Internet Service Tiering as a Market Segmentation Strategy

Qian Lv, George N. Rouskas
Department of Computer Science, North Carolina State University, Raleigh, NC 27695-8206, USA

Abstract— **We consider Internet broadband access as an elastic service whose value varies across segments of the user population. We show that introducing multiple tiers of service can be an effective market segmentation strategy that can lead to an increase of profits for the ISP. We also develop an efficient dynamic programming algorithm for the problem of determining optimally both the service tiers and their prices. Our approach provides new insights into the selection and pricing of Internet tiered services, and our results indicate that exponential tiering structures adopted by ISPs are far from optimal.**

I. Introduction

Historically, packet-switched computer networks, including the Internet and legacy networks based on ATM or Frame Relay technologies, are designed to be *continuous-rate*. Theoretically speaking, continuous-rate networks may allocate bandwidth at very fine granularities; for instance, one client may request a rate of 98.99 Megabit per second (Mbps), while another customer may ask for 99.01 Mbps. Clearly, the option of requesting arbitrary rates offers clients maximum flexibility in utilizing the available network capacity.

On the other hand, supporting bandwidth allocation at such extremely fine granularity may seriously complicate the operation and management of the network. Based on the above example, the network provider faces the problem of designing mechanisms to distinguish between the two rates (i.e., 98.99 Mbps vs. 99.01 Mbps) and enforce them in an accurate and reliable manner. However, the task of differentiating between the two users on the basis of these two rates may be extremely difficult, or even impossible to accomplish for traffic of finite duration, undermining the network's ability to support important functions such as robust traffic policing or accurate customer billing. Furthermore, link capacity across a continuous-rate network may become fragmented, posing significant challenges in terms of traffic engineering.

In practice, most network operators have developed a variety of *tiered service* models in which users may select only from a small set of service *tiers* (levels) which offer progressively higher rates (bandwidth). The main motivation for offering such a service is to simplify a wide range of core functions (including network management and equipment configuration, traffic engineering, service level agreements, billing, and customer support), enabling the providers to scale their operations to hundreds of thousands or millions of customers. Returning to the previous example, a tiered-service network might assign both users requesting 98.99 Mbps and 99.01 Mbps to the next higher available rate, say, 100 Mbps. In this case, there is no need to handle the two customers' traffic differently; also, the network operator only needs to supply policing mechanisms for a small set of rates, independent of the number of users.

This work was supported by the NSF under grant CNS-0434975.

Consider a network that offers a service characterized by a single parameter, e.g., the bandwidth of the user's access link. A tiered-service network is one that offers p levels (tiers) of service, where typically p is a small integer, much smaller than the number n of (potential) network users (i.e., $p \ll n$). Let $Z = \{z_1, z_2, \cdots, z_p\}$ denote the set of service tiers offered by the network provider. Without loss of generality, we make the assumption that the service tiers are distinct and are labeled such that $z_1 < z_2 < \cdots < z_p$. Users are limited to only these p tiers, and may subscribe to any tier depending on their needs and their willingness to pay the corresponding service fee. In particular, z_1 is the minimum and z_p the maximum amount of service that a user may receive. In the case of residential Internet access, for instance, z_1 may correspond to a minimum bandwidth for the service to be considered "broadband," while z_p may correspond to the capacity of the access link, e.g., as determined by limitations imposed by ADSL technology.

According to this definition, traditional telephone networks and transport networks based on SONET/SDH technology belong to the class of tiered-service networks. Indeed, such networks allocate bandwidth in discrete tiers that are multiples of the slot size in the underlying TDM system.

Current tiered service offerings by major ISPs can be broadly classified in two categories based on the tiering structure. The structure of service tiers targeted to business customers is based on the bandwidth hierarchy of the underlying transport infrastructure (e.g., DS-1, DS-3, OC-3, etc.), While this is a natural arrangement for the service provider, it is unlikely that hierarchical rates designed decades ago for voice traffic would be a good match for today's business data applications. The second class employs *exponential tiering* structures in which each tier offers twice the bandwidth of the previous one, as exemplified by the various ADSL tiers (e.g., 384 Kbps, 768 Kbps, 1.5 Mbps, 3 Mbps, etc.) available through several ISPs. While such simple tier structures may be an appropriate choice for marketing purposes, the relationship between these exponentially increasing levels of service (and their price) and the usage patterns (and ability to pay) of the population of potential subscribers is open to debate.

