An Analytical Framework for Pricing of Services in the Internet

Srinivasan Jagannathan
Department of Computer Science
University of California
Santa Barbara, CA 93106-5110
jsrini@cs.ucsb.edu

Abstract— **There exists a huge demand for multimedia goods and services on the Internet. In this paper, we develop an analytical framework to price such services in the Internet. As a first step, we consider a system where a server handles requests for a service on a First-Come-First-Served (FCFS) basis. We develop a model where customers can refuse the service based on their *capacity* to pay and their *willingness* to do so. We show that charging a flat price maximizes the expectation of revenue and derive the optimal price for the analytical framework we develop. We also discuss how resource constraints can prevent achieving this maximum expectation and analyze pricing strategies for such situations.**

I. Introduction

The Internet is seeing an explosive growth in commercial activities. Products ranging from cameras to cars are being sold in the Internet. At the same time, services like video-on-demand, though touted as killer applications, have failed to take off. However, content delivery in general is a major activity and is growing at a fast rate. For instance, downloadable software is highly popular on the Internet. One can think of scenarios where customers can download music, movies and even books after online transactions. In the Internet, a service provider can charge different prices from different customers for such software, as long as the customer agrees to the price. The variability in price can be attributed to the time of day at which the requests are made. At the customer end, price determines the decision to purchase the product or service. The customer can accept or reject a price based on his/her *capacity* to pay and *willingness* to do so. Choosing the right price is therefore of great importance to maximize revenue. This paper develops an analytical framework for pricing of services in such a setting.

The Internet is an example of a market where there is a potentially infinite supply of goods and services. The only limitation in supply is due to lack of distribution resources at the retailer or service provider. Let us consider an illustrative example. Consider a service provider selling downloadable CDs. The number of CDs that can be downloaded from the web-site within a given time frame is limited by the bandwidth and server resources available. Furthermore, the resources available cannot be arbitrarily increased. This is because, the demand (or request arrival process) cannot be easily predicted. For instance, a very popular music album available at an exclusive web-site may increase demand for a a short period of time, say a fortnight. Once the initial popularity wanes, demand (and hence request arrival rate) will drop. Acquiring high capacity links and server resources to meet the demand may therefore not be a practical solution. At the same time, short-term acquisition of server resources and bandwidth may not be possible. An intelligent pricing strategy on the other hand will help maximize revenue under a given set of resource constraints.

Under specific assumptions of user behavior and a system model, we analyze pricing mechanisms which maximize *expectation* of revenue. We use the term *expectation* in the statistical sense, because the revenue generated depends on a probabilistic user behavior model. To illustrate the probabilistic nature of user behavior, let us consider an example. Consider a teenager with $15 as pocket money at a video-game parlor. The latest release of a hit video-game is very attractive to him, but whether or not he chooses to play the game depends on the price associated with the game and the money he has with him. He may be very likely to play for $5, but not for $14. He may decide to wait for another month when the game is not so new and the price falls. But, if the price is greater then $15, he cannot play even if he wishes to do so. There is a probability associated with his decision to play based on the price and his capacity to pay. We can see a direct correlation between the example described here and purchasing goods/services on the Internet. In general, the probability that a customer buys the service decreases with price and increases with his or her capacity to pay.

Summarizing our discussion above, the delivery of content depends on three factors—resource availability, customer capacity to pay and customer willingness to pay. In this paper, we analyze pricing mechanisms for a FCFS system with finite resources under a *Pareto* distribution of customer capacity to pay and a probabilistic model for user willingness to pay the quoted price. We show that charging a constant price will maximize the expected revenue for any user willingness model in which user willingness decays with increasing price. We derive the constant price for the user willingness models we use in this paper. We also discuss how resource constraints can prevent achieving the maximum expectation of revenue and formulate pricing strategies for such situations. Our work is based on a video-on-demand server, but it is sufficiently general to be applied to other forms of content and services on the Internet.

The rest of the paper is organized as follows. We describe our basic system model used in this paper in Section 2. We formulate the theoretical expectation of revenue in Section 3. In Section 4, we discuss pricing strategies when resources are constrained. In Section 5, we discuss generalizations of the framework developed in this paper. We conclude the paper in Section 6.

