Assessing Network Service Profitability: Modeling From Market Science Perspective

Jin Xiao, *Student Member, IEEE*, and Raouf Boutaba, *Senior Member, IEEE*

***Abstract*—Network service providers regularly conduct network planning and upgrade processes to keep their businesses profitable. The effectiveness of a network upgrade/planning decision is intrinsically tied to the ability of a provider to retain and grow its customer population. This paper examines the crucial linkage between network performance, customer satisfaction and profitability of network service, and presents an analytical modeling approach from market science perspective. We derive a generalized forecasting model that projects service profitability from the underlying network service infrastructure and the subscriber population. Through simulation studies and analysis, we show how such approach captures key factors and trends influencing service profitability and how it can significantly improve current network planning and upgrade processes.**

***Index Terms*—Economics, network applications and services, network design and planning.**

I. Introduction

FACING a consumer market with rising demands for quality and impending saturation, network service providers (ISPs) are struggling to keep their customers satisfied and their businesses profitable. In the network service industry, network planning and upgrades are regularly exercised to achieve this objective. The practice is mostly ad hoc, where investment decisions are made based on past experiences and "rule of thumb" estimations. The lack of formal methodology can be attributed to the large process gap between the network planners and the business analysts. The network planners strive to improve network performance via fine tuning and optimizing upgrade decisions. Very little concern is given to the profitability of the resulting investments. And the business analysts have a very coarse understanding of how improved network performance can lead to future revenue generation. Better network performance directly translates to more profit is a common assumption. Considering the intricate relations among network operations, customer behaviors, and market dynamics that jointly influence service profitability, such an assumption is overly naive. A general and comprehensive analytical model linking networks, customers, and market environments to service profitability is then extremely beneficial and timely. Although research in market science and economics presents many insightful observations and empirical studies on service utility, customer behavior, and profitability, they remain descriptive and incomplete. This lack of formalization prevents the integration of key customer and market factors in network planning and upgrade analysis and produces network upgrade decisions that do not reflect customer behaviors and service dynamics, and do not give good service profitability estimates.

In this paper, we present a modeling approach relating the performance delivered by a network service infrastructure to the satisfaction of its customers and consequently to the network service provider's profit. We show that network upgrade and planning strategies should be made in accordance to their influences on customer satisfaction and the resulting changes in customer behavior. Compared with existing network research, we find that service profitability is not only dependent directly on the pricing of the service and indirectly on the performance of the service, "colored" by the customer's subjective evaluation, but is also dependent on the number of customers willing to repurchase the service and the new customers the service provider can attract. Customers' access behaviors, QoS sensitivities, service expectations, past experiences, service competitiveness and market growth are some of the key factors influencing the customers repurchase intentions and consequently a ISP's revenue generation. The prospectives on our work are presented in Section II and III.

Based on influential theories in economics and market science, we show that there is a strong ground for the derivation of well-behaved mathematical models linking the network service performance, the customer behavior and the market dynamics to profit. Henceforth we construct generalized mathematical models that formalize these theories, with parameters reflecting various network service characteristics, customer attributes, and market conditions. Through analysis and simulation, we demonstrate the applicability and effectiveness of our approach in ISP upgrade/planning operations. We find that rather than maximizing service utility, it is more important for ISPs to ensure that the service quality meets the customers' expectations.

The rest of this paper is organized as follows: Section II presents a summary of current industry practices and academic research. Section III presents our modeling approach and its rationale, while Section IV details the construction of a forecasting model following our approach. Section V analyzes the forms of our perception function and the impact of model parameters, followed by case studies and simulations in Section VI. Section VII concludes with final remarks and future prospectives.

II. Prospectives and Literature Works

In conducting formal analysis of investment decisions, it is well understood that the soundness of a decision is dependent on

Manuscript received April 29, 2005; revised May 6, 2006; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor Z. Zhang.

The authors are with the School of Computer Science, University of Waterloo, Waterloo, ON, Canada N2L 3G1 (e-mail: j2xiao@cs.uwaterloo.ca; rboutaba@uwaterloo.ca).

Digital Object Identifier 10.1109/TNET.2007.896535

the soundness of the analytical model and the value of the analyzed data. In the context of network services, we are presented with a rich reservoir of network information, ranging from statistical information gathered from Management Information Base [1] to active end-to-end measurements (e.g., pinging). Advanced tools are even capable of tracking individual traffic flows. Traditional customer management processes (e.g., customer relation management) gather vast amount of customer information in the form of customer surveys, service usage, trouble-ticket logs, etc.

For a network service provider, its customers, the sole source of revenue, are the crucial link between network performance and service profitability. The satisfaction of a customer is strongly influenced by the service performance he/she receives from the underlying network infrastructure, and influences his/her intention to reuse the service. A number of market studies on Telecom service operators world-wide have confirmed the existence of these relationships [2]–[4].

Some works in network research [5], [6] have recognized the importance of analyzing both the customer profile and the network information in a business decision process. However, the means of correlating the two aspects are missing [5] and there is no method for mapping network performance to service utility [6]. Other works on network upgrades [7], [8] give coarse treatment to customer satisfaction with no consideration for customer service perception or shifts in consumer market dynamics. In our past work [9], we have attempted to relate service performance to customer satisfaction and then to future revenue through simple linear mappings. They do not capture the intricacies among customer satisfaction, repurchase intentions and market dynamics.

Pricing is an important factor of service profit because it maximizes the monetary benefit a service provider can draw from its customers. Works on service charge maximization [10], [11] and usage based charging [12], [13] focus on how to best generate profit from customers' service usage. Our work investigates another important factor of profit: customer population. By studying the cause and effect of customer satisfaction, we bring focus and structure to some of the key factors influencing customer retention and growth. Our work is also complementary to works on bandwidth provisioning and network dimensioning [14] by assessing their impact on the customers.

Customer relations and profitability have been the subject of significant research in market science and economics. The well known expectancy-disconfirmation theory [15], [16] relates service utility to customer satisfaction, based on the classic adaptation theory from psychology. The work views expectation as an adapted reference point for the customers, upon which satisfaction is the result of customer value judgement on expectation and perception. Later finding [17] suggests strong relationships among satisfaction, perceived quality, and disconfirmation. Anderson and Sullivan [18] follow up on these works with a descriptive model relating service quality to customer repurchase intention. However, their work remains qualitative and does not address the issue of expectation adjustment and market dynamics. Bolton [19] formalizes the influence of customer satisfaction on customer retention and increased sales volume. The model is refined and tested over a 22-month period with cellular customers. The linkage between expectancy and customer retention is coarsely treated in this work and the impact of divergent service quality on customer experience is not considered. To obtain service utility, SERVQUAL [20]–[22] is the most used and proven model in market science. Traditionally, the model has been applied to the service industries. Recent study [3] has shown the application of SERVQUAL model in capturing customer quality perception of China's Telecommunication services.

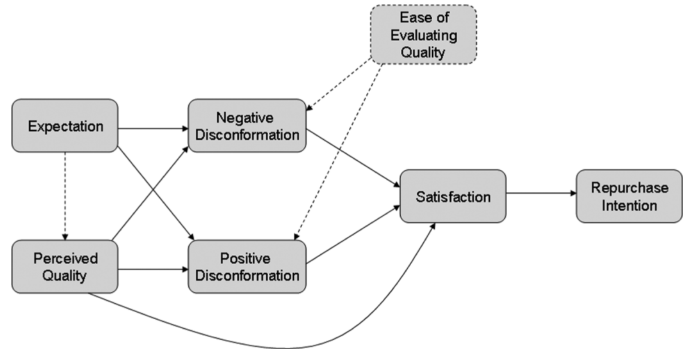


Fig. 1. Customer satisfaction model (Anderson and Sullivan).

With the vast amount of existing conceptual and empirical results from market science and economics, we believe there is a strong foundation for deriving an analytical model for network service operations. We propose a methodology for formalizing the relationships between network performance, customer satisfaction, and service profitability. The approach covers the computation of utility for network services, the derivation of customer satisfaction based on service utility, and the projection on service profitability from customer repurchase intentions and market dynamics. Following this approach, we construct a network service specific model capable of forecasting service profitability induced by network infrastructure improvements.

III. A Market Science Methodology

In this section, we show how network performance, customer satisfaction and service profitability are related in market science research, and present our modeling methodology. A key driver of our approach is the well-established expectancy-disconfirmation theory [15], [16] which relates expectation, perceived quality and disconfirmation to customer satisfaction. The perceived quality refers to the service utility a customer obtains from service usage, while expectation represents the expected utility a customer formulates before using the service. Disconfirmation is then the discrepancy between the expectation and the perceived quality. Anderson and Sullivan [18] refined this theory in a customer satisfaction framework (Fig. 1).

They consider disconfirmation to have a positive and a negative component that are influenced by expectation and perceived quality. The customer satisfaction is then a function of perceived service quality and both components of disconfirmation. The perceived quality is affected by expectation based on the observation: when the difference between expectation and perceived quality is small, customer tends to equate perception to expectation. Furthermore, the level of disconfirmation is pos-

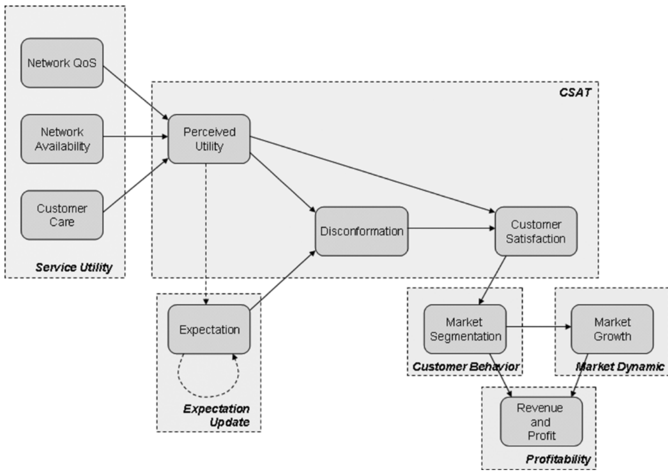


Fig. 2. Modelling service profitability.

itively related to ease of evaluating quality. For network services, the ease of evaluating quality is high as service quality can be readily measured based on the network performance and the application requirements. Hence there is very little ambiguity in customer's perception of quality, and we simplify away this factor in our customer satisfaction relationships. Furthermore, their claim that expectation influences perception is controversial as a number of important findings [19], [23]–[25] supports the theory that perceived quality influences expectation via a dynamic update process. Based on these works, we reverse Anderson and Sullivan's expectation and perception relationship and formulate an expectation update process.