An early study by Lea and Alyatama [3] investigated the benefits of "bandwidth quantization" in packet-switched networks. In their terminology, "bandwidth quantization" refers to sampling the continuous range of possible rates to select a small set of discrete bandwidth levels (tiers) that are made available to users. This work presented a heuristic based on simulated annealing to obtain a sub-optimal set of discrete bandwidth levels of service. The main contribution of this study was to demonstrate for the first time that this benefit comes almost for free, as even with a sub-optimal set of tiers the performance degradation (e.g., in terms of call blocking) compared to a continuous-rate network is negligible. In [6] we showed that the problem considered by Lea and Alyatama in [3] can in fact be solved optimally, and we presented a sophisticated optimal algorithm of linear complexity in the context of MPLS networks.

In more recent work [4] we developed an economic model for reasoning about and pricing Internet tiered services. In that work, we considered a market scenario in which all users receive the same value from the service offered by the network operator, or equivalently, all users are characterized by the same utility function $U(x)$. A market in which all users value a service (or product) similarly is said to be *inelastic* [2]. Certain essential goods (e.g., gasoline or milk) and services that everyone needs tend to be inelastic, at least in the short term. Markets for most other products and services tend to be *elastic*, in that their value may be perceived quite differently across the population of consumers. Hence, in elastic markets, consumer behavior with respect to pricing may vary widely depending on the underlying utility curve that characterizes the specific consumer.

In this paper we consider broadband Internet access as an elastic service. Specifically, we assume that users are partitioned into classes, each class characterized by a distinct utility function, and we study the problem of selecting jointly the set of service tiers and their prices so as to maximize the profit (i.e., *provider surplus* [1]) of the ISP. The paper is organized as follows. In Section II we introduce a model of user diversity. For the special case of a single tier we develop in Section III an optimal algorithm to determine both the level of service to be offered and its price. In Section IV we show that introducing multiple tiers can be an effective market segmentation strategy that may lead to an increase in profits. We present performance results in Section V, and we conclude the paper in Section VI.

II. Economic Model of User Diversity

We consider the market for broadband Internet access with one ISP and multiple users. The service of the ISP is described by the access speed x, with x taking values in the interval $[x_{min}, x_{max}]$, where x_{min} and x_{max} correspond to the lowest and highest speed, respectively, that the ISP may offer. The cost to the ISP of providing an amount x of service is given by the cost function $C(x)$. The ISP offers a tiered bandwidth service with p tiers. Let $Z = \{z_1, \ldots, z_p\}$ denote the set of distinct service tiers, labeled such that $z_1 < \ldots < z_p$. We also let $P(z_j)$ denote the price the ISP charges for tier z_j. Price is an increasing function of service x, hence, $i < j$ implies $P(z_i) < P(z_j)$. For notational convenience, we assume the existence of a "null" service tier z_0 for which $P(z_0) = 0$, and also that $P(z_{p+1}) = \infty$; the latter ensures that no user may receive service at an amount higher than the highest tier.

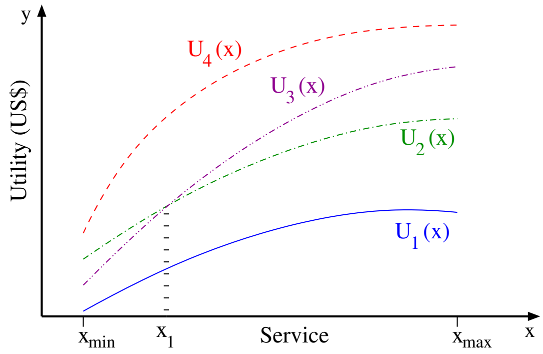

Fig. 1. Diversity of user utility functions, $T = 4$ classes of users

Users belong to one of T classes, $T > 1$. Users in class t are characterized by utility function $U_t(x)$. We express cost, price, and utility in the same units, e.g., US$. The various utility curves indicate the users' willingness to pay, and can be determined using market research tools such as surveys or conjoint analysis [5]. Fig. 1 illustrates the user diversity with respect to utility curves for $T = 4$ classes. We let $f_t, t = 1, \ldots, T$, denote the fraction of the user population that is in class t; obviously, $f_1 + \ldots + f_T = 1$. We also make the reasonable assumption that the cost $C(x)$ and utility functions $U_t(x), t = 1, \ldots, T$, are continuous, twice differentiable, and non-decreasing in the interval $[x_{min}, x_{max}]$.

