POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS

Vol. 27, No. 11, November 2018, pp. 1940–1959
ISSN 1059-1478 | EISSN 1937-5956 | 18 | 2711 | 1940
DOI 10.1111/poms.12718

Contracting Models for P2P Content Distribution

Hossein Ghasemkhani
Krannert School of Management, Purdue University, West Lafayette, Indiana 47907, USA, hossein@purdue.edu

Yung-Ming Li
College of Management, National Chiao Tung University, Hsinchu 300, Taiwan, yml@mail.nctu.edu.tw

Kamran Moinzadeh
Michael G. Foster School of Business, University of Washington, Seattle, Washington 98195, USA, kamran@uw.edu

Yong Tan
Michael G. Foster School of Business, University of Washington, Seattle, Washington 98195, USA
School of Economics and Management, Tsinghua University, Beijing, China, ytan@uw.edu

In recent years, peer-to-peer (P2P) networks have become an increasingly popular method for distributing digital content. In this study, we consider the development of optimal contracts for a P2P network by a profit-seeking provider to support the operations of an online file exchange service. By utilizing the principal-agent model of incentive theory, we propose appropriate reward and pricing schemes for profit-seeking P2P content distribution networks. We show that when peers are homogeneous, upload compensation increases with propagation delay uncertainty, maximum uploading nodes allowed, peers' provision cost and disutility of download delay, but decreases with the network size and content availability. We also characterize a general contracting model where there are a countable number of peer classes which are heterogeneous in their provisioning costs. For the case of two peer classes where optimal delays are separable, we derive the optimal upload compensations under different scenarios and show that the impact of operational parameters is quite similar to the case of homogeneous peers, lending support to the robustness of our analysis.

Key words: P2P networks; content distribution; moral hazard; principal-agent model
History: Received: April 2015; Accepted: March 2017 by Subodha Kumar, after 2 revisions.

1. Introduction

A peer-to-peer or P2P network is a social network for pooling resources, such as computing cycle, hard disk storage, network bandwidth and content, at numerous edge nodes. Business applications of P2P include grid computing, instant messaging, collaborating and file sharing. In 2008, Insight Research Corporation predicted that the global market for P2P and file sharing services would reach nearly $28 billion in revenues for carriers and ISPs by 2012 (The Insight Research Corporation 2008). While the growth in P2P file sharing experienced a decline starting in 2008 as a result of legal actions by the likes of RIAA and access to convenient alternatives like iTunes and Netflix, a 2015 report uses data from Cisco Visual Networking Index and other reports to conclude that file sharing has likely grown by around 55% from 2008 to 2015 (CVN Index 2015, Steele 2015).

In this study, we argue that there are great opportunities for a commercial P2P content distribution platform, and propose a mechanism to implement such a platform. P2P networks are self-organizing and independently operated by different parties, but the decentralized nature of these networks has made them free havens for pirated content (Kalker et al. 2004). Coming up with an effective mechanism to use the power of P2P networks for legal sharing will not only mitigate piracy, but also tap into the collective power of shared resources of a large number of users. The operational performance of a P2P network is strongly associated with the level of contribution by the peers. Therefore, in order to have a sustainable commercial P2P network, the most fundamental question is how to stimulate peers' incentives to cooperate or contribute resources, a problem that poses an obvious obstacle for P2P to be used as a distribution channel (Krishnan et al. 2006). Many computer scientists have tried to address this issue from the technical perspective. We address this problem from the economic perspective by aligning incentives.

Considering P2P networks' potential for both scalable distribution and copyright infringement, media companies face the dilemma of either suppressing the diffusion of this new distribution channel or embracing this seemingly unavoidable new technology. Incorporated with Digital Rights Management, P2P content distribution services are increasingly emerging. For example, Snocap uses a system of sound fingerprinting which allows songs traded over a P2P network to be positively identified as belonging to a particular copyright holder (Betteridge 2004). StreamerP2P applies the P2P distribution approach to deliver digital radio (StreamerP2P 2016). In 2006, Warner Brothers Home Entertainment announced the adoption of BitTorrent for distribution of protected movies and television programs (Marlowe 2006). There has been news circulating that even Netflix which is currently responsible for over 30% of all downstream traffic might be considering using P2P technology for content delivery (Hastings 2014). They have also been hiring employees to develop the capability to use the P2P technology (Van der Sar 2015). If it adopts P2P streaming, Netflix will be number one in terms of upstream bandwidth too, a position currently dominated by BitTorrent traffic (TorrentFreak 2014a). P2P applications have also expanded to mobile platforms. Popcorn Time, a mobile application which offers BitTorrent-powered streaming became an instant hit and was downloaded by 1.4 million devices in the US in a few months (TorrentFreak 2014b).

These emerging commercial P2P content distribution services indicate that resources of peer nodes could be managed by leveraging technological and economic mechanisms. With the support of appropriate service quality measurement, operating policy, and incentive mechanism, a profitable P2P content distribution network with an efficient delivery service could be realized. However, designing an economic mechanism to run a commercial P2P network calls for attention to the unique characteristics of these networks. In such a network, there are multiple agents who act as downloading or uploading nodes. There is also the platform provider who designs and implements the protocols and acts as the governing body. Also the way many existing P2P content networks are designed makes them similar to a public good, thus, giving rise to the free-riding problem. Free-riding occurs as many users take advantage of the content provided through the network, but do not contribute back by sharing the content with others. As Krishnan et al. (2006) emphasize, lacking incentives in P2P networks, free riding will likely abound.

Due to the decentralized nature of P2P networks and the stochastic characteristics of their operating environments, e.g., uncertain routing and traffic in public network domain, it is difficult to observe and measure the effort exerted by peer nodes, resulting in information asymmetry. Thus, external nodes can only estimate the overall network performance between request and provision nodes. This is particularly relevant for commercial P2P networks and gives rise to moral hazard and adverse selection problems. The moral hazard or hidden action problem exists because the upload bandwidth dedicated by the peers is not directly observable by the provider, due to the existence of propagation delay or simply because the actual realized bandwidth might be lower than the nominal bandwidth. Therefore, in the absence of a monitoring scheme, peer nodes could shirk and avoid the cost incurred during file transfer. The adverse selection or hidden knowledge problem exists because peers might have different provision costs due to differences in their cost to acquire bandwidth or differences in opportunity costs of contributing resources.

Extensive research has been done on technological design to improve P2P content distribution and on the design of incentives to alleviate the free-riding problem. However, little attention has been given to the development of a business contract under information asymmetry for profit-seeking P2P distribution channels, which is the economic approach to conquer the inherent hidden action problem. As Laffont and Martimort (2002) point out, the starting point of incentive theory is the problem of delegating a task to an agent with private information which can be of two types, moral hazard or adverse selection, and incentive theory considers when this private information is a problem for the principal and what the optimal way for the principal is to cope with it. So the special characteristics of the P2P networks make the incentive theory and principal agent model the approach of choice for designing an economic mechanism.

In this study, we develop contracts implemented between the P2P provider, as the principal, and the peer nodes, as agents, to support content distribution. According to our proposed reward scheme, the provider compensates peers for agreeing to become provision nodes by making content available for download. Moreover, an additional upload reward is paid to peers who execute the file upload. Under this compensation and payment structure, P2P provider's decision is to whether provide higher compensation to induce peers to exert higher provision quality, or to provide lower compensation to reduce operational costs. On the other hand, since only the provision nodes with the highest service quality (measured by the end-to-end transfer delay) will be chosen as provision nodes, the peers have to compete by exerting higher upload capacity. Thus, the fundamental trade-off participating peers face is allocating a higher bandwidth and increasing their chances of winning at the expense of incurring the additional cost.

Regarding the demand side, customers (content request nodes) must be charged a download fee. This fee pays for the the right to own the content, as well as the compensation awarded to uploading peers. Download fee is determined by the download QoS (i.e., overall delay), which is associated with network size, capacity contribution of peer nodes and the characteristics of the uncontrollable public network. It is important to analyze how relevant system parameters affect the development of reward and pricing mechanisms and adjust the system parameters to optimize the mechanism developed.

By utilizing the principal-agent model from incentive theory (Gibbons 1992, Green and Stokey 1983, Lazear and Rosen 1981), we formulate and solve a multilateral contracting problem for a monopolistic P2P provider. We investigate the circumstances under which a commercial P2P network is viable. This is an important problem, as existing P2P platforms are predominantly free and non-commercial. We show that our proposed commercial P2P platform is not viable for content that is extremely old or too obscure. Also if users have extremely low opportunity costs of time and do not mind very long download delays, or if it is prohibitively expensive for users to provide upload bandwidth, the commercial P2P network might not be profitable for any type of content. However, these cases are increasingly less likely given the trends in user behavior and internet infrastructure improvements. Finally, we show that when network connections are too unreliable or unpredictable, a commercial P2P content sharing network might not be viable. On the other hand, the ongoing improvements in infrastructure will help the rise of commercial P2P networks.

Our proposed contract tackles the information asymmetry issues we discussed earlier. It also eliminates the free-riding problem, as uploading peers are incentivized to contribute through a reward mechanism and downloading peers pay to acquire content. We also investigate the impact of key characteristics of P2P systems such as network size and network dispersion, content availability, and P2P technology (file swarming, i.e., the maximal number of provision nodes allowed to execute upload, and bandwidth technology) on the design of the contract and the corresponding capacity provided by the peer nodes.

We show that the population of P2P users and market structure have a significant impact on the resulting compensation and capacity selection. Optimal upload compensation always increases with propagation delay in the public network domain and the maximum number of peers allowed to upload jointly. It also always decreases with network size and content availability. These effects are shown to also hold for the case of two heterogeneous classes of peers. Upload compensation also increases or remains constant with peers' provision cost and disutility of delay, depending on values of other parameters. These same effects are also observed for the two-class model, only that each class has its own cost and the relative size of the costs also matters. As the network size increases, peer nodes provide less capacity, and individual provision delay also increases with content availability. Provision cost has a similar impact on optimal delay. On the other hand, delay decreases or remains constant with the maximum number of simultaneous uploaders, propagation delay uncertainty and disutility of delay, depending on the values of other parameters.

In this study, we connect the profit-maximizing business strategy of the P2P network provider with operational details and derive managerial insights at the operational level. This work contributes to the literature on P2P pricing as it analyzes a rank-order based compensation scheme in order to address the hidden action problem of upload bandwidth provision. The proposed contract scheme is suited to the popular P2P technological protocols which allow simultaneous download from multiple provision nodes. Though conceptually similar to the studies in the contest literature, the analyses in this study are tailored to the P2P context which possesses some unique characteristics. For example, the uncertainty due to propagation delay always affects the overall content transfer performance and there exists significant positive network externalities on the propagation delay, whereas in traditional contest models, the uncertainty (noise) plays a negligible role.

Our mechanism has three clear advantages over the free mechanism. It eliminates piracy, incentivizes peers to participate in the upload of less popular content, and eliminates free riding. One other possible mechanism is a one-part flat reward in which peers who contribute receive a flat reward regardless of content and environmental characteristics. We will argue later that even for a very simple case with minimal diversity, our model yields far superior payoffs compared to this flat reward model. Finally the subscription model used by companies like Netflix is straightforward and easy to implement, but it is not viable in our setting. First, although great for streaming, it is not suitable for downloading content to own. Second, if the reward is also paid as a fixed monthly amount, it will not incentivize peers to participate more in uploading less popular content. Finally, we cannot effectively deal with moral hazard and monitoring issues using a flat monthly compensation.

2. Literature Review

Numerous studies have looked at the technical aspects of P2P networks such as network topologies, search algorithms, and developing efficient communication protocols (Gupta and Awasthi 2010, Victer Paul et al. 2012). Many of the existing economic studies of P2P networks focus on the free-riding phenomenon and the ways to alleviate this problem. Adar and Huberman (2000) analyzed user traffic on Gnutella, one of the most popular P2P file sharing networks, and found that around 70% of Gnutella users share no files. As a decentralized system, a P2P file sharing network faces similar challenges as traditional public goods (Cornes and Sandler 1996): under-provision and over-consumption of the commons. This phenomenon has been observed and analyzed (Asvanund et al. 2004, Krishnan et al. 2004, Li et al. 2013), and many incentive mechanisms have been proposed to conquer this free-riding problem (Antoniadis et al. 2004, Ghosal et al. 2005, Ma et al. 2006). Our proposed model deals with free-riding using two mechanisms: (i) users who download content, pay for doing so, and (ii) users who contribute to file upload are compensated by an availability reward, w_0, and an upload reward, w_1.

