Lecture Notes in Computer Science 5539

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Peter Reichl Burkhard Stiller
Bruno Tuffin (Eds.)

Network Economics for Next Generation Networks

6th International Workshop on Internet Charging and QoS Technologies, ICQT 2009
Aachen, Germany, May 11-15, 2009
Proceedings

Springer

Volume Editors

Peter Reichl
Forschungszentrum Telekommunikation Wien, FTW
Donau-City-Strasse 1, 1220 Wien, Austria
E-mail: reichl@ftw.at

Burkhard Stiller
University of Zürich, UZH
Communication Systems Group, CSG
Department of Informatics, IFI
Binzmühlestrasse 14, 8050, Zürich, Switzerland
E-mail: stiller@ifi.uzh.ch

Bruno Tuffin
INRIA Rennes, Bretagne Atlantique
Campus de Beaulieu, 35042 Rennes Cedex, France
E-mail: Bruno.Tuffin@irisa.fr

Library of Congress Control Number: Applied for

CR Subject Classification (1998): C.2, H.4, H.3, J.1, K.4.4, K.6.4

LNCS Sublibrary: SL 5 – Computer Communication Networks
and Telecommunications

ISSN 0302-9743
ISBN-10 3-642-01795-9 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-01795-7 Springer Berlin Heidelberg New York

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12669301 06/3180 5 4 3 2 1 0

Preface

Starting with the imminent roll-out of the IP Multimedia Subsystem (IMS) and fourth-Generation networking technology, Next Generation Networks (NGN) are gradually becoming reality, with charging and Quality-of-Service (QoS) issues as two of the key drivers for the evolution toward the convergent all-IP network of the future. Therefore, the 6th International Workshop on Internet Charging and QoS Technology (ICQT 2009) was devoted to discussing the most recent approaches, models, and mechanisms in this highly interesting and important research area.

The present volume of the *Lecture Notes in Computer Science* series includes those papers presented at ICQT 2009—collocated this year with the IFIP Networking 2009 conference—taking place on May 15, 2009, in Aachen, Germany and hosted by the Rheinisch-Westfälische Technische Hochschule (RWTH Aachen).

For the commercial success of future QoS-enabled communication services, the emergence of viable business models, pricing schemes, and charging and accounting mechanisms is of paramount importance. Problems in this domain can only be addressed through a broad interdisciplinary approach linking together a variety of technical and economic perspectives, which are constantly driving a plethora of relevant research topics for application developers, business architects, network providers, service providers, and customers. Within the current trend toward a convergent NGN architecture, competition modeling, pricing mechanisms, and the economics of inter-domain traffic are of specific importance and urgency. Thus, they determined—in the form of three technical sessions—the core of the ICQT 2009 program. All contributions included in this volume fit perfectly into the general scope of the international ICQT workshop series, which is mainly characterized by the focus on identifying novel service charging solutions, investigating and evaluating their technical feasibility, and consolidating technical and economic mechanisms for enabling a fast, guaranteed, and efficient charging of services. This is of fundamental importance for the future evolution of convergent all-IP Next Generation Networks.

This year's ICQT followed the already established tradition of an unusually vivid workshop series on charging and QoS technology issues, which started back in 2001 with the first ICQT workshop in the framework of the Annual Meeting of the German Society for Computer Science (GI) and the Austrian Computer Society 2001 in Vienna, Austria. In 2002, ICQT was collocated with the QofIS 2002 workshop in Zürich, Switzerland, in 2003 with the NGC 2003 workshop in Munich, Germany, and in 2004 again with QofIS 2004 in Barcelona, Spain. In 2006, ICQT was hosted by the Universitary Technological Institute St. Malo, France together with ACM SIGMETRICS 2006.

As in the past, ICQT 2009 brought together researchers from the area of technology and economy in both industry and academia to discuss key improvements and to support further progress in these fields. The combination of micro-economic models, auctions, game theoretic approaches, peer-to-peer, and IMS-based charging addresses a highly interesting facet at the intersection of networking research and business

modeling. Thus, ICQT 2009 provided a truly interdisciplinary forum for analyzing topics at the overlapping of those two areas, providing for a unifying framework for all presentations included in the program of ICQT 2009.

Like all of its predecessors, ICQT 2009 provided a single-track and one-day program, which proved to be especially suitable for stimulating the interaction between and the active participation of the workshop audience. Summarized briefly, the workshop started with a keynote presentation delivered by Jim Roberts, who is internationally recognized to be one of the most distinguished research fellows in the area. The following three technical sessions included a total of nine full papers, which were selected after a thorough reviewing process out of a total number of 26 submissions. The resulting final program demonstrated again the international scope of this workshop series and included papers from Europe and Asia.

The international orientation of ICQT is also reflected in its composition of the Technical Program Committee, whose members again devoted their excellent knowledge together with many hours of their precious time to provide the basis for a highly qualified technical program and, thus, to contribute in an unfailing way to the technical and research success of ICQT. Furthermore, the editors would like to express their thanks to the ICQT 2009 webmaster for his excellent work.

Special thanks go to the organizers of the IFIP Networking 2009 conference for enabling the collocation of ICQT 2009 with their renowned event, as well as to the local organization handled in a truly exceptional way by Otto Spaniol and his enthusiastic team, including Martin Krebs, Jan Kritzner, and Alexander Zimmermann.

Finally, all three editors would like to address their thanks to Springer, and especially Anna Kramer, for a smooth cooperation on finalizing these proceedings. Additionally, special thanks go to the support of the European COST Action IS0605 “Econ@Tel” and their WG4 researchers as well as to the FP7 EU-funded project “SmoothIT” (No. 216259).

May 2009

Peter Reichl
Burkhard Stiller
Bruno Tuffin

Organization

Program Co-chairs ICQT 2009

Peter Reichl	*FTW Vienna, Austria*
Burkhard Stiller	*University of Zürich, Switzerland*
Bruno Tuffin	*INRIA Rennes, Bretagne Atlantique, France*

Program Committee ICQT 2009

Eitan Altmann	*INRIA Sophia-Antipolis, France*
Jörn Altmann	*Seoul National University, Korea*
Ragnar Andreassen	*Telenor, Norway*
Dominique Barth	*University of Versailles, France*
Nalan Gulpinar	*Warwick Business School, UK*
Peter Key	*Microsoft Research Cambridge, UK*
Patrick Maillé	*Telecom Bretagne, France*
Olivier Marce	*Alcatel-Lucent, France*
Ravi Mazumdar	*University of Waterloo, Canada*
Maurizio Naldi	*University of Rome Tor Vergata, Italy*
Andrew Odlyzko	*University of Minnesota, USA*
David Reeves	*North Carolina State University, USA*
Günter Schäfer	*University of Ilmenau, Germany*
Vasilios Siris	*ICS Forth, Greece*
George Stamoulis	*Athens University of Economics and Business, Greece*
Nicolas Stier	*Columbia University, USA*
Yoshiaki Tanaka	*Waseda University, Japan*
Tuan Trinh Anh	*BME Budapest, Hungary*
Jean Walrand	*University of California, Berkeley, USA*

Reviewers

A set of very detailed and constructive reviews for papers submitted to ICQT 2009 were provided by all of our reviewers, corresponding to the full Program Committee members as stated above and Loubna Echabbi additionally. Therefore, it is with great pleasure that the Program Co-chairs thank all those reviewers for their important work.

Table of Contents

Keynote

QoS Is Still an Issue, Congestion Pricing Is Not the Solution 1
Jim Roberts

Session 1: Competition Models

Optimization of Transmission Power in Competitive Wireless Networks 2
Patrick Maillé and Bruno Tuffin

On Competition for Market Share in a Dynamic ISP Market with Customer Loyalty: A Game-Theoretic Analysis 11
László Gyarmati and Tuan Anh Trinh

A Pricing Model for a Mobile Network Operator Sharing Limited Resource with a Mobile Virtual Network Operator 24
Hélène Le Cadre, Mustapha Bouhtou, and Bruno Tuffin

Session 2: Pricing Mechanisms

Design and Evaluation of a Combinatorial Double Auction for Resource Allocations in Grids 36
Li Li, Yuanan Liu, David Hausheer and Burkhard Stiller

A User-Influenced Pricing Mechanism for Internet Access 48
Gergely Biczók and Tuan Anh Trinh

Price Setting in Two-Sided Markets for Internet Connectivity 61
Thorsten Hau and Walter Brenner

Session 3: Economics of Inter-domain Traffic

Online Charging for IMS-Based Inter-domain Composite Services 72
Minh van Le, Frens Jan Rumph, George B. Huitema, Bert-Jan van Beijnum, and L.J.M. (Bart) Nieuwenhuis

A New Bilateral Arrangement between Interconnected Providers 85
Ruzana Davoyan, Jörn Altmann, and Wolfgang Effelsberg

Improvement of BitTorrent Performance and Inter-domain Traffic by Inserting ISP-Owned Peers 97
Ioanna Papafili, Sergios Soursos, and George D. Stamoulis

Author Index 109

QoS Is Still an Issue, Congestion Pricing Is Not the Solution

Jim Roberts
Orange Labs, France

Extended Abstract

A network clearly needs to be designed to meet user performance requirements for a wide variety of applications. In a commercial setting, Return on Investment (RoI) must be covered by the price users pay for the services provided by the network. It matters whether the investment is just in an amount of commodity infrastructure or also in complex value-added services justifying a higher profit margin. Quality-of-Service (QoS) is often seen as the basis for such added value. We discuss the issue of RoI and consider the complementary role of pricing as a QoS mechanism.

Unfortunately, none of the QoS models proposed for standardization over the past decades has provided a satisfactory solution. On one hand, it proves practically impossible to perform resource allocation so that a given flow, characterized by a traffic "profile", encounters precise performance criteria. On the other hand, it is rather easy to ensure excellent quality for all simply by providing capacity that is somewhat greater than expected demand. It is hardly possible to realize finely modulated quality levels, since performance deteriorates rapidly and unacceptably as demand exceeds capacity. We explore the scope for service differentiation based on our understanding of the stochastic nature of network traffic and discuss the limits on possible QoS control.

Overprovisioning is not a satisfactory solution for operators under the present business model where pricing is largely independent of traffic volume. Growth in demand due to the popularity of applications like file sharing and video streaming requires added investment in infrastructure but brings negligible return. Operators are, therefore, seeking to introduce a new network model giving priority to managed services whose usage is subject to a particular pricing scheme. We consider the viability of this two-tier service model and its acceptability in the light of the on-going debate on network neutrality.

QoS control would be considerably simpler, if users were made to pay in relation to the amount of congestion they cause. This is the principle of congestion pricing and a number of possible schemes have been proposed for the Internet. Despite arguments for microeconomic optimality, these schemes seem completely unworkable, if only for their obvious lack of charging transparency. A recent proposal to preserve flat rate charging for end users and to apply a form of "congestion policing" instead of congestion pricing does not appear to be a satisfactory alternative. We explain why we believe pricing should be reserved for its primary function of RoI.

Following the above discussion we conclude by proposing the outline of an alternative approach to QoS control. The essential mechanisms are network imposed fair sharing, load shedding as necessary to avoid overload, user controlled sharing of last mile resources, and simple usage-based charging. We present feasibility arguments and highlight areas needing further research.

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, p. 1, 2009.

Optimization of Transmission Power in Competitive Wireless Networks

Patrick Maillé[1] and Bruno Tuffin[2]

[1] Institut Telecom; Telecom Bretagne
2 rue de la Châtaigneraie CS 17607, 35576 Cesson Sévigné Cedex, France
Université européenne de Bretagne
patrick.maille@telecom-bretagne.eu
[2] INRIA Rennes - Bretagne Atlantique, Campus universitaire de Beaulieu
35042 Rennes Cedex, France
btuffin@irisa.fr

Abstract. Competition among providers has become an important issue in current and future wireless telecommunication networks. The providers may operate using different technologies, such as WiFi, WiMAX, UMTS... In a previous work, we have analyzed the price competition among two providers, one operating in only a subdomain of the other, due to smaller distance range. A typical situation is WiFi against WiMAX. We propose here to add a supplementary level of decision on top of that game, making use of its equilibrium: the smaller-range provider plays with its transmission power in order to attract more customers and potentially increase its revenue. We determine the optimal power in the case where energy has a negligible cost, as well as when its cost is linear in transmission power.

Keywords: Economics, Competition, Wireless Networks, WiFi, WiMAX.

1 Introduction

Telecommunication networks, and especially wireless networks, have experienced an increase in terms of traffic and subscriptions, but at the same time a fierce competition among providers. Demand for service is distributed among competitors based on the access price and available Quality of Service (QoS). Pricing strategies therefore form an important parameter in the competition. Up to now, pricing has received a large interest in the networking community, due to resurce scarcity with respect to demand. Pricing is justified by its capacity to control demand (and therefore the QoS), and/or to differentiate services [1–3]. Most models investigate optimal pricing strategies in the case of a *monopoly*, whereas an *oligopoly*, with several providers fighting for customers, could drive to substantially different results, as highlighted in [4]. Thus competition requires a deeper attention. Notable first attempts in this direction can be found, not exhaustively, in [5–8].

This paper pertains to that stream of work. We consider two providers, denoted by provider 1 and provider 2, in competition for customers, where provider 2

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 2–10, 2009.

is assumed to operate in a subdomain of provider 1's access area. This is typically the case of a WiMAX operator (provider 1) against a WiFi one (provider 2). Each provider fixes a price, and demand is distributed according to the classical Wardrop principle [9] described later on.

In a previous paper [10], the subdomain covered by provider 2 was assumed to be fixed, and existence and uniqueness of a Nash equilibrium on prices played by providers were proved. Recall that a Nash equilibrium characterizes a profile of strategies such that no provider can improve its utility (revenue) without changing *unilaterally* its own strategy. It is therefore a point from which we do not deviate in the case of selfish providers. In that paper, using a particular pricing scheme, the price of anarchy was also proved to be one, meaning that there is no loss of social welfare due to user and provider selfishness with respect to an optimal cooperative case.

The present work is built on the results in [10]. Our goal is to investigate what happens if the smaller provider can initially choose its transmission power before the above game is played, and therefore if he can reach, and potentially attract, more customers from the larger provider in order to increase its revenue. We study this problem when there is no cost related to the transmission, and when such a cost exists. The basic question we want to solve is: is there an interest for the smaller user to be competiting over the whole domain? The answer is no in general, even if there is no cost associated to transmission power: actually in this case, there is no gain of revenue, even if no loss either, after reaching a given threshold.

The paper is organized as follows. Section 2 describes the model, its assumptions, and the main results presented in [10]. Section 3 then studies the optimal decision of provider 2 (the smaller one) in terms of the proportion of area he can reach on the domain of provider 1 (directly related to transmission power), in order to maximize its revenue. Section 4 does the same kind of analysis, but in the case where there is a cost associated to the power for transmitting data. Finally we conclude and give few research directions in Section 5.

2 Model and Previous Results

2.1 Model

We consider two providers, denoted by 1 and 2, with provider 2 operating in a subdomain of provider 1, as illustrated in Figure 1. We call zone B the coverage region of provider 2. Zone A stands for the region where only provider 1 operates. This is a typical situation of a WiFi provider operating on smaller distances -tens of meters- than a WiMAX one -covering many kilometers-.

Competition is analyzed on a simplified model, where time is discretized, divided into slots. Let C_i be the capacity of provider i ($i \in \{1, 2\}$), i.e., the number of packets he can serve during one slot. Let d_i be the demand experienced by provider i in a given slot. If $d_i \leq C_i$, all packets are served but as soon as $d_i > C_i$, only C_i are served and the $d_i - C_i$ rejected ones are chosen uniformly. Compacted, each packet is served with probability $\min(C_i/d_i, 1)$.

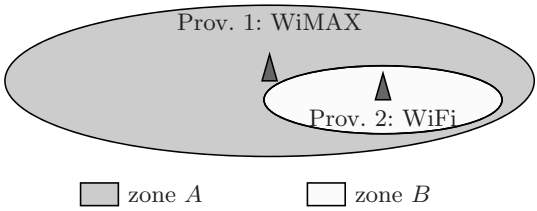


Fig. 1. The competition framework

Each provider fixes an *access* price (p_i for provider i) paid a soon as a packet is submitted, independently of being rejected or transmitted. This produces an incentive to limit the amount of sent packets. Indeed, the *perceived price* $\bar{p}_i$ per packet, i.e., the expected price to successfully send a packet, is

$$\bar{p}_i = p_i / \min(C_i/d_i, 1) = p_i \max(d_i/C_i, 1). \tag{1}$$

This kind of model was already studied in [11] for a single provider (a monopoly), when priority classes are defined and charged with different prices. The case of an oligopoly (with single service classes) is analyzed in [12]. The present model with one provider operating in a sub-area of the other is described in [10], where the smaller one does not play on its transmission range. Our basic model therefore follows [10], as described now.

Taking into account the point of view of the users, total demand on the whole domain depends on the perceived price $\bar{p}$, and is assumed to be a continuous and strictly decreasing function $D(\cdot)$ on its support $[0, p_{\max})$, with possibly $p_{\max} = +\infty$. Of course, we assume some potential congestion, i.e., $D(0) > C_1 + C_2$ to avoid uninteresting cases.

In order to introduce an additional and useful notation, remark that providing a demand function D is equivalent to providing a *marginal valuation function* $v : q \mapsto \inf\{p : D(p) \leq q\}$ (with the convention $\inf \emptyset = 0$), representing the maximum unit price at which q trafic units can be sold:

$$v(q) = \begin{cases} D^{-1}(q) & \text{if } q \in (0, D(0)) \\ p_{\max} & \text{if } q = 0 \\ 0 & \text{if } q \geq D(0). \end{cases} \tag{2}$$

Since D is nonincreasing, neither is v.

We finally define α as the proportion of the population covered by zone B. To simplify our analysis in next section, we will assume users uniformly distributed over the domain, and the domains being delimited by circles, with the restriction that the disc covered by provider 2 is always included in that of provider 1. Nonetheless, the general results of next subsections do not need such restrictions. In any case, the demand function in zone A is $(1-\alpha)D(\cdot)$, while it is $\alpha D(\cdot)$ in zone B, assuming equidistribution of users' willingness-to-pay accross subdomains.

User demand is assumed to split among providers following Wardrop's principle [9], which states all users choose the available provider with the least perceived price, and none if this perceived price is too high. Moreover, the demand on each zone is a function of the minimum perceived price available in that zone. Those conditions are a limit approximation of the Nash equilibrium conditions for the noncooperative game played among users, when the individual weight of each user tends to zero, i.e. no individual user can unilaterally impact the perceived price of the providers. This kind of game is called *nonatomic*, since it corresponds to each user being infinitely small.

Formally, let $d_{1,A}$, the demand experienced by provider 1 in zone A, and $d_{1,B}$, the demand in zone B, with $d_1 = d_{1,A} + d_{1,B}$. For given prices (p_1, p_2) set by the providers, the conditions imposed by Wardrop's principle lead to (a set of) perceived prices $(\bar{p}_1, \bar{p}_2)$, and demands $d_{1,A}, d_{1,B}, d_2$, called a Wardrop equilibrium, that satisfy

$$\bar{p}_1 = p_1 \max\left(1, \frac{d_{1,A} + d_{1,B}}{C_1}\right) \tag{3}$$

$$\bar{p}_2 = p_2 \max\left(1, \frac{d_2}{C_2}\right) \tag{4}$$

$$d_{1,A} = (1-\alpha) D(\bar{p}_1) \tag{5}$$

$$d_{1,B} + d_2 = \alpha D(\min(\bar{p}_1, \bar{p}_2)) \tag{6}$$

$$\bar{p}_1 > \bar{p}_2 \Rightarrow d_{1,B} = 0 \tag{7}$$

$$\bar{p}_1 < \bar{p}_2 \Rightarrow d_2 = 0. \tag{8}$$

Relations (7) and (8) come from zone B users choosing the cheapest provider, while (5) and (6) are the demand-price relations for each zone.

2.2 Previous Analysis

In [10], we have shown the following results:

- For each price profile (p_1, p_2), there exists at least one Wardrop equilibrium. Moreover, the corresponding perceived prices $(\bar{p}_1, \bar{p}_2)$ are unique. The only cases when demands might not be unique are when $\bar{p}_1 = p_1 = p_2 = \bar{p}_2$ and at the same time $d_1 + d_2 < C_1 + C_2$.
- We consider the non-cooperative game, where providers play with their price to maximize their own revenue $R_i(p_1, p_2) := p_i d_i$ knowing that demand will spread according to the above Wardrop equilibrium. Under the common assumption that price elasticity of demand $-\frac{D'(p)p}{D(p)}$ is strictly larger than 1 for all $p \in [\hat{p}, p_{\max})$, with $\hat{p} := \min\left(v\left(\frac{C_1}{1-\alpha}\right), v\left(\frac{C_2}{\alpha}\right)\right)$, there *exists* a *unique*

Nash equilibrium (p_1^*, p_2^*) in the price war between providers[1]. That Nash equilibrium is characterized as follows.

- If $\frac{C_1}{1-\alpha} < \frac{C_2}{\alpha}$ (i.e. $\alpha < \frac{C_2}{C_1+C_2}$), the Nash equilibrium is such that

$$p_1^* = v\left(\frac{C_1}{1-\alpha}\right) < p_2^* = v\left(\frac{C_2}{\alpha}\right). \tag{9}$$

  We then have $d_{1,A} = C_1$, $d_{1,B} = 0$ and $d_2 = C_2$, meaning that demand exactly equals capacity and zone B is left to provider 2 by provider 1.
- If $\frac{C_1}{1-\alpha} > \frac{C_2}{\alpha}$ (i.e. $\alpha \geq \frac{C_2}{C_1+C_2}$), the Nash equilibrium is such that prices are the same

$$p_1^* = p_2^* = p^* = v(C_1 + C_2). \tag{10}$$

  We then have $d_2 = C_2$, $d_{1,A} + d_{1,B} = C_1$. Again, demand exactly equals capacity, but zone B is shared by the providers (except for the limit case $\alpha = \frac{C_2}{C_1+C_2}$, when only provider 2 has customers in zone B).

– The Price of Anarchy, defined as the worst-case ratio comparing social welfare (sum of valuations of all actors) at the Nash equilibrium to the optimal value, is equal to one: social welfare is maximized even in the presence of selfish users and providers.

3 Optimal Radius/Proportion Parameter without Any Cost

We now assume that provider 2 can play with its transmission power, i.e. with the proportion parameter α representing the proportion of customers that it can actually reach (in bijection with transmission power), and for which he will be in competition with provider 1. The idea is to play strategically and *use* the information about what the Nash equilibrium is for each value of α. We therefore end up with a two-stage game where provider 2 plays first on α, and then both providers play on prices. Provider 2 acts as a leader of a kind of Stackelberg game [13], since we assume it uses the knowledge of the pricing game outcome (that we described earlier) to perform its best choice of α.

We further assume to simplify the analysis that the antennas of the two providers are located at the same point. Then, increasing the range of action of the smaller one will always let him in a subdomain of the big one (or the antennas do not need to be located at the same point provided the domain covered by 2 is included in that covered by 1).

In this section, we assume that provider 2 payoff is not affected by the power it uses. The other case is considered in next section. In this situation, from (9)

[1] It this paper we will play on α and still consider that demand elasticity is larger than 1 even if $\hat{p} = \min\left(v\left(\frac{C_1}{1-\alpha}\right), v\left(\frac{C_2}{\alpha}\right)\right)$ changes.

and (10) we have at Nash equilibrium $d_2 = C_2$, and provider 2 revenue is expressed as a function of α by

$$R_2(\alpha) = \begin{cases} v\left(\frac{C_2}{\alpha}\right) C_2 & \text{if } \frac{C_1}{1-\alpha} \leq \frac{C_2}{\alpha} \\ v(C_1 + C_2)C_2 & \text{if } \frac{C_1}{1-\alpha} > \frac{C_2}{\alpha}, \end{cases}$$

which is equivalent to

$$R_2(\alpha) = C_2 v\left(\max\left(\frac{C_2}{\alpha}, C_1 + C_2\right)\right). \tag{11}$$

It is therefore constant as soon as $\alpha \geq \frac{C_2}{C_1+C_2}$. Note also that it is a continuous function of α.

On $[0, \frac{C_2}{C_1+C_2}]$, function R_2 is increasing due to the non-increasingness of v. The optimal choice is therefore any value $\alpha \in [\frac{C_2}{C_1+C_2}, 1]$, all producing the same revenue $v(C_1 + C_2)C_2$.

It is interesting to us to remark that there is no increase in revenue after the threshold $\alpha = C_2/(C_1+C_2)$ is reached. Indeed, in that case, demand and optimal price will be the same whatever α. On the other hand, from practical reasons, $\alpha = \frac{C_2}{C_1+C_2}$ is the most relevant choice because of more limited interferences and power consumption.

4 Optimal Radius/Proportion Parameter with Transmission Power Cost

We now assume that increasing the transmission power P induces a cost linear in that power, βP with β a constant. An important remark is that for the pricing game with fixed α, this has no consequence because it is a constant cost and therefore does not change the results.

If R is the radius such that provider 2 can transmit with a minimal reception power P_{min} for a given QoS, and assuming without loss of generailty that the coverage radius of provider 1 equals 1, we have $\alpha = \pi R^2/\pi = R^2$.

Similarly, we assume that for a user located at a distance d of the antenna, the reception power is $c\frac{P}{d^\mu}$ with c and μ constants (the value of μ depends on the area -countryside, city...-, but generally $2 \leq \mu \leq 5$). In order to fulfill a required minimal value P_{min} at reception, the relation between power and radius is

$$P_{min} = c\frac{P}{R^\mu}$$

i.e., $R^\mu = \frac{c}{P_{min}}P = \alpha^{\mu/2}$, which yields $P = \frac{P_{\min}}{c}\alpha^{\mu/2}$.

The goal is therefore to find the value α maximizing the overall benefit B_2, that is the revenue (11) at Nash equilibrium minus the power cost:

$$\begin{aligned} B_2(\alpha) &= R_2(\alpha) - \beta P \\ &= C_2 v\left(\frac{C_2}{\alpha}\right) - \frac{\beta P_{min}}{c}\alpha^{\mu/2}. \end{aligned}$$

We then have the following result.

Proposition 1. *If v is derivable and concave on its support, and $\mu \geq 2$, there is a unique solution $\alpha^* \in [0,1]$ for optimizing the net revenue $B_2(\alpha)$ of provider 2. Moreover, $\alpha^* \in [0, \frac{C_2}{C_1+C_2}]$*

Proof. Any optimal value of α is necessarily in $[0, \frac{C_2}{C_1+C_2}]$ because revenue is constant in $[\frac{C_2}{C_1+C_2}, 1]$ while cost strictly increases.

On $[0, \frac{C_2}{C_1+C_2}]$, the maximization problem writes

$$\max_{\alpha \in \left[0, \frac{C_2}{C_1+C_2}\right]} v\left(\frac{C_2}{\alpha}\right) C_2 - \frac{\beta P_{min}}{c} \alpha^{\mu/2}. \tag{12}$$

The derivative of the objective function is

$$-(1/\alpha^2) v'\left(\frac{C_2}{\alpha}\right)(C_2)^2 - \frac{\beta \mu P_{min}}{2c} \alpha^{\mu/2-1},$$

which is strictly decreasing in α on the support of $v(C_2/\cdot)$, thus the objective function is strictly concave on that support. Remark that possible values beyond the support of $v(C_2/\cdot)$ are not of interest since the associated objective is strictly negative.

Trying to find if there is a value of α for which that derivative is zero gives $-v'\left(\frac{C_2}{\alpha}\right) = \frac{\beta \mu P_{min}}{2cC_2^2} \alpha^{\mu/2+1}$. Due to the strict concavity of the objective, if an interior solution exists then it is unique. Otherwise, the derivative is always of the same sign, meaning that the optimal value is obtained at one of the extremities of the interval $[0, \frac{C_2}{C_1+C_2}]$.

Proposition 1 establishes that only one α will be chosen by provider 2. We now investigate the consequences of that choice on the price perceived by providers. More precisely, from (9) and (10) we know that at the equilibrium of the price war, the perceived prices $\bar{p}_1$ and $\bar{p}_2$ respectively equal the real prices p_1 and p_2, and that

- if $\alpha \geq \frac{C_2}{C_1+C_2}$ then $p_1 = p_2 = v(C_1 + C_2)$;
- if $\alpha < \frac{C_2}{C_1+C_2}$ then $p_1 = v\left(\frac{C_1}{1-\alpha}\right) > p_2 = v\left(\frac{C_2}{\alpha}\right)$.

In other words, when provider 2 really provides service (i.e. the optimal value α^* is strictly positive), then if the optimization problem (12) has an interior solution, the equilibrium perceived price in zone B, namely p_2, is strictly below the equilibrium perceived price p_1 in zone A. Otherwise, the perceived price on both zones is the same and equals $v(C_1 + C_2)$.

Those considerations are summarized in the next proposition.

Proposition 2. *Assume that provider 2 has an interest in providing service, i.e. that there exists $\alpha \in \left(0, \frac{C_2}{C_1+C_2}\right]$ such that*

$$v\left(\frac{C_2}{\alpha}\right) C_2 - \frac{\beta P_{min}}{c} \alpha^{\mu/2} > 0.$$

Then,

- *either* $-v'(C_1+C_2) \geq \frac{\beta\mu P_{\min}}{2c} C_2^{\mu/2-1}(C_1+C_2)^{-\mu/2-1}$, *therefore* $\alpha^* = \frac{C_2}{C_1+C_2}$ *and users in zones A and B perceive the same price at equilibrium;*
- *or* $-v'(C_1+C_2) < \frac{\beta\mu P_{\min}}{2c} C_2^{\mu/2-1}(C_1+C_2)^{-\mu/2-1}$, *therefore* $\alpha^* < \frac{C_2}{C_1+C_2}$, *and users in zone B will all choose provider* 2 *and experience a strictly lower price than zone A users.*

Proof. We just express the condition for the derivative of the objective function in (12) to be positive or negative at $\alpha = \frac{C_2}{C_1+C_2}$.

Example. Assume that $v(q) = 10 - q$ over [0,10], $C_1 = C_2 = 1$, $\mu = 2$ and $\frac{\beta P_{min}}{c} = 16$ for convenience.

The objective function $R_2(\alpha) - \frac{\beta P_{min}}{c}\alpha^{\mu/2}$ becomes on $[0, 1/2]$

$$10 - \frac{1}{\alpha} - 16\alpha.$$

The optimal value of $R_2(\cdot)$ is 2, obtained at $\alpha^* = 1/4 < \frac{C_2}{C_1+C_2}$. For that value of α, the price war among providers leads to the prices $p_1 = v(C_1/(1-\alpha)) = 26/3 \simeq 8.67$ and $p_2 = v(C_2/\alpha) = 6$. Therefore all users in zone B choose (the cheaper) provider 2.

5 Conclusion

In this paper, we have worked on determining the optimal proportion of customers of a concurrent that a smaller provider should try to reach in order to maximize its revenue. We have shown, based on results on the price war from a previous paper, that if there is no cost associated to power consumption, there is a threshold over which revenue stops increasing. It means that there is no need to install competition on the whole domain, i.e. for all customers. When increasing the transmission range induces a cost, assuming concavity of the marginal valuation function, we have proved the existence and uniqueness of the optimal range value. Determing it can be done very simply numerically.

We plan to work further on that kind of model. What happens if one of the providers is not included in the domain of the other? What if there is uncertainty on demand?

Acknowledgment

This work was supported by Euro-NF Network of Excellence and the French research agency through the CAPTURES and Winem projects.

References

1. Courcoubetis, C., Weber, R.: Pricing Communication Networks—Economics, Technology and Modelling. Wiley, Chichester (2003)
2. DaSilva, L.: Pricing of QoS-Enabled Networks: A Survey. IEEE Communications Surveys & Tutorials 3 (2000)
3. Tuffin, B.: Charging the Internet without bandwidth reservation: an overview and bibliography of mathematical approaches. Journal of Information Science and Engineering 19, 765–786 (2003)
4. Gibbens, R., Mason, R., Steinberg, R.: Internet service classes under competition. IEEE Journal on Selected Areas in Communications 18, 2490–2498 (2000)
5. El-Azouzi, R., Altman, E., Wynter, L.: Telecommunications network equilibrium with price and quality-of-service characteristics. In: Proc. of 18th International Teletraffic Congress, Berlin, Germany (2003)
6. Acemoglu, D., Ozdaglar, A.: Competition and efficiency in congested markets. Mathematics of Operations Research 32, 1–31 (2007)
7. Acemoglu, D., Ozdaglar, A.: Price competition in communication networks. In: Proc. of IEEE INFOCOM (2006)
8. Hayrapetyan, A., Tardos, E., Wexler, T.: A network pricing game for selfish traffic. In: Proc. of IEEE PODC (2006)
9. Wardrop, J.: Some theoretical aspects of road traffic research. Proceedings of the Institute of Civil Engineers 1, 325–378 (1957)
10. Maillé, P., Tuffin, B.: Price war in heterogeneous wireless networks (2008) (submitted), `http://www.irisa.fr/dionysos/pages_perso/tuffin/Publis/PM-BT-heterogeneous.pdf`
11. Marbach, P.: Analysis of a static pricing scheme for priority services. IEEE/ACM Transactions on Networking 12, 312–325 (2004)
12. Maillé, P., Tuffin, B.: Analysis of price competition in a slotted resource allocation game. In: Proc. of IEEE INFOCOM (2008)
13. Fudenberg, D., Tirole, J.: Game Theory. MIT Press, Cambridge (1991)

On Competition for Market Share in a Dynamic ISP Market with Customer Loyalty: A Game-Theoretic Analysis

László Gyarmati and Tuan Anh Trinh

High Speed Networks Laboratory
Department of Telecommunication and Media Informatics
Budapest University of Technology and Economics
2 Magyar tudósok körútja, H-1117, Budapest, Hungary
{gyarmati,trinh}@tmit.bme.hu

Abstract. Customer loyalty as part of user behaviour has significant impact on the Internet Service Providers' (ISPs) price setting strategies as shown recently in [1,2,3,4]. However, the issue of a dynamic ISP market, where new ISPs enter the market and try to increase their market shares by offering favourable access prices for incumbent ISPs' loyal customers, has not been addressed yet. Furthermore, the cost of entrance is not yet properly dealt with in the previous studies. In this paper, we use the tools from game theory to understand the competition for market share in a dynamic ISP market with customer loyalty. We model the situation by a Stackelberg leader-follower game, and use the model to compute the Nash/Stackelberg equilibria of the game with customer loyalty and different cost models. For simple cost models, we give explicit formulas for the equilibria of the games. For more complex cost models, we use approximation and simulations to illustrate the dynamics of market shares in these situations.

1 Introduction

The economic interactions among service providers of different levels and end-users have been in the focus of interest for several years. These interactions will continue to get special attention, since initiatives like the NSF FIND [5] and Euro-NF [6] promote economic incentives as a first-order concern in future network design. In addition, user behaviour also has a significant impact on the design of next-generation network architectures as well as creating profitable services running them. Also, decision-makers trying to work out a plausible solution for the recently surfaced net neutrality debate would greatly benefit from an in-depth understanding of economic processes inside the user-ISP hierarchy. There is broad literature in the area of modelling interactions between ISPs with game-theoretical means [7,8,9]. While these papers introduce and analyse complex models for the interaction of ISPs at different levels of hierarchy, they mostly assume a very simple user behaviour model when investigating the market for local ISPs: end-users choose the cheapest provider assuming that the

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 11–23, 2009.

quality of the certain services is the same. This assumption could be plausible in certain scenarios, but it could be misleading if there are loyal customer segments present in the market. A vivid example of customer loyalty in practice is the loyalty contract between a service provider and a customer. The customers are charged with different price if they sign a contract and this difference depends on the length of the contract! In [10] authors use a game-theoretic framework to prove that if loyalty is an additional product of market share and penetration, customer retention strategies seem to be consequently more efficient for market leaders. Another study [11] analyses a duopolistic price setting game in which firms have loyal consumer segments, but cannot distinguish them from price sensitive consumers. [12] presents a duopolistic price setting game, where loyal and also disloyal customers are on the market. The companies set prices based on the number of their loyal customers, therefore the Nash equilibrium of the game changes resulting higher utilities.

A number of empirical studies deal with user loyalty on the ISP markets, we shortly review some of them as a global picture. The 2005 Walker Loyalty Report for Information Technology shows that 38 percent of the enterprise customers in the USA have been truly loyal to their Internet Service Providers [13], while Choice survey states that 90 percent of the household respondents had not changed their ISP in the previous 12 months including contract-users as well [14]. [1] states that customer loyalty towards ISPs does exist in Taiwan too. National communication authorities of European Union's countries carry out market research dealing with customer loyalty toward local ISPs. UK's Office of Communications' 2008 Communications Market Report states that 27 percent of broadband users have already switched at least once their provider. Furthermore, the dynamics of the ISP market is illustrated with real market shares, e.g. BSkyB entered the UK ISP market in 2005, two years later it had almost 10 percent of the connections [15]. The consumer ICT survey of Commission for Communications Regulation of Ireland reveals that 84 percent have not change their ISP in the last 12 months [16]. 81 percent of broadband customers said that they do not intend to change Internet operator in the next 12 months finds out Anacom of Portugal [17]. According to a recent report [18], in Finland 16 percent of the subscribers have switched their ISP in 2007 mainly because of a better offer from a competitor. The Malta Communications Authority' survey states that 84 percent of respondents have not switch their Internet Service Provider in the last two years [19]. We carried out a survey on the customer loyalty issue for the Hungarian ISP market, 60 percent of the questioned people do not change their ISP in the last five years. The analysis of our own survey shows that subscriber loyalty depends on the price difference of the current and the possible future service providers, users would become disloyal if the price difference is large enough.

