Economic Incentives for Adopting Congestion Accountability Protocols

Alexandros Kostopoulos, Antonis Dimakis

Athens University of Economics and Business
Network Economics and Services Research Group
{alexkosto, dimakis}@aueb.gr

Costas Courcoubetis

Singapore University of Technology and Design
costas@sutd.edu.sg

Abstract— New Internet protocols such as ECN, re-ECN and Conex, provide valuable information to ISPs about the congestion within a network. Such information could be useful in allocating network resources more equitably, as well as to employ new pricing schemes. However, in designing a new protocol one must take proper account of the incentive issues that thereby arise for its adoption. In this paper, we consider a duopoly market consisting of two access ISPs and formulate a noncooperative game over their choices of pricing strategies when one of them uses volume-based and the other congestion-based pricing. We prove, under mild conditions, the existence of a pure strategy Nash equilibrium for the competing providers. We further examine how the different types of transit charges affect the pricing strategies of the access ISPs. We find cases where the volume charging ISP is forced to leave the access market. We conclude that ISPs have economic incentives to adopt congestion accountability mechanisms, since they become more competitive due to the smoother and more predictable traffic they have to accommodate and to the lower transit charges they incur.

Keywords—multihoming; competition; pricing; game theory; accountability; congestion exposure; ECN marks; protocol adoption.

I. Introduction

The success of new web applications, the wide use of cloud or video-on-demand services, as well as the increasing number of mobile devices are only some examples revealing the significant growth of Internet traffic and congestion. However, the existing pricing policies in the current Internet do not make end users conscious of the congestion they impose upon each other when consuming bandwidth during peak demand periods.

In response, mechanisms for congestion accountability, such as ECN [1], re-ECN [2] and Conex [3] have been proposed to provide additional information about the level of congestion within ISPs' networks. Such mechanisms intend to promote cooperative traffic management and allocate efficiently the network resources, without necessarily restricting the rate of the sender.

Congestion accountability mechanisms allow new Service Level Agreements (SLAs) to be introduced. For example, instead of monitoring the traffic volume sent by an end-user, ISPs are able to replace volume allowances by congestion allowances for their customers [4]. Such SLAs are compatible with the economic principles, as they take into account the negative network externalities (in terms of congestion) a user is causing [5]. Additionally, they provide incentives for an end user to implement a rate adaptation mechanism that regulates traffic flows (including sending different traffic streams at differentiated rates if necessary) so that the congestion caused remains within their allowance.

It is natural to ask the following questions: Do ISPs have economic incentives to adopt congestion accountability protocols? Can an ISP benefit from employing congestion pricing instead of volume charging? What is the impact of congestion charging in attracting traffic within a competitive context? How do the different types of transit charges affect the pricing strategies of access ISPs?

Our paper is motivated by the need to consider the economic incentives of ISPs to adopt congestion accountability mechanisms when competing with other ISPs that use traditional tariffs. We define a model assuming two access ISPs who compete to attract traffic from multihomed users[1], employing different pricing schemes. The former ISP uses volume-based tariffs for charging its users, while the latter ISP has deployed congestion accountability mechanisms to employ dynamic congestion-based pricing. Based on the applied charges, multihomed users are able to choose which access ISP out of the two will serve their traffic. Each access ISP aims at maximizing his profits by selecting the optimal price for charging users, while users intend to minimize their cost by selecting the lowest price for sending traffic. Assuming that access ISPs forward the received traffic to a monopolistic higher-Tier transit ISP, we investigate two scenarios where the transit ISP charges both access ISPs either based on volume, or based on the 95th percentile rule. Furthermore, we extend our model by considering more layers of ISPs charging similarly.

Our key contributions can be summarized as follows:

- We define a duopoly market consisting of two access ISPs pricing according to volume and according to congestion, and we derive their expected traffic and profits. To the best of our knowledge, this is a first attempt to model the competitive advantages of congestion pricing in the ISP market.

---

[1] If we relax the multihoming assumption, then the same results hold when users respond to pricing in longer timescales.

- We formulate a noncooperative game over ISPs' choices of pricing strategies when one of them uses volume-based and the other congestion-based pricing. We analytically prove the existence of a pure strategy Nash equilibrium for the game consisting of setting prices in this competitive context.

- We discuss how the different types of transit charges, either based on volume or based on the 95th percentile rule, affect the pricing strategies of access ISPs. We consider the case in which there is a monopolistic transit ISP, as well as there are two competing higher-Tier transit ISPs. We prove, under mild conditions, that the congestion charging ISP has higher profits than the volume charging ISP. Another interesting result is that there are cases where the volume charging ISP has negative profits and is forced to leave the market, due to the bursty traffic he attracts, implying higher transit charges. The results are similar if there are more layers of ISPs charging similarly. The aforementioned remarks reveal that ISPs do have economic incentives to adopt congestion accountability mechanisms.

The reader may wonder how congestion pricing can be implemented in practice. There are currently some very interesting proposals on how to achieve this and make the real-time congestion charges visible to the traffic sources in order for them to take the appropriate actions. The *ECN* protocol [1] allows routers to mark packets that would otherwise have been dropped as having experienced congestion. The proportion of marked packets can be used by the network as a metric for charging end-users for their upstream congestion, i.e., charging based on the number of ECN marks. *Re-feedback of Explicit Congestion Notification* (Re-ECN) [2] is designed to provide the network with information about the expected level of congestion along the entire path. A *policer* within the re-ECN framework could be used for making possible flat rate, congestion allowance contracts [4], [6]. *Congestion Exposure* (Conex) [3] stands for an extension of the re-ECN, which removes the dependency on ECN by detecting the packet losses [7]-[9]. Whereas ECN provides information about upstream congestion (i.e. the congestion a traffic flow has already experienced), re-ECN and Conex additionally provide the level of downstream congestion.

The rest of the paper is organized as follows. In Section 2 we provide the related work in competition and Internet pricing models. Sections 3 and 4 describe our model, and present the game theoretic analysis for the defined scenarios. Finally, we conclude our remarks and outline the future work in Section 5.

