Computers and Mathematics with Applications 62 (2011) 419–439

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Analysis of bandwidth allocation on end-to-end QoS networks under budget control

Chia-Hung Wang, Hsing Paul Luh *

Department of Mathematical Sciences, National Chengchi University, No. 64, Sec. 2, ZhiNan Rd., Wen-Shan District, Taipei 11605, Taiwan

ARTICLE INFO

Article history:
Received 3 January 2011
Received in revised form 2 May 2011
Accepted 11 May 2011

Keywords:
Revenue management
Bandwidth allocation
Convexity
Blocking probability
Sensitivity analysis
Internet economics

ABSTRACT

This paper considers the problem of bandwidth allocation on communication networks with multiple classes of traffic, where bandwidth is determined under the budget constraint. Due to the limited budget, there is a risk that the network service providers can not assert a 100% guaranteed availability for the stochastic traffic demand at all times. We derive the blocking probabilities of connections as a function of bandwidth, traffic demand and the available number of virtual paths based on the Erlang loss formula for all service classes. A revenue/profit function is studied through the monotonicity and convexity of the blocking probability and expected path occupancy. We present the optimality conditions and develop a solution algorithm for optimal bandwidth of revenue management schemes. The sensitivity analysis and three economic elasticity notions are also proposed to investigate the marginal revenue for a given traffic class by changing bandwidth, traffic demand and the number of virtual paths, respectively. By analysis of those monotone and convex properties, it significantly facilitates the operational process in the efficient design and provision of a core network under the budget constraint.

1. Introduction

The telecommunication industry is moving toward a converged network because of the rapid growth of Internet traffic, aggressive deployment of broadband fiber optic network, advance of Voice over IP technology, and the global standardization of IP technology. The converged network uses a single global IP-based network carrying all types of network traffics to replace the traditional separated packet switching and circuit switching networks [1]. A core network is a part of the global IP-based network under a single organization or Internet Service Provider (ISP) [1]. Applications expected to produce the bulk of traffic in the future multi-service Internet can be broadly categorized as streaming or elastic according to the nature of the connections they produce [2]. Connections using the core network are typically generated by a very large population of users independently communicating with an equivalently large population of servers and correspondents for a variety of applications.

Every core network has a centralized network manager, known as the Bandwidth Broker (BB), which is aware of the network topology and status, using the underlying routing protocols. First the client (maybe a single user, a corporate network or an aggregated sub-network) negotiates its Quality of Service (QoS) requirement, known as Service Level Agreement (SLA) negotiation, with the BB in its core network. For example, real-time multimedia applications require stringent QoS guarantees, including hard bounds on bandwidth and packet loss probability [3,4]. Then the BB in the source network negotiates resource allocation with the BBs in intermediate and destination networks. The users accessing the network via the ISP are guaranteed an amount of bandwidth under the total available budget for their connections [2].

* Corresponding author.
E-mail address: slu@nccu.edu.tw (H.P. Luh).

0898-1221/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2011.05.024

For a communication network providing performance guarantees, it has to reserve resources and exercise call admission control [4]. Network users are mainly interested in getting good quality connections whenever they place requests. It is the network providers' mission to have an virtual path with suitable bandwidth. Clearly, it is too costly for the network providers to assert a 100% guaranteed availability for all connections under the budget constraint at any time. This is also not necessary since demand for connections or bandwidth capacity varies over time. Traffic flow fluctuates with time, and connections do not last forever but occur at random times and vanish in the network once the corresponding digital document has been transferred completely. This results in a random, dynamic set of active connections. Moreover, the bandwidth assigned to each connection would determine how long that connection will stay active and thus impact the evolution of the set of active connections. The network provider chooses an optimal sharing scheme for the different users under the total budget to fulfill connection requirements. In addition, the risk (probability) of rejecting connection requests due to lack of resources is supposedly kept below a negotiated level.

We deal with the problem of dimensioning bandwidth on core networks to serve all traffic demands under the budget constraint. In this paper, we aim to analyze the relationship between blocking probability, bandwidth, traffic demand and the available number of virtual paths on communication networks with service from ISPs, where requests for connections represent customers arriving at the system. As soon as requests are accepted by the system, the service begins. The installed bandwidth allocation is used to maintain a guaranteed connection availability where the blocking probability is kept below certain negotiated level. Our intent is to analyze the sensitivity of the blocking probability by system parameters, including directional monotone and convex properties. Then we study how these changes affect the long-run average revenue and find structural results. The relationship among the revenue function, the blocking probability and allocated bandwidth are investigated under the budget constraint.

Bonald et al. [5] provided a queueing analysis of three usual bandwidth allocations, namely max–min fairness, proportional fairness and balanced fairness in a communication network. Nain [6] provided a solution to the classical Erlang blocking model on circuit-switched network, where the author obtained monotone and concave properties for loss probabilities, throughput and channel occupancy in terms of traffic intensity. Antunes et al. [7] provided an analysis of loss networks with different classes of requests which move according to some routing policies. Cho et al. [2] investigated the optimal partitioning of the end-to-end QoS budget to quantify the advantage of having a non-uniform allocation of the budget over the links in a path. Jin and Jordan [3] studied the sensitivity of resource allocation and the resulting QoS to resource prices in a reservation-based QoS architecture. Güven et al. [8] formulated an optimization problem of load balancing the traffic, where multiple paths are provided between a source and a destination using application-layer overlay. Faragó [9] gave an estimated blocking probability and link utilization for general multi-rate, heterogeneous traffic, where the individual bandwidth demands may aggregate in complex ways. Maglaras and Zeevi [10] studied the equivalent behavior of communication systems in a single-class Markovian model under revenue and social optimization objectives. Bruni et al. [11] designed a connection admission control procedure for resource management in a telecommunication network. However, it was mentioned that the relationship between performance, demand and capacity has not been well investigated [12].

We assume the user population is infinite and connections join the core network according to a Poisson process [4,13–15], etc. The sojourn time during which they occupy virtual paths has an arbitrary distribution. The blocking probability is determined while allocating resources under the budget constraint. The blocking of connections occurs due to the failure of meeting the demand of each traffic class for virtual paths, where the traffic demand is considered and written as a function of the product of the occurrence rate of connections and the average connection volume. The explicit expressions of the blocking probability and the expected path occupancy are to be derived in terms of model parameters.

The originality of our work lies in derivation of bandwidth elasticity and demand elasticity of blocking in economic models by analyzing properties of the blocking probability with respect to allocated bandwidth, traffic demand and the number of virtual paths. The main contribution of the present paper is to prove the relationship between the blocking probability and allocated bandwidth under the budget constraint. Monotone and convex properties of the blocking probability are shown in both theoretical construction and numerical examples.

The remainder of the paper is organized as follows. Assumptions and problem definitions are presented in Section 2. Section 3 introduces two revenue management schemes for allocating bandwidth under the budget constraint. In Section 4, we prove monotone and convex properties of the blocking probability and expected path occupancy of connections. Through monotonicity and convexity of the blocking probability, those properties of two revenue/profit functions are studied. The optimality conditions and a solution algorithm for two revenue management schemes are presented in Section 5. Section 6 provides some applications in economic models to illustrate the use of the monotonicity and convexity of the blocking probability. Sensitivity analysis with numerical illustrations are conducted in Section 7. Concluding remarks are drawn in Section 8. We give proofs for each proposition and theorem while providing most of them in the Appendix as the supplementary material for online publication only in order not to interrupt the flow of presentation.

2. Problem definition

Consider a communication (core) network $\mathbb{G}$, the traffic occurs at the source node randomly, and it will be connected through the core network $\mathbb{G}$ if there exist available virtual paths. We presume the existence of a reservation-based QoS architecture using virtual paths for real-time applications and using scheduling policies that are capable of assigning bandwidth to aggregates of flows. Suppose there are m different traffic classes in the network $\mathbb{G}$, and denote $\mathbb{M} = \{1, 2, \ldots, m\}$ as an index set consisting of m traffic classes.

For each class $i \in \mathbb{M}$, there exist K_i available virtual (end-to-end) paths for connections of class i, where K_i is a positive integer. For any incoming connection, say j, of class $i \in \mathbb{M}$, when allowed to enter the core network $\mathbb{G}$, it will be routed through one path $p_{i,j}$ from those K_i virtual paths with allocated bandwidth x_i. In other words, data packets of incoming connection j of class $i \in \mathbb{M}$ can be transmitted along an virtual path $p_{i,j}$. In general, every virtual path of class $i \in \mathbb{M}$ is allocated by the same amount of bandwidth x_i. Every virtual path of class $i \in \mathbb{M}$ has to meet the same minimum bandwidth requirement $b_i^{\min} \geq 0$, namely,

$$x_i \geq b_i^{\min}, \quad \forall i \in \mathbb{M}. \tag{1}$$

Given limited budget B, network managers would like to determine the bandwidth allocation $\mathbf{x} = (x_1, \ldots, x_m)$ under the available number of virtual paths $\vec{K} = (K_1, \ldots, K_m)$. Due to the limited budget B on network planning, there exists the budget constraint

$$\sum_{i \in \mathbb{M}} K_i c_i x_i \leq B, \tag{2}$$

where $c_i > 0$ is the average cost of one unit bandwidth through virtual paths for class $i \in \mathbb{M}$. The goal is to determine the bandwidth allocation under the budget constraint so that the revenue earned by the network access providers is maximized. From the budget constraint (2) and the minimum bandwidth requirement (1), we can determine the possible range of available bandwidth for each class $i \in \mathbb{M}$, e.g.,

$$b_i^{\min} \leq x_i \leq \frac{B - \sum_{j \neq i} K_j c_j b_j^{\min}}{K_i c_i}, \quad \forall i \in \mathbb{M}. \tag{3}$$

The maximum throughput of those virtual paths are limited by either the total budget B or bottleneck links which lie on those virtual paths. The blocking of connections occurs if the dynamic traffic demand exceeds the maximum throughput.

Definition 1. The *maximum throughput* of class $i \in \mathbb{M}$ is defined as the maximum connection volume (in packets) that can be transmitted through those virtual paths in a unit of time [16]. Namely, the maximum throughput Θ_i of K_i virtual paths for class $i \in \mathbb{M}$, is the product

$$\Theta_i \triangleq K_i x_i. \tag{4}$$

Assume that connections of class i occur at the source node in accordance with independent Poisson processes at rate $\lambda_i > 0$ [4,14], but the connection volume in terms of packets to be transmitted has an arbitrary distribution with mean $\sigma_i > 0, i \in \mathbb{M}$. For each class $i \in \mathbb{M}$, suppose the mean sojourn time of connections on virtual paths is $1/\mu_i$, which corresponds to the packet transmission time, and it is equal to average connection volume divided by bandwidth, i.e.,

$$\frac{1}{\mu_i} \triangleq \frac{\sigma_i}{x_i}. \tag{5}$$

Suppose that connections occupy the virtual paths in the order they occur and that sojourn times are identically distributed and mutually independent.

Definition 2. The *traffic intensity* of virtual paths is defined as the fraction of the time in which virtual paths are occupied [6,15,17]. Namely, the traffic intensity of K_i virtual paths, $i \in \mathbb{M}$, is

$$\rho_i \triangleq \frac{\lambda_i}{K_i \mu_i} = \frac{\lambda_i \sigma_i}{K_i x_i}, \tag{6}$$

which is the average occupancy of those K_i virtual paths.

Definition 3. The *traffic demand* is defined as the product of the mean occurrence rate and the average connection volume [12]. Equivalently, the traffic demand y_i for class $i \in \mathbb{M}$, is the product

$$y_i \triangleq \lambda_i \sigma_i. \tag{7}$$

The principal quantity of interest is the blocking probability that all K_i virtual paths of class $i \in \mathbb{M}$ are occupied. In the situation here, the blocking is due to the failure of setting up the available number of virtual paths under limited budget B. For each class i, a connection gets dropped off upon its arrival at the source node when all K_i virtual paths are occupied. Otherwise, it will be routed through an virtual path $p_{i,j}$ with allocated bandwidth x_i. Our objective is to determine blocking probabilities in terms of bandwidth x_i.

