RJ 10196 (95070) November 2, 2000
Mathematics

Research Report

NONLINEAR PRICING OF INTERNET BANDWIDTH VIA CHANCE CONSTRAINED PROGRAMMING

Xin Guo

IBM Research Division
T.J. Watson Research Center

John Tomlin

IBM Research Division
Almaden Research Center

LIMITED DISTRIBUTION NOTICE

This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties).

IBM Research Division
Yorktown Heights, New York • San Jose, California • Zurich, Switzerland

NONLINEAR PRICING OF INTERNET BANDWIDTH VIA CHANCE CONSTRAINED PROGRAMMING

Xin Guo

IBM Research Division
T.J. Watson Research Center
email: xinguo@watson.ibm.com

John Tomlin

IBM Research Division
Almaden Research Center
email: tomlin@almaden.ibm.com

ABSTRACT: This paper presents a model for pricing of bandwidth (for internet access) which incorporates nonlinear price-demand data directly into an optimization model. This is accomplished through a chance-constrained programming formulation, leading to an equivalent nonlinear deterministic model.

Keywords: Dynamic Pricing, Bandwidth, Chance-Constrained Programming

1. Introduction

In this paper we consider the problem of a re-seller who buys bandwidth in bulk and sells it in smaller bundles to customers who may buy a certain amount of bandwidth at a price we wish to determine. The reseller must not only choose prices which will attract customers, but also make sure that these customers do not collectively exceed the bandwidth available. Since the behaviour of the end users is neither deterministic nor under the re-sellers control, we shall take this to mean that given the distribution of individual customer bandwidth consumption, the total available shall not be exceeded with some (high) probability at any time t within the planning horizon. This is accomplished by means of "chance-constraining" (see, e.g. [1],[5]) total bandwidth consumption.

Other models for network and bandwidth management using a chance- constrained approach have been discussed by Hui[2] and Medova[4]. However our emphasis is on obtaining an analytical (nonlinear) deterministic equivalent of the chance constraints, in terms of the parameters of the distributions involved, in such a way that price-demand curve data can be plugged directly into the model and used to optimize revenue subject to these constraints over some time horizon.

2. The fundamental approach

Let us begin by considering only one class of customer. For any fixed time t, let

$$Y_t = X_1 + X_2 + \cdots + X_{N(t)}$$

and assume that the X_i's are i.i.d. normal distributed with mean μ and variance σ, which represent the real usage of N(t) signed-on customers. Let us further assume that customers arrival is described by a Poisson distribution with $\lambda = \lambda(t)$, independent of X_i, representing the arrival rate of customers. Then:

$$P(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}.$$

The parameter λ will be dependent on the price set, as we shall discuss below.

In order to derive any reasonable pricing policy related to revenue and performance, we must find a way of specifying that the customers collective bandwidth consumption does not exceed that available. We do this by formulating a "chance constraint":

$$P(Y_t > b_t) \leq \delta_t \tag{1}$$

for functions b_t and δ_t representing bandwidth sold and a tolerance on capacity violation. We now examine methods for expressing this constraint in a computationally tractable way.

Consider the following moment generating function for Y_t:

$$\psi_r(Y_t) = E[e^{rY_t}]$$

where $r > 0$ is any positive real number so that the above form converges. By the well-known property of the conditional expectation, we have

$$\begin{aligned} E[e^{rY_t}] &= E[E[e^{rY_t}|N(t)]], \\ &= \sum_{k=0}^{\infty} E[e^{r(X_1+\cdots+X_k)}]P(N(t)=k). \end{aligned}$$

Recall that the moment generating function of a normal distribution with mean μ and variance σ is $e^{r\mu+\frac{\sigma^2 r^2}{2}}$. Note also that if X_i and X_j are independent, then $\psi_r(X_i + X_j) = \psi_r(X_i)\psi_r(X_j)$, hence

$$\begin{aligned} E[e^{rY_t}] &= \sum_{k=0}^{\infty} \frac{(\lambda t)^k e^{-\lambda t}}{k!} e^{kr\mu+kr\sigma^2/2} \\ &= \sum_{k=0}^{\infty} \frac{e^{-\lambda t}}{k!} (\lambda t e^{r\mu+r\sigma^2/2})^k \\ &= e^{-\lambda t} \exp(\lambda t e^{(r\mu+r^2\sigma^2/2)}) \end{aligned}$$

the last equation is by the Taylor expansion of e^x for $|x| < 1$. Here $r < A$ such that $\lambda t \exp(r\mu + r^2\sigma^2/2) < 1$. Therefore,

$$\psi_r(Y_t) = \exp(\lambda t(e^{r\mu+\sigma^2 r^2/2} - 1)) \qquad \forall |r| < A. \tag{2}$$