II. RELATED WORK

Multihoming refers to the connection of a user to more than one ISP, as a form of reliability and redundancy. The usual case is being connected to a specific ISP used as the primary provider, while another is used as backup solution. Multihoming can be viewed as an environment for access ISPs that compete to attract traffic, due to the freedom of users to

prefer an ISP. Such examples may include cases in which a mobile user chooses to connect to a specific wireless access ISP, or between different transmission platforms (i.e. WiFi, 3G/4G). There has been significant research on evaluating and realizing the benefits of multihoming [10]. Goldenberg *et al.* [11] investigate how to distribute traffic among multiple links to optimize both cost and performance. Authors in [12] investigate how to solve the ISP subscription problem optimally, as well as how ISPs adjust their pricing strategies based on users' optimal ISP subscription. Based on cost criterion, a user determines which subset of ISPs will subscribe to. It is also studied the case where ISPs provide heterogeneous services based on different levels of reliability. Contrarily, our model assumes that ISPs provide homogenous connectivity services that are perfect substitutes to the end-users. We further assume that end-users are multihomed in order to define a competitive access market. Similarly to [13], total user demand is split among providers depending on price, according to Wardrop's principle [14]. A large proportion of papers applies Game Theory to inquire the strategic interaction among competing players. In this paper, we formulate a noncooperative game over ISPs' pricing strategies. For further examples, the reader is advised to look at [15]-[21].

There has been increasing interest in Internet pricing models. Authors in [22] study the optimal combination of flat-rate and usage-based access price components for maximization of ISP's revenue. They indicate that flat pricing can lead to a significant loss of consumer net-utility, due to the inability of the ISP to adapt his charges. Based on this remark, we do not include flat-rate pricing in our model. Instead, we compare volume charging with dynamic congestion-based charging. Volume-based charging is a commonly used approach adopted by most of mobile operators. On the other hand, congestion pricing has been proposed as a solution resulting in more efficient resource allocation [23]. The selfish behavior of end-users may result in the famous "tragedy of the commons", where shared resources are overconsumed by individuals without considering the cost to society as a whole [24]. In this spirit, several models have been proposed. MacKie-Mason and Varian [5] consider the implications of flat and congestion-based pricing as a way to encourage efficient use of network resources. Moreover, they discuss the incentives of ISPs for capacity expansion planning in competitive and monopolistic environments. Similarly, Gibbens and Kelly [25] propose charging a fixed small amount for each congestion mark at overloaded resources, in order to provide the necessary information to use the network efficiently. In [26], an ISP charges a fee per service request depending on the current congestion level which affects users' demand and determines an optimal pricing strategy, intending to maximize his revenue. The use of congestion marks for charging users can be viewed as a provision of service differentiation. References [27] and [28] study the incentives for end-users to select the appropriate differentiated class of service at a queuing system.

Most of the research efforts in literature, such as [29] focus on revenue-maximizing pricing schemes for the ISPs. However, it is important to consider the cost that ISPs have to

deal with, due to the transit charges by higher-Tier network providers. For example [30] examines how transit and customer prices, as well as QoS are set in a network consisting of multiple ISPs. A common approach for transit ISPs is to charge the access ISPs based on the traffic volume. On the other hand, the 95th percentile method is an increasingly popular pricing policy used by network providers for billing lower-Tier ISPs [31], intending to avoid traffic bursts. Authors in [32] study top-percentile pricing schemes by considering a multihomed ISP seeking an efficient routing strategy to reduce his costs, while in our paper the ISP intends to solve a profit-maximization problem. Similarly to [32], we propose a probabilistic model to reflect the stochastic nature of traffic demand.

Our paper differs from the above papers due to the fact that our goal is to devise a model comparing volume with congestion-based pricing schemes within a competitive context, while taking also into account the different types of transit charges. Our work provides useful insights for considering the adoption of congestion accountability mechanisms by the ISPs.

III. PROBLEM FORMULATION

Consider a large number of small multihomed users who can send their traffic towards a destination through two different paths. Each path corresponds to a different access ISP. We assume that time is discrete and that during each epoch the total demand for traffic generated by all the users is the random variable D, where $D \sim U[0, d]$. We use the uniform distribution to indicate the key trends in the equilibria of these games.

In each epoch the total demand D will be split among the two ISPs according to the resulting prices. We model this as a Wardrop [14] type of equilibrium where D will self-adjust its fraction that will be routed through each network so that the price/bit sent through each network will be as low as possible. Note that in our model the price of ISP1 is fixed over a given epoch, while the price of ISP2 is a function of the amount x of traffic through its network. We assume that all the traffic sent will be charged at the equilibrium price. Figure 1 depicts the basic idea our model.
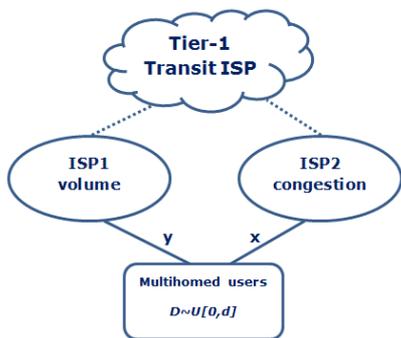

Fig. 1. Competition among access ISPs with a monopolistic transit ISP.

ISP1 applies linear charges based on the traffic volume and sets a constant price/bit p_1. The charging of ISP2 depends on the carried volume per epoch. In particular, ISP2 uses ECN marks and sets a price p_2/congestion mark. We assume that the number of congestion marks created per change epochs to periods consistently based on volume x is an increasing linear function of x. Although an exponential assumption may be also applied, the linearity assumption is simple and sensible, since it captures the notion of setting higher price/bit, as more traffic is sent through the network (which induces higher network congestion).

Coming back to our Wardrop model, during each time slot the generated demand D will be divided to infinitely many small pieces and each traffic piece will start routing its bits through the cheapest of the two networks. As traffic through ISP2 grows, the number of congestion marks and so its price/bit will increase, and might become eventually less competitive than ISP1. At the equilibrium the traffic will be split so that there is no incentive for a piece of traffic to change its selected route. This is of course an abstract model for load balancing between ISPs that assumes that users have price-sensitive multihoming capabilities, and they can adjust their routing in fast time scales reacting to the dynamic prices of ISP2. In this timescale decomposition, we assume that traffic is routed according to the Nash equilibrium that corresponds to the demand of each slot. The assumption on price-sensitive technologies at the user-end is consistent with the current research trends as in [7] and [33].