Among other economic studies on P2P networks, Johar et al. (2011) examine the use of P2P networks for delivering general-purpose content over the Web under various congestion measures, and Hosanagar et al. (2010) look at the diffusion of content in P2P networks. Herings et al. (2010) and Johar et al. (2012) study the situation where an information goods provider has to compete with a P2P network where pirated content is provided. Using peer-level data from a music sharing network, Xia et al. (2012) find that the more users benefit from the network and the more they give to the network, the more likely they are to share. Guo et al. (2013) examine the case of Comcast slowing down P2P traffic from a net neutrality standpoint. Xu et al. (2010) analyze a P2P streaming peer selection model using mobile agents and a super-peer selection strategy based on trust.

Our study can be placed under the general umbrella of the literature on pricing of information goods. However, we study digital content pricing in P2P networks where the price depends not only on the value of the content, but also on the delays incurred in the network. More specifically, our study is related to the literature on pricing of P2P services and resources. Maillé and Toka (2008) study a simple P2P storage marketplace where users can buy storage space for a fixed unit price, and sell their own memory space to the system at another price. Our research considers a commercial P2P network where the platform owner incentivizes peers to upload and downloading peers pay a fee to acquire content based on the quality of service. Kumar et al. (2011) design a pricing and allocation mechanism for a P2P network in an organization in a decentralized manner. They model the allocation of tasks to peers in a self-organized way, while we study a network where peers compete to be selected to upload content. Lang and Vragov (2005) develop a monopolistic pricing scheme to reward P2P peers for contributing, however in contrast to our study, their model is developed based on the economic growth theory and does not consider the competition among uploading peer nodes.

This work is also related to network pricing literature. Most of the research associated with communication networks use delay as the quality measure for the development of pricing schemes, such as congestion-based pricing models for Internet services (Ganesh et al. 2007, MacKie-Mason and Varian 1995). Considering the delay cost, Afèche and Mendelson (2004) study alternative price-service mechanisms for a provider that serves customers whose delay costs depend on their service valuations. Similar to such studies, we also use the end-to-end delay of file transfer as the main QoS evaluation criteria. However, we study delay in the P2P context where moral hazard and adverse selection exist.

The problem of hidden action or moral hazard, has long been of interest in economics literature on information asymmetry, and contract and agency theory. This principal-agent problem has been widely used as a representation of various standard economic relations such as in the theory of insurance under moral hazard (Arrow 1970, Spence and Zeckhauser 1971), in efficiency wage theory (Shapiro and Stiglitz 1984), and in teams with sequential hidden actions (Holmstrom 1982, Strausz 1996). More recently, Zhang and Zenios (2008) propose a general framework for a large class of multi-period principal-agent problems and develop a dynamic programming algorithm to derive optimal long-term contracts. In the context of networks, existing studies mainly focus on mechanism or contract design to conquer the hidden action problem occurring in network routing (Feigenbaum et al. 2002, Feldman et al. 2007). In this study, we propose a mechanism to overcome the hidden action problem in the context of P2P content distribution.

Previous literature on P2P content distribution has mainly studied the problems from the perspective of cooperative distribution channels, without considering the competition and information asymmetry issues among the players in the digital supply chain. In this study, considering these important factors, we present reward and pricing schemes to support the operations of commercial P2P content sharing networks.

3. The Model

We consider a content-sharing P2P network in which a profit-seeking P2P provider (firm) develops contracts to support the operations of an online file exchange service. We assume the system is in a stable state, meaning the P2P provider has been in service for some time, so there are a number of peers who have the P2P client installed on their machines. In other words, we do not model the decision of a potential peer to join the P2P network. We will study the two-sided contracts between a peer who is willing to pay to download content and the P2P provider on one hand, and the P2P provider and peers who are willing to contribute to file upload in return for some compensation on the other hand. Both the price and the amounts paid as compensation to participating peers are different for different content. The price that a downloading peer is willing to pay for a specific file is assumed to depend on the intrinsic value of the file and the end-to-end transfer delay. Intrinsic value of the file is the market price of the content where it is readily available. We also assume all files are initially provided by the provider to ensure legality and consistent quality. So uploading peers have already purchased and downloaded the content on the network. After downloading a file, peers are shown the relevant rewards and have the choice to keep the content available, decide on the bandwidth and be prepared to participate.

When a download request is initiated, the P2P provider searches the content catalog (either through a central index or through distributed search on the peer nodes), and provides download information (transfer delay estimation) of potential provision nodes. We call a node which has the requested content available an available peer node. Flow of monetary transfer is as follows: P2P provider receives the download price from the request peer node, and gives the provision nodes the upload service compensation after file transfer is complete. All available nodes receive a fixed availability payment, w_0, but only the available peers who are selected to provide the file receive the upload payment, w_1. w_0 could serve as an incentive for peers to have the file available (although, as we will discuss later, we do not model this decision). It also provides an additional means for the P2P provider to control the types of peers it wants to induce to participate, in the case where peers are heterogeneous. In the absence of information asymmetry and the ensuing moral hazard issue, the provider could simply give all the participating peers a single reward, w, and enforce an optimal level of provision capacity. However, due to information asymmetry, the risk associated with the uncertainty has to be transferred to the peers to induce them to choose the optimal provision capacity.

3.1. P2P Dynamics and Operating Policies

While a number of advanced P2P technologies have been developed, the operations of a typical P2P file sharing network can be summarized as having three main procedures: First, a peer node initiates a content request. Then, after searching the profiles of peer nodes, the P2P provider recommends a list of provision nodes that have the desired file and are “close” to the request node (determined by the expected file transfer latency between these two nodes). Once this information is passed on to the requesting and provision nodes, download occurs directly between the peers. The stated recommendation policy is commonly implemented in most of the popular P2P file sharing platforms (such as BitTorrent and KaZaA). The network dynamics and operating policy are specifically described as follows, and the model parameters and decision variables are listed in Table 1.

Content provision: Assume there are N active peer nodes in a certain P2P network. Due to the limit on storage capacity and the uncertainty of content availability at the peer nodes, we introduce binary random variables, x_{ik}, to indicate the availability of file k stored on an arbitrary node i. Specifically, let

$$\begin{aligned} &x_{ik} = 1, \text{ if node } i \text{ has file } k \text{ and,} \\ &x_{ik} = 0, \text{ otherwise}, \forall i \in \{1, \ldots, N\} \end{aligned} \tag{1}$$

Then $\beta_k = E(x_{ik})$, denotes the probability that node i has file k. We assume that this distribution is known to both the P2P provider and the peers. Later on, as we will study the contract for a given file, we will drop the k subscript.

Table 1 Model Parameters

	Description
Parameters	
N	Total number of peer nodes
$\hat{n}$	Maximum number of provision nodes allowed to upload simultaneously
τ	Upper bound of propagation delay for 1 byte of data
β	Probability of content availability
c_i	Cost of capacity provisioning for peer i
T_i	Propagation delay of peer i for 1 byte of data
Q_i	End-to-end transfer delay of peer i ($Q_i = D_i + T_i$)
v	Intrinsic value of content
γ	Disutility of transfer delay for download of 1 byte of data
Decision variables	
w_0	Availability compensation
w_1	Upload compensation
D_i	Provision delay (transmission delay of peer i for one byte of data)
π_0	Download price

Transfer delay: P2P technologies utilize the aggregate bandwidth from edge nodes for content transmission to avoid congestion at dedicated servers. Therefore, the effective bandwidth is scalable with the number of active users. End-to-end transfer delay, from here on referred to as *transfer delay* (denoted by Q) in a P2P content distribution network includes transmission delay, propagation delay, processing delay and queueing delay (Bertsekas and Gallager 1992). Transmission delay occurs at the end edge of a provision node, so transmission delay at a peer node can be improved if bandwidth capacity (upload speed) is increased. We call the transmission delay for 1 byte of data *provision delay* (denoted by D) which is equal to the inverse of the effective bandwidth. D_i is endogenously decided by node i. For ease of exposition and to highlight the most important aspect of the other delays (geographical dispersion), we use the term *propagation delay* (denoted by T) to refer to the sum of propagation, processing and queueing delays. In other words, propagation delay captures the additional delay incurred on top of D_i. In addition to network quality factors (such as congestion), propagation delay depends on the distance between peers, so the overall propagation delay depends on the number of active nodes—the more active nodes are available, the more likely it is to find provision nodes closer to the request node. Since uptime and location of participating peers are stochastic, propagation delays, T_i, are assumed to be random variables, distributed *i.i.d.* across the different provision nodes with density function f_T.

Provision policy: $\hat{n}$ is the maximum number of peer nodes allowed to jointly upload the requested file simultaneously. While a larger number of uploading peers could result in lower delay, the corresponding administrative cost of managing concurrent downloads increases with the number of uploading peers. Therefore, there is an optimal number of provision nodes that a provider can set. Here, we treat $\hat{n}$ as an exogenous parameter. If more than the maximum allowed nodes have a file available, only the ranked nodes with top-$\hat{n}$ smallest estimated transfer delays will be selected as provision nodes. In other words, when node r requests file k, the transfer delay between request node r and provision node i is Q_i, where $Q_i = D_i + T_i$. Let $S_0 = \{Q_i | x_{ik} = 1\}$ be the set of transfer delays of all available peer nodes and $|S_0|$ be the total number of available nodes. The optimal set of provision nodes suggested by the P2P provider is described as:

$$\Phi = \left\{ i | i \in I_1 \cup I_2 \cup \cdots \cup I_{\min(|S_0|, \hat{n})} \right\}, \tag{2}$$

where $I_k = \{i | Q_i = \inf S_{k-1}\}$, and $S_{k-1} = S_0 - \{Q_i | i \in I_1 \cup I_2 \cup \cdots \cup I_{k-1}\}$, for $2 \leq k \leq \hat{n}$.

As a result, if the total number of available nodes is fewer than or equal to $\hat{n}$, all of the available nodes will be chosen and we will have $|\Phi| = |S_0|$. But if there are more than $\hat{n}$ available nodes, we will have $|\Phi| = \hat{n}$. The request peer node is assumed to be rational and chooses the provision nodes with the highest service quality (fastest transfer speeds), according to the information provided by the P2P software. Note that the recommendation of provision nodes and upload activities may be executed automatically by the P2P system. The distribution of content availability is commonly known to all peer nodes. For the sake of analytical tractability, the value of variable T_i is assumed to be uniformly distributed on $[0, \tau]$, where parameter τ is the upper bound of propagation delay per byte. The value of τ has several implications for the network characteristics. Higher values of τ could be interpreted as larger position or uptime dispersion of peer nodes or higher traffic uncertainty levels in the public network domain. Other than making the analysis tractable, the uniform distribution favors all possible outcomes equally. We believe this is a reasonable assumption as most peers might not know much about the distribution of other peers and the parameters affecting their effective bandwidth.

3.2. Multilateral Contracting Model

We formulate this problem as a multilateral contracting model as follows: The principal (the P2P provider) hires many agents (provision nodes) to perform a task (file upload). In doing so, the principal observes performance (overall transfer delay $Q_i = D_i + T_i$) and not the effort exerted by agents (upload capacity b_i or equivalently, provision delay D_i). As effort is costly to agents, the principal has to compensate them for incurring this cost.