These works and recent works [2,3,4] initiate the discussion on customer loyalty and its impact on pricing strategies of ISPs. However, a number of issues are still to be solved. First, these works use simple or any cost models, usually zero cost is supposed for every Internet subscription, which does not describe

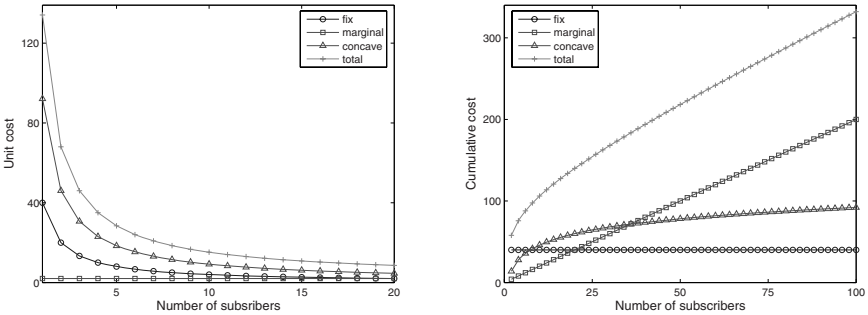
the costs which ISPs have to face (e.g. fixed and variable costs). Second, these works deal with static market scenarios when fixed number of service providers exist on the market and they compete for the customers, to our best knowledge the dynamics of the ISP market has not been examined yet. From the arguments above, in this paper we address these issues and try to give answers to these open questions. The ISP market changes dynamic if a new service provider enters to it. The new as well as the incumbent ISPs have to select their price setting strategies based on the costs of providing Internet access, on the cost of entering and on the current market situation. To model the dynamics of the ISP pricing competition, we use Stackelberg leader-follower game. Based on our game-theoretic model we quantify the effects of dynamics and costs on ISPs' price setting strategies and market shares.

The paper is structured as follows. First, in Section 2, we summarize the basics of Internet Service Providers' costs and propose an ISP cost model. Section 3 provide a game-theoretic model for pricing Internet access on a dynamic ISP market in particular when a new ISP wants to start providing access. We use Stackelberg leader-follower game to calculate the equilibrium market shares and profits of the ISPs. We present equilibrium strategies of a dynamic market with loyal customers using different cost models in Section 4, a simple and a full cost model will be examined. In addition, we quantitatively illustrate the changes based on simulation results. Finally, Section 5 concludes the paper.

2 Modelling Cost of Providing Internet Access

Providing local Internet access is a typical example for the so called economics of scale principle, which means cost advantages that a firm obtains due to expansion, because ISPs have to deploy a fixed cost large-scale transporting network. In this section we review the basics of ISPs' costs and present our ISP cost model. The cost of providing Internet access has significant effect on ISPs' pricing strategies because the prices have to be selected to cover the expenses.

The operators' costs can be partitioned based on several aspects [20]: timing (historical or current cost), association (direct, indirect, shared) or its behaviour as production increases. For our goal this last one is the most relevant because on a dynamic ISP market the amount of costs belonging to different number of subscribers determines the pricing decisions of ISPs. On the one hand, costs can be *fixed* (or volume insensitive), if they are constant for a given range of subscribers. The value of fixed costs may vary in long term, e.g. the provider decides to buy a new facility. On the other hand, *variable costs* depend on the number of users, in case of networking services the functions of variable costs are usually decreasing. In addition, the same cost can belong to different categories based on the transporting media. For example, transmission and infrastructure cost is variable in case of wireline media based on FCC's HCMP (Hybrid Cost Proxy Model) tool [21], but it can be considered as fixed cost in case of a licensed spectrum wireless service. A cost model of an ISP can include the followings: Capital Expenditures (CAPEX) e.g. deployment costs, access installation,



(a) Unit costs (b) Cumulative costs

Fig. 1. The different types of cost of Internet Service Providers

Operational Expenditures (OPEX) including maintenance, network management, billing, Internet transit cost, commercial, costumer care costs, etc.

Numerous publications examine the costs of service providers and present the relation of the number of the subscribers and the total cost, including [22,23,24,25]. [22] gives a comprehensive description of the costs of Internet Service Providers and presents ISP specific market and provider segmentation techniques and accounting methods. In addition, a new ISP cost model is proposed and applied to different ISP scenarios. [23] describes relevant Operational Expenditures elements, these elements are modelled as functions of parameters, e.g. number of customers or network devices. The presented grouped list of OPEX elements can be used to perform business case analysis. [24] investigates the costs of Internet access using cable modem and ISDN. Furthermore, it proposes a ISP cost model and based on detailed simulation results the effect of the number of subscribers on total costs is quantified for both technologies. [25] presents economic considerations of FTTX deployment, similar decreasing unit cost figures are shown in the case studies.

A common finding of these papers is the shape of the total cost function of a service provider, which is similar to a logarithmic function. However, none of these works suggest a cumulative cost function with few parameters, they either present only simulation results, or propose formulas for single type of costs. Therefore, we model the costs of the ISPs with the $F_c(n) = nc + c_l(n) + C_{fix}$ cost function where n is the number of subscribers and the three cost types are as follows:

- c marginal cost - constant cost of a subscriber
- $c_l(n)$ concave cost - a logarithmic function of the number of users
- C_{fix} fix cost - independent from the number of subscribers

In practice, the magnitude of these cost types are usually not equivalent, based on the specific scenarios, one type could be considered relatively negligible to other types. In particular, in this paper we deal with two models of ISP costs:

first, we examine the dynamic ISP market with only marginal costs as a simple model to gain insights, then the full cost model will be assumed. We illustrate the three type of costs on Figure 1. Figure 1(a) plots the unit costs of a subscriber as the number of the users increases, while Figure 1(b) shows the cumulative costs of a provider.

3 Dynamic ISP Market with Loyal Customers as a Stackelberg Game

The market of Internet Service Providers is open, namely a new firm can enter or left the market easily. ISPs have loyal customers who usually do not change their providers. It is an interesting situation when a new ISP enters the market, because the entrant ISP wants to get customers from the incumbent companies. In this paper we examine those ISP market scenarios where the number of the service providers increases.

We model the entry situation in the following way: there are N Internet Service Providers on the market and a new ISP wants to enter to provide Internet access for customers. The cost of entry is C which represents the capital expenditures (CAPEX), e.g. the price of network equipments, the cost of facilities, the cost of access devices. We suppose, that these costs have to be paid once, when the firm enters the market.

The local Internet access market is modelled as follows. The demand for Internet access is constant until an α reservation price, everyone can afford to have Internet access if there exists an ISP whose price is lower than α. Every ISP has customers who are loyal to their provider, formally ISP_i has l_i loyal subscribers. Everyone who has an Internet subscription belongs to an ISP, there do not exist independent users on the market. Different ISP customer loyalty models have been proposed earlier. [3] presents double-price reservation price loyalty models, both deterministic and stochastic one. Price difference dependent loyalty models are proposed by [4] including step based, threshold based and uniformly distributed models. We will use the last model in this paper, i.e. loyal customers are price difference dependent namely if there exist a cheaper provider a fraction of the customers leave their current ISP. We model the loyalty of the ISP_i's subscribers with a linear function, a provider looses $L_i = \frac{p_i - p_j}{\alpha} l_i$ customers if ISP_j has a lower price ($p_j < p_i$). We suppose that every ISP plays rationally, i.e. selects its profit maximising strategy.

We model the entry situation with a sequential game. First, the entrant ISP selects an action, she can enter to the market or not. If the entrant ISP decides to enter, the incumbent providers have two possibilities: set a low enough price to keep all of her customers or set a profit maximizing price. On Fig. 2 we illustrate the two-player version of the game with the utilities of the cases both in extensive and strategic form. The calculations of the profits will be presented later in this paper. We can see that the threat of ISP_1, set c as price, is non-credible, because if the new ISP enters the market the incumbent can have higher payoff if she do not play this option. Accordingly, if it is worth for ISP_2 to enter, the ISPs will

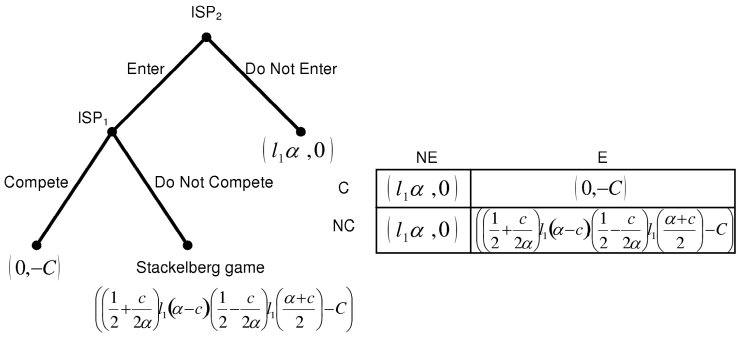


Fig. 2. Extensive and strategic form of entering game

play a Stackelberg leader-follower price setting game where ISP_1 is the leader of the game and the entrant ISP_2 is the follower. ISP_1 sets her price first when the entry of ISP_2 turns out, after that ISP_2 selects her own price. ISP_2 decides about the entry based on the number of users who she will have after the entry, if ISP_2 can earn at least C she will enter the market.

4 Game-Theoretic Analysis of Dynamic ISP Market with Customer Loyalty

In this section we apply the proposed cost, loyalty and dynamic ISP market models in a game-theoretical analysis. We examine the dynamic ISP market, where loyal customers exist, using different cost models. We start with a simple cost model where only constant marginal cost exists to gain insights of the problem then we extend the simple model to a full ISP cost model.

4.1 Dynamic ISP Market with Linear Cost

In this section we use a simple cost function to model the dynamic ISP market, where only marginal cost exists, namely the cost of providing Internet access for a single user is c. We examine different market scenarios through this section, we first examine the case when an ISP enters a monopolistic market. After that we deal with a market where N ISPs exist and a new ISP enters, finally we discuss cases where more than one ISPs enter to a market.

A simple model: one incumbent - one entrant ISP. First, only ISP_1 exists on the market who has l_1 loyal customers which is the whole demand. The cost of providing Internet access for a single user is c, while α denotes the largest price, at customers still buy Internet access. ISP_2 has to set always a smaller price than ISP_1 otherwise she would not have any customers. The payoff function of the ISPs is $\Pi_i = l_i^*(p_i - c)$, where l_i^* denotes the number of the ISP's customers at

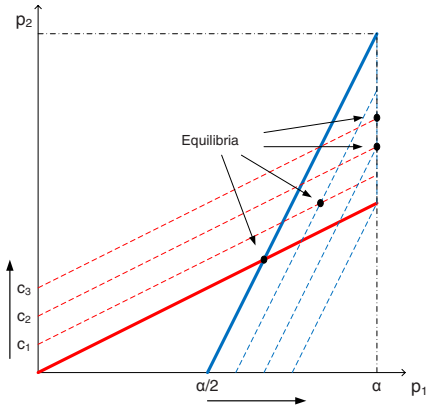


Fig. 3. Best response function of one incumbent and one entrant ISP game at different costs

the equilibrium. ISP_1 will have $l_1^* = \left(1 - \frac{p_1 - p_2}{\alpha}\right) l_1$ users after ISP_2 entered the market, while ISP_2 will have $l_2^* = \frac{p_1 - p_2}{\alpha} l_1$ customers. Accordingly, the payoff of the service providers can be expressed as $\Pi_1 = \left(1 - \frac{p_1 - p_2}{\alpha}\right) l_1 \left(p_1 - c\right)$ and $\Pi_2 = \frac{p_1 - p_2}{\alpha} l_1 \left(p_2 - c\right)$.

Based on the payoff functions the best response function of the ISPs can be calculated, the proof yields $BR_1 = \frac{c + \alpha + p_2}{2}$ and $BR_2 = \frac{c + p_1}{2}$ as best response functions. We show the best response functions and illustrate the effect of the cost (c) on Figure 3.

The ISPs play leader-follower game, where ISP_1 is the leader, she selects her price based on the best response function of the entrant ISP, after that ISP_2 - the follower - sets her price. $p_{max} = \alpha + c$ is the optimal price which maximize $\max_{p_1} \Pi_1(p_1, p_2) = \left(1 - \frac{p_1 - BR_2(p_1)}{\alpha}\right) l_1(p_1 - c)$, the payoff of ISP_1. ISP_1 can not set $\alpha + c$ as price because in this case none of the subscribers would buy Internet access from her. Thus, ISP_1 will set the highest possible price α in order to have maximal payoff. The follower ISP_2 calculates the best response price for this, which is $p_2 = \frac{\alpha}{2} + \frac{c}{2}$. Using these Stackelberg equilibrium prices we calculate the market shares and the payoffs:

$$
\begin{array}{ll}
p_1^S = \alpha & p_2^S = \frac{\alpha}{2} + \frac{c}{2} \\
l_1^S = \left(\frac{1}{2} + \frac{c}{2\alpha}\right) l_1 & l_2^S = \left(\frac{1}{2} - \frac{c}{2\alpha}\right) l_1 \\
\Pi_1 = \left(\frac{1}{2} + \frac{c}{2\alpha}\right) l_1 (\alpha - c) & \Pi_2 = \left(\frac{1}{2} - \frac{c}{2\alpha}\right) l_1 \left(\frac{\alpha}{2} + \frac{c}{2}\right)
\end{array}
$$

The results illustrate the effect of a newly entering Internet Service Provider. We note that it can be showed that the property of the Stackelberg games holds also in this case, namely the payoff of the leader at the leader-follower game is larger or equals to the Nash equilibrium payoff.

Generalized model: N incumbents - one entrant ISP. In the previous section we have seen the payoffs and the market shares if a new ISP enters to a monopolistic market. Usually there exist more than one ISP on the market when a new ISP wants to enter. In the followings we suppose that the incumbent service providers do not increase the number of their subscribers, they do not grab users from each other. Only the entrant ISP will have new customers from the incumbent companies because her price has to be the smallest otherwise she would not have any subscribers because switching subscribers select the cheapest offer. Accordingly, in the following game the entrant ISP will set the smallest price.

There exist $i = 1, \ldots, N$ ISPs on the market who are selling Internet access to their customers, ISP_i has l_i loyal customers. The new ISP_j enters the market, she does not have any subscribers at the beginning. The service providers play a leader-follower game where the incumbent companies are the leaders, they set their prices first, while the entrant ISP is the follower of the game. The payoff functions of the service providers are as follows:

$$\Pi_i = \left(1 - \frac{p_i - p_j}{\alpha}\right) l_i \left(p_i - c\right) \quad i = 1, \ldots, N \tag{1}$$

$$\Pi_j = \sum_i \frac{p_i - p_j}{\alpha} l_i \left(p_j - c\right) \tag{2}$$

The best response function of the entrant ISP maximizes her payoff, it can be expressed as $BR_j(p_1, \ldots, p_n) = \frac{c}{2} + \frac{\sum_i l_i p_i}{2 \sum_i l_i}$. The incumbent ISPs maximize their profits based on the entering ISP's best response function:

$$\begin{aligned}
\max_{p_i} \Pi_i(p_i, p_j) &= \max_{p_i} \left(1 - \frac{p_i - p_j}{\alpha}\right) l_i \left(p_i - c\right) = \\
&= \max_{p_i} \left(1 - \frac{p_i - BR_j(p_1, \ldots, p_n)}{\alpha}\right) l_i \left(p_i - c\right) = \\
&= \max_{p_i} \left(1 - \frac{p_i - \left[\frac{c}{2} + \frac{\sum_i l_i p_i}{2 \sum_i l_i}\right]}{\alpha}\right) l_i \left(p_i - c\right)
\end{aligned}$$

The calculations yield the following implicit expressions for the Stackelberg equilibrium prices:

$$p_k^* = \frac{(2\alpha + c) \sum_i l_i + \sum_{i \neq k} l_i p_i}{2 \sum_i l_i + l_k} \frac{1}{2} + \frac{c}{2} \quad k = 1, \ldots, N \tag{3}$$

$$p_j^* = \frac{c}{2} + \frac{\sum_i l_i p_i}{2 \sum_i l_i} \tag{4}$$

The equilibrium prices produce a system of linear equations, where the variables are the prices of the ISPs. The equilibrium prices of the Stackelberg ISP price setting game solve the following system of linear equations:

$$\begin{pmatrix} 1 & -\frac{1}{2}\frac{\sum_{i\neq 2} l_i}{2\sum l_i - l_2} & \cdots & -\frac{1}{2}\frac{\sum_{i\neq n} l_i}{2\sum l_i - l_N} & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ -\frac{1}{2}\frac{\sum_{i\neq 1} l_i}{2\sum l_i - l_1} & \frac{1}{2}\frac{-\sum_{i\neq 2} l_i}{2\sum l_i - l_2} & \cdots & 1 & 0 \\ \frac{l_1}{2\sum l_i} & \frac{l_2}{2\sum l_i} & \cdots & \frac{l_N}{2\sum l_i} & 1 \end{pmatrix} \begin{pmatrix} p_1^* \\ \vdots \\ p_N^* \\ p_j^* \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\frac{(2\alpha+c)\sum l_i}{2\sum l_i - l_1} + \frac{c}{2} \\ \vdots \\ \frac{1}{2}\frac{(2\alpha+c)\sum l_i}{2\sum l_i - l_N} + \frac{c}{2} \\ \frac{c}{2} \end{pmatrix}$$

The equilibrium prices determine the new market shares of the Internet Service Providers:

$$l_i^* = \left(1 - \frac{p_i^* - p_{n+1}^*}{\alpha}\right) l_i \quad i = 1, \ldots, N \tag{5}$$

$$l_j^* = \sum_i \frac{p_i^* - p_{n+1}^*}{\alpha} l_i \tag{6}$$

The equilibrium profits of the ISPs are the product of the market shares and the equilibrium prices, namely $\Pi_k^* = l_k^*(p_k^* - c), k = 1, \ldots, N, j$. In Section 4.3 we present simulation results of different market scenarios.

The previously introduced price setting strategies can also be used if more than one ISPs enter the market. In case of providing Internet access the new firms do not have to win a concession tender to start their businesses. Accordingly, we model the situation when more than one ISPs enter the market iteratively. We suppose that the entering firms do not enter the market on a flag day, but they can be order in time. Using this assumption we use our one entrant ISP model iteratively: a new company enters the market, we calculate the new market shares then the next ISP enters the market, etc.

4.2 Dynamic ISP Market with Full Cost

We have seen the equilibrium strategies of ISPs if they have simple costs. In this section we generalize the cost model, thus we examine the dynamic ISP market, where the ISPs have costs based on the full model presented in Section 2. The incumbent ISPs ($i = 1, \ldots, N$) has l_i loyal users, the entrant ISP$_j$ does not have any users at start, α is the maximal price at Internet access can be sold. The payoff function of the ISPs are as follows:

$$\Pi_i = \left(1 - \frac{p_i - p_j}{\alpha}\right) l_i \left(p_i - c\right) - C_{fix} - \ln\left[\left(1 - \frac{p_i - p_j}{\alpha}\right) l_i\right], i = 1, \ldots, N \tag{7}$$

$$\Pi_j = \sum_i \frac{p_i - p_j}{\alpha} l_i \left(p_j - c\right) - C_{fix} - \ln\left(\frac{p_i - p_j}{\alpha} l_i\right) \tag{8}$$

The incumbent ISPs are the leaders of the game, while the entrant company is the follower. It is interesting, that even on a monopolistic market, the equilibrium prices can not be formulated in closed forms. The best response

of the entrant still can be expressed, it is a quadratic equation. The incumbent ISP's best response function is a quartic equation which does not have closed form solution. If the number of incumbents increases the formula gets complicated, the derivative of the entrant ISP's profit is

$$\Pi'_j = \sum_i \frac{\alpha + l_i(c + p_i - 2p_j)(p_i - p_j)}{\alpha(p_i - p_j)}$$

which contains the prices of all incumbent ISPs. Accordingly, the best response of an incumbent ISP is a non-linear function of the best response of all the other ISPs as well because they are included in the entrant's best response function.

Accordingly, how can we model the dynamic ISP market? The first solution is that we numerically express approximating implicit equations for the equilibrium prices. In this case, the incumbent ISPs have quartic equations while the entrant ISP has quadratic one. Then systems of equations have to be made with the combination of the implicit forms, the number of the systems are $2 \cdot 4^N$ which is exponential in the number of incumbents. These are systems of non-linear equations therefore can be computed with approximation methods. At the second solution we define an order of the incumbent IPSs as a leader-follower chain. First, the entrant ISP is expressed because she is the only follower in the game. After that the lowest ordered incumbent is the follower and will be expressed, etc. Using the derived formulas the equilibrium prices can be calculated numerically. Finally, a simpler cost model can be used to model the dynamics of the ISP market as we have done in Section 4.1.

4.3 Simulation Results

We use Matlab to calculate market shares, prices and profits of dynamic ISP markets. For the simple cost model we solve the system of linear equation for the equilibrium prices. First, we present results on the change of market shares. Let us suppose that there are three incumbent ISPs on the market with different

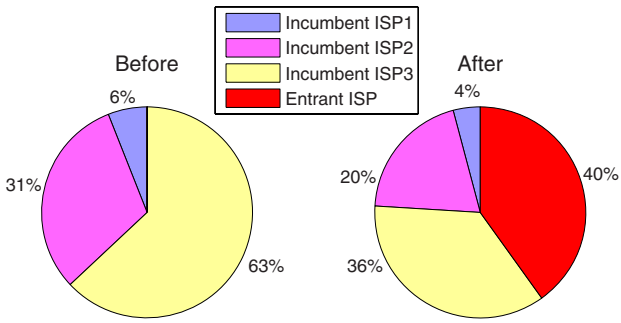


Fig. 4. Market shares before and after an ISP entered a three-incumbent ISP market with simple cost

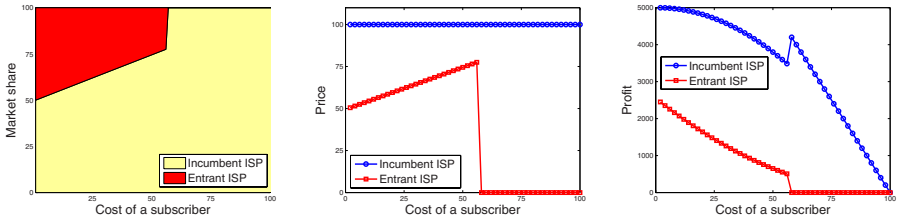


Fig. 5. The impact of unit cost on market shares, prices and profits if an ISP enters a monopolistic market with simple cost

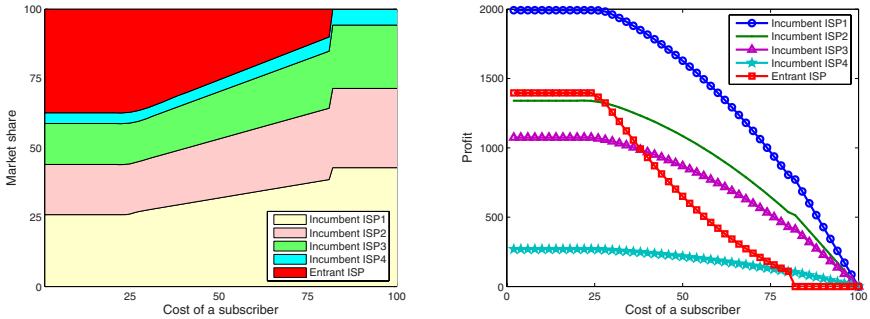
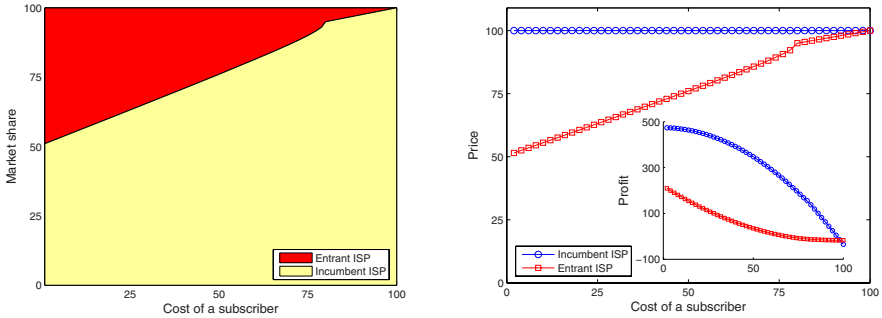


Fig. 6. The impact of unit cost on market shares, prices and profits if an ISP enters a competitive market (four incumbents) with simple cost



(a) Market shares (b) Prices and profits

Fig. 7. Dynamic ISP market with one incumbent and one entrant ISP with full cost

market shares (Fig. 4) and one new ISP enters. In this scenario the entrant ISP can grab a lot of subscribers thus she will be the largest ISP on the market. It can be seen that because the ISPs have the same linear loyalty function, the incumbents loose users based on the ration of their market shares.

We illustrate the effect of the costs on dynamic, simple cost ISP markets. We model a monopolistic market on Fig. 5 where the market shares (Fig. 5(a)), the prices (Fig. 5(b)) and the profits (Fig. 5(c)) are presented as a function of the marginal cost. The cause of the brakes is the cost of the entry (C), if the entrant ISP can not have a profit that covers the entry cost, she will not enter the market. A more competitive ISP market is shown on Fig. 6 where four incumbents exist on the market. We can see similar trends in the market shares and profits as we have seen on the monopolistic market.

We use the ordered incumbents method to calculate the effect of dynamics in case of a full cost, monopolistic ISP market. We have calculated the profit maximising prices with Matlab, the effects of increasing unit cost on market shares, profits and prices are presented on Fig. 7. As we can see the market shares (Fig. 7(a)) and the prices (Fig. 7(b)) are almost linear functions of the cost of subscribers, only around 80 are non-linear parts because of the logarithmic function. The entrant ISP has always smaller payoffs even if the market shares are almost the same, because her prices are always lower than the incumbent's in order to grab customers.

5 Conclusion

In this paper, we have demonstrated how ISPs price Internet access in a dynamic ISP market with customer loyalty, taking in account the impact of dynamics on prices, profits and market shares. We have provided an overview about customer loyalty on global ISP markets. The analysis of our own survey as well as the results from other empirical researches showed that customer loyalty in ISP markets exists. In addition, we have examined the different costs of Internet Service Providers and we have shown which type of expenses are fixed and variables costs. Local ISP markets are dynamic, the number of ISPs and their market shares change continually. We have created a game theoretic model which handles the situation when a new company wants to enter the market. Using Stackelberg's leader-follower game we have shown how much the market shares and the profits of the ISPs have changed on the dynamic market. Furthermore the effect of different costs have been illustrated, both analytically and quantitatively.

As a future research, we plan to investigate furthermore the dynamic ISP market and the effects of costs, where the service providers can have sophisticated price setting strategies (e.g. when a new ISP enters the market, incumbent ISPs select their strategies taking into account the profits of forthcoming months).

References

1. Chiou, J.S.: The Antecedents of Consumers' Loyalty Toward Internet Service Providers. Inf. Manage. 41(6), 685–695 (2004)
2. Trinh, T.A.: Pricing Internet Access for Disloyal Users: a Game-Theoretic Analysis, extended abstract. In: Workshop on Socio-Economic Aspects of Next Generation Internet, Sweden (2008)

3. Biczók, G., Kardos, S., Trinh, T.A.: Pricing Internet Access for Disloyal Users: a Game-Theoretic Analysis. In: SIGCOMM 2008 Workshop on Economics of Networked Systems (2008)
4. Gyarmati, L., Trinh, T.A.: How to Price Internet Access for Disloyal Users under Uncertainty. Annals of Telecommunications (submitted)
5. NSF: Future Internet Network Design Initiative, `http://find.isi.edu`
6. Euro-NF: Network of Excellence on the Network of the Future, `http://euronf.enst.fr/en_accueil.html`
7. He, L., Walrand., J.: Pricing and Revenue Sharing Strategies for Internet Service Providers. In: IEEE Infocom (2005)
8. Cao, X.-R., Shen, H.-X., Wirth, R.M.: Internet Pricing with a Game Theoretical Approach: Concepts and Examples. In: IEEE/ACM ToN, pp. 208–216 (2002)
9. Shakkottai, S., Srikant, R.: Economics of Network Pricing with Multiple ISPs. In: IEEE Infocom (2005)
10. McGahan, A.M., Ghemawat, P.: Competition to Retain Customers. Marketing Science 13(2), 165–176 (1994)
11. Deneckere, R., Kovenock, D., Lee, R.: A Model of Price Leadership Based on Consumer Loyalty. The Journal of Industrial Economics 40(2), 147–156 (1992)
12. Georges, C.: Econ 460: Game Theory and Economic Behavior. Course handouts and exercises, Department of Economics, Hamilton College (2008)
13. Walker: The 2005 Walker Loyalty Report for Information Technology (2005)
14. Choice: ISP Satisfaction Survey (2007)
15. Ofcom - Office of Communications: The Communications Market (2008)
16. Comreg - Commission for Communications Regulation: Consumer ICT Survey (2008)
17. ANACOM: Survey on the use of broadband (2006)
18. FICORA - Finnish Communications Regulatory Authority: Market Review (2007)
19. MCA- Malta Communications Authority: End-users perception survey - Broadband Services (2007)
20. Azcoitia, S.: Next Generation Networks Techno-Economic Insights. In: Networks Tutorial, Budapest, Hungary (2008)
21. Federal Communications Commission: Hybrid Cost Proxy Model (HCPM) (2003)
22. Reichl, P., Kurtansky, P., Gerke, J., Stiller, B.: ISP Cost Model (ICOMO) Design Market Managed Multi-service Internet (2000)
23. ECOSYS Project: OPEX Models, Celtic Initiative (2005)
24. Gillet, S.E.: Connecting Homes to the Internet: An Engineering Cost Model of Cable vs. ISDN, M.S. thesis. MIT (1995)
25. El-Sayed, M.: Fiber-to-the-X: Technologies & Economics. In: Networks Tutorial, Budapest, Hungary (2008)

A Pricing Model for a Mobile Network Operator Sharing Limited Resource with a Mobile Virtual Network Operator

Hélène Le Cadre[1], Mustapha Bouhtou[1], and Bruno Tuffin[2]

[1] Orange Labs, rue du Général Leclerc, Issy-les-Moulineaux, France
{helene.lecadre,mustapha.bouhtou}@orange-ftgroup.com
[2] INRIA - Rennes Bretagne Atlantique, Campus de Beaulieu, France
bruno.tuffin@irisa.fr

Abstract. Radio spectrum allocation is essential to the provision of mobile communication services. The spectrum is a finite resource and can accomodate a limited number of simultaneous users at one time. Due to this scarcity, allocating traditional mobile licenses to new mobile operators is unrealizable. Hence, new entrants should bargain access to the networks of the incumbents who establish contracts specifying access charge and maximum traffic volume that the MVNO is allowed to send on the MNO's network. In this article a Mobile Network operator (MNO) shares his finite network resource with a Mobile Virtual Network operator (MVNO) lacking the infrastructure. We study the game where the MVNO invests in content/advertising to compensate for the quality of service degradation. Modeling the system as a supply-chain, i.e. a logistics network consisting of the MNO, the MVNO and the consumers, we determine the access charge and the optimal traffic volume that the MVNO should be allowed to send on the MNO's network to coordinate the system.

1 Introduction

Despite the absence of common definition, Mobile Virtual Network operators (MVNOs) are characterized as being mobile operators without their own infrastructure and government issued licenses. MVNOs buy spectrum and possibly also infrastructure from primary providers, referred to as Mobile Network operators (MNOs). They increase competition in the retail market and enable MNOs to fetch niche markets. Since they are new in a highly competitive market, MVNOs invest a lot in advertising campaigns or specific contents (for instance, M6 Mobile a MVNO on the Orange network, broadcasts soccer games; while Fnac Mobile proposes live concerts or advices about expositions). However, up to date, MVNOs are not really powerfull in the retail market. Indeed most contracts linked MNOs with MVNOs are too restrictive.

Deregulation provides wireless cellular network operators incentives to sell their exceeding spectrum for secondary usage. As an example, Mutlu et al. [5] introduce two classes of consumers: primary users (PUs) whose demand function is price independent since they have long-term contracts with the operator and secondary opportunistic users(SUs). The cellular operator (MNO) aims at selling his exceeding spectrum to

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 24–35, 2009.

increase his revenue but he should be careful to not decrease the quality of service perceived by the PUs who have long term contracts with him. Indeed SU's presence can increase PU's blocking probability, generating a punishement, i.e. the loss of market shares. However they do not introduce competition between the operators on the consumers.

Sethi et al. [4] outline the growing importance of the advertising in most industries today. To increase the advertising investment level, cooperative advertising is used, i.e. producers and sellers share the cost of advertising and the resulting revenue. Contrary to Sethi et al. [4] who use a dynamic demand process based on word-of-mouth effects and advertising perception, our demand function results from a fixed point equation. Indeed it relies on the *opportunity cost* [9] which incorporates parameters such as price, delay and content/advertising investment level. Besides, contrary to them, we test various forms of contracts and there is some competition between the manufacturer (MNO) and the retailer (MVNO) on consumers.

In this article, we deal with a MVNO interconnected with a MNO, transmitting data-services including Voice-over-IP. Each operator is supposed to have a fixed market share, i.e. there is no competition on the consumers, but on the proportion of resource that is shared (bandwidth). Following [4], the system is modeled as a supply-chain. According to [7], *supply-chain management* is a set of approaches utilized to efficiently integrate suppliers (MVNOs), manufacturers (MNOs), and consumers, so that merchandise (i.e. network access) is delivered and shared at the right quantities, to the right consumers, and at the right time, in order to minimize systemwide costs while satisfying service level requirements (i.e. providing the operators sufficient profits). Relationships between the MNO and the MVNO can take many forms, both formal and informal, but often, to ensure adequate bandwith sharing and timely deliveries, MNOs and MVNOs typically agree on supply contracts. Besides these contracts can be used as powerful tools to achieve global optimization, and to motivate the supply-chain parties to reveal their true forecast of customer demand. In our case, the MNO aims at defining a contract allowing the MVNO to send some traffic on his limited resource network. The contract definition is fundamental. Indeed, if the MNO sells too much bandwidth to the MVNO, he might lack capacity for his own clients which would imply a quality of service degradation and in turn, a heavy loss of money. But if he does not sell enough bandwidth, some capacity might be unused which would generate also a heavy loss of money. Furthermore each operator wants to maximize one's revenue by determining consumers' retail prices.

We study the game when the MVNO acts as a Stackelberg leader by investing in advertsing to compensate for the quality of service deterioration, in Section 2. Then, in Section 3, we determine the maximum traffic volume that the MVNO should be allowed to send on the MNO's network and the associated access charge to coordinate the supply-chain. It means that the MNO should have no incentives to sell more or less traffic access to the MVNO and, that the MVNO should agree on the access charge that the MNO suggests.

For the sake of simplicity in all the game models, we will consider the MNO as a male player while the MVNO is female. Also, due to space constraints, most proofs are

left to the extended version [11] of the paper. We just keep the proof of Lemma 1 as it is useful to get the main principle.

2 Optimization of Operators' Revenues without Competition on Consumers When the MVNO Invests in Advertising

In this section, the operators have a fixed number of consumers. They do not compete for consumer market shares but for the proportion of resource they are going to use. To emphasize that, assuming no competition for consumers is relevant in the cases where providers have targeted specific market segments; we can for instance think about a MVNO broadcasting soccer games on mobile phones, or other specific contents: it aims at providing a compensation for the MNO's QoS degradation. The operators want to increase the usage of each consumer they have acquired because it increases their average revenue per user (ARPU). However, the total traffic they can transport is limited because they share a common resource (i.e. spectrum because MNO's consumers and MVNO's consumers are in the same coverage area). The total available bandwidth is μ. We assume in addition that the MNO does not differentiate between his consumers' traffic and the MVNO's consumers' traffic. The MNO now has to share his network with the MVNO; indeed the MVNO is allowed to send a maximum traffic volume $\Lambda_2 \in [0; \mu]$, on the MNO's network. Her access charge is $w\Lambda_2$, where w is a wholesale access price. Besides the MNO saves $\Lambda_1 = \mu - \Lambda_2$ as capacity for his own traffic. The interconnexion model is decribed in Figure 1.