Figure 2 depicts a plot, where the x-axis represents users' total traffic and the y-axis represents the price/bit charged by both ISPs. According to our discussion about traffic routing in a way to prefer the cheapest network, if $D \leq h = p_1/p_2$, then all traffic will flow through ISP2 and $y = 0$. Else, we will have that $x = h$ and $y = D–h > 0$.
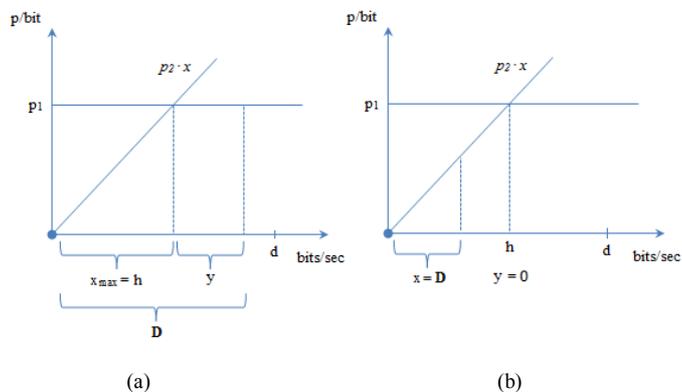

Fig. 2. How traffic is split given the pricing strategies of ISPs: (a) The total traffic is split among both ISPs when $D > h$, (b) The total traffic is shifted only through ISP2 when $D \leq h$.

We assume that both ISPs route the received traffic to a monopolistic higher-Tier transit ISP. The transit cost is considered to be the only cost for both ISPs to deal with. Furthermore, we extend our model by considering more layers of ISPs charging similarly. The different scenarios for transit charges are analytically presented in the next Section.

We formulate the interaction between the two ISPs as a best response non-zero sum game in which each ISP chooses

his optimal pricing strategy to maximize his profits. We further assume that there is a reservation price r, which is the maximum price/bit that a user is willing to pay. This prevents ISPs to charge arbitrarily high prices and not affecting the demand of their customers. Is this assumption realistic? Let us think the case in which an end-user needs to receive his e-mails or get access to a specific mobile application. Each task requires a fixed amount of bandwidth (i.e. a transaction via an e-banking application). If the price/bit set by the access ISP is higher than user's willingness to pay, he will choose not to connect to this network. Hence, this simple reservation price model has the meaning that if prices get higher than r, then no customer wants to send data or that there is a third ISP which offers r as a competitive price to both ISP1 and ISP2. We set $r = 1$, as well as $d = 1$.

IV. Game Theoretic Analysis

We investigate two pricing schemes that the monopolistic transit ISP employs. In the first one, we assume that the transit ISP charges based on the average traffic volume received from each access ISP. The second more interesting scenario assumes that the transit ISP charges based on the 95th percentile rule, which is a widely used approach to estimate the usage level of resources for long timescales (i.e. a month) [31]. This charging method penalizes traffic which has the same total volume but it is more bursty (variable). The intuitive reason is that more variable traffic is harder to multiplex and hence needs more resources. Moreover, we extend our model by considering a more complex scenario, in which there are more layers of ISPs charging similarly.

Our strategy space is for ISP1 the price/bit p_1 and for ISP2 the value $h = p_1/p_2{}^2$. The value h corresponds to ISP2 choosing the maximum level of traffic he would like to attract before traffic starts moving into the network of ISP1. It should be noted here that ISP2 has no incentive to set $h > d$, since it results in a lower price/bit to charge users, without attracting more traffic. So, we do not consider this case in our analysis.

A. Received Traffic and Revenues

The total revenues of ISP1 are given by the price/bit multiplied with the expected received traffic. The expected traffic received by ISP1 will be

$$E(y) = E(y \mid D > h) P(D > h) + E(y \mid D \le h) P(D \le h)$$
$$= \frac{d-h}{2} \frac{d-h}{d} + 0 \frac{h}{d} = \frac{(d-h)^2}{2d}.$$

In the above equation, the term $(d-h)/2$ represents the expected value of y, when $h < D \le d$. The term $(d-h)/d$ represents the probability of $h \le D \le d$. When $D \le h$, $y = 0$.

The expected revenues of ISP1 are given by

---

² We choose h to be the strategy of ISP2 instead of his congestion price p_2 since this simplifies considerably our analysis and allows us to get closed form expressions for the reaction curves.

$$R_1(p_1) = p_1 E(y) = p_1 \frac{(d-h)^2}{2d}.$$

We observe that the expected traffic and the revenues of ISP1 are linear functions with respect to p_1, which means that as p_1 increases, his expected revenues also increase. This is due to the fact that by our chose of strategies, when ISP2 chooses h, then all remaining traffic is routed to ISP1 irrespectively of his price p_1.

The expected received traffic by ISP2 will be

$$E(x) = h \frac{d-h}{d} + \int_0^h x f(x) dx \frac{h}{d}.$$

The expected revenues of ISP2 will be

$$R_2(h) = p_2 h^2 \frac{d-h}{d} + p_2 \int_0^h x^2 f(x) dx \frac{h}{d} = p_1 h - \frac{2}{3} \frac{p_1 h^2}{d}.$$

B. Scenario 1: Volume Based Pricing by the Monopolistic Transit ISP

1) Transit Cost
We assume that both access ISPs route the received traffic to one higher-Tier transit ISP. The monopolistic transit ISP charges based on the traffic volume received by the two access ISPs.

The expected cost is equal to the expected traffic of ISP1 multiplied by the price/bit charged by the transit ISP:

$$C_1(h) = \alpha E(y) = \alpha \frac{(d-h)^2}{2d},$$

where α is the price/bit charged by the transit ISP.

The expected transit cost of ISP2 will be

$$C_2(h) = \alpha E(x) = \alpha \left(h - \frac{h^2}{2d} \right).$$

2) Profits and Best Response
Since we assume that the transit cost is the only cost that ISPs have to deal with, their profits are calculated by the expected revenues minus the expected transit cost:

$$\pi_1(p_1) = R_1(p_1) - C_1(h) = (p_1 - \alpha) \frac{(d-h)^2}{2d}$$

$$\pi_2(h) = R_2(h) - C_2(h) = p_1 h - \frac{2}{3} \frac{p_1 h^2}{d} - \alpha h + \frac{\alpha h^2}{2d}.$$

Assumption A: The profits of both access ISPs are positive, in order to play the game. This is valid when $\alpha \le r$.