Fig. 2 presents our modeling approach. Our view of the customer satisfaction model (CSAT) is a modified Anderson-Sullivan model. The expectation model updates a customer's future service expectation based on past expectations and current utility perception through a recurrent process in our expectation update model. The CSAT model takes as input the service utility, referred to as the "antecedent" of customer satisfaction [18]. It is computed through a utility model that operates on network performance and service attributes. Specialized from the SERVQUAL model, we consider three aspects of network services: service quality, service availability and customer care. The "consequence" of customer satisfaction is customer's intention to repurchase [18]. It is captured in our customer behavior model, with regard to market competition and customer desire, to assess subscriber population change via a Bayesian decision process. The output of the customer behavior model is an estimation of the market segmentation: service provider retention, competitor retention, churn, and turnover from competitors. In our market dynamic model, the attractiveness of the service to new entry customers is projected using the Bass growth model [26]. The service profitability is then computed based on the revenue generating potential, derived from the market segments, and the service cost. Since network services are continuous where customers make periodic repurchase decisions (e.g., monthly for xDSL services), the entire process can be iterated through consecutive decision periods, providing long-term profitability forecasts of network service operations.

We believe a model developed from this market science methodology provides significantly better assessment on the impact of network performance on network service profitability, compared with the simple linear models used in network planning research today. In the context of this paper, we make the following assumptions: the network service market is an open market with multiple competitors; the customer is rational in his/her purchase decisions and does not exit the service market; all competitors of the network service market charge similar price, have identical technology attractiveness from the customer's perspective and employ similar advertisement strategies; and the pricing for the service is flat rate subscription based.

IV. Forecasting Service Profitability

In this section, we detail the construction of our analytical model. We first introduce the computation of utility based on network service performance (Section IV-A) and then construct the customer satisfaction model (Section IV-B), followed by a formalization of the expectation update process (Section IV-C). Formulating the outcome as a decision problem, we estimate the market segmentation in customer behavior model (Section IV-D) and then deduce the growth of new entry customers in market dynamics model (Section IV-E). Finally, service profitability is computed as a function of revenue and cost (Section IV-F).

A. Service Utility and Perceived Utility

As noted by Dabholkar [27], customer satisfaction and utility are not the same construct. Satisfaction is a customer's subjective evaluation of the service performance, while utility is its objective measurable quantification. There are two concepts of utility presented in this section: service utility and perceived utility. Service utility denotes a set of service related performance metrics that are measurable or observable. Together, these metrics yield a single quantitative evaluation of utility: the perceived utility. We first discuss service utility and its computation.

The SERVQUAL model [21] identifies tangibles, empathy, assurance, responsiveness and reliability as the five major aspects of service quality. In the context of network services, tangibles, empathy, assurance and responsiveness can be grouped together under customer care, including helpdesk support, installation, troubleshooting, billing service, on-call technical support, etc. Reliability is readily mapped to network service availability, often regarded as an essential factor in network service contract. Empirical studies done in Telecom services from Germany, US, and China [3], [4] confirm the applicability of SERVQUAL to network services and suggest network quality as an additional aspect of SERVQUAL. In accordance, we consider service utility U as consisting of three basic aspects: service quality, service availability, and customer care. The three service aspects are further documented by the TeleManagement Forum (TMF) in its SLA handbook suite [28].

To compute service quality Q, we consider factors related to network QoS of the customer's traffic flows, the application requirements, and the customer's own preferences. Let a *service path* denote an end-to-end network path carrying a customer's service traffic running a particular application. For a service path j of customer i, the service quality Q_{ij} is computed by considering a networked application to belong to one of two categories: QoS-sensitive services and QoS-insensitive services. QoS-sensitive services are applications whose satisfactory performance is contingent on fulfilling certain QoS requirements. On the other hand, QoS-insensitive services are applications that do not have specific QoS requirements. Their performance is best computed based on an overall measurement of throughput quality.

For QoS-sensitive services, we model Q_{ij} based on the concept of *defective service instances* (DSI). We define a defective service instance experienced by a customer i on a service path j, denoted by D_{ij}, as a series of consecutive network QoS measurements whose values are below the QoS requirements of the supported application. A network flux parameter is also introduced to account for performance fluctuations [9]. Let A_{ij} be the total access time of customer i on service path j, and $l(D_{ij})$ be the time length of D_{ij}, then Q_{ij} takes on the following form:

$$Q_{ij} = \frac{A_{ij} - \sum l(D_{ij})}{A_{ij}} \quad \text{for QoS-sensitive services} \tag{1}$$

For QoS-insensitive services, we model Q_{ij} based on the average throughput (P^a_{uij} for upload and P^a_{dij} for download) and the maximum bandwidth (P^o_{uij} and P^o_{dij} respectively). The maximum bandwidth is the capability limit of the customer's service offering (e.g., 32 Mb/s customer download ceiling for xDSL service). Let γ_1 and γ_2 represent the download and upload performance preferences of the customer, then Q_{ij} is computed as:

$$\begin{aligned} Q_{ij} &= \gamma_1 \frac{P^a_{dij}}{P^o_{dij}} + \gamma_2 \frac{P^a_{uij}}{P^o_{uij}} \\ \gamma_1 + \gamma_2 &= 1, \quad \text{for QoS-insensitive services} \end{aligned} \tag{2}$$

Given the above formulation of service utility, we can now define the computation of perceived utility. Let W_{ij} denote the percentage of time a service path j of customer i is deemed available, C_i be a scalar rating (between 0 and 1) of customer care service for customer i, SP_i be a set of service paths customer i uses, and α_1, α_2, and α_3 be customer i's weight preferences for service quality, service availability, and customer care respectively, the perceived utility for customer i is expressed as:

$$U_i = \frac{\alpha_1 \sum_j (Q_{ij} \times A_{ij})}{\sum_j A_{ij}} + \frac{\alpha_2 \sum_j (W_{ij} \times A_{ij})}{\sum_j A_{ij}} + \alpha_3 C_i \quad \text{where} \quad j \in SP_i \; and \; \alpha_1 + \alpha_2 + \alpha_3 = 1 \tag{3}$$

Taking as input the network and service performance of customer i's service paths, (1), (2) and (3) yield the perceived utility of customer i, normalized between 0 and 1. This is a unified quantification of the service utility according to customer's service preference and serve as the input to the CSAT model, described in Section IV-B. The utility model described here is theoretical. Work in [9] gives a pragmatic framework on how the above computations can be performed in practice. Furthermore, Section VI also provides some demonstration on how such computation could be carried out in regional networks.

B. Customer Satisfaction (CSAT)

Customer satisfaction can be modeled through the interaction between perceived utility and expectation [15], [16], expressed as a linear combination of a perception function and a disconfirmation function. Let f_1 be the perception function, f_2 be the disconfirmation function, U_{pi} and U_{ei} be the perceived utility and expected utility (i.e., expectation) of customer i, then the general form of customer satisfaction Γ_i for a customer i is given [18] as:

$$\Gamma_i = f_1(U_{pi}) + f_2(U_{pi} - U_{ei}) \tag{4}$$

The perception function gives the baseline customer satisfaction obtained from service utility, while the disconfirmation function modifies this satisfaction value based on the discrepancy between perceived utility and expectation (i.e., disconfirmation). In the subsections below, we derive general mathematical forms for the perception and the discoformation functions.

1) The Perception Function: The perception function f_1 is a mapping between perceived utility and baseline customer satisfaction. It is described in [18] as an increasing concave function starting at the origin (i.e., $f_1(0) = 0$). Its general shape is conceived based on the observation that as the utility increases, the customer becomes less sensitive to changes in utility. We express the rate of change of the perception function as:

$$f_1''(x) = \mu_2 x - \mu_1 \quad \text{where } x = U_{pi}, \; and \; \mu_1, \; \mu_2 \geq 0 \tag{5}$$

The parameters μ_1 and μ_2 control the concavity of the perception function. In Section V, we will discuss our choice of this particular form f_1''. Twice integrating (5) yields:

$$f_1(x) = \int\int f_1''(x)dx = \frac{\mu_2}{6}x^3 - \frac{\mu_1}{2}x^2 + \Psi x + C \quad \text{where} \quad x = U_{pi}, \; and \; \mu_1, \; \mu_2 \geq 0 \tag{6}$$

The constraint $f_1(0) = 0$ yields $C = 0$. We note that the domain of f_1 is bounded between 0 and 1. Moreover, we would like $f_1''(x)$ to be non-positive and $f_1'(x)$ to be non-negative for all possible values of x, and control the maximum value of f_1 via parameter ω_p (i.e., $f_1(1) = \omega_p$). We thus have the following set of constraints on parameters of f_1:

$$\begin{cases} \mu_1 \geq \mu_2 & \text{from } f_1''(x) \leq 0 \\ \Psi \geq \mu_1 - \frac{\mu_2}{2} & \text{from } f_1'(x) \geq 0 \\ \Psi = \frac{\mu_1}{2} - \frac{\mu_2}{6} + \omega_p & \text{from } f_1(1) = \omega_p \\ \mu_1, \mu_2 \geq 0, \omega_p > 0 \end{cases} \tag{7}$$