Both operators' consumers valuate the opportunity cost of consuming a unit of traffic (time wasted and price) and send traffic accordingly. We let $p_1,\ p_2 \geq 0$ be the prices paid by consumers for consuming a unit of traffic chosen by both operators. The opportunity costs perceived by consumers for one unit of traffic are $c_1(p_1, p_2) = p_1 + \alpha_1\ d(p_1, p_2)$ at the MNO and $c_2(p_1, p_2) = p_2 + \alpha_2\ d(p_1, p_2) + \beta\ \Theta(\lambda_2, \theta)$ at the MVNO, where $\alpha_1 \geq 0$ (resp. $\alpha_2 \geq 0$) measures the MNO's (resp. MVNO's) consumers' sensitivity to the QoS $d(p_1, p_2)$, and $\beta \leq 0$ describes the sensitivity to content / advertising. $\Theta(\lambda_2, \theta) = \theta\lambda_2 + r\ (\theta \leq 0,\ r > 0)$ models the MVNO's content/advertising investment level perceived by the consumer as a function of her traffic λ_2. Indeed, we assume that the MVNO's investment level is linearly decreasing in the traffic flow generated from her clients. If the consumers produce lots of traffic, it means that they think that the MVNO's perceived quality is satisfactory, hence the MVNO do not need to invest too much in content/advertising to compensate for the bad QoS. On the contrary, if the consumers do not produce much traffic, it means that the experienced QoS is too bad; consequentely the MVNO should invest in content/advertising to seduce them with an attractive brand image and make the traffic volume increase. Besides the more the MVNO invests, the better her brand image will be. Hence for the MVNO's consumers, $-\beta\Theta\Big(\lambda_2, \theta\Big)$ which represents the brand image associated to the MVNO, compensates for the bad opportunity cost of traffic excessive delay.

Demand (that is, arrival rate or average traffic λ_k, for $k \in \{1, 2\}$) is driven by the (random) utility of a unit of traffic U_k, $\lambda_k = \Lambda_k \mathbf{P}\Big[U_k \geq c_k\Big]$ with c_k the above opportunity cost. More specifically, $\lambda_1 = \Lambda_1 \bar{F}\Big(p_1 + \alpha_1 d(p_1, p_2)\Big)$ for the MNO and

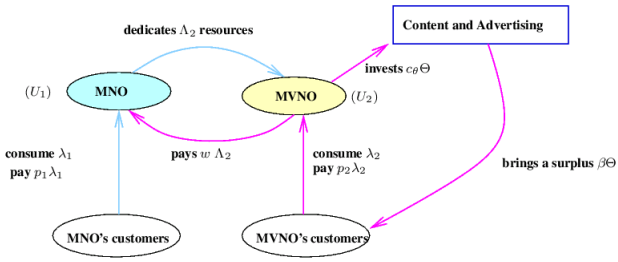


Fig. 1. Description of the interactions between the MNO and the MVNO when their market shares are fixed

$\lambda_2 = \Lambda_2 \bar{F}\Big(p_2 + \alpha_2 d(p_1, p_2) + \beta\Theta(\lambda_2, \theta)\Big)$ for the MVNO, where $\bar{F}$ is the consumers' complementary cumulative distribution function modeling their opportunity cost perception. Indeed consumers send traffic only if their opportunity cost is inferior or equal to the utility they associate to the operator's QoS. We assume that the utility U is generated according to a uniform law on $[0; 1]$. Hence, if there exists a solution with $\lambda_1, \lambda_2 > 0$ to the above fixed point equations, i.e. when the opportunity costs belong to $[0; 1]$, it takes necessarily the form,

$$\lambda_1 = \Lambda_1\Big(1 - p_1 - \alpha_1 d(p_1, p_2)\Big) \text{ for the MNO} \tag{1}$$

$$\lambda_2 = \Lambda_2\Big(1 - p_2 - \alpha_2 d(p_1, p_2) - \beta\Theta(\lambda_2, \theta)\Big) \text{ for the MVNO.} \tag{2}$$

The QoS $d(p_1, p_2)$ is measured via the system's average delay. We assume that the network bottleneck is represented by an M/M/1 queueing system so that $d(p_1, p_2) = \frac{1}{\mu - (\lambda_1 + \lambda_2)}$ where μ is the MNO's maximum bandwith volume [10]. Note that considering the steady-state delay is not a stringent constraint even if there is a feedback loop since prices depend on delay too. Indeed, if price changes operate at a much larger time scale, the queue reaches steady-state before the next price change, and the expected delay can be observed. This is assumed in the present paper.

The operators' utilities depend on their consumers' traffic and chosen prices. The MNO's utility U_1, is the sum of the revenue generated from his consumers and from the contract established with the MVNO minus the cost of his infrastructure C:

$$U_1(p_1, p_2) = p_1\lambda_1 - C + w\Lambda_2 \tag{3}$$

The MVNO's utility U_2, is the sum of the revenue generated from her consumers minus the contract and the content/advertising investment costs. We assume that the content/advertising investment costs are proportional to the investment level as perceived by each individual consumer. The cost of one unit of perceived content/advertising investment level (e.g. soccer game) is c_θ. As a result :

$$U_2(p_1, p_2) = p_2\lambda_2 - w\Lambda_2 - c_\theta\Theta(\lambda_2, \theta) \tag{4}$$

Lemma 1. *The fixed point Equations (1) and (2) have unique solutions in λ_1 and λ_2. Besides, the delay can be expressed as a function of both operators' prices p_1 and p_2:*

$$d(p_1,p_2) = \frac{\Lambda_1(1-p_1) + \frac{\Lambda_2}{1+\Lambda_2\beta\theta}\Big(1-p_2-\beta r\Big) - \mu + \sqrt{\Delta}}{2\Big[\alpha_1\Lambda_1 + \frac{\alpha_2}{1+\Lambda_2\beta\theta}\Lambda_2\Big]}, \tag{5}$$

where $\Delta = \Big[\mu - \Lambda_1(1-p_1) - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r)\Big]^2 + 4(\alpha_1\Lambda_1 + \frac{\alpha_2}{1+\Lambda_2\beta\theta}\Lambda_2)$.

Proof of Lemma 1. $\lambda_1 \mapsto \frac{\lambda_1}{\Lambda_1}$ is strictly increasing in λ_1, while $\lambda_1 \mapsto 1 - p_1 - \alpha\frac{1}{\mu-(\lambda_1+\lambda_2)}$ is decreasing in λ_1; then using the MNO's fixed point Equation (1) we infer that the traffic rate λ_1 is uniquely defined provided $p_1 + \alpha_1 d(p_1,p_2) \leq 1$ otherwise $\lambda_1 = 0$. Since $\lambda_2 \mapsto \frac{(1+\Lambda_2\beta\theta)\lambda_2}{\Lambda_2}$ is strictly increasing in λ_2 and $\lambda_2 \mapsto \bar{F}\Big(p_2 + \beta r + \alpha_2\frac{1}{\mu-(\lambda_1+\lambda_2)}\Big)$ is decreasing in λ_2; the Equation (2) tells us that the MVNO's traffic rate λ_2 is uniquely defined provided $p_2 + \alpha_2 d(p_1,p_2) \leq 1$ otherwise $\lambda_2 = 0$. The system delay can be written: $d(p_1,p_2) = \frac{1}{\mu-(\lambda_1+\lambda_2)} = \Big\{\mu - \Lambda_1(1-p_1) - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r) + (\alpha_1\Lambda_1 + \frac{\alpha_2\Lambda_2}{1+\Lambda_2\beta\theta})d(p_1,p_2)\Big\}^{-1}$. Determining analytically $d(p_1,p_2)$ is equivalent to solve a second order polynomial equation: $(\alpha_1\Lambda_1 + \frac{\alpha_2}{1+\Lambda_2\beta\theta}\Lambda_2)d(p_1,p_2)^2 + (\mu - \Lambda_1(1-p_1) - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r))d(p_1,p_2) - 1 = 0$. Since the discriminant $\Delta = \Big[\mu - \Lambda_1(1-p_1) - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r)\Big]^2 + 4(\alpha_1\Lambda_1 + \frac{\alpha_2}{1+\Lambda_2\beta\theta}\Lambda_2)$ is always non negative, we get that $d(p_1,p_2)$ is the unique positive root of the equation, which gives us Equation (5). □

In this section we assume that the contract parameters (w, Λ_2) are fixed, i.e. we deal with a simple form of wholesale contract; however it will be the purpose of Section 3 to study more complex contracts and define which one the MNO should choose depending on both operators' power relations and to optimize the contract parameters. Hence we assume that the MNO allows the MVNO to send Λ_2 as traffic volume on her network; while the MVNO pays a wholesale access price of $w\Lambda_2$ to gain access to the MNO's network. The operators' utility optimization problem is made of the two steps described below. It can be seen as a Stackelberg game where the MVNO is the leader since she has the possibility to invest in content/advertising and hence is more powerful. Indeed by investing in content/advertising, she will improve her brand image; as a result, the consumers' perceived opportunity costs will decrease, i.e. their traffic λ_2 will increase. In turn the MVNO's traffic increase implies that the MNO's network becomes more congested. Hence the MNO should have to lower his price to increase his consumers' traffic λ_1.

Operators' revenue optimization and game on content/advertising investment

1. The MVNO chooses a price p_2 and invests in content/advertising to maximize her utility, U_2;
2. the MNO chooses a price p_1 to maximize his utility, U_1;
3. the consumers compute their opportunity costs: $c_1(p_1,p_2)$ and $c_2(p_1,p_2)$, their traffics evolve according to the fixed point equations (1) and (2).

The game is solved by *backward induction*. The leader (MVNO) considers what the best response of the follower (MNO) is, i.e. how it will respond once it has observed the price p_2 and content/advertising investment level $\Theta(\lambda_2,\theta)$ of the leader. The leader then

determines a price p_2 that maximizes her payoff, anticipating the predicted response of the follower. The follower actually observes this and in equilibrium picks the expected price p_1 as a response.

Proposition 1. *The MVNO's price being fixed to p_2*[1]*, to optimize the MNO's utility, we have to solve a fourth order polynomial equation*[2] *in p_1*

$$
\begin{aligned}
& a_4 p_1^4 + a_3 p_1^3 + a_2 p_1^2 + a_1 p_1 + a_0 = 0 \\
\Leftrightarrow & \Big[-3\alpha_1 \Lambda_1^6 \Big] p_1^4 + \Big[2KL\Lambda_1^2 - 2\alpha_1^2 \Lambda_1^5 M - 2\alpha_1^2 \Lambda_1^5 (\mu - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r) - \Lambda_1) \\
- & \ 2\alpha_1^2 \Lambda_1^4 P \Big] p_1^3 + \Big[K^2 \Lambda_1^2 + 2KLP + L^2 - \Lambda_1^4 \alpha_1^2 - 2\alpha_1^2 \Lambda_1^3 MP \\
- & \ \alpha_1^2 \Lambda_1^4 (\mu - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r) - \Lambda_1)^2 - 2\alpha_1 \Lambda_1^4 N \Big] p_1^2 + \Big[K^2 P + 2KLN \\
- & \ (\Lambda_1 + \alpha_1)^2 P - 2\alpha_1^2 \Lambda_1^3 MN - 2\Lambda_1 (\mu - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r) \\
- & \ \Lambda_1) \Big] p_1 + \Big[K^2 N - \Lambda_1 \alpha_1^2 N \Big] = 0,
\end{aligned}
$$

where we have set:

$$
\begin{aligned}
K &= 2[\alpha_1 \Lambda_1 + \frac{\alpha_2}{1+\Lambda_2\beta\theta}\Lambda_2]\Lambda_1 - \alpha_1 [\frac{\Lambda_1 \Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r)], \\
L &= -4\Lambda_1 (\alpha_1 \Lambda_1 + \frac{\alpha_2}{1+\Lambda_2\beta\theta}\Lambda_2) + 2\alpha_1 \Lambda_1^2, \\
M &= \mu - \Lambda_1 - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r), \\
N &= [\mu - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r) - \Lambda_1]^2 + 4(\alpha_1 \Lambda_1 + \frac{\alpha_2}{1+\Lambda_2\beta\theta}\Lambda_2), \\
P &= 2\Lambda_1 (\mu - \frac{\Lambda_2}{1+\Lambda_2\beta\theta}(1-p_2-\beta r) - \Lambda_1).
\end{aligned}
$$

- *Besides if $0 < w\Lambda_2 - C \leq \Lambda_1(1-p_1^*)\frac{d(1,p_2^*)}{d(p_1^*,p_2^*)}$, a unique positive price maximizes the MNO's utility (3).*
- *If $\alpha_1 > \frac{w\Lambda_2 - C}{\Lambda_1 d(1,p_2^*)}$ then $p_1^* \in [0;1]$. Since the consumers' utilities belong to the interval $[0;1]$, to guarantee that the MNO's traffic rate λ_1 defined in Equation (1) does not vanish in (p_1^*, p_2^*), i.e. that $c_1(p_1^*, p_2^*) \in [0;1]$, we impose that $\alpha_1 \in]\frac{w\Lambda_2 - C}{\Lambda_1 d(1,p_2^*)}; \frac{1-p_1^*}{d(p_1^*,p_2^*)}]$.*
- *If $p_1^* \geq \frac{\Lambda_2}{\log \Lambda_2} - c_\theta \theta$ then it is sufficient to choose $\alpha_2 > \frac{p_2^* - \beta_2 \Theta(\lambda_2(p_1^*,p_2^*),\theta)}{d(p_1^*,p_2^*)}$ to guarantee that $c_2(p_1^*, p_2^*) \in [0;1]$.*

Proof of Proposition 1. The proof can be found in [11].
From this value, the optimal price for the MVNO can be derived numerically.

Numerical illustrations. Using Proposition 1, the MVNO's utility U_2 can be expressed as a function of p_2 since the MNO's consumer access price p_1 is now merely a function of p_2. We check numerically[3] in Figure 2 that U_2 has a unique solution in p_2. The

[1] It implies that the MVNO's content/advertising level is fixed according to the fixed point Equation (2).

[2] The polynomial equation coefficients ranked in decreasing power are $a_4,\ a_3,\ a_2,\ a_1,\ a_0$.

[3] In the revenue optimization problem, the parameters are defined as follows: $\alpha_1 = 0.8,\ \alpha_2 = 0.85,\ \mu = 10,\ \Lambda_1 = \Lambda_2 = 5,\ \beta = -0.4,\ \theta = -3,\ r = 16,\ C = 3,\ c_\theta = 0.5$.

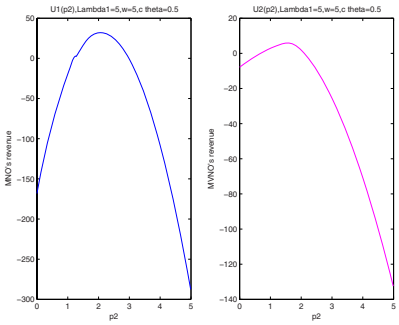


Fig. 2. The MVNO's price maximizing her utility is unique

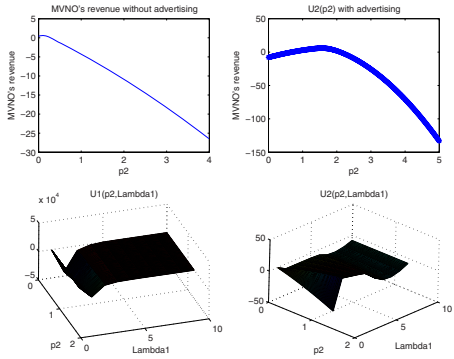


Fig. 3. Investing in content / advertising increases the MVNO's revenue and both operators' optimal revenues depend on the maximal traffic volume the MVNO is allowed to send on the MNO's network ($\Lambda_2 = \mu - \Lambda_1$)

maxima of the operators revenues are: $U_1^* = 22.01$ and $U_2^* = 5.89$. In Figure 3 we observe that both operators' maximized utilities relie on the maximum traffic volume the MVNO is allowed to send on the MNO's network[4] and that the MVNO has incentives to invest in content/advertising since maximized revenue is greater with some content/advertising than without ($U_2^*(\theta = 0, r = 0) = 1.19$).

3 Providers Power Relations and Contract Definition

The aim of this section is to define the contract between the MNO and the MVNO. Indeed we can see the MNO as a supplier of the scarce resource for both his consumers' business units and for the MVNO. Then the system composed of the MNO, the MVNO and both providers' consumers can be modeled as a supply-chain. The contracts of the supply-chain are between the MNO and the MVNO; additionally these latter determine optimal pricing strategies to sell their services on the market. The need from *coordination*, i.e. the maximization of the social-welfare using contract parameters imposed by a regulatory authority such that no provider has incentives to deviate, results from the following observation: if the MNO imposes upon the MVNO to buy $\Lambda_2^* = \arg\max_{\Lambda_2} U_1$, then the MVNO assumes all of the risk of having more bandwidth than she really needs, while the MNO takes no risk. Indeed, since the MNO takes no risk, he would like the MVNO to buy as much of his unused bandwith as possible; while the MVNO would prefer to limit the quantity due to the huge financial risk. Of course if the MVNO limits her order quantity there is a significant increase in the likelihood of spoilt capacity for the MNO. If the MNO is willing and able to share some of the risk with the MVNO, it may be profitable for the MVNO to order more bandwith, thereby reducing spoilt capacity probability and increasing profit for both of them. It turns out that a variety of

[4] It is quite natural since U_1 and U_2 are functions of Λ_1 and Λ_2.

supply-chain contracts enable this risk sharing, and therefore increase profits for both supply-chain entities [7]. In Section 2 we have supposed that the unit access charge w was fixed; i.e. we were in the simple case of a wholesale contract. The aim of this section is now to study more complex contracts such as revenue sharing, quantity discount and sales rebate [2]. The MNO wants to determine how much traffic the MVNO should be allowed to send on his network. He assumes to be in the *worst* case, i.e. that the MVNO has so much consumers that she sends the maximal allowed traffic volume on his network. More formally, we aim to determine the maximal traffic volume Λ_2 that the MVNO should buy at the access charge w such that neither of the operators can increase ones' profit by unilaterally deviating from his (her) choice.

If the MNO acts *selfishly*, he determines contract parameters: Λ_2^*, $w^* \equiv w(\Lambda_2^*)$ maximizing his own utility. However such a behavior is not optimal when looking at the social-welfare as the sum of MNO and MVNO utilities. Optimal performances can be achieved if they coordinate by contracting on a set of transfer payments such that each operator's objective becomes aligned with the supply-chain's objective. The contract coordinating the supply-chain relies on the optimal traffic volume $\Lambda_2^0 = \arg\max_{\Lambda_2} U$. In fact supply-chain's contracts play the role of an unbiased decision maker who would identify the best strategy for the entire supply-chain and allocate the whole profit between both players. This unbiased decision maker would consider the two supply-chain's partners: the MNO and the MVNO, as two members of the same organization. Hence the transfer of money between the parties would be ignored and the unbiased decision maker will maximize the supply-chain's profit. Contracts help firms to achieve global optimization [7]. The contract definition requires the four following steps:

(i) the MNO chooses a contract category between revenue sharing, quantity discount and sales rebate;
(ii) the MNO determines contract parameters w^*, Λ_2^* maximizing his utility U_1;
(iii) an unbiased decision maker computes $\Lambda_2^0 = \arg\max_{\Lambda_2} U$ and allocates the supply-chain's profit between the operators;
(iv) the MVNO is free to refuse the contract.

In practice, the MNO's share of the supply-chain's optimal profit is defined as in [2]: $\rho_1^* = \frac{U_1(\Lambda_2^*, w^*)}{U(\Lambda_2^0)}$, where the supply-chain's utility $U = U_1 + U_2$ is the sum of the MNO and MVNO's utilities. The higher this coefficient is, the more attractive the contract is for the MNO since $U_1 = \rho_1^* U(\Lambda_2^0)$ and $U_2 = (1 - \rho_1^*)U(\Lambda_2^0)$. We consider the three following types of contracts.

– If the revenue sharing contract is implemented, the MNO charges $w_\Phi(\Lambda_2)$ per unit traffic purchased plus the MVNO gives the MNO a percentage of her revenue (cf [3] for an application to the video cassette rental industry). Indeed in the sequential supply-chain, one important reason for the MVNO to refuse a contract is the high wholesale price [7]. If somehow the MVNO can convince the MNO to reduce the wholesale price, then clearly the MVNO will have an incentive to cooperate. Of course a reduction in wholesale price will decrease the MNO's profit if he is unable to sell more capacity. This is adressed in the revenue sharing contract where the MVNO and the MNO share the revenues generated from the consumers. Let $\Phi \in [0; 1]$ be the fraction of the supply-chain's revenue that the MVNO keeps;

$(1-\Phi)$ is then the fraction earned by the MNO. Depending on both operators' power relations, the sharing parameter will be defined in Theorem 1. In the revenue sharing contract, the providers' utilities are defined as:

$$U_1 = (1-\Phi)(\lambda_1 p_1 + \Lambda_2 p_2) + w_\Phi(\Lambda_2)\Lambda_2 - C \tag{6}$$
$$U_2 = \Phi(\lambda_1 p_1 + \Lambda_2 p_2) - w_\Phi(\Lambda_2)\Lambda_2 - c_\theta(\theta\Lambda_2 + r) \tag{7}$$

– In the quantity discount contract, the transfert payment is $w_d(p_2, \Lambda_2)\Lambda_2$ where w_d is a decreasing function of the MVNO's traffic Λ_2. The MVNO pays the MNO $w_d(p_2, \Lambda_2)\Lambda_2$ to send her traffic on the MNO's network [2]. In the quantity discount contract, the operators' utilities take the form:

$$U_1 = \lambda_1 p_1 + w_d(p_2, \Lambda_2)\Lambda_2 - C \tag{8}$$
$$U_2 = \Lambda_2 p_2 - w_d(p_2, \Lambda_2)\Lambda_2 - c_\theta(\theta\Lambda_2 + r) \tag{9}$$

– With a sales rebate contract, the MNO charges $w_s(p_1, p_2, \Lambda_2)$ per unit of traffic purchased by the MVNO but then gives the MVNO a rebate $r_s(\Lambda_2)$ which is a function of the traffic sent by the MVNO, if the MVNO's QoS $d(p_1, p_2)$ is above some predefined threshold d_0 [2]. In the sales rebate contract, the utilities are:

$$U_1 = \lambda_1 p_1 + w_s(p_1, p_2, \Lambda_2)\Lambda_2 - r_s(\Lambda_2)\mathbf{1}_{\{d(p_1,p_2)>d_0\}} - C \tag{10}$$
$$U_2 = \Lambda_2 p_2 - w_s(p_1, p_2, \Lambda_2)\Lambda_2 + r_s(\Lambda_2)\mathbf{1}_{\{d(p_1,p_2)>d_0\}} - c_\theta(\theta\Lambda_2 + r) \tag{11}$$

We insert the QoS measure expression, i.e. the delay $d(p_1, p_2) = \frac{1}{\mu - \left(\lambda_1 + \Lambda_2\right)}$, in the fixed point Equation (1); the MVNO's traffic being fixed to Λ_2. To determine the MNO's traffic $\lambda_1(\Lambda_2)$ we need to solve a second order polynomial equation in λ_1 whose unique positive root is:

$$\lambda_1(\Lambda_2) = \frac{1}{2}\Big\{[\mu - \Lambda_2 + \Lambda_1(1-p_1)] + \Big((\mu - \Lambda_2 + \Lambda_1(1-p_1))^2 - 4(\Lambda_1(1-p_1)(\mu - \Lambda_2) - \alpha\Lambda_1)\Big)^{\frac{1}{2}}\Big\}. \tag{12}$$

Let $b \geq 1$ and $\gamma \in \mathbb{R}$ be coefficients characterizing the operators' power relations. We assume that the supply-chain's utlility is linear in the MNO's one, i.e. $U = bU_1 + \gamma$. It implies that the contract parameters which maximize the supply-chain's utility is a Nash equilibrium for the operators [2]. Indeed let (w, Λ_2) be contract parameters maximizing the supply-chain's utility U. If we assume that there exists Λ_2' such that $\Lambda_2' \neq \Lambda_2$ and $U_1(w, \Lambda_2') > U_1(w, \Lambda_2)$ then by linearity we get $bU_1(w, \Lambda_2') + \gamma > bU_1(w, \Lambda_2) + \gamma$ i.e. $U(w, \Lambda_2) > U(w, \Lambda_2')$. By application of Nash's Lemma [8] we get that the MNO's strategy is to choose Λ_2 with probability one. If we assume that there exists w' such that $w' \neq w$ and $U_2(w', \Lambda_2) > U_2(w, \Lambda_2)$ then by linearity we get $\frac{b}{b-1}U_2(w', \Lambda_2) - \frac{\gamma}{b-1} > \frac{b}{b-1}U_2(w, \Lambda_2) - \frac{\gamma}{b-1}$ i.e. $U(w', \Lambda_2) > U(w, \Lambda_2)$. By application of Nash's Lemma [8] we get that the MVNO's strategy is to choose w with probability one. Besides depending on both operators' power, we can express the MNO's utility has a linear function of the supply-chain's one: $U_1 = \frac{1}{b}U - \frac{\gamma}{b}$ and $U_2 = \frac{b-1}{b}U + \frac{\gamma}{b}$ for the MVNO. For each contract we determine $\Lambda_2^0(RS),\ \Lambda_2^0(QD),\ \Lambda_2^0(SR)$

i.e. the maximal traffic volume that the MVNO should be allowed to send on the MNO's network to maximize the supply-chain's utility under revenue sharing, quantity discount and sales rebate contracts. Then by substitution in the contract parameters we get $w_\Phi^0 \equiv w_\Phi\Big(\Lambda_2^0(RS)\Big)$, $w_d^0 \equiv w_d\Big(p_2, \Lambda_2^0(QD)\Big)$ and $w_s^0 \equiv w_s\Big(p_1, p_2, \Lambda_2^0(SR)\Big)$.

Theorem 1. *The three contract parameters* $\Big(\Lambda_2^0(RS), w_\Phi^0\Big)$, $\Big(\Lambda_2^0(QD), w_d^0\Big)$ *and* $\Big(\Lambda_2^0(SR), w_s^0\Big)$ *which are respectively the MNO's optimal traffic to sell to the MVNO and the associated access charge for the revenue sharing, the quantity discount and the sales rebate contracts, defined with the following parameters*

$$\Phi = 1 - \frac{1}{b}, \text{ and } w_\Phi(\Lambda_2) = (\Phi - 1)\frac{\gamma + C\frac{\Phi}{\Phi-1} + c_\theta(\theta\Lambda_2 + r)}{\Lambda_2}; \tag{13}$$

$$w_d(p_2, \Lambda_2) = p_2 + \frac{c_\theta(\theta\Lambda_2 + r) + \gamma}{\Lambda_2}; \tag{14}$$

$$w_s(p_1, p_2, \Lambda_2) = \frac{(1-b)\lambda_1 p_1 + \Lambda_2 p_2}{b\Lambda_2}, \text{ and } r_s(\Lambda_2) = \frac{c_\theta(\theta\Lambda_2 + r) + \gamma}{b}; \tag{15}$$

coordinate the supply-chain (in the sense that it maximizes the total surplus of the MNO and the MVNO, i.e. the sum of their utilities).

Proof of Theorem 1. Using Equations (6) and (7), (8) and (9), (10) and (11); by identification of the relation $U = U_1 + U_2 = bU_1 + \gamma$, we infer the contract parameters values. □

We determine the optimal traffic volume that the MVNO should be allowed to send on the MNO's network to maximize the supply-chain's utility under the three contracts considered.

Proposition 2. *For the revenue sharing contract we get:*

$$\Lambda_2^0(RS) = \frac{1}{2 - \frac{1}{2[2\frac{(c_\theta\theta - \Phi p_2)^2}{\Phi p_1} + 1]^2}}\Big[2\mu - 2\Lambda_1(1-p_1) + \frac{1}{2[2\frac{(c_\theta\theta - \Phi p_2)^2}{\Phi p_1} + 1]^2} \Big(\mu - 3\Lambda_1(1-p_1)\Big) + \sqrt{\Delta_r}\Big], \tag{16}$$

where $\Delta_r = \Big[2\mu - 2\Lambda_1(1-p_1) + \frac{1}{2[2\frac{(c_\theta\theta - \Phi p_2)^2}{\Phi p_1} + 1]^2}\Big(\mu - 3\Lambda_1(1-p_1)\Big)\Big]^2 - 4\Big[\Big(2\frac{(c_\theta\theta - \Phi p_2)^2}{\Phi p_1} + 1\Big)^2\Big]\Big[\mu^2 - 2\mu\Lambda_1(1-p_1) + \Lambda_1^2(1-p_1)^2 + 4\alpha\Lambda_1 - \frac{1}{4[2\frac{(c_\theta\theta - \Phi p_2)^2}{\Phi p_1} + 1]}\Big]$.

For the quantity discount and sales rebate contracts we have:

$$\Lambda_2^0(QD) = \Lambda_2^0(SR) = \frac{1}{2}\Big[\frac{16}{(\frac{p_1}{2} - p_2 + c_\theta\theta)^2} - \Big(2\mu_2 - 2\Lambda_1(1-p_1) + 4\Lambda_1(1-p_1)\Big) + \sqrt{\Delta_q}\Big] \tag{17}$$

where $\Delta_q = \left[\left(2\mu_2 - 2\Lambda_1(1-p_1) + 4\Lambda_1(1-p_1)\right)\frac{16}{(\frac{p_1}{2} - p_2 + c_\theta\theta)^2}\right]^2 - 4\left[\mu^2 - 2\Lambda_1(1-p_1)\mu + \Lambda_1^2(1-p_1)^2 + 4\alpha\Lambda_1 + \frac{16}{(\frac{p_1}{2} - p_2 + c_\theta\theta)^2}\right]$.

To determine $\Lambda_2^*(RS)$ *(resp.* $\Lambda_2^*(QD)$*,* $\Lambda_2^*(SR)$*) we substitute* $\lambda_1(\Lambda_2)$ *obtained in Equation (12) and the contract parameters obtained in Equation (13) (resp. (14) and (15)) in the MNO's utility* U_1 *and derive it with respect to* Λ_2*. The traffic volumes that the MNO should allowed access to when he acts selfishly, are defined analytically for each contract as one of the positive roots of fourth order polynomial equations and solve* $\partial_{\Lambda_2} U_1\left(\Lambda_2^*(RS)\right) = 0$ *(resp.* $\partial_{\Lambda_2} U_1\left(\Lambda_2^*(QD)\right) = 0,\ \partial_{\Lambda_2} U_1\left(\Lambda_2^*(SR)\right) = 0$*).*

Proof of Proposition 2. The proof can be found in [11].

Table 1. Most profitable contract to use for the MNO depending on power relations and content/advertising investment level (RS means *revenue sharing*, QD means *quantity discount* and SR, *sales rebate*)

Advertising level \ MNO's power par.	$0 < \frac{1}{b} \leq 0.25$	$0.25 < \frac{1}{b} \leq 0.5$	$0.5 < \frac{1}{b} \leq 0.75$	$0.75 < \frac{1}{b} < 1$
$-0.1 \geq \theta > -0.5$	QD	SR	RS	RS
$-0.5 \geq \theta > -0.7$	QD	SR	SR	SR
$-0.7 \geq \theta \geq -1.0$	SR	SR	RS	RS

In Table 1 we identify the parameter $\frac{1}{b}$ with the MNO's power in the supply-chain; optimizing numerically both operators' consumers' retail prices, we determine which contract maximizes the MNO's share of the optimal supply-chain's profit depending on power relations and content/advertising investment level θ. The Table 1 results are logical since when the MNO is not powerful, he has to give the MVNO some incentives to collaborate, for example by using an access charge decreasing in the volume of traffic this latter sends on his network or by promising a refund if the MVNO's delay is above some predefined threshold; and when he is powerful, a revenue sharing contract is more lucrative since he gets a high fraction (equal to his power coefficient) of the supply-chain's total revenue. For example, in the last column of Table 1 (the MNO has high market power), we can explain the swapping between revenue sharing (RS) and sales rebate (SR) like this: when the investment is low, the MVNO is weak compared to the MNO so the MNO can impose risk sharing to the MVNO via a revenue sharing contract (RS); when the investment is higher, the MNO is threatened by the MVNO because it consumes more traffic (and so increases the average delay), so the MNO prefers a sales rebate contract (SR); finally when the investment is very high, the revenue of the MVNO is more attractive for the MNO than providing low delays in order to satisfy his consumers, so he would rather use a revenue sharing contract (RS).

4 Conclusions

In this paper, we have modeled the economic interactions between a MNO sharing his limited network resource and a MVNO lacking the infrastructure. We have designed an optimal pricing strategy when no traffic differentiation is introduced by the MNO while the MVNO invests in content/advertising to compensate for the QoS degradation. The access charge and the optimal maximal traffic volume that the MVNO should be allowed to send on the MNO's network to coordinate the supply-chain made of both operators and consumers, are defined for various popular contracts in the supply-chain literature (i.e. revenue sharing, quantity discount and sales rebate).

The model can be extended to include some dynamicity by introducing two competitive Markov decision processes modeling both operators' market share evolution, and uncertainty might result from the MNO's ignorance of the consumers' expected service level [6]. Besides the model could be transposed to the machine-to-machine market taking data integrity as QoS measure.

Acknowledgment. This work was supported in part by the French research agency through the CAPTURES project. We also thanks the reviewers for helpfull remarks.

References

[1] Acemoglu, D., Bimpikis, K., Ozdaglar, A.: Price and Capacity Competition. In: Games and Economic Behavior (2008) (to appear)
[2] Cachon, G.: Supply-chain Coordination with Contracts. Handbooks in Operations Research and Management Science 11, 227–339 (2007)
[3] Cachon, G., Lariviere, M.: Capacity Choice and Allocation: Strategic Behavior and Supply-Chain Performance. Management Science 45, 1091–1108 (1999)
[4] Sethi, S., He, X., Prasad, A.: Cooperative Advertising and Pricing in a Dynamic Stochastic Supply-Chain: Feedback Stackelberg Equilibrium. In: Production and Operations Management (2008) (to appear)
[5] Mutlu, H., Alanyali, M., Starobinski, D.: Spot Pricing of Secondary Spectrum Usage in Wireless Cellular Networks. In: IEEE INFOCOM (2008)
[6] Bernstein, F., Federgruen, A.: A General Equilibrium Model for Industries with Price and Service Competition. Operations Research 52, 868–886 (2003)
[7] Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E.: Designing and Managing the Supply-Chain: Concepts, Strategies and Case Studies. McGraw-Hill International Edition (2008)
[8] Myerson, R.: Game Theory: Analysis of Conflict. Harvard University Press, Cambridge (1991)
[9] Pohjola, O.-P., Kilkki, K.: Value-based methodology to analyze communication services. Netnomics 8, 135–151 (2007)
[10] Latouche, G., Ramaswami, V.: Introduction to Matrix Analytic Methods in Stochastic Modelling. In: Quasi-Birth-and-Death Processes, 1st edn., ch. 1. ASA SIAM, Philadelphia (1999)
[11] Le Cadre, H., Bouhtou, M., Tuffin, B.: A Pricing model for a Mobile Network Operator sharing limited resource with a Mobile Virtual Network Operator, Internal Report, Orange Labs/R&D/CORE/MCN/13.09

Design and Evaluation of a Combinatorial Double Auction for Resource Allocations in Grids*

Li Li[1], Yuanan Liu[1], David Hausheer[2], and Burkhard Stiller[2,3]

[1] School of Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China
[2] Communication Systems Group CSG, Department of Informatics IFI, University of Zürich, Binzmühlestrasse 14, CH-8050 Zürich, Switzerland
[3] Computer Engineering and Networks Laboratory TIK, ETH Zürich, Gloriastrasse 35, CH-8092 Zürich, Switzerland
lily82.bupt@gmail.com, yuliu@bupt.edu.cn, hausheer@ifi.uzh.ch, stiller@ifi.uzh.ch

Abstract. Offering Grid services in an open market determines an optimization case for finding the best suitable resource allocation for a given number of requests and existing resources. Thus, appropriate resource allocation schemes, supporting accounting, are required in addition to a pricing scheme, which supports financial fairness criteria. The newly developed Resource Allocation Model for the Combinatorial Double Auction (RAMCoDA) achieves these requirements, while being incentive compatible.

1 Introduction and Related Work

Commercial Grid services need to support resource allocations, which should be incentive-compatible. Therefore, resource allocations in Grids need the support of a suitable accounting and pricing (often termed billing) scheme. As economic theory tells, auctions do have the potential, if applied in a sensible manner, to achieve the incentive compatibility. With respect to the Grid services market, Combinatorial Auctions (CA) can represent satisfying characteristics. Within CAs, the user can bid for combinations of resources, on which tasks can be executed, while improving economic efficiency and maximizing the revenue of the Grid. However, existing CA-based resource allocations [1], [6] focus on the users' side only and they do not define any specific mechanisms for pricing and accounting resources. Since the CA's pricing algorithm is a computationally complex task, existing work regard the final price paid by the agent often as its bid or the simple division of the auctioneer's total income. This may lead to a misrepresentation of the agents' true valuations, which cannot ensure incentive compatibility.

* This work was supported partly by the European Union through the EC-GIN project (STREP) under Grant No. FP6-2006-IST-045256. The authors acknowledge fruitful discussions with all their project partners.

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 36–47, 2009.