The risk of blocking possible connections is analyzed according to an Erlang loss model under assumptions of Poisson arrivals, general sojourn time, preset K_i virtual paths with identical bandwidth x_i, and no waiting space [6,14]. Assume that, for each traffic class $i \in \mathbb{M}$, there exists the steady-state occupancy probabilities of n $(0 \le n \le K_i)$ connections, P_n. The unique steady-state probability is determined by

$$P_n = P_0 \frac{(\lambda_i/\mu_i)^n}{n!} = \frac{P_0}{n!}\left(\frac{\lambda_i\sigma_i}{x_i}\right)^n, \quad n = 1, 2, \ldots, K_i. \tag{8}$$

Solving for P_0 in the equation $\sum_{n=0}^{K_i} P_n = 1$, we obtain that

$$P_0 = \left[\sum_{n=0}^{K_i} \frac{1}{n!}\left(\frac{\lambda_i}{\mu_i}\right)^n\right]^{-1} = \left[\sum_{n=0}^{K_i} \frac{1}{n!}\left(\frac{\lambda_i\sigma_i}{x_i}\right)^n\right]^{-1} \tag{9}$$

and then

$$P_n = \frac{1}{n!}\left(\frac{\lambda_i\sigma_i}{x_i}\right)^n \left[\sum_{j=0}^{K_i} \frac{1}{j!}\left(\frac{\lambda_i\sigma_i}{x_i}\right)^j\right]^{-1}, \quad \text{for } n = 1, 2, \ldots, K_i. \tag{10}$$

Thus, the blocking probability of incoming connections is formulated as

$$\mathcal{P}(x_i, K_i, y_i) = P_{K_i} = \frac{1}{K_i!}\left(\frac{y_i}{x_i}\right)^{K_i} \left[\sum_{n=0}^{K_i} \frac{1}{n!}\left(\frac{y_i}{x_i}\right)^n\right]^{-1}, \tag{11}$$

where y_i is the traffic demand and is defined by (7). Moreover, the expected (virtual) path occupancy in the steady state can be formulated as

$$\mathcal{L}(x_i, K_i, y_i) = \frac{y_i}{x_i}(1 - \mathcal{P}(x_i, K_i, y_i)) \tag{12}$$

from the definition of $\mathcal{L}(x_i, K_i, y_i) = \sum_{n=0}^{K_i} nP_n$. We assume the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is twice differentiable with respect to bandwidth x_i and traffic demand y_i, respectively.

Definition 4. The *average throughput* of class $i \in \mathbb{M}$ is defined as the average connection volume transmitted through those virtual paths in a unit of time [5]. That is, the average throughput $\overline{\Theta}_i$ of K_i virtual paths for class $i \in \mathbb{M}$, is

$$\overline{\Theta}_i \triangleq \sum_{n=1}^{K_i} n x_i P_n, \tag{13}$$

where P_n is the steady-state probabilities that there are n virtual paths occupied by connections, $1 \le n \le K_i$.

Remark 1. For each class $i \in \mathbb{M}$, let $\mathcal{L}(x_i, K_i, y_i)$ be the expected virtual path occupancy in the steady state. Then we have $\overline{\Theta}_i = x_i\mathcal{L}(x_i, K_i, y_i)$.

Definition 5. The *utilization level* of virtual paths is defined as the percentage of maximum throughput [9]. That is, the utilization level of K_i virtual paths is

$$U_i \triangleq \frac{\overline{\Theta}_i}{\Theta_i} = \frac{\mathcal{L}(x_i, K_i, y_i)}{K_i}. \tag{14}$$

3. Revenue management schemes

We consider the scenario in which network users access the core network via an ISP, where users belong to m different classes and the bandwidth received by each class is determined by a revenue management scheme. In this section, we develop two queueing-based models to determine the amount of bandwidth required by a connection for each traffic class that maximizes the revenue of the ISP, subject to satisfying the budget constraint. The main aim is to determine the bandwidth allocation that maximizes the average revenue/profit for the ISP, which allocates bandwidths given each user's willingness-to-pay.

3.1. Revenue Management Scheme I

Establishing a pricing scheme that charges the network users can regulate an overwhelming number of connections during congestion times [4]. Network users will be charged by ISPs to recover the operating cost based on the duration, the volume and the distance of a connection, which is incurred for setting and maintaining virtual paths [18]. Other than recovering costs, the ISP managers may handle congestion control by charging the users for regulating the mean occurrence rate λ_i of the connections during peak periods [19]. Network users are willing to pay a price because an excessive congestion would result in the inability of providing service to critical applications.

The operating costs can be determined by the types of traffic transmitted and the QoS guaranteed for such transfer, e.g., bandwidth allocation and blocking probability [4]. As far as QoS is concerned, bandwidth allocation $\mathbf{x}$ and blocking probability $\mathcal{P}(x_i, K_i, y_i)$ are the key elements of the pricing scheme. For example, the long-run average revenue per unit time $F_i(x_i, K_i, y_i)$ for a given traffic class $i \in \mathbb{M}$ can be expressed as follows [4]:

$$F_i(x_i, K_i, y_i) = c_i^t \mathcal{L}(x_i, K_i, y_i) + c_i^b \lambda_i x_i (1 - \mathcal{P}(x_i, K_i, y_i)), \tag{15}$$

where $c_i^b > 0$ is the cost charged for using per unit of bandwidth and $c_i^t > 0$ is the cost per unit of time for the sojourn time $1/\mu_i = \sigma_i/x_i$ on those virtual paths. The total long-run average revenue is obtained by summing over (15) for all traffic classes. In this paper, we illustrate the effect of changing x_i, y_i and K_i on the average revenue (15) by investigating properties of $\mathcal{P}(x_i, K_i, y_i)$ and $\mathcal{L}(x_i, K_i, y_i)$. Those phenomena will also be illustrated in the numerical results.

The network managers may want to maximize the long-run average revenue in their bandwidth sharing policies [4,13,20,21]. Here, we consider the long-run average revenue in (15) as the objective function for each traffic class. To maximize the average revenue under the budget constraint, we propose Revenue Management Scheme I as follows:

$$F = \max \sum_{i \in \mathbb{M}} w_i F_i(x_i, K_i, y_i)$$
$$\text{s.t. budget constraint (2)},$$

where the weight $0 < w_i < 1$ is assigned to traffic class $i \in \mathbb{M}$ by network managers.

3.2. Revenue Management Scheme II

In this subsection, we provide a congestion-based pricing scheme to allocate bandwidth while taking network users' utility into account. Bandwidth sharing in a network is frequently evaluated in terms of a utility function [19,22], etc. For example, Kelly et al. [16] proposed an optimization framework in which the objective is to maximize the total utility of all network users over their transmission rates. The utility $f_i(x_i)$ of a connection of class $i \in \mathbb{M}$ is assumed to be an increasing function of its bandwidth x_i, as introduced by Kelly et al. [16]. The shape of the utility function $f_i(x_i)$ depends on the network user's behavior. For example, utility functions of risk-averse users are different from those of risk-seeking users. Examples of possible utility functions are $f_i(x_i) = \log x_i$ for class $i \in \mathbb{M}$, leading to so-called proportional fairness in [16], and $f_i(x_i) = x_i^{1-\alpha}/(1-\alpha)$, $0 < \alpha < \infty$, for α-proportional fairness defined more generally by Mo and Walrand [23]. Max–Min fairness arises in the limit $\alpha \to \infty$ while proportional fairness corresponds to $\alpha \to 1$. In the limit $\alpha \to 0$, the objective is to maximize overall throughput by the detriment of fairness. More general notions of weighted fairness can be defined by multiplying the utility function with a class-dependent weight.

As introduced in [5,19,16], etc., network managers may consider the utility function $f_i(x_i) : [b_i^{\min}, b_i^{\max}] \to [0, 1]$ for each traffic class $i \in \mathbb{M}$, where $b_i^{\min}$ is the minimum bandwidth requirement, and $b_i^{\max}$ is the upper bound of bandwidth determined from (3). The utility function $f_i(x_i)$ is assumed to be continuous, increasing, and concave [5,16,24]. For example, [24] introduced the utility function

$$f_i(x_i) = \log_{\frac{a_i}{r_i}} \frac{x_i}{r_i}, \tag{16}$$

where a_i and r_i are the aspiration level and reservation level of bandwidth for class-i users, respectively, and they have $b_i^{\min} < r_i < a_i$. The logarithmic function is intimately associated with the concept of proportional fairness [16]. Depending on the specified reference levels, a_i and r_i, this utility function can be interpreted as a measure of the user's satisfaction with the value of the i-th criteria [24]. It is a strictly increasing function of bandwidth x_i, having value 1 if $x_i = a_i$, and value 0 if $x_i = r_i$. The utility function can map the different bandwidth requirement of traffic classes onto a normalized scale of the user's satisfaction.

Network managers' economic profit consists of all revenue gained by providing bandwidth x_i to class i and the opportunity cost through calculating the risk of blocking connections/users. Suppose that, for each traffic class $i \in \mathbb{M}$, network managers gain the payoff $p_i > 0$ for achieving the utility $f_i(x_i)$ by providing bandwidth x_i. Meanwhile, we introduce the opportunity cost of blocking connections/users. Let $q_i > 0$ be the opportunity cost of increasing blocking probability $\mathcal{P}_i(x_i, K_i, y_i)$ for traffic class $i \in \mathbb{M}$. A higher blocking probability will lead to a higher opportunity loss in the network manager's revenue. Those two criteria can be combined into single objective function with payoff p_i and opportunity cost q_i for all classes $i \in \mathbb{M}$, where payoff p_i and opportunity cost q_i can be applied in designing network pricing mechanisms. By economic definition, let

$$G_i(x_i, K_i, y_i) = p_i K_i f_i(x_i) - q_i \mathcal{P}_i(x_i, K_i, y_i) \tag{17}$$

be the managers' economic profit gained from class $i \in \mathbb{M}$, which represents its payoff minus the opportunity cost. Network managers determine the optimal bandwidth allocation under the budget constraint when taking both users' utility and blocking probability into account. Thus, through the weighted sum of all traffic classes, the optimal bandwidth allocation can be determined by the following optimization model, Revenue Management Scheme II:

$$G = \max \sum_{i \in \mathbb{M}} w_i G_i(x_i, K_i, y_i)$$
$$\text{s.t. budget constraint (2),}$$

where the weight $0 < w_i < 1$ is assigned to traffic class $i \in \mathbb{M}$ by network managers.

Remark 2. If the budget B is finite and (3) holds, the feasible set is bounded since the bandwidth allocated to each class i has a upper bound, $\forall i \in \mathbb{M}$. Moreover, the feasible set shrinks to an empty set if $\|\vec{K}\|_2 = (\sum_{i=1}^{m} K_i^2)^{1/2}$ increases to a sufficiently large number, where $\| \cdot \|_2$ denotes the well-known Euclidean norm on the vector space $\mathbb{R}^m$.

4. Monotonicity and convexity

In this section, we prove some monotone and convex properties of the blocking probability and expected path occupancy of connections. Those two revenue functions (15) and (17) can be analyzed through the monotonicity and convexity of the blocking probability and expected path occupancy.

Proposition 1 shows that the blocking probability is decreasing in bandwidth. The monotonicity can be proved by deriving the first partial derivative of the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ in (11). For ease of reading, detailed proofs of the following propositions are provided in the Appendix as the supplementary material for online publication only.

Proposition 1. *The blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is a decreasing function of bandwidth x_i, given $K_i \geq 1$ and $y_i > 0$ fixed.*

Furthermore, by deriving the second partial derivative of (11), it shows that the blocking probability is convex in bandwidth for a specific region.

Proposition 2. *For each $K_i \geq 1$ and $y_i > 0$, there exists a subset (or region) $\mathbb{S}_i$ of positive real numbers such that the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is convex (concave) in bandwidth x_i for all $x_i \in (\notin)\, \mathbb{S}_i$.*

As the number of virtual paths, K_i, is huge in real-world communication systems, Proposition 2 implies that $\mathcal{P}(x_i, K_i, y_i)$ is convex in bandwidth x_i if we have $0.5 < \mathcal{P}(x_i, K_i, y_i) \leq 1$. Otherwise, there exist two inflection points x_i^* and x_i^{**} when $0 \leq \mathcal{P}(x_i, K_i, y_i) < 0.5$.

In the following result, we demonstrate the monotone property of the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ with respect to allocated bandwidth x_i, i.e., $\partial \mathcal{L}(x_i, K_i, y_i)/\partial x_i < 0$, for all traffic class $i \in \mathbb{M}$.

Proposition 3. *If the traffic intensity $\rho_i = y_i/K_i x_i > 1$ holds in the case of large $K_i \gg 1$, the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ is a decreasing function of bandwidth x_i, given $y_i > 0$ fixed.*

Remark 3. Given $y_i > 0$ and $K_i \geq 1$ fixed, there exists an inflection point $\widetilde{x}_i$ such that for all $x_i \leq (\geq) \widetilde{x}_i$ the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ is concave (convex) in bandwidth x_i.

Remark 4. It can also be observed that the utilization level U_i defined in (14) is a decreasing function of bandwidth x_i for given $y_i > 0$ and $K_i \geq 1$. This is because the utilization level U_i equals the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ divided by K_i. Meanwhile, there exists the same inflection point $\widetilde{x}_i$ as in $\mathcal{L}(x_i, K_i, y_i)$ such that for all $x_i \leq (\geq) \widetilde{x}_i$ the utilization level U_i is concave (convex) in bandwidth x_i.

