i)}{\partial r_i}$. Hence $|p_i^r - p^r| \leq \bar{c}$. In addition, $p^r \geq n\underline{c}$. Therefore

$$\frac{|p_i^r - p^r|}{p^r} \leq \frac{\bar{c}}{n\underline{c}}.$$

Define $\bar{n} \equiv \bar{c}/\epsilon\underline{c}$. By assumption 1, for n sufficiently large, $n > \bar{n}$ and so $(|p_i^x - p^x|)/p^x < \epsilon$. □

This result means that if we had the ability to charge peers for their consumption and reward them for their contribution we could maximize the social welfare without the need to apply different prices to different peers by computing good approximations of the optimal uniform prices even when complete knowledge about peers' private information is not available (see Section 6.4.4 below). However, even in this case, the existence of a benevolent system manager is required who would be capable of accounting all peers' transactions and be responsible for all money transfers which notably could lead to a deficit; and thus the system designer should be willing to finance the efficient system operation.

6.4 Reciprocity rules

In the case of realistic p2p systems it is important to develop non-priced incentives to encourage efficient behavior on the part of peers. In this section, we consider the use of rules that must be obeyed (and possibly enforced) by peers. More specifically, our attention is focussed on *simple system rules*. By simple, we mean that any policy that is used to regulate the behaviour of peers must be *anonymous*: the policy cannot be tailored to the identity of a peer, but only to its current behaviour. By rule, we mean a policy that controls *directly* behaviour, through constraints on request rates and files shared, rather than policy that operates indirectly through prices. Both restrictions are motivated by pragmatism. In practice, it will be difficult to establish unambiguously the identity of peers. And the overhead of a price system (establishing a currency, monitoring payments etc.) is likely to be too large.

6.4.1 Fully efficient rules

Any simple rule must take the form $r = \rho(f)$ i.e., the rate at which a participating peer can request files is a function of the number of files that it makes available for sharing. The function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ is, ideally, independent of the identity of a peer, and can, in principle, take any form. As a first step in the analysis, suppose that $\rho(\cdot)$ is a continuously differentiable function. Each peer therefore chooses the number of files that she shares (which then determines the rate at which she can request files) according to the first-order condition (for peer i):

$$\left(\frac{\partial u_i\left(\rho(f_i),\sum_j f_j\right)}{\partial f_i}-\frac{\partial c_i\left(\sum_j \rho(f_j), f_i\right)}{\partial f_i}\right) + \left(\frac{\partial u_i\left(\rho(f_i),\sum_j f_j\right)}{\partial r_i}-\frac{\partial c_i\left(\sum_j \rho(f_j), f_i\right)}{\partial r_i}\right)\rho'(f_i)=0 \quad (6.10)$$

for $i = 1, \ldots, n$, where $\rho'(\cdot)$ denotes the first derivative of $\rho(\cdot)$. This difference equation for $\rho(\cdot)$ is not, generally, solvable. One route for progress is to guess at a solution; for example, suppose that $\rho(\cdot)$ is linear. That is,

$$r_i \leq \alpha_i + \beta_i f_i$$

where α_i and β_i are constants, which for the moment let's assume that could be different for each peer i.

To make the matter as clear as possible, suppose that the rule binds for all peers, and thus any peer i chooses its rate of request so that $r_i = \alpha_i + \beta_i f_i$. The peer i therefore faces the (constrained) maximization problem

$$\max_{f_i} u_i\left(\alpha_i+\beta_i f_i, \sum_j f_j\right) - c_i\left(\alpha_i+\beta_i f_i+\sum_{j\neq i} r_j, f_i\right).$$

The first-order condition for the maximization is

$$\left(\frac{\partial u_i\left(\alpha_i+\beta_i f_i,\sum_j f_j\right)}{\partial f_i}-\frac{\partial c_i\left(\alpha_i+\beta_i f_i+\sum_{j\neq i} r_j, f_i\right)}{\partial f_i}\right) + \beta_i\left(\frac{\partial u_i\left(\alpha_i+\beta_i f_i,\sum_j f_j\right)}{\partial r_i} -\frac{\partial c_i\left(\alpha_i+\beta_i f_i+\sum_{j\neq i} r_j, f_i\right)}{\partial r_i}\right)=0. \quad (6.11)$$

According to (6.11), it is easy to see that if

$$\beta_i = \frac{-\left(\frac{\partial u_i(r_i^*, \sum_k f_k^*)}{\partial f_i} - \frac{\partial c_i(\sum_k r_k^*, f_i^*)}{\partial f_i}\right)}{\frac{\partial u_i(r_i^*, \sum_k f_k^*)}{\partial r_i} - \frac{\partial c_i(\sum_k r_k^*, f_i^*)}{\partial r_i}},$$

at equilibrium no peer has the incentive to deviate from the optimal strategy, (r_i^*, f_i^*) –see Section 6.4.5 for the stability properties of this model when linear rules are to be used. Moreover, from the first-order conditions of (6.9) we get that $p_i^r = \frac{\partial u_i(r_i^*, \sum_k f_k^*)}{\partial r_i} - \frac{\partial c_i(\sum_k r_k^*, f_i^*)}{\partial r_i}$ and $p_i^f = -\frac{\partial u_i(r_i^*, \sum_k f_k^*)}{\partial f_i} - \frac{\partial c_i(\sum_k r_k^*, f_i^*)}{\partial f_i}$.

So, if efficiency is to be maximized, β_i should be set so that

$$\beta_i = \frac{p_i^f}{p_i^r}, \tag{6.12}$$

where p_i^f and p_i^r as in equations (6.7) and (6.8) respectively. Then

$$\alpha_i = r_i^* - \beta_i f_i^*. \tag{6.13}$$

Hence, a linear rule of the form $r_i = \beta_i f_i + \alpha_i$ for each peer i is indeed adequate for maximizing social welfare. However, one challenge apparent in the expressions for α_i and β_i in equations (6.12) and (6.13) is the amount of information required to compute these rule coefficients. In other words, we have not yet satisfied our requirement for anonymity according to which a uniform function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ for which $r_i \leq \rho(f_i), \forall i$ is required.

6.4.2 Approximately efficient uniform rules

Ideally, as in the case of prices, we would like to compute uniform coefficients α and β for which the linear rule $r_i = \alpha + \beta f_i, \forall i$ would lead to an approximately efficient equilibrium. A similar argument to the one used previously shows that a uniform slope coefficient β can be used to approximate the fully efficient coefficients β_i in the case of large p2p systems. (The proof of the following proposition follows closely that of Proposition 3 and so is omitted.)

Proposition 4 *Let*

$$\beta \equiv \frac{\sum_{j=1}^{n} \frac{\partial u_j(r_j^*, \sum_k f_k^*)}{\partial f}}{\sum_{j=1}^{n} \frac{\partial c_j(\sum_k r_k^*, f_j^*)}{\partial r}} = \frac{p^f}{p^r}.$$

For given values of $\bar{b}, \underline{b}, \bar{c}$ and $\underline{c}$, and for a constant $\epsilon > 0$, there exists a critical number of peers $\bar{n}$ such that for a number of peers $n > \bar{n}$,

$$\frac{|\beta_i - \beta|}{\beta} < \epsilon, \quad i = 1, \ldots, n.$$

This means that we could use (6.14) to compute the uniform slope coefficient β of a uniform linear rule $\rho(.)$. However, since the computation of α_is depends on the actual request rate and files shared of each individual peer ($\alpha_i = r_i^* - \beta_i f_i^*$) and not on their sum (as in the case of β_is) a uniform intercept cannot be used, except from the very special case where the points $(r_i^*, f_i^*)_{\{i=1,\ldots,n\}}$ all lie on the same straight line with slope β. Hence, in general, the planner must set $n+1$ coefficients to implement approximately efficient rules. Formally, according to (6.10) the function $\rho(\cdot)$ must satisfy the following $2n$ equations:

$$\rho'(f_i^*) = \frac{\sum_{j \neq i} \frac{\partial u_j\left(r_j^*, \sum_k f_k^*\right)}{\partial f_i}}{\frac{\partial u_i\left(r_i^*, \sum_k f_k^*\right)}{\partial r_i} - \frac{\partial c_i\left(\sum_k r_k^*, f_i^*\right)}{\partial r_i}}; \qquad (6.14)$$

$$\rho(f_i^*) = r_i^*, \qquad (6.15)$$

which means that a linear uniform function $\rho(\cdot)$ rule wouldn't be in general sufficient to achieve full efficiency.

A two peer example

A specific very simple theoretical example helps to understand the difficulty of computing approximately efficient uniform (and linear) rules. Suppose that there are $n = 2$ peers; and that the net benefit functions of the peers are of the form

$$b_i(f_1, f_2, r_1, r_2) = v_i(f_1 + f_2 + r_i^2) - k_i((r_1 + r_2) + f_i^2)$$

where v_i and c_i are strictly positive constants. In this simple case,

$$f_1^* = \frac{v_1 + v_2}{2k_1}, \qquad f_2^* = \frac{v_1 + v_2}{2k_2},$$
$$r_1^* = \frac{k_1 + k_2}{2v_1}, \qquad r_2^* = \frac{k_1 + k_2}{2v_2}.$$

Equations (6.14) and (6.15) become

$$\rho'(\frac{v_1 + v_2}{2k_1}) = \frac{v_2}{k_2}, \qquad \rho'(\frac{v_1 + v_2}{2k_2}) = \frac{v_1}{k_1},$$
$$\rho(\frac{v_1 + v_2}{2k_1}) = \frac{k_1 + k_2}{2v_1}, \qquad \rho(\frac{v_1 + v_2}{2k_2}) = \frac{k_1 + k_2}{2v_2}.$$

Suppose first that $v_1 > v_2$, $k_1 > k_2$ and $v_1/k_1 < v_2/k_2$. In this case, the higher valuation peer (peer 1) also faces the greater cost; and has the lower benefit/cost ratio. Then the relevant information about the function $\rho(\cdot)$ is shown in Figure 6.1.

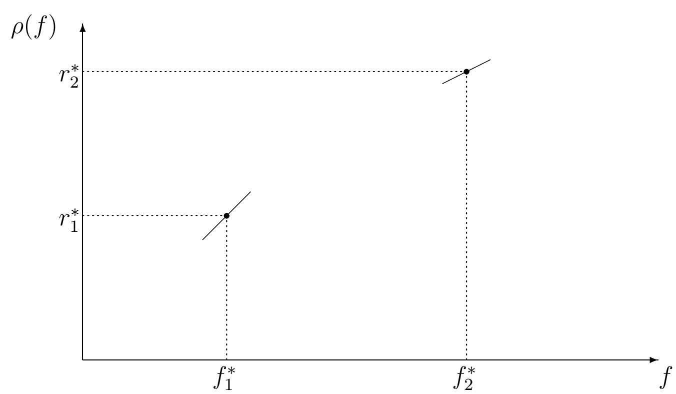


Figure 6.1: The function $\rho(\cdot)$ when $v_1 > v_2$, $k_1 > k_2$ and $v_1/k_1 < v_2/k_2$

It is natural in this case to make $\rho(\cdot)$ a concave function of f; doing so ensures that the first-order condition (6.10) is necessary and sufficient to determine an interior maximum.

Suppose now that $v_1 > v_2$ and $k_1 < k_2$, so that $v_1/k_1 > v_2/k_2$. The relevant information about the function $\rho(\cdot)$ in this case is shown in Figure 6.2. Clearly in this case, a simple concave function does not satisfy the equations for $\rho(\cdot)$. The complication then arises that the function used will not satisfy the second-order conditions that are required for equation

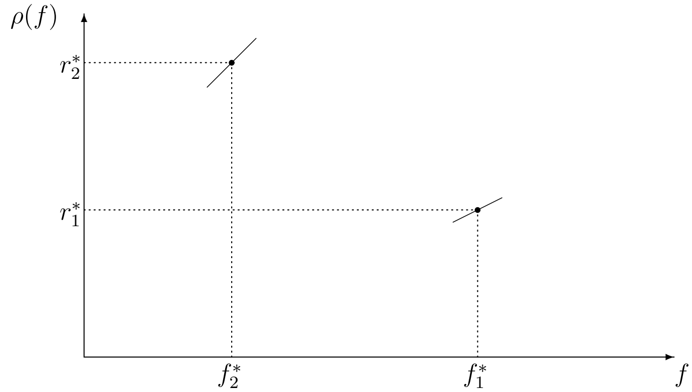


Figure 6.2: The function $\rho(\cdot)$ when $v_1 > v_2$ and $k_1 < k_2$

(6.10) to determine the global solution.[2]

6.4.3 Heuristics

If, however, the planner did want to use a uniform linear rule and compute coefficients α and β, then they should be set according to the following program:

$$\max_{\alpha \in \mathbb{R},\ \beta \in \mathbb{R}} \sum_{j=1}^{n} \left[u_j \left(\alpha + \beta f_j, \sum_k f_k \right) - c_j \left(\sum_k \langle a + \beta f_k), f_j \right) \right] \tag{6.16}$$

$$\text{s.t.} \quad f_i = \arg\max_{f \in \mathbb{R}_+} u_i \left(\alpha + \beta f, f + \sum_{j \neq i} f_j \right) - c_i \left(\alpha + \beta f + \sum_{j \neq i} r_j, f \right). \tag{6.17}$$

This program is well-specified so that a maximum certainly exists. But it is very hard to actually compute its solution even numerically. Notice that in the case of our public good model we have actually solved a similar program for computing the optimal uniform rule (the fixed contribution scheme), which has very interestingly been proven [Courcoubetis and Weber (2005)] that asymptotically achieves the second-best efficiency. Unfortunately, in this case, one should compromise with simple heuristics for the calculation of a uniform α, such as the computation of an average α based on the distributions of the coefficients of the utility and cost functions of the peers as we did for comparison purposes in Section 4.5.1

[2] In the final case, in which $v_1 > v_2$, $k_1 > k_2$ and $v_1/k_1 > v_2/k_2$, a quasi-concave function can be fitted for $\rho(\cdot)$, which would ensure that the second-order condition is satisfied.

for the case of our public good model.