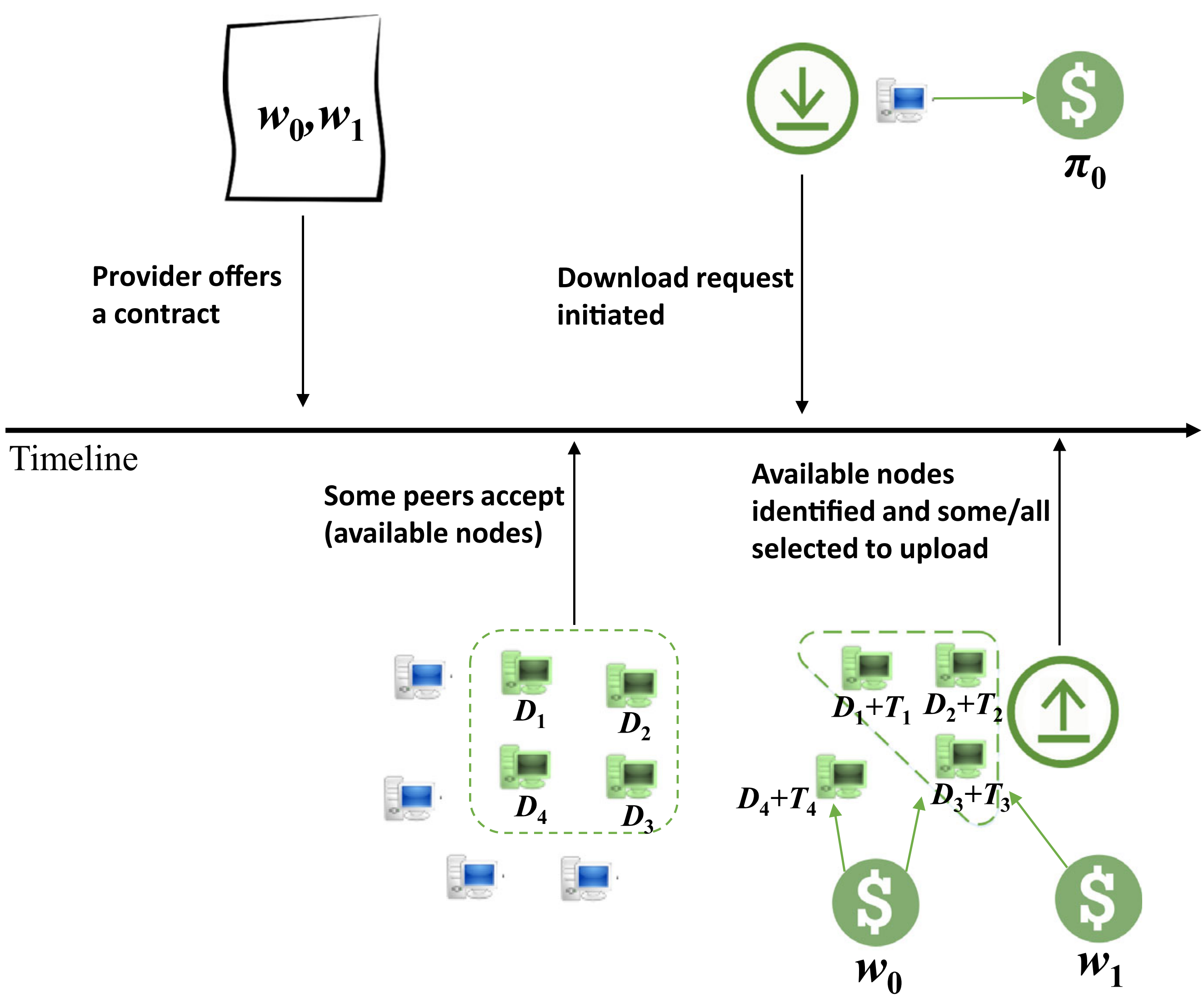
Service price and compensation: Assume the main revenue of P2P networks comes from usage-based file transfer service. The price of file transfer service is determined based on the value of the downloaded content and the corresponding transfer delay. Thus, in order to improve its revenue, the P2P provider has to offer the provision nodes sufficient rewards to induce higher effort by providing higher bandwidth capacity. Our proposed reward scheme includes two parts: a non-negative availability compensation, w_0, that is paid to all available peer nodes (the nodes that have the requested content), and a non-negative upload compensation, w_1, which is paid to those who take part in uploading the content. Whenever a download transaction is completed, all contracted available peer nodes will receive w_0. However, only provision nodes, chosen according to Equation (2) will upload the content and receive upload compensation, w_1. Therefore, peer nodes compete in providing bandwidth capacity to win the upload compensation. We assume that all participants are risk neutral and payments by guaranteeing parties are enforceable.

Stages of the contracting game: Figure 1 depicts the game stages and the decision variables of the players. In the first stage, the P2P provider makes take-it-or-leave-it offers to the peer nodes: a reward scheme W (w_0, w_1) for recruiting the provision nodes. In the second stage, P2P participants observe the offer, and decide whether to accept the offer or not. If they accept, they will have the content available, and decide on upload bandwidth capacity. In the third stage, a file download request is initiated, peers with the content are identified and a number of peers are selected to participate in uploading. Before download starts, the requesting peer pays the price for the file. In the fourth stage, the download is completed. All the peers who had the content available receive the availability reward, w_0, while the ones who contributed to uploading also receive the upload compensation, w_1.

4. Optimal Reward Scheme

In this section, we develop the optimal contract for a monopolistic P2P network, utilizing a principal-agent framework. If content uploading is compensated by a sufficiently high reward, the peer nodes become interested in winning the reward. So they engage in provision capacity competition in order to win the rewards by being selected as one of the provision nodes to upload the requested content. Therefore, in order to achieve the profit-maximization objective, the P2P provider can offer the appropriate reward scheme to induce P2P participants to exert the optimal level of effort. Using the backward induction approach, we first analyze the best response capacity provision

Figure 1 Graphical Representation of the Stages of the Contracting Game [Color figure can be viewed at wileyonlinelibrary.com]



strategy of the peer nodes, and then develop the optimal reward scheme.

It is worth mentioning that we are focusing on the peers who have the content available, and we are not modeling their decision to have the file or not. We don't model a peer's decision to acquire the content for two reasons: First, we are assuming that the P2P service has been working for some time and is at a stable state, so a reasonable number of peers who are uploaders have downloaded the file through the same P2P network. Secondly, this decision is easy to model and depends on the cost of acquiring and keeping the content and the expected demand for the content, so modeling it is not of much interest, while it makes the model more complicated. In this section, we first characterize the model for a general case where we have P classes of peers who are different with respect to their provisioning cost. Then we study the special case of homogeneous peers. The case in which we have two classes of peers will be discussed in the next section.

4.1. General Model

Let the optimal reward scheme be denoted by $W(w_0^*, w_1^*)$, and the provision delay choice of an arbitrary peer node j, by D_j. Delays of other nodes will be $D_i \ \forall i \neq j$. Notice that the decision variable of peer nodes is actually the upload bandwidth and provision delay is the inverse of the bandwidth. If the number of available nodes is fewer than or equal to $\hat{n}$, all available peer nodes will be provision nodes, regardless of their effort level.

Cost of capacity provision: The participation cost for a peer node is defined as the opportunity cost of allocating bandwidth capacity for P2P uploading service. Assuming that the participating peers are required to provide a bandwidth capacity which results in a delay no slower than a maximum, $D_{\max}$. The cost of provisioning at $D_{\max}$ is fixed and normalized to zero. We assume that the cost for providing a faster service, $D_j \leq D_{\max}$, is given by:

$$C(D_j) = \frac{1}{2}c_j(D_{\max} - D_j)^2 \tag{3}$$

where c_j is the cost parameter. In the general case, we assume that peers are heterogeneous in their costs. There are P classes of peers. P could be any number as long as it is countable, and the cost for a peer belonging to class p is c_p, where $p = 1, 2, \ldots, P$. We label the classes such that we have $c_p < c_{p+1}$ for $p = 1, \ldots, P - 1$. So a peer with c_1 has the lowest cost, and a peer with c_P the highest. Peers only know their own cost and the provider is not aware of any given peer's cost. However, they are all aware of the distribution of peer classes and the cost for each class. The distribution is Pr (peer belonging to class p) = m_p, such that $m_p > 0$ and $\Sigma_{p=1}^{P} m_p = 1$, for $p = 1, 2, \ldots, P$.

We will limit our analysis to the class of symmetric Nash equilibria. So in equilibrium, peers with the same provisioning cost will choose the same delay as their optimal decision in a way that no peer has an incentive to deviate. In other words, for the same content and given the same reward scheme, in equilibrium we will have P optimal delays distributed such that $D_p < D_{p+1}$ for $p = 1, 2, \ldots, P - 1$ and Pr $(D = D_p) = m_p$ for $p = 1, 2, \ldots, P$. So the peer class with provisioning cost c_p will choose their delay to be D_p in equilibrium.

Reward for uploading service: The expected overall compensation that an arbitrary peer node j carrying the file earns is given by:

$$W(D_j) = w_0^* + \Pr\{j \in \Phi | D_j\} w_1^*. \tag{4}$$

Here $\Pr\{j \in \Phi | D_j\}$ denotes the probability that peer node j who has the file will be selected as one of the provision nodes if she chooses provision delay D_j, and is explicitly written as:

$$\begin{aligned} \Pr\{j \in \Phi | D_j\} &= \sum_{k=0}^{\hat{n}-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \\ &+ \sum_{k=\hat{n}}^{N-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \\ &\Pr\{j \text{ among top-}\hat{n} | D_j\}. \end{aligned} \tag{5}$$

The first term on the right hand side of Equation (5) is the probability that fewer than $\hat{n}$ other peer nodes have the requested content. The second term describes the probability that peer node j's transfer speed is among the top-$\hat{n}$ fastest, given k other peer nodes ($k \geq \hat{n}$) also have the content. The probability of j being among the top-$\hat{n}$ fastest peers when an arbitrary set of k other peers also have the content can be expressed as follows:

$$\begin{aligned} \Pr\{j \text{ among top-}\hat{n} | D_j, T_j\} &= \sum_{u=0}^{\hat{n}-1} \binom{k}{u} \prod_{i=1}^{u} \Pr\{D_i + T_i \\ &< D_j + T_j\} \prod_{i=u+1}^{k} \Pr\{D_i + T_i > D_j + T_j\} \\ &= \sum_{u=0}^{\hat{n}-1} \binom{k}{u} \left(F(D_j + T_j)\right)^u \left(1 - F(D_j + T_j)\right)^{k-u}. \end{aligned} \tag{6}$$

Here, nodes $i = 1, \ldots, u$ form an arbitrary set of peer nodes which are faster than peer node j. j is among the top-$\hat{n}$ fastest as long as $u \leq \hat{n} - 1$. F is the cumulative distribution function (CDF) of convolution of provision delay, D_i, and propagation delay, T_i, for an arbitrary peer. If we plug (6) into (5), we will have:

$$\Pr\{j \in \Phi|D_j, T_j\} = \sum_{k=0}^{\hat{n}-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} + \sum_{k=\hat{n}}^{N-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \sum_{u=0}^{\hat{n}-1} \binom{k}{u} \left(F(D_j+T_j)\right)^u \left(1-F(D_j+T_j)\right)^{k-u} \tag{7}$$

As argued before, different classes of peers will select different provision delays in equilibrium. Since a peer's cost is a random variable and we are focusing on symmetric Nash equilibria, the provision delay of a peer in equilibrium will be a random variable with the same distribution as the cost. In other words, in equilibrium, peers choose their optimal delays according to the same rule d, that is, a peer with cost c_i chooses optimal delay $D_i^* = d(c_i)$. Since c is a discrete random variable, D will be a transformation of this random variable. Therefore, $D_i + T_i$ is the convolution of a continuous random variable and a discrete random variable, which is a step function. The expression in (7) is a function of T_j, the propagation delay of peer j. But the peer is not aware of the exact value of this delay and only knows the distribution of it. So we will have the following:

$$\Pr\{j \in \Phi|D_j\} = \int_0^{\tau} \Pr\{j \in \Phi|D_j, T_j\} f_T(T_j) dT_j, \tag{8}$$

f_T is the PDF of propagation delay and is assumed to be uniformly distributed. Thus, the probability of being selected as a provision node is determined by not only provision delays of all the participating peer nodes, but also the stochastic propagation delay incurred in the public network.

Optimal upload capacity: The tradeoffs a P2P participant faces are between the reward associated with winning the upload compensation and the corresponding capacity cost. Thus, the utility maximization problem of a typical P2P participant j is to select the optimal (maximum) provision delay D_j to achieve the maximum overall compensation. The objective function for participant j is formulated as follows:

$$\max_{D_j} U_j = W(D_j) - C(D_j) = w_0 + \Pr\{j \in \Phi|D_j\} w_1 - \frac{1}{2} c_j \left(D_{\max} - D_j\right)^2. \tag{9}$$

The solution to this optimization problem is presented in the following proposition.

Proposition 1. *If propagation delay T_j is uniformly distributed in $[0, \tau]$, the optimal delay selection D_j^*, $\forall j$ satisfies:*

$$\frac{w_1}{\tau} \sum_{k=\hat{n}}^{N-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \sum_{u=0}^{\hat{n}-1} \binom{k}{ur} \left(\left(F\left(D_j^*+\tau\right)\right)^u \left(1-F\left(D_j^*+\tau\right)\right)^{k-u} - \left(F\left(D_j^*\right)\right)^u \left(1-F\left(D_j^*\right)\right)^{k-u} \right) + c_j \left(D_{\max} - D_j\right) = 0, \tag{10}$$

where the CDFs are:

$$F\left(D_j^*+\tau\right) = \frac{1}{\tau} \sum_{p=1}^{P} m_p \left((\tau)\theta\left(D_j^* - D_p^*\right) + \left(D_j^* + \tau - D_p^*\right) \theta\left(D_j^* + \tau - D_p^*\right) \theta\left(D_p^* - D_j^*\right) \right) \tag{11}$$

$$F\left(D_j^*\right) = \frac{1}{\tau} \sum_{p=1}^{P} m_p \left((\tau)\theta(D_j^* - D_p^* - \tau) + (D_j^* - D_p^*)\theta(D_j^* - D_p^*)\theta(D_p^* + \tau - D_j^*) \right), \tag{12}$$

and $\theta(\cdot)$ is the step function.