Next, we prove the monotone property of the blocking probability with respect to the traffic demand. The monotonicity can be observed by deriving those partial derivatives of the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ in (11) with respect to y_i.

Proposition 4. *The blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is increasing in traffic demand y_i, given $x_i > 0$ and $K_i \geq 1$ fixed.*

When traffic demand y_i exceeds the maximum throughput Θ_i, for class i, i.e., $\lambda_i \sigma_i > K_i x_i$, the number of blocked connections increases indefinitely. It is shown that per-flow QoS depends critically on whether the traffic demand y_i is less than or greater than the maximum throughput Θ_i.

Remark 5. For fixed $x_i > 0$ and $K_i \geq 1$, the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is convex (concave) in traffic demand y_i if we have $\Theta_i \geq (\leq) y_i$. Furthermore, the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ is an increasing and concave function of the traffic demand y_i, and the upper bound of $\mathcal{L}(x_i, K_i, y_i)$ is K_i no matter what y_i increases. The utilization level U_i in (14) is also an increasing and concave function of the traffic demand y_i, and the upper bound of U_i is 1 no matter what y_i increases.

The following monotone and convex properties of the blocking probability with respect to K_i are consistent with those results of the Erlang-B function proved by Messerli [25], Jagers and Van Doorn [26], and Esteves et al. [27]. For detailed proofs, interested readers may refer to [25–27].

Proposition 5. *The blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is decreasing with respect to the number of virtual paths K_i, given $x_i > 0$ and $y_i > 0$.*

Remark 6. For fixed $x_i > 0$ and $y_i > 0$, the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is convex in K_i, which is considered specifically in relaxation for continuous functions. Moreover, the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ is an increasing function of K_i given x_i and y_i are fixed. It can also be observed that the utilization level U_i is an increasing function of K_i because the increment of $\mathcal{L}(x_i, K_i, y_i)$ is larger than the increment of K_i when both is increased with K_i by one unit.

As mentioned in [4], there exists no closed-form algebraic expression of the optimal solution in Revenue Management Scheme I. Yacoubi et al. [4] plotted the revenue function (15) only and solved it numerically. To investigate the objective function (15), we derive and prove the monotonicity of the revenue function $F_i(x_i, K_i, y_i)$ in (15) with respect to model parameters x_i and K_i, individually, in the following results.

Theorem 1. *Let $\sigma_i > 0$ and $b_i^{\min} \geq 0$ be the mean connection volume and the minimum bandwidth requirement of class i, respectively. Given costs $c_i^t > 0$ and $c_i^b > 0$ in (15), if the allocated bandwidth*

$$x_i \geq \max\left\{\sqrt{\frac{c_i^t \sigma_i}{c_i^b}}, b_i^{\min}\right\} \tag{18}$$

holds for class $i \in \mathbb{M}$, the long-run average revenue $F_i(x_i, K_i, y_i)$ is increasing in bandwidth x_i, given $K_i \geq 1$ and $y_i > 0$.

Proof. From the minimum bandwidth requirement (1), it holds that feasible solution $x_i \geq b_i^{\min}$ for $i \in \mathbb{M}$. From the condition $x_i \geq \sqrt{c_i^t \sigma_i / c_i^b}$, it implies that

$$c_i^b \lambda_i - c_i^t \frac{y_i}{x_i^2} \geq 0,$$

where the traffic demand $y_i = \lambda_i \sigma_i$. From (12), we obtain that

$$\frac{\partial \mathcal{L}(x_i, K_i, y_i)}{\partial x_i} = -\frac{y_i}{x_i^2}(1 - \mathcal{P}(x_i, K_i, y_i)) + \frac{y_i}{x_i}\left(-\frac{\partial \mathcal{P}(x_i, K_i, y_i)}{\partial x_i}\right). \tag{19}$$

By Proposition 1, we know that $\partial \mathcal{P}(x_i, K_i, y_i)/\partial x_i < 0$. Hence, if $c_i^b \lambda_i \geq c_i^t y_i / x_i^2$, the first derivative of (15) with respect to x_i is

$$\begin{aligned}\frac{\partial F_i(x_i, K_i, y_i)}{\partial x_i} &= (1 - \mathcal{P}(x_i, K_i, y_i))\left[c_i^b \lambda_i - c_i^t \frac{y_i}{x_i^2}\right] - \frac{\partial \mathcal{P}(x_i, K_i, y_i)}{\partial x_i}\left[c_i^b \lambda_i x_i + c_i^t \frac{y_i}{x_i}\right] \\ &\geq 0,\end{aligned} \tag{20}$$

for all $K_i \geq 1$ and $y_i > 0$. So, the long-run average revenue $F_i(x_i, K_i, y_i)$ is an increasing function of bandwidth x_i given that the inequality (18) holds. □

Theorem 2. *The long-run average revenue $F_i(x_i, K_i, y_i)$ is an increasing function of the number of virtual paths K_i for each class $i \in \mathbb{M}$, given $x_i > b_i^{\min}$ and $y_i > 0$.*

Proof. From Proposition 5, we know that the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is a decreasing function of the number of virtual paths K_i. Moreover, the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ is an increasing function of K_i. Given $x_i > b_i^{\min}$ and $y_i > 0$, the increment of long-run average revenue (15) with respect to K_i is

$$\begin{aligned}F_i(x_i, K_i + 1, y_i) - F_i(x_i, K_i, y_i) &= c_i^t[\mathcal{L}(x_i, K_i + 1, y_i) - \mathcal{L}(x_i, K_i, y_i)] + c_i^b \lambda_i x_i[\mathcal{P}(x_i, K_i, y_i) - \mathcal{P}(x_i, K_i + 1, y_i)] \\ &\geq 0,\end{aligned} \tag{21}$$

for all $K_i \geq 1$. So, the long-run average revenue $F_i(x_i, K_i, y_i)$ is increasing in K_i for class $i \in \mathbb{M}$. □

Theorem 1 implies that the objective function of Revenue Management Scheme I is increasing in bandwidth x_i. Theorem 2 shows that the long-run average revenue $F_i(x_i, K_i, y_i)$ in Revenue Management Scheme I is also an increasing function of the number of virtual paths K_i. Those structural results on the long-run average revenue function can be helpful in the problem of maximizing the long-run average reward in communication networks with dynamic pricing [3,13,21].

Next, in the following results, we prove the monotonicity and convexity of the profit function $G_i(x_i, K_i, y_i)$.

Theorem 3. *The profit function $G_i(x_i, K_i, y_i)$ is increasing in bandwidth x_i, given $K_i \geq 1$ and $y_i > 0$.*

Proof. By Proposition 1, we have proved that $\partial \mathcal{P}(x_i, K_i, y_i)/\partial x_i < 0$. It can be derived that the first partial derivative of (17) with respect to x_i is

$$\frac{\partial G_i(x_i, K_i, y_i)}{\partial x_i} = \frac{p_i K_i}{x_i \log(a_i/r_i)} - q_i \frac{\partial \mathcal{P}(x_i, K_i, y_i)}{\partial x_i} \geq 0, \tag{22}$$

for all $K_i \geq 1$ and $y_i > 0$. So, the profit function $G_i(x_i, K_i, y_i)$ is an increasing function of bandwidth x_i. □

Theorem 4. *The profit function $G_i(x_i, K_i, y_i)$ is an decreasing function of the traffic demand y_i for each class $i \in \mathbb{M}$, given $x_i > b_i^{\min}$ and $K_i \geq 1$.*

Proof. By Proposition 4, we know that $\partial \mathcal{P}(x_i, K_i, y_i)/\partial y_i > 0$. Then the first partial derivative of (17) with respect to y_i is

$$\frac{\partial G_i(x_i, K_i, y_i)}{\partial y_i} = -q_i \frac{\partial \mathcal{P}(x_i, K_i, y_i)}{\partial y_i} \leq 0 \tag{23}$$

for all $K_i \geq 1$ and $x_i > b_i^{\min}$. So, $G_i(x_i, K_i, y_i)$ is an decreasing function of traffic demand y_i. □

Theorem 5. *The profit function $G_i(x_i, K_i, y_i)$ is an increasing function of the number of virtual paths K_i for each class $i \in \mathbb{M}$, given $x_i > b_i^{\min}$ and $y_i > 0$.*

Proof. From Proposition 5, we know that the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is a decreasing function of the number of virtual paths K_i. Given $x_i > b_i^{\min}$ and $y_i > 0$, the increment of the profit function (17) with respect to K_i is

$$G_i(x_i, K_i + 1, y_i) - G_i(x_i, K_i, y_i) = p_i f_i(x_i) + q_i[\mathcal{P}(x_i, K_i, y_i) - \mathcal{P}(x_i, K_i + 1, y_i)] \geq 0 \tag{24}$$

for all $K_i \geq 1$. So, the profit function $G_i(x_i, K_i, y_i)$ is increasing in K_i for class $i \in \mathbb{M}$. □

Theorem 6. *For each $K_i \geq 1$ and $y_i > 0$, there exists a region $\mathbb{S}_i$ of positive real numbers such that the profit function $G_i(x_i, K_i, y_i)$ is concave in bandwidth x_i for all $x_i \in \mathbb{S}_i$.*

Proof. From Proposition 2, there exists a region $\mathbb{S}_i$ of positive real numbers such that the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is convex in bandwidth x_i for all $x_i \in \mathbb{S}_i$. That is, for all $x_i \in \mathbb{S}_i$, we have $\partial^2 \mathcal{P}(x_i, K_i, y_i)/\partial x^2 \geq 0$. Then the second derivative of $G_i(x_i, K_i, y_i)$ with respect to x_i is

$$\frac{\partial^2 G_i(x_i, K_i, y_i)}{\partial x_i^2} = -\frac{p_i K_i}{x_i^2 \log(a_i/r_i)} - q_i \frac{\partial^2 \mathcal{P}(x_i, K_i, y_i)}{\partial x_i^2} \leq 0 \tag{25}$$

for all $K_i \geq 1$ and $y_i > 0$. So, the profit function $G_i(x_i, K_i, y_i)$ is concave in bandwidth x_i for all $x_i \in \mathbb{S}_i$. □

Remark 7. If network managers allocate bandwidth in a specific region $\mathbb{S}_i$ such that $\partial^2 \mathcal{P}(x_i, K_i, y_i)/\partial x_i^2 \geq 0$, the profit function $G_i(x_i, K_i, y_i)$ is concave in bandwidth x_i, which implies that the marginal profit is decreasing in such a region $\mathbb{S}_i$. Otherwise, network managers could gain insight into the convexity of $G_i(x_i, K_i, y_i)$ by determining the opportunity cost q_i if bandwidth is allocated such that $\partial^2 \mathcal{P}(x_i, K_i, y_i)/\partial x_i^2 < 0$. That is, in the case of $\partial^2 \mathcal{P}(x_i, K_i, y_i)/\partial x_i^2 < 0$, $G_i(x_i, K_i, y_i)$ becomes convex in bandwidth x_i if the opportunity cost q_i is sufficiently large, i.e.,

$$q_i > \frac{p_i K_i}{x_i^2 \log(a_i/r_i)(-\partial^2 \mathcal{P}(x_i, K_i, y_i)/\partial x_i^2)};$$

otherwise, $G_i(x_i, K_i, y_i)$ is still concave in bandwidth x_i.

5. Solution analysis

In this section, we present the optimality conditions for Revenue Management Scheme I and Revenue Management Scheme II, respectively. First, for Revenue Management Scheme I, we introduce the Lagrangian multiplier υ_1 and Lagrangian function $\psi_1(\mathbf{x}, \upsilon_1)$, where $\mathbf{x} = (x_1, \ldots, x_m)$. The Lagrangian function $\psi_1(\mathbf{x}, \upsilon_1)$ is defined as

$$\psi_1(\mathbf{x}, \upsilon_1) = \sum_{i \in \mathbb{M}} w_i F_i(x_i, K_i, y_i) + \upsilon_1 \left(\sum_{i \in \mathbb{M}} K_i c_i x_i - B \right). \tag{26}$$

Then, we can obtain

$$\frac{\partial \psi_1(\mathbf{x}, \upsilon_1)}{\partial x_i} = w_i \frac{\partial F_i(x_i, K_i, y_i)}{\partial x_i} + \upsilon_1 K_i c_i \tag{27}$$

for all $i \in \mathbb{M}$, where $\partial F_i(x_i, K_i, y_i)/\partial x_i$ is determined in (20), and

$$\frac{\partial \psi_1(\mathbf{x}, v_1)}{\partial v_1} = \sum_{i \in \mathbb{M}} K_i c_i x_i - B. \tag{28}$$

From $\partial \psi_1(\mathbf{x}, v_1)/\partial x_i = 0$ and $\partial \psi_1(\mathbf{x}, v_1)/\partial v_1 = 0$, we obtain the following Proposition 6. We can determine the optimal solution of Revenue Management Scheme I by solving the system of Eqs. (29) and (30) in Proposition 6.