6.4.4 A numerical example

We present below a specific numerical example to show how the efficiency achieved by such heuristics is affected by the heterogeneity of peers. Suppose that there are n peers; and that the net benefit functions of the peers are of the form

$$b_i(\mathbf{f}, \mathbf{r}) = v_i(\sqrt{r_i} + \sum_j f_j) - k_i(\sum_j r_j + f_i^2)$$

where v_i and k_i are strictly positive constants. Equilibrium is given by

$$\hat{f}_i = \frac{v_i}{2k_i}, \qquad \hat{r}_i = \frac{v_i^2}{4k_i^2}$$

while the efficient solution is given by

$$f_i^* = \frac{\sum_j v_j}{2k_i}, \qquad r_i^* = \frac{v_i^2}{4\left(\sum_j k_j\right)^2}$$

for $i = 1, \ldots, n$. The fully efficient, non-uniform prices and the approximately efficient, uniform prices from Section 6.3.2 are

$$p_i^r = \sum_{j \neq i} k_j, \qquad p_i^f = -\sum_{j \neq i} v_j;$$
$$p^r = \sum_j k_j, \qquad p^f = -\sum_j v_j.$$

Equations (6.12) and (6.13) give

$$\beta_i = \frac{\sum_{j \neq i} v_j}{\sum_{j \neq i} k_j}, \qquad \alpha_i = \frac{1}{2\sum_j k_j}\left(\frac{v_i^2}{2\sum_j k_j} + \frac{\left(\sum_j v_j\right)^2}{2k_i}\right);$$

the uniform slope coefficient β identified in proposition 4 is

$$\beta = \frac{\sum_j v_j}{\sum_j k_j}.$$

So, a simple heuristic for setting a uniform linear rule is to use the above value for the coefficient β and compute an average value of the optimal a_is based on the distribution from which the values of $\{v_i, k_i\}$ are drawn. We present in the following the results we obtained following the same procedure described in detail in Section 4.5.2 for computing numerically, based on the simple example of this section, the efficiency loss from using average approximate prices and rules as a function of the degree of peer heterogeneity. We assumed that the parameters $\{v_i, k_i\}$ are drawn from the lognormal distributions $V(0.5, \sigma)$ and $K(1, \sigma)$ respectively and computed their efficiency loss compared to the complete information optimal for different values of their common variance, σ, which expresses different degrees of peer heterogeneity. As shown in Figure 6.3, and was to be expected, the performance of the uniform rule decreases significantly as the degree of heterogeneity grows, while the performance of the average price stays very close to the optimal.

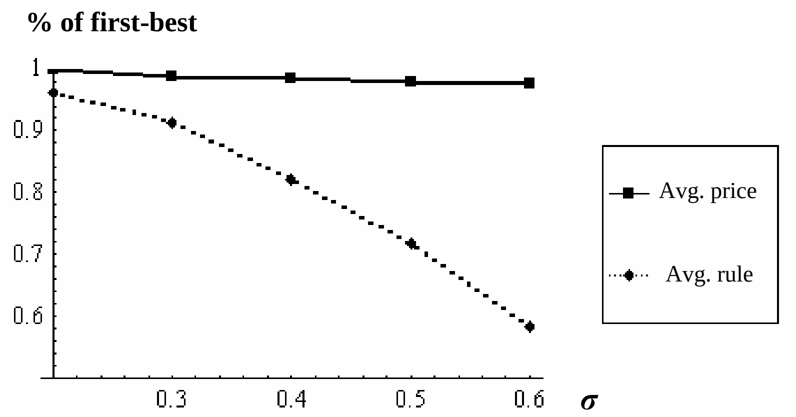


Figure 6.3: Efficiency loss of the uniform linear rule as a function of the common variance of the lognormal distributions V and K, σ, for $n = 100$.

6.4.5 Stability of rules

In this section we discuss the stability properties of our simple model when linear rules are used to control peer behaviour. The reason stability is an important issue is because peers choose how to update their level of consumption and contribution locally, which in turn affects the decisions of other peers. We will prove the stability of the system[3] when each peer updates her f_i assuming a rule of the form $r_i = \alpha_i + \beta_i f_i$. Then, her net benefit

[3] We do this using the more accurate net benefit function for peer i $u_i(r_i, \sum_{j \neq i} f_j) - c_i(f_i, \sum_{j \neq i} r_j)$.

becomess

$$g_i = u_i(\alpha_i + \beta_i f_i, \sum_{j\neq i} f_j) - c_i(f_i, \sum_{j\neq i}[\alpha_j + \beta_j f_j])\,,$$

and he should be increasing f_i according to the derivative

$$\phi_i = \frac{\partial g_i}{\partial f_i} = \frac{\partial u_i}{\partial r_i}\beta_i - \frac{\partial c_i}{\partial f_i}\,.$$

We assume that as a function of time, $\dot{f}_i = \phi_i$, and choose as a potential Liapunov function $V = \sum_i \phi_i^2$. Now[4],

$$\dot{\phi}_i = \frac{\partial^2 u_i}{\partial r^2}\beta_i^2\phi_i + \sum_{j\neq i}\frac{\partial^2 u_i}{\partial r\partial f}\beta_i\phi_j - \frac{\partial^2 c_i}{\partial f^2}\phi_i - \sum_{j\neq i}\frac{\partial^2 c_i}{\partial r\partial f}\beta_j\phi_j\,.$$

Then

$$\dot{V} = \sum_i 2\phi_i\dot{\phi}_i = 2\sum_i[\frac{\partial^2 u_i}{\partial r^2}\beta_i^2 - \frac{\partial^2 c_i}{\partial f^2}]\phi_i^2 + \sum_i\sum_{j\neq i}[\frac{\partial^2 u_i}{\partial r\partial f}\beta_i - \frac{\partial^2 c_i}{\partial r\partial f}\beta_j]\phi_i\phi_j\,.$$

Let

$$A_{ii} = \frac{\partial^2 u_i}{\partial r^2}\beta_i^2 - \frac{\partial^2 c_i}{\partial f^2},\ A_{ij} = \frac{\partial^2 u_i}{\partial r\partial f}\beta_i - \frac{\partial^2 c_i}{\partial r\partial f}\beta_j\,.$$

Then

$$\begin{aligned}
\frac{1}{2}\dot{V} &= \sum_i A_{ii}\phi_i^2 + \sum_i\sum_{j\neq i} A_{ij}\phi_i\phi_j \\
&\leq -\min|A_{ii}|\sum_i \phi_i^2 + \max|A_{ij}|\sum_i\sum_{j\neq i}\phi_i\phi_j \\
&\leq -\min|A_{ii}|\sum_i \phi_i^2 + \max|A_{ij}|(n-1)\sum_i \phi_i^2\,.
\end{aligned}$$

Hence, a sufficient condition for stability is

$$\min|A_{ii}| > (n-1)\max|A_{ij}|\,.$$

The above condition is satisfied if the effects of f and r to the benefit and cost function have a small correlation, while $u(\cdot)$ and $c(\cdot)$ are strictly concave and convex respectively (if

[4]In what follows, we denote by r and f the first and second (second and first) argument of the functions u_i (c_i) respectively.

these effects are independent, then $|A_{ij}| = 0$).

6.4.6 A model with quality of service

In the p2p file sharing model presented above, the utility function u_i of peer i depended on her request rate r_i and the sum of available resources $\sum_j f_j$, assuming a constant quality of service offered by all peers, independent from factors such as the concurrent requests made to a specific peer, her technical capabilities or her serving policy.

In this section we formulate a simple extention of this model where peers choose the quality level of the service they provide to the other peers. Let $u_i(\{Q_j\}, \{r_{ij}\})$ be the rate of benefit obtained by peer i when the rate of service requests she routes to peer j is r_{ij} ($r_{i\cdot} = (r_{i1}, \ldots, r_{in})$), and these are served by peer j with quality Q_j ($Q = (Q_1, \ldots, Q_n)$). One could imagine quality of service Q_j as the blocking probability experienced by clients requesting service from peer j.

Personalized service requests model the case of a file sharing system where specific peers store particular content and thus according to the tastes of an individual peer not all other peers have equally desirable content or the case of a peering wireless LANs system where customers of a particular peer network have a certain mobility pattern and thus have different demand for access services from different networks. The model presented in this section was actually originally proposed for the peering WLAN application [Antoniadis et al. (2003b)] (see also Section 3.6.2 for a short description) but it is also applicable to file sharing assuming that not all content is equally useful to a certain peer (as in the model of Buragohain et al. (2003)).

Hence, in our model, peer i controls the rate of service requests she makes to her peers[5]. In most cases, there may be substitution between elements of the vector $\{r_{ij}\}$, captured in the definition of the function u_i. This models the case of the same content being available by more than one peers, and the case of access services of different peers spanning the same geographic location. One could also assume that this rate can not exceed some maximum rate $\bar{r}_{ij}$ that corresponds for example to the rate of roaming customers of peer i which physically enter the area covered by peer j.

The rate of cost incurred by peer i is denoted by $c_i(Q_i, \sum_k r_{ki})$, where Q_i is the quality level maintained for serving the requests of her peers, when these arrive with a total rate

[5]In the case of WLAN peering this would be accomplished by not allowing all her roaming customers to request service at a remote location.

$\sum_k r_{ki}$ (to simplify notation, we assume as in our previous models that the sum is for all k, i included). This cost is a function of the resources allocated by peer i for serving the other peers with quality Q_i. In many cases this amount of resources is only known privately to peer i. This motivates us to write it as a function of the observed (by third parties) quantities r_{ki} and Q_i. Then the net benefit of peer i is $u_i(\{Q_j\}, \{r_{ij}\}) - c_i(Q_i, \sum_k r_{ki})$[6], and the social welfare is

$$SW = \sum_i [u_i(\{Q_j\}, \{r_{ij}\}) - c_i(Q_i, \sum_k r_{ki})]\,. \tag{6.18}$$

Our first goal is to seek prices under which the maximum is achieved in (6.18). Taking derivatives of SW with respect to Q_i and r_{ki} and setting these equal to zero we obtain

$$\sum_{j \neq i} \frac{\partial u_j}{\partial Q_i} - \frac{\partial c_i}{\partial Q_i} = 0\,, \tag{6.19}$$

and

$$\frac{\partial u_j}{\partial r_{ji}} - \frac{\partial c_i}{\partial r} = 0\,. \tag{6.20}$$

The above equations suggest that in order to make the socially optimal point an equilibrium we should use lindahl prices, for quality and request rates (different for each peer i)

$$p_i^q = \sum_{j \neq i} \frac{\partial u_j}{\partial Q_i}, \quad p_i^r = \frac{\partial c_i}{\partial r}\,, \tag{6.21}$$

where the derivatives are computed at the optimum of (6.18). Using these prices, peer i offering quality level Q_i and having request rates $r_{i\cdot}$ is rewarded a negative charge (receives) $p_i^q Q_i$, and incurs a positive charge (pays) $\sum_{j \neq i} r_{ij} p_j^r$.

Such prices motivate the use of a rule for peer i of the form

$$Q_i \geq \sum_{j \neq i} \beta_{ij} r_{ij} + \alpha_i\,, \tag{6.22}$$

where the vector of weights$\{\beta_{ij}, \alpha_i\}$ is defined for each peer i using the optimal prices as we did in Section 6.4.1 above for our original model.