The solution to the above set yields the constraint on μ_1:

$$0 \leq \mu_1 \leq 2\omega_p + \frac{2}{3}\mu_2 \tag{8}$$

Equation (8) suggests that the upper bound of μ_1 is positively related to the upper bound of μ_2. As we would like the concavity parameter μ_1 to have the largest possible value range, and given the constraint $\mu_1 \geq \mu_2$, then we obtain the maximum value range of μ_1 when $\mu_1 = \mu_2$. This leads to a desirable simplification of f_1. Fig. 3 demonstrates the general characteristics of the

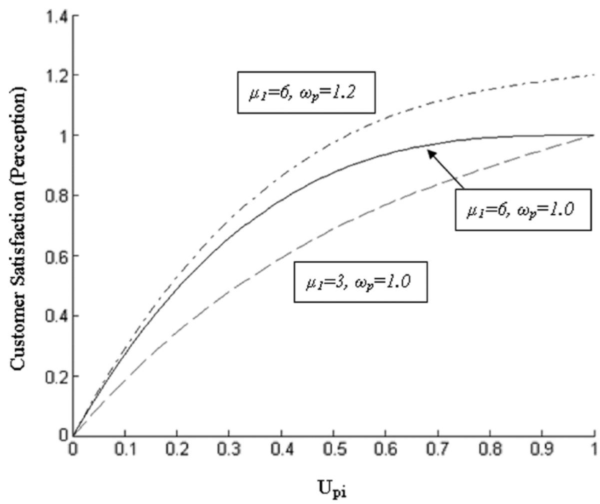


Fig. 3. The perception function.

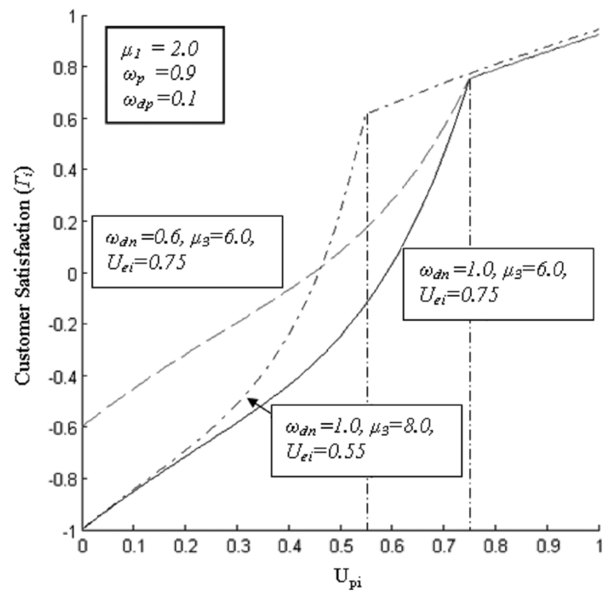


Fig. 4. The customer satisfaction function.

perception function. In summary, f_1 is a function of perceived utility, with the following form:

$$f_1(x) = \frac{\mu_1}{6}x^3 - \frac{\mu_1}{2}x^2 + \left(\frac{\mu_1}{3} + \omega_p\right)x$$
$$\text{where} \quad \mu_1 \leq 6\omega_p,\ \mu_1 \geq 0,\ \text{and}\ \omega_p > 0 \quad (9)$$

2) The Disconfirmation Function: The disconfirmation function f_2 accounts for the subjectivity of customer evaluation given a reference point (i.e., expectation). Tversky and Kahneman [29] found that "losses relative to a reference value looms larger than gains". Grounded on this psychological theory, Anderson and Sullivan [18] suggest that customer satisfaction is mildly increasing when perceived utility exceeds expectation and is significantly reduced when perceived utility falls below expectation. We formalize this interaction as a two-piece increasing function:

$$f_2(x) = \begin{cases} \omega_{dp}x & x \geq 0 \\ \omega_{dn}(x+1)^{\mu_3} - \omega_{dn} & x \leq 0 \end{cases}$$
$$\text{where}\ x = U_{pi} - UTIL_{ei}, \mu_3 \geq 0,\ \text{and}\ \omega_{dp}, \omega_{dn} > 0 \quad (10)$$

We note that the domain of f_2 is bounded between -1 and 1. The function is continuous (i.e., the two piece-wise functions converge at $x = 0$). The parameter ω_{dp} controls the maximum value of f_2 (i.e., maximum positive disconfirmation) while ω_{dn} controls the minimum value of f_2 (i.e., maximum negative disconfirmation). μ_3 regulates the impact of negative disconfirmation on customer satisfaction.

3) Customer Satisfaction: From (4), (9) and (10), we observe that Γ_i is bounded between $-\omega_{dn}$ and $\omega_p + \omega_{dp}$. In general, the choice of ω parameters should follow: $\omega_p \geq \omega_{dn} > \omega_{dp}$. ω should be fixed for all customers of a service and ω_{dp} should be small compared to ω_p.

As we observe in Fig. 4, the rate of change in customer satisfaction differs significantly when perceived utility falls below and exceeds expectation. The rate and severity of dissatisfaction (controlled by ω_{dn} and μ_3 respectively) reflect different customer's tolerance to negative disconfirmation. Our formalization of the customer satisfaction fits a rational customer's subjective evaluation of the service utility, and conforms to empirical findings [15], [16], [18], [19]. In addition, we offer a set of well-defined control parameters to fit different service characteristics, and individual customer's preferences and sensitivities.

C. Expectation Update

Empirical studies [25], [30] suggest that a customer adjusts his/her future expectation of service utility based on current expectation and perception. The studies also find favorable disconfirmation increases future expectation while unfavorable disconfirmation has the opposite effect. Through an expectation update process, we deduce a customer's future expectation as a function of the customer's current expectation and disconfirmation, subject to two psychological factors: assimilation and experience. When the relative level of disconfirmation is small, a customer tends to equate the perceived utility to the expected utility, due to assimilation effect [19]. Furthermore, as a customer perceives consistent service utility over time, he/she gains experience with the service, and consequently is less sensitive to short-term utility fluctuations [25]. In other words, the customer gradually establishes long-term reputation of the service.

Let κ_a be the assimilation factor, a customer i's future expectation U_{ei}^* has the following form:

$$U_{ei}^* = \begin{cases} U_{ei} & \left|\frac{U_{pi}-U_{ei}}{U_{ei}}\right| \leq \kappa_a \\ h(U_{ei})[U_{pi} - U_{ei}] + U_{ei} & \text{otherwise} \end{cases}$$
$$U_{ei}^* \text{ is constrained to } 0 \leq U_{ei}^* \leq 1 \quad (11)$$

The parameter κ_a is constrained $(0 \leq \kappa_a \leq 1)$ and should be a very small value (e.g., 0.01). The function $h(U_{ei})$ adjusts the expectation as a factor of the disconfirmation. Our general form of (11) is established based on assimilation theory of economics [24], where new information are assimilated as an aggregate quantity over time. According to works of [19], [25], [29], (11) should exhibit three characteristics. First, given the same expectation, a negative disconfirmation is weighed much more heavily than a positive disconfirmation [29]. This effect is similarly reflected in the construct of disconfirmation function. Second, a positive disconfirmation has a greater impact on U_{ei}^* as U_{ei} decreases, and conversely a negative disconfirmation has a greater impact on U_{ei}^* as U_{ei} increases [19]. Third, the longer a customer experiences consistent utility, the less impact on expectation should a short-term utility fluctuation have [25]. Based on these characteristics, we construct $h(U_{ei})$ as such:

$$h(U_{ei}) = \begin{cases} \beta_G \beta_M^m (1 - U_{ei}) & \frac{U_{pi} - U_{ei}}{U_{ei}} > \kappa_a \\ \beta_L \beta_M^m U_{ei} & \frac{U_{pi} - U_{ei}}{U_{ei}} < -\kappa_a \end{cases}$$
$$\text{where} \quad 0 \le m \le \Upsilon, 0 \le \beta_M \le 1, \text{ and } 1 \le \beta_G < \beta_L \quad (12)$$

β_G and β_L are the positive and negative disconfirmation factors respectively. β_M is the memory factor and m the memory length. The term β_M^m controls the significance of the new information (i.e., current disconfirmation) on the aggregate (i.e., expectation). As m increases, β_M^m decreases. We use integer values for m, representing the number of repurchase evaluations the customer underwent while using the service. The constant Υ represents the maximum memory length a customer keeps track of. The value of m is updated (m^*) based on the following equation:

$$m^* = \begin{cases} m + 1 & \left| \frac{U_{pi} - U_{ei}}{U_{ei}} \right| \le \kappa_a \\ m - 1 & \text{otherwise} \end{cases}$$
$$m^* \text{ is constrained to } 0 \le m^* \le \Upsilon \quad (13)$$

The initial value of m is set to 0. We infer from (13) that as customer regularly experiences consistent service performance, he/she is more insulated from short-term performance fluctuations. Conversely, when performance significantly fluctuates over time, the customer is unable to make an experienced evaluation of the service, and hence his/her reliance on new information does not diminish with time (i.e., the divergence effect).

