Variation of:
“The effect of caching on a model of content and access provider revenues in information-centric networks”*

G. Kesidis
CS&E and EE Depts
Pennsylvania State University
University Park, PA, USA
gik2@psu.edu

June 13, 2014

Abstract

This is a variation of the two-sided market model of [10]: Demand $D$ is concave in $\tilde{D}$ in (16) of [10]. So, in (5) of [10] and after Theorem 2, take the parametric case $0 < a \leq 1$. Thus, demand $D$ is both decreasing and concave in price $p$, and so the utilities ($U = pD$) are also concave in price. Also, herein a simpler illustrative demand-response model is used in Appendix A and B.

1 Introduction

In this paper, we consider a game between an Internet Service (access) Provider (ISP) and content provider (CP) on a platform of end-user demand. A price-concave demand-response is motivated based on the delay-sensitive applications that are expected to be subjected to the assumed usage-priced

*This research was supported by NSF CNS grant 1116626.
priority service over best-effort service. Thus, we are considering a two-sided market with multiclass demand wherein one class (that under consideration herein) is delay-sensitive. Both the Internet and proposed Information Centric Network (ICN, encompassing Content Centric Networking (CCN)) scenarios are considered. For our purposes, the ICN case is basically different in the polarity of the side-payment (from ISP to CP in an ICN) and, more importantly here, in that content caching by the ISP is incented.

Pricing congestible commodities have been extensively studied. For example, in [7] a demand model is based on a “cost” that is the sum of a price and latency term. We herein take this relationship to be an implicit one in which the latency factor is also an increasing function of demand. The resulting price-concave demand-response model is extended to account for content caching. The corresponding Nash equilibria are derived as a function of the caching factor.

2 Problem Set-Up: The Internet model

Suppose there are two providers, one content (CP indexed 2) and the other access (ISP indexed 1), with common consumer demand-response [6]. First suppose that the demand response to price is linear:

\[ D = D_{\text{max}} - d(p_1 + p_2), \]  

(1)

where \( d \) is demand sensitivity to the price, \( p_1 \) and \( p_2 \) are, respectively, the prices charged by the ISP and CP, and \( D_{\text{max}} > 0 \) is the demand at zero usage based price \( p_1 \). Suppose the revenue of the ISP is

\[ U_1 = (p_1 + p_s)D, \]  

(2)

where \( p_s \) is the side payment from content to access provider. Similarly, the revenue of the CP is

\[ U_2 = (p_2 - p_s)D. \]  

(3)

---

1Leader-follower dynamics, rather than simultaneous play at the same time-scale, are considered in [11]. For the problem setting of this paper, leader-follower dynamics were considered by us in [1] and provider competition in [4, 9].

2Note that ISPs are continuing to depart from pure flat-rate pricing (based on maximum access bandwidth) for unlimited monthly volume, e.g., [12, 3].
Consider a noncooperative game played by the CP and ISP adjusting their prices, respectively $p_2$ and $p_1$, to maximize their respective revenues, with all other parameters fixed. In particular, the fixed side-payment $p_s$ is here assumed regulated. Note that the utilities are linear functions of $p_s$ so that if $p_s$ were under the control of one of the players, $p_s$ would simply be set at an extremal value.

The following simple result was shown in [1, 4].

**Theorem 1.** The interior Nash equilibrium\(^3\) is

$$p_1^* = \frac{D_{\text{max}}}{3d} - p_s \quad \text{and} \quad p_2^* = \frac{D_{\text{max}}}{3d} + p_s$$

when

$$|p_s| < \frac{D_{\text{max}}}{3d},$$

\(^3\)In this paper, we do not consider boundary Nash equilibria, where at least one player is selecting an extremal value for one of their control parameters, often resulting in that player essentially opting out of the game, or maximally profiting from it at the expense of the other player. The boundary equilibria are also specified in [1].
with player utilities

\[ U_1^*, U_2^* = \frac{D_{\text{max}}^2}{9d}. \]

Note that this result allows \( p_s < 0 \), i.e., net side payment is from ISP to CP (remuneration for content instead of access bandwidth). But in the Internet setting, we take \( p_s > 0 \), whether there is direct side-payment from CP to ISP (or, again, indirectly by payment through the peering contract between the residential ISP and the ISP of the CP - a contract that penalizes for asymmetric traffic exchange neutrally based on aggregate traffic volume).

In \([4, 8]\), we showed that the ISP may actually experience a reduction in revenue/utility with the introduction of side payments, using a communal demand model that had different demand-sensitivity-to-price parameters \( d \) per provider type and also multiple providers of each type (i.e., provider competition). Such a model was also considered in \([2]\).