[6]As in our previous models, $u_i(\cdot, \cdot) \geq 0$ is assumed to be increasing and strictly concave (expressing saturation effects), and $c_i(\cdot, \cdot) \geq 0$ to be increasing and strictly convex, in both their arguments and for all i. Since peers gain no benefit when success probability is zero, let $u_i(0, \cdot) = 0$.

More specifically it can be easily proven that for

$$\alpha_{ij} = \frac{p_j^r}{p_i^q}, \beta_i = Q_i^* - \frac{\sum_{j\neq i} p_j^r r_{ij}^*}{p_i^q} \tag{6.23}$$

the above rules make the optimum of (6.18) to be an equilibrium if peers are left free to choose their levels of quality and request rates. We can also prove, as in Section 6.4.5, that under suitable second order conditions involving the utility and cost functions, this equilibrium is stable.

In such a model the rule involves the quantities Q_i and r_{ij}. Both quantities may not be easy to measure accurately since these involve parties with conflicting interests. For instance, in the case of Q_i representing the probability a service request to be accepted by peer i, how would one measure such a quantity? Asking peers about their perceived Q_i offered by peer i may not lead to truthful answers since these may have the incentives to downgrade peer i in order to force him to raise his performance. One may consider the possibility of using reputation for motivating peers to answer truthfully. For instance, after collecting answers from all peers, the peers with answers being outliers in the above statistical sample may be punished and have their reputation lowered. To offer the right incentives we could make (6.22) depend on the reputation of the peer. For instance, we could use

$$q_i Q_i \geq \sum_{j\neq i} \beta_{ij} r_{ij} + \alpha_i\,, \tag{6.24}$$

where q_i is the reputation of peer i.

In this simple reputation model we assume that peer i offers the same quality Q_i to all peers, i.e., she can not discriminate against individual peers (for instance, by using cryptography to hide the particular identity of the roaming customers in the peering WLANs application). But the model can easily be extended to allow for such a service differentiation, where peers can offer different blocking probability to different peers. In any case, as already discussed in Section 2.6, providing reliable information in fully distributed p2p systems through reputation mechanisms is an open and very challenging research question and it is out of the scope of this dissertation.

6.5 Summary and discussion

In this chapter, we formulated a more detailed economic model capturing uploading costs in addition to file sharing costs. We then introduced a general class of reciprocity rules for enforcing the efficient amount of consumption and contribution of resources by peers according to this model. The notion of reciprocity is very fundamental in p2p systems since it is generally considered as a fair and realistic alternative to prices for incentivizing peers to contribute their resources to the system. Notably, the simplest such reciprocity rule is the one that equates consumption and contribution in terms of actual uploads and downloads[7], and a large percentage of the research on p2p economics addresses issues that are related to its enforcement in an untrusted p2p environment. But in this part of our work we explore the design of reciprocity rules enforcing the relation between consumption and contribution of resources (not necessarily 1-1) that would maximize the economic efficiency of the system under our generalized model. This means that the optimal relation between consumption and contribution of each peer would in general depend on the utility and cost functions of all participating peers.

We showed that a linear such rule can maximize efficiency but for the computation of its coefficients we require complete information and they would need to be in general personalized. So, ideally, a different relation between consumption and contribution should be enforced for each peer, which is impractical even in the case of centralized systems. Hence, one should compromise as far as economic efficiency is concerned if a uniform such rule is to be employed. Notably, the optimal uniform rule under incomplete information in this context is unknown and only simple heuristics could be considered, which although they could be employed to increase a system's efficiency, they have limited theoretical interest.

However, a rule dictating an arbitrary relation between consumption and contribution of resources, even if it was uniform, it would be very difficult to implement in a realistic p2p system, where accounting and enforcement face serious attacks, especially when peers can easily change identities. To see this, notice that the most intuitive mechanism for equating consumption with contribution is the use of virtual currency (tokens), which again requires

[7]Notice that in our model, contribution is measured in terms of content availability —number of files shared per unity of time— rather than actual uploads performed. But if we assume that files have equal probability to be requested (see also Section 4.6.2) then it is easy to translate files shared per unity of time to actual uploads per unity of time.

trusted entities and strong (or not cheap) identities to be successfully implemented. Otherwise, peers could always acquire the necessary amount of credit without contributing anything by creating a new identity each time their initial endowment is consumed. Additionally, it is to see that any other token exchange ratio than 1-1 would lead to inflation or deflation and the virtual economy would collapse. So, techniques based on tokens or in general virtual currencies, even if successfully implemented, are inadequate for enforcing different than 1-1 relationships between uploads and downloads. To achieve this there should be some way to independently enforce the efficient amount of consumption and contribution for each peer which is an even more challenging accounting and enforcement task.

But note that even though the equilibrium achieved using uniform reciprocity rules would be probably inefficient under the assumption that uploading imposes significant costs on peers, such rules constitute a powerful tool for incentivizing contribution. Actually, in the main body of our work, we have considered uploading costs as having a second-order effect in the contribution of a peer in a p2p file sharing system (and most economic models in the literature do not take account such costs as well) and showed that a memory-less reciprocity rule who seeks to enforce the efficient amount of contribution during the time peers are consuming resources, can achieve levels of content availability which could be comparable to the optimal ones under certain assumptions. Most importantly, in contrast with reciprocity rules equating consumption and contribution in terms of actual uploads/downloads, this contribution is to be computed based on specific system parameters such as the size of the system and the distribution of peer types and thus some sort of regulation is employed in order for the overall efficiency to be maximized (see Chapter 5).

Thus, in any case, reciprocity rules that relate somehow consumption with contribution would force peers to contribute a certain amount of resources in order to satisfy their own demand and thus achieve a certain level of content availability is achieved. Another alternative approach for incentivizing contribution if simple system rules are to be employed, is to provide increased QoS to peers that contribute more; modelled for example as a higher probability that a providing peer will accept their request (as in the work of Buragohain et al. (2003) and Feldman et al. (2004)) or as having access to a larger percentage of the overall amount of files made available (as in our analysis for group self-selection in Section 4.6.1).

Chapter 7

A Market model

Until now we have explored ways to incentivize the provision of content and peer availability in a p2p file sharing system in order to achieve efficient levels of content availability. In this chapter we wish to demontrate the negative effect of relatively cheap distribution of content amongst the members of a p2p community on the incentives to purchase and introduce costly items into the community. We formulate a suitable market model for content distribution and study the underlying trade-offs in terms of the availability of information and show that in many cases specific privileges should be granted to peers in order to be beneficial for them to purchase and share expensive content items. See also [Antoniadis and Courcoubetis (2002)].

7.1 Introduction

In contrast with centralized or hierarchical content distribution paradigms such as multicasting, caching and content distribution networks (CDNs), in the p2p model the content is delivered in a fully distributed fashion. Each peer after receiving the content can act as a content provider making the content distribution much more efficient. As initially conceived, p2p networks assumed insignificant costs in terms of obtaining and transporting content and deployed minimum control mechanisms. In such a simple model, an agent that needs the content will always obtain it, and will make it freely available to others.

But content may be costly to obtain from an external source (even if it could be then cheaply distributed among the members of the p2p community), since the bandwidth of the network may be restricted at places causing delay costs over particular distribution routes or content is costly itself to obtain. As an example imagine a peer having to download a large file from a remote location (suffering the large delay) and then this content become easily accessible to peers residing in the same LAN.

If such costs become significant and peers have no mechanisms to recover their costs, content may fail to be introduced into the p2p community. Deploying a market mechanism within the peer community may solve these problems and reduce unnecessary waste of resources resulting from peers requesting content for which their value is less than the cost imposed to others. In a market, peers are modeled as economic agents having incentives for obtaining revenue from distributing content. We may assume that there is a common cost W to bring the content to a p2p group, resembling to the case of a copyright cost or to the case where there is a substantial cost for transporting the content to any peer from its external source. We also assume that there is some intra-group cost for transporting the content, different in general for each pair of communicating peers and for each service direction. This can model delays or link performance degradation when bandwidth is a scarce resource, as in the case of network access. From the time that an agent has downloaded a piece of content from its source, all the other agents in its group will have the ability to acquire the specific content suffering only the intra-group cost and avoiding the possibly high initial cost.

Each time an agent requests a specific content we assume there exists a lookup service that provides cost and price information and can control the number of competitive offers allowed for each request. We compare two types of content distribution paradigms:

1. A *restricted competition* model, where at most d agents are randomly selected from the set S of all candidate content providers. This is an oligopoly model, where prices will depend on the number d of competing peers.

2. A *privileged provider* model, where the first agent to acquire the content[1] and hence to incur the initial cost W is granted the privilege to be the unique seller of this content within the group. The above agent has the power of a monopolist that can do personalized pricing. Granting such exclusivity rights corresponds to the caching content distribution paradigm where a specialised node (the cache) is the only one known —within an administrative domain— to be able to provide 'cheap' (relative to the initial content cost) copies of the content.

[1]Since this agent has different characteristics from the rest of the group we will assign him index 0.

We would like to investigate the effects of such information restrictions on the proliferation of the content given the fact that the economic agents maximize their net benefit and there is some initial cost for obtaining the content. When this cost is high, it may prohibit the purchase of the content from its external source.

In the following we formulate and analyze a specific economic model in order to quantify the influence of restricting content availability information to the spreading of the content in the peer community. Granting different market power to agents selling the content affects their expected profits, and hence influences their decisions about buying the content and further reselling it. Allowing a highly competitive market has the positive effect of bringing prices low and resulting in higher social welfare since more agents will finally buy the content. On the other hand, it is not obvious that excessive competition will always benefit the system as a whole since low prices make it difficult for the agent that acquires the content to recover her cost, and hence this could reduce her motivation to do so. It is thus reasonable to assume that the end result may depend on the size of the system and on the costs of initially acquiring the content and further distributing it within the group compared to the value of the content to the agents.

We propose a simple economic model that takes into account all the above parameters and results in tractable analytic solutions of the underlying games. It can be used to analyze the sensitivity of various performance metrics such as the social welfare and the expected net benefit of the agents with respect to parameters such as the initial cost, the value and popularity of the content, the available information, the transport costs, etc. These parameters affect the prices and the expected net benefit of the agents which in turn influence their decisions and the resulting content distribution.

7.2 Related work

There is a significant amount of research work on p2p content distribution. There are two main categories of research questions that are addressed in the literature. The first considers the creation of application-level overlay networks (e.g. [Banerjee et al. (2002); Castro et al. (2002)] or even full meshes (e.g. [Y.Chu et al. (2000); Cohen (2003)]) for the distributed content distribution of a single content item (either with real-time characteristics, like in p2p streaming applications, or not, like in BitTorrent) via multicast paths formed and/or bandwidth sharing amongst the receivers. The second studies the efficient replication of content amongst the members of a group [Leff et al. (1993)] in order to take advantage of locality of requests and thus minimize the overall content distribution costs[2]. In both cases, one has to address the often conflicting goals of reduced management complexity and communication overhead, scalability, stability, and efficient resource usage.

Regarding the distributed content distribution of a single content item, there are in general two main challenging research issues. First, one should decide on the overlay network to be formed between the participating nodes and according to which the content is to be distributed. Numerous solutions have been proposed in the context p2p streaming, others based on a 2-tier hierarchical approach [Jannotti et al. (2000); Chawathe (2000)], others building application-level multicast trees with a single root [Banerjee et al. (2002); Chu et al. (2003)], others building full meshes [Y.Chu et al. (2000); Francis (2001)], and others multiple multicast trees [Castro et al. (2003); Padmanabhan et al. (2003)], each one addressing more efficiently different aspects of the problem. The second research issue is how the source will distribute the different fragments of the file to the different nodes of the network, which is more relevant in the case of non real-time content, where the whole file to be distributed is initially available. This is also known as the 'simultaneous send/receive broadcast problem' [Bar-Noy et al. (2000)], which is adapted and analyzed in depth for the case of BitTorrent-like applications by Mundinger and Weber (2004). Interestingly, Bullet [Kostic et al. (2003)] combines a standard single-tree structure with a mesh made of random connections (orthogonal to the tree) which are used to exchange the bulk of the data among peers which are far away in the tree hierarchy exploiting as much as possible the available bandwidth offered by the mesh architecture.

Two different placement problems can be defined in the context of a distributed replication group. The first is the proxy (or cache) placement problem: the selection of appropriate physical network locations (e.g. routers) for installing content proxies [Krishnan et al. (2000); Li et al. (1999); Qiu et al. (2001); Cronin et al. (2002)]. The second, is the object placement problem: the placement of objects to the different nodes in the network, under given nodes' physical locations and capacities [Leff et al. (1993); Korupolu et al.