D. Customer Behavior: Repurchase Decision and Market Segmentation

Repurchase intention is the direct consequence of customer satisfaction [18]. Research in inter-temporal planning (e.g., [23]) state that customers re-estimate purchase decisions periodically based on previous estimates and new information. Furthermore, there exists a strong linkage among customer satisfaction, future expectation, and repurchase intention [18], [19]. We formulate the customer's repurchase intention as a decision problem whereby the customer's decision to use a service from a particular service provider is primarily influenced by the customer's current level of satisfaction and the expected future utility. The finding of [30] suggests that when a customer chose a service brand that meets his/her desire, he/she is likely to choose the same service brand again regardless if the brand has the highest expected performance in the market or not. Therefore, we consider a customer i will stay with a service provider if his/her customer satisfaction at the end of the current service period is above such a threshold Γ_i^D. If below Γ_i^D, the action of choosing a new service provider is a decision problem in which the customer attempts to maximize his/her future satisfaction based on his/her future expectations of similar services. Let U_{eiv} be the future expectation of service v estimated by a customer i, let k be the service customer i has just used, let Φ be the set of all similar services, and let κ_r be the resistance factor of customer i, then we can express the decision problem as:

$$max\,\{f_1(U_{eik}) + \kappa_r, f_1(U_{eiv})\} \quad \forall(v \in \Phi, v \ne k) \quad (14)$$

The parameter κ_r is a small satisfaction modifier representing the extra effort (e.g., service switching time, etc.) customer i has to spend in order to switch service provider. Eq. (14) relies on precise knowledge of a customer's future expectations. In practice, a customer's expectation of services he/she has not used can at best be estimated from service reputation with some uncertainty. Hence, we reformulate (14) as a Bayesian decision problem [31]. Let $F_{ev}(\mu_{ev}, \sigma_{ev})$ be the probability distribution of expectation of service v, with mean μ_{ev} and standard deviation σ_{ev}, our decision problem can be expressed as:

$$max \left\{ f_1(U_{eik}) + \kappa_r, \int f_1(\mu_{ev}) dF_{ev}(\mu_{ev}, \sigma_{ev}) \right\}$$
$$\forall(v \in \Phi, v \ne k) \quad (15)$$

In (15), $f_1(\mu_{ev})$ is the loss function and $F_{ev}(\mu_{ev}, \sigma_{ev})$ is the prior distribution. Given overall customer turnover rate in the market, (15) could also be used in a random sampling process to forecast service switching decisions of customers from other competitors. Ultimately, applying this decision process to all consumers in the market classifies the consumer population into three disjoint partitions: the set of customers with intention to repurchase the same service k (Ω_R), the set of customers choosing not to use service k (Ω_D), and the set of customers switching to service k from another service provider (Ω_P).

E. Market Dynamic

Up to this point, we have considered the partitioning of the existing consumer market. The entry of new consumers in the market could be described by the Bass growth model [26]. This model is applicable to network service industry, as suggested by the techno-economic studies on European xDSL market penetration [32]. The Bass model categorizes new consumers that enter the market into two categories: innovators and imitators. The innovators enter the market without any incentives and they are the main consumer faction during the inception of the market; the imitators are attracted to the market by the innovators and they are the main consumer faction as the market matures. The hazard function of the Bass model, describing the conditional probability of new consumers entering the market, is formally expressed [26] as:

$$\frac{f(T)}{1 - F(T)} = p + qF(T)$$
$$\text{where} \quad 0 < p < q < 1 \text{ and } p + q = 1 \quad (16)$$

$f(T)$ is the probability density function over time T, while $F(T)$ is the cumulative function over T. The parameter p is the coefficient of innovators and q is the coefficient of imitators. In general, p is much smaller than q.

Let S be the market potential (i.e., the maximum number of consumers), and $L(t)$ be the mapping function that maps real time t to time domain T of the Bass model, then we can represent the number of entry customers that choose service k as:

$$\Omega_N = \frac{\Omega_R + \Omega_P}{\Omega_R + \Omega_P + \Omega_D} S\left[F\left(L(t_{c+1})\right) - F\left(L(t_c)\right)\right] \quad (17)$$

The time value t_c denotes the end of current service period (i.e., current evaluation time), while t_{c+1} denotes the time of next evaluation. The last term in (17) is the cumulative probability of new consumers entering the market from current time to next evaluation time. We estimate that a fraction of them will choose service k based on the competitiveness of service k at current time. This is represented by the first term of (17).

In the above discussion, we have considered a single service market. For ISPs that has multiple service offerings, a market dynamic should be established per service. We will show an example of this in Section VI.

F. Service Profitability

From our forecast of consumer market segmentations at time t_c, the revenue generating potential R_k of service k in $[t_c, t_{c+1}]$ time interval is:

$$R_k = (\Omega_N + \Omega_P) \times \xi_N + \Omega_R \times \xi_R \quad (18)$$

The parameters ξ represent the price of service k to new customers ξ_N and old customers ξ_R in time interval $[t_c, t_{c+1}]$. The profitability of service k is then:

$$PROF_k = R_k - COST_k - PEN_k \quad (19)$$

The parameters $COST_k$ and PEN_k are the cost of running service k and the estimated monetary penalties (e.g., due to contract violation, etc.) from time t_c to time t_{c+1}.

V. Model Analysis

In Section IV-B, we have formalized the customer satisfaction function. In the context of network services, we now discuss our particular choice of the perception function and analyze the impact of the parameters in the model.

A. Choice of the Perception Function

In constructing the perception function, we also considered two other simple equation forms ((20) and (21)). Both of them are concave increasing functions in the domain of 0 to 1.

$$f_1(x) = \omega_p x^{\mu_1} \quad [0 \le \mu_1 \le 1, \omega_p \ge 0, 0 \le x \le 1] \quad (20)$$

$$f_1(x) = 1 - e^{-\mu_1 x} + (e^{-\mu_1} - 1 + \omega_p)x \quad [\mu_1, \omega_p \ge 0, 0 \le x \le 1] \quad (21)$$

Similar to (9), the μ_1 parameter controls the concavity and $f_1(1) = \omega_p$. Fig. 5 illustrates the characteristics of (9), (20), and (21).

The solid curves are the forms of (20) with varied concavities. These forms are useful in modeling services that have high customer satisfaction even when utility is low, and the effect of desensitization is not significant when utility is high. The dot slash curve is the perception function of (9) with maximum concavity.

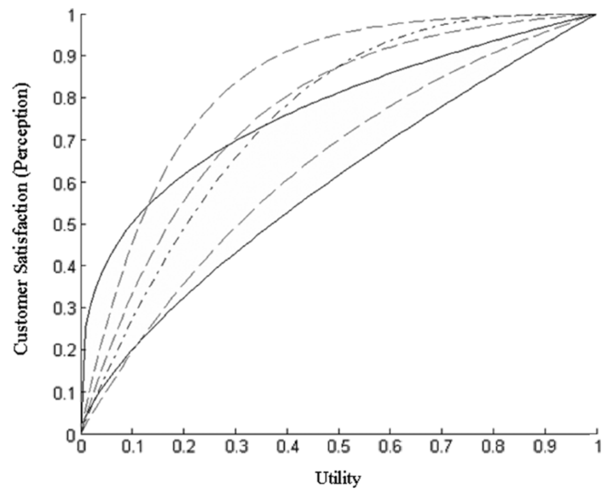


Fig. 5. Forms of perception functions.

The dash curves are the forms of (21) with varied concavities. Unlike (9), the forms of (21) do not place constraint on μ_1. However, for curves with similar concavity, the forms generated by (9) has near linear customer satisfaction growth when utility is low ($U < 0.7$) and the effect of desensitization becomes significant thereafter. These characteristics seem to fit the network services best. In addition, the concavity factor of (9) is more meaningful to analysis (i.e., f'' is in linear form). Higher orders of polynomials are also considered, but they do not add significant control to concavity. In practice, the choice of a best form should be network service specific and be determined based on empirical data gathered for the analyzed service.

B. Impact of the Perception and Disconfirmation Parameters

The parameters ω_p, ω_{dp} and ω_{dn} define the range of customer satisfaction values. The maximum disconfirmation parameter ω_{dp} should be much smaller than the maximum perception parameter ω_p as utility above expectation does not induce significant satisfaction improvement from customers. The combination of ω_p and ω_{dp} gives the maximum ceiling value of customer satisfaction Γ. A value above 1 is not meaningful as Γ is bounded between 0 and 1. However, a value of below 1 is quite feasible, as Γ may be influenced by non-service related factors (e.g., a chronical complainer is unlikely to be fully satisfiable regardless of delivered service utility). The ω_{dn} controls the maximum impact a negative disconfirmation has on perception. When ω_{dn} is large, the degree of negative disconfirmation is also large. As Γ is non-negative, ω_{dn} should be at most as large as ω_p.

Customer satisfaction is particularly sensitive to the choice of μ_1 when the utility value is moderate (i.e., 0.4 $\sim$ 0.8). When μ_1 is high (i.e., $\mu_1 \ge 3$), it is more beneficial for the network service provider to keep service utility at a moderate range (i.e., $U \approx 0.8$). However, this observation holds only if the effect of disconfirmation is low (i.e., the customer expectation is met or the customer has high tolerance to negative disconfirmation). The impact of negative disconfirmation on customer satisfaction is controlled by Parameter μ_3. When utility is fixed, an increase in μ_3 exponentially decreases customer satisfaction. However when negative disconfirmation is very low (i.e., below 0.1), increases in μ_3 approximately result in a linear reduction of customer satisfaction.

For network services, where both μ_1 and the customer expectation are high (i.e., $\mu_1 \geq 3$ and $U_{ei} \geq 0.7$), customer satisfaction does not differ significantly when perceived utility exceeds expectation. However, when expectation is not met, negative disconfirmation will have a significant impact on customer satisfaction. Hence, to retain customers, it is sufficient for a network service provider to deliver service at a quality level matching the expectations of the customers, without maximizing their perceived utilities. However, low service quality delivery (even if it meets customer expectation) adversely influence the customer's future expectations and when the overall customer satisfaction is below the threshold Γ_i^D, the customer will seek other providers. We suggest the parameters of Γ to be acquired through data fitting techniques. The network performance of customers could be obtained in conjunction with customer satisfaction surveys overtime. The computed service utility should be plotted against customer satisfaction and then use best-fit techniques to determine the most appropriate function parameters.