Therefore, for any $b_t > 0, r > 0$, we have

$$\begin{aligned} P(Y_t \geq b_t) &= P(\exp(rY_t) \geq \exp(rb_t)) \\ &\leq \frac{E[e^{rY_t}]}{e^{rb_t}} \quad (\forall 0 < r < A) \\ &= \exp(\lambda t(e^{r\mu+\sigma^2 r^2/2} - 1) - rb_t) \quad \forall 0 < r < A, \end{aligned}$$

i.e.

$$P(Y_t \geq b_t) \leq \inf_{0<r<A} \exp(\lambda t(e^{r\mu+\sigma^2 r^2/2} - 1) - rb_t).$$

Solving the right side of the inequality, we have

$$\exp(\lambda t(\exp(r_0\mu + \sigma^2 r_0^2/2) - 1) - r_0 b_t) = \inf_{0<r<A} \exp(\lambda t(\exp(r\mu + \sigma^2 r^2/2) - 1) - rb_t),$$

where r_0 is the solution of

$$\lambda t(\mu + \sigma^2 r/2)(\exp(r\mu + \sigma^2 r^2/2) - 1) = b_t.$$

I.e.,

$$(\mu + \sigma^2 r/2)(\exp(r\mu + \sigma^2 r^2/2) - 1) = b_t/\lambda t.$$

An alternative and simpler method of bounding $P(Y_t > b_t)$ is as follows. By definition

$$\frac{\partial^2 \psi_r(Y_t)}{\partial r^2} |_{r=0} = E[Y_t^2].$$

thus,

$$E[Y_t^2] = \lambda t \sigma^2 + \lambda t \mu^2 + (\lambda t)^2 \mu^2.$$

hence,

$$P(Y_t > b_t) \leq \frac{E[Y_t^2]}{b_t^2} = \frac{\lambda t \sigma^2 + \lambda t \mu^2 + (\lambda t)^2 \mu^2}{b_t^2}.$$

Note that a constant λ can be replaced by some function, say $\lambda(q)$, assuming a Poisson distribution with $\lambda(t)$ expected arrivals on $[0, t]$. We shall use this below, when q is the price and $\lambda(q)$ is a function of that price.

3. Fixed duration of contracts

So far we have not considered the length of a customer contract. Let us suppose that for the single type of customer we are (so far) considering the fixed contract length is D. We also suppose that at time $t = 0$ we have n_0 initial customers of this type, and that at time $t < D$ there are n_t of these remaining. In this case the generating function becomes:

$$\bar{\psi}_r = \sum_{k=0}^{\infty} E[\exp(r(X_1 + \cdots + X_{k+n_t}))] P(N(t) = k). \tag{3}$$

which by the same arguments as above leads to:

$$\bar{\psi}_r(Y_t) = e^{n_t(r\mu + r^2\sigma^2/2)} \exp(\lambda t(e^{r\mu + \sigma^2 r^2/2} - 1)) \qquad \forall |r| < A. \tag{4}$$

Then for $t \leq D$

$$E[Y_t^2] = \frac{\partial^2 \bar{\psi}_r(Y_t)}{\partial r^2} |_{r=0} \tag{5}$$

$$= \lambda t \mu^2 + (n_t + \lambda t)\sigma^2 + (n_t + \lambda t)^2 \mu^2 \tag{6}$$

hence for $t, < D$

$$P(Y_t > b_t) \leq \frac{E[Y_t^2]}{b_t^2} = \frac{\lambda t \mu^2 + (n_t + \lambda t)\sigma^2 + (n_t + \lambda t)^2 \mu^2}{b_t^2} \tag{7}$$

For the case when $t \geq D$ we need only consider arrivals since time $t - D$, Hence for $t \geq D$, and appropriate r:

$$\hat{\psi}_r(Y_t) = \exp(\lambda D(e^{r\mu + \sigma^2 r^2/2} - 1)) \tag{8}$$

Thus in this case:

$$E[Y_t^2] = \frac{\partial^2 \hat{\psi}_r(Y_t)}{\partial r^2} |_{r=0} \quad (9)$$
$$= \lambda D\sigma^2 + \lambda D\mu^2 + (\lambda D)^2\mu^2. \quad (10)$$

and so for $t \geq D$

$$P(Y_t > b_t) \leq \frac{E[Y_t^2]}{b_t^2} = \frac{\lambda D\sigma^2 + \lambda D\mu^2 + (\lambda D)^2\mu^2}{b_t^2}.$$