If the price set for a product is *below* the value of this product to a consumer, then the consumer will purchase the product. On the other hand, if the price of the product is *higher* than the consumer's perceived value of the product, then they will not make the purchase. Consequently, users in class t will subscribe to the highest tier z_j for which the price charged does not exceed its value $U_t(z_j)$ to the users. More formally, given the utility and price functions, there is an implied mapping $h : \{1, \ldots, T\} \rightarrow Z$ from the set of user classes to the set of tiers, where $h(t) = z_j$ if and only if:

$$P(z_j) \leq U_t(z_j) < P(z_{j+1}), t = 1, \ldots, T; j = 0, 1, \ldots, p. \quad (1)$$

Note that, if the price of the lowest tier is higher than the utility of some class of users, then, from (1) these users are forced to "subscribe" to the "null" service tier z_0, which implies that they will not use the service. Fig. 2 illustrates the mapping of $T = 2$ classes of users to $p = 5$ service tiers based of the given price structure imposed by the step pricing function $P(x)$. Specifically, users in class 1 and class 2 are mapped to tiers z_4 and z_2, respectively, consistent with expression (1).

From the point of view of the ISP, there is a clear tradeoff in setting the price for the service tiers. If the price for some tier is high, the ISP will lose revenue as some customers may decide to subscribe to a lower tier or not use the service at all. On the other hand, if the ISP prices the tiers conservatively, it may attract some low-utility customers, but may also forego a significant amount of revenue from customers with high utility

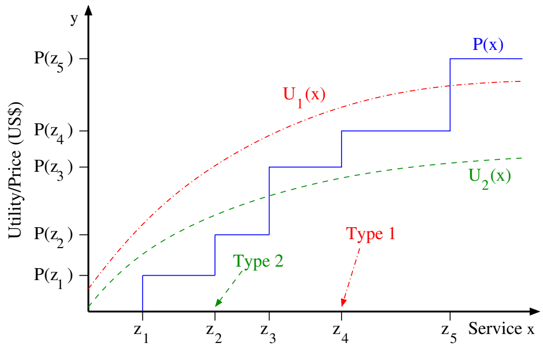

Fig. 2. Mapping of $T = 2$ classes of users to $p = 5$ service tiers based on the price structure $P(x)$

who would be willing to pay more for the service. Therefore, the objective is to select jointly the p service tiers to be offered and their prices so as to maximize the provider surplus. We will refer to this problem as the *surplus maximization (MAX-S)* problem, and formally define it as follows.

Problem 2.1 (MAX-S): Given the cost function $C(x)$, an integer T, the fraction f_t of users in class t and their utility $U_t(x), t = 1, \ldots, T$, and the domain $[x_{min}, x_{max}]$ of the cost and utility functions, find a set $Z = \{z_1, \ldots, z_p\}$ of p service tiers and their respective prices $P(z_j)$ that maximize the following objective function (provider surplus):

$$S_{pr}(z_1, \ldots, z_p) = \sum_{j=1}^{p} [P(z_j) - C(z_j)] \sum_{h(t)=z_j} f_t \quad (2)$$

under the constraints ($t = 1, \ldots, T$; $j = 0, \ldots, p$):

$$h(t) = z_j \text{ iff } P(z_j) \leq U_t(z_j) < P(z_{j+1}) \quad (3)$$

$$0 = z_0 < x_{min} \leq z_1 < z_2 < \ldots < z_p \leq x_{max} \quad (4)$$

The following lemma states that the price of each service tier in the optimal solution may take one of T distinct values.