C. Impact of Other Model Parameters

Service utility parameters for QoS-insensitive services influence a customer's preference on the proportioning of upstream throughput performance γ_1 and downstream throughput performance γ_2. This is very application and user dependent. For P2P applications, we typically expect γ_2 to be much higher than γ_1, whereas for FTP-based applications, the proportioning depends on customer's access behavior. For customer preference parameters α_1, α_2 and α_3, the proportioning fundamentally influences the degree of impact each aspect of service utility has on the overall customer satisfaction, and therefore service profitability. For example, when customer care preference α_3 is high, enhancing network infrastructure does not constitute good investment strategy. All of the service utility parameters should be acquired via structured SERVQUAL studies. Methods used in Telecom customer satisfaction studies [3], [4] could serve as guidelines.

In the expectation update process, the assimilation factor κ_a controls the likelihood of the assimilation effect. The perception of difference does not differ significantly between humans and should be small [24] (e.g., $\kappa_a = 0.01$). The parameters β_G and β_L are modifiers of disconfirmation. If we consider $U_{ei} = 0.5$, the impact of disconfirmation on expectation is differentiated by β_G and β_L. The value of β_G and β_L should be equal to or greater than 1, with β_L larger than β_G. The parameter β_M considers the effect of past experience. A larger β_M causes current disconfirmation to be evaluated more significantly on expectation with regard to experience, and the dissipation of this impact is slower as experience accumulates. More specifically, a customer without prior experience is not influenced by β_M (i.e., β_M^0). The cumulation of experience (i.e., $m \geq 1$) rapidly lessens the impact of current disconfirmation on expectation, represented by

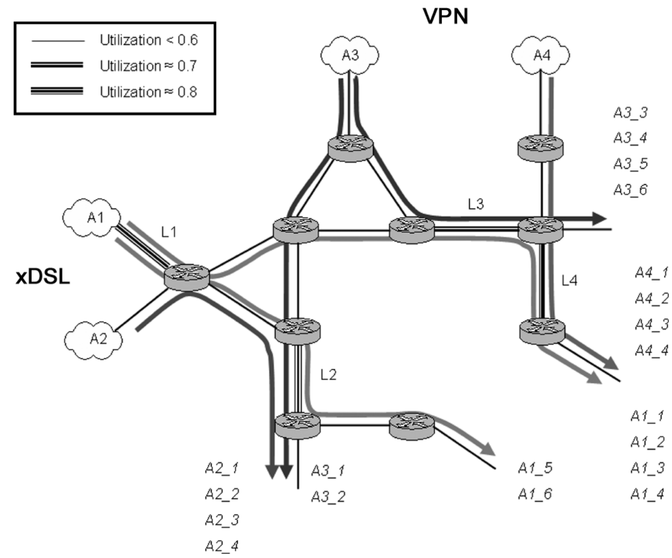


Fig. 6. Simulation topology.

β_M^m. We further notice that experience is accumulated with consistent performance, whether good or bad, and can be destroyed by inconsistency.

The Bass model parameters p, q and S govern the general market growth pattern in a service market. The parameter values are determined via techno-economic studies (e.g., [32]) or growth analysis of similar service markets in the past.

VI. Case Studies and Simulation

In this section, we demonstrate the application and effectiveness of our approach through two sets of case studies and simulations. First, we show how our models can help in a network upgrade decision process and illustrate how key economic, customer and market factors that influence network service planning are captured in our models. Then we analyze the performance of a typical regional ISP network through simulation and show that by representative flow tracking, our model can be applied to WANs. A comparative analysis of the network infrastructure is conducted from three different perspectives: network utilization, customer traffic flow, and customer satisfaction.

As the basis of our first discussion, we simulate a network infrastructure and customer population that is representative of a real world network planning scenario, onto which we offer three equally promising upgrade strategies. The network infrastructure as depicted in Fig. 6 is realized in ns2. The link capacity is set to 24 Mbps. Two regions of xDSL customers ($A1$ and $A2$) and two regions of VPN customers ($A3$ and $A4$) generate a total of 20 service paths (e.g., $A1\_1$, $A2\_1$, etc.). Each service path is used by 5 customers with similar application characteristics, QoS requirements, and access time (Table I). However, each customer varies in performance preferences and service expectations. The routing in this setup is static, such that all customer flows will follow pre-established paths. As the customers in regions $A1$ and $A2$ represent xDSL customers, we have chosen the traffic mix to be primarily a mix of web based light traffic (web browsing, email, etc.) and regular data

TABLE I
CUSTOMER SERVICE PATHS

Path ID	Traffic Type	Requirements / Opt. BW	Start Time	End Time
A1_1	FTP	800 Kb/s Max.	18:00	23:00
A1_2	FTP	800 Kb/s Max.	16:00	23:00
A1_3	FTP	800 Kb/s Max.	6:00	24:00
A1_4	FTP	800 Kb/s Max.	0:00	24:00
A1_5	HTTP	120Kb/s, 100ms	18:00	23:00
A1_6	HTTP	120Kb/s, 100ms	16:00	23:00
A2_1	HTTP	120Kb/s, 100ms	16:00	23:00
A2_2	HTTP	120Kb/s, 100ms	18:00	23:00
A2_3	FTP	800 Kb/s Max.	6:00	24:00
A2_4	FTP	800 Kb/s Max.	0:00	24:00
A3_1	CBR	450Kb/s, 100ms	6:00	24:00
A3_2	CBR	450Kb/s, 100ms	0:00	24:00
A3_3	CBR	450Kb/s, 100ms	9:00	17:00
A3_4	CBR	450Kb/s, 100ms	11:00	19:00
A3_5	CBR	450Kb/s, 100ms	6:00	24:00
A3_6	CBR	450Kb/s, 100ms	0:00	24:00
A4_1	CBR	450Kb/s, 100ms	9:00	17:00
A4_2	CBR	450Kb/s, 100ms	11:00	19:00
A4_3	CBR	450Kb/s, 100ms	6:00	24:00
A4_4	CBR	450Kb/s, 100ms	0:00	24:00

transfer occurring in prime time after work. We also disperse sources of semi-permanent and permanent FTP traffic to reflect the P2P and other constant data transfer traffic often present in the ISP networks. For the VPN customers, CBR traffics are used to represent their constant resource demand. For realism, the simulated network infrastructure and customer behaviors are designed to facilitate traffic intermixing among customers of different service classes and introduce a varied mix of customer access behaviors. Additional background aggregate traffics modeled as Pareto flows establish the desired average link utilization (Fig. 6), and a daily wave pattern with high traffic from 12:00 to 22:00 (peak at 20:00), modeled after a real regional ISP aggregate load trace. Details on the setup of aggregate traffics are presented in the latter simulation case. In our experiments, the 24 hour daily cycle is mapped to 120 min. simulation time so that each simulation run is within reasonable time bound. We consider the base case and three upgrade options: no upgrade (base case), upgrade links L1 and L2 to 48 Mbps (option 1), upgrade links L2 and L3 to 48 Mbps (option 2), and upgrade links L2 and L4 to 48 Mbps (option 3). All three upgrade options have the same cost. Option 1 is an aggressive upgrade strategy aimed at pleasing xDSL customers, while option 2 and 3 are more balanced strategies. Each of the upgrade options is simulated multiple times in ns2 through a 24 hour day (i.e., 120 min. simulation time) and the QoS measurements (i.e., delay and throughput) are taken at 1 min. intervals. The average measurements of each interval across runs are used to compute network performance for the customers.

When taking into account the customer access behaviors (i.e., customers' service path and service time), we can compute the perceived utility of each upgrade option. For simplicity, we assume the customers only care about service quality (i.e., $\alpha_1 = 1$, α_2, $\alpha_3 = 0$). This then implies that the perceived utility is analogous to the ratio of average throughput to optimum bandwidth for FTP traffic, and to the percentage of non-defective service time for the other traffic. The solid lines in Fig. 7 present the perceived utility of the customers under each upgrade option. In each region, the customers are ordered on the graphs by their path IDs. In the base case of no upgrades, we notice that xDSL customers accessing path $A1\_3$ and $A1\_4$ have significantly better perceived utilities compared to other xDSL customers because they do not access link L1 and are not "prime time" traffic. Option 1 significantly increases the service performance of the xDSL customers at the expense of some moderate performance drop from the VPN customers, which seems quite acceptable. The perceived utility of some customers from region $A1$ remains low due to L3 link load. Option 2 and 3 are roughly equivalent as they improve the service performance of region $A1\_3$ and $A1\_4$ respectively. Although informative from a network performance point of view, this analysis provides few insight into which upgrade option is in fact the most beneficial, as all of the options improve performance for some customers.