In order to find the pure strategy in the competitive game between the two access ISPs, we compute the derivatives of their profit functions.

Proposition 1: Given that the transit ISP charges α per traffic unit transferred, under Assumption A, there exists a

pure strategy of the two access ISPs which is $p_1 = 1$ and

$$h = \frac{3d(p_1 - \alpha)}{4p_1 - 3\alpha}.$$

Proof: The profit function of ISP1 is linear increasing with respect to p_1. However, since there is a reservation price for users, the best response is to set $p_1(h) = r = 1$.

The second derivative of the profit function of ISP2 with respect to h is $\partial^2 \pi_2 / \partial h = -4p_1 / 3d + \alpha/d$. Since $p_1 > \alpha$ we have $\partial^2 \pi_2 / \partial h < 0,$ for all $h \in [0, d]$. Solving the equation $\partial \pi_2 / \partial h = 0$, we find out that the best response strategy of ISP2 is to set $h(p_1) = \frac{3d(p_1 - \alpha)}{4p_1 - 3\alpha}.$ ∎

Remark: Figure 3 shows the optimal h^* chosen by ISP2 depending on α, assuming that ISP1 chooses his optimal strategy. Since $p_1^* = 1$ and $d = 1$, we only need to consider $h(1)$. We observe that the best h is decreasing as long as α is increasing. In particular, if $\alpha = 1$, then $h^* = 0$. Otherwise, when $\alpha = 0$, then $h^* = 0.75d$. Hence, we observe that ISP2 will choose an h^* such that $0 \le h^* \le 0.75d$.
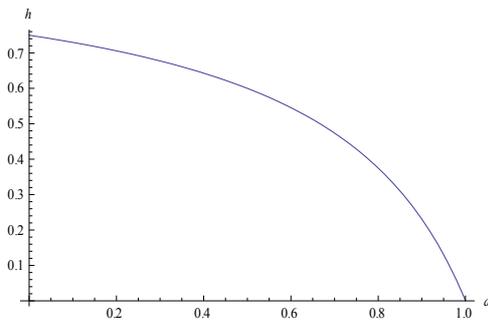

Fig. 3. The optimal h^* chosen by ISP2 under volume-based transit cost when ISP1 sets $p_1^* = 1$, $d = 1$, as a function of the cost parameter α.

The following proposition shows the relation of the profits of the two ISPs for small and for large values of the transit price α. ISP1 is more profitable than ISP2 only if this price is very large. For large values of α, ISP2 chooses a very small h, which means that ISP1 attracts most of the traffic.

Proposition 2: Given the pricing strategies of both access ISPs at the equilibrium under volume-based transit cost:

- $\pi_2(h^*) \ge \pi_1(p_1^*)$, when $\alpha \le (7 - \sqrt{5})/6$

- $\pi_2(h^*) < \pi_1(p_1^*)$, when $\alpha > (7 - \sqrt{5})/6$.

Proof: We define the function $G(\alpha) = \pi_2(h^*) - \pi_1(p_1^*)$.

We investigate the sign of $G(\alpha)$. Due to Assumption A, the only accepted root for $G(\alpha) = 0$ is $\alpha = (7 - \sqrt{5})/6$. So,

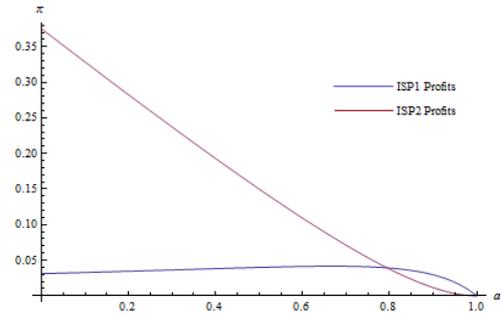$G(\alpha) \ge 0$ when $\alpha \le (7 - \sqrt{5})/6$. ∎

Fig. 4. Profits of the competing ISPs under volume-based transit cost at the equilibrium. ISP1 is more profitable than ISP2 only if the transit price α is very large, where the average profit/bit $(p_1 - \alpha)$ of both ISPs is very low.

Figure 4 represents the profits of both ISPs at the equilibrium, depending on α. Both ISPs have positive profits for any value of α. We also notice that the profits of ISP2 are higher than those of ISP1, except for the case where $\alpha \in ((7 - \sqrt{5})/6, 1]$. When α is high, ISP1 has quite higher profits than ISP2. ISP2 will choose to attract lower traffic, since the optimal h^* selected by ISP2 is decreasing as long as α is increasing. As the value of α tends to 1, the average profit per bit $(p_1 - \alpha)$ of both ISPs is very low, and their expected profits tend to zero.

C. Scenario 2: Pricing Based on the 95th Percentile Rule by the Monopolistic Transit ISP

1) Transit Cost

We now turn to investigate what happens when access ISPs are charged by the transit ISP based on the 95th percentile rule. Who wins? Do both ISPs always have positive profits?

In case of charging based on the 95th percentile rule, the bandwidth used by a lower-tier ISP during a specific period (i.e. 10 minutes) is estimated and divided by the time period (in seconds) resulting in a single bps transfer measurement. At the end of the billing cycle period, all measurements are sorted in decreasing order and the top 5% of these measurements are thrown out. The next highest measurement is the 95th%, and the customer will be billed based on that volume.

We denote l_1^* as the 95th% measurement of the traffic received by the ISP1. Then l_1^* is defined as follows:

$$l_1^* = \max\{l_1 \mid P(y > l_1) = 0.05\}.$$

Based on the aforementioned definition of the 95th percentile rule, we can make an interesting observation. If an ISP very rarely attracts traffic (more than 95% of the traffic measurements is zero), he will have zero transit cost. So, we seek the probability of an access ISP not being charged by the transit ISP.