II. System Model

We consider a system where requests are satisfied if resources are available and the customer agrees to pay the quoted price. Resources are modeled as *logical channels*. Every request which is satisfied occupies a channel for some finite amount of time. For a video-on-demand server we can think of the channels as the number of movies that can be served simultaneously. In this paper we do not focus on how the channel is allocated or how an allocated channel is managed. These issues have been treated in detail in earlier work [1], [2]. We mainly focus on the interaction between the system and the customer before a channel is allocated. The sequence of actions resulting in content-delivery is depicted in Figure 1. In this paper, we assume that there is no distinction or classification in the service. For a video-on-demand system, this can be thought of as a model in which all movies (both popular as well as unpopular) are treated uniformly, i.e., price is independent of the popularity of the movie. The framework we develop can be suitably modified to handle the general case of multiple classes of service offered by the same service provider. We defer a discussion on this subject to later sections.

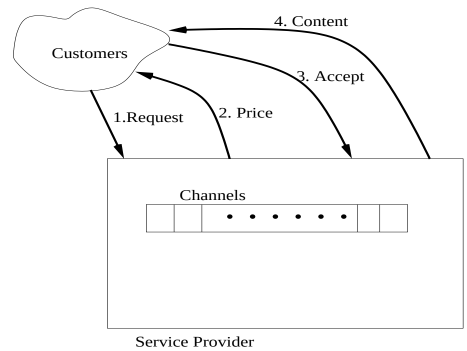

Fig. 1. System Model

Economic theory has established that there are a large number of customers with a small income and a very small number of customers with a very large income [3]. It is reasonable to assume that customers' capacities to spend will follow a similar behavior. Currently, two probability distribution models – *Pareto* and *log-normal* are used to represent the distribution of incomes[3], [4]. In this paper, we use the Pareto distribution to represent the capacity to spend. Every customer has the capacity to pay based on a Pareto distribution with two parameters–shape α and scale b. All customers have capacities at least as large as b. The shape α determines how the capacities are distributed. The larger the value of α, the fewer the people with a very large capacity to pay. The Pareto density function is defined as shown below:

$$f_{\wp}(x) = \begin{cases} \frac{\alpha b^{\alpha}}{x^{\alpha+1}} & , \quad x \geq b. \\ 0 & , \quad x < b \end{cases} \tag{1}$$

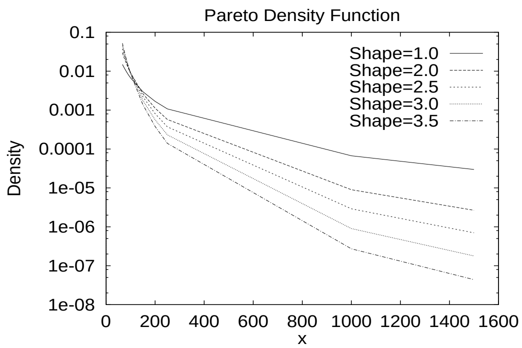

Fig. 2. Pareto Density

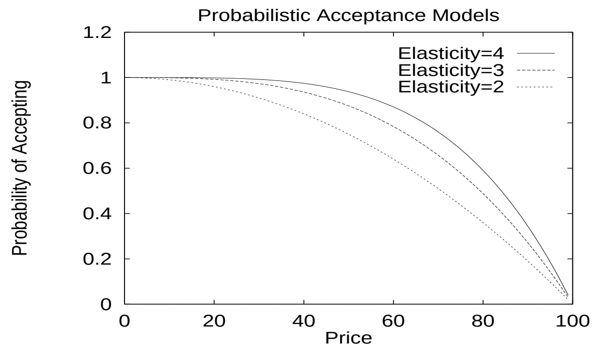

Fig. 3. Probabilistic User Willingness

Figure 2 illustrates the Pareto density function for different values of shape α, and scale $b = 67$[1]. Let us consider an illustrative example to understand the Pareto distribution of capacities. Consider a video-on-demand server. We can expect all customers to have a capacity to pay at least some money for the movie. We call the largest such amount that can be paid by all the customers as the scale of the distribution of their capacities and denote it as b. We would expect most of the customers to be able to pay only about this amount. There will be very few customers who can pay a lot more than the scale. This information is captured by the shape of the distribution, which we denote as α. The greater the value of α, the fewer the customers who can pay a lot more than b. When $\alpha \to \infty$, all customers have the same capacity b. For systems like video-on-demand servers, we would expect the shape to be a very large but finite number.