We now carry the computation of these options through the rest of our model. For this scenario, let the expectation of the customers be normal distributions with means of 0.65, 0.7, 0.8 and 0.9 respectively for $A1$, $A2$, $A3$ and $A4$, and a standard deviation of 0.05. This distribution is used to reflect the diversity in customer expectations and the relative differentiation between the xDSL customers and the VPN customers. Fig. 7 illustrates the perceived utilities, expectations, and customer satisfactions of the customers under each upgrade option. The choice of customer satisfaction parameters are taken so that the customer satisfaction function exhibits its general form as observed in empirical studies. In the base case, the xDSL customers have significant negative disconfirmations and consequently low satisfactions. Option 1 significantly improves the satisfactions of xDSL customers, although some of the customers from region $A1$ are still dissatisfied due to L3 link load. Interestingly, the additional influx of $A1$ traffics on L3 and L4 reduced the perceived utilities of VPN customers, just enough to make their perceived utilities to fall below expectations. Therefore, our computations of Γ indicate option 1 may not be a good upgrade option. Regarding option 2 and option 3, our model suggests that negative disconfirmations are eliminated in region $A3$ and $A4$ respectively with each option. When factoring in the Γ^D level, we further observe that option 2 seems to generate more dissatisfied customers than option 3 due to the high expectation rating in region $A4$. To accentuate our case, we consider a saturated xDSL and VPN market where our network service provider does not offer the best service. This market condition could be represented in our model with zero market growth and a 0.0 market turnover rate (from other competitors). According to our market segmentation model, this implies that more customers will leave the provider in option 2 compared with option 3. By determining the number of customers remaining with the service after each upgrade option, we can compute the retention rate and project the future profitability from our model as presented in Fig. 8. With the cost of upgrades being equal, the outcome of our an-

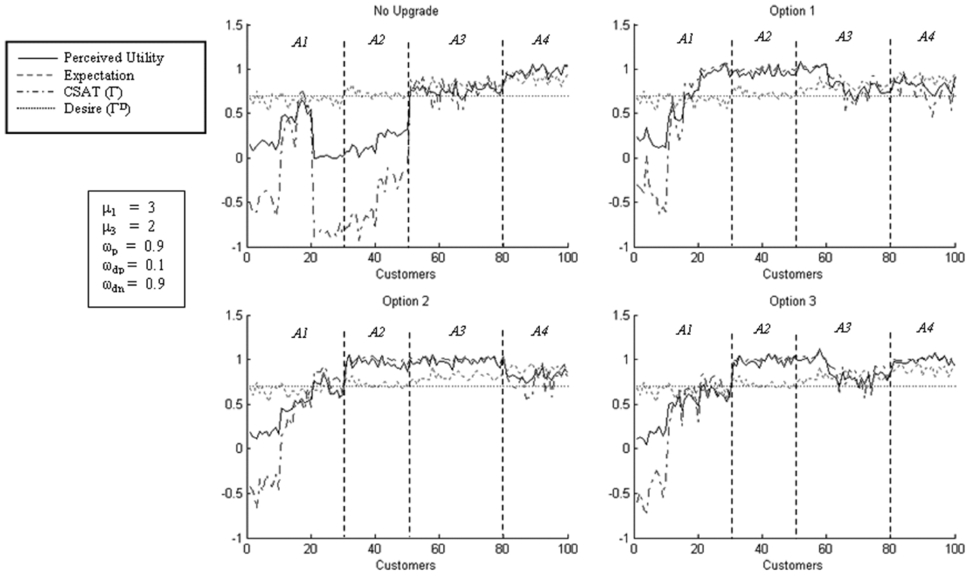


Fig. 7. Customer satisfaction of network upgrades.

alytical model indicates option 3 is the more profitable option as long as the service charge for VPN is higher than the service charge for xDSL. By following our model and factor in the customer's expectations, satisfaction conditions and market dynamics, we are able to arrive at a much more informed upgrade decision using service profitability as an indicator.

A commonly observed principle in market science states that service profitability is maximized with respect to customer's satisfiability. Rather than satisfy each customer, a business should strive to satisfy each satisfiable customer, and only if it is profitable to do so [33]. When we conduct the computation of Γ for upgrade option 2 with optimistic VPN customers ($\omega_p = 0.9$) and pessimistic VPN customers ($\omega_p = 0.75$), the results confirm with this principle: the upgrade raises the customer satisfaction of the optimistic customers and fails to impress the pessimistic customers.

The long-term interactions among service performance, customer experience and future expectation is a well studied topic in market science and economic psychology. Our model effectively captures many of their key observations. Consider a customer who has stayed with the service provider for 20 evaluation periods and has experienced consistent service performance, we subject the customer to low and inconsistent service performance for the next 50 evaluation periods and trace his/her service expectations (Fig. 9) obtained from our expectation update process. The parameters β_L and β_G are general values taken based on our discussion in Section V-C, and κ_a is set to 0 (i.e., no assimilation) for simplicity. As shown in Fig. 9(a), in the short term (first 7 iterations), the customer's future expectations are not significantly influenced by perceived utilities as the customer has experience with consistent service delivery in the past. However, the customer gradually loses confidence with the service (iterations 8 to 17) and expectations become heavily dependent on short-term perceived utilities. This trend confirms with the observations on expectation and customer experience [34], [35]. When a customer is dissatisfied due to poor service performance in the short term, an experienced customer (whose future expectation is not significantly reduced) is more likely to be loyal than an inexperienced customer. The works on expectation further suggest that when customer perceives disconfirmation, the degree of adjustment to expectation is determined by the uniqueness of the event and the strength of previous expectation. In the first few iterations of our illustrated case Fig. 9(a), the impact of disconfirmation on expectation adjustment is low. As the occurrence of disconfirmation increases, its impact is significantly more severe. The parameter β_M controls the weight of current disconfirmation on expectation. A higher β_M indicates a lower strength of the past expectation. In Fig. 9(b) where β_M is higher, the impact of disconfirmation on expectation is significantly more severe even in the presence of long past experience. It is apparent that the interaction among expectation, performance, and customer satisfaction is a significant factor influencing the service profitability of ISP operations and should be considered in the

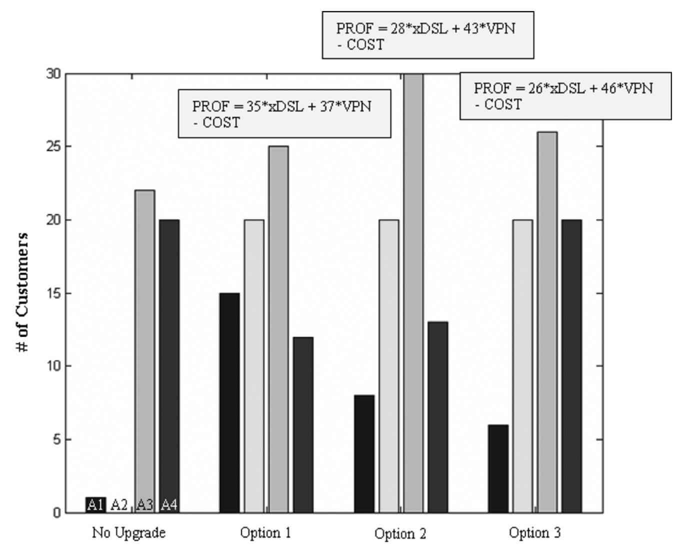


Fig. 8. Customer retention of different upgrade options.

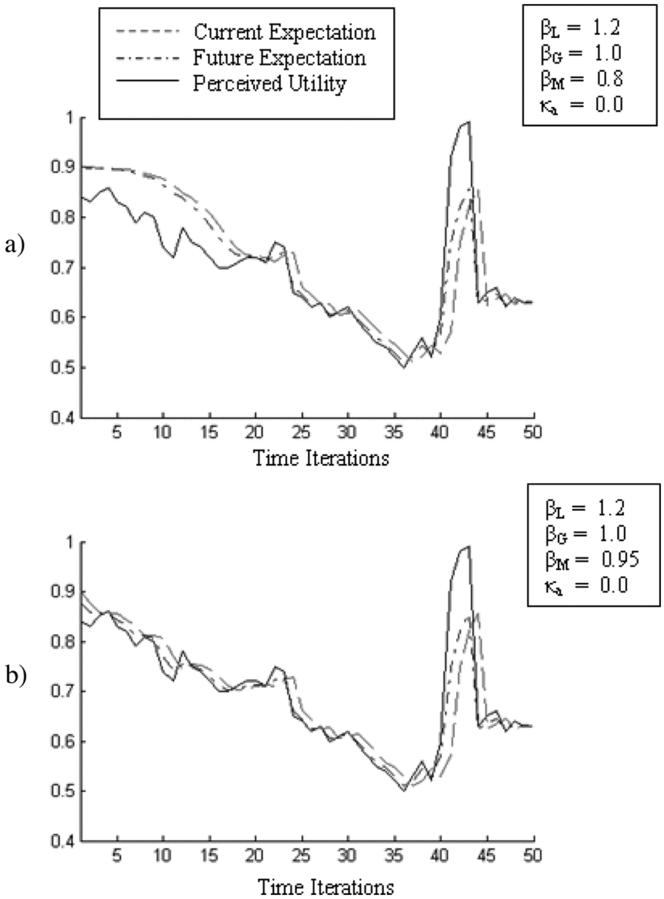


Fig. 9. Experience and long-term service expectation.

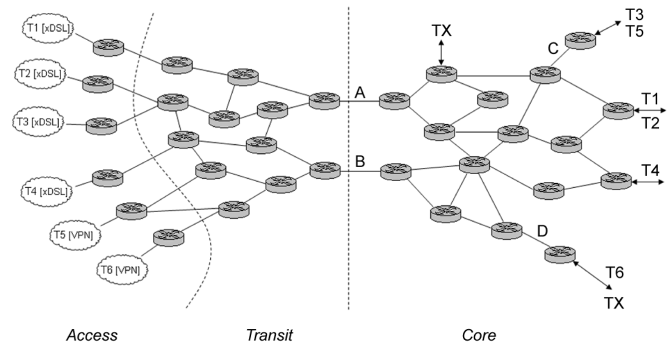


Fig. 10. Simulated topology of a regional service provider network.

network upgrade decision process. The trends captured by our model integrate such factors in the network upgrade decision process.

Finally, we examine how varying market conditions can affect service profitability as presented in our model. Regarding the aforementioned three upgrade options, suppose the current consumer market size is 100 for the xDSL and 100 for the VPN service. Furthermore, suppose the VPN market is fully saturated while the xDSL market is estimated to grow by 80 customers (in practice, this value is projected by the Bass model). When conducting the modeling analysis on the same flow data, we find that the aggressive xDSL strategy (option 1) is able to attract more xDSL customers by pushing for better service performance, and hence better future expectation. In contrast to our previous analysis based on saturated markets, upgrade option 1 now becomes a more profitable option.