The condition where ISP1 will attract bursty traffic for which he will not be charged when $l_1^* = 0$ is:

$$P(h < D \le d) = 1 - \frac{h}{d} < 0.05 \Leftrightarrow h > 0.95d.$$

When $h \leq 0.95d$, ISP1 will be charged for the traffic attracted. The 95th% measurement l_1^* is calculated as follows:

$$P\left(D - h > l_1^*\right) = \frac{d - \left(h + l_1^*\right)}{d} = 0.05 \Leftrightarrow l_1^* = 0.95d - h.$$

The transit cost of ISP1 is given if we multiply the level of traffic l_1^* by the price/bit α charged by the transit ISP. Hence, the expected transit cost of ISP1 will be the following:

$$C_1(h) = \begin{cases} 0.95d\alpha - h\alpha, & h \leq 0.95d \\ 0, & h > 0.95d. \end{cases}$$

Similarly, the traffic of ISP2 for which he will be charged by the transit ISP based on the 95th% measurement, is defined as follows:

$$l_2^* = \max\left\{l_2 \mid P(x > l_2) \geq 0.05\right\}.$$

The 95th% measurement of ISP2 is equal to h, when $P(D > h) = \frac{d - h}{d} \geq 0.05$. This implies that $l_2^* = h$, when $h \leq 0.95d$. Otherwise, $l_2^* = 0.95d$, when $h > 0.95d$.

The expected transit cost of ISP2 will be

$$C_2(h) = \begin{cases} ha, & h \leq 0.95d \\ 0.95da, & h > 0.95d. \end{cases}$$

It should be noted here that in contrast to our previous scenario where the volume transit cost was α, in this case the access ISPs have different transit costs.

2) Profits and Best Response
The expected profits of the access ISPs, under the 95th percentile rule-based transit cost, will be

$$\pi_1(p_1) = \begin{cases} p_1 \dfrac{(h - d)^2}{2d} - 0.95d\alpha + \alpha h, & h \leq 0.95d \\ p_1 \dfrac{(h - d)^2}{2d}, & h > 0.95d. \end{cases}$$

$$\pi_2(h) = \begin{cases} p_1 h - \dfrac{2}{3} \dfrac{p_1 h^2}{d} - \alpha h, & h \leq 0.95d \\ p_1 h - \dfrac{2}{3} \dfrac{p_1 h^2}{d} - 0.95d\alpha, & h > 0.95d. \end{cases}$$

Assumption B: The profits of ISP1 are positive, in order to play the game. This is valid when

$$a \leq \frac{(h - d)^2}{2d(0.95d - h)} r.$$

Proposition 3: Given that the transit ISP charges α for the 95th% traffic measurement, under Assumption B, there exists a pure strategy of the two access ISPs which is $p_1 = 1$ and

$$h = \frac{3d(p_1 - \alpha)}{4p_1}.$$

Proof: Similarly to the case of fixed transit volume cost, the profit function of ISP1 is linear increasing with respect to p_1. However, since there is a reservation price for users, the best response is to set $p_1(h) = r = 1$.

We compute the derivative of the profit function of ISP2 to find the pricing strategy that maximizes his profits, as a function of p_1 set by ISP1. When $h > 0.95d$, the second derivative of the profit function of ISP2 with respect to h is $\partial^2 \pi_2 / \partial h = -4p_1/3d$, which is negative. Solving the equation $\partial \pi_2 / \partial h = 0$, we find out that the profit function is maximized when $h = 0.75d \notin (0.95d, d]$. Solving $\partial \pi_2 / \partial h = 0$ when $h \leq 0.95d$, we find out that the best response strategy of ISP2 is $h(p_1) = \dfrac{3d(p_1 - \alpha)}{4p_1}$. ∎

Figure 5 shows the optimal h^* chosen by ISP2 depending on α, assuming that ISP1 chooses his optimal strategy $p_1^* = 1$. The optimal h^* is linearly decreasing as long as α is increasing. Similarly to the Scenario 1, ISP2 will choose an h^* such that $0 \leq h^* \leq 0.75d$.
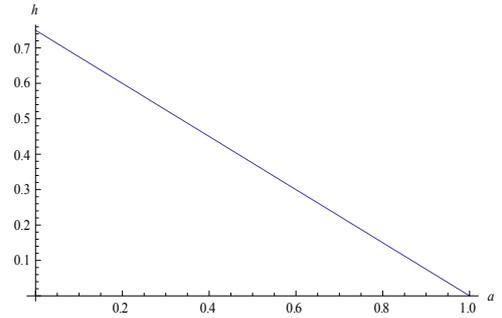

Fig. 5. The optimal h^* chosen by ISP2 under 95th percentile rule-based transit cost when ISP1 sets $p_1^* = 1$, $d = 1$, as a function of the cost parameter α.

The following proposition shows the relation of the profits of the two ISPs, depending on the transit cost parameter α.

Proposition 4: Given the pricing strategies of both access ISPs at equilibrium under 95th percentile rule-based transit cost, $\pi_1(p_1^*) < \pi_2(h^*)$, for all $a \in [0,1]$.

Proof: $G(\alpha) = \pi_2(h^*) - \pi_1(p_1^*)$ is a quadratic function of α that has a negative discriminant, and hence it has no real roots and is positive for all $a \in [0,1]$. ∎

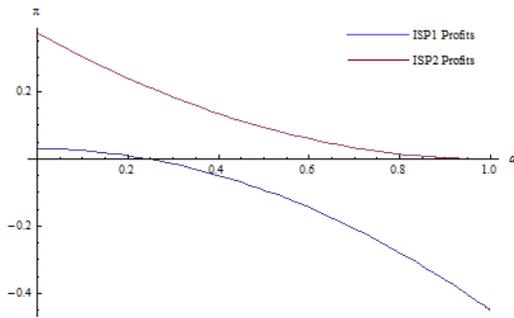Figure 6 represents the profits of both ISPs depending on α, at the equilibrium.

Fig. 6. Profits of the competing ISPs under 95th percentile rule-based transit cost at the equilibrium. The profits of ISP2 clearly dominate these of ISP1.