4.2. Symmetric Model with Homogeneous Peers

We will first study the problem for the case in which the peers are homogenous, meaning that we have a single class of peers and the provisioning cost is the same for everyone.

Lemma 1. *The optimal delay selection is:*

$$D^*(w_1) = D_{\max} - \frac{\eta}{c\tau} w_1, \tag{13}$$

where $\eta = \sum_{k=\hat{n}}^{N-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k}$ is the probability that at least $\hat{n}$ other peers have the content available.

It is indicated in Equation (13) that a peer's optimal delay selection is a decreasing function of w_1. After calculating the best response provision delay function $D^*(w_1^*)$ in the second stage, the P2P provider chooses the optimal compensations, w_0^* and w_1^*, in the first stage to maximize its profit.

Optimal reward scheme: Assume that the value of a downloaded content is v. The highest price a peer node is willing to pay is $\tilde{\pi}_0 = v - \gamma(D+T)$, where γ represents the sensitivity to delay.

Therefore, the expected profit for the P2P provider is:

$$\pi = H\pi_0 - N\beta w_0 - \kappa w_1, \tag{14}$$

where $\pi_0 = E(\tilde{\pi}_0) = v - \gamma(D + \tau/2)$, and $H = 1-(1-\beta)^N$ denotes the probability that a requested file is available in the network and $\kappa = \sum_{k=1}^{N} \binom{N}{k} \beta^k (1-\beta)^{N-k} \min(k, \hat{n})$ is the expected number of nodes selected to jointly upload the file. $N\beta$ is the expected number of peers who have the content available. The P2P provider chooses w_0 and w_1 to maximize the expected profit, subject to the constraint that a peer receives a utility no less than U_0, her reservation utility. This can be interpreted as the utility a peer node will achieve if she chooses to do other activities instead of participating in content provisioning. For convenience, we assume $U_0 = 0$.

LEMMA 2. *The equilibrium expected probability of a peer being selected as a provision node is:*

$$\Pr\{j \in \Phi | D_j\} = \frac{\kappa}{N\beta}. \tag{15}$$

Equation (15) gives an intuitive result that in equilibrium, the probability of an available peer node (one who has the content) being selected as a provision node is the ratio of the expected number of provision nodes to the expected number of available nodes. Hence, the expected utility of a peer node is a follows:

$$\begin{aligned} U &= w_0 + \frac{\kappa}{N\beta} w_1 - \frac{1}{2} c (D_{\max} - D^*)^2 \\ &= w_0 + \frac{\kappa}{N\beta} w_1 - \frac{\eta^2}{2c\tau^2} w_1^2, \end{aligned} \tag{16}$$

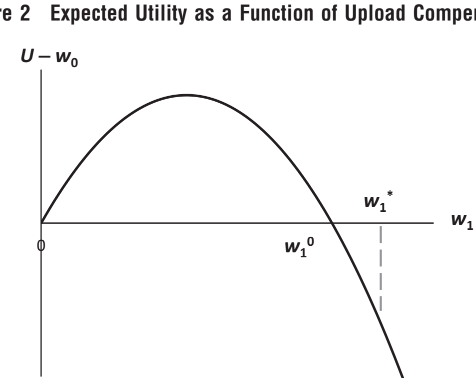
which should be non-negative. Figure 2 plots $U - w_0$ as a function of upload compensation, w_1.

The expected utility increases with w_1 initially as the benefit from upload compensation outweighs the cost of providing the bandwidth. But as w_1 increases, the competition becomes fiercer and the cost of the bandwidth grows more significantly compared to the benefit from provisioning. $U - w_0$ reaches zero at w_1^0, where

$$w_1^0 = \frac{2c\tau^2\kappa}{N\beta\eta^2}, \tag{17}$$

and then becomes negative.

Figure 2 Expected Utility as a Function of Upload Compensation



PROPOSITION 2. *The optimal reward scheme (availability compensation and upload compensation), and delay are given by the following:*

$$i.\ \text{if } \kappa > \frac{\gamma H\eta}{c\tau},\ W(w_0^*, w_1^*) = (0, 0),\ D^* = D_{\max}; \tag{18}$$

$$\begin{aligned} &ii.\ \text{if } \frac{\gamma H\eta}{2c\tau} < \kappa \le \frac{\gamma H\eta}{c\tau},\ W(w_0^*, w_1^*) = (0, w_1^0), \\ &\quad D^* = D_{\max} - \frac{2\tau\kappa}{N\beta\eta}; \end{aligned} \tag{19}$$

$$\begin{aligned} &iii.\ \text{if } \kappa \le \frac{\gamma H\eta}{2c\tau},\ W(w_0^*, w_1^*) = \left(\frac{\gamma^2H^2}{2N^2\beta^2c} - \frac{\gamma H\tau\kappa}{N^2\beta^2\eta}, \frac{\gamma\tau H}{N\beta\eta}\right), \\ &\quad D^* = D_{\max} - \frac{\gamma H}{N\beta c}. \end{aligned} \tag{20}$$

The expected profit expressed in (14) indicates that the P2P provider should decrease availability compensation w_0 to zero, whenever possible. While paying out κw_1 to peers, the P2P provider gains $(\gamma H\eta/c\tau)w_1$ which is a result of a reduction in propagation delay. If the latter is smaller than the former, that is Case *i* in Proposition 2, there is no incentive for the P2P provider to provide any compensation. This happens when the number of provision nodes is relatively large. If the relative expected number of provision nodes decreases to the range specified in case *ii*, the provider should increase w_1 to gain more net benefit. For this case, $w_0 = 0$ and w_1 must be increased to w_1^0 for the peers' expected utility to be non-negative. If, $\kappa \le \gamma H\eta/2c\tau$, w_1 must be further increased. However, the provider also needs to provide a positive availability reward, that is $w_0 > 0$, so that the participation constraints for the peers is satisfied ($U = 0$).

The effects of system parameters on the reward scheme and capacity provision are presented in Table 2 and summarized as follows. Upload compensation, w_1^*, increases with the maximum number of peers allowed to upload jointly, $\hat{n}$, the dispersion of peer positions (propagation delay uncertainty), τ, and could increase with disutility of delay, γ, and provision cost, c, but decreases with network size, N, and content availability β. Provision delay of a peer node, D^*, increases with the network size and content

Table 2 Impact of System Parameters on Reward Scheme in Cases *ii* and *iii*

	N	$\hat{n}$	τ		c	γ
w_1^*	↓	↑	↑	↓	*ii*: ↑, *iii*: –	*ii*: –, *iii*: ↑
D^*	↑	*ii*: ↓, *iii*: –	*ii*: ↓, *iii*: –	↑	*ii*: –, *iii*: ↑	*ii*: –, *iii*: ↓

↑Increase; ↓Decrease; –No effect.

availability, and could increase with provision cost, while it decreases or remains constant with the maximum number of peers allowed to upload jointly, propagation delay uncertainty and disutility of delay.

In what follows, we will study these effects in more depth and draw insights.

Impact of network size: There exist both negative and positive scale effects. Competition intensifies as network size increases and more peers compete for the reward. This according to Equation (13) would induce the peers to choose a lower delay for a constant w_1 (η is increasing in N). However, based on Equations (19) and (20) in cases *ii* and *iii*, upload reward (w_1^*) decreases with N, and so does the chance to win the reward. This effect is stronger, consequently peer nodes exert less effort in capacity provision. So increasing competition among peers has a negative impact on the resulting provision delay. On the other hand, P2P network can benefit from a larger number of provision nodes due to the reduction of propagation delay. This result is in line with the empirical findings of Asvanund et al. (2004). Although their study focuses on existing non-commercial P2P networks, they also find that both positive and negative externalities exist in P2P networks.

Impact of file swarming: Advancement in file swarming technology allows a larger number of provision nodes to upload simultaneously. In cases *ii* and *iii* where $\hat{n}$ is not too high (κ is increasing in $\hat{n}$ while η is decreasing in $\hat{n}$), as $\hat{n}$ increases, competition among peer nodes decreases, which will enforce the P2P provider to increase the upload compensation. This increase, in case *iii*, offsets the reduced competition effect, so optimal delay remains constant. In case *ii*, the rate of increase in w_1^* is even faster, which induces the peers to lower their optimal delay.

Impact of public network: In cases *ii* and *iii* where uncertainty is not too high ($\tau \leq \gamma H\eta/c\kappa$), as the uncertainty increases, the impact of information asymmetry increases which induces the peers to raise their optimal delay, and to counter this effect, the provider has to increase the upload reward. In case *iii* where uncertainty is lowest $\tau \leq \gamma H\eta/2c\kappa$, this increase offsets the information asymmetry effect and optimal delay remains constant. For medium levels of uncertainty in case *ii* ($\gamma H\eta/2c\kappa < \tau \leq \gamma H\eta/c\kappa$), the provider responds to increases in uncertainty by increasing the provision reward such that this increase more than compensates for the information asymmetry effect and the peers start lowering their optimal delays. However, when uncertainty is too high as in case *i* ($\tau > \gamma H\eta/c\kappa$) it is no longer profitable for the provider to offer rewards.

Impact of content availability: In case *i*, when availability is too low (both H and η are increasing in β), content availability has no impact. But for medium and high levels of content availability in cases *ii* and *iii*, as β goes up, the probability that a request for content is satisfied increases. Hence, more peers have the content and the competition for being selected as provision nodes intensifies, so the P2P provider does not need to incentivize peers to participate as aggressively, and lowers the upload reward. The effect of lowering the reward is stronger than the competition effect, so individual peers increase their optimal delay. In other words, when content availability is low, the provider will offer higher rewards to induce peers to set a higher bandwidth.

Impact of provision cost: When the cost is low (case *iii*, $c \leq \gamma H\eta/2\tau\kappa$), optimal delay intuitively increases with the cost, but upload reward is unaffected. When $\gamma H\eta/2\tau\kappa < c \leq \gamma H\eta/\tau\kappa$ (medium levels of cost in case *ii*), provider starts compensating the peers when provision cost increases and it offsets the effect of increasing cost, so optimal delay remains constant. When the cost is too high, ($c > \gamma H\eta/\tau\kappa$ in case *i*), it is not profitable for the provider to try to incentivize the peers, so upload reward is set to 0 and optimal delay is $D_{\max}$.

Impact of delay disutility: If disutility of delay is small enough as in cases *i* and *ii*, it has no impact on upload reward and delay. However, in case *iii*, as the downloading peers' disutility of delay increases past $2c\tau\kappa/H\eta$, the P2P provider increases the upload reward to induce peers to choose a higher bandwidth and the peers will respond by choosing a lower provision delay.

5. Heterogeneous Peers

Now we look at the problem for the case in which we have two types of peers: low-cost peers and high-cost peers. The bandwidth costs for the two classes are, respectively, c_1 and c_2 such that $c_1 < c_2$. The equations to solve for optimal delay selections D_1^* and D_2^* can be derived from Proposition 1 and are presented explicitly in the Appendix A. However, the closed form solutions for a general setting cannot be readily obtained, therefore we focus our attention on the case where classes are separable, that is, $D_2^* - D_1^* \geq \tau$. This will not only assure analytical tractability but also provide insights for the scenarios which are quite different from the case of homogeneous peers (if the delays are not separable, it means the difference between the two classes is not as significant as the case where they are separable).