Consider a concave demand response to price, e.g.,

\[ D = \left( \frac{1}{D_{\text{max}} - (p_1 + p_2)d + a} \right)^{-1}, \quad a \geq 0, \]  

where

\[ p_{\text{max}} = D_{\text{max}}/d \quad \text{when } a = 0. \]

The following is a simple extension of Theorem 1 accommodating (5).

**Theorem 2.** For utilities (2) and (3), the interior Nash equilibrium for a strictly concave demand response \( D \) is

\[ p_1^* + p_s = p^*/2 = p_2^* - p_s, \]  

where \( p^* = p_1^* + p_2^* \) solves

\[ D(p^*) + D'(p^*)p^*/2 = 0. \]  

and \( |p_s| < p^*/2. \)

For the example of (5) with \( a \geq 0 \),

\[ p^* = \frac{4aD_{\text{max}} + 3 - \sqrt{8aD_{\text{max}} + 9}}{4ad}. \]  

and |\( p_s | < p^*/2. \)
Note that simply by L’Hopital’s rule, \[ \lim_{a \to 0} p^* = \frac{2}{3} D_{\text{max}}/d = \frac{2}{3} p_{\text{max}}, \] which is consistent with Theorem 1. Again, under communal demand response with only one provider of each type, neither \( p^* = p_1^* + p_2^* \) nor \( U_1^* \) depend on the side payment \( p_s \). In an illustrative example of Appendix A, the parameter

\[ a = (B - \lambda)^{-1}, \]

where \( B \) is a possibly congested bandwidth resource, parameter \( 0 \leq \lambda < B \), and demand \( D \leq \min\{B - \lambda, D_{\text{max}}\} \).

### 3 ICN model

Again, in an ICN, residential users request content (or, more generally, information regarding application services) of the ISP/resolver, and the ISP/resolver decides the content provider. Therefore in an ICN, it’s reasonable to assume that the side-payment is from ISP to CP, \( i.e., p_s < 0 \). Also, the ISP is motivated to cache content, unlike for our simple Internet case, to reduce the side payment (\( i.e., \) avoid paying for, \( e.g., \) the networking costs of the ISP-selected CP to transmit the user-requested content). Suppose that the ISP decides to cache a fraction \( \kappa \) of the content and this results in lower delay between the CP and ISP, and a lower required side-payment to the CP, \( c.f., (9) \). If we model mean delay as \( 1/(B - D) \) \[^{[13]}\], where \( B \) is the service capacity between CP and ISP, then with caching factor \( \kappa \), this delay is reduced to \( 1/(B - (1 - \kappa)D) \). For the models of Appendix B, the demand response:

- is increasing in caching factor \( \kappa \),
- is concave in price for \( \kappa \in [0, 1) \), and
- tends to linear in price \( ([1]) \) as \( \kappa \to 1 \).

In an illustrative example of Appendix B, the demand parameter in \( (5) \) is

\[ a = (1 - \kappa)(B - \lambda)^{-1}. \]

Note that neither \( D_{\text{max}} \) nor \( p_{\text{max}} \) are assumed dependent on \( \kappa \). Because of ISP caching, the ISP and CP utilities generalize to

\[
\begin{align*}
U_1 &= (p_1 + (1 - \kappa)p_s)D - c(\kappa), \\
U_2 &= (p_2 - (1 - \kappa)p_s)D,
\end{align*}
\]

\[^{(9)}\]
again with \( p_s < 0 \), where \( c(\kappa) \) is the cost of caching borne by the ISP. In Appendix C, we argue that \( c \) is convex in \( \kappa \).

Note that the caching cost \( c \) component of \( U_1 \) does not depend on \( p_2 \) or \( p_1 \), and \( |p_s| < p^*/2 \) implies \( |(1 - \kappa)p_s| < p^*/2 \). So, we can use the results of Theorems 2 and 6 here, with parameters \((1 - \kappa)p_s \) and \( a = (1 - \kappa)(B - \lambda)^{-1} \) instead of \( p_s \) and \( a = (B - \lambda)^{-1} \), respectively, to obtain the utilities \( U_1^*, U_2^* \) at Nash equilibrium, \( p_1^*, p_2^* \). We can then consider how \( U_1^*, U_2^* \) depend on the caching factor \( \kappa \).

**References**

[1] E. Altman, P. Bernhard, S. Caron, G. Kesidis, J. Rojas-Mora, and S. Wong. A study of non-neutral networks under usage-based pricing. *Telecommunication Systems Journal Special Issue on Socio-economic Issues of Next Generation Networks*, 2011.

[2] E. Altman, A. Legout, and Y. Xu. Network non-neutrality debate: An economic analysis. In *Proc. IFIP Networking*, 2011.

[3] K. Bode. AT&T to impose caps, overages. [http://www.dslreports.com/shownews/Exclusive-ATT-To-Impose-Caps-Overages-113149](http://www.dslreports.com/shownews/Exclusive-ATT-To-Impose-Caps-Overages-113149), Mar. 13, 2011.