Proposition 6. *In solving the optimal solutions for Revenue Management Scheme I, the optimal bandwidth allocation* $\mathbf{x} = (x_1, \ldots, x_m)$ *and Lagrangian multiplier* v_1 *satisfy*

$$v_1 = -\frac{w_i}{K_i c_i} \frac{\partial F_i(x_i, K_i, y_i)}{\partial x_i}, \ \forall i \in \mathbb{M}, \tag{29}$$

and

$$x_i = \frac{B - \sum_{j \neq i \in \mathbb{M}} K_j c_j x_j}{K_i c_i}, \quad \forall i \in \mathbb{M}, \tag{30}$$

where $\partial F_i(x_i, K_i, y_i)/\partial x_i$ *is determined in* (20).

Next, we introduce the Lagrangian multiplier v_2 and Lagrangian function $\psi_2(\mathbf{x}, v_2)$ for Revenue Management Scheme II. The Lagrangian function $\psi_2(\mathbf{x}, v_2)$ is defined as

$$\begin{aligned} \psi_2(\mathbf{x}, v_2) &= \sum_{i \in \mathbb{M}} w_i G_i(x_i, K_i, y_i) + v_2 \left(\sum_{i \in \mathbb{M}} K_i c_i x_i - B \right) \\ &= \sum_{i \in \mathbb{M}} w_i p_i K_i f_i(x_i) - \sum_{i \in \mathbb{M}} w_i q_i \mathcal{P}_i(x_i, K_i, y_i) + v_2 \left(\sum_{i \in \mathbb{M}} K_i c_i x_i - B \right). \end{aligned} \tag{31}$$

Then, we can obtain

$$\frac{\partial \psi_2(\mathbf{x}, v_2)}{\partial x_i} = \frac{w_i p_i K_i}{x_i \log(a_i/r_i)} - w_i q_i \frac{\partial \mathcal{P}_i(x_i, K_i, y_i)}{\partial x_i} + v_2 K_i c_i \tag{32}$$

for all $i \in \mathbb{M}$, and

$$\frac{\partial \psi_2(\mathbf{x}, v_2)}{\partial v_2} = \sum_{i \in \mathbb{M}} K_i c_i x_i - B. \tag{33}$$

From $\partial \psi_2(\mathbf{x}, v_2)/\partial x_i = 0$ and $\partial \psi_2(\mathbf{x}, v_2)/\partial v_2 = 0$, we can determine the optimal solution of Revenue Management Scheme II in the following Proposition 7.

Proposition 7. *In solving the optimal solutions for Revenue Management Scheme II, the optimal bandwidth allocation* $\mathbf{x} = (x_1, \ldots, x_m)$ *and Lagrangian multiplier* v_2 *satisfy*

$$v_2 = -\frac{w_i}{K_i c_i} \left[\frac{p_i K_i}{x_i \log(a_i/r_i)} - q_i \left(\frac{y_i}{x_i} - K_i \right) \frac{y_i^{K_i} e^{-y_i/x_i}}{K_i! x_i^{K_i+1}} \right], \quad \forall i \in \mathbb{M}, \tag{34}$$

and

$$x_i = \frac{B - \sum_{j \neq i \in \mathbb{M}} K_j c_j x_j}{K_i c_i}, \quad \forall i \in \mathbb{M}. \tag{35}$$

In practice, it is complicated to solve the system of Eqs. (34) and (35) because the blocking probability $\mathcal{P}_i(x_i, K_i, y_i)$ in (11) is a nonlinear function of bandwidth x_i. If we omit temporarily the consideration of blocking probability $\mathcal{P}_i(x_i, K_i, y_i)$ in the profit function $G_i(x_i, K_i, y_i)$ defined in (17), the objective function of Revenue Management Scheme II will be reduced to the weighted sum of utility functions only. Then, Revenue Management Scheme II can be reduced to a utility maximization problem as studied in [28]. It implies that the objective function of Revenue Management Scheme II can be rewritten according to [28] as follows:

$$\sum_{i \in \mathbb{M}} w_i p_i K_i f_i(x_i) = \sum_{i \in \mathbb{M}} K_i \left(w_i p_i \log_{\frac{a_i}{r_i}} \frac{x_i}{r_i} \right) = \sum_{i \in \mathbb{M}} K_i \left(\mathcal{D}_i \log x_i - \mathcal{C}_i \right),$$

where $\mathcal{D}_i = w_i p_i / \log(a_i/r_i)$ and $\mathcal{C}_i = w_i p_i \log r_i / \log(a_i/r_i)$ are constant for all $i \in \mathbb{M}$. Hence, the utility maximization problem under budget constraint can be formulated as follows:

$$\begin{aligned} \max \quad & \sum_{i \in \mathbb{M}} K_i \left(\mathcal{D}_i \log x_i - \mathcal{C}_i\right) \\ \text{s.t.} \quad & \text{budget constraint (2).} \end{aligned} \tag{36}$$

Then we may relax the result in Proposition 7 to the following Corollary 1, which is an optimal solution to the utility maximization problem (36).

Corollary 1. *In the case of omitting the blocking probability from Revenue Management Scheme II, the optimal bandwidth allocation* $(\hat{x}_1, \ldots, \hat{x}_m)$ *and Lagrangian multiplier* $\hat{v}_2$ *for the utility maximization problem (36) satisfy*

$$\hat{v}_2 = -\frac{\sum_{i \in \mathbb{M}} K_i \mathcal{D}_i}{B}, \tag{37}$$

and

$$\hat{x}_i = \frac{B\mathcal{D}_i}{c_i \sum_{i \in \mathbb{M}} K_i \mathcal{D}_i}, \quad \forall i \in \mathbb{M}, \tag{38}$$

where $\mathcal{D}_i = w_i p_i / \log(a_i/r_i)$ *is constant.*

Note that Corollary 1 is the same as the result of Proposition 3 studied in Guan et al. [28]. The feasible set of the utility maximization problem in Guan et al. [28] is the same as that of Revenue Management Scheme II. Observe that the utility function $f_i(x_i)$ is increasing in bandwidth x_i. Moreover, in Proposition 1, we have shown that the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is a decreasing function of bandwidth x_i. It implies that the objective function of Revenue Management Scheme II is increasing in bandwidth x_i. Hence, we can determine the optimal solution of Revenue Management Scheme II efficiently from the optimal solution of the utility maximization problem (36).

According to Corollary 1, we present an optimal solution $\hat{x}_i$ to the utility maximization problem (36) given parameters K_i, c_i and p_i corresponding to each class i, $\forall i \in \mathbb{M}$. Let $\hat{\mathbf{x}}$ and $\mathbf{x}$ denote vectors with their component $\hat{x}_i$ and x_i respectively. For the Lagrangian function $\psi_2(\mathbf{x}, v_2)$ defined in (31), we denote its gradient vector by $\nabla\psi_2(\mathbf{x}, v_2)$ and Hessian matrix by $\nabla^2\psi_2(\mathbf{x}, v_2)$ at a point $(\mathbf{x}, v_2) \in \mathbb{R}^{m+1}$. By using (32) and (33), we can determine

$$\frac{\partial^2 \psi_2(\mathbf{x}, v_2)}{\partial x_i^2} = \frac{-K_i \mathcal{D}_i}{x_i^2} - w_i q_i \frac{\partial^2 \mathcal{P}_i(x_i, K_i, y_i)}{\partial x_i^2}, \tag{39}$$

$$\frac{\partial^2 \psi_2(\mathbf{x}, v_2)}{\partial x_i \partial v_2} = \frac{\partial^2 \psi_2(\mathbf{x}, v_2)}{\partial v_2 \partial x_i} = K_i c_i, \tag{40}$$

and

$$\frac{\partial^2 \psi_2(\mathbf{x}, v_2)}{\partial v_2^2} = 0. \tag{41}$$

From the above equations, we can compute the gradient vector $\nabla\psi_2(\mathbf{x}, v_2)$ and Hessian matrix $\nabla^2\psi_2(\mathbf{x}, v_2)$. Therefore, for a given $\hat{\mathbf{x}}$ obtained from Corollary 1, we can determine an optimal objective value and the corresponding optimal bandwidth $\mathbf{x} = (x_1, \ldots, x_m)$ as shown in (34) and (35) by using the following algorithm.

A solution algorithm: *(Optimal bandwidth of Management Scheme II)*

Initialization Set iteration $k = 0$. Apply Corollary 1, we set the initial solution $x_i^0 = B\mathcal{D}_i/(c_i \sum_{i \in \mathbb{M}} K_i \mathcal{D}_i)$, $\forall i \in \mathbb{M}$ and $v_2^0 = -\sum_{i \in \mathbb{M}} K_i \mathcal{D}_i / B$. Set $\mathbf{x}^0 = (x_1^0, \ldots, x_m^0)$, and choose a sufficiently small number $\xi > 0$.

Step 1 Compute the gradient vector $\nabla\psi_2(\mathbf{x}^k, v_2^k)$ at point $(\mathbf{x}^k, v_2^k)$.

Step 2 If $\nabla\psi_2(\mathbf{x}^k, v_2^k) = 0$ or

$$\|\psi_2(\mathbf{x}^k, v_2^k) - \psi_2(\mathbf{x}^{k-1}, v_2^{k-1})\| < \xi,$$

then stop with an optimal (or near optimal) solution $\mathbf{x}^k$ and v_2^k. Otherwise, go to the next step.

Step 3 Compute the Hessian matrix $\nabla^2\psi_2(\mathbf{x}^k, v_2^k)$ and its inverse

$$\mathbf{H}^k = (\nabla^2\psi_2(\mathbf{x}^k, v_2^k))^{-1}.$$

Step 4 Perform the translation

$$(\mathbf{x}^{k+1}, v_2^{k+1}) \leftarrow (\mathbf{x}^k, v_2^k) - \mathbf{H}^k \nabla\psi_2(\mathbf{x}^k, v_2^k).$$

Set $k \leftarrow k + 1$ and go to Step 1.

This algorithm defines a series of $(\mathbf{x}^k, \upsilon_2^k)$'s, starting from an initial solution $(\hat{\mathbf{x}}, \hat{\upsilon}_2)$, such that the sequence converges toward an optimal solution which satisfies $\nabla\psi_2 = 0$. The initial solution $(\hat{\mathbf{x}}, \hat{\upsilon}_2)$ obtained from Corollary 1 is a feasible solution to Revenue Management Scheme II because it satisfies the budget constraint (2). In Step 1, the gradient vector $\nabla\psi_2(\mathbf{x}^k, \upsilon_2^k)$ can be easily calculated through Eqs. (32) and (33) given a point $(\mathbf{x}^k, \upsilon_2^k)$. From Eq. (39)-(41), we can construct the symmetric Hessian matrix $\nabla^2\psi_2(\mathbf{x}^k, \upsilon_2^k)$ and determine its inverse $\mathbf{H}^k$ in Step 3, where $\nabla^2\psi_2(\mathbf{x}^k, \upsilon_2^k)$ and $\mathbf{H}^k$ are $(m+1)$-by-$(m+1)$ matrices. Note that the computational time needed to find $\nabla^2\psi_2(\mathbf{x}^k, \upsilon_2^k)$ and $\mathbf{H}^k$ is short because the total number of traffic classes, m, is small in the core network. The descent direction is given by $-\mathbf{H}^k\nabla\psi_2(\mathbf{x}^k, \upsilon_2^k)$ at the point $(\mathbf{x}^k, \upsilon_2^k)$. With the good initial point $(\hat{\mathbf{x}}, \hat{\upsilon}_2)$, the sequence $(\mathbf{x}^k, \upsilon_2^k)$ defined in Step 4 converges to the root of $\nabla\psi_2(\mathbf{x}, \upsilon_2) = 0$. The stopping criterion in Step 2 implies that we terminate the algorithm with an optimal solution which satisfies the optimal condition in Proposition 7. Or, we terminate the algorithm early when the improvement of $\psi_2(\mathbf{x}^k, \upsilon_2^k)$ in (31) is less than a sufficiently small number ξ. The convergence of the algorithm is quadratic, and it converges quickly to the (near) optimal solution because the initial point $(\hat{\mathbf{x}}, \hat{\upsilon}_2)$ is easily directed to the root of $\nabla\psi_2 = 0$ since the monotone property is applied to the case.