Lemma 2.1: Let $Z = \{z_1, \ldots, z_p\}$ be an optimal solution to the MAX-S problem and $P(z_j), j = 1, \ldots, p$, be the price of the corresponding tiers. Then:

$$\exists\, t \in \{1, \ldots, T\} : P(z_j) = U_t(z_j), \quad j = 1, \ldots, p. \quad (5)$$

Proof. By contradiction. Assume that the price of tier z_j is such that $U_s(z_j) < P(z_j) < U_t(z_j)$ for some classes $s \neq t$. In other words, class-t users (and, perhaps, users of some class q with $U_q(z_j) > U_t(z_j)$) subscribe to tier z_j, whereas users of class s and any class r with $U_r(z_j) < U_s(z_j)$ do not subscribe to tier z_j. Therefore, the price of the tier can be raised to $P'(z_j) = U_t(z_j)$ without affecting the set of users subscribing to this or any other tier. This increase in price will result in an increase to the provider surplus, contradicting the assumption about optimality of the original price $P(z_j)$. ∎

III. The Single Tier Case

Let us first consider the simpler case of $p = 1$, i.e., a single level of service z_1. Due to Lemma 2.1, we know that $P(z_1) = U_t(z_1)$ for some t, and our goal is to determine an optimal value for $z_1 \in [x_{min}, x_{max}]$ and a corresponding optimal price. To this end, we distinguish two cases.

Case 1. The T utility functions $U_t(x), t = 1, \ldots, T$, and the cost function $C(x)$ do not pairwise intersect anywhere in their domain $[x_{min}, x_{max}]$. Without loss of generality, we make the assumption that $C(x)$ lies below all of the T utility functions in the same interval. If that is not true, we can ignore the utility functions that lie below $C(x)$ and only consider the $T' < T$ functions that lie above $C(x)$. Doing so will not affect optimality, since setting the price below cost will result in a loss for the provider.

Now let us relabel the T utility functions such that:

$$C(x) < U_1(x) < U_2(x) < \ldots < U_T(x), \forall x \in [x_{min}, x_{max}],$$

and define $F_t = \sum_{s=t}^{T} f_s, t = 1, \ldots, T$, as the fraction of users falling in the classes with utilities equal to, or higher than, that of class t.

If the provider offers a single tier in the amount of z_1^t and prices it according to the corresponding utility of class-t users, then we can write the provider surplus from (2) as:

$$S_{pr}^t(z_1^t) = F_t \left[U_t(z_1^t) - C(z_1^t)\right], \quad t = 1, \ldots, T. \quad (6)$$

Therefore, we can find the optimal tier $z_1 \in [x_{min}, x_{max}]$ and its price using these two steps:

1) For each class t, determine the value of $z_1^t \in [x_{min}, x_{max}]$ that maximizes the quantity $S_{pr}^t(z_1^t)$ in (6).
2) Set the tier z_1 to the quantity z_1^q, and its price to $U_s(z_1^q)$, where q is such that $S_{pr}^q(z_1^q)$ is maximum among the T quantities computed in Step 1.

Case 2. Some of the utility and cost functions pairwise intersect in one or more points within their domain $[x_{min}, x_{max}]$. In this case, it is always possible to partition this interval into sub-intervals within which none of the functions intersect. Returning to Fig. 1, we can see that the domain of the utility functions can be divided into two sub-intervals, $[x_{min}, x_1]$ and $[x_1, x_{max}]$, within which the functions do not intersect. Therefore, we can obtain the optimal value for z_1 and the corresponding price by following the following steps:

1) Divide the interval $[x_{min}, x_{max}]$ into K non-overlapping sub-intervals $e_k, k = 1, \ldots, K$, such that no two utility or cost functions intersect within each sub-interval e_k.
2) For each sub-interval e_k, find the optimal value $z_1^q(k)$ and optimal price $U_q(z_1^q(k))$, as in Case 1 above.
3) Set the tier z_1 and its price to the corresponding values for the interval e_k with the maximum provider surplus $S_{pr}^q(z_1^q(k))$ among all the intervals in Step 2.

Based on the above discussion, in order to find the optimal solution to the MAX-S problem for $p = 1$ tier, we need to determine the maximum of the provider surplus function in expression (6) in any sub-interval $[x_1, x_2]$ of $[x_{min}, x_{max}]$. Let us define function $g(x)$ for some class t as:

$$g(x) = U_t(x) - C(x), \quad x \in [x_1, x_2]. \quad (7)$$

Since the utility and cost functions are continuous and twice differentiable throughout their domain, then function $g(x)$

is continuous and twice differentiable in any sub-interval $[x_1, x_2]$. Therefore, we can find its maximum as follows:

1) The second derivative $g''(x) \leq 0$ everywhere in $[x_1, x_2]$. Then, $g(x)$ is concave in the sub-interval, and its maximum can be found by solving the equation $g'(x) = 0$.
2) The second derivative $g''(x) \geq 0$ everywhere in $[x_1, x_2]$. Then, $g(x)$ is convex in the sub-interval, and its maximum values occur at either x_1 or x_2.
3) The second derivative changes sign in the sub-interval. In this case, we subdivide $[x_1, x_2]$ into intervals such that the second derivative g'' is either non-negative or non-positive everywhere in the smaller intervals. We obtain the maximum of $g(x)$ within each smaller interval according to either case 1 or case 2 above, from which we can select the overall maximum in $[x_1, x_2]$.

IV. The Multiple Tier Case: Market Segmentation

Let us now return to the general case of a tiered service with $p > 1$ tiers. Such a service can be viewed as a *market segmentation* strategy [7], whereby the ISP splits the market into several segments with the goal of increasing profitability. A typical example of market segmentation is when providers offer a "premium" service at a high price for the high end of the market, and a "standard" service at a lower price for the rest of the market. An important issue that arises in the market segmentation is process is determining how to segment the market and how to differentiate among the services to be offered to the various segments so as to maximize profitability. A tiered service and price structure obtained as a solution to the MAX-S problem resolves this issue since the tiers and corresponding prices uniquely identify the market segments that optimize the provider surplus (profit).

We also note that market segmentation follows the law of diminishing returns [7] in that, after an initial increase in profits, creating an additional market segment may have a negligible effect in overall profitability. Therefore, the number p of market segments (service tiers) will, in general, be less than the number T of user classes, especially if T is relatively large. In other words, an optimal market segmentation strategy may combine several user classes into a single segment. On the other hand, because of Lemma 2.1, in an optimal solution to the MAX-S problem the price of each tier $z_j, j = 1, \ldots, p$, is equal to the utility $U_t(z_j)$ of some class t. Therefore, the solutions we develop are for the general case $p \leq T$.

For simplicity, in the remainder of this paper we make the assumption that the T utility curves do not intersect anywhere in the domain $[x_{min}, x_{max}]$ and are labeled such that $U_1(x)$ is the lowest and $U_T(x)$ the highest curve. This is a reasonable assumption, since if some user A values an amount of service x_1 more than a user B, then an amount $x_2 > x_1$ of service is likely to have more value for A than for B. On the other hand, the cost function $C(x)$ may intersect with some of (or all) the utility curves within the interval $[x_{min}, x_{max}]$, but may intersect at most once with a given utility function.

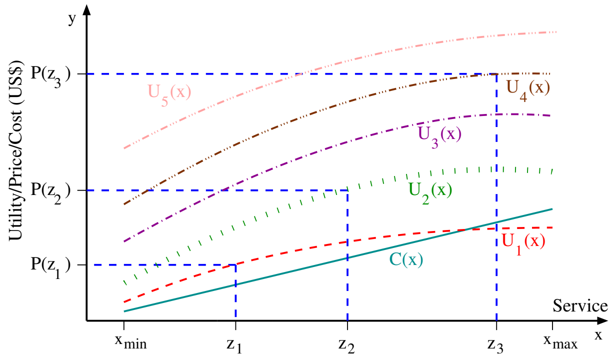

Fig. 3. Feasible price structure for an instance of MAX-S with $T = 5$ classes of users and $p = 3$ fixed service tiers

A. The MAX-S Problem with Fixed Tiers

Let us first consider a restricted version of the MAX-S problem in which the p tiers are provided as input to the problem and are not subject to optimization. This problem variant arises naturally for the simple uniform and exponential tiering structures we discuss in Section V. Given the minimum and maximum service levels, x_{min} and x_{max}, respectively, and the number p of tiers, the p service levels are completely specified under uniform and exponential tiering, hence only the prices of these levels need to be optimized.

Due to the assumption that utility curves do not intersect and that price is an increasing function of service x, a feasible solution to MAX-S is such that, for $j = 1, \ldots, p-1$:

$$P(z_j) = U_s(z_j) \text{ and } P(z_{j+1}) = U_t(z_{j+1}) \Rightarrow s < t. \quad (8)$$

as illustrated in Fig. 3. Therefore, we can obtain an optimal price structure for the MAX-S problem with p fixed tiers using the dynamic programming algorithm described next.