Thus, a novel resource allocation model based on Combinatorial Double Auctions (CoDA) is proposed, which is suitable for accounting and pricing purposes. The CoDA combines both advantages of a Double Auction (DA) and a CA. Analytical experiments of the newly developed Resource Allocation Model for the Combinatorial Double Auction (RAMCoDA) show that the new scheme is effective and incentive-compatible. Existing pricing schemes in related work can be classified into three main categories: Bargaining Models [2], Commodity Market Models [5], [10], and Auction Models [6], [4], [9]. Auction Models include either one-to-many or many-to-many interactions. The DA is widely used for many-to-many auctions, in which buyers and sellers are treated symmetrically with buyers and sellers submitting bids. Instead of selling items of resources individually, in a CA the seller allows bids on bundles of items, enabling bidders to deal with entities of direct interest and avoiding the risk of obtaining incomplete bundles. The auction based pricing schemes include an autonomous pricing mechanism as proposed by [9], in which prices are decided by Grid traders within the trading process, a type of autonomous pricing mechanism proposed by [4], in which consumers and producers are autonomous agents that make their own decisions according to their capabilities and their local knowledge, and another pricing strategy for combinatorial Grids introduced in [6], where resource agents administrate available memory, CPU, network bandwidth, and disk capacities on the supply side. Further details on related auctions and pricing schemes can be obtained in [3].

The distinguishing characteristic of RAMCoDA as herein proposed is determined by the use of CoDA onto resource allocation and pricing in Grids. The newly designed model enables to obtain the complete allocation information and trade price as described in the following sections.

While Section 2 introduces the underlying resource allocation model of RAMCoDA, Section 3 defines the corresponding pricing algorithm. Both are simulated in Section 4 and a summary is given in Section 5.

2 CoDA-Based Resource Allocation Model

For the novel CoDA-based resource allocation model (cf. Fig. 1.) each Grid user operates a User Broker (UB). Within the UB, the Resource Discovery component is responsible for finding resources according to users' requirements; it contacts the Grid Information Service (GIS) to obtain the list of resources that matches these requirements. The Auction Agent in UB is responsible for generating resource combinations based on the list of resources returned by the GIS. For each combination of resources, it generates a bid within the user's budget and submits the bid and the corresponding combination to the Grid Market Auctioneer (GMA). The Price Depository component is responsible for storing price information related to the task. At a later stage, this information can be used for accounting. The Job Management Agent is responsible for sending the user job to resources and collecting the results.

The GIS provides resource registration services and maintains a list of resources available in the Grid. The GIS may be implemented in a decentralized

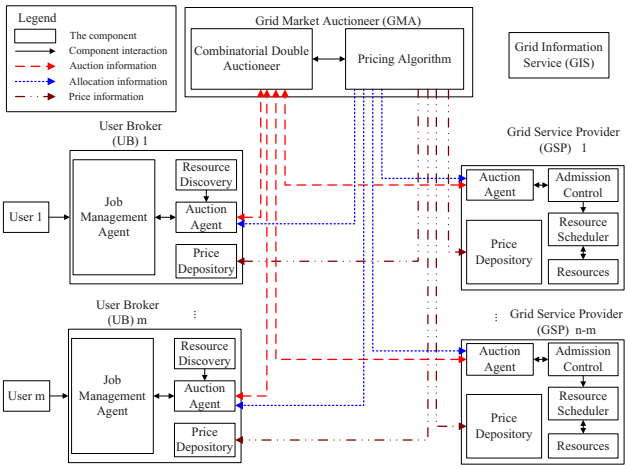


Fig. 1. Resource Allocation Model for the Combinatorial Double Auction (RAMCoDA)

manner, *e.g.*, based on a Distributed Hash Table (DHT). Each Grid Service Provider (GSP) will register all resources it can provide at the GIS. The Auction Agent in GSP is responsible for generating the resource combination it would sell. For each resource combination, it generates a bid and submits it together with the corresponding combination description to the GMA. The Admission Control component receives the auction result from the Auction Agent and decides, whether the tasks sent from the UB will be managed or not. The Resource Scheduler is responsible for allocating all tasks to the corresponding resources. Within the GMA, the Combinatorial Double Auctioneer is responsible for collecting the combination of resources and corresponding bids sent by UBs and GSPs. Based on the that information it runs the CoDA algorithm to determine the winner UBs and GSPs. Additionally, it sends the result obtained to UBs and GSPs. Finally, the Pricing Algorithm is responsible for generating particular allocation results and corresponding price information. The price calculated will be sent to all UBs and GSPs, who did participate in the trade.

RAMCoDA mainly differs from the allocation model in [1] in terms of the applied auction model and the pricing algorithm (cf. Section 3) in the GMA and the Price Depository components in the UB and GSP. Consequently, interactions between the components are different. In [1], only UBs submit their bids, while in RAMCoDA both UBs and GSPs submit their bids and offers, respectively.

3 The Pricing Algorithm

The pricing algorithm is a key part of any DA approach. Thus, for the resource allocation the algorithm proposed follows five steps to determine the best suited resource price.

1. Assuming that there are n participants in a CoDA, including m users and $n-m$ GSPs, UBs and GSPs submit resource combinations (bundles) and bids, in form of B_j, to the GMA. B_j can be specified as (a_j, p_j), where $a = (a_{1j}, ..., a_i\ , ...a_k\ )$, a_i is the quantity of item i requested $(a_i\ > 0)$ or supplied $(a_i\ < 0)$ in the bundle j. Here suppose a UB/GSP is allowed to ask for one bundle in the auction, so the subscript j could identify both a bundle and a UB/GSP. The symbol k denotes that there are k resource items to be considered jointly in the auction. K is the set of the items. p_j is the amount the bidder is willing to pay for the bundle j, if $p_j > 0$, it is regarded as a buying price; otherwise it is regarded as a selling price. The combinatorial double auctioneer runs the CoDA algorithm represented in (1) and receives the result x_j [8]. The objective of CoDA is to maximize the total trade surplus, while satisfying the constraint that the number of units selected by buying bundles does not exceed the number provided by the selected selling bundles for each item.

$$\begin{aligned} &max \sum_{j=1}^{n} p_j x_j, \quad s.t. \sum_{j=1}^{n} a_i\ x_j \leq 0 \quad \forall i \in K, \\ &with\ x_j \in \{0,1\}\ \forall j \in \{1, ..., n\}. \end{aligned} \tag{1}$$

   Afterwards, the GMA informs UBs and GSPs about the bid's acceptance or rejection and requests the winner GSPs to reserve the resources awarded. Loser participants can renew their bids in the next round. The winner participants are denoted as the traders in the following.
2. For each trader, the average bid is calculated as

$$p_j^a = \frac{p_j}{\sum_{i=1}^{k} a_i}. \tag{2}$$

   All traders are buyers or sellers. Buyers are ranked in the decreasing order of the average bid; the result is denoted as the buyer list b^l. Accordingly, sellers are ranked in the increasing order of the average bid; the result is the seller list s^l. All sellers are classified by the resource item. The algorithm achieves s_i^l, $i \in \{1, ..., k\}$ for each item of the resource, while q_i^l represents the resource quantity list corresponding to the s_i^l.
3. Generating the average trade price matrix p^m, in which $p^m(s,t)$ represents the average trade price, when the sth buyer in the list b^l trades with the tth seller in the list s^l, the value is calculated as follows in (3):

$$p^m(s,t) = \frac{p_{b^l(s)}^a + p_{s^l(t)}^a}{2}. \tag{3}$$

4. The resource allocation and pricing is done in order from the first buyer in b^l to the last one. Buyers will be matched with the first seller holding the required resources in the s^l for the allocation. If requirements of this buyer can be satisfied, the algorithm will update results and go to the next

buyer, otherwise, the buyer will be matched with the next seller for surplus requirements. Moreover, R_i^a and R_i^p, which are both $g \times h$ matrices, represent particular allocation and pricing results of the resource i, respectively. Here g denotes the number of buyers in b^l, while h denotes the number of sellers in s^l. The algorithm is defined as follows:

Input:$B, b^l, s^l, s_i^l, q_i^l$ Output:$R_i^p, R_i^a; i \in \{1, ..., k\}$

(a) Initialization:
$s = 1$, $i = 1$, $R_i^p = [0]_{g \times h}$, $R_i^a = [0]_{g \times h}$.

(b) Inquire the average trade price matrix p^m; compare the buyer's requirement $a_{ib^l(s)}$ with $q_i^l(m)$
$m \leftarrow$ the location of the first non-zero quantity in the list q_i^l
$t \leftarrow$ the location of seller $s_i^l(m)$ in the list s^l
If $q_i^l(m) \geq a_{ib^l(s)}$
$R_i^p(s,t) = R_i^p(s,t) + a_{ib^l(s)} \cdot p^m(s,t)$;
$R_i^a(s,t) = R_i^a(s,t) + a_{ib^l(s)}$;
$q_i^l(m) = q_i^l(m) - a_{ib^l(s)}$;
$a_{ib^l(s)} = 0$;
Go to Step (c);
else
$R_i^p(s,t) = R_i^p(s,t) + q_i^l(m) \cdot p^m(s,t)$;
$R_i^a(s,t) = R_i^a(s,t) + q_i^l(m)$;
$a_{ib^l(s)} = a_{ib^l(s)} - q_i^l(m)$;
$q_i^l(m) = 0$;
Repeat Step (b);

(c) Store the temporary allocation result R_i^a and pricing information R_i^p, determine whether all the resource requirements of the sth buyer are satisfied,
If all requirements are satisfied
Go to Step (d);
else
$i = i + 1$;
Go to Step (b);

(d) Determine whether all buyers' requirements are satisfied, i.e., whether get to the end of the b^l,
If end of the b^l
EXIT;
else
$s = s + 1$;
$i = 1$;
update R_i^p, R_i^a, q_i^l;
Go to Step (b);

Finally, each trader will receive a trade price represented by the vector p^b or p^s:

$$p^b = (\sum\nolimits_{t=1}^{h} \sum\nolimits_{i=1}^{k} [R_i^p]_{g \times h})^T,$$
$$p^s = \sum\nolimits_{s=1}^{g} \sum\nolimits_{i=1}^{k} [R_i^p]_{g \times h}, \tag{4}$$

where T denotes the transposition of the matrix.

5. Finally, the GMA sends the related information in those vectors p^b, p^s, and R_i^a, R_i^p to each trader.

4 Simulation Results and Discussion

Based on the model and pricing algorithm presented above, a performance investigation has been undertaken. While the functionality of RAMCoDA shows the interactions needed to obtain the result of a corresponding price, the performance needs an analytical approach. Thus, a simulation has been performed. Experiments were run in Matlab on a Pentium D dual-core CPU 2.8 GHz with 1 GB memory. Key rules for the parameter selection are as follows: for buyers, the value of each resource item is within the range $[l_i^b, u_i^b]$, for sellers, the value of each resource item is within the range $[l_i^s, u_i^s]$, with the constraint $l_i^s \leq l_i^b$, $u_i^s \leq u_i^b$ to lower the possibility that too many bids of sellers are higher than bids of buyers. Here $l_i^b, u_i^b, l_i^s, u_i^s$ are all the fixed parameters of the simulation. The demand and supply quantity of resource i from each participant, i.e.$|a_{ij}|$, is uniformly distributed over the interval $[0, d_i]$ and $[0, s_i]$, respectively. Based on the assumption above, each buyer can give its valuation of his resource combination within the range $\left[\sum_{i=1}^{k} a_{ij} l_i^b, \sum_{i=1}^{k} a_{ij} u_i^b\right]$, while seller's valuation is within $\left[\sum_{i=1}^{k} a_{ij} u_i^s, \sum_{i=1}^{k} a_{ij} l_i^s\right]$. In the following simulation, all participants bid for resource combinations according to their true valuation, which means bids are equal to those valuations.

The particular simulation undertaken considers 3 items of resources in the Grid, denoted as A, B, and C. These resources may include in a real-world scenario storage, access bandwidth, and a certain software library needed to run the Grid's task of a weather simulation. In the case considered, 20 participants are involved in the CoDA, including 10 UBs and 10 GSPs. For UBs, the value ranges of 3 items of resources are [5, 10], [10, 15], [15, 20] respectively, while for GSPs, the ranges are [4, 8], [8, 12], [12, 16] respectively. Moreover, $d_i = s_i = 20$ has been selected. The parameter settings can be found in Tab.1. Through solving the CoDA represented in (1), participants 1, 2, 3, 5, 6, 7, 8, 9, 10 and 11, 12, 13, 14, 16, 17, 19, 20 are the winning bidders. The bids of 4, 15 and 18 are rejected in this round.

The allocation and pricing results are shown in Fig. 2. Buyer list b^l has the same order as the row order of R_i^a and R_i^p, while the seller list s^l has the same order as the column order of these matrices. It can be seen that Fig. 2 gives the complete information of the allocation and all trade prices. Take buyer 3 for example: his demand is (0 2 7) – determining the request for 2 units for resource B and 7 units for resource C, respectively – and the bid is 147. His trade information can be obtained from the corresponding row based on his position in b^l, i.e. the first row, in the R_i^a and R_i^p matrices. As marked in R_i^a in Fig. 2, all requirements of buyer 3 are satisfied by 3 sellers jointly, thus, here the first seller in s^l, i.e. 17, provides 2 units of resource B and 2 units of resource C; the second seller in s^l, i.e. 16, provides 2 units of resource C; and the third seller in s^l, i.e. 14, provides 3 units of resource C. They will charge this buyer 23, 23, 24.8 and 38.3 monetary units, respectively, as marked in R_i^p in Fig. 2. Similarly, the trade information of each seller can be gained from the corresponding column

Table 1. Parameter settings for each participant

No.	A	B	C	p	No.	A	B	C	p
1	3	10	4	253	11	-13	-20	-11	-461
2	6	16	8	417	12	-12	-1	-19	-326
3	0	2	7	147	13	-7	-15	-3	-250
4	5	13	14	399	14	-7	-20	-16	-396
5	11	14	20	683	15	-11	-12	-7	-336
6	8	9	16	509	16	-19	-15	-2	-303
7	15	9	9	392	17	-14	-10	-2	-174
8	16	6	1	200	18	-10	-1	-20	-366
9	15	16	6	442	19	-7	-14	-13	-349
10	13	19	1	431	20	-13	-9	-6	-272

$$R_1^a = \begin{bmatrix} 0 & & & & & & & \\ 8 & & & & & & & \\ 6 & 5 & & & & & & \\ 0 & 3 & 0 & & & & & \\ 0 & 6 & 0 & \ddots & & & & \\ & 5 & 7 & 1 & & & & \\ & & 0 & 12 & 3 & \ddots & & \\ & & & 0 & 4 & 11 & 0 & \\ & & & & 0 & 1 & 7 & 8 \end{bmatrix} \quad R_1^p = \begin{bmatrix} 0 & & & & & & & \\ 88.5 & & & & & & & \\ 65.6 & 59 & & & & & & \\ 0 & 34.9 & 0 & & & & & \\ 0 & 67 & 0 & \ddots & & & & \\ & 53.7 & 77.9 & 11.4 & & & & \\ & & 0 & 130 & 32.9 & \ddots & & \\ & & & 0 & 43.8 & 121.4 & 0 & \\ & & & & 0 & 9.4 & 66.4 & 76.7 \end{bmatrix}$$

$$R_2^a = \begin{bmatrix} 2 & 0 & & & & & & \\ 8 & 1 & \ddots & & & & & \\ 0 & 14 & 0 & & & & & \\ & 0 & 10 & 0 & \ddots & & & \\ & & 10 & 6 & 0 & & & \\ & & & 3 & 15 & 1 & 0 & \\ & & & & 0 & 0 & 14 & 2 \\ & & & & & 0 & 0 & 9 \\ & & & & & & 0 & 6 \end{bmatrix} \quad R_2^p = \begin{bmatrix} 23 & 0 & & & & & & \\ 88.5 & 11.9 & \ddots & & & & & \\ 0 & 165.2 & 0 & & & & & \\ & 0 & 120.5 & 0 & \ddots & & & \\ & & 115.5 & 70.8 & 0 & & & \\ & & & 34.2 & 173 & 11.6 & 0 & \\ & & & & 0 & 0 & 155.5 & 22.4 \\ & & & & & 0 & 0 & 100.6 \\ & & & & & & 0 & 57.5 \end{bmatrix}$$

$$R_3^a = \begin{bmatrix} 2 & 2 & 3 & 0 & & & & \\ 0 & 0 & 13 & 3 & 0 & & & \\ & \ddots & \ddots & 3 & 3 & 14 & & \\ & & & & 0 & 4 & 0 & \\ & & & \ddots & \ddots & 1 & 7 & 0 \\ & & & & & 0 & 1 & 0 \\ & & & & & \ddots & 5 & 1 \\ & & & & & & 0 & 9 \\ & & & & & & & 1 \end{bmatrix} \quad R_3^p = \begin{bmatrix} 23 & 24.8 & 38.3 & 0 & & & & \\ 0 & 0 & 160.1 & 37.7 & 0 & & & \\ & \ddots & \ddots & 37.3 & 37.8 & 177.6 & & \\ & & & & 0 & 50.1 & 0 & \\ & & & \ddots & \ddots & 12 & 84.6 & 0 \\ & & & & & 0 & 11.7 & 0 \\ & & & & & \ddots & 55.5 & 11.2 \\ & & & & & & 0 & 100.6 \\ & & & & & & & 9.6 \end{bmatrix}$$

$$b^l = [3 \quad 6 \quad 5 \quad 1 \quad 2 \quad 10 \quad 9 \quad 7 \quad 8] \quad p^b = [109.1 \quad 386.7 \quad 542.4 \quad 205.5 \quad 350 \quad 373.4 \quad 407.5 \quad 366.3 \quad 219.6]$$

$$s^l = [17 \quad 16 \quad 14 \quad 20 \quad 13 \quad 12 \quad 19 \quad 11] \quad p^s = [288.6 \quad 416.4 \quad 512.4 \quad 321.4 \quad 287.4 \quad 382.2 \quad 373.6 \quad 378.6]$$

Fig. 2. Selected Results of the Allocation and Pricing Simulation

in the matrices. The total payment of each buyer and the total income of each seller are shown in p^b and p^s. As marked, the total payment of buyer 3 is 109.1.

It can be seen that RAMCoDA is very suitable for resource allocation in Grid. This is due to the following reasons: (a)comparing with the allocation based on the single-item auction, which needs multiple rounds of auctions for a single item of resource to be auctioned off, RAMCoDA performs more efficiently, *e.g.*, the proposed pricing algorithm determines the resource allocation and pricing of 17 traders in a single round in the above simulation; (b)the combination of both advantages of the DA and the CA, including the prevention of a monopoly, while

considering requirements of both sides and permitting bidding for the resource combination requested by the task; and (c)it can achieve the explicit allocation and trade price information as well as the incentive compatibility characteristic, which are missing in a pure CA-based approach.

The achieved incentive compatibility characteristic is explained in the following. Due to the assumption that all participants will bid for resource combinations according to their true valuations, for buyers it holds that *utility = true valuation - trade price = bid - trade price*, while for sellers it holds that *utility = trade price - true valuation = trade price - bid* [7]. If there is a buyer who attempts to obtain a better utility through misrepresenting the true valuation, the results can be classified as follows: If $bid < true\,valuation$, then he might be (1) rejected by the CoDA algorithm because his bid is too low, then his utility becomes 0; (2) still one of the winner bidders, but due to the reason that this bid change will lead to the change of his location within list bl, this causes him to trade with the higher bid sellers and finally obtain the lower utility; (3) still one of the winner bidders, and can get the lower final trade price because of his lower bid, even though he has to trade with the higher bid sellers. If $bid > truevaluation$, he will generally obtain the lower utility except the situation in which he has the chance to trade with an extremely low bid seller. Few participants would like to act in this way.

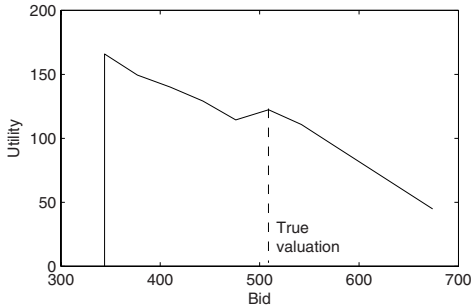


Fig. 3. The utility as a function of the bid

Fig. 3 takes buyer 6 in the previous experiment for example, and shows the utility as a function of his bid. Here assuming all other participants still hold their original bids as shown in Tab.1. It can be seen from Fig. 3, when buyer 6 is under-reporting the true valuation, the fluctuation in the figure is corresponding to the three situations discussed above. He can only get the higher utility under situation (3); however, because there is still the risk of getting a lower utility or even being rejected by the CoDA, and the buyer has no idea about other participants' bids, it will be difficult for him to decide how much to under-report, and arbitrarily under-reporting may actually lower his utility, as shown in Fig. 3. The situation is similar for the sellers. Therefore, bidding the true valuation is still the optimal strategy for each participant.

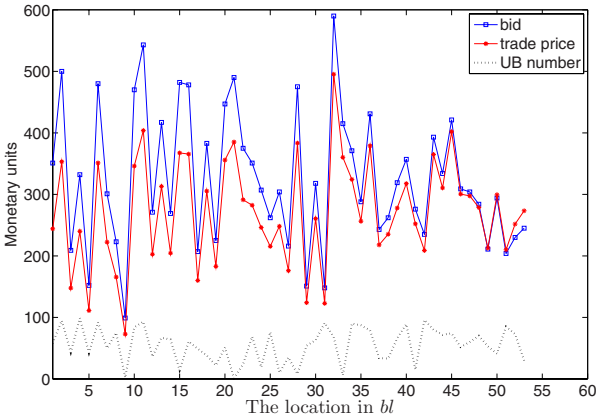


Fig. 4. Results of the Pricing Algorithm (for buyers,n=200)

In order to demonstrate the performance of the proposed pricing algorithm, a resource allocation on a larger scale has been simulated. It is assumed that 200 participants are involved in the CoDA, including 100 UBs and 100 GSPs. For all participants, the value ranges of 3 items of resources are [5, 20], [10, 25], [15, 30] respectively, and $d_i = s_i = 10$ is selected. Since this selection has been done randomly and through solving the CoDA, 53 UBs and 54 GSPs are the winning bidders. In any other case of randomly selected parameter values, different UBs and GSPs would be winning, however, the algorithm proposed would work the same way. Therefore, these results of the algorithm are illustrated in Fig. 4 and Fig. 5, representing a demonstration case. It presents the comparison between the original bid and the trade price. The horizontal axes in Fig. 4 and Fig. 5 represent the corresponding position in b^l and s^l, respectively. Dotted lines in the figure do not relate to the price, they determine the original UB or GSP number corresponding to the number on the horizontal axis.

For example, 21 on the horizontal axis in Fig. 4 denotes the 21th buyer in b^l; from the dotted line, it can be seen that it represents the first (no. 1) UB in the simulation. The original bid and trade price of this UB are 490 and 385.1 monetary units, respectively. It can be seen that the algorithm proposed can complete the resource pricing efficiently, the trend of the trade price obtained is similar to that of the original bid. Moreover, the buyers (sellers) in the front of the list $b^l(s^l)$ can receive a compensation, while the last set of buyers (sellers) will pay (get) the higher (lower) price for the service than their original valuations, i.e., they will get negative utilities. If the trader only enters into a transaction under a positive utility, the trader with the negative utility can quit the current resource allocation, and shall renew its bid in the next round.

Fig. 6 shows the efficiency of RAMCoDA. Three kinds of parameter settings are considered here. For parameter setting 1, the rules of these parameter value

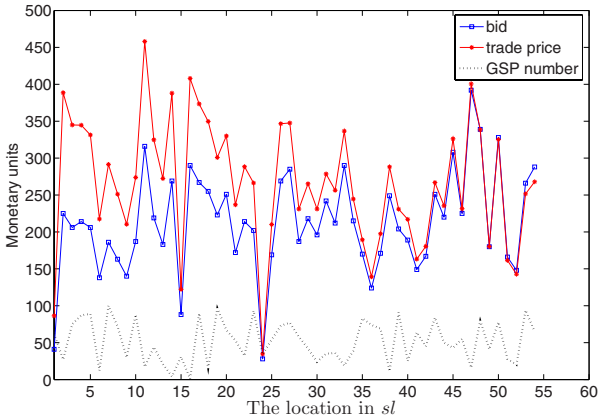


Fig. 5. Results of the Pricing Algorithm (for sellers,n=200)

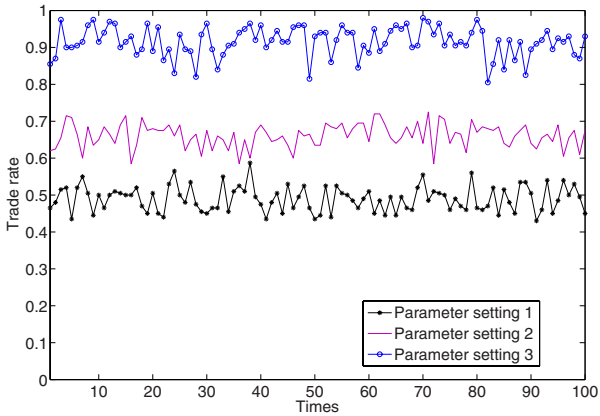


Fig. 6. The trade rate for 3 kinds of parameter settings(100 simulation rounds)

selections are the same as those in Fig. 4 and Fig. 5. For parameter setting 2 and 3, the resource value ranges of UBs change to [10, 25], [15 30], [20, 35] and [20, 30], [25, 35], [30, 40]. Trade rate is defined as the number of final traders divided by the number of the original participants. It can be seen that RAMCoDA can obtain a high trade rate in a one round auction, and the trade rate will increase remarkably when the resource value ranges of UBs is higher than those of GSPs. Since the bids of the UBs are generally higher than those of the GSPs, it is possible for an auction to get more matches between the demand and supply.

Fig. 7 gives the average trade price level for the buyers and sellers for 100 simulation rounds. All the parameters follow the rules of parameter setting 1. The average trade price level is defined as the mean of all the buyers' or sellers' average trade prices. While for each trader, the average trade price equals to

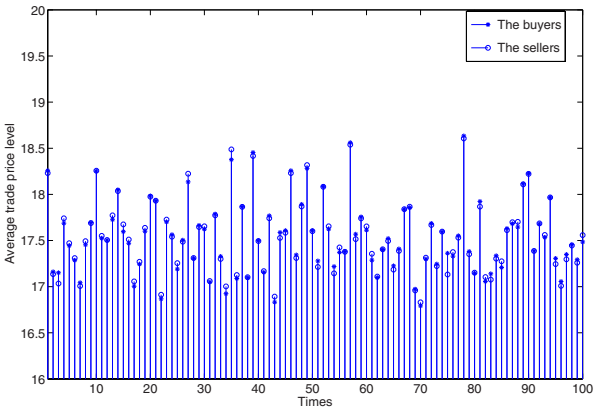


Fig. 7. The average trade price level for the buyers and sellers

the trade price of this trader divided by his total demand or supply. It can be noticed in Fig. 7 that the average trade price levels for the buyers and sellers are similar for a certain time, and in 100 simulation rounds, the fluctuation range of the level is less than 2, which can be regarded as a stable price level since the fluctuation range is less than 13.3% of the average price level.

5 Summary and Conclusions

The newly developed RAMCoDA determines an effective means in support of Grid services resource allocation, accounting, and pricing. RAMCoDA's pricing algorithm is incentive compatible. Analytical investigations presented that the allocation and pricing results need a single round of auctions only, while all requirements of both users and GSPs are taken into account. Finally, the algorithm achieves an explicit allocation and trade price, a high trade rate and a stable price level.

Therefore, RAMCoDA offers a valuable approach for Grid service providers within a commercial situation to market their services. The scheme's simplicity and effectiveness determine reasonable arguments for a practical approach. In addition, the resource usage obtained is fair and can be applied to accounting purposes, thus, GSPs and users will benefit from the approach proposed.

References

1. Das, A., Grosu, D.: Combinatorial auction-based protocols for resource allocation in Grids. In: 19th IEEE International Parallel and Distributed Processing Symposium, Colorado, U.S.A., pp. 23–30 (2005)
2. Ghosh, P., Roy, N., Das, S.K., Basu, K.: A pricing strategy for job allocation in mobile Grids using a non-cooperative bargaining theory framework. Journal of Parallel and Distributed Computing 65(11), 1366–1383 (2005)

3. Li, L., Liu, Y.A., Stiller, B.: Combinatorial double auction-based scheme for resource allocation in Grids. University of Zürich, Department of Informatics, CSG, Technical Report No. 2008.05, Zürich, Switzerland (2008)
4. Pourebrahimi, B., Bertels, K., Kandru, G.M., Vassiliadis, S.: Market-based resource allocation in Grids. In: Second IEEE international Conference on e-Science and Grid Computing, Amsterdam, The Netherlands, pp. 80–88 (2006)
5. Stuer, G., Vanmechelen, K., Broeckhove, J.: A commodity market algorithm for pricing substitutable Grid resources. Future Generation Computer Systems 23(5), 688–701 (2007)
6. Schwind, M., Gujo, O., Stockheim, T.: Dynamic resource prices in a combinatorial Grid system. In: 8th IEEE International Conference on E-Commerce Technology and 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, San Francisco, California, U.S.A., pp. 49–54 (2006)
7. Xia, M., Koehler, G.J., Whinston, A.B.: Pricing combinatorial auctions. European Journal of Operational Research 154(1), 251–270 (2004)
8. Xia, M., Stallaert, J., Whinston, A.B.: Solving the combinatorial double auction problem. European Journal of Operational Research 164(1), 239–251 (2005)
9. Yang, J., Yang, S.B., Li, M.S., Fu, Q.F.: An autonomous pricing strategy toward market economy in computational Grids. In: International Conference on Information Technology: Coding and Computing, Las Vegas, Nevada, U.S.A., pp. 793–794 (2005)
10. Zhao, X.G., Xu, L.T., Wang, B.: A dynamic price model with demand prediction and task classification. In: The Sixth International Conference on Grid and Cooperative Computing, Urumchi, P.R. China, pp. 775–782 (2007)

A User-Influenced Pricing Mechanism for Internet Access

Gergely Biczók and Tuan Anh Trinh

High Speed Networks Laboratory
Dept. of Telecommunications and Media Informatics
Budapest University of Technology and Economics
{biczok,trinh}@tmit.bme.hu

Abstract. Proper pricing schemes are vital components to the continuing success of the Internet. In this paper, we propose a new pricing scheme for Internet access called user-influenced pricing. Our main contribution is threefold: first, we show how user-influenced pricing can provide the ISP with calculable revenues, while giving the users a chance to lower their costs via voting for their preferred pricing scheme. Second, we develop a cooperative weighted voting game which models the decision-making process, and we derive equilibrium solutions to analyze possible outcomes of the vote. Third, we investigate the distribution of power and we show that users with medium generated traffic volume are pivotal to the outcome. Finally, we discuss the practical feasibility of the proposed mechanism regarding user population, revenue planning and charging.

1 Introduction

Advances in networking technology and affordable service prices are continuing to make the Internet a success story both for users and network providers. However, the recently emerged net neutrality debate has shed light on some problems of Internet Service Providers (ISPs) [1]. Since flat-rate billing is dominant and user traffic keeps on growing [14], ISPs get lower profits per data unit carried. An increasing number of news and studies report on the techniques ISP are beginning to look at and use to keep themselves profitable: these include traffic discrimination, introducing download caps and experimenting with alternative pricing schemes (e.g., usage-based pricing, three-part tariffs and charging content providers) [2]. In parallel, there is an ongoing global economic crisis of unseen proportions folding out in the recent months. This downturn makes people think twice about spending more than they absolutely have to. Consequently, ISPs may have to face the fact of decreasing popularity of their services among users. Since economic analysts cannot really predict the length of the global crisis, ISPs have to prepare for a user demand-driven market resulting in diminishing profits, and similarly, customers have to minimize their Internet access costs for an extended period of time.

There is extensive research work dedicated to pricing network services. Some of the papers propose sophisticated pricing models for ISPs to extract consumer

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 48–60, 2009.

surplus [4] [5] [7]. Others argue that simple pricing plans are the only viable ones, since there is a clear user preference towards them [3] [6]. In [8] authors establish the "Price of Simplicity" (PoS) referring to the difference in revenues between a simple pricing scheme (flat-rate) and the maximum achievable revenue. Furthermore, they characterize a range of environments, where PoS is low, i.e., flat-rate pricing is efficient. The authors of [9] show how ISPs can charge content providers for terminating their traffic at their users creating extra income if no net neutrality is enforced.

We take a different approach: our goal is to give ISPs the benefit to plan their revenues, while giving a freedom of choice to the users to shape their own monthly cost. In a certain sense this approach has something in common with packet auctions [5] [7]: we involve users in the pricing process. On the other hand, we do not use a sophisticated auction scheme which makes it harder both for ISPs and users to plan/estimate their revenues and costs, respectively [3]. We also note that in these economically hard times users generating a low traffic volume have a strong incentive to be billed proportionally to traffic volume, contradicting the findings of [6]. Heavy users, of course, prefer sticking to flat rates.

In this paper we propose a user-influenced pricing scheme for ISPs. First, the ISP determines the amount of income it wants to collect in the next billing period, and announces it to the forum of its users. At the same time, it announces the pricing schemes the users can choose from. In this paper we restrict the selection to simple flat-rate and usage-based schemes due to space constraints and tractability. Second, users vote for their preferred pricing scheme. Simple majority decides the outcome of the vote. Finally, the ISP implements the chosen billing method and bills its service accordingly. This simple scheme enables ISPs to get a fixed revenue that can be planned in advance, and gives incentive to users of the same traffic category to cooperate in order to achieve lower monthly costs. We analyze the possible outcomes of the vote in the presence of different user distributions, where different class of users dominate the population.

The remainder of this paper is structured as follows. We introduce the concept of user-influenced pricing in Section 2. A game-theoretical model of the voting process is proposed in Section 3. We study the equilibrium solutions in Section 4. The distribution of voting power is derived using the Shapley value approach in Section 5. Section 6 discusses practical issues, and finally, Section 7 concludes the paper.

2 User-Influenced Pricing

Here we describe how a service provider can use user-influenced pricing to bill its customers. As a first step, the ISP has to set a goal for the next billing cycle (e.g., one month), how much revenue R it wants to collect. This depends on a number of factors. From Section 3 in this paper we do not consider multiple ISPs competing for the same set of users, rather a single ISP in a monopolistic setting. Nevertheless, here we mention that choosing a very high R would certainly drive users away, so there is an incentive to keep the expectations reasonable.

Second, the ISP announces R to its users along with the possible billing options: flat-rate (F) and usage-based (U). Then users vote for the billing scheme they like. We assume that voting is mandatory, non-voting users are punished to pay according to the pricing scheme that is worse for them (e.g., usage-based for non-voting heavy users). The ISP summarizes the votes and announces the chosen pricing scheme for the upcoming billing cycle. During the vote, users can motivate other users to vote with them. We assume that users can utilize financial incentives (side-payments) to sweeten the deal for others, while still profiting from the outcome of the vote.

Third, users use their subscription and pay according to the implemented pricing scheme chosen by the user community. We assume that the decision on the applied billing method does not affect user behavior during the billing cycle. Note that in the rest of the paper we put the voting at the beginning of the billing period because of conformity; however, putting it at the end of the billing period would anneal the need for the above assumption on user behavior.

3 The Game

This section presents a game-theoretical representation of the user-influenced pricing game. Suppose that there is a single ISP on the Internet access market selling a single-tier service. There are n customers, each of them with a fixed monthly traffic amount τ_i measured in bytes. The ISP's goal is to get a monthly revenue of R while serving a traffic volume of T, and it does not care about how users share this total cost. The ISP lets the users decide on the applied pricing scheme: it can be either flat-rate (F) or usage-based (U). The simple majority wins and their preferred pricing scheme will be used to bill all customers. We use a cooperative game with transferable payoffs to model this decision-making process.

3.1 Players

Today's typical ISP has a very diverse set of users. Some users download massive amounts of data via peer-to-peer file sharing systems such as BitTorrent, watch streaming videos frequently through sites like YouTube and play multiplayer online games (e.g., World of WarCraft). Those customers are considered heavy users, they can impose a monthly traffic amount of several hundred of gigabytes on the ISP's network. An other category consists of light users: they just browse the Web and send a couple of e-mails. Light users usually have a monthly traffic amount around 5-10 gigabytes. Somewhat forgotten, between the above categories are people who use their Internet access in an "average" sense. That means an occasional movie download, contacting their relatives via some VoIP application (e.g., Skype), using one or two social networking sites, such as Facebook or MySpace, to keep in touch with friends and colleagues. Those customers are referred to as medium users.

These three groups have different interests when it comes to pricing schemes applied. Obviously, heavy users want to pay a fixed monthly rate, since their

traffic volume is high, so paying per byte would result in huge bills for them. Conversely, light users are interested in paying on-the-go. Since it is likely that they never really consume the bandwidth equivalent of the flat-rate price, they prefer to pay proportionally to their traffic volume. We assume that medium users are indifferent: they pay more or less the same price regardless of the applied pricing scheme.

To reduce the complexity of the game and to provide intuitive results, we model this voting as a three-player game [10]. Player 1 represents the heavy users preferring flat-rate pricing. Let the ratio of heavy users among all consumers be $0 \leq w_1 \leq 1$. Similarly, the ratio of the whole monthly traffic volume imposed on the ISP by heavy users is $0 \leq t_1 \leq 1$.

Player 2 stands for the class of medium users. Their ratio compared to the whole customer population is $0 \leq w_2 \leq 1$. They generate a traffic ratio of $0 \leq t_2 \leq 1$.