6. Applications

In this section, we present connection's blocking elasticity to bandwidth for each traffic class to illustrate how the expressions of $\mathcal{P}(x_i, K_i, y_i)$ can be used with some applications. The term elasticity was introduced to the networking research community by Shenker [29]. Based on the investigation of elasticity, one can develop distributed pricing algorithm that takes user's elasticity into consideration [10,18,19,30]. By using the concept of elasticity, we can define the elasticity of the blocking probability with respect to bandwidth as follows.

Definition 6. The *bandwidth elasticity of blocking* is defined as

$$\varepsilon_i^b \triangleq \frac{\Delta\mathcal{P}(x_i, K_i, y_i)/\mathcal{P}(x_i, K_i, y_i)}{\Delta x_i/x_i}, \tag{42}$$

where x_i is the allocated bandwidth, Δx_i is the change of allocated bandwidth, $\mathcal{P}(x_i, K_i, y_i)$ is the blocking probability given predetermined number of virtual paths K_i, and $\Delta\mathcal{P}(x_i, K_i, y_i)$ is the change of the blocking probability.

The bandwidth elasticity of blocking (42) can be rewritten as

$$\varepsilon_i^b = \frac{x_i}{\mathcal{P}(x_i, K_i, y_i)}\frac{\partial\mathcal{P}(x_i, K_i, y_i)}{\partial x_i}. \tag{43}$$

The elasticity ε_i^b represents the percentage change of the blocking probability in response to a percent change of bandwidth. Proposition 8 shows the phenomenon that the blocking probability will decrease if the allocated bandwidth increases.

Proposition 8. *The bandwidth elasticity of blocking ε_i^b is negative for all $x_i > 0$.*

As $K_i \gg 1$, Proposition 9 shows that the bandwidth elasticity of blocking ε_i^b will decrease when the allocated bandwidth x_i increases.

Proposition 9. *The bandwidth elasticity of blocking ε_i^b is decreasing in bandwidth x_i for all $x_i > 0$ as $K_i \gg 1$.*

As applications of elasticity in economics, we present the demand elasticity of blocking ε_i^d and the capacity elasticity of blocking ε_i^c for each traffic class $i \in M$ in the following. The *demand elasticity of blocking* is formulated as

$$\varepsilon_i^d \triangleq \frac{\Delta\mathcal{P}(x_i, K_i, y_i)/\mathcal{P}(x_i, K_i, y_i)}{\Delta y_i/y_i} = \frac{y_i}{\mathcal{P}(x_i, K_i, y_i)}\frac{\partial\mathcal{P}(x_i, K_i, y_i)}{\partial y_i}, \tag{44}$$

where Δy_i is the change in the traffic demand. In addition, the *capacity elasticity of blocking* is written as

$$\varepsilon_i^c \triangleq \frac{\Delta\mathcal{P}(x_i, K_i, y_i)/\mathcal{P}(x_i, K_i, y_i)}{\Delta K_i/K_i} = \frac{K_i}{\mathcal{P}(x_i, K_i, y_i)}\frac{\Delta\mathcal{P}(x_i, K_i, y_i)}{\Delta K_i}, \tag{45}$$

where ΔK_i is the change in the number of virtual paths. Similarly, the properties of the demand elasticity of blocking ε_i^d and the capacity elasticity of blocking ε_i^c are derived as follows.

Proposition 10. *The demand elasticity of blocking ε_i^d is non-negative for all traffic demand $y_i \geq 0$.*

Proposition 11. *The capacity elasticity of blocking ε_i^c is non-positive for all $K_i \geq 1$ and decreasing as the number of virtual paths K_i increases.*

Proposition 10 infers that the blocking probability will increase as the traffic demand increases. Proposition 11 concludes that the blocking probability is decreasing as increasing the number of virtual paths. Those properties can be easily derived from the results of Propositions 4 and 5.

Next, as another application, we present monotone and convex properties of the blocking probability with respect to traffic intensity ρ_i defined in (6). In (11), if we replace $y_i/(K_i x_i)$ by traffic intensity ρ_i, the equivalent expression of (11) is

$$\mathcal{P}(\rho_i, K_i) = \frac{(K_i\rho_i)^{K_i}}{K_i!}\left[\sum_{n=0}^{K_i}\frac{(K_i\rho_i)^n}{n!}\right]^{-1}. \tag{46}$$

Meanwhile, the expected path occupancy (12) can be rewritten as

$$\mathcal{L}(\rho_i, K_i) = K_i\rho_i(1-\mathcal{P}(\rho_i, K_i)). \tag{47}$$

After a little algebra, we derive the first and second derivatives of (46) with respect to ρ_i as follows:

$$\frac{\partial \mathcal{P}(\rho_i, K_i)}{\partial \rho_i} = \frac{K_i\mathcal{P}(\rho_i, K_i)}{\rho_i}[1-\rho_i+\rho_i\mathcal{P}(\rho_i, K_i)] \tag{48}$$

and

$$\frac{\partial^2 \mathcal{P}(\rho_i, K_i)}{\partial \rho_i^2} = \frac{K_i\mathcal{P}(\rho_i, K_i)}{\rho_i^2}[1-\rho_i+\rho_i\mathcal{P}(\rho_i, K_i)]^2\mathcal{H}(\rho_i), \tag{49}$$

where

$$\mathcal{H}(\rho_i) = K_i + \frac{K_i\rho_i\mathcal{P}(\rho_i, K_i)}{1-\rho_i+\rho_i\mathcal{P}(\rho_i, K_i)} - \frac{1}{[1-\rho_i+\rho_i\mathcal{P}(\rho_i, K_i)]^2}. \tag{50}$$

The monotone and convex properties of the blocking probability in (46) are listed below, which are consistent with those properties shown by Harel [17]. The blocking probability in (46) is convex in the traffic intensity ρ_i if ρ_i is below certain inflection point ρ_i^* and concave if ρ_i is greater than ρ_i^*. For detailed derivations, interested readers may refer to Harel [17], Bakry [31].

Theorem 7. *For each $K_i \ge 1$, the blocking probability $\mathcal{P}(\rho_i, K_i)$ is an increasing function of traffic intensity $\rho_i > 0$.*

Theorem 8. *For each $K_i \ge 1$, there exists an inflection point ρ_i^* such that for all $\rho_i < (>)\rho_i^*$, the blocking probability $\mathcal{P}(\rho_i, K_i)$ is convex (concave) in traffic intensity $\rho_i > 0$.*

Remark 8. For each $K_i \ge 1$, the expected path occupancy $\mathcal{L}(\rho_i, K_i)$ is increasing and concave in traffic intensity ρ_i, and the upper bound of $\mathcal{L}(\rho_i, K_i)$ is K_i. Moreover, the utilization level U_i is also increasing and concave in traffic intensity ρ_i, and the upper bound is 1.

In real world cases, the number of connections on networks is always huge. Next, we investigate the blocking probability defined in (11) in the case of large $K_i \gg 1$. If the traffic intensity $\rho_i = y_i/(K_i x_i) < 1$ holds, Eq. (11) can be rewritten as

$$\mathcal{P}(x_i, K_i, y_i) = \frac{(y_i/x_i)^{K_i}}{K_i!}\left[e^{y_i/x_i} - \mathcal{R}(K_i)\right]^{-1} \approx \frac{y_i^{K_i}e^{-y_i/x_i}}{K_i!x_i^{K_i}}, \quad \text{as } K_i \gg 1, \tag{51}$$

where $\mathcal{R}(K_i)$ is the K_ith-degree Taylor remainder term of e^{y_i/x_i}. The remainder term $\mathcal{R}(K_i) \approx 0$ as $K_i \gg 1$. Moreover, we can conclude that

$$\mathcal{L}(x_i, K_i, y_i) \approx \frac{y_i}{x_i}\left(1 - \frac{y_i^{K_i}e^{-y_i/x_i}}{K_i!x_i^{K_i}}\right), \quad \text{as } K_i \gg 1. \tag{52}$$

From (51) and (52), it implies that

$$\frac{\partial \mathcal{P}(x_i, K_i, y_i)}{\partial x_i} = \left(\frac{y_i}{x_i^2} - \frac{K_i}{x_i}\right)\cdot\mathcal{P}(x_i, K_i, y_i), \tag{53}$$

$$\frac{\partial \mathcal{P}(x_i, K_i, y_i)}{\partial y_i} = \left(\frac{K_i}{y_i} - \frac{1}{x_i}\right)\cdot\mathcal{P}(x_i, K_i, y_i), \tag{54}$$

$$\begin{aligned}\frac{\partial^2 \mathcal{P}(x_i, K_i, y_i)}{\partial x_i\partial y_i} &= \frac{\partial^2 \mathcal{P}(x_i, K_i, y_i)}{\partial y_i\partial x_i}\\ &= \frac{1}{x_i}\left(1+2K_i-\frac{y_i}{x_i}-\frac{K_i^2x_i}{y_i}\right)\cdot\mathcal{P}(x_i, K_i, y_i);\end{aligned} \tag{55}$$

$$\frac{\partial \mathcal{L}(x_i, K_i, y_i)}{\partial x_i} = -\frac{y_i}{x_i^2}\left[1-\mathcal{P}(x_i, K_i, y_i)\cdot\left(1+K_i-\frac{y_i}{x_i}\right)\right], \tag{56}$$

$$\frac{\partial \mathcal{L}(x_i, K_i, y_i)}{\partial y_i} = \frac{1}{x_i}\left[1-\mathcal{P}(x_i, K_i, y_i)\cdot\left(1+K_i-\frac{y_i}{x_i}\right)\right], \tag{57}$$

and

$$\frac{\partial^2 \mathcal{L}(x_i, K_i, y_i)}{\partial x_i \partial y_i} = \frac{\partial^2 \mathcal{L}(x_i, K_i, y_i)}{\partial y_i \partial x_i}$$
$$= -\frac{1}{x_i^2}\left\{1 - \mathcal{P}(x_i, K_i, y_i) \cdot \left[\left(1 + K_i - \frac{y_i}{x_i}\right)^2 - \frac{y_i}{x_i}\right]\right\} \quad (58)$$

in the case of large $K_i \gg 1$ and traffic intensity $\rho_i < 1$.

In the case of large $K_i \gg 1$, Proposition 1 can be restated as follows.

Corollary 2. *If $K_i \gg 1$ and $\rho_i < 1$ holds, the first derivative of blocking probability $\mathcal{P}(x_i, K_i, y_i)$ with respect to bandwidth x_i is always negative for all $x_i > 0$, i.e.,*

$$\frac{\partial \mathcal{P}(x_i, K_i, y_i)}{\partial x_i} = \left(\frac{y_i}{x_i} - K_i\right) \frac{y_i^{K_i} e^{-y_i/x_i}}{K_i! x_i^{K_i+1}} < 0. \quad (59)$$

As a final application, we present a budget ratio to investigate the budget allocation among different classes in the network management schemes. Using the maximum throughput Θ_i defined in (4) for class $i \in \mathbb{M}$, the budget constraint (2) can be represented as

$$\sum_{i \in \mathbb{M}} \Theta_i + \tau = \Phi(\vec{K}, B, \mathbb{G}), \quad (60)$$

where $\tau \geq 0$ is the reserved bandwidth, and $\Phi(\vec{K}, B, \mathbb{G})$ is the total available bandwidth purchased with limited budget B for preset numbers of virtual paths $\vec{K} = (K_1, \ldots, K_m)$ on the core network $\mathbb{G}$. Then, the budget ratio for each traffic class $i \in \mathbb{M}$ is given below.

Definition 7. A *budget ratio* allocated to class i is defined as the fraction of total available bandwidth $\Phi(\vec{K}, B, \mathbb{G}) < \infty$ allocated for class $i \in \mathbb{M}$. That is, the budget ratio allocated for class $i \in \mathbb{M}$ is

$$B_i \triangleq \frac{\Theta_i}{\Phi(\vec{K}, B, \mathbb{G})}, \quad (61)$$

where the maximum throughput $\Theta_i = K_i x_i$.

If there exists sufficient bandwidth, the reserved bandwidth $\tau > 0$ may be shared among all the ongoing connections with budget ratio B_i for each class $i \in \mathbb{M}$ at the online process. That is, any remaining bandwidth is shared according to the budget ratio B_i. In case there is no bandwidth reserved ($\tau = 0$), the allocated bandwidth x_i will decrease if K_i increases when those preset numbers of virtual paths, $K_{i'}$, $i' \neq i$, of other classes are fixed.

Remark 9. If the maximum throughput Θ_i increases, from (60), the value $\sum_{i' \neq i} \Theta_{i'}$ will decrease. Since those numbers $K_{i'}$ are fixed for all $i' \neq i$, the bandwidth $x_{i'}$ will decrease. That is, for certain class $i' \neq i$, the bandwidth $x_{i'}$ may be snatched by class i as the maximum throughput Θ_i of class i becomes larger.