Let $\Lambda(t, l)$ denote the optimal provider surplus when there are t classes of users and l service tiers. For an instance of MAX-S with T classes and p tiers, we can compute $\Lambda(T, p)$ recursively using the following expressions (recall that the quantity F_t was defined in the previous section):

$$\Lambda(t, 1) = \max_{1 \leq s \leq t} \{F_s [U_s(z_1) - C(z_1)]\}, t = 1, \ldots, T \quad (9)$$

$$\Lambda(t, l+1) = \max_{s=l,\ldots,t-1} \{\Lambda(s, l) + F_{s+1} [U_{s+1}(z_{l+1}) - C(z_{l+1})]\}$$
$$l = 1, \ldots, p-1, \; t = 2, \ldots, T. \quad (10)$$

Expression (9) states that, if there is a single service tier fixed at z_1, the price is set to the utility that maximizes the provider surplus (refer also to the definition of provider surplus for a single tier in expression (6)). The recursive expression (10) is derived from the observation that if the price of the $(l+1)$-th tier is set to the utility of the $(s+1)$-th class, then all users in this class and all classes of higher utility will subscribe to this tier. The second term within the brackets in the right-hand side of (10) represents the contribution of this tier to the provider surplus. The first term in brackets in the right-hand side of (10) represents the optimal surplus for s

classes of users and l tiers, $r \geq l$. Taking the maximum over all values of s yields the overall maximum. The running time complexity of this algorithm is $O(pT^2)$.

B. Approximate Solution to the MAX-S Problem

We now turn our attention to the original version of the MAX-S problem whereby both the service level at each tier and its price are subject to optimization. We solve this problem approximately by employing a discretization technique. Specifically, we divide the interval $[x_{min}, x_{max}]$ into $K > T$ segments of equal length, and impose the additional constraint that the p tiers, $z_1, \ldots, z_p$, may only take values from the set $\{e_k, k = 1, \ldots, K\}$, where $e_k = x_{min} + \frac{k(x_{max} - x_{min})}{K}$ is the right endpoint of the k-th interval. As $K \to \infty$, this discrete version of MAX-S approaches the original version in which $z_j, j = 1, \ldots, p$, are continuous variables.

Let $\Delta(k, t, l)$ denote the optimal solution to this discrete version of MAX-S when there are k points, t classes, and l tiers. We can compute $\Delta(K, T, p)$ recursively as follows:

$$\Delta(k,t,1) = \max_{1 \leq m \leq k} \left\{ \max_{1 \leq s \leq t \leq m} \{F_s [U_s(e_m) - C(e_m)]\} \right\}$$
$$k = 1, \ldots, K, \ t = 1, \ldots, T, \ t < k \qquad (11)$$

$$\Delta(k,t,l+1) = \max_{m=l,\ldots,k-1} \left\{ \max_{s=l,\ldots,t-1} \{\Delta(m,s,l) + \max_{r=m+1,\ldots,k} \{F_{s+1} [U_{s+1}(e_r) - C(e_r)]\} \} \right\}$$
$$l = 1, \ldots, p-1, \ t = 2, \ldots, T, \ k = 2, \ldots, K, \ t < k \qquad (12)$$

When there is only one service tier, it is placed at some endpoint e_m and its price is set at the utility of some class s that maximizes the provider surplus, hence we have expression (11). In the general case of k points, t classes, and $l+1$ tiers, the optimal value is obtained by (1) considering the best placement and pricing of l tiers in $m < k$ points and $s < t$ classes, given by $\Delta(m, s, l)$, in which case the best placement and price for tier $(l+1)$ is given by the second line of (12), and (2) then taking the maximum over all possible values of m and s, yielding the recursive expression (12). The running time complexity of this algorithm is $O(pT^2K^3)$.

We have conducted a large number of experiments (omitted due to space constraints) which indicate that $K = 100$ is sufficient for the algorithm to converge; hence we use this value in the performance study we present in the next section.