PROPOSITION 3. *The optimal delay selections are:*

$$D_1^* = D_{\max} - \frac{\eta_1}{c_1\tau}w_1,\ D_2^* = D_{\max} - \frac{\eta_2}{c_2\tau}w_1, \tag{21}$$

where $D_2^* - D_1^* \geq \tau$ *is satisfied, and*

$$\eta_1 = \sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\sum_{u=\hat{n}}^{k}\binom{k}{u}{m_1}^u{m_2}^{k-u},\ \text{and} \tag{22}$$

$$\begin{aligned}\eta_2 &= \eta - \eta_1\\ &= \sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\sum_{u=0}^{\hat{n}-1}\binom{k}{u}{m_1}^u{m_2}^{k-u}\end{aligned} \tag{23}$$

Similar to η in the case of homogenous peers, η_1 is interpreted as the probability that at least $\hat{n}$ other peers have the content available and among them at least $\hat{n}$ are class 1 peers. This is the probability for the situation where not only class 1 peers have to compete to be selected, but they have to compete with other class 1 peers (not all available class 1 peers are selected). η_2 is the probability to have competition with at least $\hat{n}$ other available peers, but class 1 peers in this case are automatically selected, since there is fewer than $\hat{n}+1$ class 1 peers and they have smaller delay than class 2 peers. In order for the condition $D_2^* - D_1^* \geq \tau$ to be satisfied, we require that $\eta_1/c_1 > \eta_2/c_2$, and the difference between η_1/c_1 and η_2/c_2, or w_1, should be large enough. In other words $\eta_1 > \eta_2 c_1/c_2$ (η_1 should not be significantly smaller than η_2). Looking at Equations (13) and (21), we can see that compared to the homogeneous peers model, since $\eta_1 < \eta$, in the presence of class 2 peers, class 1 peers select a worse quality of service (higher delay).

To obtain the optimal availability and upload compensations, we first write the expected profit for the P2P provider:

$$\pi = H\Big(v - \gamma\Big(\frac{\kappa_1}{\kappa}D_1^* + \frac{\kappa_2}{\kappa}D_2^* + \frac{\tau}{2}\Big)\Big) - N\beta w_0 - \kappa w_1. \tag{24}$$

Similar to the case of homogeneous peers, κ is the expected number of nodes selected, and among them κ_1 are of class 1 while κ_2 belong to class 2. The explicit expressions are as follows:

$$\kappa_1 = \sum_{k=1}^{N}\binom{N}{k}\beta^k(1-\beta)^{N-k}\sum_{u=0}^{k}\binom{k}{u}{m_1}^u{m_2}^{k-u}\min(u,\hat{n}) \tag{25}$$

and $\kappa_2 = \kappa - \kappa_1$. Note that since the separability constraint holds, that is, $D_2^* - D_1^* \geq \tau$, a class 2 peer cannot be selected unless there are fewer than $\hat{n}$ class 1 peers that have the content.

Similarly as in Lemma 2, it can be shown that in equilibrium, the expected probability of a class 1 (class 2) peer, which has the content, being selected as an uploading node is $\kappa_1/m_1N\beta$ ($\kappa_2/m_2N\beta$). This allows us to write the expected utility for the two classes, explicitly:

$$\begin{aligned}U_1 &= w_0 + \frac{\kappa_1}{m_1N\beta}w_1 - \frac{1}{2}c_1\left(D_{\max} - D_1^*\right)^2\\ &= w_0 + \frac{\kappa_1}{m_1N\beta}w_1 - \frac{{\eta_1}^2}{2c_1\tau^2}{w_1}^2;\end{aligned} \tag{26}$$

$$\begin{aligned}U_2 &= w_0 + \frac{\kappa_2}{m_2N\beta}w_1 - \frac{1}{2}c_2\left(D_{\max} - D_2^*\right)^2\\ &= w_0 + \frac{\kappa_2}{m_2N\beta}w_1 - \frac{{\eta_2}^2}{2c_2\tau^2}w_1^2.\end{aligned} \tag{27}$$

Before we proceed to the P2P provider's optimization problem, we state:

LEMMA 3. *The following inequality holds:*

$$\frac{\kappa_1}{m_1} > \frac{\kappa_2}{m_2}. \tag{28}$$

This lemma indicates, intuitively, that there is a higher probability for a class 1 peer to be selected. With the help of this lemma, we need to consider three possible cases which are depicted in Figure 3. In Case A, $U_1 - w_0$ becomes negative first and afterwards is always below $U_2 - w_0$. In Case B, $U_1 - w_0 > U_2 - w_0$. In Case C, $U_2 - w_0$ becomes negative first, but will be higher than $U_1 - w_0$ after $w_1 > w_1^c$, where

$$w_1^c = \frac{2\tau^2}{N\beta}\left(\frac{\kappa_1}{m_1} - \frac{\kappa_2}{m_2}\right)\left(\frac{\eta_1^2}{c_1} - \frac{\eta_2^2}{c_2}\right)^{-1}. \tag{29}$$

The P2P provider's expected profit function can be rewritten as:

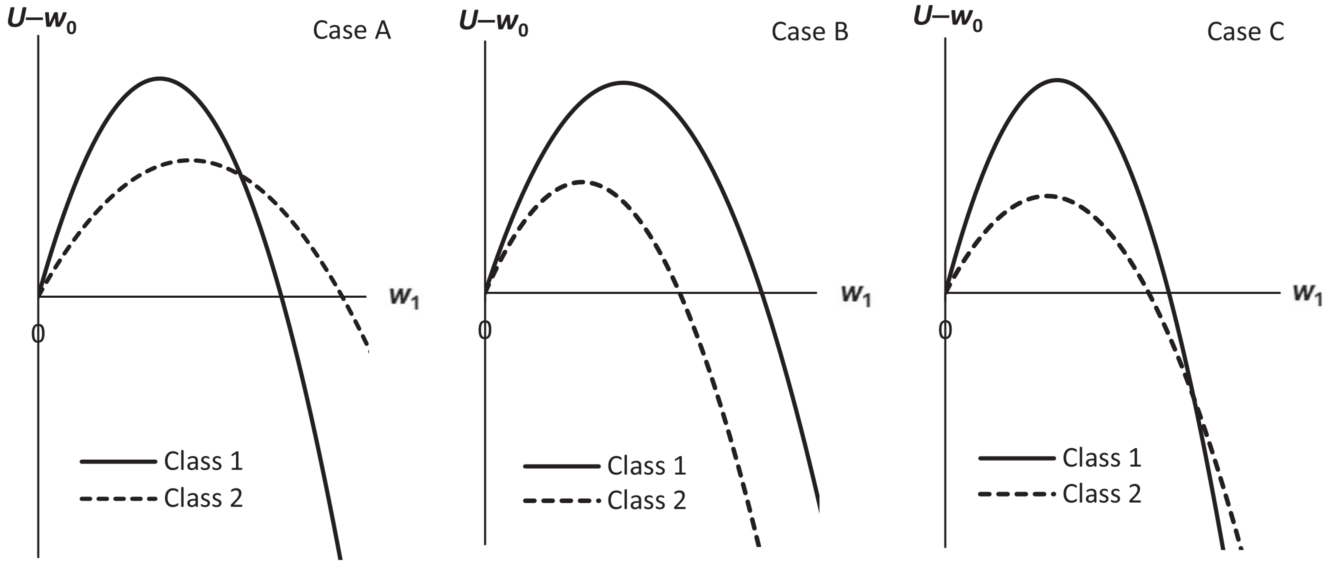
$$\begin{aligned}\pi = {} & H\Big(v - \gamma\Big(D_{\max} + \frac{\tau}{2}\Big)\Big) + \frac{\gamma H}{\kappa\tau}\left(\frac{\kappa_1\eta_1}{c_1} + \frac{\kappa_2\eta_2}{c_2}\right)w_1\\ & - N\beta w_0 - \kappa w_1.\end{aligned} \tag{30}$$

In the following, we only consider the case where the benefit from delay reduction outweighs the upload compensation, explicitly:

$$\Gamma \equiv \frac{\gamma H}{\kappa\tau}\left(\frac{\kappa_1\eta_1}{c_1} + \frac{\kappa_2\eta_2}{c_2}\right) - \kappa > 0; \tag{31}$$

otherwise the optimal solution is $W(w_0^*, w_1^*) = (0, 0)$ which violates the separability constraint. We summarize the solutions in the following proposition.

Figure 3 Three Possible Scenarios for Expected Utilities of the Two Classes



PROPOSITION 4. *The optimal compensation scheme is given by:*

i. *Case A where* $\frac{\kappa_1 c_1}{m_1 \eta_1^2} < \frac{\kappa_2 c_2}{m_2 \eta_2^2}$ *(which according to Lemma 3 implies* $\frac{\eta_1^2}{c_1} > \frac{\eta_2^2}{c_2}$*):*

$$W(w_0^*, w_1^*) = \begin{cases} \left(0, w_{1,A}^0\right), & 0 \le \Gamma < \frac{\kappa_1}{m_1}; \\ \left(w_{0,A}^*, w_{1,A}^*\right), & \frac{\kappa_1}{m_1} \le \Gamma, \end{cases} \tag{32}$$

where $w_{1,A}^0 = \frac{2c_1\tau^2\kappa_1}{m_1 N\beta\eta_1^2}$, $w_{1,A}^* = \frac{c_1\tau^2}{N\beta\eta_1^2}\left(\Gamma + \frac{\kappa_1}{m_1}\right)$, *and* $w_{0,A}^* = \frac{w_{1,A}^*}{2N\beta}\left(\Gamma - \frac{\kappa_1}{m_1}\right)$;

ii. *Case B where* $\frac{\kappa_1 c_1}{m_1 \eta_1^2} > \frac{\kappa_2 c_2}{m_2 \eta_2^2}$ *and* $\frac{\eta_1^2}{c_1} < \frac{\eta_2^2}{c_2}$*:*

$$W(w_0^*, w_1^*) = \begin{cases} \left(0, w_{1,B}^0\right), & 0 \le \Gamma < \frac{\kappa_2}{m_2}; \\ \left(w_{0,B}^*, w_{1,B}^*\right), & \frac{\kappa_2}{m_2} \le \Gamma, \end{cases} \tag{33}$$

where $w_{1,B}^0 = \frac{2c_2\tau^2\kappa_2}{m_2 N\beta\eta_2^2}$, $w_{1,B}^* = \frac{c_2\tau^2}{N\beta\eta_2^2}\left(\Gamma + \frac{\kappa_2}{m_2}\right)$, *and* $w_{0,B}^* = \frac{w_{1,B}^*}{2N\beta}\left(\Gamma - \frac{\kappa_2}{m_2}\right)$;

iii. *Case C where* $\frac{\kappa_1 c_1}{m_1 \eta_1^2} > \frac{\kappa_2 c_2}{m_2 \eta_2^2}$ *and* $\frac{\eta_1^2}{c_1} > \frac{\eta_2^2}{c_2}$*:*

$$W(w_0^*, w_1^*) = \begin{cases} \left(0, w_{1,B}^0\right), & 0 \le \Gamma < \frac{\kappa_2}{m_2}; \\ \left(w_{0,B}^*, w_{1,B}^*\right), & \frac{\kappa_2}{m_2} \le \Gamma < \Gamma_2; \\ \left(w_0^c, w_1^c\right), & \Gamma_2 \le \Gamma < \Gamma_1; \\ \left(w_{0,A}^*, w_{1,A}^*\right), & \Gamma_1 \le \Gamma, \end{cases} \tag{34}$$

where $w_0^c = \frac{w_1^c}{N\beta}\left(\frac{\eta_1^2}{c_1} - \frac{\eta_2^2}{c_2}\right)^{-1}\left(\frac{\kappa_1}{m_1}\frac{\eta_2^2}{c_2} - \frac{\kappa_2}{m_2}\frac{\eta_1^2}{c_1}\right)$ *and the two thresholds are:*

$$\Gamma_2 = 2\left(\frac{\kappa_1}{m_1} - \frac{\kappa_2}{m_2}\right)\left(\frac{\eta_1^2}{c_1} - \frac{\eta_2^2}{c_2}\right)^{-1}\frac{\eta_2^2}{c_2} - \frac{\kappa_2}{m_2}, \text{ and}$$

$$\Gamma_1 = 2\left(\frac{\kappa_1}{m_1} - \frac{\kappa_2}{m_2}\right)\left(\frac{\eta_1^2}{c_1} - \frac{\eta_2^2}{c_2}\right)^{-1}\frac{\eta_1^2}{c_1} - \frac{\kappa_1}{m_1}.$$

The solutions in Proposition 4 seem involved, but the intuition behind them is quite simple and surprisingly similar to what we observed for the case of homogeneous peers in section 4.2. First let us consider the situation where it is not profitable for the platform owner to offer compensation to peers for uploading, i.e., the case where the commercial P2P platform is not viable. This happens when we have $\Gamma < 0$ or in other words, $\kappa > \frac{\gamma H}{\kappa\tau}\left(\frac{\kappa_1\eta_1}{c_1} + \frac{\kappa_2\eta_2}{c_2}\right)$. So this happens when the number of expected provision nodes is relatively large. Under these circumstances, free P2P file sharing succeeds and there is no need for incentivizing peers. We see this phenomenon in the real world where very popular content is efficiently shared in free P2P platforms. In the homogeneous peers case, this happened when we had $\kappa > \gamma H\eta/c\tau$, so the structure of the conditions is quite similar, except in this case, we have to account for the fact that there are two classes of peers.