Even though customers *can* spend, they may not be *willing* to do so. To adequately describe the willingness of customers to pay, we define a family of probability functions. Consider an arbitrary customer with capacity χ. We denote his/her decision to purchase the service, by the random variable Υ which can take two values–1 for accept and 0 for reject. As discussed in the example in the previous section, the probability that the customer accepts the price ψ, denoted by $P\{\Upsilon = 1 \mid \psi\}$ depends on his/her capacity χ, and the price ψ. In this paper, we work with a simple model, where $P\{\Upsilon = 1 \mid \psi\}$ is defined as shown in Equation 2:

$$P\{\Upsilon = 1 \mid \psi\} = \begin{cases} 1 - \left(\frac{\psi}{\chi}\right)^{\delta} & , \quad 0 \leq \psi \leq \chi \\ 0 & , \quad \psi > \chi \end{cases} \qquad (2)$$

By varying the parameter δ in Equation 2, we can make the willingness as *elastic* as desired. The higher the value of δ, the more willing are customers to spend money. We show three different willingness models for a customer having capacity 100, with δ values 2, 3 and 4 respectively in Figure 3. As can be seen, the model with $\delta = 4$ makes the customer much more willing to spend money than in the case of the other two models. In fact, as δ increases to around 4.0 or greater, the willingness begins to resemble a "step-function", i.e., the customer is willing to pay as much as his/her capacity to pay. Even though different customers shall in reality have different degrees of willingness, in this paper we shall assume that all customers have the same degree of willingness. This makes the analysis simpler while at the same time illustrating the overall customer behavior.

Typically, economic analysis of customer behavior involves utility function models. In these models, it is assumed that every customer associates a specific *value* to the product and the price is compared with this value. The model we use in this paper can be shown to be similar to one such utility function based model. We chose our model of customer behavior instead of a utility function based model because it is difficult to quantify the value a customer associates with the content. Choosing a probabilistic model on the other hand makes the analytical framework easier while at the same time capturing typical customer behavior.

[1]For, b=67 and $\alpha = 3$, the mean of the Pareto distribution is 100.

Notation	Description
α	Shape of Pareto distribution
b	Scale of Pareto distribution
Υ	Decision to purchase (0 or 1)
δ	Elasticity of willingness
γ	Revenue per customer
Γ	Total revenue
ψ	Price
χ	Capacity of arbitrary customer
λ	Request arrival rate
n	Number of channels
d	Mean service time
ρ	System Utilization

TABLE I
SYMBOLS USED

III. Expectations of Revenue and Acceptance

In this section, we discuss the dynamics of the user capacity model and user willingness model and how it affects revenue. For the purposes of our analysis, we have used a number of variables and terms. Table I lists our notation for the variables.

Intuitively, if we do not know how much customers are capable or willing to pay, it makes sense to charge a constant amount of money from each customer. This is because we have no means of predicting which customer to charge a high price and which a lower price. By choosing a constant price we maximize the chances that they accept. We have proven this intuition to be correct. The proof assumes that at high prices (all prices greater than an arbitrary price denoted by ψ_∞), the probability that a customer will purchase the service is zero. This is a reasonable assumption since every customer has only a finite capacity to pay. For instance, ψ_∞ can be assigned the asset value of the richest person on earth. This means that the expectation of the decision to purchase given price ψ, denoted by $E[\Upsilon \mid \psi]$, is zero for ψ greater than or equal to ψ_∞.