We proved that the profits of ISP2 are always higher than those of ISP1 at the equilibrium. Furthermore, it is obvious to notice that the profit of ISP1 under the 95th percentile rule-based transit cost is always less than the case where the transit ISP charges based on volume. This seems reasonable, since in the second scenario ISP1 attracts sporadic and bursty traffic, implying higher transit charges. An interesting observation is that the expected profits of ISP1 become negative as α increases, since he cannot set a price p_1 higher than the reservation price r. ISP1 participates in the market when he has positive profits. This is not possible when $2da\left(0.95d-h^*\right)/\left(h^*-d\right)^2 \geq r$. We can compute the region of parameters that this occurs. Upon substituting $h^* = 3d\left(r-\alpha\right)/4r$, $r = 1$ and $d = 1$, this gives $a \geq 0.2452$. In this case we do not have ISP1 participating in the market.

This does not affect the strategy of ISP2, even if ISP1 decides not to play the game. Due to the existence of the reservation price, ISP2 chooses his optimal h as a reaction to r.

D. Scenario 3: Competition Among Higher-Tier Transit ISPs

1) Transit Cost

We extend our model by considering a more complex scenario. In particular, we assume there are two competing higher-Tier transit ISPs, say L (Left) and R (Right), instead of a monopolistic transit ISP, as well as the existence of the reservation price $r = 1$ at all levels. The transit ISPs L and R send their received traffic to a Tier-1 transit ISP. ISP_L charges based on volume, while ISP_R based on congestion. Repeating the previous analysis for the case of these two Tier-2 ISPs receiving in a competitive way the traffic from the Tier-3 (access) ISPs, ISP_L will choose an optimal $\overline{p} = 1$, and ISP_R an optimal $\overline{h}$, depending on the price/bit α and also on how the Tier-1 transit ISP charges (either based on volume or based on the 95th percentile rule). The total traffic generated by the lower-Tier ISPs is equal to the total traffic generated by the actual customers, i.e., equal to D. Thus, the expected revenues of both Tier-3 ISPs remain the same as in the previous scenarios, but they have different transit costs. Figure 7 illustrates this scenario.
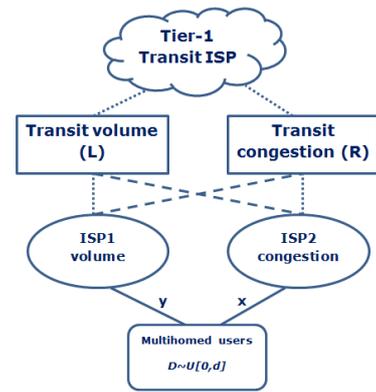

Fig. 7. Competition among access ISPs with competing higher-Tier ISPs charging similarly

We focus on an indicative scenario, assuming that the reservation price exists at all levels and that $p_1 = \overline{p} = r$. Given the optimal strategies of the Tier-2 ISPs and ISP1, we like now to derive the strategy of ISP2, as well as to compare the profits of both access ISPs.

Figure 8 depicts the price/bit charged by the Tier-2 ISPs, when ISP2 sets (a) $h \leq \overline{h}$, and (b) $h > \overline{h}$.
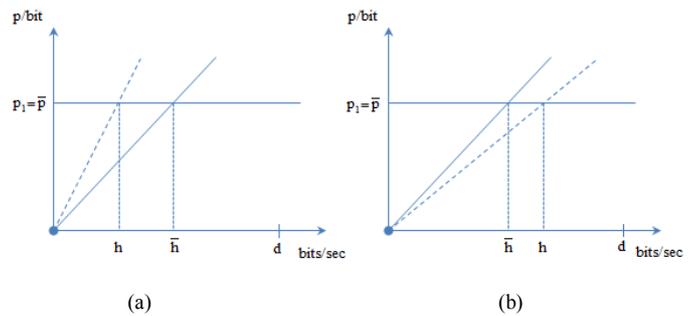

(a)   (b)

Fig. 8. The price/bit charged by ISP2, given the optimal strategies of the Tier-2 ISPs ($\overline{p}$ and $\overline{h}$ correspondingly), and $\overline{p} = p_1$: ISP2 sets (a) $h \leq \overline{h}$, and (b) $h > \overline{h}$.

The expected cost of ISP1, when $h \leq \overline{h}$ will be

$$
\begin{aligned}
C_1(h) &= E\left(C_1 \mid D < h\right)P\left(D < h\right) \\
&+ E\left(C_1 \mid h \leq D < \overline{h}\right)P\left(h \leq D < \overline{h}\right) \\
&+ E\left(C_1 \mid D \geq \overline{h}\right)P\left(D \geq \overline{h}\right) \\
&= 0 + \frac{1}{\overline{h}-h}\int_h^{\overline{h}}(x-h)x\frac{\overline{p}}{\overline{h}}dx\frac{\overline{h}-h}{d} \\
&+ \overline{p}\int_{\overline{h}}^d(x-h)f(x)dx\frac{d-\overline{h}}{d},
\end{aligned}
$$

else if $h > \overline{h}$, then

$$C_1(h) = \overline{p} \int_h^d (x-h) f(x) dx \frac{d-h}{d}.$$

The transit cost of ISP2, when $h \leq \overline{h}$ will be

$$C_2(h) = \frac{\overline{p}}{h} \int_0^h x^2 f(x) dx \frac{h}{d} + h \frac{\overline{p}}{h} \int_h^{\overline{h}} x f(x) dx \frac{\overline{h}-h}{d}$$
$$+ h \overline{p} \frac{d-\overline{h}}{d},$$

else if $h > \overline{h}$, then

$$C_2(h) = \frac{\overline{p}}{h} \int_0^{\overline{h}} x^2 f(x) dx \frac{\overline{h}}{d} + \overline{p} \int_{\overline{h}}^h x f(x) dx \frac{h-\overline{h}}{d} + h \overline{p} \frac{d-h}{d}.$$

2) Profits and Best Response
The expected profits of the access ISPs will be

$$\pi_1(h) = \begin{cases} \dfrac{\overline{p}}{2d}\left(h^2 + \dfrac{\overline{h}^2}{3} - h\overline{h} - \dfrac{h^3}{3\overline{h}}\right), & h \leq \overline{h} \\[2ex] 0, & h > \overline{h}. \end{cases}$$

$$\pi_2(h) = \begin{cases} \dfrac{\overline{p}}{6\overline{h}d} h^3 - \dfrac{2\overline{p}}{3d} h^2 + \dfrac{\overline{h}\overline{p}}{2d} h, & h \leq \overline{h} \\[2ex] -\dfrac{\overline{p}}{6d} h^2 + \dfrac{\overline{p}\overline{h}^2}{6d}, & h > \overline{h}. \end{cases}$$

Proposition 5: Given the optimal strategies of the Tier-2 ISPs and ISP1 ($\overline{p} = r = 1$, $\overline{h} = h(r)$ and $p_1 = r$), there exists an optimal strategy of ISP2, which is $h = \overline{h}(4 - \sqrt{7})/3$.