Now let's look at the three different cases, A, B and C, where it is profitable to have a commercial P2P network. Again, the intuitive explanation of the results is much simpler than the equations in Proposition 4. Therefore, here we focus on Figure 3 instead of the equations, and compare the findings here with the findings for the homogeneous-peers model in Figure 2. Case A in Figure 3 and Proposition 4 represents a situation where there is a relatively fierce competition between class 1 peers. This could happen when there is a relatively large number of class 1 peers who have the content available. In this case, the platform owner takes advantage of the competition between class 1 peers by setting their expected utility to zero, while class 2 peers gain positive expected utility.

In Case B, the situation is reverse. As there are not as many class 1 peers present compared to class 2 peers, the platform owner is forced to give them positive expected rewards. On the other hand, all the expected value of class 2 peers is extracted in this case. Cases A and B are each similar to cases *i* and *ii* in Proposition 2, and the optimal solution in each case is determined by the relative number of available peers from the more dominant class. One difference between the heterogeneous and homogeneous peer models is that in both cases A and B, the platform owner has to give one of the classes positive expected payoffs, so the expected profits for the platform owner is lower when peers are heterogeneous.

Case C in Figure 3 is a middle ground between cases A and B, as the competition between class 1 peers is not as fierce as in Case A. However, if the upload reward is sufficiently large, it will be enough to instigate a fierce competition between class 1 peers, so the situation could become similar to case A. In such a situation, the platform owner compares different scenarios to choose the optimal reward structure, and the result might be completely extracted expected value for either or both classes of peers.

Although the structure of the optimal solution for the heterogeneous model is more involved than the homogeneous model, as we can see in Table 3, the impact of parameters on the optimal upload compensation is quite similar. Therefore, we do not provide a detailed discussion here as we did in section 4.2 and instead refer the readers to section 4.2. Instead we focus on the optimal upload rewards under different cases in Table 3 and compare the findings to the findings in the homogeneous-peers model. As we see in Equation (21), the structure of the optimal delay for both classes is similar to that of homogeneous peers

Table 3 Impact of System Parameters on Reward Scheme with Two Classes of Peers

	N	$\hat{n}$	τ	β	c_1	c_2	γ
$W^0_{1,A}$	↓	↑	↑	↓	↑	–	–
$W^*_{1,A}$	↓	↑	↑	↓	↑	↓	↑
$W^0_{1,B}$	↓	↑	↑	↓	–	↑	–
$W^*_{1,B}$	↓	↑	↑	↓	↓	↑	↑
W^c_1	↓	↑	↑	↓	↑	↓	–

↑Increase; ↓Decrease; –No effect.

in Equation (13), so the impact of parameters on optimal delay will also be very similar to those for homogeneous peers.

Despite having different optimal upload rewards depending on the relative size of parameters, the impact of network size, propagation delay uncertainty, content availability, and maximum number of peers allowed to upload simultaneously are exactly similar to the homogeneous-peers model for all five possible values of rewards in Table 3. The impact of delay disutility is also similar, as it either has a positive impact under some cases, or no impact. The effect of the cost parameters seems to be different from the homogeneous-peers model. However, upon closer examination, we see that the effect is still the same. Only that in this case, the relative size of the costs as well as each cost parameter matter. More importantly, it matters which class of peers is the one having a more fierce competition and therefore having all their value extracted. In case A and sometimes in case C, an increase in bandwidth cost of class 1 peers results in an increase in upload reward. On the other hand, in case B and under certain circumstances in case C, when class 2 peers are the ones in a more fierce competition, an increase in class 2 peers' bandwidth cost results in an increase in upload reward. So again, the result is similar to the homogeneous model, as the effect of an increase in bandwidth cost of the dominant class is the same.

6. Conclusion and Discussion

Utilizing the principal-agent model, this study presents a contracting scheme for P2P file-sharing networks which allows a profit-seeking P2P provider to induce peer nodes to provide appropriate upload capacity. In this proposed reward scheme, each contracted peer node is awarded an availability compensation, and receives an upload compensation after the upload is complete. The interactions among technological issues, economic issues, and Internet environment in a typical P2P network are investigated. In addition, we propose a price schedule which is determined based on the performance of the P2P download service.

Compared to prior studies, the contract scheme we develop complies with P2P technological protocols which allow file swarming—simultaneous download from various provision nodes, and recommends provision nodes according to transfer delay performance in which available peer nodes are ranked according to their estimated transfer speed. The contract between the provider and an uploading peer is effectively initiated as the downloading peer agrees to make a payment. At the end of the download, all available peers receive the availability reward while the peers who were selected to participate in upload also receive the upload reward. Therefore, every instance of file download is associated with a set of contracts between the provider and the available peers. We are implicitly assuming that if a peer goes offline or removes the content in the midst of a download, they will not receive the upload reward. This is a reasonable simplifying assumption as with today's fast internet connections, even large video files could be downloaded in a few hours. For example, assuming a 10 Mbps download speed, a video file as large as 5 GB could be downloaded in about an hour. An extension of this research can study more complicated scenarios where users get partial rewards for partially contributing.

In sections 3 and 4, we discussed the context and our assumptions, and characterized the model for the generic case with heterogeneous peers and solved the model and derived the closed form solution for the case of homogeneous peers. We then offered comparative statics and discussed what is driving the effects. In section 5, we extended our analysis to the case with two heterogeneous types of peers and through the analysis and comparative statics, we showed that the results are quite consistent with those for homogeneous peers when the classes are separable. Here, we will summarize our discussions in sections 4 and 5 and further discuss the managerial implications of our results.

In Table 2, we focused on the impact of changes in important peer-level and environmental factors on upload compensation and the optimal delay chosen by individual peers when upload reward is offered by the P2P provider (cases ii and iii). For example, we saw that the reward offered by the provider to peers to upload niche content is larger than that offered for new and popular content. We also saw that for P2P networks with a smaller user base, the provider has to pay larger upload rewards to incentivize the peers. As the network grows, the provider will be able to lower the rewards. This indicates increasing returns to scale for the provider in this context which could serve as an additional motivation for a commercial P2P provider to invest in growth. We also considered the impact of uncertainty and observed that in the presence of more uncertainty, the provider has to pay a larger reward to peers which is partly information rent.

We are assuming that the P2P platform has been in operation for a while and has reached a relatively stable state. Therefore, arguably there should not be sudden major changes in the characteristics of the environment. Under such circumstances, the platform owner is able to react to the changes in the market by updating the contracts offered to peers. The optimal contracts derived in the paper allow the platform owner to easily adjust the values of prices and rewards as a function of changes in various environmental factors such as technological constraints or the size of the network.

Another important way we can look at the problem is by investigating when it is profitable for the P2P provider to offer rewards and when the provider prefers to set the rewards to zero. This has important practical implications, as it essentially offers guidelines about the circumstances where commercial P2P networks are effective. In order to do so, we focus on Proposition 2 and the conditions under which each case is optimal. What determines the solution in equilibrium is the marginal profit for offering rewards for the provider which is calculated as $\gamma H\eta/c\tau - \kappa$. So the important factor is the size of the expected number of uploading peers relative to $\gamma H\eta/c\tau$ which represents the marginal revenue. If the expected number of peer nodes selected for the contract is relatively too large, it is not profitable for the P2P provider to offer rewards to the peers. Focusing on this distinction, we can draw the following insights.

One of the circumstances where this happens is when availability is too low. So our proposed commercial P2P platform does not support content that is either extremely old or too obscure. The other important factor is the relative size of disutility of delay to the cost of capacity provision. If this ratio is too small, it is possible that the operation of the commercial P2P network is not profitable for any type of content, regardless of popularity. This happens when the users have extremely low opportunity costs of time, so they do not mind very long download delays, or when it is prohibitively expensive for users to provide upload bandwidth. The latter case is increasingly less likely as available internet bandwidth increases and bandwidth costs go down. It also indicates that the platform is more effective in areas where better and cheaper internet connections are available. Finally, when uncertainty is too large, we might again have a situation where the commercial P2P provider is not profitable. Uncertainty here is represented by the upper bound of propagation delay which in our broad definition is affected by geographic dispersion and network quality. So when network connections are too unreliable or unpredictable, a commercial P2P content sharing network might not be viable. On the other hand, the ongoing improvements in infrastructure will help the rise of commercial P2P networks.

The flip side of the above discussion is when the expected number of peer nodes selected is small enough in relative terms, which encompasses the circumstances under which a commercial P2P content sharing platform is viable. Therefore a profitable commercial P2P provider supports content files which are not too scarce, and has a user base which is sufficiently sensitive to download delays. In relative terms, for a sufficiently small expected number of uploading peers, it is not only profitable to offer upload reward, the provider also offers a fixed availability reward to whoever carries the content and enters into the contract. This will be the case if customers are quite averse to download delays, or when capacity provision is inexpensive or delay uncertainty is small.

Given the small size of music files, our proposed mechanism might be more effective for movies. However, there has been real-world efforts to implement commercial P2P music sharing. For example, Grooveshark compensated users for distributing music, such that, whenever a user purchased a track, a small payment (25 cents) was made to the user (Bryce 2007). Because of the small size of music files, a very small number of uploaders (even one) is enough to achieve a high QoS. Accounting for such differences, our model can be adjusted and applied to both video and music sharing.

We proposed a two-part contracting mechanism for the operation of a commercial P2P content sharing network. Although this is not the only mechanism possible, we argue that for this context, it is superior to many other common mechanisms. Our mechanism has three clear advantages compared to the free mechanism. It eliminates piracy of copy-righted content. It incentivizes peers to participate in the upload of less popular and niche content, and it eliminates free riding. Some P2P platforms based on the popular protocol, BitTorrent, offer non-monetary incentives to peers to participate, e.g., they give priority in downloading to nodes which contribute to uploading. Another way to implement our proposed contract is to give participating peers credit points that could be used to download content. These points work like money and can ideally be bought or redeemed for money. So this implementation is similar to our proposed model, as long as the value of the points is exogenous and doesn't depend upon the dynamics of the system. However, in our proposed model as opposed to BitTorrent, money is transferred to the original content owner. Moreover, our contract encourages participation in sharing less popular content.

Another alternative mechanism is a one-part flat reward mechanism where peers who contribute to the upload receive a flat reward regardless of content and environmental characteristics. Since the provider pays the same reward for all content, he faces a tradeoff between setting a high reward to include more content and excluding some content. We performed a numerical analysis to compare our model to the one-part flat reward model, and even for the simple case of only two content files where everything except availability of the files is the same, the increase in profits from switching to our proposed model is quite significant. Our model also has clear advantages for P2P content sharing over the subscription-based model which is used by companies like Netflix. Although the subscription-based model is simple and straightforward, it has three problems for our setting. First, this model is appropriate for streaming and not downloading content. Second, if the reward is also paid as a fixed monthly reward, similar to the case for flat rewards, it will not incentivize the peers to participate more in uploading less popular content. Finally, having a flat monthly compensation is not practical because of moral hazard and monitoring problems.

Our study shares similarities with traditional contest models. However, at the operational level where we derive managerial implications, the results from the contest literature are not readily applicable. Specifically, compared to existing works in contest literature (e.g., Kalra and Shi 2001, Lazear and Rosen 1981, Nalebuff and Stiglitz 1983), our study has several significant and unique contributions. First, in existing contest studies, the uncertainty factor (or noise) plays a small role in affecting the long-run output as both positive and negative noise will be balanced out to be zero and the number of contesters will not affect the impact of uncertainty factor on the long-run performance. However, in the context of P2P networks, the uncertainty factor (propagation delay) always affects the overall expected content transfer performance. Furthermore, there exists significant positive network externalities on the propagation delay (the larger the number of participants, the smaller expected propagation delay). Second, while many existing works study the effect of the number of players, uncertainty, and the number of prizes on tournaments, the number of participants is deterministic and exogenous. On the contrary, in the context of P2P networks, the number of participants is stochastic and depends on content availability. We believe this research expands the boundary of contest concept appropriately to the context of P2P networks, and adds value to the fields of P2P networks and contest literature.

This work can be extended in a number of directions. For example, it is interesting to study the model developed under other propagation delay distribution functions. One could also study the impact of competition between multiple P2P providers. Designing experiments to empirically test the model in a controlled environment is another very interesting avenue for future research.

Acknowledgments

Yong Tan is the Chang Jiang Scholar Visiting Chair Professor at Tsinghua University, and was supported in part by the National Science Foundation of China (grants 71490723, 71531013, and 71572004). The authors have made equal contributions, and their names are listed in alphabetical order.