[2]If a content item is locally available, its request incurs a minimal distribution cost. Otherwise, if it is made available by some other node in the group it will be fetched at a potentially higher, but small, cost. If the content item is not available within the group, it is retrieved from its origin server, with maximum distribution cost. Note that in this context the abstraction of a replication group [Leff et al. (1993)] differs from distributed caching in that caching involves the definition of a content replacement algorithm, while replication assumes a fixed placement of content (at least until a new invocation of the placement algorithm takes place).

(1999); Kalpakis et al. (2001); Kangasharju et al. (2002); Baev and Rajaraman (2001); Loukopoulos and Ahmad (2004)]. See also [Laoutaris et al. (2004); Laoutaris et al. (2005)] for joint formulations of the aforementioned problems. These approaches propose specific object and proxy placement algorithms that optimize the system's performance in terms of the overall delay and bandwidth usage due to replication. The corresponding algorithms are executed by a central authority which has the power to enforce the corresponding replication decisions and they are thus suitable for applications such as web mirroring and CDNs.

When no central control can be assumed, another important problem should be addressed: the provision of the suitable incentives to self-interested agents for ensuring socially optimal object replication decisions in replication groups (see the current work of Laoutaris et al. (2005)). Obviously, this is also an issue in the case of content distribution of a single item. BitTorrent [Cohen (2003)] is a realistic application that addresses the problem of free riding in the case of non-real-time content, through 'direct exchange' of upload capacity, as described in Section 2.6.4. In the context of p2p streaming, Chu et al. (2004) propose the distribution of 'wealth' (upload capacity) amongst peers using taxation assuming that the capabilities of each peer are observable by the source.

We should stress that our approach is not concerned with the efficient content distribution inside a p2p group (neither economic nor in terms of performance) but studies the provision of the suitable incentives for introducing new items into a community from an external source, assuming content distribution costs between the members of the community are predefined but much smaller than the cost for acquiring the content from its external source. There is not, to our knowledge, other research work investigating this specific game that may arise in p2p content distribution applications.

7.3 A simple p2p content distribution market model

Lets assume a group of N peer agents which may request some content, each agent i obtaining value u_i by using the content, where u_i is uniformly distributed in $[0, V]$. We assume that initially the content belongs to an external provider who charges a substantial fee W, high enough relative to V, to any agent that will purchase the content. So, once some agent purchased the content, it makes more sense for him to resale the content at lower prices within the group in order to recover his cost and even make some profit, instead of more agents paying this high fee to the content provider. This potential profit allows for agents with $u_i - W < 0$ to take the risk and purchase the content. The larger the size of the group, the greater the value $\max_i\{u_i - W_i\}$ will be, hence increasing the probability for such a decision. Clearly, the desicion whether to pay this fee W depends on the estimate of the average revenue that can be obtained by reseling it, and on the degree of agents' risk aversion.

On this level there is an interesting game to be played. Each agent waits for another agent to do this first move and pay the fee W. This 'free rider' problem may in some cases result in low probability of purchasing the content. We discuss next the part of the model that describes the internal market of the p2p group, which is used to compute the average revenue obtained by content resale.

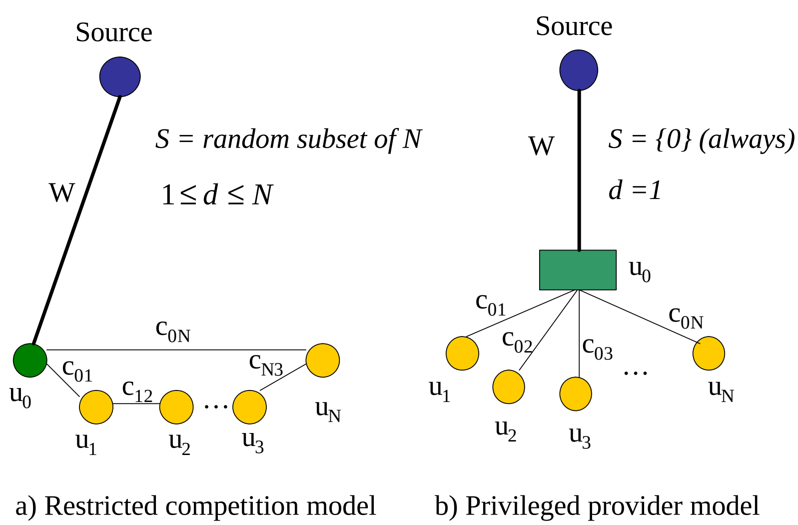


Figure 7.1: Content distribution based on information

An important aspect of our model is the cost associated with moving content around. We denote by c_{ij} the cost for moving content from agent i to agent j. This may be direct communication cost such as delay, information loss or payments to transport service providers, or indirect cost corresponding to performance deterioration of the access service of the agent giving away the content. Such a cost will be charged to the agent purchasing the content in addition to the payment for the content itself. Depending on the situation, c_{ij} may be known a priori only to j, or only to i, or to both. For instance, agent j may get back from the network the identity of agent i and the average throughput of a connection to this agent, while the agent i may only know that there is a potential customer. This uncertainty and asymmetry about the costs is crucial for deriving positive revenues. For instance, if agent j requests the content and there are two competing providers i and k which already have the content, where $c_{ij} = c_{kj} = 0$, then the equilibrium price is zero (see [Varian (1992)])! On the other hand, if costs are known to all parties involved and $c_{ij} > c_{kj}$, then k will win posting the price $c_{ij} - c_{kj}$. If the distribution of these costs is random with a known distribution, then one may calculate the average revenue from such a transaction. Our model allows various types of information asymmetry: costs may be random and known only by the customers, or each seller knows only his cost.

Depending on the assumptions, there are different games to be solved. A seller makes a bid (the price for the content) trying to maximize his expected net benefit (expected revenue). His strategy takes into account the information about the distribution of the value of the content to the buyer, the distribution of the transport cost of his competitors, and the number of such competitors. We will discuss some of the simplest games that one can solve which capture most of the above aspects. It is clear that the expected revenue greatly depends on the degree of competition. Building mechanisms allowing full information and hence fierce competition will result in low prices and revenues, increasing the risk for paying the high fee W for bringing the content within the group.

There are other important parameters such as content popularity, which can be modeled by the probability of an agent to request the content or by the number of agents that will eventually request the content (the equivalent size of the group with respect to the above content). Reputation may be modelled by associating different values for the same content obtained from different sellers. For instance, the value of agent i may be u_{ij} where j is the seller of the content.

We have already mentioned that when an agent requests the content, the system provides him with information regarding the potential sellers and possibly the cost of the transport. The amount of information disclosed by the network to buyers is a crucial issue. The number of sellers that are allowed to participate in such a content auction is also a crucial parameter that defines prices of content resale. From our discussion above, controlling the degree of competition in the internal p2p market may be critical for achieving economic efficiency. In our case, we measure efficiency by the sum of the value generated from all agents that will pruchase the content. We are interested in investigating the effect of such 'information policies' to the propagation of the content. There may be regimes (defined for some range of parameters) where restricting information and granting exclusivity for content resale may be beneficial. One suspects that this may be the case when content is not very popular but has high value for the agents that need it. Similarly, if content is more popular but $W >> V$, excessive competition may again prohibit efficient content distribution. On the other hand, if content is very popular and W reasonably close to V, there may be a significant chance that at least one agent having such a high utility will purchase the content, even if the revenue from resale may be small. Upper bounds on revenue may be obtained by taking $N \to \infty$. As the number of peers increases, the chance of a peer winning the auction goes to zero, and if competition is unrestricted prices will also drop to zero.

7.3.1 Calculation of social welfare and expected revenue

We discuss next the market models and the resulting prices and revenues corresponding to different levels of competition. As we show, the degree of competition influences the agents' decisions and the social welfare of the system. Our goal is to convince the reader that our model leads in many interesting cases to analytically tractable solutions, which we hope that can provide better insight to the sensitivity of the performance with respect to the key parameters introduced.

The case of a privileged provider

This is the extreme case of a monopoly, where agent 0 (the one who decides to purchase the content from its external source) is granted the exclusivity right to resale the content after purchasing it for the initial price W. This agent, given the available information, must choose the optimal price w to charge for the content in order to maximize his expected revenue $\bar{R}_m$ obtained by selling the content to the peer group. Then he must compare this expected revenue to $W - u_0$ and decide if it is worth to him to undertake this effort.

This simple decision rule makes sense when the agent is risk neutral. In general such a decision will depend on the fluctuations of the random variable R_m. In a more sophisticated setup, this decision may be probabilistic and correspond to the optimal strategy of the agent in the corresponding game where all agents are facing the same decisions. Although the agent may have a positive net benefit by choosing to be the exclusive content provider, he may choose to wait for an other agent to take this risk. The reason is that then he may be able to purchase the content at a much lower price than W, possibly leading to a larger net benefit. This is an instance of the 'free riding' problem: there is a probability that no agent decides to take the risk and hence all agents loose by obtaining zero value.

We briefly describe how one may compute the expected revenue from resale. Let c_i be the cost to transport the content to agent i from agent 0. We assume for simplicity that the c_{ij}s are iid[3] with uniform distribution in $[0, C]$. We also assume that the value of c_i is made known only to agent i that requests the content[4] whereas only its distribution is known to agent 0. Similarly, the value u_i of the content to agent i is iid and uniform in $[0, V]$. Let w the price that agent 0 uses to respond to the request of agent i. If $u_i < w + c_i$, then agent i will refuse to buy the content resulting in zero revenue for agent 0. The optimization problem faced by agent 0 becomes

$$\max_w R(w) = wP[u > c + w]\,, \tag{7.1}$$

where for simplicity we have omitted the subscript i. By solving (7.1) we obtain

$$0 \le C \le 2V/3 \quad : \quad w^* = V/2 - C/4,\ R^* = \frac{[V - C/2]^2}{4V}\,, \tag{7.2}$$

$$2V/3 \le C \le V \quad : \quad w^* = V/3,\ R^* = 2V^2/27C\,. \tag{7.3}$$

The total average revenue $\bar{R}_m(N)$ is now obtained by multiplying R^* by N.

To compute the social welfare we must compute the average value of a peer that accepts such an offer, and then multiply it by the average number of such peers. In order to do this, one must compute the conditional distribution of the value of a peer given the fact that he accepted price w^*. Such quantities can be computed analytically. For instance, if $C \le 2V/3$, the conditional expectation of u_i is equal to $\frac{7C^2+12CV-36V^2}{24(C-2V)}$. The average number of peers that will eventually accept the offer is $NP[u > c + w^*]$. This allows us to compute the social welfare, given that the content is purchased,

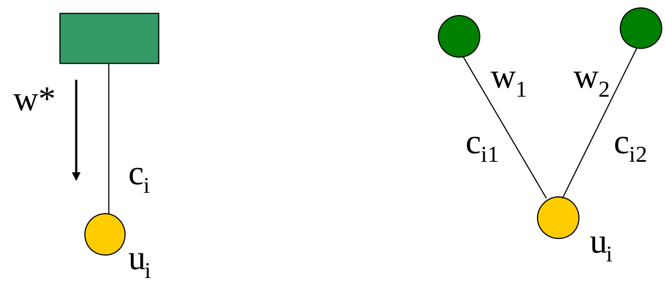
$$SW_m = NP[u > c + w^*]\frac{7C^2 + 12CV - 36V^2}{24(C - 2V)}\,, \tag{7.4}$$

for $C \le 2V/3$, and similarly for $V \ge C \ge 2V/3$. Multiplying with the probability that at

[3]indepent identically distributed

[4]One may make different assumptions about the disclosure of such information.

least one agent will decide to initially buy the content, $P_b = [1-(1-P[u > W+\bar{R}_m(N)])^N]$, we can compute the expected social welfare $S\bar{W}_m = P_b SW_m$.



(a) Privileged: Agent i buys the content if $u_i > c_i + w^*$.

(b) Restricted (for $d = 2$): Agent i selects $\min_j(c_{ij} + w_j)$ such that $u_i > c_{ij} + w_j$, $j = 1, 2$.

Figure 7.2: Privileged vs. restricted competition markets

The case of restricted competition

In this case we allow d peers to compete for every content request. These are chosen randomly among the peers that have already purchased the content. We consider the illustrative case where $d = 2$. To simplify notation consider the case where peers 1 and 2 compete for a given request by some third peer. Let c_i and w_i, $i = 1, 2$, be the costs and prices corresponding to each of these peers.