7. Numerical illustrations

7.1. Experimental setting

In this section, we present numerical results to show the optimal bandwidth allocation for different traffic classes under the budget constraint. Here, we select four traffic classes as test examples from statistical data monitored at the Cooperative Association for Internet Data Analysis (CAIDA) [32]. Connections of class 4 have the highest priority, and traffic class 1 is given the lowest priority. The number of connections in the high-priority traffic class is often less than that in the low-priority class, but the traffic demand and bandwidth requirement of high-priority traffic class are always larger than those of low-priority traffic class. Those parameters are summarized in Table 1, including class weight w_i, minimum bandwidth requirement $b_i^{\min}$ (Mbps), aspiration level a_i (Mbps), reservation level r_i (Mbps), the average cost c_i (cents) of one unit bandwidth through class i's virtual paths, number of virtual paths K_i, mean occurrence rate λ_i, the connection volume σ_i (Mb), cost charged for using per unit of bandwidth c_i^b (cents), cost per unit of sojourn time c_i^t (cents), payoff p_i (cents) and opportunity cost q_i (cents).

7.2. Comparison between management scheme I and management scheme II

The allocated bandwidth is determined by solving Revenue Management Scheme I and Revenue Management Scheme II, respectively. Table 2 shows those optimal bandwidth allocation and optimal values of Management Scheme I and

Table 1
Characteristics of each traffic class.

i	w_i	$b_i^{\min}$	a_i	r_i	y_i	c_i	K_i	λ_i	σ_i	c_i^t	c_i^b	p_i	q_i
1	0.1	0.8	1.32	0.84	59.1	1.0	96	96	0.62	4	130	130	4
2	0.2	1.0	1.85	1.18	96.5	1.3	75	75	1.29	6	180	180	6
3	0.3	2.0	2.79	1.73	185.4	1.8	42	42	4.41	12	320	320	12
4	0.4	2.5	3.58	2.28	324.3	2.5	27	27	12.01	20	500	500	20

Table 2
Budget versus optimal bandwidth allocation x_i.

Budget B	Scheme I		Scheme II	
	Optimal bandwidth allocation (x_1, x_2, x_3, x_4)	Revenue F	Optimal bandwidth allocation (x_1, x_2, x_3, x_4)	Profit G
500	(0.80, 1.00, 2.08, 2.50)	10,141	(0.80, 1.00, 2.00, 2.59)	1,412
550	(0.80, 1.00, 2.74, 2.50)	12,898	(0.80, 1.08, 2.08, 3.11)	4,480
600	(0.80, 1.00, 3.40, 2.50)	16,254	(0.80, 1.20, 2.30, 3.44)	7,187
650	(0.80, 1.00, 4.06, 2.50)	19,997	(0.80, 1.31, 2.52, 3.77)	9,646
700	(0.80, 1.00, 4.72, 2.50)	23,787	(0.80, 1.42, 2.74, 4.10)	11,900
750	(0.80, 1.00, 5.38, 2.50)	27,284	(0.80, 1.54, 2.97, 4.43)	13,980
800	(0.80, 1.00, 6.04, 2.50)	30,379	(0.80, 1.65, 3.19, 4.76)	15,911
850	(0.80, 1.00, 6.71, 2.50)	33,203	(0.82, 1.76, 3.40, 5.07)	17,714
900	(0.80, 1.00, 7.37, 2.50)	35,911	(0.87, 1.87, 3.60, 5.37)	19,412
950	(0.80, 1.00, 8.03, 2.50)	38,584	(0.92, 1.97, 3.80, 5.67)	21,018
1000	(0.80, 1.00, 8.69, 2.50)	41,247	(0.97, 2.07, 4.00, 5.97)	22,542

Table 3
Budget versus blocking probability of allocated bandwidth.

Budget B	Blocking probability in Scheme I $(\mathcal{P}(x_1, K_1, y_1), \mathcal{P}(x_2, K_2, y_2), \mathcal{P}(x_3, K_3, y_3), \mathcal{P}(x_4, K_4, y_4))$	Blocking probability in Scheme II $(\mathcal{P}(x_1, K_1, y_1), \mathcal{P}(x_2, K_2, y_2), \mathcal{P}(x_3, K_3, y_3), \mathcal{P}(x_4, K_4, y_4))$
500	(0.0020, 0.2504, 0.5389, 0.7938)	(0.0020, 0.2504, 0.5553, 0.7869)
550	(0.0020, 0.2504, 0.4004, 0.7938)	(0.0020, 0.1977, 0.5372, 0.7440)
600	(0.0020, 0.2504, 0.2720, 0.7938)	(0.0020, 0.1311, 0.4903, 0.7174)
650	(0.0020, 0.2504, 0.1616, 0.7938)	(0.0020, 0.0768, 0.4442, 0.6910)
700	(0.0020, 0.2504, 0.0795, 0.7938)	(0.0020, 0.0383, 0.3989, 0.6648)
750	(0.0020, 0.2504, 0.0311, 0.7938)	(0.0020, 0.0158, 0.3547, 0.6386)
800	(0.0020, 0.2504, 0.0096, 0.7938)	(0.0020, 0.0054, 0.3119, 0.6127)
850	(0.0020, 0.2504, 0.0025, 0.7938)	(0.0010, 0.0016, 0.2725, 0.5881)
900	(0.0020, 0.2504, 0.0006, 0.7938)	(0.0002, 0.0005, 0.2368, 0.5649)
950	(0.0020, 0.2504, 0.0001, 0.7938)	(0.0000, 0.0001, 0.2030, 0.5419)
1000	(0.0020, 0.2504, 0.0000, 0.7938)	(0.0000, 0.0000, 0.1714, 0.5191)

Management II, where the total budget B varies from \$500 to \$1000. For each available budget B, according to (38), the initial solutions $x_i^0 = B\mathcal{D}_i/(c_i \sum_{i=1}^{4} K_i\mathcal{D}_i)$ for all $i = 1, \ldots, 4$, are given in the optimization process. In can be observed from Table 2 that those bandwidth allocations and optimal values are increasing when enlarging the available budget. From Figs. 1–3, we graphically illustrate the effect of changing budget B on optimal bandwidth allocation and optimal values of those two schemes. In Fig. 1, it shows that almost all the available resources are allocated in the direction of class 3 in Management Scheme I, and the others are allocated to satisfy the minimum bandwidth requirements only. However, in Management Scheme II, it can be seen in Fig. 2 that the available resource is allocated to all classes proportionally. Both optimal values of Management Scheme I and Management Scheme II are increasing in the total budget B. In Fig. 3, we find that there exists an inflection point such that the optimal revenues of Management Scheme I are concave up when the budget B is smaller than the inflection point, and the optimal revenue is concave down if the budget exceeds the inflection point. However, the optimal profit of Management Scheme II is expressed in logarithmic form. The marginal (optimal) profit obtained by solving Management Scheme II is decreasing with respect to the available budget B.

Next, we compare bandwidth sharing policies between Scheme I and Scheme II by showing the blocking probability and budget ratio for four traffic classes. Table 3 presents those blocking probabilities determined by those optimal bandwidth allocation of two management schemes. For Management Scheme I, it can be observed from Fig. 4 that the blocking probability of class 3 is decreasing when the budget B increases from \$500 to \$1000 while other classes' blocking probabilities remain unimproved. However, in Management Scheme II, Fig. 5 shows that those blocking probabilities of four traffic classes are decreasing proportionally when increasing the total budget.

Given a budget B, the budget ratio for each traffic class is determined from the definition in (61) as follows: $B_i = K_ic_ix_i/\sum_{i=1}^{4} K_ic_ix_i$, for $i = 1, \ldots, 4$, where bandwidth x_i is determined by solving Management Scheme I and Management Scheme II, respectively. Table 4 summarizes those numerical results of the budget ratio for four traffic classes when increasing the available budget B from \$500 to \$1000. It can be seen from Fig. 6 that most of the available budget is

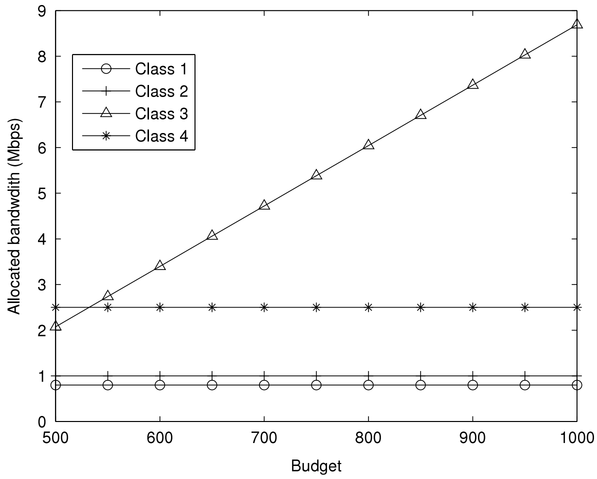

Fig. 1. Budget versus optimal bandwidth allocation of Revenue Management Scheme I.

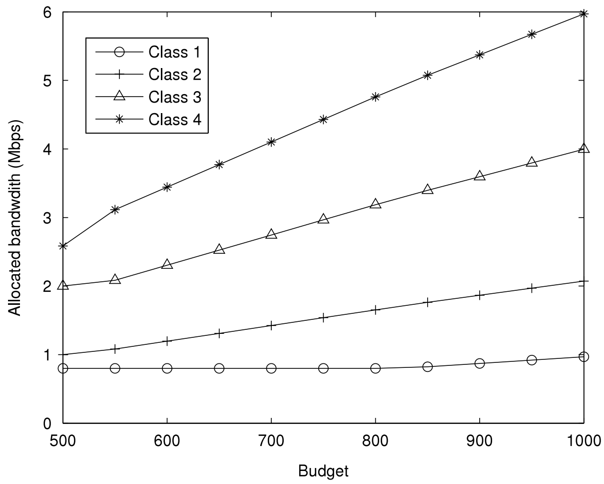

Fig. 2. Budget versus optimal bandwidth allocation of Revenue Management Scheme II.

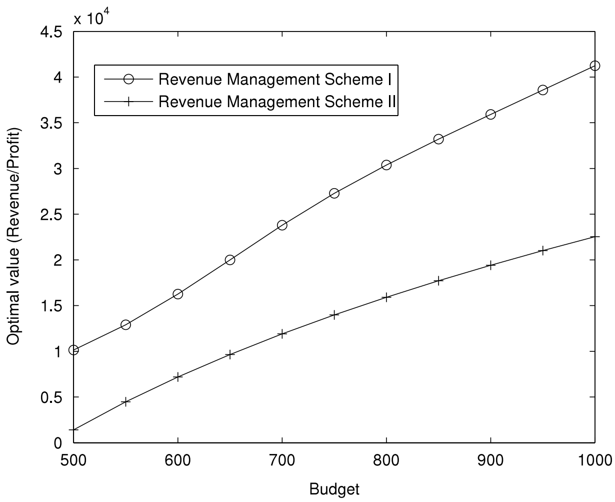

Fig. 3. Budget versus optimal values of two revenue management schemes.

Table 4
Budget versus budget ratio for four traffic classes.

Budget B	Budget ratio in Scheme I (B_1, B_2, B_3, B_4)	Budget ratio in Scheme II (B_1, B_2, B_3, B_4)
500	(15.36%, 19.50%, 31.39%, 33.75%)	(15.36%, 19.50%, 30.24%, 34.90%)
550	(13.96%, 17.73%, 37.63%, 30.68%)	(13.96%, 19.17%, 28.65%, 38.22%)
600	(12.80%, 16.25%, 42.82%, 28.13%)	(12.80%, 19.43%, 29.03%, 38.73%)
650	(11.82%, 15.00%, 47.22%, 25.96%)	(11.82%, 19.65%, 29.36%, 39.17%)
700	(10.97%, 13.93%, 50.99%, 24.11%)	(10.97%, 19.84%, 29.64%, 39.55%)
750	(10.24%, 13.00%, 54.26%, 22.50%)	(10.24%, 20.00%, 29.89%, 39.87%)
800	(9.60%, 12.19%, 57.12%, 21.09%)	(9.60%, 20.14%, 30.10%, 40.16%)
850	(9.04%, 11.47%, 59.64%, 19.85%)	(9.29%, 20.21%, 30.20%, 40.29%)
900	(8.53%, 10.83%, 61.88%, 18.75%)	(9.29%, 20.21%, 30.20%, 40.29%)
950	(8.08%, 10.26%, 63.89%, 17.76%)	(9.29%, 20.21%, 30.20%, 40.29%)
1000	(7.68%, 9.75%, 65.69%, 16.88%)	(9.29%, 20.21%, 30.20%, 40.29%)

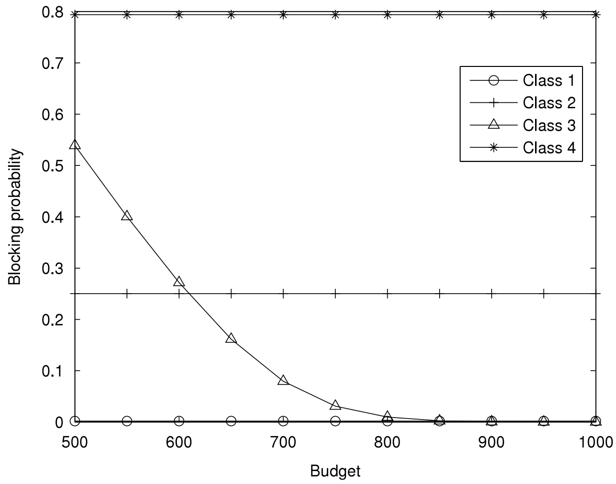

Fig. 4. Budget versus blocking probability determined by optimal solutions of Revenue Management Scheme I.