Player 3 represents the class of light users preferring usage-based pricing. Their ratio among all users is $0 \leq w_3 \leq 1$, while their traffic ratio is $0 \leq t_3 \leq 1$.

Note that we classify every user as heavy, medium or light, therefore $w_1 + w_2 + w_3 = 1$ (all users are represented), furthermore $t_1 + t_2 + t_3 = 1$ (all traffic is accounted for). An interesting question is how the actual values of parameters w_i and t_i should be chosen. We do not make any further assumptions in our analysis to maintain the generality of our model, but we discuss realistic parameters in Section 6.

Certainly, we lose some behavioral details by introducing our assumptions and simplifications, e.g., by assuming that medium users are indifferent to the actual pricing scheme. Therefore, our results are intended to be qualitative, i.e., we concentrate on the rough behavior of the pricing mechanism and the players.

3

We treat the user-inferred pricing problem as a majority voting game. In our case there is one significant difference to a general cooperative game: the strongly opposed interests of two players, i.e., heavy and light users, induce some non-cooperative aspects referred to as *quarreling*.

The possible coalitions in a general three-player cooperative game are: $\{\{1\}, \{2\}, \{3\}, \{12\}, \{13\}, \{23\}, \{123\}\}$. In our game, heavy users (Player 1) and light users (Player 3) are strategically opposed, thus they will never be a part of the same coalition. Additionally, since there are only two pricing methods offered by the ISP, medium users (Player 2) will always cast a vote, either for flat-rate or usage-based pricing. These constraints eliminate the chance of forming a *grand coalition*, the coalition of $\{2\}$ and also the coalition of the two extremists. The remaining possible coalitions are: $\{\{1\}, \{3\}, \{12\}, \{23\}\}$.

Heavy users clearly choose flat-rate pricing (F), on the other hand, light users always prefer usage-based pricing (U). Since Player 2 is indifferent in choosing either side, the other two players have to give him some incentive to join forces. We model this as a side-payment, which reduces the costs of Player 2. Giving a large side-payment can be crucial to winning the voting game, nevertheless none

of the two quarreling players can pay more for the vote of Player 2 than their payoff expected from the ISP implementing their preferred pricing scheme. Heavy users can offer a side-payment s_1 in the range $[0, (t_1 - w_1)R) \equiv S_1$, where S_1 is the strategy set of Player 1 in the voting game. It is easy to see that the upper limit of the side-payment corresponds to Player 1's profit due to flat-rate pricing. Similarly, the side-payment offered by Player 3 is $s_3 \in [0, (w_3 - t_3)R) \equiv S_3$, where the upper limit is Player 3's profit due to usage-based pricing and S_3 is the strategy set of Player 3. Considering Player 2, we assume that the vote and the side-payment are exchanged at the same time ensuring that Player 2 can only accept one side-payment and it has to vote accordingly. So Player 2's strategy set is $S_2 \equiv \{F, U\}^{S_1 \times S_3}$, i.e., all functions mapping side-payments to votes.

We can now define the payoffs of each player formally. The payoff of heavy users (Player 1) is:

$$\Pi_1(s_1, s_2, s_3) = (t_1 - w_1)RI_1 - s_1 I_2 \tag{1}$$

where

$$I_1 = \begin{cases} 1 & \text{if Player 1 wins} \\ 0 & \text{otherwise} \end{cases}$$

and

$$I_2 = \begin{cases} 1 & \text{if } s_2 = F \\ 0 & \text{if } s_2 = U \end{cases}$$

The payoff of light users (Player 3) is:

$$\Pi_3(s_1, s_2, s_3) = (w_3 - t_3)R(1 - I_1) - s_3(1 - I_2) \tag{2}$$

Note that indicator variables are complemented due to opposing conditions.

Player 2's payoff is the following:

$$\Pi_2(s_1, s_2, s_3) = \begin{cases} s_1 & \text{if } s_2 = F \\ s_3 & \text{if } s_2 = U \end{cases} \tag{3}$$

Now we formulate the characteristic function using the standard approach, keeping in mind that certain coalitions of players are not reasonable because of quarreling. Those coalitions receive zero utility, formally:

$$\nu(H) = 0 \quad | \quad C \notin \{\{1\}, \{3\}, \{12\}, \{23\}\} \quad \text{and} \quad H \in 2^N. \tag{4}$$

For the reasonable coalitions the corresponding utilities are:

$$\begin{aligned} \nu(\{1\}) &= \max_{s_1} \min_{s_2, s_3} \Pi_1(s_1, s_2, s_3) \\ \nu(\{3\}) &= \max_{s_3} \min_{s_1, s_2} \Pi_3(s_1, s_2, s_3) \\ \nu(\{12\}) &= \max_{s_1, s_2} \min_{s_3} [\Pi_1(s_1, s_2, s_3) + \Pi_2(s_1, s_2, s_3)] \\ \nu(\{23\}) &= \max_{s_2, s_3} \min_{s_1} [\Pi_2(s_1, s_2, s_3) + \Pi_3(s_1, s_2, s_3)] \end{aligned} \tag{5}$$

Table 1. Characteristic function for the user-influenced pricing game (w_1 and t_1 are the population ratio and traffic ratio of heavy users, while w_3 and t_3 are those of the light users, respectively)

Characteristic function	Heavy user regime ($w_1 > 1/2$)	Balanced regime ($w_1 < 1/2, w_3 < 1/2$)	Light user regime ($w_3 > 1/2$)
$\nu(\{1\})$	$(t_1 - w_1)R$	0	0
$\nu(\{3\})$	0	0	$(w_3 - t_3)R$
$\nu(\{12\})$	$(t_1 - w_1)R$	$(t_1 - w_1)R$	0
$\nu(\{23\})$	0	$(w_3 - t_3)R$	$(w_3 - t_3)R$
$\nu(H)$ for all other $H \in 2^N$	0	0	0

Using Equations 4 and 5 we compile the characteristic functions presented in Table 1. Different columns represent different user distributions in the population. If heavy users are a majority ($w_1 > 1/2$) they will dominate voting (heavy user regime). If light users are a majority ($w_3 > 1/2$) they will be the dominant player (light user regime). If neither of the above are true ($w_1 < 1/2, w_3 < 1/2$, but due to constraints of w_i, $w_1 + w_2 > 1/2, w_2 + w_3 > 1/2$), the players enter a balanced regime, where the outcome of the pricing game will be decided by the offered side-payments.

4 Equilibrium Solutions

Here we derive the equilibrium solutions for the user-influenced pricing game G. Since G includes players that will never form a coalition, we employ the notion of ψ-allowable coalition structures [12]. Let P be a partition of N, called a *coalitional structure*. The possible partitions are: $\{\{1\},\{2\},\{3\}\}, \{\{12\},\{3\}\}, \{\{1\},\{23\}\}, \{\{123\}\}$. Then we define the set of *allowable coalitional structures* ($\psi(P)$) that satisfy the constraints imposed by quarreling. For G

$$\psi(P) = \psi = \{(\{12\},\{3\}),(\{1\},\{23\})\}. \tag{6}$$

For a given $P \in \psi$, let $X(P)$ be the set of imputations as follows:

$$X(P) = \{(x_1, x_2, x_3) \in R^3 \mid \sum_{i \in H} x_i = \nu(H) \text{ for all } H \in P \text{ and } x_i \geq \nu(\{i\}) \text{ for } i = 1,2,3\} \tag{7}$$

Intuitively an imputation is a distribution of the maximum side-payment such that each player receives at least the same amount of money that they can get if they choose to stay alone (individual rationality), and each coalition in the structure P receives the total side-payment they can achieve (group rationality).

Now, we restrict the set of imputations to the *core* $C(P)$. The core is defined to be the set of undominated imputations. To put it differently, the core is the

set of imputations under which no coalition has a value greater than the sum of its members' payoffs. Formally:

$$C(P) = \left\{ (x_1, x_2, x_3) \in X(P) | \sum_{i \in H} x_i \geq \nu(H) \text{ for all } H \in \bigcup \{J \in P \mid P \in \psi\} \right\} \tag{8}$$

Considering our game G, Equation 8 is equivalent to the standard core (since $\nu(H) = 0$ for unreasonable coalitions), so

$$C(P) = \left\{ (x_1, x_2, x_3) \in X(P) \mid \sum_{i \in H} x_i \geq \nu(H) \text{ for all } H \in 2^N \right\} \tag{9}$$

As it can be noticed the core is dependent on a certain coalitional structure P. For us to determine which structure will emerge when playing the game, we define a ψ-stable pair $[\overline{x}, P]$:

$$[\overline{x}, P] \mid \overline{x} \in C(P), P \in \psi \tag{10}$$

Now applying this solution to the characteristic function $\nu(H)$ in Table 1, we have three different cases depending on user regimes.

4.1 Heavy User Regime

In this case heavy users are dominant in the population, thus $w_1 > 1/2$. The only possible imputation is $\overline{x} = ((t_1 - w_1)R, 0, 0)$ hence there are two ψ-stable pairs:

$$[((t_1 - w_1)R, 0, 0), \{\{12\}, \{3\}\}] \text{ and } [((t_1 - w_1)R, 0, 0), \{\{1\}, \{23\}\}]$$

Note that both coalitional structures are possible, since it does not matter which side medium users take.

In words, this means heavy users dominate the voting, no side-payment is transferred. Considering the individual user's point of view, let c_i denote the cost of a single user i. Flat-rate pricing is implemented by the ISP, Internet access costs are shared per capita, hence the cost for a single user is independent of his traffic and equal for every user is

$$c_i = \frac{R}{n} \text{ for all } i \in N \tag{11}$$

4.2 Light User Regime

Here light users have the absolute majority across the population ($w_3 > 1/2$). Following the same line of thought as in Section 4.1 we derive the ψ-stable pairs for this case:

$$[(0, 0, (w_3 - t_3)R), \{\{1\}, \{23\}\}] \text{ and } [(0, 0, (w_3 - t_3)R), \{\{12\}, \{3\}\}]$$

As expected light users dominate the voting, no side-payment is made to medium users. From a single user's perspective, let τ_i denote the traffic volume of user i. Since usage-based pricing is implemented by the ISP, Internet access costs are shared proportionally to traffic volume. Therefore the access cost for user i is

$$c_i = \frac{\tau_i}{T} R \text{ for all } i \in N. \tag{12}$$

4.3 Balanced Regime

In this case cooperation is explicitly needed to form a winning coalition, since $w_1 < 1/2, w_3 < 1/2$, and $w_1 + w_2 > 1/2, w_3 + w_2 > 1/2$. Side-payments determine the outcome of the voting game. For easier analysis let $s_1^{\max} = (t_1 - w_1)R$ and $s_3^{\max} = (w_3 - t_3)R$ be the maximum reasonable side-payment possibly offered by Player 1 and Player 3, respectively. The imputations and the core for any $s_1^{\max}, s_3^{\max}$ are :

$$X(\{1\},\{23\}) = \{(x_1, x_2, x_3) \in R^3 \mid x_1 = 0, x_2 \geq 0, x_3 \geq 0, x_2 + x_3 = s_3^{\max}\}$$

$$C(\{1\},\{23\}) = \begin{cases} \emptyset, & \text{if } s_1^{\max} > s_3^{\max} \\ (0, s_1^{\max} + \epsilon, s_3^{\max} - s_1^{\max} - \epsilon), & \text{if } s_3^{\max} \geq s_1^{\max} \end{cases}$$

where $0 \leq \epsilon \leq s_3^{\max} - s_1^{\max}$. Furthermore:

$$X(\{12\},\{3\}) = \{(x_1, x_2, x_3) \in R^3 \mid x_1 \geq 0, x_2 \geq 0, x_3 = 0, x_1 + x_2 = s_1^{\max}\}$$

$$C(\{12\},\{3\}) = \begin{cases} \emptyset, & \text{if } s_1^{\max} < s_3^{\max} \\ (s_1^{\max} - s_3^{\max} - \epsilon, s_3^{\max} + \epsilon, 0), & \text{if } s_1^{\max} \geq s_3^{\max} \end{cases}$$

where $0 \leq \epsilon \leq s_1^{\max} - s_3^{\max}$.

Let us first study the coalitional structure $(\{1\},\{23\})$. The core is empty if the maximum side-payment of Player 3 is smaller than that of Player 1. This is due to the fact that Player 2 wants to form a coalition with Player 1 and get more money than $s_3^{\max}$, but the constraint on imputations prevents this. On the other hand, if the maximum side-payment of Player 3 is greater than Player 1's, than the core is non-empty with Player 3 (the light users) winning, and the game G is *balanced*. Player 3 pays $s_1^{\max} + \epsilon$ to Player 2 and retains $s_3^{\max} - s_1^{\max} - \epsilon$. A similar (but opposing) explanation applies for the coalitional structure $\{\{12\},\{3\}\}$.

The solution of the user-influenced pricing game is given as ψ-stable pairs in Table 2. Note that the ψ-stable concept does not restrict the possibilities. In the first row of the table heavy users win (flat-rate pricing is chosen), but a side-payment of at least $s_3^{\max}$ has to be paid. According to the third row, light users win by paying at least $s_1^{\max}$ to medium users. If the maximum side-payments are equal, the outcome is indeterminate.

Now, let us take a look at how individual users can share the burden of side-payments in the balanced regime. Let $H, M, L \subset N$ be the set of heavy, medium and light users.

Table 2. ψ-stable pairs in the balanced regime

Side-payment parameters	Core solution	Coalitional structure
$0 < s_3^{\max} < s_1^{\max}$	$(s_1^{\max} - s_1, s_1, 0)$	$(\{12\}, \{3\})$
$0 < s_3^{\max} = s_1^{\max}$	$(0, s_1^{\max}, 0)$	$(\{12\}, \{3\})$ or $(\{1\}, \{23\})$
$0 < s_1^{\max} < s_3^{\max}$	$(0, s_3, s_3^{\max} - s_3)$	$(\{1\}, \{23\})$

Flat-rate pricing. Suppose that $s_3^{\max} < s_1^{\max}$, hence heavy and medium users team up to implement flat-rate pricing. A suitable division of side-payments among heavy users would be to share the additional cost equally, resulting in a payment of $\frac{s_1}{|H|}$ for each heavy user i. Also by choosing the flat-rate approach, medium users share the profit from the side-payments equally, each medium user getting $\frac{s_1}{|M|}$.

Now we can give the monthly cost of a single user:

$$c_i = \begin{cases} \frac{R}{N} + \frac{s_1}{|H|} & \text{if } i \in H \\ \frac{R}{N} - \frac{s_1}{|M|} & \text{if } i \in M \\ \frac{R}{N} & \text{if } i \in L \end{cases} \tag{13}$$

Usage-based pricing. Suppose that $s_1^{\max} < s_3^{\max}$, therefore light and medium users join forces to achieve usage-based pricing. A suitable division of side-payments among light users would be to share the the additional cost proportional to traffic volume, resulting in a payment of $\frac{s_3 \tau_i}{t_3 T}$ for each light user i. Also by choosing the usage-based approach, medium users can agree to benefit from the side-payments proportionally to their traffic volume, so each medium user j user gets $\frac{s_3 \tau_j}{t_2 T}$.

Now we can give the monthly cost of a single user:

$$c_i = \begin{cases} \frac{R\tau_i}{T} & \text{if } i \in H \\ \frac{R\tau_i}{T} - \frac{s_3 \tau_i}{t_2 T} & \text{if } i \in M \\ \frac{R\tau_i}{T} + \frac{s_3 \tau_i}{t_3 T} & \text{if } i \in M \end{cases} \tag{14}$$

5 Distribution of Power

This section reveals the distribution of voting power in the game G. The usual approach is to calculate the *Shapley value*:

$$\phi_k[\nu] = \sum_{S \subset N} \gamma(n, s) V_k(S) \tag{15}$$

with

$$\gamma(n, s) = \frac{(s-1)!(n-s)!}{n!}, \text{ and } V_k(S) = \nu(S) - \nu(S \setminus \{i\}) \tag{16}$$

where $s = |S|$ and $n = |N|$.

Here, the existence of quarreling players prevents us to use the original Shapley value. Fortunately, a *modified Shapley value* incorporating quarreling has been developed in [11]. This modified value represents an expected distribution of side-payments (x_1, x_2, x_3) in G, when the players arrive in random fashion to join coalitions, and receive their marginal worth to the coalition. The modified Shapley value employs the constraint of a quarreler not joining a coalition where an other quarreler is already present: then he receives no payoff. Formally for Player j:

$$\phi_j^\star[Q,\nu] = \sum_{S\cap Q=j} \gamma(n,s)V_j(S) + \frac{q-1}{q}\nu(j),\ j \in Q \tag{17}$$

and

$$\phi_j^\star[Q,\nu] = \sum_{S\cap Q=\emptyset} \gamma(n,s)V_j(S) + \sum_{|S\cap Q|=1} \frac{\gamma(n-q,s-1)-\gamma(n,s-1)}{q} V_j(S),\ j \notin Q \tag{18}$$

where Q is the set of quarrelers and $q = |Q|$.

Now, we can calculate the modified Shapley values. For Player 1:

$$\begin{aligned}\phi_1^\star &= \frac{1!\cdot 2!}{3!}\nu(\{1\}) + \frac{1!\cdot 1!}{3!}(\nu(\{12\}) - \nu(\{2\})) + \frac{1}{2}\nu(\{1\}) = \\ &= \frac{5\nu(\{1\}) + \nu(\{12\}) - \nu(\{2\})}{6}\end{aligned} \tag{19}$$

Similarly for Player 3:

$$\phi_3^\star = \frac{5\nu(\{3\}) + \nu(\{23\}) - \nu(\{2\})}{6} \tag{20}$$

Finally, for non-quarreling Player 2:

$$\phi_2^\star = \frac{\nu(\{12\}) + \nu(\{23\}) + \nu(\{2\}) - \nu(\{1\}) - \nu(\{3\})}{3} \tag{21}$$

Evaluating the modified Shapley values for the different regimes depending on the relation of $s_1^{\max}$ and $s_3^{\max}$, we get the distributions of power based on Table 1.

Under the heavy user regime the modified Shapley value is $(s_1^{\max}, 0, 0)$, heavy users have all the power. Similarly for the light user regime, the value is $(0, 0, s_3^{\max})$, light users are in total control. In the balanced regime the power is shared with a modified Shapley value of $(\frac{s_1^{\max}}{6}, \frac{s_1^{\max}+s_3^{\max}}{3}, \frac{s_3^{\max}}{6}$. Note that for $s_1^{\max} > s_3^{\max}$ heavy users have more power than light users, and for $s_3^{\max} > s_1^{\max}$ the opposite is true. Most importantly, irrespective of the maximum offered side-payments, Player 2 is the most powerful since he is a *pivotal* player. His power grows twice as fast as the other players if side-payments begin to grow.

An other measure of power is the *Shapley-Shubik index* [13]. Suppose that voters arrive in a random order, until a pivotal player turns a losing coalition

into a winning one. The Shapley-Shubik index is then the proportion of orders where the player is pivotal, formally:

$$\phi_i^{SS} = \frac{p_i}{n!} \tag{22}$$

where p_i is the number of occasions where Player i is pivotal. Note that we restrict possible coalitions to those where quarreling players are not together. It is easy to see that under the heavy user and light user regimes the Shapley-Shubik power index is $(1, 0, 0)$ and $(0, 0, 1)$, respectively. In the balanced regime, side-payments determine the outcome: if $s_1^{\max} > s_3^{\max}$ the index is $(0.5, 0.5, 0)$; symmetrically if $s_1^{\max} < s_3^{\max}$ the index is $(0, 0.5, 0.5)$. To put in words, the Shapley-Shubik index shows the importance of a player in a winning coalition: medium users are just as important as heavy users if flat-rate pricing is voted for, and they have the same importance as light users if usage-based pricing prevails.

6 Discussion on Feasibility

Here we give an outlook at the practical issues that can be raised by the actual implementation of the proposed pricing scheme.

Composition of user population. During our analysis in Section 3 to 5 we have not assumed any particular composition of the user population, and we studied the entire parameter space. In practice though, the composition can decide the outcome by itself, hence the notions of heavy and light user regimes. Of course, a lot depends on how different user classes are defined. An exact definition is out of scope for this paper, but a rule of thumb is presented in Section 3. Note that heavy-hitters still dominate overall traffic, but their shares are decreasing, while other users are catching up due to multimedia content, resulting in a more balanced user distribution based on generated traffic volume [14]. In other words, the existence of a balanced regime is highly likely.

Planning income. First of all, can an ISP efficiently estimate its future revenues? It is common sense that companies do plan their revenues and expenses in advance. The difference here is that the ISP actually gets the exact amount of money they planned for. Coming up with a single number every month is not straightforward; a small provider has some advantage over its larger counterpart in this sense, since smaller ISPs tend to have a simpler business and service structure.

Voting and charging. How can the announcement process be implemented? Also, is there a reasonable method to distribute side-payments among the users? We believe if an ISP uses the proposed method, it is in its best interest to provide for the announcement, negotiations and the voting process. The voting process can be entirely web-based. This requires strong identities and a secure infrastructure. Since such infrastructures already exist, it is reasonable to assume

that the ISP can fulfill all the requirements. It is important to emphasize that users do not have to explicitly administer side-payments: the ISP can calculate and incorporate side-payments into their monthly bill.

Future work. We see the presented mechanism as a first step towards a user-controlled pricing system. We plan to plug slightly more complicated schemes, such as three-part tariffs, to the framework of user-influenced pricing introducing further benefits both for the user and the provider. Also, we plan to conduct a simulation study on a competitive market setting where multiple ISPs are present. This will enable us to evaluate the proposed scheme without the limitations introduced by the analytical model.

7 Conclusion

In this paper we presented results on a novel pricing mechanism for Internet Service Providers called user-influenced pricing, where users can vote for the exact pricing scheme implemented in the next billing cycle. Our assumptions were that ISPs want plannable revenues, while users want to keep their costs low. We showed that under user-influenced pricing, users of different traffic volumes (heavy, medium, light) can cooperate to achieve lower costs utilizing side-payments. We modeled this process as a weighted cooperative voting game, and derived the equilibrium solutions and payoffs on the individual user and group level. We showed how the ratio of different users and maximum reasonable side-payments affect the outcome of the voting game. We also derived the distribution of power in various regimes of the game. Results indicate that medium users are pivotal in the decision-making process.

References

1. Peha, J.M.: The Benefits and Risks of Mandating Network Neutrality, and the Quest for a Balanced Policy. In: Proc. of TPRC (2006)
2. Dhamdere, A., Dovrolis, C.: Can ISPs be profitable without violating "network neutrality"? In: Proc. of ACM NetEcon (2008)
3. Odlyzko, A.M.: Internet pricing and the history of communications. Computer Networks 36, 493–517 (2001)
4. Shenker, S., Clark, D., Estrin, D., Herzog, S.: Pricing in Computer Networks: Reshaping the Research Agenda. ACM Computer Communication Review 26, 493–517 (1996)
5. MacKie-Mason, J.K., Varian, H.R.: Pricing the Internet. In: Public Access to the Internet, pp. 269–314. MIT Press, Cambridge (1995)
6. Varian, H.R.: The Demand for Bandwidth: Evidence from the INDEX Project. In: Crandall, R.W., Alleman, J.H. (eds.) Broadband: Should We Regulate High-Speed Internet Access?, pp. 39–56. Aei-Brookings Joint Center for Regulatory Studies, American Enterprise Institute (2003)
7. Yang, X.: Auction, but Don't Block. In: Proc. of ACM NetEcon (2008)
8. Shakkottai, S., Srikant, R., Ozdaglar, A., Acemoglu, D.: The Price of Simplicity. IEEE JSAC 26(7), 1269–1276 (2008)

9. Cheng, H., Bandyopadhyay, S., Guo, H.: The Debate on Net Neutrality: A Policy Perspective, Working Paper (2008)
10. Jarrow, R.A., Leach, J.C.: Bribes, Power, and Managerial Control in Corporate Voting Games. Theory and Decision 26, 235–251 (1989)
11. Kilgour, D.M.: A Shapley Value for Cooperative Games with Quarrelling. In: Game Theory as a Theory of Conflict Resolution. D. Reidel Publishing Company, Dordrecht (1974)
12. Luce, R.D., Raiffa, H.: Games and Decisions. John Wiley, Chichester (1957)
13. Shapley, L.S., Shubik, M.: A Method for Evaluating the Distribution of Power in a Committee System. American Political Science Review 48, 787–792 (1954)
14. Cho, K., Fukuda, K., Esaki, H., Kato, A.: Observing Slow Crustal Movement in Residential User Traffic. In: Proc. of ACM CoNEXT (2008)

Price Setting in Two-Sided Markets for Internet Connectivity

Thorsten Hau and Walter Brenner

Institute of Information Management, University of St. Gallen
Müller-Friedberg-Strasse 8, CH-9000 St. Gallen, Switzerland
Tel.: +41 (0)71 224 3803; Fax: +41 (0)71 224 3296
thorsten.hau@unisg.ch

Abstract. Due to a lack of incentives, Internet peerings are a notorious bandwidth bottleneck. Through the use of direct interconnection and content delivery networks, content providers are able to provide better services to their customers. These technologies have a profound impact on the business models of internet service providers. Instead of competing for consumers and keeping uplink connection costs low, ISPs face a two-sided market in which they compete for EUs and generate revenues on the CP side of the market. This work presents a formal model for the providers' pricing decision towards content providers and discusses consequences for the Internet.

Keywords: Quality of Service, Network Economics, Peering, Internet.

1 Introduction

The Internet is made up of many independent sub-networks – so called "autonomous systems" (AS). Generally speaking these ASs correspond to different carriers or Internet service providers (ISPs): firms that own and operate the infrastructure (cables and routers) that make up the Internet. These ISPs have customers who are either content providers (CP) with mostly outgoing traffic or end-users (EU) with mostly incoming traffic. To form the Internet, each ISP offers all of its customers' connectivity to all of the other ISPs' customers. In order to uphold this universal connectivity, the ISPs have to exchange user traffic, an activity that is governed by contractual agreements between the ISPs and physically enabled by infrastructure that interconnects their networks. Thus, even though the Internet consists of many independent sub-networks, each user can reach every website on the web. However, it is also well known, that ISPs are not very cooperative in their peering behavior [7]. The decision to interconnect usually is more expensive for one party than for the other and therefore peerings tend to have smaller capacity than what would be optimal.

The rules regulating the exchange of traffic between ISPs have been subject to extensive treatment in the literature. Issues like hot potato routing [22, 24] and determination of access charges have been extensively studied and are quite well understood [8, 10, 17]. However, the available literature studies an idealized Internet in which

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 61–71, 2009.

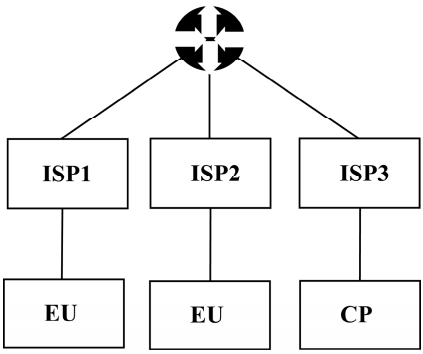


Fig. 1. Physical Interconnection 1

there are EUs, CPs and ISPs that have relations as depicted in fig. 1. The Internet is modeled as a strictly hierarchical system in which traffic flows from a CP to its ISP, is exchanged with the requesting user's ISP and is then sent to the EU.

A key characteristic of this setup is that traffic is exchanged through a peering point. Depending on the contract between two ISPs this traffic exchange may happen in exchange for a payment or "for free". Due to lacking incentives to extend peering capacity sufficiently, these peerings represent major traffic bottlenecks [1, 7, 15].

In contrast to the available work [17, 18, 23], this paper focuses on two important variations of this idealized model of Internet infrastructure as shown in figures 2 and 3. The existing literature has ignored the possibility that content providers and terminating internet service providers interconnect directly.

There are two modes of "direct interconnection" that we will consider. Firstly, a content provider can directly buy transit from the terminating ISPs, thus effectively paying them for preferential access to end-users. This practice shown in fig. 2 is called multi-homing (MH) and contributes to exponential growth of routing tables [5]. Secondly, content delivery networks (CDNs) shown in fig. 3 are a popular way to enhance the flow of information on the Internet. A CDN uses local caches to keep distributed images of content close to EUs without the need to traverse several ISPs' networks [26]. Both technologies provide viable means to improve the speed and reliability of data transport from a CP's website to EUs. They allow bypassing peerings and gaining more direct access to the EUs, thus increasing the probability of timely delivery of data to the end-user. The motivation to use CDN or MH is to provide better quality of service (QoS) with the following chain of causality: Traversing peerings degrades user experience by creating delays → QoS access to EUs through CDN and MH creates better user experience → more visitors to a website → higher revenues from selling ad space on the website.

This paper uses the economic theory of two-sided markets to understand the pricing decision an ISP has to make with respect to the charges levied on the CP side of the market in settings such as those in Fig. 2 and 3. Neither the Internet as a whole, nor individual internet service providers (ISPs) can straightforwardly be considered platforms that optimize their revenue from two sides of a market. With the standard Internet business model, each terminating ISP lacks the power to charge content

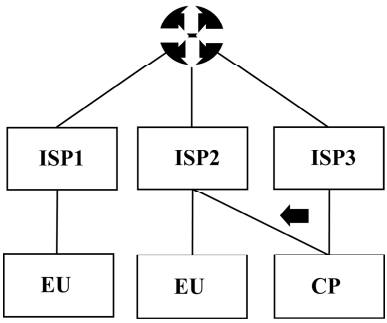


Fig. 2. Physical Interconnection 2

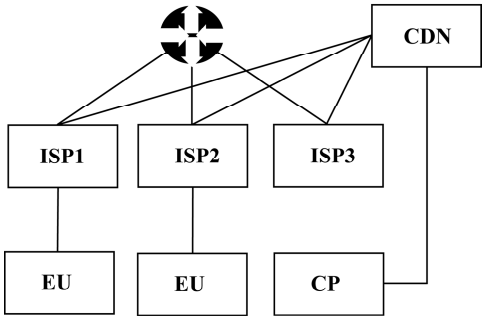


Fig. 3. Physical Interconnection 3

providers that are signed up with another ISP. There are technical as well as contractual barriers to charge some remote content provider for single data packets it sends to an ISP's network. The access charge (the interconnection fee) exchanged between two ISPs is only an imperfect tool to exploit an access monopoly on the Internet due to the fact that it is often reciprocal or zero for external reasons [2, 13, 17]. This is a key difference between the Internet and telephone services (PSTN), where for each call, sender and receiver can be identified and billed per unit of time and a per unit settlement between providers is possible [4, 6, 10]. With CNDs or MH, the property of PSTN that the participating parties can be identified (and are billable customers) is recreated.

The CDN is a third party mediating between CP and ISP but the ISP can charge the CDN for delivery of traffic which will pass this cost on to its CPs. This situation is different from the case of access charges between different providers (as analyzed by [17] with a focus on the Internet or [9] with a focus on PSTN) because in a CDN relationship there is no reciprocity or two way access which is an important condition for that model to be applicable. For the rest of this article we simplify the role that CDNs play on the Internet by treating them as pure mediators between atomic CPs and ISPs. They aggregate CP demand but do not engage in strategic actions. This simplification allows us to model the situation of fig.1 and fig. 2 in the same way. In the last section of this work we sketch a path to relaxing this rather strong assumption.

The paper is structured as follows: Firstly we review the relevant literature on two-sided markets and related topics in telecommunications pricing. Then we explain the abstracted situation we wish to understand and motivate our use of a two-sided market model. Thirdly, we present a formalized model for an ISP facing a two-sided market, deriving results from the market setup. We derive optimal prices charged by ISPs to CPs and CDNs that wish to directly interconnect with them. Lastly we summarize our findings and discuss implications and future research topics.

2 Literature Review

Armstrong's discussion of competition in two-sided markets [3] provides much of the foundation for this work. Two-sided markets are markets where a platform optimizes profit across two distinct sets of customers instead of just one. In the credit

card industry, the card issuing company would be the platform and the merchants accepting the card constitute one group of customers while the buyers using the card to pay form the other. Armstrong analyzes three distinct settings with different customer behaviors and levels of platform competition. The situation relevant for this work is termed "competitive bottleneck": One group of customers can use any number of platform providers simultaneously, while the other group chooses only one of the competing platforms. In our problem, this situation corresponds to EUs being subscribed to only one single ISP while CPs can deal with any number of ISPs at the same time.

Rochet et al. [21] provide a comprehensive overview of the current literature on two-sided markets. They define two-sided markets as markets in which not only the total price but also the price structure influences the number of transactions on the market. For the case at hand, the ISP provides the platform on which transactions between EUs and CPs can take place. They also provide definitions for membership and usage externalities. In the first case one party profits from the sheer presence of the other, while a usage externality is a network effect that arises in an transaction between members of the two sides. They also discuss the effects of fixed and variable prices on the platform. Since variable prices reduce the externality exerted by one group of customers on the other, participation incentives are reduced.

Laffont et al. [17] are not directly concerned with two-sided markets. This work analyzes the access charge paid from one ISP to another for passing traffic on to that ISP's network. In their model the ISP optimizes the prices it charges to CPs and EUs subject to the access charges it pays (for sending traffic to an EU on another ISP's network) and receives (for terminating traffic with its own EUs). In their model the access charge turns out to be a pure tool for reallocating termination costs between EUs and CPs. In the common case of zero access charges all termination costs are born by the EU which corresponds to a subsidy to CPs.

Musacchio et al. [19] compare the effects of single and multi-homing of CPs. They provide explicit formulations of welfare under both regimes and offers results for an economy with many ISPs. However, they do not model EU and CP demand separately but base their model on the assumption of click rates of EUs as a measure of demand for both customer groups and only differentiate CPs from EUs via the per-click price.

3 Problem Description

This work uses the theory on two-sided markets to explore two special cases of interconnection that are different from the symmetric and reciprocal case studied by [17]. The standard model of Internet traffic exchange as shown in fig. 1 follows the pattern $\mathrm{CP} \rightarrow \mathrm{ISP_o} \leftrightarrow_{\mathrm{a}} \mathrm{ISP_t} \leftarrow \mathrm{EU}$ (t=terminating, o=originating, a=access charge) as shown in fig. 4. CP and EU pay their respective ISP and the ISPs exchange traffic for a fee a. This scheme ignores the source of the CP's funding and emphasizes the analysis of the inter-ISP settlement a, which has an influence on the prices paid to the ISPs. By contrast, this work focuses on the setup $\mathrm{EU} \rightarrow \mathrm{ISP}_t \leftarrow \mathrm{CP} \leftarrow Adv.$ as shown in fig. 5. EUs derive utility from high quality of service (QoS) access to CPs' websites while

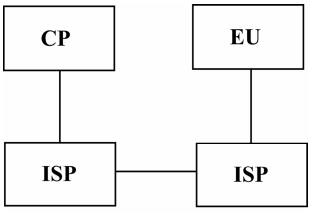


Fig. 4. Business View 1

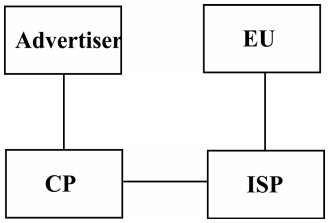


Fig. 5. Business View 2

CPs generate profits from selling ad space to third parties. There is no monetary flow between CPs and EUs. Both, however, may exchange money with the ISP, which acts as a profit maximizing platform. This situation corresponds to the majority of today's Internet business models. CPs create websites that appeal to many EUs, thus generating page views that translate into value of ad-space on that site. (Fig. 1, 2 and 3 focus on physical interconnection while Fig. 4 and 5 depict the business view on connection relationships.)

This CP business model has received wide attention in the two-sided markets literature as it corresponds to the business model of newspapers [3, 20, 21]. This work, however, does not consider the business model of the CP but that of the ISP. CPs pay a transaction-independent price for direct connection to EUs through buying bandwidth from the terminating ISP. EUs on the other hand pay a flat rate fee to the ISP to be connected to the Internet and no transaction based fee for viewing content. There is no payment between EU and CP. The case with payments between the CPs and the EUs has been analyzed in [12].

In the sense of the two-sided market literature we have the following setup: Platform = ISP, single homing side = EU, multi-homing side = CP. The platform charges both sides a lump sum fee for facilitating transactions. This is more reasonable than a linear fee since for EUs, flat rates are the common pricing model and CPs commonly buy a certain bandwidth or a fixed traffic volume. Furthermore the price for Internet services delivered by an ISP might depend on the data volume but rarely on the value of a transaction. Therefore we assume that there is no linear payment that reduces the size of the externalities exerted on the other side of the market, respectively.

4 ISPs as Platforms in Two-Sided Markets

The analysis in this section is related to the competitive bottleneck case of [3]. Competitive bottlenecks arise, when one firm has a monopoly over access to certain customers. Suppose there are two ISPs in a region denoted $\mathrm{ISP}_i, i\epsilon\{1;2\}$. There are also two groups of agents. Group one agents are called end-users (EUs) while group two members are called content providers (CPs) or websites. There is a fraction n_j^i of agents of Group j participating on platform i. In other words, $\mathrm{ISP_i}$ has n_1^i subscribed EUs and n_2^i directly interconnecting customers from the CP side.