Prices w_i are the solutions of the following game. Consider peer 1. Assuming that peer 2 has chosen price w_2, he chooses his price $w_1(w_2)$ in order to maximize his expected revenue $R_1(w_1) = w_1 P[w_1 + c_1 < w_2 + c_2 \wedge w_1 + c_1 < u]$. For simplicity assume that $V >> C$ in which case we can disregard the second event in the calculation of the above probability, and hence compute an approximation of (higher than) the optimal price. Simple calculations show that $w_1^* = \frac{C+w_2}{3}$ and since in equilibrium both agents will post the same price, we will have that $w_1^* = w_2^* = \frac{C}{2}$. This means that each seller has equal probability to win depending on whether his connection incurs a lower cost compared to his adversary. In other words, each one of them will finally sell the content with probability 1/2 (e.g. for peer 2, the probability that $c_{i1} > c_{i2}$). Hence

$$w_1^* = \frac{C}{2},\ R_1(w_1^*) = \frac{C}{4}P[w_1^* + c < u] = \frac{C(V-C)}{4V}\,. \tag{7.5}$$

Note that again in this case the content will not always be sold, even if the price is lower than in the case of (7.1). The probability for selling the content if C is small compared to V, say $C \leq 2V/3$, is $P_u = (V-C)/V$.

Lets evaluate now the revenue of agent 0. The first time he will resell the content he will use (7.1) since he is the unique agent that can sell the content. After that he will always use price w_1^*. How much revenue will he make in total? Assuming that agents are randomly chosen among the set S of candidate providers, the probability that agent 0 is chosen in the competing pair is $2/|S|$.

Conditioned on the event that M agents will eventually buy the content, where M has a binomial distribution with parameter P_u, the expected revenue of agent 0 in this duopoly case is

$$\bar{R}_d(M) = \frac{[V-C/2]^2}{4V} + \frac{C}{2}\frac{1}{2}\sum_{k=2}^{M}\frac{2}{k} \approx \frac{[V-C/2]^2}{4V} + \frac{C}{2}[\gamma + \ln M - 1]\,, \tag{7.6}$$

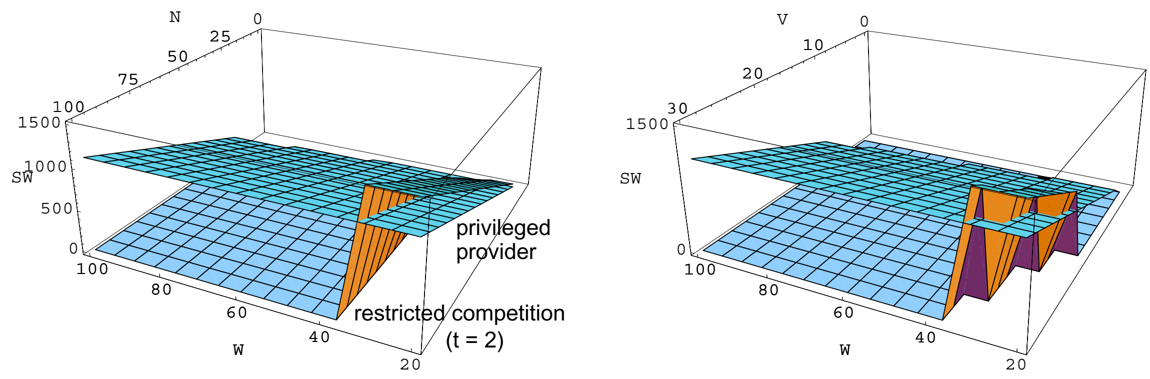
where γ the Euler-Mascheroni constant and $M = NP[w_1^* + c < u]$. From (7.6) we can compute an upper bound of the total expected revenue of agent 0 in the case of restricted competition:

$$\bar{R}_d(N) = \frac{[V-C/2]^2}{4V} + \frac{C}{2}[\gamma + \ln[NP_u] - 1]. \tag{7.7}$$

Similarly to the case of the privileged provider we can compute the total value generated in this system. For instance, in the case where $C \leq 2V/3$, $SW_d = N\frac{13C^2 - 12V^2}{24V}$ and thus

$$S\bar{W}_d = [1 - (1 - \frac{V - W + \bar{R}_d(M)}{V})^N]SW_d\,. \tag{7.8}$$

In Figure 7.3 we present the results we have obtained comparing the above two cases based on the analysis above. Both subfigures depict the same, expected, phenomenon. When V is not close to W, the probability that an agent will decide to take the risk and buy the content of the restricted competition case is almost zero and so is the expected revenue. Thus, in this range of parameters, privileged provider is the only viable scenario and the resulting social welfare increases the more peers are participating, since the expected



(a) As a function of (W,N), for $V = 30$ (b) As a function of (W,V), for $N = 100$

Figure 7.3: Comparison of the expected social welfare between the 'privileged provider' and 'restricted competition' cases

revenue of agent 0 and the total value increases, and the probability of the initial purchase increases as well.

On the other hand, when W is close to V (in our range of parameters when it is less than 40, V being 30) this probability becomes positive in the restricted competition scenario as well, and since prices will be much lower in this case, more peers will finally purchase the content and thus more total welfare will be achieved compared to this of the privileged provider scenario, with increasing difference as N and V grows.

7.4 Discussion and future work

The market model presented in this chapter may seem inapplicable in a realistic problem of content distribution in today's Internet. Its most important ingredient that constitutes such a criticism viable is the fact that it requires the implementation of all the necessary functionality for making a distributed market feasible, such as micro-payments, accounting, etc. Moreover, in addition to the very challenging research issues involved in the implementation of such a functionality in a p2p environment, dynamic payments (or micropayments) are in general undesirable by users, especially in the Internet, who in general prefer predictable cost and quality of service in their transactions. Finally, in cases where the initial cost W has to do with buying copyrighted material, content's redistribution would be subject to legal threats, especially since the implementation of a market mechanism requires the existence of trusted central entities which would be vunerable to possible lawsuits of copyright holders.

However, we believe that the underlying game is theoretically interesting and it is possible that there will be situations in the future which could constitute such an analysis very helpful. Notice that we have to deal with a cost sharing game, similar to multicast cost sharing and our public good model (assuming the quantity of the good is fixed): peers should decide how to share the cost of an expensive item, which once bought by one of them it would be accessible to everybody with small cost. But there are also some unique characteristics of our formulation of the game.

First, the good (in our case the content item) cannot be made available concurrently to all peers (as is the case in multicasting for example). In practice, this could be due to the fact that technically only one peer can purchase the item or because membership in the community is evolving over time (or peers are anonymous). So, in our game formulation a single agent is responsible for taking the risk of initially purchasing the content hoping that she will recover her cost and thus eventually share it with the rest of the members of the community. This means that peers cannot coordinate their decision and for example run a resource provision mechanism (like the ones discussed concerning the private provision of a public good —see Section 3.3), which would decide how the cost should be shared amongst them given their declared utility for the item.

Second, in our game formulation intra-group distribution costs are assumed to be non-zero. This characteristic also motivates the sequential content purchase but also constitutes the implementation of a market for content distribution meaningful, as already explained. Of course, the assumptions about the available information concerning these costs play a very important role on the properties of the resulting game. For example, unlike our analysis above, we could assume that the intra-group distribution costs are also visible to the content providers. In this case, the content item will be always sold by the provider with the smaller cost, since he could post a price that equals the smallest cost of its competitors, which we could also assume for simplicity that it is independent from the potential buyer. Hence, agent 0 will be able to sell the content until an agent with smaller cost buys it. Again, one could perform similar computations with our analysis above to estimate the social welfare achieved under different scenarios concerning the competition allowed.

Another interesting future direction of this work is to consider multiple levels in the content distribution hierarchy, where a significant cost W_{ij} is required to move content from level i to level j. Then it would be interesting to study how deep into the hierarchy a content item will manage to travel, starting from a source at level 0 according to different assumptions on peer interarrival items, cost and utility distribution and number of peers. One could also consider a mixed game where coalitions of agents with zero distribution costs amongst them could agree to buy the content and share the revenue.

Finally, it is particularly interesting to consider other types of rewards for agents deciding to purchase a content item for the first time such as reputation and credits. This would be feasible if there is some way to remember the identity of the peer that introduced an item in a community whose reputation or credit account would increase the more this item is distributed amongst the community members. Then agents who have collected a high reputation or number of credits could receive better QoS as consumers, gain social recognition, or even get a larger share of fixed entry fees that could be required for peers to join into the community. In such a system not only there will be no need for payments to be made between peers but also content would be more efficiently distributed since there would be no need for restricting competition, provided of course that there is a way to ensure that the peer who introduced an item takes the credit from its further distribution, which is probably the most challenging issue that needs to be addressed in this context.

Chapter 8

Conclusion

8.1 Summary

The p2p paradigm, although not really new as a concept has received significant attention from the Internet and the research community over the last years mainly due to the increased capabilities of the end-hosts and the expansion of broadband access networks which enable the scalable deployment of many interesting applications. However, until now file sharing remains the killer p2p application and the only widely deployed amongst the large variety of possible applications enabling the exploitation of different types of local resources such as computing power, storage, access bandwidth, packet forwarding (e.g. when mobile devices form ad-hoc networks), or even cheap dialing capabilities (e.g. Asterisk), and more. The fundamental difference from traditional distributed systems is that in p2p systems the nodes make their decisions based in principle on their own self-interest (and not towards achieving a common goal) since their resources are under their own control, and clearly the algorithms optimizing system's performance should take this important dimension into account.

The fact that self-interested behaviour (e.g. free-riding) could be in many cases detrimental for the overall system's efficiency constitutes the provision of the suitable incentives to peers to contribute their resources and act towards improving the global efficiency, rather than their personal benefit, a very challenging and critical task. The theory of economics studies for several decades now exactly such issues but there are many unique characteristics of p2p systems that attach new interest to this problem.

However, the resources and actions involved in the participation of a peer in a p2p community form a very complicated modelling problem which depends on many technology and application-specific characteristics. So, it would be impossible for a tractable economic model to capture all the details of all the transactions and resources involved in the system's operation. Hence, an important part of our work was to consider the most important aspects of the peers' behaviour in a p2p system, understand and categorize the main concepts and related work concerning the provision of the suitable incentives, and then make the correct abstractions so as to formulate an economic model that would be both meaningful and tractable.

We focused on the p2p file sharing and identified its public good aspect (content availability) as its most important characteristic, which we believe that it is also a dominant aspect of other p2p systems as well such as grid computing and peering WLANs under certain assumptions. Three characteristics of p2p systems that played important role in our analysis were the following: the very large number of participants, the high degree of heterogeneity, and the very challenging implementation issues due to the fully distributed and untrusted environment. Notably, the latter is the one that gives new interest to this classic, and very challenging, problem in the economics literature.

The large size in general reduces the incentives to peers to contribute the efficient amount of resources for the provision of the public good. However, when exclusions are possible, as suggested by recent important asymptotic results (see Courcoubetis and Weber (2005) and references therein), a fixed contribution scheme is within $O(1/n)$ from the maximum social welfare that could be achieved under incomplete information as the number of peers in the system becomes large. We demonstrated the very attractive properties of the fixed contribution scheme for the case of our public good model for file sharing compared to other alternative mechanisms and discussed certain theoretical issues of practical interest such as the importance of heterogeneous file popularity, the conditions that are required for system stability, and possible ways to discover system parameters that are originally assumed known.

However, the performance of the fixed contribution scheme is decreased the more heterogeneous are peers in terms of their preference parameters. In this case categorizing peers into different groups could be greatly beneficial. So, we studied under which assumptions one could exploit such additional information and the incentives that need to be given to peers to agree to be distinguished and declare their group type if this is not observable.

Finally, the fully distributed and untrusted p2p environment constitutes even such a simple mechanism (as is the fixed contribution scheme) very difficult to be enforced in practice since some sort of accounting is required, which as we saw faces very challenging attacks in this context. Towards this end, we proposed a memory-less enforcement mechanism which ensures that peers contribute to the system as far as content availability is concerned while they are consuming resources for themselves by dictating a fixed (no too high) throughput for uploading files. We presented the basic system requirements for supporting the necessary functionality and formulated a suitable economic model focusing on peer availability as the main cost generator of a peer's contribution in a p2p file sharing system and assessed its performance in terms of economic efficiency, which under certain assumptions (i.e. that the request rate of a peer has a convex relation with her preference parameter) is comparable to the optimal.

Moreover, this model provided the means to compare our proposed mechanism to other alternatives incentivizing contribution by imposing constraints on the consumption of resources such as the popular in the literature system rule equating downloads and uploads performed by each peer. Interestingly, we show that the fact that in our mechanism we can tune appropriately the critical system parameter (the amount of time peers should stay connected per download) leads to increased efficiency compared to the one achieved using the downloads=uploads rule. Moreover, its inherent elasticity in terms of the ratio of the actual uploads and downloads performed per peer ensures a better system stability since when there is an initial amount of content available (e.g. from super peers) a different ratio than 1-1 should be enforced in a virtual market for file uploads for the system to be stable, which is in general difficult to implement.