Theorem 1: If the expectation of the decision to buy given price ψ, $E[\Upsilon \mid \psi]$, is 0 $\forall\ \psi \geq \psi_\infty$, then the expectation of revenue per customer, $E[\gamma]$, is maximum when ψ is a constant.

Proof Outline: We shall assume that $E[\Upsilon \mid \psi]$ is defined $\forall \psi \in [0, \infty)$. Suppose that different prices are charged with a probability density p_Ψ. Then the expectation of revenue per customer is given by:

$$\begin{aligned} E[\gamma] &= \textstyle\int_0^\infty \psi E[\Upsilon \mid \psi] p_\Psi(\psi) d\psi \\ &= \textstyle\int_0^{\psi_\infty} \psi E[\Upsilon \mid \psi] p_\Psi(\psi) d\psi \\ &+ \textstyle\int_{\psi_\infty}^\infty \psi E[\Upsilon \mid \psi] p_\Psi(\psi) d\psi. \end{aligned} \quad (3)$$

Since $E[\Upsilon \mid \psi] = 0, \forall \psi \geq \psi_\infty$, the integral defined above is non-zero only over a finite interval $[0, \psi_\infty)$. Since $E[\Upsilon \mid \psi]$ is defined at all points in this interval, $\exists \psi_{max} \in [0, \psi_\infty)$ at which the function $\psi E[\Upsilon \mid \psi]$ is maximum. If there are many such points, we arbitrarily choose one of them. The expectation of the function $\psi E[\Upsilon \mid \psi]$ is maximized if the probability density at ψ_{max} is the highest. This will be the case when p_Ψ is the Dirac delta function $\delta(\psi - \psi_{max})$. In other words, the expectation of revenue is maximized when ψ has a constant value ψ_{max}. $\square$

We now derive an expression for $E[\Upsilon \mid \psi]$, for a Pareto distribution of customer capacities and the willingness model described Equation 2. We shall assume that the shape α of the Pareto distribution is greater than or equal to 1. This is because, when α is less than 1, Pareto distributions do not have a finite mean. $E[\Upsilon \mid \psi]$ denotes the mean rate at which customers accept the price ψ for the service. All customers are assumed to have the same willingness parameter, denoted by δ.

Theorem 2: For a Pareto distribution of customer capacities, with shape α and scale b, $\alpha \geq 1$, $b > 0$, and the customer willingness defined in Equation 2, the expectation of the variable Υ given price ψ, $E[\Upsilon \mid \psi]$ is as follows.

$$E[\Upsilon \mid \psi] = \begin{cases} 1 - \frac{\alpha}{\alpha+\delta}\left(\frac{\psi}{b}\right)^\delta & , \quad 0 \leq \psi \leq b \\ \frac{\delta}{\alpha+\delta}\left(\frac{b}{\psi}\right)^\alpha & , \quad \psi > b \end{cases} \quad (4)$$

Proof Outline: Let χ denote the capacity of a customer and $f_{\wp}(\chi)$ the probability density at χ. If the price $\psi \leq b$, then all customers have the capacity to pay that price. Therefore, the expectation of Υ, given ψ is:

$$\int_b^\infty f_{\wp}(\chi) P\{\Upsilon = 1 \mid \psi\} d\chi. \tag{5}$$

$$= \int_b^\infty \frac{\alpha b^\alpha}{\chi^{\alpha+1}} \left(1 - \left(\frac{\psi}{\chi}\right)^\delta\right) d\chi. \tag{6}$$

When the price $\psi > b$, then some of the customers do not have the capacity to pay this price. According to Equation 2, the probability of purchase is zero for all the customers whose capacity to pay is less than the price ψ. Therefore, when $\psi > b$, the expectation of Υ, given ψ is:

$$\int_b^\infty f_{\wp}(\chi) P\{\Upsilon = 1 \mid \psi\} d\chi. \tag{7}$$

$$\begin{aligned} &= \int_b^\psi f_{\wp}(\chi) 0 d\chi \\ &+ \int_\psi^\infty f_{\wp}(\chi) \left(1 - \left(\tfrac{\psi}{\chi}\right)^\delta\right) d\chi \end{aligned} \tag{8}$$

$$= \int_\psi^\infty \frac{\alpha b^\alpha}{\chi^{\alpha+1}} \left(1 - \left(\frac{\psi}{\chi}\right)^\delta\right) d\chi. \tag{9}$$

The integrals in Equations 6 and 9 can be evaluated using the standard rules of integration to obtain the expression in Equation 4. □

Theorem 2 tells us that the mean rate of acceptance is at least $\frac{\delta}{\alpha+\delta}$ if the price is less than b, and at most $\frac{\delta}{\alpha+\delta}$ if the price is greater than b.