The setup is such that two ISPs are present in a market and serve two distinct groups of EUs with Internet connectivity. EUs are single-homing with their ISP. This

means that they are only subscribed with one ISP at a given time. CPs on the other hand multi-home. They may be connected to zero, one or two ISPs in order to reach potential customers (EUs).

To analyze this situation we start by modeling the target function of two ISPs that compete in a market for EUs. The ISPs maximize their respective profits. Symbolically,

$$\pi^i = n_1^i p_1^i + n_2^i p_2^i - C^i(n_1^i, n_2^i), i \in \{1;2\} \tag{1}$$

which is a function of the number of EUs times the price they have to pay, plus the number of CPs times the price they have to pay minus the cost for connecting the two types of customers. The fraction $n_1^i \in [0,1]$ of EUs that are customers of ISP$_i$ is given as a function of the utilities offered by the two ISPs:

$$n_1^i = \phi^i(u_1^i u_1^j) = \frac{1}{2} + \left(\frac{u_1^i - u_1^j}{2t}\right), \forall i \neq j\,. \tag{2}$$

The function ϕ^i is thus increasing in the first argument and decreasing in the second. Note that $n_1^i + n_1^j = 1$ holds since EUs do not multi-home. To specification of EU-demand in equation (2), i.e. the fraction of EUs that are signed up with either ISP is described in a Hotelling [14, 25] way. This implies that the two ISPs share the market equally if they are undifferentiated from the consumers' point of view. If, however, one ISP offers superior utility, it can capture more than half of the market.

The utility EUs get from subscribing to ISP$_i$ is given by

$$u_1^i = U^i(n_2^i) - p_1^i = \alpha_1 n_2^i - p_1^i\,. \tag{3}$$

It equals the gross utility they get from being connected with superior QoS to n_2^i directly interconnected CPs minus the price they have to pay for that connection. The function U^i is increasing in n_2^i since more content in better quality is always better than less. The parameter α_1 can be interpreted as the utility an EU derives from being able to reach one high QoS CP. The EUs perceive the ISP with more CPs connected with QoS as providing a better connection to the Internet.

The fraction $n_2^i \in [0,1]$ of CPs that is connected to ISP$_i$ is given by

$$n_2^i = \phi_2^i(n_1^i, p_2^i) = 1 - \boldsymbol{F}(\gamma^i) \text{ with } \gamma^i = \frac{p_2^i}{n_1^i}\,. \tag{4}$$

It is a function of the number of EUs that can be reached through ISP$_i$ and the price charged. The number of CPs the ISP can persuade to directly interconnect depends on the parameter $\gamma^i = p_2^i / n_1^i$. This parameter is calculated as the fraction of the fixed price for connectivity over number of reachable EUs. Thus it can be interpreted as the perceived price per EU. The distribution $\boldsymbol{F}(\gamma^i)$ then yields the fraction of CPs that are not willing to pay that price and $\boldsymbol{1 - F}$ yields the fraction of CPs that are willing to pay that price because their expected revenue per EU covers the expense.

The CPs do not deal exclusively with a single ISP but may be connected to zero, one or two ISPs, depending on their participation constraint being fulfilled. Therefore in general $n_2^i + n_2^j \neq 1$.

While equation (4) only depends on factors under control of ISP_i, equation (2) also depends on factors controlled by the other ISP. This reflects the fact that there is competition for EUs, but none for CPs.

Costs for interconnection are defined as

$$C^i(n_1^i, n_2^i) = cn_2^i \quad . \tag{5}$$

This implicitly includes the assumption that the cost of the access network is not part of the considerations for interconnecting with CPs. This assumption is justified by the fact that access networks largely represent sunk costs.

Now, in order to solve the ISPs' optimization problem $\max \Pi^i$, assume that the platforms have reached an equilibrium and offer utility $\hat{u}_1^i$ to their $\hat{n}_1^i$ EUs, respectively. That is, we keep these values fixed while varying the others. This corresponds to today's situation in many markets for DSL or cable. There is some churn, but by and large networks operate in saturated markets with stable customer numbers. Since (4) defines n_2^i as a function of p_2^i, we can eliminate p_2^i and only have n_2^i left as a dependent variable. Thus, given an equilibrium $(\hat{u}_1^i, \hat{n}_1^i)$, we can solve for the optimal number of CPs n_2^i.

Rewriting equation (3) as $p_1^i = U^i(n_2^i) - u_1^i$ we can insert this expression into (1) to get

$$\begin{aligned} \Pi^i &= \hat{n}_1^i (U^i(n_2^i) - u_1^i) + p_2^i \overbrace{(1 - F(p_2^i / \hat{n}_1^i))}^{n_2^i} - C(\hat{n}_1^i, n_2^i) \\ &= \hat{n}_1^i (\alpha_1 n_2^i - \hat{u}_1^i) + \left(p_2^i - c\right) n_2^i \quad . \end{aligned} \tag{6}$$

This expression shows that given an arbitrary equilibrium we can explicitly write the profit of the platform as a function of the price charged to its group two customers (i.e. CPs). The platform can thus easily calculate the optimal price and the resulting number of CPs, given its current competitive situation on the EU side of the market.

To give a concrete example, we define the distribution $\boldsymbol{F}$ and explicitly calculate the profit maximizing price p_2^i. Let the distribution function $\boldsymbol{F}$ be given by the probability density function $f(\gamma) = 1/\tau, \forall \gamma \in [0; \tau]$ of the uniform distribution. γ represents the expected revenue from ad-clicks per EU and τ represents the maximum price a CP is willing to pay for access to such an EU. The corresponding cumulated distribution function is

$$\boldsymbol{F} = \gamma / \tau = p_2^i / n_1^i \tau \quad . \tag{7}$$

Any other distribution function would work as well. However, the normal distribution for example is not easily manipulated and thus would only allow a numerical solution to the problem at hand.

Now we insert (4) and (7) into (6)

$$\begin{aligned}\Pi^i &= \hat{n}_1^i(\alpha_1 n_2^i - \hat{u}_1^i) + (p_2^i - c)n_2^i \\ &= \hat{n}_1^i(\alpha_1 n_2^i - \hat{u}_1^i) + \left[(1-n_2^i)\tau\hat{n}_1^i - c\right]n_2^i\end{aligned} \tag{8}$$

and find the maximizer of the resulting expression:

$$\begin{aligned}\frac{\partial \Pi^i}{\partial n_2^i} &= \hat{n}_1^i\alpha_1 + (1-2n_2^i)\tau\hat{n}_1^i - c \overset{!}{=} 0 \\ n_2^i &= (1-\frac{c-\hat{n}_1^i\alpha_1}{\tau\hat{n}_1^i})\frac{1}{2}\ .\end{aligned} \tag{9}$$

This is the optimal number of CPs the ISP should allow on its platform (since the 2nd order condition for a maximum holds). Together with (4) and (7) this yields the optimal price to CPs

$$p_2^i = \frac{1}{2}(c + \hat{n}_1^i\tau - \hat{n}_1^i\alpha_1)\ . \tag{10}$$

Therefore, CPs pay a price that is calculated on the basis of the cost they cause, increased by a factor relating to their per-EU-valuation and decreased by the externality they exert on the EUs. The factor $1/2$ should not be over-interpreted since it is an artifact of the definition of the distribution function in (7).

We thus have calculated the optimal number of CPs and the optimal price that an ISP should charge atomistic and ex-ante identical CPs for quality interconnection.

5 Conclusions and Further Research

To sum up, we have firstly explained two phenomena of the Internet that fundamentally change the way CPs and EUs are interconnected. CDNs and MH foster more direct links between these two user groups with only one mediating ISP instead of many. Employing the theory of two-sided markets we then went on to show how direct interconnection puts the ISP into a position to charge CPs directly. In the main section we showed how the optimal price p_2^i can be calculated for any given equilibrium on the EU side of the market.

While today it is uncommon to explicitly charge content providers for delivering traffic to their customers, there are clearly developments in the marketplace that can be understood in the above context. Google's effort to provide free W-Lan to customers in the US is only one example. Google wants to control the platform over which its content is delivered so that the profits it makes on the advertisement side cannot be extracted by ISPs.

To interpret the results obtained, lets first compare the predicted price to today's bill and keep regime. In today's peering agreements between ISPs, the fee for carrying traffic is very often zero. As [17] point out, this corresponds to a subsidy to CPs, since EUs carry most of the transmission cost. In our two-sided market framework on the other side, the CPs have to bear the cost they cause. They may be furthermore charged by the ISP, depending on their willingness to pay. This charging is balanced

by a "bonus" for the externality they exert on the ISP's EUs. Since the difference between being subsidized and paying bottleneck prices can be quite large, there will probably be a transitional period before ISPs can leverage their whole power in charging CPs. However, the presence of charges to content providers in itself does not represent a market failure. As long as ISPs are competing in the market for EUs, the profits they make on the CPs are used to compete in the EU market [3]. A waterbed effect might occur, but would merely be a sign of imperfect competition in the EU market [11].

Secondly, the last term in (10) illustrates a very interesting result. Imagine that the ISP could perfectly discriminate between two different groups of CPs. The group that exerts a higher externality on the customers through its presence would pay a lower price than the group with the lower externality effect. Thus, CPs that are very important to EUs will pay a low price to the bottleneck ISP, while those CPs, the presence of which is less valued by EUs, will pay a high price for access to EUs. Thus, a power balance could develop, in which CPs are charged by the network if they have low market power; or charge the network, if their content is highly desired by EUs.

Lastly, look again at the externality exerted by CPs. Here might lie an interesting option for future ISP business models. The ISP could try to capture some of the externality. This could happen for example through transaction dependent charges. Aside from contractual problems this would fulfill many ISPs' long standing vision to capture some of the profits of the content business. This development can already be witnessed in the mobile sector where Vodafone provides high quality ad-financed content to its customers.

An important aspect of this work that requires further research is the effect of the two-sided markets phenomena on the quality of standard peerings. As it stands today, peerings do not generate revenue for ISPs but only costs. With revenues from direct interconnection there is obviously a strong incentive for ISPs to move as many CPs as possible to a paying interconnection model. The ultimate consequence of this would be that, in order to foster a self selection process, standard peering quality would be considerably degraded to make sure that all customers with a willingness to pay are in the paying group. While such price discrimination is welfare enhancing, it is crucially important the market for EUs is competitive since otherwise, ISPs are in a position to appropriate rents.

This paper demonstrates the use of two-sided market theory to analyze the decision problems faced by Internet service providers in more complex setups then the standard peering scenario examined in earlier works. A first analysis demonstrates that new business models such as content delivery networks and multi-homing can fundamentally change the rules for interconnection pricing. This work thus extends the work on Internet interconnection [17] and the work on voice interconnection such as [16] or [2] (as well as the references cited therein).

As this is only a first step towards a thorough understanding of the new rules of interconnection pricing brought about by new interconnection regimes, there remains considerable work to be done:

Firstly, the presented analysis cuts short some more in depth equilibrium analysis by assuming a market equilibrium as given.

Furthermore, a more thorough analysis of the effects of the ISPs actions on the secondary markets for advertisements would be interesting. How do two vertically dependent two-sided markets interact?

In a similar line of thought, the aspect that CDNs are intermediaries between ISPs and CPs has been used as a starting point of the analysis but is then abstracted from in the further analysis. This can be justified by assuming that CDNs only pass on costs but their role certainly deserves more attention, especially since CDNs are potent players in the Internet market. A further topic to be analyzed is the role of peer to peer traffic.

The paper has shown an aspect of the quality of service debate that has been under-researched. The market for Internet interconnection has a considerable influence on the deliverable quality of Internet services. Understanding these markets (the contribution of this work) and "engineering" them to function better (future research) propose challenging research topics that might shape the next generation of networks.

References

1. Akella, A., Seshan, S., Shaikh, A.: An empirical evaluation of wide-area Internet bottlenecks. In: Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement, Miami Beach, FL, USA (2003)
2. Armstrong, M.: Network Interconnection in Telecommunications. The Economic Journal 108, 545–564 (1998)
3. Armstrong, M.: Competition in two-sided markets. RAND Journal of Economics 37, 668–691 (2006)
4. Berger, U.: Bill-and-keep vs. cost-based access pricing revisited. Economics Letters 86, 107–112 (2005)
5. Bu, T., Gao, L., Towsley, D.: On characterizing BGP routing table growth. Computer Networks 45, 45–54 (2004)
6. Cambini, C., Valletti, T.M.: Network competition with price discrimination: 'bill-and-keep' is not so bad after all. Economics Letters 81, 205–213 (2003)
7. Cremer, J., Patrick, R., Tirole, J.: Connectivity in the Commercial Internet. The Journal of Industrial Economics 48, 433–472 (2000)
8. DeGraba, P.: Reconciling the off-net cost pricing principle with efficient network utilization. Information Economics and Policy 16, 475–494 (2004)
9. Economides, N., Lopomo, G., Woroch, G.: Strategic Commitments and the Principle of Reciprocity in Interconnection Pricing. Stern School Business, NYU (2008)
10. Gans, J.S., King, S.P.: Using bill and keep interconnect arrangements to soften network competition. Economics Letters 71, 413–420 (2001)
11. Genakos, C., Valletti, T.: Testing the 'Waterbed' Effect in Mobile Telephony. CEIS Working Paper No. 110 (2008)
12. Hau, T., Wulf, J., Zarnekow, R., Brenner, W.: Economic Effects of Multi Homing and Content Delivery Networks on the Internet. In: Proceedings of the 19th ITS European Regional Conference, Rome (2008)
13. Hermalin, B.E., Katz, M.L.: Sender or Receiver: Who Should Pay to Exchange an Electronic Message? The RAND Journal of Economics 35, 423–448 (2004)
14. Hotelling, H.: Stability in Competition. Economic Journal 39, 41–57 (1929)
15. Kushman, N., Srikanth, K., Katabi, D.: Can you hear me now?! it must be BGP. SIGCOMM Comput. Commun. Rev. 37, 75–84 (2007)
16. Laffont, J.-J., Rey, P., Tirole, J.: Network Competition: I. Overview and Nondiscriminatory Pricing. The RAND Journal of Economics 29, 1–37 (1998)

17. Laffont, J.-J., Marcus, S., Rey, P., Tirole, J.: Internet Interconnection and the Off-Net-Cost Pricing Principle. The RAND Journal of Economics 34, 370–390 (2003)
18. Laffont, J.-J., Rey, P., Tirole, J.: Competition between telecommunications operators. European Economic Review 41, 701–711 (1997)
19. Musacchio, J., Walrand, J., Schwartz, G.: Network Neutrality and Provider Investment Incentives. In: Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, 2007. ACSSC (2007)
20. Rochet, J.C., Tirole, J.: Platform Competition in Two-Sided Markets. Journal of the European Economic Association 1, 990–1029 (2003)
21. Rochet, J.-C., Tirole, J.: Two-sided markets: a progress report. The RAND Journal of Economics 37, 645–667 (2006)
22. Roughgarden, T.: Selfish Routing and the Price of Anarchy. MIT Press, Cambridge (2005)
23. Shakkottai, S., Srikant, R.: Economics of network pricing with multiple ISPs. IEEE/ACM Trans. Netw. 14, 1233–1245 (2006)
24. Teixeira, R., Shaikh, A., Griffin, T., Rexford, J.: Dynamics of hot-potato routing in IP networks. In: Proceedings of the joint international conference on Measurement and modeling of computer systems (2004)
25. Tirole, J.: The Theory of Industrial Organization. MIT Press, Cambridge (1988)
26. Vakali, A., Pallis, G.: Content delivery networks: status and trends. IEEE Internet Computing 7, 68–74 (2003)

Online Charging for IMS-Based Inter-domain Composite Services

Minh van Le[1], Frens Jan Rumph[2], George B. Huitema[2], Bert Jan van Beijnum[3], and L.J.M. (Bart) Nieuwenhuis[1]

[1] Faculty of Information Systems and Change Management, University of Twente, the Netherlands
[2] TNO, the Netherlands
[3] Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, the Netherlands
van-minh.le@atosorigin.com, frens_jan.rumph@tno.nl, george.huitema@tno.nl, beijnum@ewi.utwente.nl, l.j.m.nieuwenhuis@utwente.nl

Abstract. In order to manage financial risks online charging of composite services is becoming increasingly important for service providers to support service delivery in inter-domain environments. The 3rd Generation Partnership Project (3GPP) has developed a framework for off-line and online charging of IMS-based services. At service level, delivery of composite services often involves many service providers, where each service provider is responsible for the delivery of one or more service components (e.g. access service, IMS communication service, content service, etc.). Current Online Charging System (OCS) specified by 3GPPP does not support an online charging function for composite services. This paper discusses a number of implications of online charging of composite services, in particular inter-domain composite services. It addresses important shortcomings of the current 3GPP online charging architecture and suggests a way to overcome these shortcomings. The contribution of this paper is twofold. Firstly, it proposes an information model to support charging of inter-domain composite services. The proposed information model is based on the NGOSS and SID concepts of the TeleManagement Forum. Secondly, it proposes additional functionalities required for existing OCS in IMS to handle online charging of inter-domain composite services.

Keywords: Online charging, composite services, IP Multimedia Subsystem (IMS), inter-domain.

1 Introduction

Most telecom service providers are currently implementing multi-service IP infrastructures to cope with huge customer demand for advanced multimedia services over fixed and wireless broadband networks. Rapid penetration of smart phones with comfortable large screen sizes also contributes to a positive customer experience. The combination of these two trends results in a significant growth of usage of mobile

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 72–84, 2009.

services such as Voice, SMS, MMS, email, location-based services, instant messaging, community gaming, IPTV, etc. In particular, these multimedia services are composed of services that are delivered by *different* providers in different administrative domains. At service level, delivery of composite multimedia services may involve different service providers such as e.g. telecom service providers, content providers or game providers. We therefore speak of *inter-domain composite* services.

The 3rd Generation Partnership Project (3GPP) has developed the IP Multimedia Sub-system (IMS) [1] to support multimedia services. Besides the IMS architecture, the 3GPP also provides a framework for off-line and online charging. Off-line charging implies that charging for service usage occurs after a service event or service session has occurred. Online charging implies that charging occurs right after a service event has occurred or during a service session usage.

Now, customers of composite services want (near) real-time charging and billing information to manage their expenses while using these services. Service providers also have the same needs in order to manage their financial risks. However, today's charging systems are not capable of dealing with requirements related to composite services. In particular, current online charging systems specified by 3GPPP [2] do not support online charging functions for inter-domain composite services.

In this paper a number of implications are discussed of online charging of inter-domain composite services. Note that here inter-domain is the generic case. The case where one deals with composite services in the same administrative domain can be derived from it. Hence, in this paper significant shortcomings of the current 3GPP online charging architecture are described while also a way to overcome these shortcomings is given. The major contribution of this paper is twofold. *Firstly*, an information model is proposed to support charging of inter-domain composite services. The proposed information model is based on the TeleManagement Forum's (TMF) programme of New Generation Operations Systems and Software (NGOSS) [3, 4]. S*econdly*, in this paper additional functionalities are suggested for the current Online Charging System (OCS) of IMS in order to handle online charging of composite services.

The remainder of this paper is organized as follow: Section 2 provides a brief overview of related work. Section 3 discusses a motivating example in which composite services are delivered across different domains. Section 4 presents the problem domain of online charging of composite services. Section 5 proposes a service composition information model. Section 6 discusses important functionalities for coping with the charging of composite services. Finally, Section 7 presents the conclusions and contribution of this paper together with some future directions.

2 Related Work

In this section related work with respect to the service composition of inter-domain services and its charging is described.

The TMF launched an initiative on their Service Delivery Framework (SDF) in 2007, which aims at specifying standards for managing service delivery across multiple administrative domains. The ultimate objective of these future standards is to enhance partnerships among many kinds of service providers and to achieve efficient

interoperation between different domains. In such an environment the service management processes of all service providers involved need to be orchestrated to ensure smooth and consistent service delivery to end-users. The work in progress of TMF [5] shows clear direction of the SDF to standardize "overlay" service delivery platforms that will rely on existing and future networks such as the IMS core infrastructure.

The 3GPP has proposed the concept of a *Service Broker* to enable service composition and orchestration [6, 7]. In the context of IMS a service broker has two fundamental tasks: 1. service execution and orchestration, which ensures that sub-services (being part of a composite service) are deployed and co-exist in harmony; and 2. service offering coordination, which ensures that service composition is conducted in real-time in order to fulfill requests of customers. One can think of three alternative service broker models, i.e. a centralized, distributed or hybrid model. A centralized service broker model implies that different application servers interact with a single S-CSCF (Serving-Call Session Control Function) via a single service broker. A distributed service broker model implies only a one-to-one relationship between an application server and a service broker. In turn, service brokers communicate with a single S-CSCF. Finally, a hybrid service broker model allows one-to-one as well as many-to-one relationships between application servers and service brokers. In turn, service brokers communicate with a single S-CSCF.

With respect to charging, the 3GPP has proposed two reference charging models for different service scenarios [2], namely: an off-line and an online charging model. The latter charging model covers near real-time charging issues such as charging authorization, credit control during service sessions. However, it lacks a service composition model to deal with charging of composite services. For instance, a video conference can fall back from video plus voice to only voice due to some network problem. This introduces a change in charging of the voice component for the rest of the conversation. In this kind of situation, it is necessary for a charging system to keep track of the service composition information to adapt charging accordingly.

In [12], the authors introduced a concept called Time Interval Calculation Algorithm (TICA) for online charging to deal with performance issues. That is, to avoid large overhead caused by credit checks. TICA supports flexible tariff functions to cope with sophisticated business relationships between the involved business partners. TICA and the proposed solution in this paper are complementary. Hence, a combination of TICA and our proposed solution helps to tackle performance issues of online charging of inter-domain composite services.

3 An Inter-domain Service Scenario

Below we will describe a motivating example of a service scenario with inter-domain service composition and related charging. In our scenario a young business professional Jane leaves her office after a working day and takes the train back home. During the journey she wants to catch the evening news from a national television channel. Jane uses her smart phone to connect to the 4G mobile network of her service provider and requests on-demand TV service. The service provider responds by offering her two options: 1. premium TV-on-demand (TVoD) service without mobile advertising for a premium price and 2. TVoD service with mobile advertising for a sponsored price. As

she likes discounts Jane first chooses mobile advertising integrated in the TVoD service. At some point in time she is bored with this service and switches the advertising banner off.

In our example scenario the Service Provider is assumed to be capable of providing core IMS services (e.g. access, mobile internet, VoIP, Push-to-talk, Messaging, etc.) using an IMS-based infrastructure. Here the TVoD services and the mobile advertising service are provided by third parties in other administrative domains. The business rationale behind this scenario is as follows. On the service delivery side: the Service Provider combines the TVoD services and mobile advertising to create attractive service offerings to customers. The Service Provider purchases the TVoD services from the TV-on-demand Provider. On the paying side: the Mobile Advertising Provider pays the Service Provider for inserting mobile advertising into each TVoD service session. Furthermore, Jane pays for the composite service to the Service Provider (all-in). However, Jane receives discounts for the TVoD services from the Service Provider by accepting mobile advertising. If Jane switches off the mobile advertising during the TVoD service session, the premium price is applied from that point forwards.

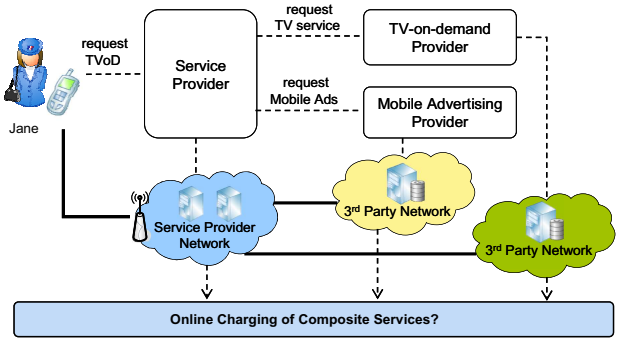


Fig. 1. Inter-domain service scenario depicting functional relationships (dash lines) and the physical paths of the delivered data (solid lines). The arrows and the blue box denote the problem domain of online charging over multiple administrative domains.

Figure 1. depicts the above described inter-domain service scenario. In this service scenario the Service Provider is responsible for the *user-facing* charging and billing of Jane, as well as the *3rd party-facing* charging and billing of the TVoD Provider and the Mobile Advertising Provider. Here, charging is expected to be online (near-real-time) in order to manage the financial risks for both the user Jane and the Service Provider.

4 Problem Domain of Online Charging of Composite Services

This section deals with the scope of our research into the problem domain of online charging of composite services. Moreover a summary of the research questions we deal with is given.

4.1 Research Scope

This paper focuses on online charging of composite services. Depending on the implementation strategy online charging may involve different distributed Online Charging Systems (OCS). However, we abstract from a concrete distribution and focus on one single OCS. We assume that composite services will be delivered across several administrative domains and across several delivery platforms. Hence, a combination of web services and IMS services based on SOA (Service Oriented Architecture) [7] is considered.

In order to master the complexity of such a service delivery, the concept of service broker as advocated in [6, 7] is used. There are different possible configurations of service brokers in an actual deployment. In this paper, two types of service broker are considered: 1. An IMS Service Broker within the IMS domain and 2. An Inter-domain Service Broker within the Web-Services domain. Here, the Inter-domain Service Broker is leading and is responsible for the end-to-end service composition for the user.

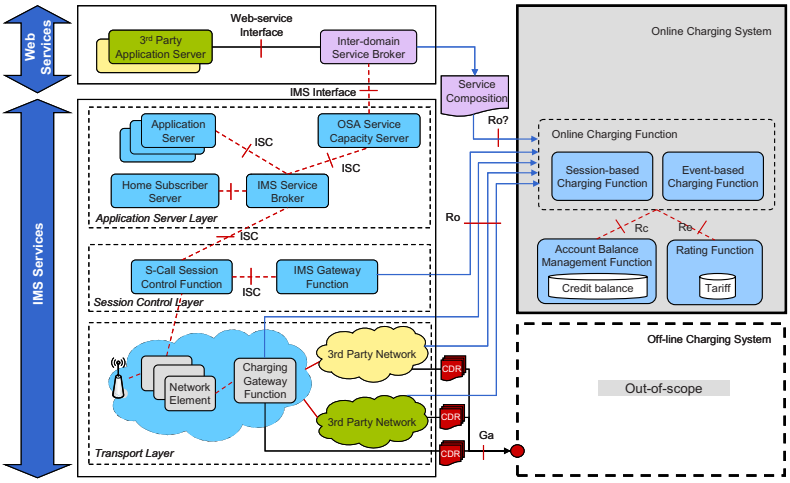


Fig. 2. Provisioning and Charging of Composite Services in an Inter-domain Environment

Figure 2 above illustrates a possible model for provisioning and charging of composite services in a multi-domain environment. A composite service is delivered across different administrative domains: the service provider domain, which is supported by IMS core infrastructure, and the third party domains, which are not necessarily supported by an IMS but some different network infrastructures. The service composition and orchestration occurs at the web-service level through the Inter-domain Service Broker. This implies that the Inter-domain Service Broker has the knowledge of the service composition and needs to communicate this information with the OSC for online charging purposes. Note that here off-line charging is out-of-scope.

4.2 Research Questions

The 3GPP framework describes online charging for both events and sessions. However, there exists no standard model yet for service composition within 3GPP specifications. Although an IMS service may consist of different service components (e.g. VoIP over IP-access), current online charging systems do not correlate charges between these service components. As a result, when a service component is added or removed from an incurred composite service session, adjustment of online charging due to possible tariff changes (e.g. zero rate bearer usage when a VoIP session is active) cannot be handled. Since the composition of a composite service may change at run time, online charging needs to adapt to this dynamic behavior as well. This implies that the OCS needs to have knowledge about how the ongoing service composition is built up and the corresponding tariff of an individual service component. Furthermore, charging policies need to be enforced appropriately according to some pre-defined service level agreement between the end-user and the service provider.

Hence we come up to the following research questions regarding online charging of composite services:

1. What is the service composition model used by the OCS? This involves information exchange between an inter-domain service broker function and the OCS.
2. Which technical details should be included in the service composition model? A charge request should contain enough details to enable the OCS to conduct credit reservations and correlate different charges.
3. How do the charging processes of individual service components influence each other during an ongoing service session? There must be a way to keep the state of service composition within the OCS. In addition, the cross dependency of tariff between different service components depends strongly on imposed charging policies. Hence, the OCS also needs to take charging policies in to account.

The above questions will be addressed in the sections 5 and 6 below where we propose solutions.

5 Service Composition Information

This section presents a service composition information model to deal with online charging of inter-domain composite services. An example is also given to concretize the proposed solution.

5.1 Service Composition in the Context of TMF's SDF

According to [5], there are three steps to arrive at the eventual service delivery, namely: product design, service creation and service execution. During the product design phase, a service designer from the service provider domain can look up available service components in a catalog and chooses the necessary service components to form a composite service. During the creation phase, the designed composite service is tested throughout. When a composite is accepted, a meta information model of the composite service is created and stored in a composite service catalog. In the last

phase, whenever a user requests a composite service, an instance to the corresponding meta information model is generated. Figure 3 depicts when a meta service composition model comes into existence.

In order to compose a composite service, the service provider sometimes needs to acquire external services from third party providers. A service composition model therefore must express the relationship between the constituent service components. The TMF has been working on the Shared Information/Data (SID) model [4], which provides guidelines for the modeling of information/data for the purpose of product design, service construct and service provisioning. Currently, the SID model is widely accepted as standard in the industry. Among many aspects, the SID model addresses the basic entities: product, service, end-user-facing service and provider-facing service and their relationships. The next section will discuss a service composition information model based on SID, which can be used for online charging of composite services.

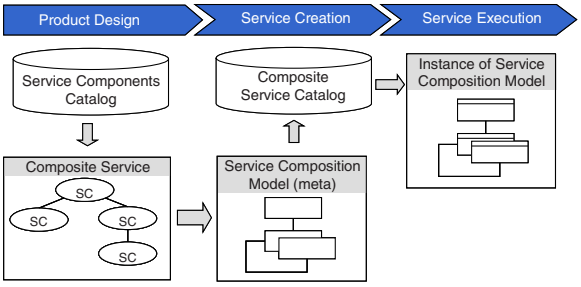


Fig. 3. Service Composition Information in the context of TMF's SDF

5.2 Service Composition Model

A `Product` is a particular "item" that an end-user can buy. For example, the end-user can browse through a list of products (e.g. video's) and pick out a preferred one. A `Service` is part of a `Product`. A `Service` cannot exist by itself, but is bound to a `Product`. An end-user can only buy a product, not a service. For example, the end-user buys 30 minute of ToVD of channel A (as a product) and experience high quality news (as a service). A `Service` represents the things, which are experienced by an end-user. A `Provider-facing Service` represents the resources which are needed to support the `End-user-facing Service`, which is to the `Service Provider` but invisible to an end-user.

The heart of the service composition is the `Service`, which is distinguished into an `End-user-facing Service` and a `Provider-facing Service`. The `End-user-facing Service` is linked to `Product`, which an end-user can choose. The `Provider-facing Service` consists of one or more `Atomic Services`, which can be `Provider Services` (i.e. internal resources), or `Partner-facing services` (i.e. external resources) or both. The relationship between the `Provider-facing Service` and the `Atomic Service` is a transformation

duality relationship [8, 10]. Hence, in order to arrive at a `Provider-facing Service`, the `Service Provider` needs to compose a service session from different service components. Figure 4 below depicts the service composition model.

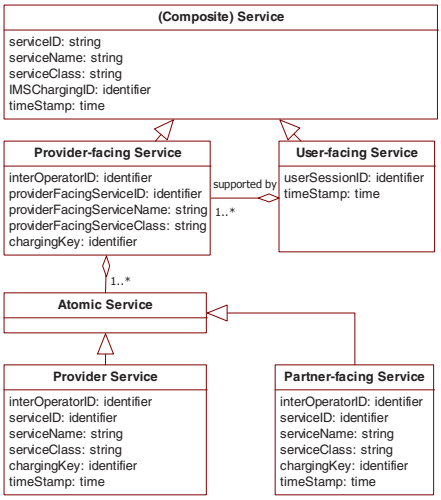


Fig. 4. Service Composition Model ChargingKey

The separation between the `End-user-facing Service` and `Provider-facing Service` makes it possible to construct service session compositions which contain detailed information about how a service session is built and what service components are used in a service session. The *how* and the *what* are expressed by `Provider-facing Service`. What an end-user "experiences" is the `End-user-facing Service`, which is transparent and abstracted from detailed business information intended only to the `Service Provider`.

The service composition information model contains necessary detail information to ensure the correlation of service components and their corresponding charge. The following pieces of information are crucial:

- `serviceID` - an unique identifier of a provided composite service or a service component.
- `interOperatorID` – a unique identifier of a service provider or a 3rd party provider.
- `chargingKey` - an identifier used by the OCS to determine the tariff of a composite service or a service component.

The combination of an `interOperatorID`, the corresponding `serviceID` and the `chargingKey` allows for an appropriate credit authorization request at the OCS.

Figure 5 below shows an example of an instance of a service composition information created for a TVoD request as described in the service scenario of Section 2.

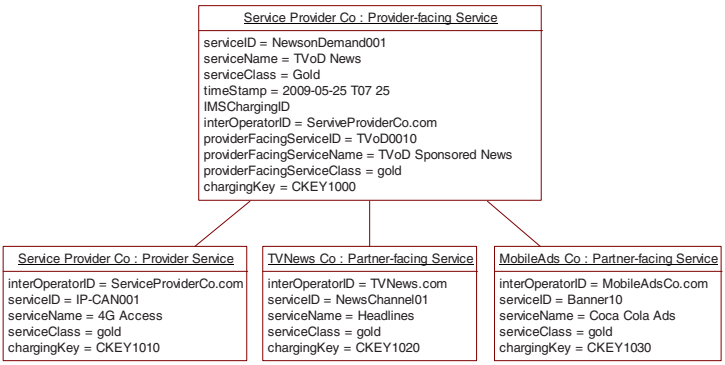


Fig. 5. Example of an instance of service composition sent to the OCS

6 Online Charging Functions for Supporting Composite Services

This section discusses the functionalities required for existing IMS's OCS to handle online charging of inter-domain composite services. In particular, it focuses on credit control mechanisms to deal with financial risks.

6.1 Credit Control for Composite Service Sessions

The main objective of online charging is to provide service providers with a mechanism to control user credit allowance. A credit allowance can be a pre-defined upper limit of a postpaid account (e.g. parents setting up spending limits for their children) or a current amount of money of a prepaid account. For this purpose, it is necessary to check the user credit balance prior to service provisioning. Moreover, if the user credit balance is sufficient, credit reservation must be made for the requested composite service session. In some cases, it will be necessary to create separate credit reservations for individual service components. To this extent, the Diameter Credit Control Application [13] is suitable to support the credit authorization requests for (inter-domain) composite service sessions. The charging of composite services can be divided into three phases:

1) *Charging request* – The Inter-domain Service Broker (ISB) sends a request to the OCS asking for charge authorization of a composite service session. The service composition information must be included in this request to enable the OCS determining the required credit reservation. At this point, the OCS creates a "parent claim" for the requested composite service. We note that the service session information sent from the ISB to the OCS must be combined with necessary information from the customer domain such as subscriber identifier and user identifier to enable the OSC to look up the appropriate account (not shown in this paper).
2) *Charging initiation* – If the user credit balance is sufficient, the ISB continues to initiate the required service components. For each service component a credit authorization request can be sent to the OCS, e.g by an individual

Application Server (AS). Each credit authorization request is followed by a corresponding credit authorization response which includes (among other things) an assigned usage quota (e.g. data or time unit).

3) *Service Charging* – During the service session usage, credit re-authorization might be necessary when the usage quota is approaching zero. An individual AS can undertake this action with the OCS independently. When the provisioning of a service component is terminated (this can be user-initiated, AS-initiated or OCS-initiated caused a credit constraint), the involved AS sends a final charge report to the OCS and a service termination message to the ISB.

Figure 6 shows the interaction between the ISB and OCS in the charging request phase. To avoid unnecessary load caused by credit requests for individual service components, we propose to conduct the "parent claim" for each request using a

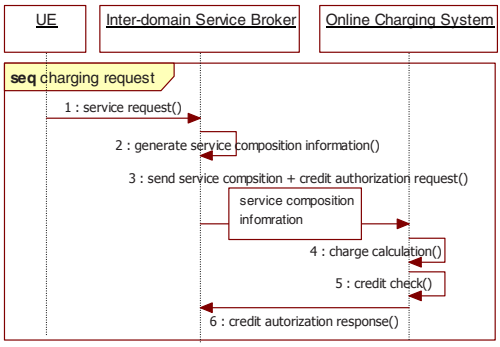


Fig. 6. Exchange of Service Communication Information

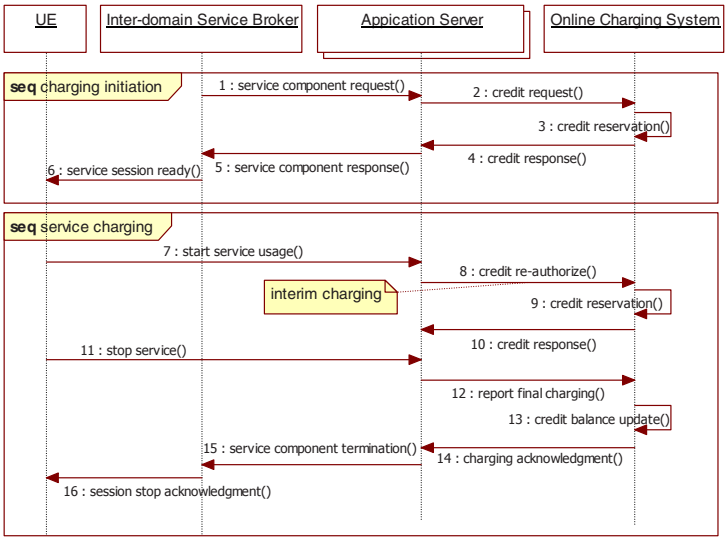


Fig. 7. Online Charging Initiation and Online Service Charging

limited data set of information contained in the service composition. Moreover, the combination of {`interOperatorID's` + `serviceID's` + `chargingKey's`} and their mutual relationships provide enough information for the OCS to determine the "parent claim". Hence, the ISB does not need to gather detailed technical information from the involved AS before being sure that the service provisioning is allowed.