We believe that the above major part of our contribution provides a sound basis for understanding the theoretical and practical issues concerning the provision of incentives for resource contribution in p2p systems with a strong public good aspect, as is p2p file sharing. The specific solutions proposed for increasing system's efficiency lead to some interesting and practical approaches for providing incentives for content availability in p2p file sharing systems. Additionally, during our first modelling efforts, we also proposed two interesting economic models focusing on different aspects of the activity in a p2p file sharing system presented in the last two chapters.

First, we proposed a more elaborate economic model capturing uploading costs in addition to file sharing costs. We then introduced a general class of reciprocity rules for enforcing a desirable relation between consumption and contribution of resources according to this model and demonstrated the difficulties that arise due to this more detailed modelling approach. More specifically, we showed that a linear such relation could maximize efficiency but for the computation of its coefficients we would require complete information and most importantly they would need to be in general personalized (different for different peers).

And second, we defined a pure market model for p2p content distribution and identified an interesting content distribution game concerning the introduction of costly items into a p2p community. We formulated a suitable market model for content distribution and studied under which circumstances specific privileges, and of which kind, should be granted to peers who choose to purchase and share expensive content items in order to be beneficial for them to do so.

8.2 Future Work

Besides the many possible extentions of our market model for content distribution which have been left for future research, and which are analyzed in detail in the previous chapter, we believe that there are two main important avenues for future work motivated by this dissertation.

The first is the further study of the theoretical and implemetation issues that arise in the context of our public good model for content availability and the proposed incentive mechanisms for improving the economic efficiency of a p2p system. An important practical, and theoretical, aspect we have not addressed in our modelling work is the highly dynamic environment of p2p systems. Hence, one could explore the design of suitable algorithms, based on measurements of the participation, activity, and mixture of peers in the system, so as to dynamically tune the basic system parameters which we propose that should be regulated (e.g. the value of a fixed uploading throughput) in order for the system to reach efficient equilibria. It would be then interesting to assess whether such elaborate mechanisms have significant effects on the overall system performance.

Moreover, the in depth study of more realistic behavioural models is required in order for one to verify some important simplifying assumptions made throughout our analysis and decide on specific system requirements that could lead to an actual implementation of a p2p file sharing system incorporating the proposed mechanisms for incentivizing peers to contribute their resources. To this respect one should also take account the social prespective of the participation of peers in a p2p system which as real-life implementations show could play a very important role. In our work we have tried to devise mechanisms that wouldn't harm significantly the community spirit inherent in most p2p applications, but much more work is needed in order to formalize such concepts from social sciences in this context and formulate suitable socio-economic models that explicitly take them into account.

The second major avenue for future work is the application of the main ideas and concepts developed in this dissertation in the context of file sharing for other types of p2p systems. We have already pointed out the connection of our public good model with WLAN peering and scientific grids and identified for the latter several application-specific issues that need to be addressed. The most important of them is to explore the extention of our simple public good model for scientific grids for the case where congestion plays an important role but not to an extend that would wash out its public good aspect. Additionally, one could also try to identify other current or future p2p applications for which a 'contribute while consuming' mechanism could be appropriate for providing the suitable incentives and devise application-specific techniques for ensuring that a peer is contributing to the system as a whole while consuming resources.

Finally, a general interesting issue concerning the design of practical incentive mechanisms for p2p systems is the high-level strategies available to peers. More specifically, in most cases services offered by p2p systems have also an outside option (e.g. buy a cd). So, one should ensure that the costs imposed by an incentive mechanism to a peer in order for her to acquire a certain service by participating in a p2p system shouldn't exceed the cost of this service through some other means.

Ad-hoc networks actually provide a very interesting example of this complicated high-level game. More specifically, nodes participating in an ad-hoc network have two additional strategies available: mobility and tranmission range tuning. The decisions of a peer concerning her position in the network and the transmission range used are considered as predefined in existing work on incentive mechanisms for packet forwarding in ad-hoc networks. However, depending on the incentive mechanism and on the commitment of the user, moving around and selecting appropriately her transmission range could be part of her strategy (both at the routing and at the application level) and should not be disregarded.

As far as transmission range is concerned, nodes have always the possibility to increase it (within the limits of technology) in order to reach as many one-hop destinations as possible and thus avoid the costs incurred by the incentive mechanism when participating in a multi-hop network. These would be the fixed costs for running the underlying protocol and the usage based ones according to the degree of reciprocity enforced (in the mechanism sketched in Section 5.5.1 this would be the extra power required for transmiting the required piggybacked packets). So, nodes have the option to either to participate in the p2p community using the smallest possible transmission range or select use a suitable (larger) transmission range to reach directly the desired destination. Note that transmitting with maximum range could create significant congestion in many settings and thus reduce the performance of all nodes. Thus the cost incurred by a node in order to transmit a single packet in both cases depends also on the strategies of all the other peers.

Thus, according to the net benefit acquired, which is a function of the full strategy space and the incentive mechanism applied, each node will decide whether she would get involved in the multi-hop scheme and how she will behave in each case (in terms of position and range used). This is a very complex game. We hope to study some interesting cases. For example solve the problem without considering mobility or simplify it by assuming that mobile users could choose between a small predefined set of positions. Such models could provide useful insights to system designers to assess the efficiency of certain incentive mechanisms and tune appropriately their parameters. For example, one could estimate the optimal number of packets that should be piggybacked under various contexts, in terms of network performance or economic efficiency.

Bibliography

Adar, E. 2005. Drawing Crowds and Bit Welfare. *ACM SIGecom Exchanges 5,* 4 (July), 31–40.

Adar, E. and Huberman, B. 2000. Free riding on gnutella. *First Monday 5,* 10 (October).

Akerlof, G. and Yellen, J. 1985. Can small deviations from rationality make significant differences to economic equilibria? *The American Economic Review 75,* 4, 708–720.

Al-Najjar, N. I. and Smorodinsky, R. 2000. Pivotal players and the characterization of influence. *Journal of Economic Theory 92(2)*, 318–342.

Alexander, R. D. 1987. The biology of moral systems. *New York: Aldine de Gruyter*.

Anagnostakis, K. G. and Greenwald, M. B. 2004. Exchange-based Incentive Mechanisms for Peer-to-Peer File Sharing. In *Proceedings of ICDCS04: 24th IEEE International Conference on Distributed Computing.*

Anderson, C. 2004. The Long Tail. Wired Magazine, http://www.wired.com/wired/archive/12.10/tail.html.

Andrade, N., Mowbray, M., Lima, A., Wagner, G., and Ripeanu, M. 2005. Influences on Cooperation in Bittorrent Communities. In *ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems.*

Andreoni, J. 1995. Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics 110(1)*, 1–21.

Antoniadis, P. and Courcoubetis, C. 2002. Market models for P2P content distribution. In *AP2PC'02, Bologna, Italy.*

Antoniadis, P., Courcoubetis, C., Efstathiou, E., Mason, R., Papaioannou, T., Polyzos, G., Siris, V., Stamoulis, G. D., Strulo, B., and Weber, R. 2003.

Specification of Market Management Models for Peer-to-Peer Services (Final Version). MMAPPS deliverable.

Antoniadis, P., Courcoubetis, C., Efstathiou, E., Polyzos, G., and Strulo, B. 2003a. The case for peer-to-peer wireless lan consortia. 12th IST Summit on Mobile and Wireless Communications, Aveiro, Portugal.

Antoniadis, P., Courcoubetis, C., Efstathiou, E. C., Polyzos, G. C., and Strulo, B. 2003b. Peer-to-Peer Wireless Consortia: Economic Modelling and Architecture. In *Proceedings of Third IEEE International Conference on Peer-to-Peer Computing (P2P 2003).*

Antoniadis, P., Courcoubetis, C., and Mason, R. 2004a. Comparing Economic Incentives in Peer-to-Peer Networks. *Special Issue on Network Economics, Computer Networks, Elsevier 45,* 1, 133–146.

Antoniadis, P., Courcoubetis, C., and Strulo, B. 2005. Incentives for Content Availability in Memory-less Peer-to-Peer File Sharing Systems. *ACM SIGecom Exchanges 5,* 4 (July), 11–20.

Antoniadis, P., Courcoubetis, C., and Weber, R. 2004b. An Asymptotically Optimal Scheme for P2P File Sharing. 2nd Workshop on Economics of Peer-to-Peer Systems, Harvard University.

Ausubel, L. M. and Milgrom, P. R. 2005. *The Lovely but Lonely Vickrey Auction.* MIT Press.

Axelrod, J. 1984. The evolution of cooperation. *N.Y.: Basic Books*.

Baev, I. D. and Rajaraman, R. 2001. Approximation algorithms for data placement in arbitrary networks. In *Proceedings of the 10th Annual Symposium on Discrete Algorithms (ACM-SIAM SODA).*

Banerjee, S., Bhattacharjee, B., and Kommareddy, C. 2002. Scalable application layer multicast. In *ACM SIGCOMM, Pittsburgh, PA, USA*.

Bar-Noy, A., Kipnis, S., and Schieber, B. 2000. Optimal multiple message broadcasting in telephone-like communication systems. *Discrete Applied Mathematics 100*, 1–15.

Batten, C., Barr, K., Saraf, A., and Treptin, S. 2001. pstore: A secure peer-to-peer backup system. Technical Memo MIT-LCS-TM-632, MIT Laboratory for Computer Science.

Bawa, M., Garcia-Molina, H., Gionis, A., and Motwani, R. 2003. Estimating aggregates on a peer-to-peer network. Technical report, Stanford University.

Benevenuto, F., Jr, J. I., and Almeida, J. 2004. Quantitative evaluation of unstructured peer-to-peer architectures. In *International Workshop on Hot Topics in Peer-to-Peer Systems (HOT-P2P'04)*.

Bertsekas, D. and Gallager, R. 1992. *Data Networks*. Prentice-Hall, Englewood Cliffs, New Jersey.

Buchegger, S. and LeBoudec, J.-Y. 2002. Performance Analysis of the CONFIDANT Protocol: Cooperation of nodes — Fairness in Dynamic Ad-hoc Networks. In *Proceedings of IEEE/ACMSymposium on Mobile Ad Hoc Networking and Computing(MobiHOC),Lausanne, CH*.

Buchegger, S. and LeBoudec, J.-Y. 2004. A Robust Reputation System for Peer-to-Peer and Mobile Ad-Hoc Networks. In *Second Workshop on Economics of Peer-to-Peer Systems, University of Harvard*.

Bungale, P., Goodell, G., and Roussopoulos, M. 2005. Conservation vs. consensus in peer-to-peer preservation systems. In *Proceedings of IPTPS 2005*.

Buragohain, C., Agrawal, D., and Suri, S. 2003. A Game Theoretic Framework for Incentives in P2P Systems. In *Proceedings of the Third IEEE International Conference on Peer-to-Peer Computing (P2P 2003), Linkoping, Sweden*.

Buttyan, L. and Hubaux, J.-P. 2003. Stimulating cooperation in self-organizing mobile ad hoc networks. *ACM/Kluwer Mobile Networks and Applications (MONET) 8,* 5 (October).

Castro, M., Druschel, P., Ganesh, A., Rowstron, A., and Wallach, D. 2002. Security for Structured Peer-to-Peer Overlay Networks. In *Proceedings of Multimedia Computing and Networking 2002 (MMCN '02)*.

Castro, M., Druschel, P., Kermarrec, A.-M., Nandi, A., Rowstron, A., and Singh, A. 2003. Splitstream: High-bandwidth multicast in cooperative environments. In *ACM SOSP, Bolton Landing, NY, USA*.

Castro, M., Druschel, P., Kermarrec, A.-M., and Rowstron, A. 2002. SCRIBE: A Large-scale and Decentralized Application-level Multicast Infrastructure. *IEEE Journal on Selected Areas in Telecommunications 20,* 8 (October), 100–110.

Chaum, D., Fiat, A., and Naor, M. 1990. Untraceable electronic cash. In *Proceedings of Advances in cryptology (CRYPTO88)*.

Chawathe, Y. 2000. Scattercast: An architecture for internet broadcast distribution as an infrastructure service. Ph.D Thesis, University of California, Berkeley.

Cheng, A. and Friedman, E. 2005. Sybilproof reputation mechanisms. In *ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems*.

Christin, N., Grossklags, J., and Chuang, J. 2004. Near Rationality and Competitive Equilibria in Networked Systems. In *Proc. SIGCOMM workshop on Practice and Theory of Incentives and Game Theory in Networked Systems*.