In the second simulation case, we examine how network performance influences customer satisfaction and show an example of how our model could be used in practice. The simulation setup depicts a typical regional service provider network. Three comparative performance analysis of the network infrastructure are presented, each from a different view: link utilization, QoS performance of customer flows, and customer satisfaction. Fig. 10 shows the regional service provider network, as simulated in ns2. The typical access, transit and core network topology is recreated. The links in this simulated network are identical in their characteristics to the first simulation study. Six customer groups and one transit traffic from a peer provider are studied. Each customer group has a mixture of service types and customer access times with daily traffic shapes also similar to the first simulation study. We track a representative customer flow from each population assuming they are a subscriber of either the xDSL or VPN service (noted between brackets in Fig. 10) and analyze their behaviors under various conditions. The customer flows from each service population are modeled as an aggregate Pareto flow from customer access to their respective traffic exchange point at the edge of the core network. The flows are routed through least joint paths. When solely considering the link utilization, four potential bottleneck links: A, B, C and D are identified. However, this data alone do not indicate the impact of these congestive links on the performance of customer flows.

To analyze the impact of the network utilization on customer flows, we trace a representative customer flow from each customer population. A representative xDSL customer is traced in each of the populations T1 to T4 and a representative VPN customer is traced in each of the populations T5 and T6. We consider in this case study that the xDSL customers are offered 330 kb/s (maximum throughput) service while the VPN customers are offered 680 kb/s service. Each traced flow is modeled as FTP in the simulation and the dotted lines in Fig. 11 shows the application level throughput measured over a 24 hour period at five minutes sampling intervals. Delay is not monitored in this case because round trip delays within regional network seldom exceed application requirements. From the throughput trace, it is apparent that the congestions at link A and B during prime time of the day have significant impact on customer flows. Comparing the throughput of T1 to T4, T4 seems the least impacted because link A is the only bottleneck link along the flow and the trace from T4 has the least shared path with other flows. Comparing T5 and T6, T6 fares significantly worse since in addition to the bottleneck at link B, its traffic also shares link D with transit traffic TX from a peer provider. Upgrade link A and link B appears to be sensible from this analysis. The dashed lines in Fig. 11 present the throughput traces after link A and B are upgraded.

Aside from conducting link upgrades which are cost prohibitive, resource provisioning mechanisms such as service differentiation and network dimensioning could be equally beneficial. To illustrate this, we create service differentiation across link C and D into premium and standard classes (60% and 40%

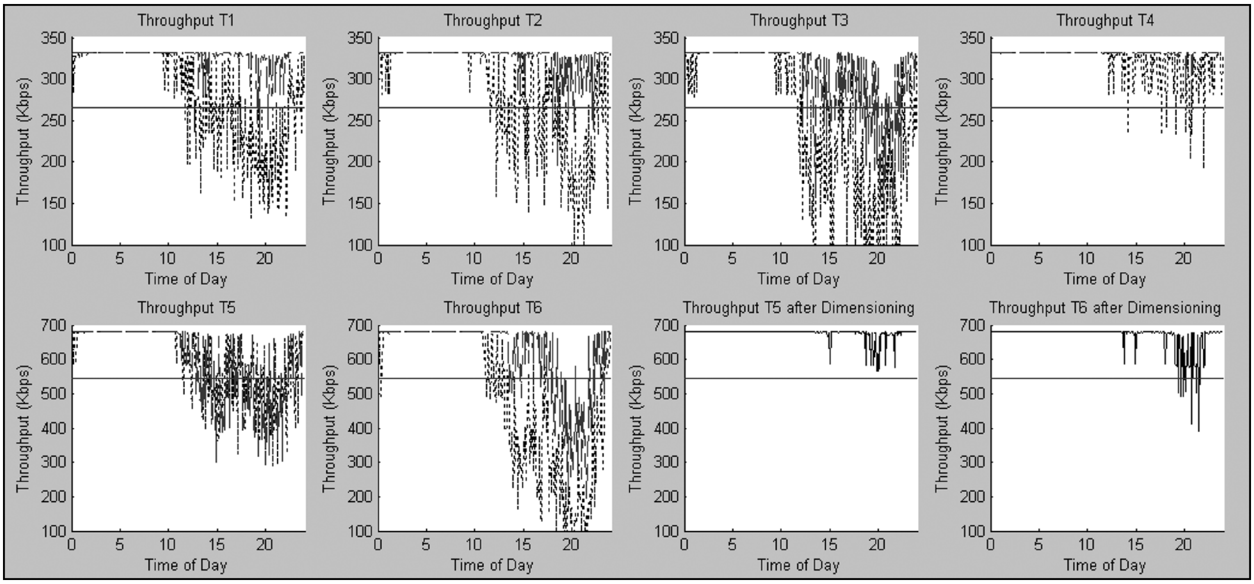


Fig. 11. Throughput performance of representative customer flows before and after upgrades.

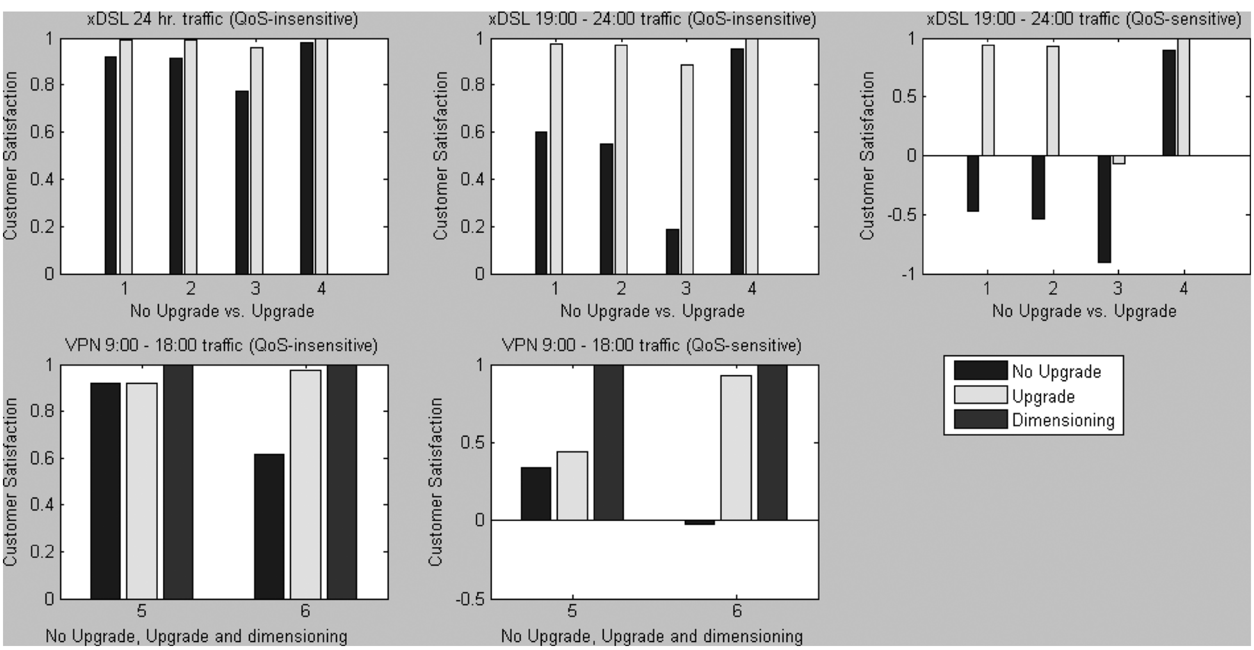


Fig. 12. Effect of upgrades and dimensioning on customer satisfaction.

of the link capacity is dimensioned for each class respectively). Traffic from T5 and T6 is thus given precedence over traffic from T3 and TX. Fig. 11 shows the result of such dimensioning.

Improvement in network performance does not produce proportional improvement in customer satisfaction. In analyzing the customer satisfaction of the above trace traffics under varied customer conditions, we demonstrate this property. Fig. 12 presents the customer satisfaction under different customer access patterns and QoS sensitivity. For QoS-insensitive traffic (e.g., FTP and P2P) the perceived utility is computed as the ratio between obtained throughput over maximum throughput. For QoS-sensitive traffic (e.g., multimedia traffic), throughput of 266 kpbs (for xDSL) and 544 kpbs (for VPN) are used as the defective thresholds, corresponding to roughly 80% of the maximum throughput (depicted in Fig. 11 as solid horizontal lines). We compute customer satisfaction with the same modeling parameters as used in the first simulation study and the customer expectation is set to 0.8. As illustrated in Fig. 12, the raw computation of customer satisfaction could yield negative values. In practice, these negative values should be set to 0 to obtain the normalized value of Γ, nevertheless they are left here for comparison. The first three graphs consider xDSL customers from population T1 to T4. We see that for customers that are QoS-insensitive and access the network 24 hours a day (representation of the permanent P2P population often prevalent in xDSL service), performing link upgrade is of little consequence. This category of customers is satisfied as long as their achievable daily average throughput remains reasonable.

However, for the other xDSL users that are the bulk of "prime time" traffic, their satisfaction is severely impacted by link congestion and hence they benefit the most from link upgrade. We note that because T4 does not access the same transit-core link as T1, T2 and T3, it was not as severely influenced by prime time traffic as the others. Hence our analysis of the xDSL customer satisfaction suggests that upgrade link A is quite effective given a large mix of prime time xDSL users in population T1 to T3 (which should be the case in practice). For the VPN customers, their access times are generally during business hours. For QoS-insensitive customers, upgrade link B only improves the performance of T6 somewhat, while service differentiation does not yield any visible result. For QoS-sensitive customers, The link upgrade and service differentiation strategies creates very different customer responses. It illuminates a prevailing theory in our model: customer satisfaction is a subjective, comparative evaluation between perception and expectation. In the case of T5, performing network upgrade alone does not raise the customer's perceived utility to a level that meets the customer's expectation and hence there is very little improvement in customer satisfaction. In the case of T6, the performance improvement due to link upgrade already meets the customer's expectation, conducing service differentiation does not further influence the customer's opinion of the service. In general, network planning and upgrade strategies should be made with respect to the particularities and expectations of the customers.