V. Numerical Results

To evaluate the performance of tiered service as a market segmentation strategy, we consider the market for broadband Internet access. We let the minimum service x_{min} =256 Kbps and the maximum service x_{max} =12 Mbps, roughly corresponding the range of broadband speeds in the U.S. For all instances of the MAX-S problem we investigate in this study, we assume the existence of $T = 50$ classes of users characterized by the family of utility curves:

$$U_t(x) \ = \ \lambda_t x^\gamma \log(x), \quad t = 1, \ldots, T = 50, \qquad (13)$$

we use the linear cost function $C(x) = \mu x$, and we set the values for parameters μ, γ and λ_t to:

$$\mu = 0.3, \ \gamma = 0.5, \ \lambda_t = 10 + .1(t-1), \ t = 1, \cdots, T = 50,$$

such that the T utility curves do not intersect in the domain [256 Kbps, 12 Mbps] and are labeled from lowest to highest.

We consider three distributions of users into classes:

- a *uniform* distribution, in which each class contains an equal fraction of the user population: $f_t = \frac{1}{T}$,
- an *increasing* distribution, such that the fraction of users in a given class increases with utility: $f_t = ct$, where $c = \frac{1}{1275}$ is a constant that ensures that $\sum_{t=1}^{T} f_t = 1$, and
- a *decreasing* distribution, in which the fraction of users in a given class increases with utility: $f_t = c(T + 1 - t)$, where $c = \frac{1}{1275}$.

A. Tier Structure Comparison

We compare the performance of four tiered structures:

1) **Optimal:** the tiered structure obtained from the dynamic programming algorithm (11)-(12).
2) **Optimal-rounded:** the tiered structure derived from rounding the values of the optimal tiers to the nearest multiple of 256 Kbps.
3) **Uniform:** the p tiers are spread uniformly across the domain [256 Kbps, 12 Mbps].
4) **Exponential:** the p tiers divide the domain [256 Kbps, 12 Mbps] into intervals that double in length (from left to right).

The uniform and exponential are simple solutions similar to structures employed by major ISPs in which the $p > 1$ service tiers are completely defined; hence, their prices were optimized using the approach we described for fixed tiers in Section IV-A. For the optimal structure, on the other hand, we obtained both the $p > 1$ service levels and their prices using the dynamic programming algorithm in Section IV-B. However, for $p = 1$, we obtained the optimal service level and its price following the algorithm in Section III, and we use this value for the curves of *all four* tiering structures.

The three Figs. 4-6 plot the provider surplus against the number p of tiers and correspond to the decreasing, uniform, and increasing distribution of users into classes, respectively. The three figures show four curves, each corresponding to one of the four tiering structures. As we can see, the tiering structure (referred here as "optimal") obtained from the approximate solution to the MAX-S problem and the corresponding optimal-rounded structure outperform the uniform and exponential tiering structures across the range of values for p and across the user distributions into classes. Therefore, network providers would benefit by applying the dynamic programming solutions to determine the tiered structures to offer. Furthermore, although the uniform and exponential tiering structures uniquely define the various tiers to be offered, the prices for these tiers are determined by the dynamic programming algorithm (9)-(10) so as to optimize the provider