Christin, N., Weigend, A., and Chuang, J. 2005. Content Availability, Pollution and Poisoning in Peer-to-Peer File Sharing Networks. In *Proceedings of ACM E-Commerce Conference (EC'05)*.

Chu, Y., Ganjam, A., Ng, T. S. E., Rao, S. G., Sripanidkulchai, K., Zhan, J., and Zhang, H. 2003. Early experience with an internet broadcast system based on overlay multicast. Technical Report CMU-CS-03-214, Carnegie Mellon University.

Chu, Y.-H., Chuang, J., and Zhang, H. 2004. A Case for Taxation in Peer-to-Peer Streaming Broadcast. In *Proceedings of the ACM SIGCOMM workshop on Practice and Theory of Incentives in Networked Systems*.

Clarke, E. 1971. Multipart pricing of public goods. *Public Choice 1*, 17–33.

Clarke, I. 1999. Freenet white paper. Division of Informatics, University of Edinburgh.

Cohen, B. 2003. Incentives Build Robustness in BitTorrent. In *Workshop on Economics of Peer-to-Peer Systems, Berkeley, CA*.

Cooper, B. and Garcia-Molina, H. 2005. Ad hoc, self-supervising peer-to-peer search networks. *ACM Trans. Inf. Syst. 23,* 2, 169–200.

Cooper, B. F. and Garcia-Molina, H. 2002. Peer-to-peer resource trading in a reliable distributed system. In *Proceedings of the First International Workshop on Peer-to-Peer Systems, Cambridge, MA*.

Corbo, J. and Parkes, D. 2005. The price of selfish behavior in bilateral network formation. In *Proceedings of ACM PODC*.

Courcoubetis, C. and Weber, R. R. 2003. *Pricing Communication Networks : Economics, Technology and Modelling.* Wiley Europe.

Courcoubetis, C. and Weber, R. R. 2004. Asymptotics for Provisioning Problems of Peering Wireless LANs with a Large Number of Participants. In *Proceedings of WiOpt'04 workshop, University of Cambridge, UK*.

Courcoubetis, C. and Weber, R. R. 2005. Incentives for large p2p systems. accepted for publication in IEEE Journal on Selected Areas in Telecommunications, available at `http://nes.aueb.gr/p2p.html`.

Cox, L. and Noble, B. 2003. Samsara: Honor among thieves in peer-to-peer storage. In *Proceedings of the ACM Symposium on Operating Systems Principles*.

Cox, L. P., Murray, C. D., and Noble, B. D. 2002. Pastiche: Making backup cheap and easy. In *Proceedings of the 5th Symposium on Operating Systems Design and Implementation*.

Cramton, P., Gibbons, R., and Klemperer, P. 1987. Dissolving a partnership efficiently. *Econometrica 55*, 615–632.

Cronin, E., Jamin, S., Jin, C., Kurc, A. R., Raz, D., and Shavitt, Y. 2002. Constraint mirror placement on the internet. *IEEE Journal on Selected Areas in Communications 20,* 7 (September).

Crowcroft, J., Gibbens, R., Kelly, F., and Ostring, S. 2003. Modelling incentives for collaboration in mobile ad hoc networks. In *Proceedings of WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*.

D. Zeinalipour-Yazti, V. K. and Gunopulos, D. 2004. Information retrieval techniques for peer-to-peer networks. *Computing in Science and Engineering 06,* 4, 20–26.

Dellarocas, C. 2002. The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Systems - draft. MIRC Workshop Carlson School of Management, University of Minnesota, U.S.A.

Dellarocas, C. 2005. Reputation mechanisms. Handbook on Economics and Information Systems. Available at `http://faculty.haas.berkeley.edu/hender/ISEcon/ISecon.htm`.

DIgnazio, A. and Giovannetti, E. 2004. From exogenous to endogenous economic networks: Internet applications. University of Cambridge Economics Working Paper No. CWPE 0445.

Douceur, J. 2002. The sybil attack. In *Proceedings of the 1st International Workshop on Peer-to-Peer Systems, Boston, MA*.

Eagle, N. and Pentland, A. 2005. Social serendipity: Mobilizing social software. *IEEE Pervasive Computing 4,* 2, 28–34.

Efstathiou, E. C., Frangoudis, P. A., and Polyzos, G. C. 2006. Stimulating Participation in Wireless Community Networks. In *Proceedings of IEEE INFOCOM 2006, Barcelona*.

Efstathiou, E. C. and Polyzos, G. C. 2003. A Peer-to-Peer Approach to Wireless LAN Roaming. In *Proc. 1st ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots (WMASH 2003), San Diego, CA)*.

Efstathiou, E. C. and Polyzos, G. C. 2005. Self-Organized Peering of Wireless LAN Hotspots. *European Transactions on Telecommunications (special issue on Self-Organization in Mobile Networking)*.

Fabrikant, A., Luthra, A., Maneva, E., Papadimitriou, C., and Shenker, S. 2003. On a network creation game. In *Proceedings of ACM PODC*.

Feeney, L. M. 2001. An energy consumption model for performance analysis of routing protocols for mobile ad hoc networks. *Mobile Networks and Applications (MONET) 6,* 3, 239–249.

Feeney, L. M. and Nilsson, M. 2001. Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In *Proceedings of IEEE INFOCOM*. Anchorage AK, USA, 1548–1557.

Fehr, E. and Gachter, S. 2002. Altruistic Punishment in Humans. *Nature 415*, 137–140.

Feigenbaum, J., Papadimitriou, C., Sami, R., and Shenker, S. 2002. A BGP-based Mechanism for Lowest-Cost Routing. In *ACM Symposium on Principles of Distriubted Computing (PODC)*.

Feigenbaum, J., Papadimitriou, C., and Shenker, S. 2001. Sharing the cost of multicast transmissions. *Journal of Computer and System Sciences 63*, 21–41.

Feigenbaum, J. and Shenker, S. 2002. Distributed algorithmic mechanism design: Recent results and future directions. In *Proceedings of the International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*.

Feldman, M. and Chuang, J. 2005. The Evolution of Cooperation Under Cheap Pseudonyms. In *Proceedings of the Seventh International IEEE Conference on E-Commerce Technology*.

Feldman, M., Chuang, J., Stoica, I., and Shenker, S. 2005. Hidden-Action in Multi-Hop Routing. In *Proceedings of ACM Conference on Electronic Commerce (EC'05)*.

Feldman, M., Lai, K., Chuang, J., and Stoica, I. 2003. Quantifying Disincentives for Collaboration in Peer-to-Peer Networks. 1st Workshop on Economics of Peer-to-Peer Systems, University of Berkeley.

Feldman, M., Lai, K., Stoica, I., and Chuang, J. 2004. Robust Incentive Techniques for Peer-to-Peer Networks. ACM E-Commerce Conference (EC'04).

Feldman, M., Papadimitriou, C., Chuang, J., and Stoica, I. 2004. Free-Riding and Whitewashing in Peer-to-Peer Systems. In *Proceedings of 3rd Annual Workshop on Economics and Information Security (WEIS04).*

Figueiredo, D., Shapiro, J., and Towsley, D. 2005. Incentives to Promote Availability in Peer-to-Peer Anonymity Systems. In *Proceedings of International Conference on Network Protocols (ICNP 2005), Boston.*

Foster, I., Kesselman, C., and Tuecke, S. 2001. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *International Journal of Supercomputer Applications 15,* 3.

Francis, P. 2001. Yoid: Extending the internet multicast architecture. Technical Report, ACIRI.

Friedman, E. and Resnick, P. 2001. The Social Cost of Cheap Psedonyms. *Economics and Management Strategy 10,* 2, 173–199.

Fundenberg, D. and Tirole, J. 1992. *Game Theory.* Cambridge: The MIT Press.

Gigerenzer, G. and Selten, R. 2001. *Bounded Rationality.* MIT press.

Golder, S. A. and Huberman, B. A. 2005. The structure of collaborative tagging systems. Information Dynamics Lab, HP Labs.

Golle, P., Leyton-Brown, K., Mironov, I., and Lillibridge, M. 2001. Incentives for Sharing in Peer-to-Peer Networks. In *Proceedings of WELCOM'01.*

Green, J. R. and Laffont, J.-J. 1979. *Incentives in Public Decision Making.* North Holland: Amsterdam.

Groves, T. 1973. Incentives in teams. *Econometrica 41*, 617–663.

Gu, B. and Jarvenpaa, S. 2004. Are Contributions to P2P Technical Forums Private or Public Goods? - An Empirical Investigation. 2nd Workshop on Economics of Peer-to-Peer Systems, Harvard University.

Hales, D. and Edmonds, B. 2003. Evolving Social Rationality for MAS using 'Tags'. In *J.S., et al (ed.) Proceedings of the 2nd International Conference on Autonomous Agents and Multiagent Systems, Melbourne*.

Hales, D. and Patarin, S. 2005. How to Cheat BitTorrent and why nobody does. Technical Report UBLCS-2005-12, University of Bologna. May.

Hardin, G. 1968. The Tragedy of the Commons. *Science 162,* 3859, 1243–1248.

Hausheer, D., Liebau, N., Mauthe, A., Steinmetz, R., and Stiller, B. 2003. Token-based Accounting and Distributed Pricing to Introduce Market Mechanisms in a Peer-to-Peer File Sharing Scenario. In *Proceedings 3rd IEEE International Conference on Peer-to-Peer Computing, Linkoping, Sweden*.

Hausheer, D. and Stiller, B. 2005. PeerMart: The Technology for a Distributed Auction-based Market for Peer-to-Peer Services. In *Proceedings of the 40th IEEE International Conference on Communications (ICC 2005), Seoul, Korea*.

Herzog, J., Shenker, S., and Estrin, D. 1997. Sharing the "cost" of multicast trees: An axiomatic analysis. *IEEE/ACM Transactions on Networking 5,* 6, 847–860.

Hindriks, J. and Pancs, R. 2001. Free riding on altruism and group size. Working paper No. 436.

Hogg, L. M. and Jennings, N. R. 1997. Socially rational agents. In *Proceedings of AAAI Fall symposium on Socially Intelligent Agents, Boston*.

Holland, J. 1993. The effect of lables (tags) on social interactions. Santa Fe Institute Working Paper 93-10-064. Santa Fe, NM.

Holmstrom, B. 1979. Groves schemes on restricted domains. *Econometrica 47*, 1137–1144.

Huang, E., Crowcroft, J., and Wassell, I. 2004. Rethinking Incentives for Mobile Ad Hoc Networks. In *Proceedings of the ACM SIGCOMM workshop on Practice and theory of incentives in networked systems*.

Jackson, M. and Wolinsky, A. 1996. A strategic model of social and economic networks. *Journal of Economic Theory 71*, 44–74.

Jakobsson, M., Hubaux, J.-P., and Buttyan, L. 2003. A Micro-Payment Scheme Encouraging Collaboration in Multi-Hop Cellular Networks. *Financial Cryptography*.

Jannotti, J., Gifford, D. K., Johnson, K. L., Kaashoek, M. F., and O'Toole, J. W. 2000. Overcast: Reliable multicasting with an overlay network. In *Proceedings of the Fourth Symposium on Operating Systems Design and Implementation.*

Josang, A., Hird, S., and Faccer, E. 2003. Simulating the Effect of Reputation Systems on e-Markets. In *Proceedings of the 1st International Conference on Trust Management, Crete, Greece.*

Jun, S. and Ahamad, M. 2005. Incentives in bittorrent induce free riding. In *ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems.*

Kalpakis, K., Dasgupta, K., and Wolfson, O. 2001. Optimal placement of replicas in trees with read, write, and storage costs. *IEEE Transactions on Parallel and Distributed Systems 12,* 6 (June), 628–637.

Kamvar, S., Schlosser, M., and Garcia-Molina, H. 2003a. The EigenTrust Algorithm for Reputation Management in P2P Networks. In *Proceedings of the Twelfth International World Wide Web Conference.*

Kamvar, S. D., Schlosser, M. T., and Garcia-Molina, H. 2003b. EigenRep: Reputation Management in P2P Networks. In *Proceedings of the Twelfth International World Wide Web Conference, Budapest, Hungary.*

Kangasharju, J., Roberts, J., and Ross, K. W. 2002. Object replication strategies in content distribution networks. *Computer Communications 25,* 4 (March), 376–383.

Kempe, D., Dobra, A., and Gehrke, J. 2003. Gossip-based computation of aggregate information. In *Proceedings of FOCS.*

Kenyon, C. 2003. Creating Services with Hard Guarantees from Cycle-Harvesting Systems. In *Proceedings of CCGrid 2003, Tokyo.* IBM Research Report RZ 3461 11/11/2002.