Proof: We compute the derivative of the profit function of ISP2 to find the pricing strategy that maximizes his profits, for a given $\overline{h}$ set by ISP$_R$, assuming $\overline{p} = p_1 = 1$. The second derivative of the profit function of ISP2 with respect to h when $h > \overline{h}$, is $\partial^2 \pi_2 / \partial h = (3\overline{p} - 4p_1)/3d$, which is negative. Solving $\partial \pi_2 / \partial h = 0$ we obtain $h^* = 0$. The second derivative of the profit function of ISP2 with respect to h when $h \leq \overline{h}$, is $\partial^2 \pi_2 / \partial h = \overline{p}h / \overline{h}d - 4p_1/3d$, which is negative. Solving $\partial \pi_2 / \partial h = 0$ we find out that the best response strategy of ISP2 is $h = \overline{h}(4 - \sqrt{7})/3$. ∎

The following proposition shows the relation of the profits of the two access ISPs, when $\overline{p} = p_1 = 1$ and $h = \overline{h}(4 - \sqrt{7})/3$, depending on the optimal strategy $\overline{h}$ set by ISP$_R$.

Proposition 6: At the Nash equilibrium we have $\pi_2(h^*) \geq \pi_1(p_1^*)$, for all $\overline{h} \in [0,1]$.

Proof: We define the function $G(\overline{h}) = \pi_2(h^*) - \pi_1(p_1^*)$. Upon substituting $h = \overline{h}(4 - \sqrt{7})/3$ and $\overline{p} = p_1 = 1$, this gives

$$G(\overline{h}) = \frac{1 + 2\sqrt{7}}{81} \overline{h}^2 \geq 0.$$ ∎

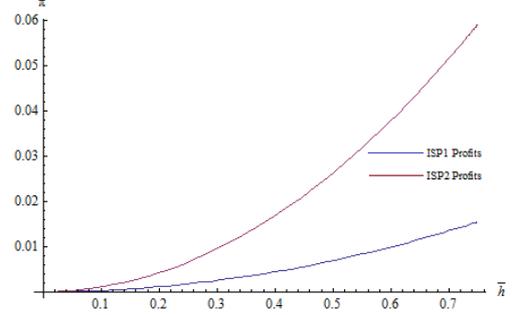Figure 9 shows the profits of both ISPs depending on $\overline{h}$, at the equilibrium.

Fig. 9. The profits of ISP2 clearly dominate these of ISP1 under competition among higher-Tier transit ISPs at the equilibrium.

V. CONCLUSIONS AND FUTURE WORK

Assuming a duopoly market consisting of two access ISPs pricing according to volume and according to congestion, we have characterized how demand will be split between those providers. We derived the expected traffic and profits of both ISPs, and we determined the existence and uniqueness of the Nash equilibrium of the pricing game we formulated. We found out that the volume charging ISP always chooses to charge with the reservation price, which represents end-users' maximum willingness to pay. On the other hand, the congestion charging ISP sets a price based on his trade-off between attracting more traffic and increasing transit costs. Furthermore, we realized that ISP2 will never choose to attract more than the 75% of the total demand.

We discussed how the different types of transit charges, based on volume and based on the 95th percentile rule, affect the pricing strategies of the access ISPs. When the transit ISP charges based on volume, then both ISPs have positive profits. We found out that the congestion charging ISP has almost always higher profits than the volume charging ISP, except for the case where the average profit per bit of both ISPs is very low. In the scenario where the transit ISP charges based on the 95th percentile rule, we conclude that the profits of the congestion charging ISP clearly dominate these of the volume charging ISP. Since the volume charging ISP attracts more bursty traffic, the 95th percentile rule leads to higher transit charges due to the bursty traffic he attracts. Similarly, the transit cost of the congestion charging ISP is significantly lower than competitor's due to the smoother and more predictable traffic he has to accommodate and to the lower transit charges they incur. Another interesting result is that there are cases where the volume charging ISP has negative profits and is forced to leave the market. The results are similar if there are more layers of ISPs charging similarly.

Protocols such as ECN, re-ECN and Conex are some very interesting implementation paradigms of congestion pricing in practice. Such mechanisms make the real-time congestion charges visible to the traffic sources in order for them to take the appropriate actions. In our model we considered a variation of scenarios, and we proved that it is beneficial for ISPs to charge based on congestion, since they become more competitive in the Internet market by gaining higher profits. Thus, based on the aforementioned remarks, we conclude that ISPs do have economic incentives to adopt congestion accountability mechanisms for charging their end-users.

Our future work includes the consideration of other types of dynamic pricing. During our initial work, we find out similar results when the price/bit announced by the congestion charging ISP increases faster than linearly (exponentially) which is the case when using ECN marks to account for congestion. Our model assumed a linear congestion price function to make our computations tractable. A price function with a faster than linear increase can only make our results stronger and provide a smoother traffic for ISP2.

REFERENCES

[1] K. Ramakrishnan, S. Floyd, and D. Black, "The Addition of Explicit Congestion Notification to IP", RFC 3168, 2001.

[2] B. Briscoe, A. Jacquet, C. Di Cairano-Gilfedder, A. Salvatori, A. Soppera, and M. Koyabe, "Policing Congestion Response in an Internetwork using Re-feedback", in Proc. of ACM SIGCOMM, Philadelphia, PA, USA, 2005.

[3] B. Briscoe, R. Woundy, T. Moncaster, and J. Leslie, "Congestion Exposure Mechanism (Conex)", IETF Internet Draft, draft-conex-mechanism-00, 2010.

[4] A. Jacquet, B. Briscoe, and T. Moncaster, "Policing Freedom to Use the Internet Resource Pool", CoNext ReArch'08, Madrid, Spain, 2008.

[5] J. MacKie-Mason, and H. Varian, "Pricing Congestible Network Resources", IEEE Journal on Selected Areas in Communications vol. 13, pp. 1141-1149, 1995.