[4] S. Caron, G. Kesidis, and E. Altman. Application neutrality and a paradox of side payments. In *Proc. ACM Re-Architecting the Internet (ReArch) Workshop, Philadelphia*, Nov. 2010.

[5] G. Dan and N. Carlsson. Power-law revisited: A large scale measurement study of P2P content popularity. In *Proc. IPTPS*, 2010.

[6] N. Economides. Net neutrality: Non-discrimination and digital distribution of content through the Internet. *I/S: A Journal of Law and Policy*, 4:209–233, 2008.

[7] R. Johari, G.Y. Weintraub, and B. Van Roy. Investment and market structure in industries with congestion. *Operations Research*, 58(5), 2010.
Appendix A: Explanation of concave demand response

Consider a price-concave demand response $\tilde{D}$. In particular, for price $p \in [0, p_{\text{max}}]$, consider the linear case let

$$\tilde{D}(p) := D_{\text{max}} - pd = D_{\text{max}}(1 - p/p_{\text{max}}) \geq 0. \quad (10)$$

Suppose that the demand $D$ satisfies

$$D = [g(D)\tilde{D}]^+, \quad g(D) = \frac{1 - \lambda/(B - D)}{1 - \lambda/B},$$

where $g$ is a decreasing and concave factor, not dependent of price, accounting for demand loss due to congestion. For example [10],

[8] G. Kesidis. Side-payment profitability under convex demand-response modeling congestion-sensitive applications. In Proc. IEEE ICC, Ottawa, Canada, June 2012.

[9] F. Kocak, G. Kesidis, and S. Fdida. Network neutrality with content caching and its effect on access pricing. In Smart Data Pricing. S. Sen and M. Chiang (Eds.), Wiley, 2013.

[10] F. Kocak, G. Kesidis, T.-M. Pham, and S. Fdida. The effect of caching on a model of content and access provider revenues in information-centric networks. In Proc. ASE/IEEE EconCom, Washington, DC, Sept. 2013.

[11] J. Musacchio, G. Schwartz, and J. Walrand. A two-sided market analysis of provider investment incentives with an application to the net-neutrality issue. Review of Network Economics, 8(1), 2009.

[12] A. Shin. Who’s the bandwidth bandit? [Online]. Available: [http://blog.washingtonpost.com/thecheckout/2006/10/bandwidth_bandit.html](http://blog.washingtonpost.com/thecheckout/2006/10/bandwidth_bandit.html), Oct. 4, 2006.

[13] R.W. Wolff. Stochastic Modeling and the Theory of Queues. Prentice-Hall, Englewood Cliffs, NJ, 1989.
where the term $1/(B - D)$ is taken from the queueing delay of an M/M/1 queue with mean arrival rate $D$ and mean service rate $B > D$. More simply, we can take

$$g(D) = \frac{B - \lambda - D}{B - \lambda}, \quad (11)$$

where $B$ is the available bandwidth resource, and parameter $0 \leq \lambda < B$. Note that for these examples:

- $g(0) = 1,$
- $g(B - \lambda) = 0$, so that
- $0 \leq D \leq \min\{B - \lambda, D_{\text{max}}\}$, and
- $g$ is non-negative, decreasing and concave.

**Lemma 1.** If $D = g(D)\tilde{D} > 0$ with $g$ nonnegative, decreasing and concave and $\tilde{D} \geq 0$, then $D$ is increasing and concave in $\tilde{D}$.

**Proof.** Let $g' = dg/dD$ and $D' = dD/d\tilde{D}$. By direct differentiation with respect to $\tilde{D}$:

$$D' = g + \tilde{D}g'D' \Rightarrow D' = \frac{g}{1 - \tilde{D}g'} \geq 0$$

$$\Rightarrow D'' = \frac{(1 - \tilde{D}g')g' + (\tilde{D}g'' + g')g}{(1 - \tilde{D}g')^2}D' \leq 0,$$

\[\square\]

**Corollary 1.** Under Lemma 1 and if $\tilde{D}$ is decreasing and concave in price $p$, then $D$ is non-negative and decreasing in $p$, and both $D$ and $pD$ are concave in $p$.