Kenyon, C. 2004. Grid Resource Commercialization: Economic Engineering and Delivery Scenarios. In *Grid Resource Management: State of the Art and Research Issues.* Editors: J. Nabrzyski, J. Schopf and J. Weglarz.

Kling, R., Lee, Y., Teich, A., and Frankel, M. 1999. Assessing anonymous communication on the internet: Policy deliberations. *The Information Society 15,* 2.

Konrad, K. A. 1994. The strategic advantage of being poor: Private and public provision of public goods. *Economica, New Series 61,* 241, 79–92.

Korupolu, M. R., Plaxton, C. G., and Rajaraman, R. 1999. Placement algorithms for hierarchical cooperative caching. In *Proceedings of the 10th Annual Symposium on Discrete Algorithms (ACM-SIAM SODA).*

Kostic, D., Rodriguez, A., Albrecht, J., and Vahdat, A. 2003. Bullet: High bandwidth data dissemination using an overlay mesh. In *ACM SOSP*.

Krauter, K., Buyya, R., and Maheswaran, M. 2001. A Taxonomy and Survey of Grid Resource Management Systems for Distributed Computing. *Internationl Journal of Software: Practice and Experience (SPE)*.

Krishnan, P., Raz, D., and Shavit, Y. 2000. The cache location problem. *IEEE/ACM Transactions on Networking 8,* 5, 568–581.

Krishnan, R., Smith, M., Tang, Z., and Telang, R. 2003. The Virtual Commons: why Free-Riding can be Tolerated in Peer-to-Peer Networks. Workshop on Information Systems and Economics.

Laoutaris, N., Telelis, O., Zissimopoulos, V., and Stavrakakis, I. 2005. Distributed selfish replication. *IEEE Transactions on Parallel and Distributed Systems*. accepted for publication.

Laoutaris, N., Zissimopoulos, V., and Stavrakakis, I. 2004. Joint object placement and node dimensioning for internet content distribution. *Information Processing Letters 89,* 6 (March), 273–279.

Laoutaris, N., Zissimopoulos, V., and Stavrakakis, I. 2005. On the optimization of storage capacity allocation for content distribution. *Computer Networks 47,* 3 (February), 409–428.

Leff, A., Wolf, J. L., and Yu, P. S. 1993. Replication algorithms in a remote caching architecture. *IEEE Transactions on Parallel and Distributed Systems 4,* 11 (November), 1185–1204.

Li, B., Golin, M. J., Italiano, G. F., Deng, X., and Sohraby, K. 1999. On the optimal placement of web proxies in the internet. In *Proceedings of the Conference on Computer Communications (IEEE Infocom), New York*.

Liang, J., Kumar, R., and Ross, K. W. 2004. Understanding kazaa. submitted.

Liang, J., Kumar, R., Xi, Y., and Ross, K. W. 2005. Pollution in P2P File Sharing Systems. In *Proceedings of IEEE Infocom, Miami, FL, USA*.

Liebau, N., Darlagiannis, V., Mauthe, A., and Steinmetz, R. 2005. Token-based Accounting for P2P-Systems. In *Proceedings of Kommunikation in Verteilten Systemen KiVS 2005*. 16–28.

Lillibridge, M., Elnikety, S., Birrell, A., Burrows, M., and Isard, M. 2003. A cooperative internet backup scheme. In *Proceedings of the USENIX Annual Technical Conference, San Antonio, TX*.

Loukopoulos, T. and Ahmad, I. 2004. Static and adaptive distributed data replication using genetic algorithms. *Journal of Parallel and Distributed Computing 64,* 11 (November), 1270–1285.

Lua, E., Crowcroft, J., Pias, M., Sharma, R., and Lim, S. 2005. A survey and comparison of peer-to-peer overlay network schemes. *IEEE Communications Surveys & Tutorials 7,* 2, 72–93.

Lui, S., Lang, K., and Kwok, S. 2002. Participation Incentive Mechanisms in Peer-to-Peer Subscription Systems. In *Proceedings of the 35th Hawaii International Conference on System Sciences*.

M. F. Hellwig. 2003. Public-Good Provision with Many Participants. *Review of Economic Studies 70*, 589–614.

Mailath, G. and Postlewaite, A. 1990. Asymmetric information bargaining problems with many agents. *Review of Economic Studies 57*, 351–368.

Marx, G. 1999. What's in a name? some reflections on the sociology of anonymity. *The Information Society 15,* 2.

Mas-Colell, A., Whinston, M. D., and Green, J. R. 1995. *Microeconomic Theory.* Oxford University Press, New York.

Mas-Collel, A., Whinston, M. D., and Green, J. R. 1995. *Microeconomic Theory.* Oxford: Oxford University Press.

Mason, R. A. and Valentinyi, A. 2003. Independence, heterogeneity and uniqueness in interaction games. Mimeo. Available at http://www.soton.ac.uk/~ram2.

Maymounkov, P. and Mazieres, D. 2002. Kademlia: A peer-to-peer information system based on the xor metri. In *Proceedings of 1st International Workshop on Peer-to-Peer Systems (IPTPS)*.

MMAPPS-Consortium. 2004. Market management of peer-to-peer services - white paper. Available at http://www.mmapps.info/.

Moreton, T. and Twigg, A. 2003. Trading in Trust, Tokens, and Stamps. 1st Workshop on Economics of Peer-to-Peer Systems, University of Berkeley.

MUNDINGER, J. AND WEBER, R. R. 2004. Efficient file dissemination using peer-to-peer technology. Technical Report, Statistical Laboratory Research Reports 2004-01.

MYERSON, R. B. AND SATTERTHWAITE, M. A. 1983. Efficient mechanisms for bilateral trading. *Journal of Economic Theory 29*, 265–281.

NG, A., ZHENG, A., AND JORDAN, M. 2001. Link Analysis, Eigenvectors, and Stability. In *International Joint Conference on Artificial Intelligence (IJCAI-01)*.

NGAN, T.-W. J., WALLACH, D. S., AND DRUSCHEL, P. 2003. Enforcing fair sharing of peer-to-peer resources. In *2nd International Workshop on Peer-to-Peer Systems (IPTPS)*. Berkeley, California.

NISAN, N. AND RONEN, A. 1999. Algorithmic Mechanism Design. In *Proceedings of the 31st Symposium on Theory of Computing*.

NORMAN, P. 2004. Efficient mechanisms for public goods with use exclusions. *Review of Economic Studies 7*, 1163–1188.

NOWAK, M. A. AND SIGMUND, K. 1998. The dynamics of indirect reciprocity. *Journal of Theoretical Biology 194*, 561–574.

ODLYZKO, A. 2001. Internet Pricing and the History of Communications. *Computer Networks 36,* 5–6, 493–517.

PADMANABHAN, V. N., WANG, H. J., AND CHOU, P. A. 2003. Resilient peer-to-peer streaming. In *IEEE ICNP, Atlanta, GA, USA*.

PAPAIOANNOU, T. AND STAMOULIS, G. 2004. Effective Use of Reputation in Peer-to-Peer Environments. In *Proc. of IEEE/ACM CCGRID 2004 (Workshop on Global P2P Computing)*.

PAPAIOANNOU, T. AND STAMOULIS, G. 2005. An Incentives' Mechanism Promoting Truthful Feedback in Peer-to-Peer Systems. In *Proc. of IEEE/ACM CCGRID 2005 (Workshop on Global P2P Computing)*.

QIU, D. AND SRIKANT, R. 2004. Modeling and Performance Analysis of BitTorrent-Like Peer-to-Peer Networks. In *Proceedings of ACM SIGCOMM 2004 Conference, Portland, OR, USA*.

QIU, L., PADMANABHAN, V., AND VOELKER, G. 2001. On the placement of web server replicas. In *Proceedings of the Conference on Computer Communications (IEEE Infocom), Anchorage, Alaska*.

Radner, R. 1980. Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives. *Journal of Economic Theory 22*, 136–154.

Ranganathan, K., Ripeanu, M., Sarin, A., and Foster, I. 2003. To Share or Not to Share: An Analysis of Incentives to Contribute in Collaborative File Sharing Environments. 1st Workshop on Economics of Peer-to-Peer Systems, University of Berkeley.

Ratnasamy, S., Francis, P., Handley, M., Karp, R., and Shenker, S. 2000. A scalable content addressable network. Tech. Rep. TR-00-010, Berkeley, CA.

Resnick, P., Zeckhauser, R., Friedman, E., and Kuwabara, K. 2000. Reputation systems. *Communications of the ACM 43,* 12.

Richarson, M., Agrawal, D., and Domingos, P. 2003. Trust Management for the Semantic Web. In *Proceedings of the Second International Semantic Web Conference.*

Rivest, R. L. and Shamir, A. 1996. Payword and micromint: two simple micropayment schemes. In *Proceedings of International Workshop on Security Protocols.*

Rosenschein, J. S. and Zlotkin, G. 1994. *Rules of Encounter.* MIT Press.

Rosenthal, R. W. 1973. A class of games possessing pure-strategy nash equilibria. *InternationalJournal of Game Theory 2*, 65–67.

Rowstron, A. and Druschel, P. 2001. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In *IFIP/ACM International Conference on Distributed Systems Platforms (Middleware), Heidelberg, Germany.*

Saroiu, S., Gummadi, P., and Gribble, S. 2002. A measurement study of peer-to-peer file sharing systems. In *Proceedings of Multimedia Conferencing and Networking, San Jose.*

Saroiu, S., Gummadi, P., and Gribble, S. 2003. Measurement, modelling, and analysis of a peer-to-peer file sharing workload. In *Proceedings of SOSP.*

Segal, I. 2003. Optimal pricing mechanisms with unknown demand. *The American Economic Review 93,* 3, 509–529.

Srinivasan, V., Nuggehalli, P., Chiasserini, C., and Rao, R. 2003. Cooperation in wireless ad hoc networks. In *Proceedings of Infocom 2003.*

Stoica, I., Morris, R., Liben-Nowell, D., Karger, D., Kaashoek, M., Dabek, F., and Balakrishnan, H. 2003. Chord: A scalable peer-to-peer lookup protocol for internet applications. *IEEE/ACM Transactions on Networking 11,* 1, 17–32.

Strahilevitz, L. 2003. Charismatic Code, Social Norms, and the Emergence of Cooperation on the File-Swapping Networks. Virginia Law Review, Vol. 89.

Sun, Q. and Garcia-Molina, H. 2004. SLIC: A Selfish Link-Based Incentive Mechanism for Unstructured Peer-to-Peer Networks. In *ICDCS*.

Teich, A., Frankel, M., Kling, R., and Lee, Y. 1999. Anonymous communication policies for the internet: Results and recommendations of the aaas. *The Information Society 15,* 2.

Varian, H. 1992. *Microeconomic Analysis*. Norton.

Varian, H. 1995. Economic Mechanism Design for Computerized Agents. In *Proc. of Usenix Workshop on Electronic Commerce*.

Vickrey, W. 1961. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance 16*, 8–37.

Vishnumurthy, V., Chandrakumar, S., and Sirer, E. G. 2003. Karma: A Secure Economic Framework for Peer-to-Peer Resource Sharing. 1st Workshop on Economics of Peer-to-Peer Systems, University of Berkeley.

Wallach, D. S. 2002. A survey of peer-to-peer security issues. In *International Symposium on Software Security, Tokyo, Japan*.

Wayner, P. 1997. *Digital Cash: Commerce on the Net*. Morgan Kaufmann. 2nd edition.

Wilson, R. 1985. *Reputations in Games and Markets*. Cambridge University Press, Cambridge, UK.

Yang, B. and Garcia-Molina, H. 2003. Designing a super-peer network. In *Proc. the 19th International Conference on Data Engineering, Bangalore, India*.

Y.Chu, Rao, S., and Zhang, H. 2000. A case for end system multicast. In *Proceedings of ACM SIGMETRICS, Santa Clara, CA*.

Zhao, B., Huang, L., Stribling, J., Rhea, S., Joseph, A., and Kubiatowicz, J. 2004. Tapestry: A Resilient Global-scale Overlay for Service Deployment. *IEEE Journal on Selected Areas in Communications 22,* 1 (January).

Zhong, S., Chen, J., and Yang, Y.-R. 2003a. Sprite: A simple, cheat-proof, credit-based system for mobile ad-hoc networks. In *Proceedings of Infocom 2003*.

Zhong, S., Chen, J., and Yang, Y. R. 2003b. Sprite: A Simple, Cheat-Proof, Credit-Based System for Mobile Ad-Hoc Networks. In *22nd Annual Joint Conference of the IEEE Computer and Communications Societies*.