Finally, Figure 7 below shows the interactions during the charging initiation and service charging phase between the User Equipment (UE), Inter-domain Service Broker (ISB), Application Server (AS) and Online Charging System (OCS). It is worthy to note that in case a service component is removed from a service session, the ISB needs to inform the UE about the new tariff of the remaining service session (not shown in Figure 7).

6.2 Dealing with Dynamic Change of Service Components

Online charging of composite services becomes complex when the composition of the ongoing composite service session changes (i.e. adding or removing service components) and when there is a tariff dependency between the involved service components. In such a situation, two major impacts on online charging are observed: 1. possible tariff changes of the remaining service components; 2. adjustment of user credit balance. To deal with these issues, we can think of three charging strategies.

The *first* strategy is to apply a tariff-dependent charging scheme. Here, the OCS must conduct credit reauthorization for the involved service components whenever the service composition changes. The advantage of this strategy is that it allows the OCS to adjust tariffs in near real-time, which can be desirable from a business viewpoint. The trade-off is that this strategy might induce extra load on the OCS. The *second* strategy is to use a tariff-independent charging scheme. Here, the OCS can apply a fixed tariff for each chargeable service component and a fixed tariff for each awardable service component (i.e. component from which an user receives compensation such as advertising). The advantage of this tariff-independent scheme is to avoid tariff recalculation, thus avoids extra load on the OCS. The *third* strategy is to apply a hybrid charging scheme where a combination of the two strategies is used. For instance, the tariff of a connectivity service component can be fixed, whereas the tariff of a TVoD service component depends on the rewarding of an advertising service component. In case the advertising component is removed from the service session, the OCS only needs to adjust one tariff for the TVoD component.

Moreover, in order to support the above charging strategies, existing functions within the OCS such as SBCF (Session-Based Charging Function), ABMF (Account Balance Management Function) and RF (Rating Function) need to take into account the dynamic character of the composite services and the tariff dependency amongst the service components. For instance, the ABMF needs to manage both "credit claims" of chargeable service component as well as the "rewarding prospect" of awardable service component.

6.3 Impact on Existing 3GPP Interfaces

Regarding the interfaces of the OCS, the main impact of our proposed solution in 6.1 and 6.2 is on the Ro reference point between the ISB and the OCS [11]. The current

specification of the Ro reference is only capable of supporting "flat" structure of service components. Thus, no distinction can be made between a composite service as a whole and its (sub) service components. However, the hierarchical structure of service components and their corresponding charging keys are critical for the determination of the charging dependencies between service components. As a result, some adaptation should be made at the Ro reference point to enable the exchange of service composition information sent from the ISB to the OCS. Altogether, having a hierarchical structure of service components, their corresponding charging keys and an adapted Ro reference point will allow us to apply flexible charging policies. Hence, this extension of the capability of the OCS will support a broad variety of business models between service partners in different domains.

7 Conclusion and Future Work

This paper addresses the problems of online charging for composite services which are delivered across IMS-based and web-based infrastructures. One of the main challenges is the lack of a service composition information model that expresses the hierarchical structure of the service components and the relationships between their corresponding charging keys. To overcome this problem, a service composition model based on the SID framework is proposed. Furthermore, this paper discusses the implication of credit control when dealing with composite services and tariff dependency between the involved service components. To this extent, three charging strategies have been discussed in order to tackle the dynamic changes of service composition during run time. The impact of the proposed solution on the existing 3GPP charging reference architecture [2] is limited. Minor adaptation at the Ro reference point is required to include the proposed service composition information in the charging request phase.

Future work will study the correlation function that should be introduced to the OCS [11]. Further impact of online charging of composite services on existing functions such as OCF (Online Charging Function), ABMF (Account Balance Management Function) and RF (Rating Function) will be examined.

References

1. Cuevas, A., Moreno, J.I., Vidales, P., Einsiedler, H.: The IMS service platform: a solution for next-generation network operators to be more than bit pipes. IEEE Communication Magazine 44(8) (August 2006)
2. 3GPP, Charging Architecture and Principles, TS 32.240, version 8.4.0 (September 2008)
3. Strassner, J., Fleck, J., Huang, J., Faurer, C., Richardson, T.: TMF Whitepaper on NGOSS and MDA, version 1 (November 2003)
4. TMForum, Shared Information/Data (SID) Model – Business View Concepts, Principles and Domains, release 6 (November 2005)
5. TMForum, Service Delivery Framework Overview, release 2.0 (September 2008)
6. 3GPP, Study on Architecture Impacts of Service Brokering, TR 23.810, release 8.0 (September 2008)

7. Huslak, N.S., McQuaide, A.C.: Service Brokering: Opportunities and Challenges. In: International Conference on Intelligent Network (ICIN), Bordeaux, France, October 20-23 (2008)
8. Le, M.V., Beijnum, B.J.F., Niewenhuis, L.J.M., Huitema, G.B.: An Enterprise Model for Real-time Inter-domain Billing of Services. In: Workshop on ODP Enterprise Computing (WODPEC), Muchen (September 2008)
9. de Laat, C., Gross, G., Gommans, L., Vollbrecht, J., Spence, D.: Generic AAA Architecture, IETF RFC 2903 (August 2000)
10. Geerts, G.L., McCarthy, W.E.: An ontological analysis of the economic primitives of the extended-REA enterprise information architecture. International Journal of Accounting Information Systems 3, 1–16 (2002)
11. 3GPP, Online Charging System (OCS): Applications and interfaces, TS 32.296, version 8.2.0 (June 2008)
12. Kurtansky, P., Reichl, P., Stiller, B.: The Evaluation of the Efficient Prepaid Scheme TICA for All-IP Networks and Internet Services. In: 10th IEEE/IFIP Symposium on Integrated Network Management (IM 2007). IEEE, Munich (2007)
13. 3GPP, PS Domain Charging, TS 32.251 v8.3.0 (September 2008)
14. Hakala, H., Mattila, L., Koskinen, J.P., Stura, M., Loughney, J.: Credit Control Application, IETF RFC 4006 Diameter (August 2005)

A New Bilateral Arrangement between Interconnected Providers

Ruzana Davoyan[1], Jörn Altmann[2], and Wolfgang Effelsberg[1]

[1] Department of Mathematics and Computer Science
University of Mannheim, 68131 Mannheim, Germany
[2] TEMEP, School of Industrial and Management Engineering, College of Engineering
Seoul National University, South-Korea
{ruzana.davoyan,wolfgang.effelsberg}@informatik.uni-mannheim.de,
jorn.altmann@acm.org

Abstract. Cost allocation between interconnected networks is based on measured traffic flows. This principle, however, does not provide a fair way for sharing costs. In this paper, a new bilateral model, called Differentiated Traffic-based Interconnection Agreement (DTIA) for intercarrier compensation is presented. In particular, the approach aims to determine the original initiator of a transmission by means of traffic differentiation into two types and to compensate the interconnection costs. Unlike the existing financial settlements, under which the payments are made based on the traffic flows, the proposed method suggests costs compensation according to the differentiated traffic flows. Further, in order to support the described payment scheme, a simple and scalable traffic management mechanism was designed. The results obtained from the comparative analysis showed that determination of a transmission initiator induces cost sharing between all parties and therefore, reduces the interconnection payments between providers.

Keywords: Interconnection arrangement, intercarrier compensation, Internet economics.

1 Introduction

The Internet is a system of interconnected networks, which are connected either through a direct link or through an intermediate point, called Internet exchange point (IXP) to exchange traffic. Historically, the Internet provides two types of interconnections: peering and transit [1]. Peering is the arrangement of traffic exchange on a free-settlement basis, called bill-and-keep (BAK), so that the Internet service providers (ISPs) do not pay each other and derive revenues from their own customers [2]. It is fair and efficient under symmetry of traffic flows, termination charges, and costs. Under the transit model, a customer provider pays a transit provider to deliver the traffic between the customers. The outcome of the negotiation process of being a transit or peered customer reflects on the assessment of the actual cost of traffic exchange [3-4]. Peering offers several advantages in terms of interconnection costs and quality of data transmission, but gives access to a part of the entire Internet. According to the estimates in [5], 80% of the Internet traffic is routed via private peering. In

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 85–96, 2009.

some cases, however, in order to recover the infrastructure costs, instead of peering with the smaller ISPs, the larger ISPs offer transit arrangements at a certain rate, providing access to the whole Internet. In addition to this, new types of interconnection models, such as paid peering and partial transit, emerged in the market [6].

Traditionally, before interconnecting, a provider calculates whether the interconnection benefits would outweigh the costs [7]. The simple economic principle suggests sharing the costs between all parties. The survey and discussion on interconnection with two-sided benefits are provided in [8-9]. In the case of telephony, the study [10] argued that both calling and called parties benefit from the call, and consequently, should share the interconnection costs. In the Internet, under symmetry of traffic flows, the termination costs are set to zero, since it is assumed that the termination fees are roughly the same, and a peering arrangement is used. However, because no termination cost is charged, BAK is considered inefficient in terms of the cost compensation [11]. Generally, if providers are asymmetric in terms of size, peering model is not appropriate, since providers incur different costs and benefit differently.

Therefore, if traffic is unbalanced, interconnection arrangement is governed by the financial compensation in a bilaterally (paid peering) or unilaterally (transit) negotiated basis to recover the costs of the network. In bilateral settlements, the payments are done based on the net traffic flow. Considering the Internet hierarchical structure, Internet backbone providers (IBPs) sell the wholesale services to competitive ISPs. As a result, in unilateral settlement agreements, a customer provider pays for sent and received traffic, even though traffic flows in both directions. As cited in [5], it was recommended to establish bilateral arrangements and to compensate each provider for the costs that it incurs in carrying traffic generated by the other network. However, the study [5] argued that traffic flows are not a reasonable indicator to share the costs, since it is not clear who originally initiated any transmission and therefore, who should pay for the costs. In other words, compensation between providers cannot be solely performed based on the traffic flows, which provide a poor basis for cost sharing.

Various aspects of interconnection of ISPs have been analyzed by [10], [12-16]. When analyzing economics of interconnection, existing literature considers intercarrier compensation based on the flows of traffic. Analytical studies provided in [17-18] investigated the impact of determination of an original initiator of a transmission on intercarrier compensation, demand as well as profits of the providers in the case of private peering arrangement.

This paper follows the problem of cost sharing between providers and presents a new intercarrier compensation model, called *differentiated traffic–based interconnection agreement* (*DTIA*). The key aspect of the described model is based on the determination of a transmission initiator by means of traffic differentiation into two types, referred to as *native* that is originally initiated by the provider's own customers and *stranger*, which is initiated by the customers of any other network. In comparison to the existing bilateral or unilateral settlements [3], under which the payments are based on the traffic flows, this study proposes to compensate the interconnection costs according to the differentiated traffic flows. In particular, each provider is compensated fully for the costs that it incurs in carrying native traffic and partially for the costs that it incurs in carrying stranger traffic. Unlike telephony, the proposed model does not consider a transmission initiator as a cost causer, who should cover the joint costs. Instead, all parties share the entire costs.

Further, a simple and scalable traffic management mechanism that supports the traffic differentiation approach was designed. A similar mechanism only for private peering arrangements was presented in our earlier work [19]. The major advantage of the described mechanism is that providers have not to inspect the IP header of a packet in order to determine how it should be accounted. The proposed mechanism introduces a membership label, which allows accounting the volume of a particular traffic type. Hence, a significant reduction in computational costs is achieved by using a membership label.

Finally, a comparative study of the agreements based on the traffic flows and differentiated traffic flows compensation was provided. The obtained results demonstrated that the determination of the original initiator of a transmission reduces the interconnection payment between networks.

The rest of the paper is organized as follows. Section 2 discusses the financial settlements between providers. Section 3 describes the motivation for traffic differentiation. Section 4 presents the design of the traffic management mechanism for interconnection arrangements. Section 5 provides analytical studies. Finally, Section 6 concludes this paper.

2 Financial Settlements

Generally, providers arrange financial settlements in order to determine the distribution of the interconnection costs [3], [20]. Before examining financial settlements within the Internet, we consider the telephony system. As an example, assume the scenario, where Alice makes a call to Bob. Accepting the call, Bob incurs termination costs to its provider that should be covered either directly by billing Bob or indirectly by billing the calling party's carrier. As cited in [11], "existing access charge rules and the majority of existing reciprocal compensation agreements require the calling party's carrier, […], to compensate the called party's carrier for terminating the call". Thus, an initiator of the call, i.e. Alice pays to the subscribed provider for the entire call, since Alice asked to reserve the circuit. In contrast to the telephony example, establishing a connection in the Internet does not require any reservation of the circuit. Usually packets between Alice and Bob are routed independently, sometimes even via different paths. Therefore, as cited in [12], "it is very important to distinguish between the *initiator* and the *sender*, and likewise between the *destination* and the *receiver*". The initiator is the party that initiates a call or a session, and the destination is the party that receives a call. In comparison, the sender (the originator) is the part that sends traffic, and the receiver (the terminator) is the part that receives traffic.

In telephony, the initiator is considered to be the originator and is charged based on the transaction unit, namely a "call minute" for using the terminating network. On the Internet, it might be argued that a TCP session can be considered as a call, where the initiator of a session pays for the entire traffic flow. However, considering the actual use of the network resources, financial settlement should be done at the IP level, accounting each packet of a flow. In addition to this, session-based accounting, which faces technical difficulties, is more complicated than simple packet-based accounting, under which the volume of the exchanged traffic in both directions is measured.

Therefore, generally, under unilateral business relationships, providers adopt *service-provider* settlements, where a customer ISP pays to a provider ISP for sent and received traffic, and under bilateral relationship, providers accept *negotiated-financial* settlements, where the payments are based on the net traffic flow. For detailed discussion see [3], [12], [21].

3 Motivation for Traffic Differentiation

The principle that we follow is that both parties derive benefits from the exchange of traffic and, therefore, should share the interconnection costs. Considering a system without externalities [3], [22], the costs should be shared based on the benefits obtained by each party. However, in the real world, it is impossible to measure the benefits of parties and so to share the costs. If content is not equally distributed between providers, traffic imbalance occurs, and hence, costs and revenues are not shared evenly. Indeed, the network that sends more traffic incurs lower cost than the network that receives more traffic [23]. As cited in [24], traffic flow is dominant towards a customer requested the content and generates 85% of the Internet traffic. This implies that inbound traffic is much more compared to outbound traffic of content request.

It was recommended to compensate each provider for the costs that it incurs in carrying traffic based on the traffic flows. However, according to [5], traffic flows are not a good meter for costs sharing, since "it is impossible to determine who originally initiated any given transmission on the Internet" and therefore, provide a poor basis for cost sharing. Furthermore, providers are unwilling to inspect the IP header of a packet, since "the cost of carrying an individual packet is extremely small, and the cost of accounting for each packet may well be greater than the cost of carrying the packet across the providers" [21].

In order to determine the party that originally initiated the transmission, we differentiate traffic into two types, referred to as *native*, which is originally initiated by the provider's own customers, and *stranger* that is originally initiated by the customers of any other network. Indeed, outgoing traffic of ISP_i that is the same as adjacent provider's incoming traffic may be i) either a part of a transmission initiated by a customer of ISP_i, ii) or a part of a transmission initiated by a customer of any other network. In particular, we suggest that a provider compensates the incurred costs i) fully, if the exchanged traffic is native, and ii) partially, if the originated traffic is stranger. More specifically, interconnected networks settle DTIA, whereby each partner is compensated for the costs, which it incurs in carrying traffic according to the differentiated traffic flows.

4 Traffic Management Mechanism

The traffic management mechanism for interconnection arrangements, which we propose, allows recognizing the packet type throughout the network. The key aspect of the proposed mechanism is the identification the type of traffic based on a two-bit field in the IP packet header, referred to as the *Membership Label* (*ML*).

4.1 Packet Marking by a Transmission Initiator

We assume that all nodes within the network support packet marking, where each node sets the *first bit* of the ML field of native packet to *'1'* and the packet of stranger traffic to *'0'*. The *assignment of the first bit of the label to '1' is done once*, when a node originally initiates a transmission.

A consumer can request a webpage either from a subscribed network or from any other network. This implies that a transmission endpoint, such as the destination can belong to the same network as the transmission initiator or to any other network. Therefore, a packet that appears in the network can be originated either by a local transmission endpoint or by an endpoint, which is located in any other network. Hence, we distinguish the location of a transmission endpoint originated a packet with respect to the network, where the packet appears.

The *second bit* of the label set to *'1'* indicates that the endpoint is local, and *'0'* shows that one is located in another network. The *assignment of the second bit of ML to '1' is done once*, when an endpoint of transmission originates a packet. Consequently, an original initiator of a transmission sets the ML field to *'11'*. Table I presents the description of the four available values of the label, which will be discussed latter in this section.

4.2 Outgoing Packet Re-marking

It is obvious, that native traffic with regard to one network is stranger with regard to the other. Hence, it is necessary to differentiate the exchanged traffic between networks. In order to achieve that we distinguish provider's border nodes, which are trust boundaries and maintain connection with an adjacent network, and refer to as the *Provider-to-Provider Border* (*PPB*) nodes. For calculating the *first bit* of the membership label of outgoing traffic, a PPB node performs the XOR logical operation on both bits of the ML label. Obviously, that the PPB nodes set the *second bit* to *'0'*. Even though packets within a domain can be marked by any available value of ML, interdomain traffic can take on only *'00'* or *'10'* values of the label (i.e. stranger or native traffic originated by a transmission endpoint located in any network).

In addition, in order to carry out intercarrier compensation based on the differentiated traffic (DT) flows, each PPB node keeps two counters (one for inbound and another for outbound traffic), which *calculate the volume of a particular type of traffic*, i.e. native or stranger *with regard to its network*. The volume of the other type of traffic can be easily determined by subtracting the counted volume from the total one. Table 2 demonstrates the logic of the PPB nodes for outgoing packet re-marking and for counting outgoing native traffic.

4.3 Incoming Packet Re-marking

As mentioned before, the website requested by a consumer can be subscribed either to the local network or to any other network. As a result, traffic originated by the endpoint of transmission (e.g. destination), can be part of a transmission initiated either by the network's customer or by the customer of any other network. Therefore, the identification the type of traffic (i.e. native or stranger) originated by the transmission endpoint is necessary. For *incoming traffic that is destined to the network* (i.e. destination

IP address is local), the PPB nodes perform the NOT logical operation on the *second bit* of the label and do not change the *first bit*.

A transmission endpoint does not re-examine the label. It sends *response* packets with the same ML field (i.e. the values '01' or '11' are copied from the request packet). It is obvious that incoming network traffic with the first bit set to '1' and destined to the network is a part of a transmission initiated by its own customers. Table 3 shows the logic of the PPB nodes for incoming traffic and for counting incoming native traffic. An example that helps to understand how the described traffic management mechanism works is described below.

Table 1. Available values of the ML field

Values of ML	Description
00	Stranger packet, originated by the endpoint located in another network
01	Stranger packet, originated by the local endpoint
10	Native packet, originated by the endpoint located in another network
11	Native packet, originated by the local endpoint

Table 2. Outgoing packet re-marking and counting

Input	Output	Counter
00	00	counter1
01	10	counter1
10	10	counter1
11	00	counter1++

counter shows the current value of the counter for outgoing traffic

Table 3. Incoming packet re-marking and counting

Input	Output (Counter)	
	If destination IP address is local	Otherwise
00	01 (counter2)	00 (counter2)
10	11 (counter2++)	10 (counter2)

counter shows the current value of the counter for incoming traffic

As an example, consider a model consisting of ISP_i, ISP_j, and their customers as well as the transit network ISP_k, where each provider calculates the volumes of native traffic. Assume that a customer of ISP_i requests data available on ISP_j. Let *N1* be the PPB node of ISP_i, which receives a packet marked by *'11'*. Before forwarding it to ISP_k, *N1* performs the XOR operation on the ML field of the outgoing packet (i.e. sets the label to *'00'*), and increases the counter for outgoing native traffic. The PPB node *N2* of ISP_k reads the destination IP address, however does not re-mark the label (since the packet is not destined to its network), and then forwards the packet to PPB node *N3*, which maintains connectivity with ISP_j. *N3* node performs the XOR operation on the outgoing packet label (as a result, the ML value remains the same, i.e. *'00'*) and forwards it to PPB node *N4* of ISP_j. *N4* node reads the destination IP address, and since the packet is destined for its network, applies the NOT operation on the label of the incoming packet (i.e. sets ML to *'01'*) and forwards it to the destination, e.g. the *N5* node. After receiving the packet, *N5* sends a packet stream with the requested data, where the label remains the same (*'01'* i.e. stranger traffic, which is originated locally). The similar procedure follows on the inverse path with only one difference that ISP_i considers the incoming traffic as native, initiated by its own customers.

4.4 Incentive Compatibility

It is well known that strategic agents have an incentive not to be truthful and, therefore, end-systems or the defined PPBs nodes can perform mendacious packet marking. However, there are several favorable reasons to adopt our approach. First, we considered that PPB is a trust boundary, and therefore, its operations can be recorded and then audited. Second, applying commonly used pricing scheme, such as flat-rate creates no incentive to the end-systems to perform untruthful packet marking, since it does not affect fees and quality of service. Finally, interconnection is a long-term and repeated process, arranged under mutual benefits, and, therefore, sustainable cooperation between interconnected ISPs is a reasonable and natural solution. Nevertheless, we intend to address incentive compatibility in our future work.

5 A Simple Benchmark

In our analysis two types of the customers, namely consumers and websites are considered. Actually, traffic is exchanged 1) between consumers, 2) between websites, 3) from websites to consumers, and 4) from consumers to websites. Generally, traffic between websites and from consumers to websites is negligible. Recently, peer-to-peer (P2P) traffic has increased rapidly. The significant part of the Internet traffic, comprised of FTP, Web, and streaming media traffic, is from websites to consumers. In order to investigate the impact of determination of the transmission initiator on the intercarrier compensation in its simplest way, we focus on traffic exchange i) from consumers to websites, and ii) from websites to consumers. Traffic between consumers and between websites is neglected, since it does not have any significant impact on the results of the analysis. It is worth noting that according to the proposed approach, a node in a P2P network can be considered as a consumer as well as a website simultaneously, since it can act as a server and a client. To simplify the analytical studies the following assumptions were made throughout the paper:

Assumption 1. Let $\alpha_i \in (0,1)$ be a network's market share for consumers and $\beta_i \in (0,1)$ its market share for websites. The market consists of only one transit provider and two customer networks, *i* and *j*, where *i*≠*j*, and $\alpha_i + \alpha_j = 1$, $\beta_i + \beta_j = 1$.

Assumption 2. The number of consumers and the number of websites in the market are denoted as N and M respectively. Each customer chooses only one provider to join, because of homogeneity of the services.

Assumption 3. For simplicity, a balanced calling pattern, where each consumer requests any website in any network with the same probability is considered. Each consumer originates one unit of traffic per request of website and downloads a fixed amount of content.

We examine a scenario, in which ISP_i and ISP_j exchange traffic through the transit provider ISP_k. The amount of differentiated traffic originated from ISP_i with destination ISP_k is given by

$$t_{ik}^{nat} = \alpha_i \beta_j NM \tag{1}$$

$$t_{ik}^{str} = \alpha_j \beta_i NMx \tag{2}$$

where t_{ik}^{nat} denotes the amount of outgoing native traffic (exchanged from consumers to websites) and t_{ik}^{str} the amount of outgoing stranger traffic (exchanged from websites to consumers) with respect to ISP$_i$. The variable x denotes the average amount of traffic caused by requesting a website.

Similarly, the DT volumes originated by ISP$_j$ and destined to ISP$_i$ are given by

$$t_{jk}^{nat} = \alpha_j \beta_i NM \tag{3}$$

$$t_{jk}^{str} = \alpha_i \beta_j NMx \tag{4}$$

Here, t_{jk}^{nat} represents the outgoing native traffic and t_{jk}^{str} represents the outgoing stranger traffic with respect to ISP$_j$. The total amount of traffic from ISP$_i$ and ISP$_j$ are calculated as

$$t_{ik} = t_{ik}^{nat} + t_{ik}^{str} \tag{5}$$

$$t_{jk} = t_{jk}^{nat} + t_{jk}^{str} \tag{6}$$

Since this paper is not about examining how the access charges are defined, therefore, we assume for the purpose of simplicity that access charges between providers are set by an industry regulator and then applied reciprocally. Let ISP$_i$ (ISP$_k$) charges ISP$_k$ (ISP$_i$) a_k^i (a_i^k) and b_k^i (b_i^k) for every unit of received native and stranger traffic respectively, where $a_k^i > b_k^i$ ($a_i^k > b_i^k$), since the providers compensate partially the costs of terminating stranger traffic. For the case of symmetric access charges $a_k^i = a_i^k = a_k^j = a_j^k = a$ and $b_k^i = b_i^k = b_k^j = b_j^k = b$, whereas $b = \varepsilon a$ and $0.5 \le \varepsilon < 1$. However, in order to simplify analysis, we fix $\varepsilon = 0.5$. The net interconnection payments from ISP$_i$ to the transit provider and vice versa are denoted by q_{ik} and q_{ki} correspondingly

$$q_{ik} = a t_{ik}^{nat} + b t_{ik}^{str} \tag{7}$$

$$q_{ki} = b\left(t_{jk}^{nat} + t_{jk}^{str}\right) \tag{8}$$

From (8), it can be noticed that the transit provider is charged based on the rate for stranger traffic, because it does not have any customers of its own. Similarly, the net transfers from ISP$_j$ to the transit provider and vice versa are denoted by q_{jk} and q_{kj} respectively

$$q_{jk} = a t_{jk}^{nat} + b t_{jk}^{str} \tag{9}$$

$$q_{kj} = b\left(t_{ik}^{nat} + t_{ik}^{str}\right) \tag{10}$$

The costs of ISP$_i$ (ISP$_j$) can be interpreted as a composition of two independent components i) one for native traffic business, and ii) another for stranger traffic business.

Proposition 1. *If $\alpha_i = \alpha_j$ and $\beta_i = \beta_j$, then the costs of the customer network providers are the same.*

Proof: From the conditions (1)-(4) follows that $t_{ik}^{nat} + t_{ik}^{str} = t_{jk}^{nat} + t_{jk}^{str}$. As a result, using (7) and (9) it can be obtained that $q_{ik} = q_{jk}$.

Proposition 2. *If* $\alpha_i = \alpha_j$ *and* $\beta_i > \beta_j$*, then the costs of* ISP_i *are higher than the costs of* ISP_j*.*

Proof: Observing conditions (1)-(4) it can be obtained that $t_{ik}^{nat} + t_{ik}^{str} > t_{jk}^{nat} + t_{jk}^{str}$. Consequently, from the conditions (7) and (9) follows that $q_{ik} > q_{jk}$.

Proposition 3. *If* $\alpha_i > \alpha_j$ *and* $\beta_i = \beta_j$*, then the costs of* ISP_i *are lower than the costs of* ISP_j*.*

Proof: From the conditions (1)-(4) follows that $t_{ik}^{nat} + t_{ik}^{str} < t_{jk}^{nat} + t_{jk}^{str}$. Hence, from the conditions (7) and (9), it can be obtained that $q_{ik} < q_{jk}$.

When $\alpha_i > \alpha_j$ and $\beta_i > \beta_j$, the following cases for traffic volumes are obtained from the conditions (5) and (6): 1) $t_{ik} > t_{jk}$, 2) $t_{ik} < t_{jk}$, and 3) $t_{ik} = t_{jk}$. The cases 1) and 2) are analogous to those described above. The last case when $t_{ik} = t_{jk}$ is analyzed below.

Proposition 4. *If* $\alpha_i > \alpha_j$, $\beta_i > \beta_j$*, and* $t_{ik} = t_{jk}$*, then* $\alpha_i = \beta_i$.

Proof: The result is obtained from the conditions (1)-(6).

Corollary 1. *If* $\alpha_i > \alpha_j$, $\beta_i > \beta_j$*, and* $t_{ik} = t_{jk}$*, then* $t_{ik}^{nat} = t_{jk}^{nat}$ *and* $t_{ik}^{str} = t_{jk}^{str}$.

Proposition 5. *If* $\alpha_i > \alpha_j$, $\beta_i > \beta_j$*, and* $t_{ik} = t_{jk}$ *then the costs of the customer providers are equal.*

Proof: The result is obtained from the conditions (7) and (9).

Proposition 6. *If* $\alpha_i > \alpha_j$ *and* $\beta_i < \beta_j$*, then the costs of* ISP_j *are higher than the costs of* ISP_i*.*

Proof: Considering conditions (1)-(4) it can be obtained that $t_{ik}^{nat} + t_{ik}^{str} < t_{jk}^{nat} + t_{jk}^{str}$. As a result, from the conditions (7) and (9) follows that $q_{ik} < q_{jk}$.

Table 4. Results of DTIA

Case	α	β	t^{nat}	t^{str}	q
I	$\alpha_i = \alpha_j$	$\beta_i = \beta_j$	$t_{ik}^{nat} = t_{jk}^{nat}$	$t_{ik}^{str} = t_{jk}^{str}$	$q_{ik} = q_{jk}$
II	$\alpha_i = \alpha_j$	$\beta_i > \beta_j$	$t_{ik}^{nat} < t_{jk}^{nat}$	$t_{ik}^{str} > t_{jk}^{str}$	$q_{ik} > q_{jk}$
III	$\alpha_i > \alpha_j$	$\beta_i = \beta_j$	$t_{ik}^{nat} > t_{jk}^{nat}$	$t_{ik}^{str} < t_{jk}^{str}$	$q_{ik} < q_{jk}$
IV	$\alpha_i > \alpha_j$	$\beta_i > \beta_j$	*If* $t_{ik}^{nat} = t_{jk}^{nat}$	*If* $t_{ik}^{str} = t_{jk}^{str}$	$q_{ik} = q_{jk}$
V	$\alpha_i > \alpha_j$	$\beta_i < \beta_j$	$t_{ik}^{nat} > t_{jk}^{nat}$	$t_{ik}^{str} < t_{jk}^{str}$	$q_{ik} < q_{jk}$

Tables 4 and 5 summarize the outcomes of the analytical studies. Table 4 shows how the interconnection payments of the customer providers depend on the DT flows. In addition to this, the results demonstrate the influence of providers' market shares on intercarrier compensation.

Table 5. Comparative results of the agreements based on traffic flow (TF) and DTIA compensation

Case	α_i	β_i	t_{ik}^{nat}	t_k^{str}	t_{jk}^{nat}	t_{jk}^{str}	q_{ik}		q_{jk}		$q_k = q_{ik} + q_{kj}$		π_k	
							DTIA	TF	DTIA	TF	DTIA	TF	DTIA	TF
I $\alpha_i = \alpha_j$ $\beta_i = \beta_j$	0.5	0.9	1500	52500	1500	52500	27750	108000	27750	108000	54000	0	1500	216000
II $\alpha_i = \alpha_j$	0.5	0.9	300	94500	2700	10500	47550	108000	7950	108000	54000	0	1500	216000
$\beta_i > \beta_j$	0.5	0.8	600	84000	2400	21000	42600	108000	12900	108000	54000	0	1500	216000
	0.5	0.7	900	73500	2100	31500	37650	108000	17850	108000	54000	0	1500	216000
	0.5	0.6	1200	63000	1800	42000	32700	108000	22800	108000	54000	0	1500	216000
III $\alpha_i > \alpha_j$	0.9	0.5	2700	10500	300	94500	7950	108000	47550	108000	54000	0	1500	216000
$\beta_i = \beta_j$	0.8	0.5	2400	21000	600	84000	12900	108000	42600	108000	54000	0	1500	216000
	0.7	0.5	2100	31500	900	73500	17850	108000	37650	108000	54000	0	1500	216000
	0.6	0.5	1800	42000	1200	63000	22800	108000	32700	108000	54000	0	1500	216000
IV $\alpha_i > \alpha_j$	0.9	0.9	540	18900	540	18900	9990	38880	9990	38880	19440	0	540	77760
$\beta_i > \beta_j$	0.8	0.8	960	33600	960	33600	17760	69120	17760	69120	34560	0	960	138240
	0.7	0.7	1260	44100	1260	44100	23310	90720	23310	90720	45360	0	1260	181440
	0.6	0.6	1440	50400	1440	50400	26640	103680	26640	103680	51840	0	1440	207360
V $\alpha_i > \alpha_j$	0.9	0.2	4320	4200	120	151200	6420	159840	75720	159840	79920	0	2220	319680
$\beta_i < \beta_j$	0.8	0.25	3600	10500	300	126000	8850	140400	63300	140400	70200	0	1950	280800
	0.7	0.35	2730	22050	630	95550	13755	120960	48405	120960	60480	0	1680	241920
	0.6	0.4	2160	33600	960	75600	18960	112320	38760	112320	56160	0	1560	224640

q_k = total costs of the transit provider
π_k = $q_{ik} + q_{jk} - q_k$ provider k's profit obtained from interconnection

The comparison results between the unilateral settlements based on the traffic flows compensation and DTIA, where payments are made on the DT flows are presented in Table 5. In order to calculate specific outcomes, we imposed the following parameter values $a=1$, $x=35$, $N=100$, and $M=60$. The following observations can be made from the obtained results. Firstly, in comparison to the unilateral settlement, under which the customer providers compensate equally, in DTIA the customer provider that sends more traffic compensates more. Secondly, considering transit provider costs, it can be noticed that in DTIA, the transit provider along with the customer providers carries the burden of the interconnection costs. In particular, in contrast to the classical model, compensations of the transit ISP according to the proposed model are different from zero. As a result, determination of a transmission initiator induces reduction in the interconnection payments subsidized by the customer ISPs. And finally, under bilateral DTIA, the profits of the transit provider obtained from interconnection are decreased, since costs are allocated between all parties.

6 Conclusions

In this paper we described a new bilateral model, called Differentiated Traffic-based Interconnection Agreement (DTIA), for intercarrier compensation between providers. We proposed to differentiate traffic into two types, referred to as *native* and *stranger* in order to determine an original initiator of a transmission for calculating intercarrier compensation. In comparison to the existing financial settlement agreements, under which the payments are based on traffic flows, the described model governs cost compensation according to the differentiated traffic flows. More specifically, each provider

is compensated fully for the costs incurred from delivering native traffic, which is originally initiated by its own customers, and partially for the costs incurred from carrying stranger traffic that is originally initiated by the customers of any other network.

For supporting DTIA, we designed a traffic management mechanism, in which only border nodes perform packet management. The main advantage of the presented mechanism is its simplicity and scalability that is a basic requirement for a deployment in the Internet. In particular, the provider has not to maintain a complex identification process of transmission initiator and to inspect the IP header of packets in order to determine and record all subsequent packets of the transmission. Instead, the defined membership label (ML) allows accounting the volume of the appropriate traffic type and, therefore, leads to low computational complexity (see Table 1). The logic of the border nodes for packet marking and counting is demonstrated as well (see Tables 2 and 3).

Our analytical studies showed how the interconnection payments differ to the existing solution (see Tables 4 and 5). The comparative analysis between the classical model and DTIA indicated that determination of a transmission initiator reduces the payments of the customer providers. This is achieved due to the fact that the transit provider along with the customer ISPs shares the interconnection costs. Overall, it can be concluded that the DTIA model is beneficial for the customer providers, since it outperforms the classical model in terms of payments, which are relatively small and unequal.

Acknowledgment

The authors would like to thank Kristina Davoian from University of Münster for her useful comments and helpful discussions.