Using the expression in Equation 4, for a Pareto distribution of customer capacities, the expected price acceptance rate, $E[\Upsilon \mid \psi]$, will be very small but non-zero even for very large values of ψ. Theorem 1 can therefore not be applied to this distribution of capacities. However, it can be shown that even for a Pareto distribution of capacities, the expectation of revenue is maximized for some constant price ψ_{max}. We omit the proof for reasons of space.

For a constant price ψ, the expected revenue per customer, denoted by $E[\gamma \mid \psi]$, is given by the product of the price ψ and the expectation of the decision to purchase, $E[\Upsilon \mid \psi]$. Note that $E[\Upsilon \mid \psi]$ denotes both the expectation of the decision to purchase for an arbitrary customer as well as the mean rate at which customers accept the price. Using Equation 4, the expression for revenue per customer can be written as:

$$E[\gamma \mid \psi] = \psi E[\Upsilon \mid \psi] \tag{10}$$

$$= \begin{cases} \psi - \psi \frac{\alpha}{\alpha+\delta} \left(\frac{\psi}{b}\right)^\delta & , \quad 0 \leq \psi \leq b \\ \psi \frac{\delta}{\alpha+\delta} \left(\frac{b}{\psi}\right)^\alpha & , \quad \psi > b \end{cases} \tag{11}$$

To find the price at which revenue is maximum, we differentiate the expectation of revenue given price, $E[\gamma \mid \psi]$, with respect to the price ψ, and solve the equation so obtained when the first derivative is zero. Using standard methods of calculus, it can be easily shown that the expecation of revenue per customer is maximum when ψ is equal to $\left[\frac{(\alpha+\delta)}{(\delta+1)\alpha}\right]^{\frac{1}{\delta}} b$. Figure 4 shows the expected revenue per customer as a function of price for different willingness elasticity parameters, a Pareto shape of 3.0 and scale 67. The price at which the expected revenue per customer is maximum is also shown. It can be verified that the expected revenue is indeed maximum when the price is $\left[\frac{(\alpha+\delta)}{(\delta+1)\alpha}\right]^{\frac{1}{\delta}} b$. Figure 4 also shows how the acceptance rate, $E[\Upsilon \mid \psi]$ varies with price, ψ, for the same capacity distribution. It can be verified that the expectation of revenue per customer is maximized when the acceptance rate is $\frac{\delta}{\delta+1}$.

IV. Maximum Revenue Under Resource Constraints

In the previous section we derived an expression for the expected revenue per customer. In this section, we discuss how in a FCFS system, the maximum expectation of total revenue is affected by resource constraints. We then derive an expression for the optimal price which maximizes total revenue under resource constraints.

We begin our formulation by defining system utilization for a FCFS system. System utilization is the