**Proof.** By the above lemma, $D(\tilde{D}) \geq 0$ is increasing and concave. Again, by direct differentiation:

$$\frac{\partial D(\tilde{D})}{\partial p} = D'(\tilde{D})\frac{\partial \tilde{D}}{\partial p} \leq 0$$

$$\frac{\partial^2 D(\tilde{D})}{\partial p^2} = D''(\tilde{D})\left(\frac{\partial \tilde{D}}{\partial p}\right)^2 + D'(\tilde{D})\frac{\partial^2 \tilde{D}}{\partial p^2} \leq 0$$

\[\square\]
For linear demand-response to price \( (10) \) and the linear congestion factor \( (11) \),

\[
D(p) = (\tilde{D}(p)^{-1} + (B - \lambda)^{-1})^{-1} = (B - \lambda) \left( 1 - \frac{1}{1 + (D_{\text{max}} - dp)/(B - \lambda)} \right)
\]

which is decreasing and concave in \( p \), with

\[
D(0) = (D_{\text{max}}^{-1} + (B - \lambda)^{-1})^{-1} \leq \min\{B - \lambda, D_{\text{max}}\}.
\]

It’s also easy to see that

\[
\lim_{B \to \infty, \ p \to 0} D = D_{\text{max}}.
\]

**Appendix B: Explanation of demand increasing in caching factor**

As a result of ISP caching, only a fraction \((1 - \kappa)\) of the demand \( D \) is transmitted through the the bandwidth \( B \) between ISP and CP. So, the congestion factor \((11)\) is modified to

\[
g_\kappa(D) = \frac{B - \lambda - (1 - \kappa)D}{B - \lambda} = \frac{(B - \lambda)/(1 - \kappa) - D}{(B - \lambda)/(1 - \kappa)}
\]

So, solving \( D = \tilde{D}g_\kappa(D) \) results in \((12)\) with \( B - \lambda \) replaced by \((B - \lambda)/(1 - \kappa)\):

\[
D(p) = (\tilde{D}(p)^{-1} + (1 - \kappa)(B - \lambda)^{-1})^{-1}.
\]

Thus, if positive \( \kappa < 1 \), the demand is concave in price \( p \) and increasing in \( \kappa \). On the other hand, as \( \kappa \to 1 \), the demand tends to linear in price \((1)\).

**Lemma 2.** Generally, if the congestion factor \( g \) is a decreasing function, then the demand \( D \) increases with caching factor \( \kappa \).
Proof. First note that $g_\kappa(D) = g_0((1 - \kappa)D) := g((1 - \kappa)D)$, is decreasing in $(1 - \kappa)D$ (hence increasing in caching factor $\kappa$). Consider the solution
\[ D_\kappa = \tilde{D}g_\kappa(D_\kappa) \] (14)
and note that $D_\kappa \geq D_0$. Now,
\[ D_0 = \tilde{D}g_0(D_0) < \tilde{D}g_0((1 - \kappa)D_0) = \tilde{D}g_\kappa(D_\kappa). \]
So, if $D_\kappa \leq D_0$, then we would have
\[ D_\kappa \leq D_0 < \tilde{D}g_\kappa(D_0) \leq \tilde{D}g_\kappa(D_\kappa), \]
which contradicts the definition of $D_\kappa$ in (14). \qed

Appendix C: Convexity of cost of caching as a function of caching factor

Assume that the cost of caching is proportional to the number of cached items (content), in turn proportional to the (mean) amount of memory required to store them. For a fixed population of $N$ end-users (a proximal group served by an ISP), let $\pi(j)$ be the proportion of the items that will soon be of interest to precisely $j$ end-users. Finally, suppose the ISP naturally prioritizes its cache to hold the most popular content. So, a “caching factor” $\kappa$, based on all-or-none decisions to cache content of the same popularity, would satisfy
\[ \kappa \propto \sum_{j=N-f(\kappa)}^{N} j\pi(j). \]
for some $f(\kappa) \in \{0, 1, 2, \ldots, N\}$. The cost of caching would be proportional to the number of cached items, i.e.,
\[ c(\kappa) \propto \sum_{j=N-f(\kappa)}^{N} \pi(j). \]

Suppose that the great majority of potentially desired content is only minimally popular, i.e., $\pi(j)$ is decreasing.\footnote{Note that this general assumption obviously accommodates the empirically observed Zipf distribution for content popularity, e.g., \cite{5}.} We now argue that the caching cost
$c(\kappa)$ is convex and increasing for the simplified continuous scenario ignoring the (positive) constants of proportionality:

$$\kappa = \int_{N-f(\kappa)}^{N} z \pi(z) dz \quad \text{and} \quad c(\kappa) = \int_{N-f(\kappa)}^{N} \pi(z) dz,$$

with $c(0) = 0$ and $c(1) = 1$. By differentiating successively, we get

$$1 = (N - f(\kappa)) \pi(N - f(\kappa)) f'(\kappa) \quad (15)$$

$$c'(\kappa) = \pi(N - f(\kappa)) f'(\kappa)$$

$$\Rightarrow 1 = (N - f(\kappa)) c'(\kappa)$$

$$\Rightarrow c''(\kappa) = f'(\kappa)(N - f(\kappa))^{-2} \quad (16)$$

Note that $f' > 0$ by (15) and therefore $c'' > 0$ by (16).