4. Multiple types of contract

The next requirement is to generalize the model to multiple types of contracts; with different distributions $N(\mu_i, \sigma_i)$ and arrival rates λ_i, which are themselves functions of the prices q_i at which the contracts are offered. For the moment we ignore the contract duration discussed in the previous section. Again assuming independence we can define $Z_t = Y_{1t} + \cdots + Y_{mt}$, where Y_{it} has the moment generating function ψ_{it} for contract type i at time t. Then,

$$\psi_r(Z_t) = \prod_{i=1}^{m} \psi_r(Y_{it})$$

Correspondingly we have:

$$\frac{\partial \psi_r(Z_t)}{\partial r} = \sum_{i=1}^{m} \frac{\partial \psi_r(Y_{it})}{\partial r} \prod_{j \neq i}^{m} \psi_r(Y_{jt})$$

and

$$\frac{\partial^2 \psi_r(Z_t)}{\partial r^2} = \sum_{i=1}^{m} \left[\frac{\partial^2 \psi_r(Y_{it})}{\partial r^2} \prod_{j \neq i}^{m} \psi_r(Y_{jt}) + \frac{\partial \psi_r(Y_{it})}{\partial r} \sum_{j \neq i} \frac{\partial \psi_r(Y_{jt})}{\partial r} \prod_{k \neq i,j}^{m} \psi_r(Y_{kt}) \right]$$

Now we have

$$\psi_0(Y_{it}) = 1, \frac{\partial \psi_r(Y_{it})}{\partial r} |_{r=0} = \lambda_i \mu_i t, \frac{\partial^2 \psi_r(Y_{it})}{\partial r^2} |_{r=0} = \lambda_i t(\mu_i^2 + \sigma_i^2) + (\lambda_i t \mu_i)^2$$

and so

$$E[Z^2] = \frac{\partial^2 \psi_r(Z_t)}{\partial r^2} |_{r=0}$$
$$= \sum_{i=1}^{m} \left[\frac{\partial^2 \psi_r(Y_{it})}{\partial r^2} |_{r-0} + \frac{\partial \psi_r(Y_{it})}{\partial r} |_{r-0} \sum_{j \neq i} \frac{\partial \psi_r(Y_{jt})}{\partial r} |_{r=0} \right] \quad (11)$$

For the simple case, when contract duration is ignored, we then have

$$\begin{aligned} E[Z_t^2] &= \sum_{i=1}^{m}\left[\lambda_i t(\mu_i^2+\sigma_i^2)+(\lambda_i t\mu_i)^2+\lambda_i\mu_i t\sum_{j\neq i}\lambda_j\mu_j t\right] \\ &= \sum_{i=1}^{m}[\lambda_i t(\mu_i^2+\sigma_i^2)]+(\sum_{i=1}^{m}\lambda_i t\mu_i)^2 \end{aligned} \quad (12)$$

which leads us to the bound

$$P(Z_t > b_t) \leq \frac{\sum_{i=1}^{m}[\lambda_i(t)\mu_i^2+\lambda_i(t)\sigma_i^2]+(\sum_{i=1}^{m}\lambda_i(t)\mu_i)^2}{b_t^2} \quad (13)$$

Reintroducing the contract durations D_i and plugging $\bar{\psi}_r$ or $\hat{\psi}_r$ into (11) for ψ_r as appropriate, we obtain the general form:

$$\begin{aligned} E[Z_t^2] &= \sum_{i|t<D_i}\left[\lambda_i t\mu_i^2+(n_{it}+\lambda_i t)^2\sigma_i^2+(n_{it}+\lambda_i t)^2\mu_i^2\right] \\ &\quad+\sum_{i|t\geq D_i}\left[\lambda_i D_i(\mu_i^2+\sigma_i^2)]+(\lambda_i D_i\mu_i)^2\right] \end{aligned} \quad (14)$$

5. Discrete time periods

To render our model computationally tractable, we now use a discrete representation of t, with time periods $\tau = 1, ..., T$ of equal length Δ. Correspondingly we must define contract duration D_i as a multiple of Δ, i.e. $D_i = d_i\Delta$. New customers are assumed to start their usage at the *start* of each time period, hence our discrete version of (14) becomes:

$$\begin{aligned} E[Z_\tau^2] &= \sum_{i|\tau<d_i}\left[\lambda_i\tau\Delta\mu_i^2+(n_{i\tau}+\lambda_i\tau\Delta)^2\sigma_i^2+(n_{i\tau}+\lambda_i\tau\Delta)^2\mu_i^2\right] \\ &\quad+\sum_{i|\tau\geq d_i}\left[\lambda_i D_i(\mu_i^2+\sigma_i^2)]+(\lambda_i D_i\mu_i)^2\right] \end{aligned} \quad (15)$$

Using this expression in the discrete chance constraint:

$$E[Z_\tau^2] - b_\tau^2\delta_\tau \leq 0$$

we are now ready to formally specify the model.