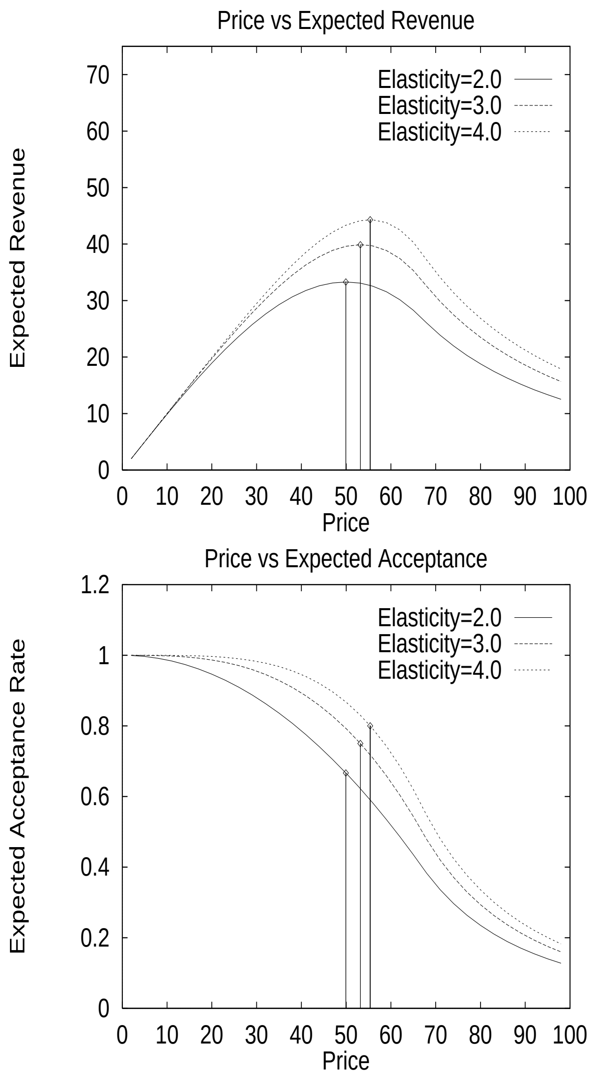

Fig. 4. Price, Expected Revenue and Expected Acceptance Rate

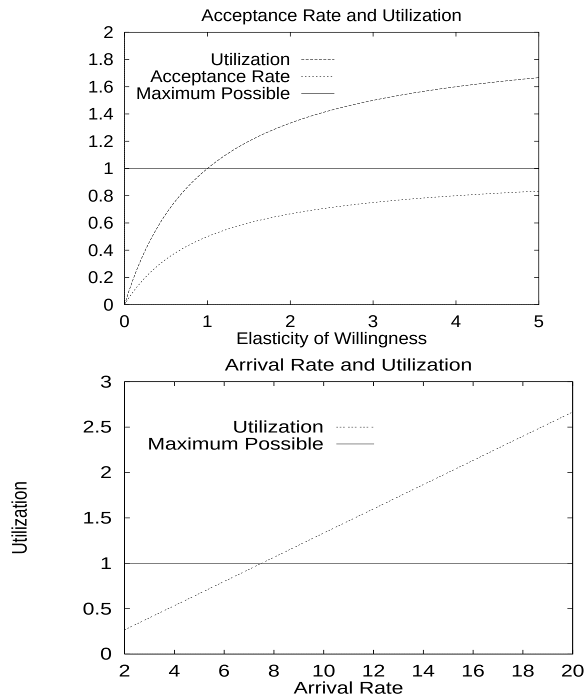

Fig. 5. Price, Arrival Rate, Acceptance Rate and System Utilization

relative fraction of time for which the channels are busy servicing requests. Let λ be the arrival rate of requests, n the number of channels, and d, the average time to serve a request. Then, the system utilization, ρ, is defined as the ratio of the number of requests entering the system per unit time to the number of serviced requests exiting the system per unit time. The mathematical expression for system utilization, when we charge a price ψ, is given by:

$$\rho(\psi) = \frac{\lambda E[\Upsilon \mid \psi] d}{n} \qquad (12)$$

Note that the system utilization is bounded above by 1. This upper bound reflects resource constraints. We now discuss how this upper bound on system utilization can affect the expectation of total revenue. Let ψ_{max} be the optimal price that maximizes the expected revenue per customer. $E[\Upsilon \mid \psi_{max}]$ is the rate at which customers accept the price ψ_{max}. Effectively, in a time interval t, $\lambda t E[\Upsilon \mid \psi_{max}]$ customers enter the system. Thus the maximum expectation of overall revenue in time t is given by $\psi \lambda t E[\Upsilon \mid \psi_{max}]$. However, on the average, the maximum number of requests that can be served by a system with n channels in a time interval t is $\frac{nt}{d}$, where d is the mean service time. When $\lambda E[\Upsilon \mid \psi_{max}]t$ is greater than $\frac{nt}{d}$, there are more requests entering the system than can be accomodated. Note that the system utilization is greater than 1 when this condition occurs. In this case the maximum expectation of revenue derived above cannot be achieved due to resource constraints. This phenomenon is illustrated in Figure 5. Figure 5 shows how the acceptance rate, $E[\Upsilon \mid \psi_{max}]$, and the system utilization for price ψ_{max}, ρ_{max}, vary with the elasticity of willingness, δ, for a system with 500 channels, a mean service time of 100min and an arrival rate of 10 requests/min. Also shown, is how ρ_{max} varies with the arrival rate for the same system when the elasticity, δ, is 2.0. The figure indicates that resource constraints may prevent achieving the maximum expectation of revenue when either the arrival rate is high, or willingness to pay is very high.

Intuitively, we should increase the price when there is high demand. In fact, we should increase the price to such an extent that only as many customers accept the price as can be accomodated by the system. We show that this maximizes total revenue. In our proof, we use the notion of concave functions. Concave functions have only a global maxima and no local maxima. A function is concave with respect to a variable, if its second derivative with respect to that variable is non positive. The revenue function derived in this paper is an example of such a function. It has only a global maxima and no local maxima. This means that as we increase price, the revenue increases upto a certain point (the global maxima) after which it decreases monotonically.

Theorem 3: If number of customers accepting the optimal price exceeds the system capacity, then the revenue is maximized for the highest price which results in maximum predicted system utilization.

Proof Outline: We derive a simple expression for the expectation of revenue per unit time, $E[\Gamma \mid \psi]$, in terms of system utilization ρ and price, ψ:

$$\begin{aligned} E[\Gamma \mid \psi] &= \lambda E[\Upsilon \mid \psi]\psi && (13) \\ &= \lambda \rho(\psi) \frac{n}{d} \frac{1}{\lambda} \psi && (14) \\ &= \rho(\psi) \psi \frac{n}{d} && (15) \end{aligned}$$

Since $\frac{n}{d}$ is a constant, the expectation of revenue is proportional to the product of the system utilization and the price ψ. The system utilization is a monotonically decreasing function of price. The expectation of revenue is a concave function of price because its second derivative with respect to ψ is nonpositive. When the number of customers accepting the optimal price ψ_{max} exceeds system capacity, i.e., $\lambda E[\Upsilon \mid \psi_{max}] > \frac{n}{d}$, $\rho(\psi_{max})$ is greater than 1. In this case, we want to find that price ψ at which $E[\Gamma \mid \psi]$ is highest and $\rho(\psi)$ less than or equal to 1. Since $\rho(\psi)$ is monotonically decreasing, this means that system utilization will decrease when we increase the price. In addition, since $E[\Gamma \mid \psi]$ is concave, it is monotonically decreasing for all ψ greater than ψ_{max}. Therefore, to the right of ψ_{max}, both $\rho(\psi)$ and $E[\Gamma \mid \psi]$ are decreasing. Clearly, the highest revenue while ensuring that $\rho(\psi)$ is less than or equal to 1 is earned when $\rho(\psi)$ is equal to 1. □

Using Theorem 3, when the arrival rate exceeds system capacity, the maximum revenue is earned when the predicted acceptance rate is $\frac{n}{d\lambda}$. This is obtained by substituting a value of 1 for system utilization in Equation 12 and solving for $E[\Upsilon \mid \psi]$. Using this acceptance rate, we can compute the price that needs to be charged using Theorem 2. We summarize this in the following theorem, which we state without proof.