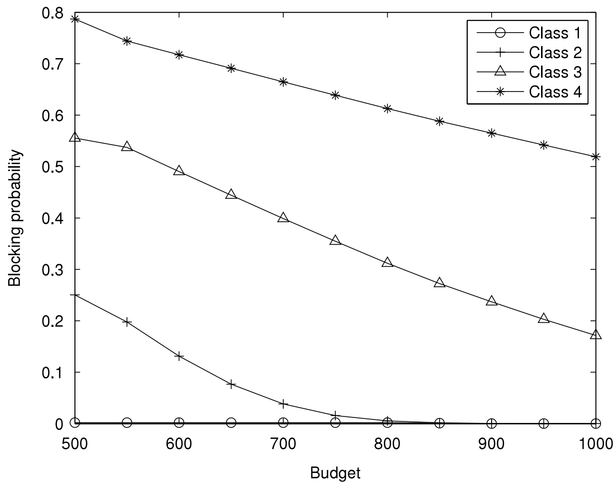

Fig. 5. Budget versus blocking probability determined by optimal solutions of Revenue Management Scheme II.

allocated to class 3, and the other classes only get the minimum bandwidth to satisfy the feasibility. This is because the marginal improvement of the objective function in Management Scheme I is the largest in the direction of class 3, i.e., $\partial w_3 F_3(x_3, K_3, y_3)/\partial x_3 \geq \partial w_i F_i(x_i, K_i, y_i)/\partial x_i$ for $i = 1, 2, 4$. Fig. 7 shows that all the budget is allocated to four traffic classes proportionally. The budget ratios for four traffic classes are almost invariable when varying budget B from \$500 to \$1000.

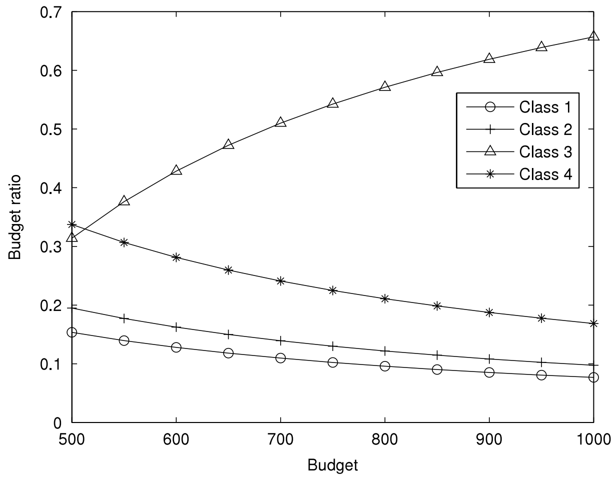

Fig. 6. Budget versus budget ratio determined by optimal solutions of Revenue Management Scheme I.

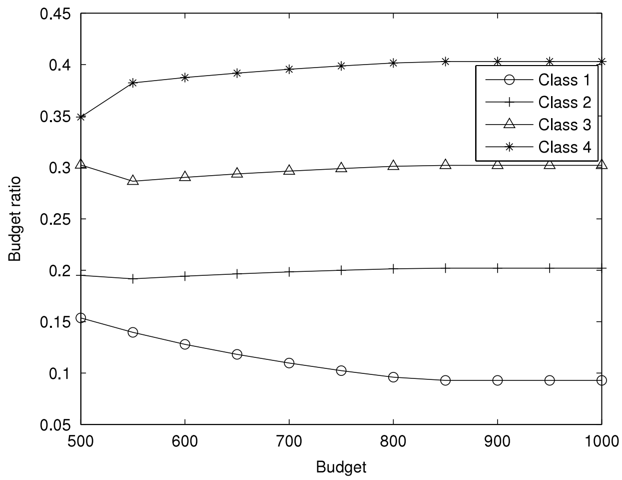

Fig. 7. Budget versus budget ratio determined by optimal solutions of Revenue Management Scheme II.

7.3. Sensitivity analysis

In this subsection, we present sensitivity analysis of the blocking probability and system utilization to illustrate numerically those monotone and convex properties of objective functions in Management Schemes I and II. First, we observe the effect of changing bandwidth x_i on the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ defined in (11). To conduct the sensitivity analysis, we check the bandwidth x_i from 0.1 Mbps to 8 Mbps, and other parameters remain fixed as listed in Table 1. It can be seen from Fig. 8 that the blocking probability is decreasing and convex when increasing bandwidth, which are consistent with those theoretical results given in Propositions 1 and 2.

Next, we show the effect of changing bandwidth x_i on the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$, average throughput $\overline{\Theta}_i$ and utilization level U_i, respectively. These three performance measures have been represented as functions of the blocking probability according to (12)–(14). From Fig. 9, we find that the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ in (12) is a decreasing function of bandwidth x_i. In addition, it can be seen from class 1 or class 2 in Fig. 9 that there exists an inflection point $\widetilde{x}_i$ such that for all $x_i \leq (\geq)\widetilde{x}_i$, the expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ is concave (convex) in bandwidth x_i. Those monotone and convex properties have been summarized in Proposition 3 and Remark 3. Furthermore, it can be observed from Fig. 10 that average throughput $\overline{\Theta}_i$ defined in (13) is increasing in bandwidth x_i for four traffic classes. In Fig. 11, it shows the effect of changing bandwidth on the utilization levels of preset virtual paths for four traffic classes. Proposition 3 infers that those utilization levels U_i defined in (14) are decreasing when enlarging bandwidth x_i, which can be seen numerically in Fig. 11. Moreover, it can be observed clearly from class 1 or class 2 that there exists an inflection point $\widetilde{x}_i$ such that U_i is concave (convex) in bandwidth x_i for all $x_i < (>)\widetilde{x}_i$.

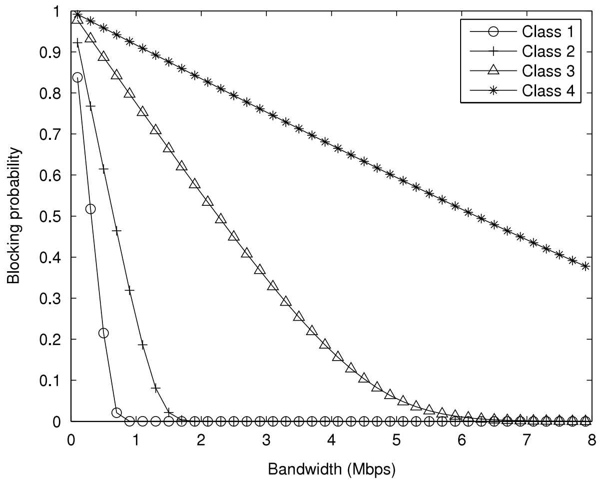

Fig. 8. Blocking probability $\mathcal{P}(x_i, K_i, y_i)$ versus bandwidth x_i for four traffic classes.

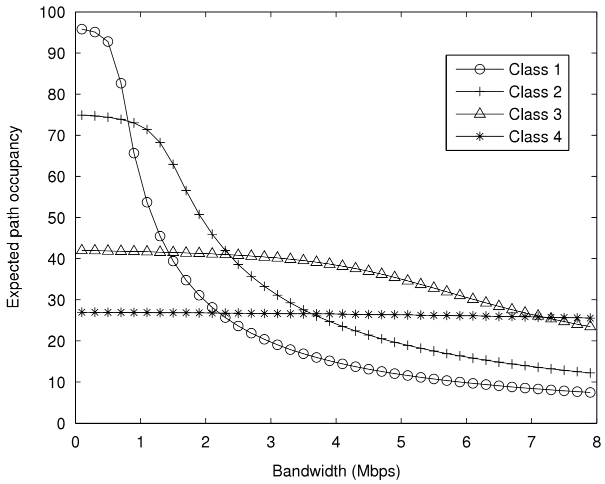

Fig. 9. Expected path occupancy $\mathcal{L}(x_i, K_i, y_i)$ versus bandwidth x_i for four traffic classes.

In the following, we present numerical analysis of revenue function (15) in Management Scheme I and profit function (17) in Management Scheme II, individually. Those numerical results are shown in Figs. 12 and 13. The results can graphically illustrate monotone and convex relationships which have been proven in Theorems 1–6.

It has been proven in Theorem 1 that the objective function $F_i(x_i, K_i, y_i)$ of Revenue Management Scheme I is increasing in bandwidth x_i if it satisfies the inequality (18). It can be seen from Fig. 12 that $F_1(x_1, K_1, y_1)$ is increasing in bandwidth x_1 for all bandwidth $x_1 \geq \max\{\sqrt{c_1^t \sigma_1 / c_1^b}, b_1^{\min}\} = 0.8$ Mbps. Similarly, $F_2(x_2, K_2, y_2)$ is increasing in bandwidth x_2 for all bandwidth $x_2 \geq 1$ Mbps, and so on. We find that the convexity of average revenue (15) fluctuates when bandwidth x_i is small corresponding to other system parameters. Proposition 2 infers that, for each traffic class i, there exists a region $\mathbb{S}_i$ of bandwidth such that the blocking probability $\mathcal{P}(x_i, K_i, y_i)$ is convex (concave) for all $x_i \in (\notin)\mathbb{S}_i$, where the region $\mathbb{S}_i$ can be constructed from the proof of Proposition 2. From numerical experiments, we find that if the budget B or bandwidth x_i is large enough, those revenue function $F_i(x_i, K_i, y_i)$ will become increasing and concave.

Finally, we illustrate the effect on the profit $G_i(x_i, K_i, y_i)$ in Revenue Management Scheme II when increasing bandwidth x_i. It can be observed from Fig. 13 that the economic profit $G_i(x_i, K_i, y_i)$ defined in (17) increases for all bandwidth x_i, which has already been proved in Theorem 3. Theorem 6 infers that the profit $G_i(x_i, K_i, y_i)$ is concave for all $x_i \leq 8$ Mbps, which can be seen obviously in Fig. 13.

7.4. Summary

Two revenue management schemes have been investigated theoretically and numerically to determine the amount of bandwidth required by a connection for each traffic class. Given network users' willingness-to-pay and other system parameters, our aim is to determine the bandwidth allocation that maximizes the average revenue/profit for the ISP under the budget constraint.

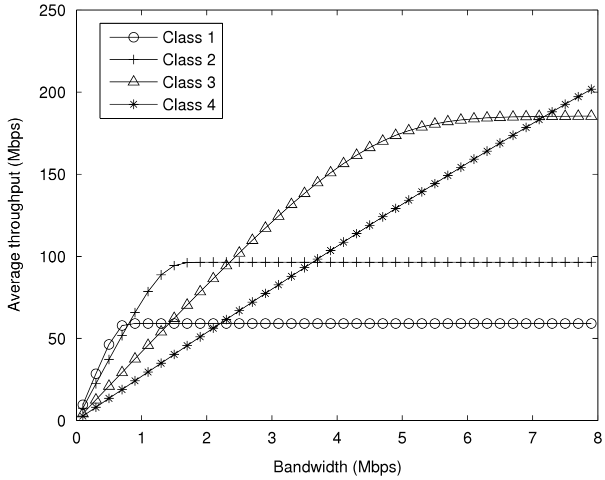

Fig. 10. Average throughput $\overline{\Theta}_i$ versus bandwidth x_i for four traffic classes.