[6] B. Briscoe, A. Jacquet, T. Moncaster, and A. Smith, "Re-ECN: Adding Accountability for Causing Congestion to TCP/IP", IETF Internet-Draft, draft-briscoe-tsvwg-re-ecn-tcp-07, 2009.

[7] M. Mathis, and B. Briscoe, "Congestion Exposure (ConEx) Concepts and Abstract Mechanism", IETF Internet-Draft, draft-ietf-conex-abstract-mech-11.txt, 2014.

[8] M. Kühlewind, and R. Scheffenegger, "TCP Modifications for Congestion Exposure", IETF Internet-Draft, draft-ietf-conex-tcp-modifications-04, 2013.

[9] D. Wagner, and M. Kühlewind, "Auditing of Congestion Exposure signals", IETF Internet-Draft, draft-wagner-conex-audit-01, 2014.

[10] A. Akella, J. Pang, A. Shaikh, S. Seshan, and B. Maggs, "A comparison of overlay routing and multihoming route control", in Proc. of ACM SIGCOMM, Portland, Oregon, USA, 2004.

[11] D. Goldenberg, L. Qiu, H. Xie, Y. R. Yang, and Y. Zhang, "Optimizing cost and performance for multihoming", in Proc. of ACM SIGCOMM, Portland, Oregon, USA, 2004.

[12] H. Wang, H. Xie, L. Qiu, A. Silberschatz, and Y. R. Yang, "Optimal ISP Subscription for Internet Multihoming: Algorithm Design and Implication Analysis", in Proc. of IEEE INFOCOM, Miami, Florida, USA, 2005.

[13] P. Maille, and B. Tuffin, "Analysis of Price Competition in a Slotted Resource Allocation Game", in Proc. of IEEE INFOCOM, Phoenix, AZ, USA, 2008.

[14] J. G. Wardrop, and J. I. Whitehead, "Correspondence. Some Theoretical Aspects of Road Traffic Research", ICE Proceedings: Engineering Divisions, vol. 1, issue 5, pp. 767-768, 1952.

[15] X. Cao, H. Shen, R. Milito, and P. Wirth, "Internet Pricing With a Game Theoretical Approach: Concepts and Examples", IEEE/ACM Transactions on Networking, vol. 10, no. 2, pp. 208-216, 2002.

[16] J. Hui Wanga, D. Ming Chiub, J. C.S. Luic, "A Game–Theoretic Analysis of the Implications of Overlay Network Traffic on ISP Peering", Computer Networks, vol. 52, issue 15, pp. 2961–2974, 2008.

[17] I. Papafili, S. Soursos, and G. D. Stamoulis, "A Novel Game-Theoretic Framework for Modeling Interactions of ISPs Anticipating Users' Reactions", in Proc. of VALUETOOLS, Cargese, France, 2012.

[18] D. DiPalantino, and R. Johari, "Traffic Engineering vs. Content Distribution: A Game Theoretic Perspective", in Proc. of IEEE INFOCOM, pp. 540-548, Rio de Janeiro, 2009.

[19] G. Biczok, S. Kardos, T. Anh Trinh, "Pricing Internet Access for Disloyal Users: a Game-Theoretic Analysis", in Proc. of SIGCOMM NetEcon '08, pp. 55-60, Seattle, WA, USA, 2008.

[20] D. Niyato, and E. Hossain, "WLC04-5: Bandwidth Allocation in 4G Heterogeneous Wireless Access Networks: A Noncooperative Game Theoretical Approach", GLOBECOM, 2006.

[21] H. Shen, and T. Basar, "Hierarchical Network Games with Various Types of Public and Private Information", 45th IEEE Conference on Decision and Control, pp. 2825- 2830, 2006.

[22] P. Hande, Mung Chiang, R. Calderbank, and J. Zhang, "Pricing under Constraints in Access Networks: Revenue Maximization and Congestion Management", in Proc. of IEEE INFOCOM, San Diego, USA, 2010.

[23] S. Kunniyur, and R. Srikant, "End-to-End Congestion Control Schemes: Utility Functions, Random Losses and ECN Marks", IEEE/ACM Transactions on Networking, vol. 11, pp. 689-702, 2003.

[24] T. Henderson, J. Crowcroft, and S. Bhatti, "Congestion Pricing: Paying Your Way in Communication Networks", Internet Computing, IEEE, vol. 5, issue 5, pp. 85-89, 2001.

[25] R. Gibbens, and F. P. Kelly, "Resource Pricing and the Evolution of Congestion Control", Automatica 35, 1999.

[26] I. Paschalidis, and J. Tsitsiklis, "Congestion-Dependent Pricing of Network Services", IEEE/ACM Transactions on Networking, vol. 8, no. 2, 2000.

[27] H. Mendelson, S. Whang, "Optimal Incentive-Compatible Priority Pricing for the M/M/1 Queue", Operations Research, vol. 38, no. 5, pp. 870-883, 1990.

[28] M. Mandjes, "Pricing Strategies Under Heterogeneous Service Requirements", in Proc. of IEEE INFOCOM, California, USA, 2003.

[29] T. Basar, and R. Srikant, "Revenue-Maximizing Pricing and Capacity Expansion in a Many-Users Regime", in Proc. of IEEE Infocom, New York City, USA, 2002.

[30] S. Shakkottai, and R. Srikant, "Economics of Network Pricing with Multiple ISPs", IEEE/ACM Transactions on Networking, vol. 14, issue 6, pp. 1233–1245, 2006.

[31] X. Dimitropoulos, P. Hurley, A. Kind, and M. Stoecklin, "On the 95-Percentile Billing Method", Passive and Active Network Measurement, LNCS, vol. 5448, pp. 207-216, 2009.

[32] J. Levy, H. Levy, and Y. Kahana, "Top Percentile Network Pricing and the Economics of Multi-homing", Annals of Operations Research, vol. 146, issue 1, pp. 153-167, 2006.

[33] G. Fortetsanakis, M. Papadopouli, G. Karlsson, M. Dramitinos, E. A. Yavuz, "To Subscribe, or not to Subscribe: Modeling and Analysis of Service Paradigms in Cellular Markets", in Proc. of IEEE DySPAN, Bellevue, Washington, USA, 2012.