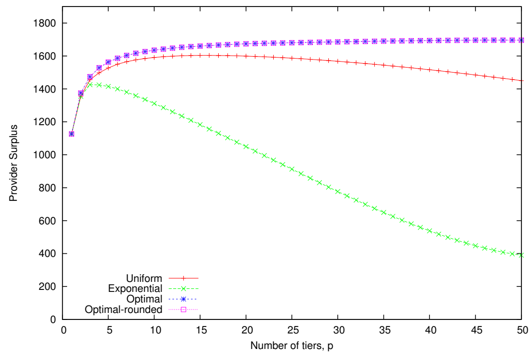

Fig. 4. Tier structure comparison, decreasing distribution of users, $T = 50$

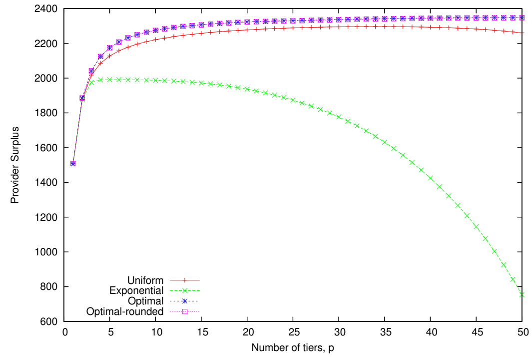

Fig. 6. Tier structure comparison, increasing distribution of users, $T = 50$

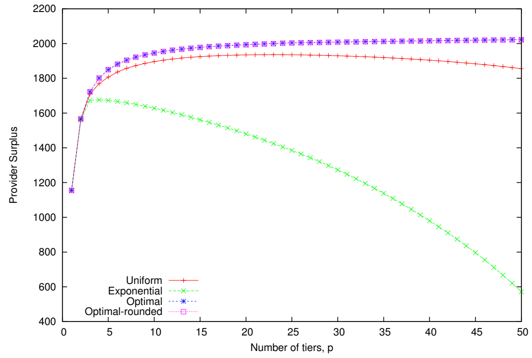

Fig. 5. Tier structure comparison, uniform distribution of users, $T = 50$

surplus for the given tiers. Any other *ad hoc* pricing scheme would result in a lower surplus, hence even for these structures the providers would benefit from the tier pricing methodology presented earlier.

There are several more important observations we make from these three figures. First, it is clear that the curves for the optimal and optimal-rounded structures increase rapidly with p initially, but flatten out once the number of tiers increases beyond $p = 10-15$; the latter behavior implies that additional tiers provide diminishing returns beyond a point. This result is consistent with economic theory which predicts that there is a limit to the benefits that can be achieved by segmenting a market; it also provides further confirmation to the thesis of our earlier work [6] that a relatively small number of service tiers is sufficient to capture most of the benefits of tiering.

We also observe that the exponential tiering structure performs poorly overall, and that its curves reach a well-defined maximum: the surplus achievable under such structure peaks at a small value of p and starts to decline rapidly thereafter. This behavior can be explained by noting that most of the tiers in an exponential structure are grouped together at the leftmost part of the service domain $[x_{min}, x_{max}]$ =[256 Kbps, 12 Mbps], and the few tiers that cover the remaining of the interval do not provide fine enough granularity to capture the benefits of market segmentation. The curves corresponding to the uniform tiering structure are below those for the optimal and optimal-rounded structures, but higher than those for exponential tiering. These results indicate that tiering structures with equally spaced tiers would be better for the service provider than exponential ones. Furthermore, we can see that the uniform tiering curves also reach a maximum at a certain value of p that depends on the user distribution, and start to decline as p increases further. This behavior demonstrates that simply adding more tiers but placing them into specific points in the domain of the service is not an effective market segmentation strategy; hence, to achieve the maximum benefits of market segmentation the service provider must optimize both the size and price of each tier.

Finally, we note that the overall provider surplus increases as we move from the decreasing distribution of users into classes (Fig. 4) to the uniform distribution (Fig. 5) to the increasing distribution (Fig. 6). This is expected, since the fraction of users characterized by high utility functions (i.e., willing to pay higher prices) is lowest for the decreasing distribution and highest for the increasing distribution. As a result, the surplus that the provider is able to extract through market segmentation is higher in the latter case.

VI. Concluding Remarks

We have investigated tiered service as a market segmentation strategy for increasing ISP profits under the assumption that consumer behavior with respect to pricing varies across the user population. We developed an efficient dynamic programming algorithm for determining optimally both the service tiers and their prices. Our approach provides new insight into the selection and pricing of Internet tiered services.

References

[1] R. Bade and M. Parkin. *Foundations of Microeconomics*. Addison-Wesley, 2nd edition, 2004.
[2] K. E. Case and R. C. Fair. *Principles of Economics*. Pearson Education, 7th edition, 2004.
[3] C-T. Lea and A. Alyatama. Bandwidth quantization and states reduction in the broadband ISDN. *IEEE/ACM Trans. Netw.*, 3:352-360, June 1995.
[4] Q. Lv and G. N. Rouskas. An economic model for pricing tiered-service networks. In *Proceedings of IEEE ICC 2009*, June 2009.
[5] L. Moutinho and A. Meidan. Quantitative methods in marketing. In *M. J. Baker (Editor), The Marketing Book*, pages 197–244. Butterworth-Heinemann, 2002.
[6] G. N. Rouskas and N. Baradwaj. A framework for tiered service in MPLS networks. *Proc. of IEEE INFOCOM 2007*, pages 1577–1585, May 2007.
[7] A. Weinstein. *Market Segmentation*. Probus Publishing Co., 1993.