Appendix: Proofs

PROOF OF PROPOSITION 1. Given other participants' optimal delay selection, D_i^*, $\forall i \neq j$, peer j's best choice of provision delay, D_j^*, is determined by solving the following:

$$\frac{\partial U_j}{\partial D_j} = w_1 \sum_{k=\hat{n}}^{N-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \times \frac{\partial}{\partial D_j} \left(\int_0^\tau \sum_{u=0}^{\hat{n}-1} \binom{k}{u} \left(F(D_j+T_j)\right)^u \left(1-F(D_j+T_j)\right)^{k-u} f_T(T_j) dT_j \right) + c_j (D_{\max} - D_j) = 0. \tag{A1}$$

Given that T_j is uniformly distributed, the above expression can be simplified to Equation (10) in Proposition 1 by differentiating and integrating. Next, we derive the cumulative distribution function, F, which results from a convolution of two delays. The PDF for the total delay (the sum of provision and transmission delays) can be expressed as:

$$f(x) = \frac{1}{\tau} \sum_{p=1}^{P} m_p \theta(x - D_p) \theta(D_p + \tau - x), \tag{A2}$$

where

$$\theta(y) = \begin{cases} 1, & \text{if } y > 0; \\ 1/2 & \text{if } y = 0; \\ 0, & \text{if } y < 0, \end{cases} \tag{A3}$$

is a step function. Hence, integrating (A2) will yield the CDF:

$$F(x) = \frac{1}{\tau} \sum_{p=1}^{P} m_p ((\tau)\theta(x - D_p - \tau) + (x - D_p)\theta(x - D_p)\theta(D_p + \tau - x)). \tag{A4}$$

Plugging $D_j + \tau$ and D_j in Equation (A4) gives (11) and (12) in Proposition 1. □

PROOF OF LEMMA 1. For homogeneous peers, Equations (11) and (12) in Proposition 1 give $F(D+\tau) = 1$ and $F(D) = 0$. □

PROOF OF LEMMA 2.

$$\begin{aligned} \Pr\{j \in \Phi | D_j\} &= \sum_{k=0}^{\hat{n}-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \\ &\quad + \sum_{k=\hat{n}}^{N-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \cdot \sum_{u=0}^{\hat{n}-1} \binom{k}{u} \int_0^\tau \left(\frac{T_j}{\tau}\right)^u \left(1 - \frac{T_j}{\tau}\right)^{k-u} \frac{1}{\tau} dT_j \\ &= \sum_{k=0}^{\hat{n}-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \\ &\quad + \sum_{k=\hat{n}}^{N-1} \binom{N-1}{k} \beta^k (1-\beta)^{N-1-k} \frac{\hat{n}}{k+1} \\ &= \sum_{k=1}^{\hat{n}} \binom{N}{k} \beta^k (1-\beta)^{N-k} \frac{k}{N\beta} \\ &\quad + \sum_{k=\hat{n}+1}^{N} \binom{N}{k} \beta^k (1-\beta)^{N-k} \frac{k}{N\beta} \frac{\hat{n}}{k} = \frac{\kappa}{N\beta}. \end{aligned} \tag{A5}$$

PROOF OF PROPOSITION 2. The expected profit function for the P2P provider is:

$$\pi = H\left(v - \gamma\left(D_{\max} + \frac{\tau}{2}\right)\right) + \frac{\gamma H \eta}{c\tau} w_1 - N\beta w_0 - \kappa w_1, \tag{A6}$$

which is a strictly decreasing function of availability compensation w_0. The effect of upload compensation w_1 depends on the sign of:

$$\frac{\gamma H \eta}{c\tau} - \kappa, \tag{A7}$$

which measures the difference between the benefit from delay reduction and cost of paying the peers. If Equation (A7) is negative, the P2P provider will decrease both compensations to 0. The expected utility of peers:

$$\begin{aligned} U &= w_0 + \frac{\kappa}{N\beta} w_1 - \frac{1}{2} c (D_{\max} - D^*)^2 \\ &= w_0 + \frac{\kappa}{N\beta} w_1 - \frac{\eta^2}{2c\tau^2} {w_1}^2 \end{aligned} \tag{A8}$$

is zero, and individual participation constraint is satisfied.

If Equation (A7) is positive, the P2P provider should increase w_1 but decrease w_0, as long as peers receive non-negative utility. While keeping $w_1 = 0$, w_1 can be increased all the way to w_1^0, given in Equation (17), where peers' expected utility reaches zero.

w_1 can be further increased, however, availability compensation w_0 has to be set at a positive value, such that Equation (A8) is binding, that is $U = 0$. w_0 becomes a function of w_1, and after being substituted into Equation (A6), gives the optimal solution w_1^* described in Equation (20). In order for the optimal solution $w_1^* > w_1^0$, we require that $\frac{\gamma H\eta}{c\tau} > 2\kappa$.

The optimal delay can be found by plugging in the expression for w_1^*. □

Proof of Proposition 3. We use Equations (11) and (12) in Proposition 1 and plug them into Equation (10) to derive the optimal delay choice of each type of peer. For the low-cost peers, we have:

$$-\frac{w_1}{\tau}\sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\left(1-\sum_{u=0}^{\hat{n}-1}\binom{k}{u}\left(m_1+m_2\frac{\tau-\Delta D}{\tau}\theta(\tau-\Delta D)\right)^u\times\left(m_2-m_2\frac{\tau-\Delta D}{\tau}\theta(\tau-\Delta D)\right)^{k-u}\right)+c_1(D_{\max}-D_1)=0. \tag{A9}$$

For the high-cost peers, we have:

$$-\frac{w_1}{\tau}\sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\sum_{u=0}^{\hat{n}-1}\binom{k}{u}\left(m_1\left(\theta(\Delta D-\tau)+\left(\frac{\Delta D}{\tau}\right)\theta(\tau-\Delta D)\right)\right)^u\times\left(1-\left(m_1\left(\theta(\Delta D-\tau)+\left(\frac{\Delta D}{\tau}\right)\theta(\tau-\Delta D)\right)\right)\right)^{k-u}+c_2(D_{\max}-D_2)=0. \tag{A10}$$

The closed form solutions for the general case are not readily available. We focus on the scenario where the two classes are separated, that is $\Delta D = D_2 - D_1 \geq \tau$. The optimal delay selections are:

$$-\frac{w_1}{\tau}\sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\sum_{u=\hat{n}}^{k}\binom{k}{u}{m_1}^u{m_2}^{k-u}+c_1(D_{\max}-D_1)=0,\ \text{and} \tag{A11}$$

$$-\frac{w_1}{\tau}\sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\sum_{u=0}^{\hat{n}-1}\binom{k}{u}{m_1}^u{m_2}^{k-u}+c_2(D_{\max}-D_2)=0, \tag{A12}$$

respectively, where

$$\eta_1=\sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\sum_{u=\hat{n}}^{k}\binom{k}{u}{m_1}^u{m_2}^{k-u},\ \text{and} \tag{A13}$$

$$\eta_2=\sum_{k=\hat{n}}^{N-1}\binom{N-1}{k}\beta^k(1-\beta)^{N-1-k}\sum_{u=0}^{\hat{n}-1}\binom{k}{u}{m_1}^u{m_2}^{k-u}. \tag{A14}$$

It can be easily verified that $\eta = \eta_1 + \eta_2$. Then the two optimal delay selections D_1^* and D_2^* can be solved as given by Equation (21).

Proof of Lemma 3. The expected number of class 1 nodes selected is:

$$\begin{aligned}\kappa_1&=\sum_{k=1}^{N}\binom{N}{k}\beta^k(1-\beta)^{N-k}\sum_{u=0}^{k}\binom{k}{u}m_1^u{m_2}^{k-u}\min(u,\hat{n})\\&=\sum_{k=1}^{\hat{n}-1}\binom{N}{k}\beta^k(1-\beta)^{N-k}m_1k+\sum_{k=\hat{n}}^{N}\binom{N}{k}\beta^k(1-\beta)^{N-k}\\&\quad\times\left(\sum_{u=0}^{\hat{n}}\binom{k}{u}{m_1}^u{m_2}^{k-u}u+\sum_{u=\hat{n}+1}^{k}\binom{k}{u}{m_1}^u{m_2}^{k-u}\hat{n}\right).\end{aligned} \tag{A15}$$

The expected number of class 2 nodes selected is $\kappa_2 = \kappa - \kappa_1$, and we can express it as:

$$\begin{aligned}\kappa_2&=\sum_{k=1}^{\hat{n}-1}\binom{N}{k}\beta^k(1-\beta)^{N-k}m_2k\\&\quad+\sum_{k=\hat{n}}^{N}\binom{N}{k}\beta^k(1-\beta)^{N-k}\sum_{u=0}^{\hat{n}}\binom{k}{u}{m_1}^u{m_2}^{k-u}(\hat{n}-u).\end{aligned} \tag{A16}$$

In the following, we evaluate:

$$\begin{aligned}\frac{\kappa_1}{m_1}-\frac{\kappa_2}{m_2}&=\sum_{k=\hat{n}}^{N}\binom{N}{k}\beta^k(1-\beta)^{N-k}\left(\sum_{u=1}^{\hat{n}}\binom{k}{u}{m_1}^{u-1}{m_2}^{k-u}u\right.\\&\quad\left.+\hat{n}\sum_{u=\hat{n}+1}^{k}\binom{k}{u}{m_1}^{u-1}{m_2}^{k-u}\right)-\sum_{k=\hat{n}}^{N}\binom{N}{k}\beta^k\left(1-\beta\right)^{N-k}\\&\quad\times\sum_{u=0}^{\hat{n}}\binom{k}{u}{m_1}^u{m_2}^{k-u-1}\left(\hat{n}-u\right)=\sum_{k=\hat{n}}^{N}\binom{N}{k}\beta^k(1-\beta)^{N-k}\\&\quad\times\left(k\sum_{u=0}^{\hat{n}-1}\binom{k-1}{u}{m_1}^u{m_2}^{k-u-1}\frac{k-\hat{n}}{k-u}\right.\\&\quad\left.+\hat{n}\sum_{u=\hat{n}+1}^{k}\binom{k}{u}{m_1}^{u-1}{m_2}^{k-u}\right)>0.\end{aligned} \tag{A17}$$

Proof of Proposition 4. The expected price charged for each request is:

$$\pi_0 = E(v - \gamma(D+T)) = v - \gamma\Big(\frac{\kappa_1}{\kappa}D_1^* + \frac{\kappa_2}{\kappa}D_2^* + \frac{\tau}{2}\Big), \tag{A18}$$

which gives the expected profit function:

$$\pi = H\Big(v - \gamma\Big(D_{\max} + \frac{\tau}{2}\Big)\Big) + \frac{\gamma H}{\kappa\tau}\left(\frac{\kappa_1\eta_1}{c_1} + \frac{\kappa_2\eta_2}{c_2}\right)w_1 - N\beta w_0 - \kappa w_1. \tag{A19}$$

The P2P provider optimizes the expected profit by choosing two compensations, subject to individual participation constraints of two classes:

$$\begin{aligned} U_1 &= w_0 + \frac{\kappa_1}{m_1 N\beta}w_1 - \frac{1}{2}c_1\left(D_{\max} - D_1^*\right)^2 \\ &= w_0 + \frac{\kappa_1}{m_1 N\beta}w_1 - \frac{\eta_1^2}{2c_1\tau^2}w_1^2 \geq 0, \end{aligned} \tag{A20}$$

$$\begin{aligned} U_2 &= w_0 + \frac{\kappa_2}{m_2 N\beta}w_1 - \frac{1}{2}c_2\left(D_{\max} - D_2^*\right)^2 \\ &= w_0 + \frac{\kappa_2}{m_2 N\beta}w_1 - \frac{\eta_2^2}{2c_2\tau^2}w_1^2 \geq 0. \end{aligned} \tag{A21}$$

Similarly, a higher w_0 decreases the expected profit, but the effect of w_1 depends on the sign of:

$$\Gamma = \frac{\gamma H}{\kappa\tau}\left(\frac{\kappa_1\eta_1}{c_1} + \frac{\kappa_2\eta_2}{c_2}\right) - \kappa. \tag{A22}$$

To obtain the separating equilibrium $D_2 - D_1 \geq \tau$, we require $\Gamma > 0$ such that w_1 is set at a positive value.