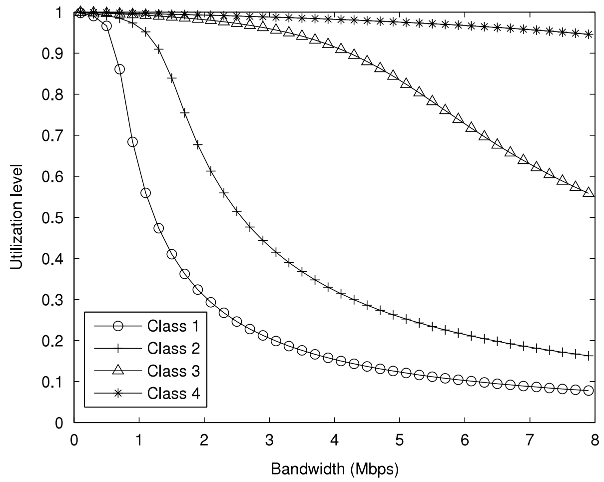

Fig. 11. Utilization level U_i versus bandwidth x_i for four traffic classes.

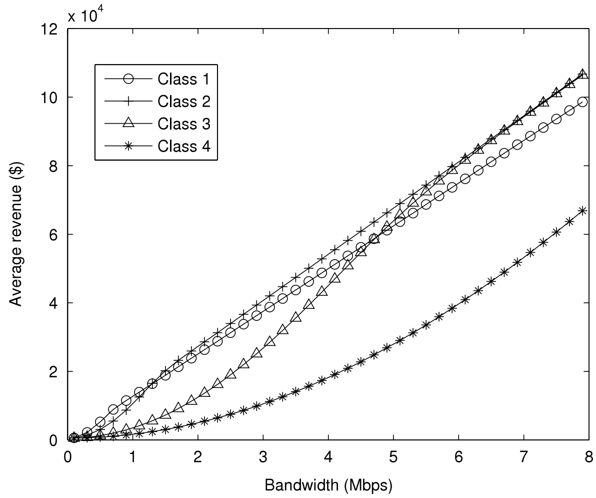

Fig. 12. Average revenue $F_i(x_i, K_i, y_i)$ versus bandwidth x_i for four traffic classes.

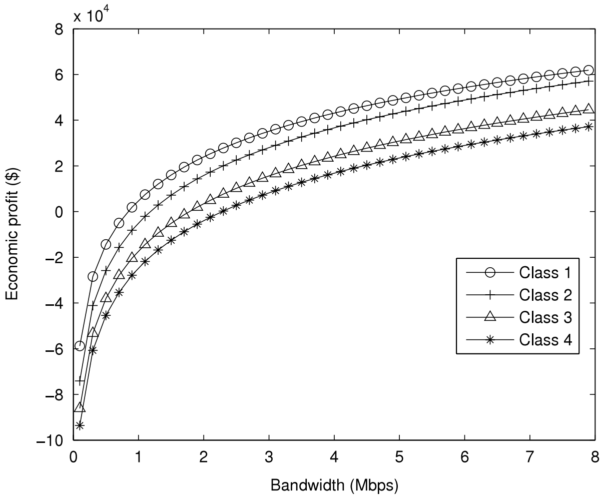

Fig. 13. Economic profit $G_i(x_i, K_i, y_i)$ versus bandwidth x_i for four traffic classes.

In Management Scheme I, almost all the resources are allocated to only one class whose marginal revenue is the largest, and the remainder are allocated to other classes to meet their feasibility only. That is, most of the available budget is allocated to certain traffic class i with the largest marginal improvement $\partial w_i F_i(x_i, K_i, y_i)/\partial x_i$ in Management Scheme I. Network managers may apply Management Scheme I to allocate limited resources among competing classes in order to maximize the weighted sum of average revenue. On the other hand, by solving Management Scheme II, all resources are allocated proportionally to four traffic classes. With the help of the utility function in (16), we can achieve the proportional fairness by allocating bandwidth through Management Scheme II.

To investigate these two bandwidth allocation policies, monotone and convex properties of the revenue/profit function as well as the blocking probability have been proven theoretically in previous sections and illustrated numerically in this section. Those phenomena in numerical experiments are consistent with theoretical results. In practice, those results may help network managers to determine their optimal/acceptable bandwidth allocation according to one of those two management schemes.

8. Conclusions

In this paper, we consider the revenue management problems on communication networks with multi-class traffic under the budget constraint. Two revenue management schemes have been investigated through the monotone and convex properties of the blocking probability and expected path occupancy of connections. We analyze the sensitivity of the blocking probability to model parameters, where the parameters change one-at-a-time. Under general assumptions, we have proved that the blocking probability is directionally (i) decreasing in bandwidth, (ii) convex in bandwidth for specific regions, (iii) increasing in traffic demand, and (iv) decreasing in the number of virtual paths. We also demonstrate the monotone and convex relations among the expected path occupancy and those model parameters. Furthermore, we prove that for a fixed number of virtual paths, the blocking probability is increasing and convex in traffic intensity for specific regions.

The optimality conditions are derived to obtain an optimal bandwidth allocation for two revenue management schemes. A solution algorithm is also developed to allocate limited budget among competing traffic classes. We have conducted the sensitivity analysis of the average revenue function and the economic profit function for a given traffic class by changing bandwidth allocation, traffic demand and the available number of virtual paths respectively. Those results have also been verified with numerical examples interpreting the blocking probability, utilization level, average revenue, etc. The relationship between blocking probability and bandwidth allocation can help network managers to design network pricing mechanisms for sharing bandwidth in terms of blocking/congestion costs.

The contribution of the current paper is the analysis of those monotone and convex relations among model parameters and performance measures of interest. The results of this work may be helpful in the operational processes involved in the efficient set-up and usage of a core network under the budget constraint, e.g., network design and provisioning purposes. One application of the relationship between blocking probability and bandwidth allocation may be referred to as designing network pricing mechanisms for sharing bandwidth in terms of blocking/congestion costs, whose examples were given by Yacoubi et al. [4] and Anderson et al. [19], etc. The closed-form expression of the blocking probability in terms of bandwidth can also be used to investigate the optimal buffer size in capacitated communication systems so that the blocking probability is kept below a specific threshold [15]. Another application of this work is used to consider the admission control in networks under different bandwidth sharing policies including throughput maximization, max–min fairness, proportional fairness and balanced fairness, etc. Interested readers may refer to Bonald et al. [5], Nilsson and Pióro [33], Jordan [34], etc.

In addition, we present three elasticities to investigate the effect of changing model parameters on the average revenue in analysis of economic models. The sensitivity results derived here could be used to guide development of congestion-based pricing of network resources, and to adjust bandwidth in the optimal proportion in response to changes in desired levels of blocking probability. Future work will be conducted in the direction of further investigation for the network revenue management schemes. Much additional work would have to be done in the future to make such an approach practical, e.g., design of reservation protocols, scheduling policies, measurement algorithms, and feedback algorithms to guarantee convergence.

Acknowledgements

The authors thank two anonymous referees for their valuable and constructive comments. This research was supported in part by National Science Council, Taiwan, R.O.C., under grant number NSC-98-2221-E-004-001-MY2.

Appendix. Proofs of propositions

Detailed proofs of propositions and corollaries can be found online at doi:10.1016/j.camwa.2011.05.024.

References

[1] E. Mingozzi, G. Stea, M.A. Callejo-Rodríguez, J. Enríquez-Gabeiras, G. García-de-Blas, F.J. Ramón-Salquero, W. Burakowski, A. Beben, J. Sliwinski, H. Tarasiuk, O. Dugeon, M. Diaz, L. Baresse, E. Monteiro, EuQoS: end-to-end quality of service over heterogeneous networks, Computer Communications 32 (2009) 1355–1370.
[2] H. Cho, A. Girard, C. Rosenberg, On the advantages of optimal end-to-end QoS budget partitioning, Telecommunication Systems 34 (2007) 91–106.
[3] N. Jin, S. Jordan, The effect of bandwidth and buffer pricing on resource allocation and QoS, Computer Networks 46 (2004) 53–71.
[4] M. Yacoubi, M. Emelianenko, N. Gautam, Pricing in next generation networks: a queuing model to guarantee QoS, Performance Evaluation 52 (2003) 59–84.
[5] T. Bonald, L. Massoulié, A. Proutière, J. Virtamo, A queueing analysis of max–min fairness, proportional fairness and balanced fairness, Queueing Systems 53 (2006) 65–84.
[6] P. Nain, Qualitative properties of the Erlang blocking model with heterogeneous user requirements, Queueing Systems 6 (2) (1990) 189–206.
[7] N. Antunes, C. Fricker, P. Robert, D. Tibi, Analysis of loss networks with routing, The Annals of Applied Probability 16 (4) (2006) 2007–2026.
[8] T. Güven, R.J. La, M.A. Shayman, B. Bhattacharjee, A unified framework for multipath routing for unicast and multicast traffic, IEEE/ACM Transactions on Networking 16 (5) (2008) 1038–1051.
[9] A. Faragó, Efficient blocking probability computation of complex traffic flows for network dimensioning, Computers & Operations Research 35 (2008) 3834–3847.
[10] C. Maglaras, A. Zeevi, Pricing and capacity sizing for systems with shared resources: approximate solutions and scaling relations, Management Science 49 (8) (2003) 1018–1038.
[11] C. Bruni, F.D. Priscoli, G. Koch, I. Marchetti, Resource management in network dynamics: an optimal approach to the admission control problem, Computers & Mathematics with Applications 59 (2010) 305–318.
[12] T. Bonald, A. Proutière, On performance bounds for balanced fairness, Performance Evaluation 55 (2004) 25–50.
[13] T. Aktaran-Kalayci, H. Ayhan, Sensitivity of optimal prices to system parameters in a steady-state service facility, European Journal of Operational Research 193 (2009) 120–128.
[14] M.Kr. Dutta, V.K. Chaubey, Performance analysis of all-optical WDM network with wavelength converter using Erlang *C* traffic model, Communications in Computer and Information Science 70 (2010) 238–244.
[15] J.M. Smith, $M/G/c/K$ blocking probability models and system performance, Performance Evaluation 52 (2003) 237–267.
[16] F.P. Kelly, A.K. Maulloo, D.K.H. Tan, Rate control for communication networks: shadow prices, proportional fairness and stability, Journal of the Operational Research Society 49 (1998) 237–252.
[17] A. Harel, Convexity properties of the Erlang loss formula, Operations Research 38 (3) (1990) 499–505.
[18] S. Shakkottai, R. Srikant, Economics of network pricing with multiple ISPs, IEEE/ACM Transactions on Networking 14 (6) (2006) 1233–1245.
[19] E. Anderson, F.P. Kelly, R. Steinberg, A contract and balancing mechanism for sharing capacity in a communication network, Management Science 52 (1) (2006) 39–53.
[20] I.Ch. Paschalidis, Y. Liu, Pricing in multiservice loss networks: static pricing, asymptotic optimality, and demand substitution effects, IEEE/ACM Transactions on Networking 10 (3) (2002) 425–438.
[21] G. Zachariadis, J.A. Barria, Dynamic pricing and resource allocation using revenue management for multiservice networks, IEEE/ACM Transactions on Networking 5 (4) (2008) 215–226.
[22] S. van Hoesel, Optimization in telecommunication networks, Statistica Neerlandica 59 (2) (2005) 180–205.
[23] J. Mo, J. Walrand, Fair end-to-end window-based congestion control, IEEE/ACM Transactions on Networking 8 (5) (2000) 556–567.
[24] C.H. Wang, H. Luh, A fair QoS scheme for bandwidth allocation by precomputation-based approach, International Journal of Information and Management Sciences 19 (3) (2008) 391–412.
[25] E. Messerli, Proof of a convexity property of the Erlang *B* formula, Bell System Technical Journal 51 (1972) 951–953.
[26] A.A. Jagers, E.A. Van Doorn, On the continued Erlang loss function, Operations Research Letters 5 (1) (1986) 43–46.
[27] J.S. Esteves, J. Craveirinha, D. Cardoso, Computing Erlang-*B* function derivatives in the number of servers, Communications in Statistics–Stochastic Models 11 (2) (1995) 311–331.
[28] Y. Guan, W. Yang, H. Owen, D.M. Blough, A pricing approach for bandwidth allocation in differentiated service networks, Computers & Operations Research 35 (2008) 3769–3786.
[29] S. Shenker, Fundamental design issues for the future Internet, IEEE Journal on Selected Areas in Communications 13 (1995) 1176–1188.
[30] M. Yuksel, S. Kalyanaraman, Elasticity considerations for optimal pricing of networks, in: Proceedings of IEEE Symposium on Computers and Communications I, 2003, pp. 163–168.
[31] S.H. Bakry, A new method for computing Erlang-*B* formula, Computers & Mathematics with Applications 19 (2) (1990) 73–74.
[32] The Cooperative Association for Internet Data Analysis. [Online]. Available: http://www.caida.org/home.
[33] P. Nilsson, M. Pióro, Solving dimensioning tasks for proportionally fair networks carrying elastic traffic, Performance Evaluation 49 (2002) 371–386.
[34] S. Jordan, A recursive algorithm for bandwidth partitioning, IEEE Transactions on Communications 58 (4) (2010) 1026–1030.