Theorem 4: Let customer capacities be Pareto distributed with shape α, $\alpha > 1$, and scale b. Let their willingness to pay be as defined in Equation 2. Consider a system with n channels serving content with mean service time d. Let λ be the request arrival rate. The expectation of revenue for the system is maximum when the content provider charges a price ψ_{MAX} defined as follows:

$$\psi_{MAX} = \begin{cases} \left[\frac{\alpha+\delta}{(\delta+1)\alpha}\right]^{\frac{1}{\delta}} b & , \quad \frac{\delta}{\delta+1} \leq \frac{n}{d\lambda} \\ \left[\left(\frac{\alpha+\delta}{\alpha}\right)\left(1-\frac{n}{d\lambda}\right)\right]^{\frac{1}{\delta}} b & , \quad \frac{\delta}{\delta+1} > \frac{n}{d\lambda} \geq \frac{\delta}{(\alpha+\delta)} \\ \left[\frac{\lambda d\delta}{n(\alpha+\delta)}\right]^{\frac{1}{\alpha}} b & , \quad \frac{\delta}{\delta+1} > \frac{\delta}{(\alpha+\delta)} > \frac{n}{d\lambda} \end{cases} \quad (16)$$

In summary, the analytical framework developed in this section enables us to set an optimal price based on the characteristics of customer behavior, the system resource constraints and the request arrival process.

V. Discussion

In this section, we place our framework in the context of the current Internet infrastructure and service model. Our framework focusses on a video-on-demand system, but is sufficiently general to be applied to other forms of content. A key issue relates to service class differentiation. How can the framework be extended to handle multiple classes of service? Different classes of service will target different classes of customers with different capacities and willingness to pay. Thus, we can model this scenraio by considering different user capacity distributions, with different shapes and scales—one for each class of service provided. Corresponding to these classes, the elasticity of willingness parameter, δ, will also be different.

Another advantage of using different elasticity of willingness parameters is that, it will account for popularity of content. Customers who want hot movies will be more willing to pay more money. Thus, the willingness elasticity parameter for such customers will be high. Similarly, customers requesting cold movies can be modelled with a small elasticity of willingness.

A drawback of the model is that the exact values of the elasticity of willingness for different classes of service or movies may not be known. Similarly, the exact values of the parameters of the customer capacity distribution are difficult to obtain. To address this problem, one can develop algorithms which perform probability experiments to estimate these parameters. We have developed such algorithms in some of our later work[5], [6].

Another potential drawback of our framework is that we assume the all customers subscribing to the same class of service or movie have the same willingness parameter. We are currently developing a model where the customers willingness parameter is exponentially distributed. Our analysis indicates that the results obtained using this model will not be significantly different.

VI. Conclusions

In this paper, we focussed on pricing models for a content delivery system. We introduced the idea of probabilistic user behavior and analyzed its impact on revenue. We developed a theoretical framework for maximizing revenue assuming a FCFS content delivery system. Specifically, we developed pricing strategies for conditions of high demand as well as for low demand. An interesting application of Theorem 2, which relates price and acceptance rate, is to use price to control system utilization. An eventual goal of our work is to allow customer and provider to negotiate the price. Theorem 4 provides a baseline for controlling the negotiation. The impact of batching, content popularity, and temporal changes in user behavior need to be studied in greater detail.

Acknowledgements

I would like to thank my advisor, Dr. Kevin C. Almeroth for the useful discussions I had with him. I would also like to thank J. Nayak, S. Regunathan, K.S. Prashanth and A. Agarwal of the Signal Processing Laboratory, University of California Santa Barbara, for their helpful comments and insights.

References

[1] A. Dan, D. Sitaram, and P. Shahabuddin, "Scheduling policies for an on-demand video server with batching," in *ACM Multimedia*, (San Francisco, California, USA), October 1994.

[2] T. Little and D. Venkatesh, "Prospects for interactive video-on-demand," *IEEE Multimedia*, pp. 14–23, Fall 1994.

[3] B. C. Arnold, *Pareto Distributions*. Burtonsville, Maryland: International Co-operative Publishing House, 1983.

[4] E. L. Crow and K. Shimuzu, *Lognormal distrbutions: theory and application*. New York: Dekker, 1988.

[5] S. Jagannathan and K. C. Almeroth, "An adaptive pricing scheme for content delivery systems," in *submitted for review*.

[6] S. Jagannathan and K. C. Almeroth, "The dynamics of price, revenue and system utilization," in *submitted for review*.