We need to analyze three distinctive cases which show the interplay between the two classes. The utility functions for the two classes are plotted in Figure 3. Case A is specified by the following condition:

$$\frac{\kappa_1 c_1}{m_1\eta_1^2} < \frac{\kappa_2 c_2}{m_2\eta_2^2}, \text{ (which implies } \frac{\eta_1^2}{c_1} > \frac{\eta_2^2}{c_2}\text{)}. \tag{A23}$$

When we increase w_1, the utility for class 1 reaches zero first and remains below that of class 2. Therefore, $U_2 \geq 0$ always holds while U_1 remains nonnegative. Similar to the homogeneous case, we solve $U_1 = 0$ for w_0, express it as a function of w_1, and plug it in Equation (A19) to find the optimal $w_{1,A}^*$. If $w_{1,A}^* < w_{1,A}^0$, where $U_1\left(w_0 = 0, w_1 = w_{1,A}^0\right) = 0$, then we set the optimal upload compensation to $w_{1,A}^0$, and availability compensation to zero. The condition for $w_{1,A}^* \geq w_{1,A}^0$ is given by:

$$\frac{\gamma H}{\kappa\tau}\left(\frac{\kappa_1\eta_1}{c_1} + \frac{\kappa_2\eta_2}{c_2}\right) - \kappa \geq \frac{\kappa_1}{m_1}, \text{ or } \Gamma \geq \frac{\kappa_1}{m_1}. \tag{A24}$$

Case B is characterized by:

$$\frac{\kappa_1 c_1}{m_1\eta_1^2} > \frac{\kappa_2 c_2}{m_2\eta_2^2}, \text{ and } \frac{\eta_1^2}{c_1} < \frac{\eta_2^2}{c_2}. \tag{A25}$$

When we increase w_1, the utility for class 2 reaches zero first and remains below that of class 1. The analyses can be conducted similarly to Case A.

Case C occurs when the following condition is satisfied.

$$\frac{\kappa_1 c_1}{m_1\eta_1^2} > \frac{\kappa_2 c_2}{m_2\eta_2^2}, \text{ and } \frac{\eta_1^2}{c_1} > \frac{\eta_2^2}{c_2}, \tag{A26}$$

The utility functions for the two classes cross at the following point:

$$w_1^c = \frac{2\tau^2}{N\beta}\left(\frac{\kappa_1}{m_1} - \frac{\kappa_2}{m_2}\right)\left(\frac{\eta_1^2}{c_1} - \frac{\eta_2^2}{c_2}\right)^{-1}. \tag{A27}$$

We can solve the optimization problem by assuming that $U_2 = 0$ and $U_1 > 0$ to $w_{1,B}^*$, or $U_1 = 0$ and $U_2 > 0$ to $w_{1,A}^*$, and compare them. When we increase Γ, both $w_{1,A}^*$ and $w_{1,B}^*$ increase, but $w_{1,A}^*$ grows slower than $w_{1,B}^*$ because of Equation (A26).

Class 2 peers have $U_2 = 0$ till $w_{1,B}^*$ reaches w_1^c (or $\Gamma = \Gamma_2$) where both classes receive zero utility. As Γ further increases, $w_{1,B}^*$ is larger than w_1^c, but $w_{1,A}^*$ remains below w_1^c. The only feasible solution is to set the compensations at $\left(w_0^c, w_1^c\right)$, till $\Gamma \geq \Gamma_1$ when $w_{1,A}^* > w_1^c$. □

References

Adar, E., B. A. Huberman. 2000. Free riding on Gnutella. *First Monday*. Rishab Aiyer Ghosh. Available at http://firstmonday.org/ojs/index.php/fm/article/view/792/701 (accessed date March 22, 2017).

Afèche, P., H. Mendelson. 2004. Pricing and priority auctions in queueing systems with a generalized delay cost structure. *Management Sci.* **50**(7): 869–882.

Antoniadis, P., C. Courcoubetis, R. Mason. 2004. Comparing economic incentives in Peer-to-Peer networks. *Com. Nets.* 133–146.

Arrow, K. J. 1970. *Essays in the Theory of Risk Bearing*. North-Holland, Amsterdam.

Asvanund, A., K. Clay, R. Krishnan, M. D. Smith. 2004. An empirical analysis of network externalities in Peer-to-Peer music sharing networks. *Inf. Syst. Res.* **15**(2): 155–174.

Bertsekas, D., R. Gallager. 1992. *Data Networks*, 2nd edn. Prentice-Hall International, Englewood Cliffs, NJ.

Betteridge, I. 2004. Napster Founder's Snocap: DRM for P2P download sites. *ExtremeTech*. Retrieved March 22, 2017. Available at http://www.extremetech.com/article2/0,1558,1736556,00.asp (accessed date May 15, 2016).

Bryce, D. 2007. P2P music sharing service Grooveshark Ups compensation. *TechCrunch*. Available at http://techcrunch.com/2007/10/10/p2p-music-sharing-service-grooveshark-ups-compensation (accessed date December 23, 2016).

Cornes, R., T. Sandler. 1996. *The Theory of Externalities, Public Goods and Club Goods*. Cambridge University Press, Cambridge.

CVN Index. 2015. Cisco visual networking index: Forecast and methodology, 2014–2019. CISCO White Paper.

Feigenbaum, J., C. Papadimitriou, R. Sami, S. Shenker. 2002. A BGP-based mechanism for lowest-cost routing. Proceedings of the ACM Symposium on Principles of Distributed Computing, Monterey, CA, USA.

Feldman, M., J. Chuang, I. Stoica, S. Shenker. 2007. Hidden-action in network routing. *IEEE J. Sel. Areas in Comms.* **25**(6): 1161–1172.

Ganesh, A., K. Laevens, R. Steinberg. 2007. Congestion Pricing and Noncooperative Games in Communication Networks. *Oper. Res.* **55**(3): 430–438.

Ghosal, D., B. K. Poon, K. Kong. 2005. P2P contracts: A framework for resource and service exchange. *Future Generation Comput. Syst.* **21**(3): 333–347.

Gibbons, R. 1992. *Game Theory for Applied Economists*. Princeton University Press, Princeton, NJ.

Green, J. R., N. Stokey. 1983. A comparison of tournaments and contracts. *J. Polit. Econ.* **91**: 349–364.

Guo, H., H. K. Cheng, S. Bandyopadhyay. 2013. Broadband network management and the net neutrality debate. *Prod. Oper. Manag.* **22**(5): 1287–1298.

Gupta, A., L. Awasthi. 2010. Toward a quality-of-service framework for Peer-to-Peer applications. *Int. J. Dist. Sys. & Tech* **1**(3): 1–23.

Hastings, R. 2014. Internet tolls and the case for strong net neutrality. *Netflix Blog*, March 20, 2014. Available at http://blog.netflix.com/2014/03/internet-tolls-and-case-for-strong-net.html (accessed date May 15, 2016).

Herings, P. J. J., R. Peeters, M. S. Yang. 2010. Competition against Peer-to-Peer networks. *Info. Econ. Pol.* **22**(4): 315–331.

Holmstrom, B. 1982. Moral hazard in teams. *Bell J. Econ.* **13**: 324–340.

Hosanagar, K., P. Han, Y. Tan. 2010. Diffusion models for Peer-to-Peer (P2P) media distribution: On the impact of decentralized, constrained supply. *Inf. Syst. Res.* **21**(2): 271–287.

Johar, M., S. Menon, V. S. Mookerjee. 2011. Analyzing sharing in Peer-to-Peer networks under various congestion measures. *Inf. Syst. Res.* **22**(2): 325–345.

Johar, M., N. Kumar, V. S. Mookerjee. 2012. Content provision strategies in the presence of content piracy. *Inf. Syst. Res.* **23**(3-part-2): 960–975.

Kalker, T., D. H. J. Epema, P. H. Hartel, R. L. Lagendijk, M. van Steen. 2004. Music2Share - copyright-compliant music sharing in P2P systems. *Proc. IEEE* **92**(6): 961–970.

Kalra, A., M. Shi. 2001. Designing optimal sales contests: A theoretical perspective. *Market. Sci.* **20**(3): 170–193.

Krishnan, R., M. D. Smith, Z. Tang, R. Telang. 2004. The impact of free-riding on Peer-to-Peer networks. Proceedings of 37th Hawaii International International Conference on Systems Science 70199c, Big Island, HI.

Krishnan, R., M. D. Smith, Z. Tang, R. Telang. 2006. Digital business models for Peer-to-Peer networks: Analysis and economic issues. Heinz Research Paper 50. Available at http://repository.cmu.edu/heinzworks/50 (accessed date April 29, 2017).

Kumar, C., K. Altinkemer, P. De. 2011. A mechanism for pricing and resource allocation in Peer-to-Peer networks. *Electron. Commer. Res. Appl.* **10**(1): 26–37.

Laffont, J. J., D. Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press, Princeton, NJ.

Lang, K. R., R. Vragov. 2005. A pricing mechanism for digital content distribution over computer networks. *J. Manage. Inf. Syst.* **22**(32): 121–139.

Lazear, E. P., S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *J. Polit. Econ.* **89**: 841–864.

Li, Y.-M., Y. Tan, P. De. 2013. Self-organized formation and evolution of Peer-to-Peer networks. *INFORMS J. Comp.* **25**(3): 502–516.

Ma, R. T. B., S. C. M. Lee, J. C. S. Lui, D. K. Y. Yau. 2006. Incentive and service differentiation in P2P networks: a game theoretic approach. *IEEE/ACM Trans. Netw.* **14**(5): 978–991.

Maillé, P., L. Toka. 2008. Managing a Peer-to-Peer data storage system in selfish society. *IEEE J. Sel. Areas Commun.* **26**(7): 1295–1301.

MacKie-Mason, J. K., H. R. Varian. 1995. Pricing congestible network resources. *IEEE J. Sel. Areas Commun.* **13**(7): 1141–1149.

Marlowe, C. 2006. Warners rolling with BitTorrent. *The Hollywood Reporter*, May 9. Available at http://www.hollywoodreporter.com/hr/search/article_display.jsp?vnu_content_id=1002464855 (accessed date May 15, 2016).

Nalebuff, B., J. Stiglitz. 1983. Prices and incentives: Towards a general theory of compensation and competition. *Bell J. Econ.* **14**(1): 21–43.

Shapiro, C., J. E. Stiglitz. 1984. Equilibrium unemployment as a worker discipline device. *Am. Econ. Rev.* **74**: 433–441.

Spence, A. M., R. J. Zeckhauser. 1971. Insurance, information, and individual action. *Am. Econ. Rev.* **61**: 380–387.

Steele, R. 2015. If you think piracy is decreasing, you haven't looked at the data…. *Digital Music News*. Available at http://www.digitalmusicnews.com/2015/07/16/if-you-think-piracy-is-decreasing-you-havent-looked-at-the-data-2 (accessed date December 24, 2016).

Strausz, R. 1996. Moral hazard in sequential teams. Department Working, Free University of Berlin.

StreamerP2P. 2016. http://www.streamerp2p.com (accessed date October 12, 2016).

The Insight Research Corporation. 2008. Peer-to-Peer and file sharing services market. Available at http://www.insight-corp.com/reports/ipappsp2p.asp (accessed date March 22, 2017).

TorrentFreak. 2014a. Netflix considers P2P-powered streaming technology. September 1, 2014. Available at https://torrentfreak.com/netflix-considers-p2p-powered-streaming-technology-140426 (accessed date March 22, 2017).

TorrentFreak. 2014b. Popcorn time installed on 1.4 million devices in the U.S. September 1, 2014. Available at https://torrentfreak.com/popcorn-time-installed-1-4-million-devices-u-s-140901 (accessed date March 22, 2017).

Van der Sar, E. 2015. Netflix needs bittorrent expert to implement P2P streaming. Retrieved December 11, 2015. Available at https://torrentfreak.com/netflix-needs-bittorrent-expert-to-implement-p2p-streaming-150520 (accessed date March 22, 2017).

Victer Paul, P., N. Saravanan, S. K. V. Jayakumar, P. Dhavachelvan, R. Baskaran. 2012. QoS enhancements for global replication management in peer to peer networks. *Future Generation Comp. Syst.* **28**(3): 573–582.

Xia, M., Y. Huang, W. Duan, A. B. Whinston. 2012. To continue sharing or not to continue sharing? An empirical analysis of user decision in Peer-to-Peer sharing networks. *Inf. Syst. Res.* **23**(1): 247–259.

Xu, H., S. Wang, R. Wang, Y. Rao, X. Shao. 2010. Improving QoS in Peer-to-Peer streaming media system. *J. Comput. Infor. Syst.* **6**(5): 1387–1395.

Zhang, H., S. Zenios. 2008. A dynamic principal-agent model with hidden information: Sequential optimality through truthful state revelation. *Oper. Res.* **56**(3): 681–696.