VII. Conclusion

In this paper, we have presented a market science approach to assess service profitability of network upgrade and planning decisions. Our approach captures the intricacies among network performance, customer behavior, and market dynamics and is founded on many theoretical and empirical studies from market science, economics and psychology. The result is a series of mathematical processes that are concrete and well-behaved. Following this approach, we created a generalized analytical model for forecasting network upgrade and planning decisions, providing a set of meaningful parameters to model wide varieties of network service characteristics, customer attributes, and market conditions. Through analysis and simulations, we have shown that the best network upgrade option cannot be determined solely based on performance improvements, but is also service, customer, and market dependent.

In the simulation studies, we applied our model in regional networks and used representative flow tracing to conduct customer satisfaction analysis. Our prior work [9] suggests that through aggregation and pruning, customer performance at the network level is obtainable without per customer tracking or full scale network simulation. Given the intricacies among the modeling parameters, it is imperative to conduct validation and tuning over time in real world ISP operations. In market science, when faced with complex models and hypotheses, much of the validation work is carried out over large data sets across long periods of time, where statistical analysis is often helpful in deducing trends and linkages among metrics. We think a similar approach should be taken here. Whereas simulation studies and numerical analysis could shed some light on the sensitivity of the modeling parameters, much of the work relies on market data such as customer satisfaction and service turnover rate, information trackable by ISPs in business practices. Through an iterative validation and parameter tuning process, the model and its parameters could be evolved and refined over time to suite the particularities of the service provider and the market. The mathematical forms that we have introduced in this paper are kept general for sake of analysis, and offer flexible parameters to facilitate operational tuning.

A number of future works extend from this paper, such as the incorporation of performance related service charging and penalty functions in the profitability computation, the effect of provider competitions, etc. We believe our approach brings a unique perspective to the network upgrade and planning research that is much needed and the resulting models are general enough to benefit many network service related analysis processes.

References

[1] M. Rose and K. McCloghrie, Concise MIB definitions RFC 1212, Mar. 1991.

[2] L. L. Tung, "Service quality and perceived value's impact on satisfaction, intention and usage of short message service (SMS)," *Information Systems Frontiers*, vol. 6, no. 4, pp. 353–368, 2004.

[3] Y. Wang, H. P. Lo, and Y. Yang, "An integrated framework for service quality, customer value, satisfaction: Evidence from China's telecommunication industry," *Information Systems Frontiers*, vol. 6, no. 4, pp. 325–340, 2004.

[4] B. Leisen and C. Vance, "Cross-national assessment of service quality in the telecommunication industry: Evidence from the USA and Germany," *Managing Service Quality*, vol. 11, no. 5, pp. 307–317, 2001.

[5] P. Flynn, K. Curran, and T. Lunney, "A decision support system for telecommunications," *Int. J. Network Management*, vol. 12, no. 2, pp. 69–80, 2002.

[6] S. Khan, K. F. Li, and E. G. Manning, "The utility model for adaptive multimedia systems," in *Proc. Int. Conf. Multimedia Modeling*, Singapore, Nov. 1997, pp. 111–126.

[7] Y. d'Halluin, P. Forsyth, and K. R. Vetzal, "Managing capacity for telecommunications networks under uncertainty," *IEEE/ACM Trans. Networking*, vol. 10, no. 4, Aug. 2002.

[8] S. Jagannathan, J. Altmann, and L. Rhodes, "A revenue-based model for making resource investment decisions in IP networks," in *Proc. IEEE/IFIP Symp. Integrated Network and System Management (IM)*, Colorado Springs, CO, Mar. 2003, pp. 185–197.

[9] J. Xiao and R. Boutaba, "Customer-centric network upgrade strategy: Maximizing investment benefits for enhanced service quality," in *Proc. IEEE/IFIP Network Operations and Management Symp. (NOMS)*, Seoul, Korea, Apr. 2004, vol. 1, pp. 759–772.

[10] D. Mitra, K. G. Ramakrishnan, and Q. Wang, "Combined economic modeling and traffic engineering: Joint optimization of pricing and routing in multi-service networks," in *Proc. 17th Int. Teletraffic Congress (ITC-17)*, Amsterdam, The Netherlands, 2001, pp. 73–85.

[11] S. Shakkottai and R. Srikant, "Economics of network pricing with multiple ISPs," in *Proc. INFOCOM*, Miami, FL, Mar. 2005, vol. 1, pp. 184–194.

[12] V. A. Siris, D. J. Songhurst, G. D. Stamoulis, and M. Stoer, "Usage-based charging using effective bandwidths: Studies and reality," in *Proc. 16th Int. Teletraffic Congress (ITC-16)*, Edinburgh, U.K., Jun. 1999.

[13] C. Courcoubetis, F. P. Kelly, V. A. Siris, and R. Weber, "A study of simple usage-based charging schemes for broadband networks," *Telecommun. Syst.*, vol. 15, no. 3–4, pp. 323–343, 2000.

[14] Z. Duan, Z. Zhang, and Y. Hou, "Service overlay networks: SLAs, QoS, and bandwidth provisioning," *IEEE/ACM Trans. Networking*, vol. 11, no. 6, Dec. 2003.

[15] R. L. Oliver, "Effects of expectations and disconfirmation on postexposure product evaluations," *J. Appl. Psychol.*, vol. 62, Apr. 1977.

[16] R. L. Oliver, "A cognitive model of the antecedents and consequences of satisfaction decisions," *J. Marketing Research*, vol. 17, Nov. 1980.

[17] R. L. Oliver and W. DeSarbo, "Response determinants in satisfaction judgements," *J. Consumer Research*, vol. 14, Mar. 1988.

[18] E. W. Anderson and M. W. Sullivan, "The antecedents and consequences of customer satisfaction of firms," *Marketing Science*, vol. 12, no. 2, 1993.

[19] R. N. Bolton, "A dynamic model for the duration of the customer's relationship with a continuous service provider: The role of satisfaction," *Marketing Science*, vol. 17, no. 1, 1998.

[20] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "A conceptual model of service quality and its implications for future research," *J. Marketing*, vol. 49, pp. 41–50, 1985.

[21] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality," *J. Retailing*, vol. 64, no. 1, pp. 12–40, 1988.

[22] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "Refinement and reassessment of the SERVQUAL scale," *J. Retailing*, vol. 67, no. 4, pp. 420–450, 1991.

[23] A. Deaton and J. Muellbauer, *Economics and Consumer Behavior*. Cambridge, U.K.: Cambridge Univ. Press, 1980.

[24] P. Slovic and S. Lichenstein, "Comparison of Bayesian and regression approaches to the study of information processing in judgement," *Organizational Behavior and Human Performance*, vol. 6, pp. 649–744, 1971.

[25] W. Boulding, R. Staelin, A. Kalra, and V. Ziethaml, "A dynamic process model of service quality: From expectations to behavioral intentions," *J. Marketing Research*, pp. 7–27, Feb. 1993.

[26] F. M. Bass, "New product growth model for consumer durables," *Management Science*, vol. 15, p. 215, 1969.

[27] P. A. Dabholkar, "Customer satisfaction and service quality: Two constructs or one?," in *Enhancing Knowledge Development in Marketing*, D. W. Cravens and P. R. Dickson, Eds. Chicago, IL: American Marketing Assoc., 1993.

[28] T. Forum, SLA Handbook Solution Suite v2.0, Volume 2: Concepts and Principles TMF Report, 2005.

[29] A. Tversky and D. Kahneman, "Availability: A heuristic for judging frequency and probability," *Cognit. Psychol.*, vol. 4, pp. 207–232, 1973.

[30] R. T. Rust, J. Inman, and J. Jia, "Customer expectation updating: a dynamic model, theoretical implications and empirical evidence," Vanderbilt Univ., Nashville, TN, Working Paper, 1997.

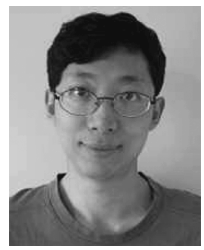
[31] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. New York: Springer-Verlag, 1985.

[32] Y. R. e. a. D. Lecrosnier and P. Grosso, "Implementation strategies for advanced access networks," UROSCOM Project P614 Main Report, 1998.

[33] R. Hallowell, "The relationships of customer satisfaction, customer loyalty, and profitability: An empirical study," *Int. J. Service Industry Management*, vol. 7, no. 4, pp. 27–42, 1996.

[34] R. B. Woodruff, E. R. Cadotte, and R. L. Jenkins, "Modeling consumer satisfaction processes using experience-based norms," *J. Marketing Research*, vol. 20, no. 3, 1983.

[35] R. L. Oliver and R. S. Winer, "A framework for the formation and structure of consumer expectations: Review and propositions," *J. Economic Psychology*, vol. 8, no. 4, 1987.



Jin Xiao received the B.Sc. degree (first class honors) in computer science from the University of Calgary, Canada, in 2001. He is currently a Ph.D. candidate at the School of Computer Science at the University of Waterloo, Canada.

He conducts research in the areas of network and distributed systems management, economic modeling, network service quality assurance, and autonomous management system design.

Raouf Boutaba received the MSc. and Ph.D. degrees in computer science from the University Pierre and Marie Curie, Paris, France, in 1990 and 1994, respectively.

He is currently an Associate Professor of computer science at the University of Waterloo, Ontario, Canada. His research interests include resource and service management in wired and wireless networks.

Dr. Boutaba is the founder and Editor-in-Chief of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT and is on the editorial boards of several journals. He is currently a distinguished lecturer of the IEEE Communications Society, and the chairman of the IEEE Technical Committee on Information Infrastructure and the IFIP Working Group 6.6.