References

1. Norton, W.: Internet Service Providers and Peering. Draft 2.5, Equinix White Papers (2001)
2. Courcoubetis, C., Weber, R.: Pricing Communication Networks: Economics, Technology and Modeling. Wiley, Chichester (2003)
3. Huston, G.: Interconnection, Peering, and Settlement. Part II, Internet Protocol Journal 2(2), 2–23 (1999b)
4. Norton, W.B.: A Business Case for ISP Peering (2001)
5. Kende, M.: The Digital Handshake: Connecting Internet Backbones. Federal Communications Commission (FCC), OPP Working Paper No. 32 (September 2000)
6. Faratin, P., Clark, D., Gilmore, P., Bauer, S., Berger, A., Lehr, W.: Complexity of Internet Interconnections: Technology, Incentives and Implications for Policy. In: Thirty Fifth Telecommunication Policy Research Conference (TPRC) (September 2007)
7. Laffont, J.J., Tirole, J.: Competition in Telecommunications. MIT Press, Cambridge (2000)
8. Armstrong, M.: The Theory of Access Pricing and Interconnection. In: Cave, M.E., Mujumdar, S.K., Vogelsang, I. (eds.) Handbook of Telecommunications Economics, vol. 1. North-Holland, Amsterdam (2002)

9. Armstrong, M.: Competition in Two-sided Markets. RAND Journal of Economics (2006)
10. DeGraba, P.: Bill and Keep at the Central Office as the Efficient Interconnection Regime. FCC, OSP, Working Paper 33 (December 2000)
11. Federal Communications Commission. In the Matter of Developing a Unified Intercarrier Compensation Regime. Notice of Proposed Rulemaking in CC Docket No. 01-92, April 27 (2001)
12. Yoon, K.: Interconnection Economics of all-IP Networks. Review of Network Economics V 5 (2006)
13. Laffont, J.J., Marcus, S., Rey, P., Tirole, J.: Internet Interconnection and the Off-net-cost Pricing Principle. RAND Journal of Economics 34(2), 370–390 (2003)
14. Carter, M., Wright, J.: Interconnection in Network Industries. Review of Industrial Organisation (1999)
15. Carter, M., Wright, J.: Asymmetric Network Interconnection. Review of Industrial Organization 22(1), 27–46 (2003)
16. Besen, S., Milgrom, P., Mitchell, B., Srinagesh, P.: Advancing in Routing Technologies and Internet Peering Agreements. American Economic Review 91, 292–296 (2001)
17. Davoyan, R., Altmann, J.: Investigating the Influence of Market Shares on Interconnection Settlements. In: Proceedings of IEEE Globecom (2008)
18. Davoyan, R., Altmann, J.: Investigating the Role of Transmission Initiator in Private Peering Arrangements. In: Proceedings of IFIP/IEEE IM. IEEE Computer Society Press, Los Alamitos (2009) (accepted)
19. Davoyan, R.: DTIA: Differentiated Traffic-based Interconnection Agreement. In: Proceedings of ISCIT. IEEE Press, Los Alamitos (2008)
20. Stiller, B.: A survey of Charging Internet Services. In: Aidarous, S., Plevyak, T. (eds.) Managing IP Networks: Challenges and Opportunities, September. Wiley-IEEE Press (2003)
21. Huston, G.: Interconnection and Peering. Broad Band Satellite (November 2000)
22. Economides, N.: The Economics of Networks. International Journal of Industrial Organization 14(6), 673–699 (1996)
23. Marcus, J.: Interconnection in an NGN Environment. In: Background paper for ITU Workshop on What rules for IP-enabled NGNs? (2006)
24. Huston, G.: ISP Survival Guide: Strategies for Running a Competitive ISP. Wiley, Chichester (1998)

Improvement of BitTorrent Performance and Inter-domain Traffic by Inserting ISP-Owned Peers

Ioanna Papafili, Sergios Soursos, and George D. Stamoulis

Athens University of Economics and Business,
Department of Informatics,
76 Patission Str., Athens 10434, Greece
{iopapafi,sns,gstamoul}@aueb.gr

Abstract. Large volumes of Internet traffic are nowadays generated by BitTorrent. In this article, we propose the insertion of high-bandwidth ISP-owned peers as an optimization approach to improve end-users' performance and reduce inter-domain traffic. An ISP-owned peer participates in BitTorrent swarms in order to download chunks and subsequently serve regular peers. We have run simulations on the ns-2 platform showing that our approach results in considerable reduction of both inter-domain traffic and the downloading times of users. We also show that the insertion of an ISP-owned peer can complement effectively the use of locality awareness, and lead to further performance improvements.

Keywords: BitTorrent, insertion, interdomain, QoE, locality.

1 Introduction

File-sharing applications are used widely by Internet users to share content such as music tracks, movies or even software releases. Because of their high popularity and the large size of the files that are shared, file sharing applications generate huge volumes of traffic in the Internet. This in turn implies a change of traffic patterns and an increase of costs (in terms of both CAPEX and OPEX) for the ISPs. In addition, monetary penalties related to the existing interconnection agreements are incurred and the nature of such agreements might also change, e.g. a peering agreement may have to be converted to a transit one, due to the change of traffic ratio between the peering ISPs. Consequently, it is necessary for the underlay network to take the overlay applications and their generated traffic into serious account in order to achieve efficient traffic management and optimal resource utilization in the underlying network. However, the ISP should meet his objectives in a way that is incentive compatible for the overlay provider, i.e. improving (or at least retaining) the overlay application performance. This is actually the topic of FP7-ICT project SmoothIT [1].

BitTorrent being the most popular file sharing application is the source for up to 60% of the overall traffic in the Internet. The BitTorrent protocol [2] was originally designed and implemented with the objective to disseminate one large file or a composition of large files to a large number of users without the original distributor

P. Reichl, B. Stiller, and B. Tuffin (Eds.): ICQT 2009, LNCS 5539, pp. 97–108, 2009.

incurring entirely the costs of hardware, hosting and bandwidth resources. BitTorrent can be deployed either by using trackers, or by using structured lookup overlays without trackers, the so-called trackerless BitTorrent. The tracker is a centralized component which stores information about all peers that participate in a swarm. Its main role is helping peers to discover other peers.

An optimization approach that has been employed in several research works is the so-called locality awareness. This amounts to biased selection of peers based on locality criteria; e.g. being in the same autonomous system as the requesting peer. In this paper, we propose and investigate an innovative approach to achieve a more efficient operation of the underlying network and therefore a cost reduction for the ISPs together with performance improvements for the BitTorrent users; namely, the insertion of a high-bandwidth ISP-owned peer, which aims to reduce both inter-domain traffic and downloading-completion times. The motivation for this approach stems from BitTorrent's tit-for-tat mechanism; due to the high upload capacity of the ISP-owned peer, regular peers establish connections to the IoP with higher probability than to other peers, thus resulting in performance improvements. Moreover, we consider the combination of the ISP-owned peer insertion together with locality awareness. In this case, even further performance improvement is expected due to the fact that local peers are more likely to select the ISP-owned peer to download from. The performance improvements attained by these approaches are evaluated by means of simulation experiments, which verify the aforementioned arguments. We also explain that pure locality awareness without the ISP-owned peer may not always be a beneficial approach for the ISP, contrary to what is widely argued in the literature. It should be noted that the insertion of an ISP-owned peer coincides neither with the use of an intervening cache as proposed in [10], nor with the enforcement of biased selection of peers as studied in the various articles overviewed in Section 2.

The paper is organized as follows: in Section 2, we present research works related to optimization of BitTorrent performance. In Section 3, we propose the insertion of ISP-owned peer as an optimization approach and discuss implementation issues. In Section 4, we describe our simulation model. In Section 5, we present and discuss our simulation results. Finally, in Section 6, we further discuss conclusions and open issues to be studied in the future.

2 Related Work

Peer-assisted content distribution is a cost-effective and bandwidth-intensive solution for ISPs. Indeed, peer-assisted, decentralized and self-organized systems such as BitTorrent provide significant benefits to end-users and content providers. However, such systems create their own logical networks and perform their own routing based on performance metrics, without taking into account the underlying topology. In this sense, overlay paths might end up reusing unnecessarily physical links or even containing circles. On the other hand, network management may not take into account the requirements of overlay applications, leading to a tussle between ISPs and Overlay Providers. This is due to information asymmetry, and may cause an increase of traffic on ISPs both intra- and inter-domain links and, as a result, higher costs. In [3] and [4], it is shown that this objective misalignment of ISPs and P2P networks in combination with the information asymmetry lead to performance degradation both for the underlay

and the overlay. An ISP-friendly peer-assisted content distribution protocol that would exploit topology information is expected in [5] to alleviate ISPs' induced costs and as well as to improve P2P users QoE.

It can be intuitively expected that using topological information in a P2P system would significantly improve network performance (both overlay and underlay), if a better selection of a "good" server or a close-by peer, in terms of latency, were performed. This approach has been undertaken in several research works. In particular, in order to achieve this, a distributed, scalable binning scheme that requires a small number of landmark machines spread across the Internet is proposed in [6]. Because of landmarking being not self-organizing though, a new model is proposed in [7]. The model uses Geographical Longest Prefix Matching (Geo-LPM) and RTT to organize nodes into clusters each of which is a group of nodes that share a common prefix and are close to each other. The fact that clustered Geo-LPM is combined with an appropriate RTT threshold ensures that a node entering the network will find other overlay nodes that belong to the same physical domain. However, in case of clusters that share a common prefix, a solution is given by means of Geo-Partitioning. Furthermore, in [8] a 2-D Euclidean space model of the connectivity among BitTorrent peers has been proposed with the objective to evaluate BitTorrent's topology. Parameters such as number of peers in the swarm, maximum number of unchokes, etc., have been also taken into account. Moreover, in [9], a lightweight approach to reduce inter-ISP costs is proposed that exploits network information derived at low cost from CDN queries.

In [10], biased neighbor selection is studied as an approach to enhance BitTorrent traffic locality, in which a peer is enforced by the tracker to select the majority of his neighbors from peers within the same ISP and only a few (namely, k neighbors) that are outside the ISP. Additionally, the peer is modified to request a new list of neighbors whenever its peer list has less local peers than a specific threshold. This locality-awareness scheme can be implemented either by modifying tracker and client, or by situating P2P shaping devices along-side the edge routers of the ISPs, so that deep packet inspection is used to identify P2P traffic and manipulate it accordingly by intercepting and modifying the exchanged messages. Instead of enforcing locality, in [11] an 'oracle' is proposed that ranks peers according to some metric, e.g. proximity, bandwidth, etc., and provides this underlay information to users so that they can choose appropriate neighbors. In order to reduce downloading times of BitTorrent networks, also alternative chunk selection policies have been proposed in [12], while in [13] a cost-aware model to reduce both ISPs' costs and distribution time is proposed. The latter approach employs also alternative peer selection policies based on chunk availability on each peer. Finally, in [14] alternative peer selection based on RTT and number of hops is considered, which seems to reduce ingress inter-domain traffic as well as downloading times.

3 Insertion of ISP-Owned Peers

An **I**SP-**o**wned **p**eer (IoP) is an entity that aims at increasing the level of traffic locality within an ISP and at improving the performance enjoyed by the users of peer-to-peer applications. The IoP, either belongs to an ISP's infrastructure and is controlled by the ISP itself; or is a regular but **h**ighly **a**ctive **p**eer (HAP) that is granted by the ISP with extra resources, e.g. higher downlink/uplink bandwidth, at no extra cost. In

principle, if dynamic adjustment of the end-user's bandwidth is possible, then the end-users might even not be aware of this enhancement. However, agreement between the ISP and the HAP is also meaningful in order for the approach to be more effective; e.g., in order to assure extended seeding time by the HAP. In any case, the most important issue and what differentiates IoP from other related approaches, is that IoP runs the standard overlay protocol, e.g. BitTorrent, like every other peer in the swarm; yet, there are introduced certain changes in some parameters of the protocol that serve IoP's purposes and that are beneficial for other peers as well. In particular, the IoP is capable of unchoking more peers than the regular ones, in order to exploit its extra uplink capacity. Since the IoP runs the overlay protocol, it is also assumed that is capable of storing the content that it downloads and of course offering it back to the network. In other words, until an IoP has a complete copy of a file, it is considered to be a leecher in that file's swarm; subsequently, it is considered to be a seed. Henceforth, we use only the term ISP-owned peer, and include the HAP in this as well. The term "HAP' is used only when certain subtleties of this approach are discussed. Below, we distinguish two approaches for deploying an ISP-owned peer:

A. Plain insertion of IoP in a BitTorrent network: All peers are assumed to run the original BitTorrent protocol. No other mechanism such as locality awareness is employed by the ISP, and no agreement with the overlay provider is considered. Thus, the overlay, e.g. the tracker, is *not* aware of the IoP's existence as a special entity but treats it as a regular peer. In this case, the IoP is expected to be preferred by other peers due to the tit-for-tat mechanism employed by BitTorrent's unchoking algorithm and because of its high uplink capacity. The IoP follows here the tit-for-tat rule exploiting the immediate incentives of the latter that are directly related to the underlay [14].

B. Combination of IoP with locality-awareness mechanisms: The use of locality-awareness mechanisms that affect the overlay network's structure is considered here as being imposed by the ISP. Furthermore, depending on their implementation, these mechanisms could be either: a) transparent to the peers, i.e., they run along with the original protocol, or b) non-transparent i.e., they are introduced along with a modified version of the protocol. Metrics that can be used as proximity criteria are RTT and number of hops associated with remote peers, peers' autonomous system identity, BGP information, etc. Due to these locality-awareness mechanisms, the IoP would be mostly preferred by peers that are 'closer' to it according to one or more of the proximity criteria.

Below important issues regarding implementation are addressed:

Dimensioning of the IoPs: Dimensioning is expressed in terms of downlink/uplink bandwidth and storage capacity the IoP should be equipped with. Recall that the aim of the approach is twofold: meet the objectives of the ISP while coping with the users' performance requirements from the overlay application, since it is highly important not to downgrade their completion times.

Number of IoPs: Increasing the number of IoPs up to a certain number implies improvement of performance but also increases the CAPEX of the ISP. Additionally, the more IoPs exist in an ISP's network, the more intra-domain traffic is generated. After this traffic exceeds a threshold, more congestion on intra-domain links may result, thus leading to deteriorated performance and increase of OPEX for the ISPs. Thus, the number of IoPs should be carefully selected.

Physical location of the IoPs: The ISP should decide, based on the overlay traffic patterns, the physical locations where the IoPs should be deployed, e.g. 1) one "large" IoP in a specific location (centralized approach), 2) multiple "smaller" IoPs in a specific location (moderately centralized) or 3) multiple "smaller" IoPs in different locations (decentralized). Terms such as "large" or "small" refer to resources capacity. Before the selection of such an approach many issues require to be addressed such as availability, content duplication, etc. The location of the IoPs is related but not identical to problems on cache dimensioning and placement. Related techniques from that field could be employed.

Generally, both the number and the location of the IoPs within an ISP have to be decided by the ISP itself taking under consideration traffic measurements on inter- and intra-domain links, as well as impact of the traffic on the interconnection costs.

Content Selection. The ISP has also to make certain decisions that are expected to have impact on the efficiency of the IoP. First, the ISP should decide on which content will the IoP be downloading, i.e., in which swarms to participate. The selection of the content can be deployed either in a centralized or in distributed way. In the centralized cases, it could be performed with or without human intervention. In the distributed case, it would probably be more efficient, if it were performed automatically. In particular, content selection approaches could be:

1) Trial-and-error: The IoP could join randomly selected swarms in popular trackers, monitor whether his intervention has the desired impact for the ISP and decide whether to maintain its position, and/or when to leave a swarm etc.
2) Swarm-size based: The selection of content to be downloaded would greatly benefit from information provided by the overlay, e.g. trackers keeping statistics about the number of peers that participate in each file's swarm.
3) Popularity-based: The underlying idea is that the IoP should download a file that is expected to become popular before other peers start asking for it even of the swarm size is originally small but expected to become larger.

Content legality. In the case of IoPs (excluding HAPs) the content downloaded is stored in ISP's equipment. Thus, only licensed or non-copyrighted content can be downloaded by the ISP. Additionally, the ISP could establish agreements with content providers, e.g. content distribution networks, software vendors, music industry, movies distributors, TV channels, etc. and they should consequently establish agreements with the overlay in order to download, store and serve licensed content. On the other hand, in the case of HAPs no licenses or agreements are required since the content is stored in the users' premises.

4 Simulation Model

Our simulation experiments were performed on the ns-2 simulator [16] using the BitTorrent patch [17] implemented by K. Eger. This patch contains four classes that implement a simplified version of the BitTorrent protocol that was not originally implemented for the ns-2 platform. We have modified several methods of the BitTorrent classes in order to deploy a locality-aware BitTorrent protocol, which is employed in half of the experiments.

As a base topology, we use the Dumbbell Topology, e.g. a complex topology that comes from the interconnection of two simple star topologies (Fig. 1). Each star topology represents an AS; the left network is AS0 and the right network is AS1. We considered both the symmetric case (the two ASs have the same number of peers), e.g. two Tier-2 ISPs, and the asymmetric case (one AS has many more peers than the other), e.g. corresponding to a Tier-2 ISP and a Tier-3 ISP. Each peer within an AS is considered to be a regular BitTorrent peer and is connected to the router of its AS via a duplex asymmetric intra-domain link. One of the peers belonging to AS0 is considered to be the unique seed of the swarm and appears in the system at time 0. Furthermore, each peer becomes a seed after finishing its download.

In Table 1, we present all parameters used in the simulation experiments and their respective values.

Table 1. Simulation parameters

Description	Value	Description	Value
Number of regular BT peers	50, 52	Number of local peers replied by tracker	20, 15
Number of initial seed	1	Number of autonomous systems (AS)	2
File size	20M	Number of IoP(s)	1, 2
Upload capacity of reg. peers	512K	Download capacity of regular peers	8x512K
Unchoked connections permitted per reg. peer	4	Unchoked connection permitted per IoP	20
Number of open connection per peer	20	Up-/download capacity of IoP	20x512K
Total number of peers replied by tracker	25	Inter-AS capacity (each direction)	30x512K

The main performance metrics of interest are the users' downloading time and the ingress inter-domain traffic to both ASs. Downloading time is defined as the difference between the time when the peer received the last chunk of the file and the time when the peer sent a request to the tracker to get a first list of neighbors. Ingress inter-domain traffic is measured in the inter-domain link, i.e. the link that interconnects the two star topologies. We monitor traffic to both directions in order to make conclusions for interconnection costs following specific charging schemes. Namely, we measure all TCP traffic moving towards AS1 and denote this as ingress inter-domain traffic to AS1. The traffic is calculated from the original trace file generated by ns-2. We considered different scenarios in a variety of cases:

- **Pure BitTorrent scenario (BT):** The tracker replies a random list of peers to each peer's request.
- **BitTorrent with locality awareness:** We assume that the tracker replies a localized list according to *Biased Neighbor Selection* approach of [9].
- **Insertion of IoP in BitTorrent:** One ISP-owned peer (down/up-load capacity: 20x512K, 20 unchokes) is inserted in AS1 in pure BT (Fig. 1).
- **Insertion of IoP in BitTorrent with locality awareness:** Combination of the two aforementioned scenarios (Fig. 2).

Below, we present further variations of the above scenarios:

- **Symmetric:** AS0 and AS1 have 25 regular peers each. The seed is in AS0.
- **Asymmetric:** AS0 has 35 regular peers, while AS1 has only 15.
- **All-together:** All peers' starting times are selected according to the uniform distribution U(0,10); note that 10 coincides with the choking interval.
- **Split:** The 5 latest peers' starting times of each AS are selected according to U(150,300); the IoP always follows U(0,10).
  The experimental results, e.g. the downloading times and the traffic volumes on inter-domain links, are presented in trace figures in the section 5.

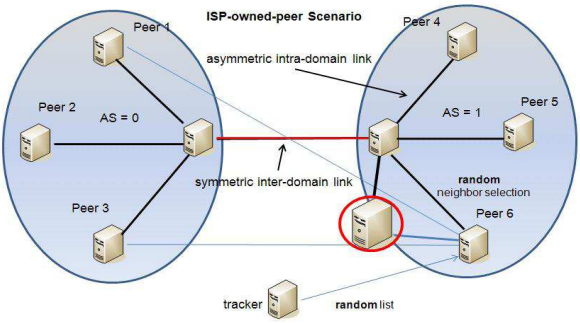


Fig. 1. Insertion of ISP-owned peer in pure BT

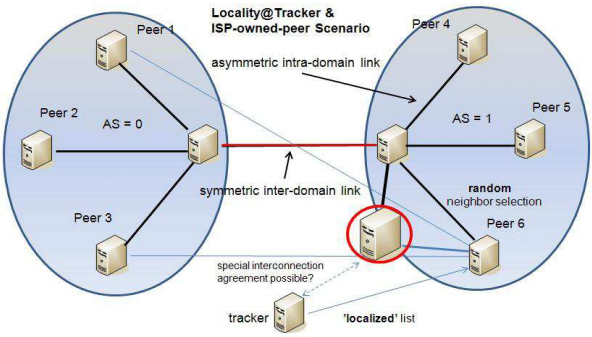


Fig. 2. Insertion of ISP-owned peer in BT with locality awareness

5 Results

Reduction of inter-domain traffic: The insertion of ISP-owned peers achieves important reduction of the inter-domain traffic that enters the AS which deploys the IoP, namely in our scenario AS1. On the other hand, due to the fact that no constraints are posed, the IoP can serve peers outside its domain; thus, inter-domain traffic that exits

AS1 towards AS0 will be increased. In Fig. 1 and Fig. 2, the ingress inter-domain traffic to AS1, in symmetric and asymmetric cases respectively, is shown. We note here that we consider only the case where peers start "all together", but similar results hold also for the "split" case. In each case, we compare all four scenarios: pure BitTorrent, insertion of IoP in pure BitTorrent, BitTorrent and locality awareness and insertion of IoP in BitTorrent with locality. In the symmetric scenario, we observe that the insertion of the IoP achieves up to 35% inbound traffic reduction both in pure BitTorrent scenario (red-cyan lines) and in BitTorrent employing locality awareness (blue-green lines). Overall, we see that the combination of the IoP with locality awareness brings up to 53% improvement of the traffic compared to the pure BitTorrent scenario where no locality or IoP are considered. Note also that the IoP insertion on its own achieves similar results to pure locality awareness. On the other hand, in the asymmetric scenario, we observe up to 31% and 37% reduction of inbound traffic by the IoP insertion, in pure BitTorrent and BitTorrent with locality awareness respectively. In this case the IoP insertion on its own is *more* effective than pure locality. The gain reaches up to 44% when comparing the IoP combined with locality awareness and the pure BitTorrent scenarios. On the other hand, the traffic that enters AS0 increases up to 10-20% (not presented here due to space limitations) when compared with the respective non-IoP scenario, due to aforementioned reasons.

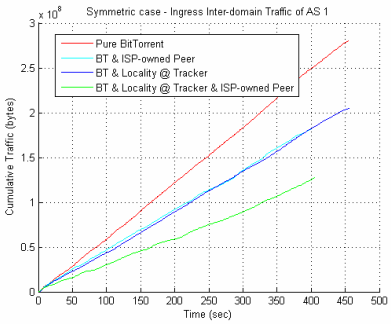


Fig. 1. Ingress inter-domain traffic to AS1 (symmetric case)

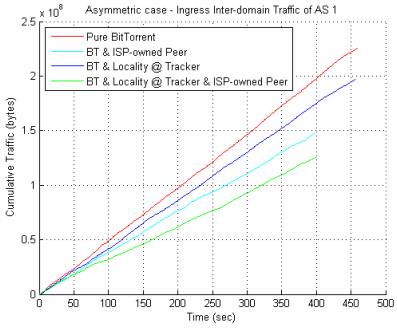


Fig. 2. Ingress inter-domain traffic to AS1 (asymmetric case)

Reduction of end-users' completion times: In Fig. 3 and Fig. 4, we compare the end-users' completion times for pure BitTorrent vs. insertion of IoP in pure BitTorrent, and BitTorrent with locality vs. insertion of IoP in BitTorrent with locality, respectively. In each figure are presented: (a) the simulation times and (b) the relative improvement of the completion times of the two scenarios compared. We have considered here the symmetric and "split" case, where the two ASs have same number of peers and some of the peers (peers with id 20 to 25 from AS0 and peers with id 45 to 50 from AS1) enter the swarm later. We observe that the insertion of IoP has important impact on the completion times. In particular, reduction of times up to 15% for peers starting along with the IoP (similar reduction is achieved also in "all-together" case where all peers start along with the IoP), and reduction up to 35-40% for those starting later (spikes in Fig. 3 and Fig. 4).

Insertion of IoP vs. locality awareness; comparison w.r.t. reduction of charge for inter-domain traffic: As already noted, insertion of the IoP results in a higher reduction of inbound traffic, than locality awareness, particularly in the asymmetric scenario (see Fig. 2), which fits better to cases of transit agreements. Clearly, under a charging scheme for the inter-domain traffic that is based on statistics of the inbound traffic, the IoP insertion would lead to a higher reduction of charge than locality awareness. Furthermore, we turn attention to compare the effectiveness of the two approaches under charging models that are based on the *difference* between inbound and outbound traffic, e.g. using the 95th percentile rule[1]. Note that this 95th percentile charging scheme, like any difference-based scheme, is sensitive to asymmetric changes only, while symmetric changes have no direct impact on the costs.

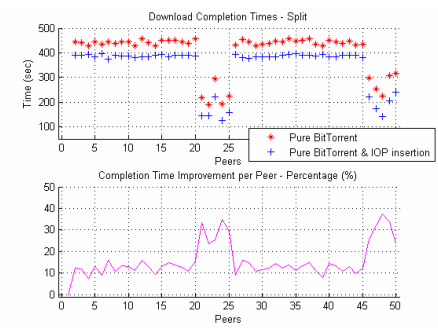


Fig. 3. End-users' completion times (scenarios without locality)

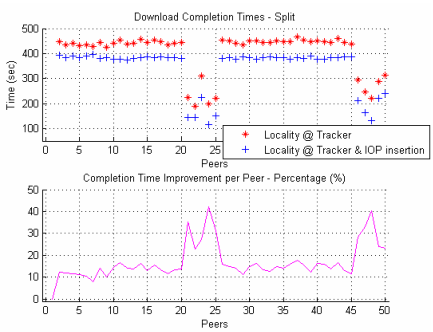


Fig. 4. End-user's completion times (scenarios with locality)

To this end, we present in Fig. 5 and Fig. 6 the instantaneous *difference* of inbound and outbound traffic to AS1, for pure BitTorrent and BitTorrent with locality, respectively. Again, we have considered here the asymmetric scenarios, i.e. AS0 has 35 peers and AS1 has 15 peers. We can observe that locality awareness on its own achieves more or less symmetric reduction of inbound and outbound inter-domain traffic. Indeed, the corresponding difference curve fluctuates around zero (Fig. 6 – top curve). This is due to the tit-for-tat mechanism that assures that the amount of traffic transferred to both directions is equivalent. On the contrary, the insertion of IoP in AS1 achieves asymmetric traffic reduction, regardless of whether locality is employed or not (Fig. 5 & Fig. 6 – bottom curves), due to the fact that the IoP quickly turns into a seed that serves peers regardless of the tit-for-tat mechanism which does not apply for seeds by the definition of the BitTorrent protocol. While locality does not affect this difference, the IoP clearly shifts the traffic difference in the favor of the AS deploying the IoP. When interpreting AS1 as a Tier 3-ISP and AS0 as a Tier 2-ISP the use of an IoP is beneficial for the Tier 3-ISP. Whether there is an actual monetary benefit depends on the OPEX and CAPEX for the IoP insertion, on the parameters of the charging model considered in each case and the achieved traffic reduction.

[1] The 95th percentile rule is applied every 5 seconds. The difference of inbound and outbound traffic is calculated and the upper 5% of that difference is cut away. The rest is what the ISP is charged for.

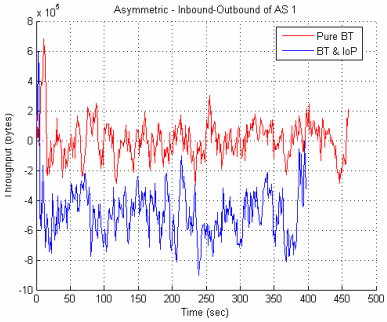


Fig. 5. Pure BitTorrent: Instantaneous difference between inbound-outbound to AS1

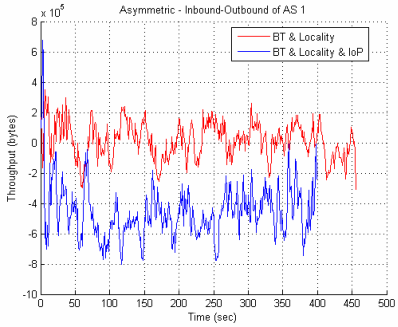


Fig. 6. BitTorrent with locality: Instantaneous difference between inbound-outbound to AS1

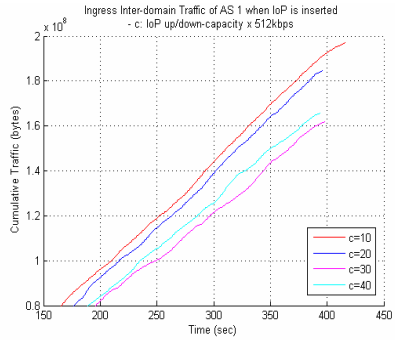


Fig. 7. Ingress traffic to AS1 for IoP capacity, c = {10, 20, 30, 40}x512kbps

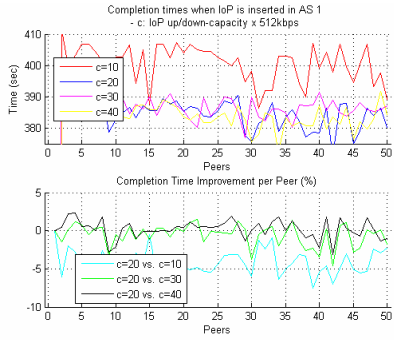


Fig. 8. End-user's completion times for IoP capacity, c = {10, 20, 30, 40}x512 kbps

Impact of the IoP dimensioning on performance: Figures 7 and 8 present the ingress inter-domain traffic to AS1 and end-users' completion times, respectively, when an IoP is inserted in pure BitTorrent, for different values of capacity *c* assigned to the IoP. Recall that the number of unchokes of the IoP is equal to 20 and that its download and upload capacities are considered to be symmetric. In particular, in Fig. 7 traffic curves for $c = 10, 20, 30, 40$ x512 kbps are depicted. We observe that the traffic that enters the AS1 is generally decreasing when *c* increases. However, for $c = 40$ (cyan line), we see that the traffic is slightly higher than that for $c = 30$. This tradeoff is due to the fact that the IoP downloads more content from external peers before it becomes a seed. Furthermore, in Fig. 8 (top curve), the completion-time curves for $c = 10, 20, 30, 40$ and in Fig. 8 (bottom curve), the relative difference (%) of the completion times achieved for $c = 10, 30, 40$ compared to the times achieved for $c = 20$ are depicted. We can observe that, for $c = 10$, 5% worse times are achieved, whereas for $c = 20, 30, 40$, the completion times are similar. To summarize, providing more resources to the IoP is beneficial only up to a certain point.

6 Conclusions

In this paper, we have proposed and investigated the insertion of ISP-owned peers both in pure BitTorrent networks and in BitTorrent networks where locality awareness is also employed. The objective is to achieve both reduction of the inter-domain traffic caused by BitTorrent and reduction of downloading-completion times. Furthermore, we have conducted simulations for several scenarios in order to evaluate the performance implications of the IoP insertion and presented related results.

Simulations have shown that the insertion of the IoP achieves significant reduction of the inter-domain traffic that enters the AS where it is deployed. Further improvements are achieved when the IoP insertion is combined with locality-aware mechanisms. Moreover, the insertion of IoP in a pure BitTorrent network leads to higher inter-domain traffic reduction than just the use of locality awareness. On the other hand, the insertion of IoP achieves reduction of end-users' completion times in all cases that have been studied, whereas sole locality awareness implies slight performance degradation for end-users, as it was shown in [10] and was also observed in our simulation experiments. Furthermore, the symmetric reduction of inter-domain traffic achieved by locality awareness has no impact on interconnection costs when charging models based the difference of inbound-outbound traffic. On the contrary, the IoP achieves important asymmetric traffic reduction, which is expected to have also important impact on interconnection costs. Additionally, even when only inbound traffic is taken into account by the charging scheme, the IoP achieves further improvement and cost reduction than locality awareness.

The idea of the IoP insertion is related to the insertion of caches by the ISP, which store the content that is downloaded by peers, as considered in [5], [10]. However, the difference is that the solution of caches should be combined with interception of peers' messages whereas the IoP is part of the overlay itself. That is, it runs the overlay protocol, without requiring any enforcement. Therefore, communication between regular peers and the IoP is optional rather than being not enforced either at the application level or by means of special hardware. In this sense, the insertion of the IoP is an innovative idea. Of course, similarly to the case of caches, the IoP should only deal with legal content.

Furthermore, the insertion of IoP could be combined with bilateral agreements between the ISP and the overlay provider, or the ISP and the content provider (see also end of Section 3), which are also in line with the aforementioned legal issue. For instance, the overlay provider could favor the IoP when replying to peers' requests, e.g. by means of an IoP-aware overlay tracker. On the other hand, if the ISP has established some kind of agreement with a content provider, then its content can be stored directly in the IoPs and the torrent file generated would immediately contain as meta-info the IP addresses of the respective IoP. Essentially, the IoP acts as a seed, rather than as a cache that intercepts the requests. These kinds of agreements and related business models are under investigation.

Last, in this paper, we have restricted attention to the insertion of IoP in a BitTorrent file-sharing overlay. Investigation of the applicability of the IoP to the optimization of BitTorrent-like real-time or streaming applications is also currently in progress.

Acknowledgments. This work has been accomplished in the framework of the EU ICT Project SmoothIT (FP7-2007-ICT-216259). The authors would like to thank all SmoothIT partners for useful discussions on the subject of the paper.

References

1. Simple Economic Management Approaches of Overlay Traffic in Heterogeneous Internet Topologies, European Seventh Framework Project FP7-2008-ICT-216259-STREP, http://www.smoothit.org/
2. Cohen, B.: Incentives build robustness in BitTorrent. In: Proceedings of the First Workshop on the Economics of Peer-to-Peer Systems, Berkeley, CA, USA (June 2003)
3. Zhang, H., Liu, T., Gong, W., Towsley, D.: Understanding the Interaction Between Overlay Routing and Traffic Engineering (2004)
4. Liu, Y., Zhang, H., Gong, W., Towsley, D.: On the Interaction Between Overlay and Underlay Routing. In: Proc. IEEE INFOCOM (2005)
5. Karagiannis, T., Rodriguez, P., Papagiannaki, K.: Should Internet Service Providers Fear Peer-Assisted Content Distribution? In: Proceedings of the Internet Measurement Conference 2005 (2005)
6. Ratnasamy, S., Handley, M., Karp, R., Shenker, S.: Topologically-Aware Overlay Construction and Server Selection. In: INFOCOM (2002)
7. Le, H., Hoang, D., Simmonds, A.: A Self-Organizing Model for Topology-Aware Overlay Formation, Seoul (2005)
8. Jia, J., Chen, C.: GMT: A Geometry-based Model on BitTorrent-like Topology. In: 8th International Conference on Signal Processing (2006)
9. Choffnes, D.R., Bustamante, F.E.: Taming the Torrent: A Practical Approach to Reducing Cross-ISP Traffic. In: Peer-to-Peer Systems, SIGCOMM (2008)
10. Bindal, R., Cao, P., Chan, W., Medved, J., Suwala, G., Bates, T., Zhang, A.: Improving Traffic Locality in BitTorrent via Biased Neighbor Selection. In: 26th IEEE International Conference on Distributed Computing Systems, p. 66 (2006)
11. Aggrawal, V., Feldmann, A., Scheideler, C.: Can ISPs and P2P Users Cooperate for Improved Performance? In: SIGCOMM 2007 (2007)
12. Wu, C.J., Li, C.Y., Ho, J.M.: Improving the Download Time of BitTorrent-Like Systems. In: IEEE International Conference on Communications, pp. 1125–1129 (2007)
13. Yamazaki, S., Tode, H., Murakami, K.: CAT: A Cost-Aware BitTorrent. In: 32nd IEEE Conference on Local Computer Networks, pp. 226–227 (2007)
14. Li, W., Chen, S., Yu, T.: UTAPS: An Underlying Topology-Aware Peer Selection Algorithm in BitTorrent. In: 22nd International Conference on Advanced Information Networking and Applications, pp. 539–545 (2008)
15. Oechsner, S., Soursos, S., Papafili, I., Hossfeld, T., Stamoulis, G.D., Stiller, B., Callejo, M.A., Staelhe, D.: A framework of economic traffic management employing self-organization overlay mechanisms. In: International Workshop on Self-Organizing Systems, Vienna (December 2008)
16. The Network Simulator – ns-2: http://www.isi.edu/nsnam/ns/
17. Eger, K.: Simulation of BitTorrent Peer-to-Peer (P2P) Networks in ns-2, http://www.tu-harburg.de/et6/research/bittorrentsim/

Author Index

Altmann, Jörn 85

Biczók, Gergely 48
Bouhtou, Mustapha 24
Brenner, Walter 61

Davoyan, Ruzana 85

Effelsberg, Wolfgang 85

Gyarmati, László 11

Hau, Thorsten 61
Hausheer, David 36
Huitema, George B. 72

Le Cadre, Hélène 24
Li, Li 36
Liu, Yuanan 36

Maillé, Patrick 2

Nieuwenhuis, L.J.M. (Bart) 72

Papafili, Ioanna 97

Roberts, Jim 1
Rumph, Frens Jan 72

Soursos, Sergios 97
Stamoulis, George D. 97
Stiller, Burkhard 36

Trinh, Tuan Anh 11, 48
Tuffin, Bruno 2, 24

van Beijnum, Bert-Jan 72
van Le, Minh 72