6. The nonlinear programming model

The chance-constrained optimization model for bandwidth management is defined as follows:

Indices

$i = 1, ..., I$	customer class, defined by a distribution $N(\mu_i, \sigma_i)$
$\tau = 1, ..., T$	time periods, each of length Δ

Data

δ_τ	tolerance on capacity violation
C_τ	cost per unit of buying new capacity in period τ
d_i	duration of contract (number of time periods) for customer class i
D_i	actual duration of contract ($d_i\Delta$) for customer class i
$n_{i\tau}$	number of existing contracts of type i still active at start ofperiod τ
$L_{i\tau}$	lower bound on contract price
$U_{i\tau}$	upper bound on contract price

Variables

b_τ	the bandwidth available in period τ
a_τ	bandwidth purchased in period τ
$q_{i\tau}$	price to new (or renewing) customers for a new standard length contract of type i in period τ

User Supplied Functions

$\lambda_i(q_{i\tau})$	the expected number of new customers of type i arriving in any period if the price for a constract is set at $q_{i\tau}$. (This reflects the elasticity of demand).

Constraints

$$b_\tau = b_{\tau-1} + a_\tau \qquad (\tau = 1, ..., T) \tag{16}$$

$$L_{i\tau} \leq q_{i\tau} \leq U_{i\tau} \quad (i = 1, ..., I; \tau = 1, ..., T) \tag{17}$$

$$\sum_{i|\tau<d_i} \left[\lambda_i\tau\Delta\mu_i^2 + (n_{i\tau} + \lambda_i\tau\Delta)^2\sigma_i^2 + (n_{i\tau} + \lambda_i\tau\Delta)^2\mu_i^2\right] + \sum_{i|\tau\geq d_i} \left[\lambda_i D_i(\mu_i^2 + \sigma_i^2)] + (\lambda_i D_i\mu_i)^2\right] - \delta_\tau b_\tau^2 \leq 0 \quad \forall\tau \tag{18}$$

Objective

$$\text{Maximize} \quad \sum_{i,\tau} q_{i\tau}\lambda_i(q_{i\tau}) - \sum_\tau C_\tau a_\tau$$

7. Using the model

In formulating this model we have made a number of choices about generality, some of which are relevant to use of the model. As with most multi-time-period optimization models, we envisage rerunning the model as frequently as is practical to reflect new data on hand. In particular the current number of customers of each type is constantly changing, and if the model is not producing accurate forecasts the price-demand curve data may need to be revised. For this reason the functions λ_i are assumed to be independent of τ, as we do not expect to be able to make sufficiently accurate long term forecasts of these functions to make them time varying, and expect frequent revision. However, the price set per time period ($q_{i\tau}$) is allowed to be time dependent and controllable by the bounds $L_{i\tau}, U_{i\tau}$. It is a simple matter to modify the model to enforce a uniform price per contract type over the model horizon. We expect the specification of good price-demand functions to be one of the most challenging aspects of actual use of the model.

The tolerance on exceeding capacity (δ_τ) has also been specified as time-dependent, but it is even more trivial to make it constant. Likewise the cost (C_τ) of acquiring new bandwidth for resale.

The overall model has a modest number of constraints per time period (depending on the number of contract types I) , however the nonlinear constraint (18) is quite complex, and it remains to be seen how available nonlinear programming solvers handle them.

8. Further research

In building this model we made several fundamental assumptions. In particular we assumed that customers actual usage can be modeled by the normal distribution. Strictly speaking we should assume some positive distribution, e.g. a gamma distribution. Computational experiments should confirm whether the normal assumption is adequate.

The other major assumption is that the customers usage pattern is independently distributed. This is clearly not absolutely true. The question is how well the model formulated with this assumption works in practice. This too can only be determined by experiment.

References

[1] M.A.H. Dempster (ed.), *Stochastic Programming*, Academic Press, London and NY (1980)

[2] J.Y. Hui, "Resource allocation for Broadband Networks", *IEEE J. on Selected Areas in Communications* **6**, 1598-1608, (1988).

[3] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol 1 (3rd edition), (1968) and Vol 2, Wiley, NY (1971).

[4] E.A. Medova, "Chance-constrained stochastic programming for integrated services network management", *Annals of Operations Research* **81**, 213-229 (1998).

[5] R. Wets, "Stochastic programming: Solution techniques and approximation schemes", in A. Bachem, M. Grotschel and B. Korte (eds.), *Mathematical Programming The State of the Art*, Springer, 566-603 (1983).