Optimization of Bandwidth Allocation in Communication Networks with Penalty Cost[*]

Jun Wu[1,3], Wuyi Yue[2], and Shouyang Wang[3]

[1] Institute of Intelligent Information and Communications Technology,
Konan University, Kobe 658-8501 Japan
`wujun@iss.ac.cn`
[2] Department of Information Science and Systems Engineering,
Konan University, Kobe 658-8501 Japan
`yue@konan-u.ac.jp`
[3] Academy of Mathematics and Systems Science,
Chinese Academy of Sciences, Beijing 100080 China
`sywang@amss.ac.cn`

Abstract. In this paper, we present an analysis on optimization and risk management in Communication Networks (CNs). The model is proposed for offline traffic engineering optimization, which takes a centralized view of bandwidth allocation, performance control, and risk control of network profit shortfall. First, we introduce a linear penalty cost in the CN optimization model and derive the optimal bandwidth capacity with the penalty cost. Then, we use the mean-variance approach to analyze the profit shortfall risk in CNs. Finally, numerical results are shown to reveal the impact of the penalty cost on the CNs performance.

1 Introduction

Traffic engineering is a process to optimize resource utilization and network performance [1], [2]. It has greatly improved network performance by using the emerging technologies, such as Multi-Path Label Switching and Optical Channel Trails [3], [4].

There are two forms of traffic engineering: online planning and offline planning. In past works, the offline optimization problem was formulated as a deterministic Multi-Commodity Flow (MCF) model, where demand of each channel was given as a fixed quantity [5], [6].

Recently, there were some works concerning with the stochastic traffic engineering. Mitra and Wang presented a stochastic traffic engineering framework for optimizing bandwidth provisioning and path selection in CNs [7]. Mitra and Wang also developed an optimization framework for the network service provider

to manage profit in a two-tier market [8]. Mitra and Wang furthered their studies in [7], [8] and developed the efficient frontier of mean revenue and revenue risk [9]. Wu, Yue and Wang presented a stochastic model for optimizing bandwidth allocation in [10], which took a centralized view of bandwidth allocation, performance control, and risk of network profit shortfall. They analyzed the loss rate constraint and risk averseness in the CN optimization model and showed the impact of the loss rate constraint and the risk averseness on the network performance.

In this paper, based on the model presented in [10], we introduce a linear penalty cost in the optimization model for network bandwidth allocation. Whenever there is unsatisfied traffic demand with the limitation of network bandwidth, a linear penalty function will be added in the objective function. Next, we study the risk of the network profit shortfall by using the mean-variance approach [11], [12]. Finally, numerical results are shown to reveal the impact of the penalty cost on the network performance.

The rest of this paper is organized as follows. In Section 2, we present the system model that we consider in this paper and present the notations and preliminaries. In Section 3, we formulate the optimization model and derive the optimal bandwidth capacity with the linear penalty cost. In Section 4, we analyze the network profit shortfall risk by using the mean-variance approach. In Section 5, we give some numerical results to show the impact of the penalty cost on the network performance. Conclusions are given in Section 6.

2   System Model

A Communication Network (CN) is supposed to have users and a service provider. The CN should derive its unit revenue by serving demand including voice, packet data, image and full-motion video. For unit bandwidth capacity allocated to the network, a unit cost will be charged. For unsatisfied traffic demand with the limitation of network bandwidth, a linear penalty cost will be added in the objective function. The objective of this system is to maximize the CN mean profit of the network.

Let (N, L) denote a CN composed of nodes v_i $(v_i \in N, 1 \leq i \leq N)$ and links l $(l \in L)$, where N is the set of all nodes and L is the set of all links. Let V denote the set of all node pairs, $v \in V$ denote an arbitrary node pair where $v = (v_i, v_j)$ and $v_i, v_j \in N$, C_l denote the maximal bandwidth capacity of link l, $R(v)$ denote an admissible route set for $v \in V$, ξ_s $(s \in R(v))$ denote the amount of capacity provisioned on route s, D_v $(v \in V)$ denote the traffic load between node pair $v \in V$, b_v $(v \in V)$ denote the amount of bandwidth capacity provisioned between node pair v, which can be routed on one or more routes, then $b_v = \sum_{s \in R(v)}(\xi_s)$.

In this paper, we consider the CN to be a whole system. We let $b > 0$ denote the amount of bandwidth capacity provisioned in the CN, then we have $b = \sum_{v \in V}(b_v)$. Let $D > 0$ denote the traffic demand in the whole CN, then we have $D = \sum_{v \in V}(D_v)$, which is characterized by a random distribution with

its probability density function $f(x)$ and cumulative distribution function $F(x)$. $b \wedge D$ is the actual traffic load transmitted in the CN, where $\wedge$ represents the choice of the smaller between b and D. Let r denote the unit revenue by serving the traffic demand, so the total revenue of the CN is $r \times (b \wedge D)$. Let c denote the unit cost for unit bandwidth capacity allocated in the CN, so the total cost is $c \times b$. Let q denote the linear penalty cost for each unsatisfied traffic demand, so the total penalty cost is $q \times (D - b)^+$, where "$+$" represents the choice of the positive part of $(D - b)$. To avoid unrealistic and trivial cases, we assume that $r > q > 0$, $r > c > 0$.

3   Optimal Bandwidth Capacity in the CN with Penalty Cost

Based on the model presented in [10], in this paper, we add a linear penalty cost in the optimization model for network bandwidth allocation to evaluate the system performance. We let q denote the unit penalty cost for the unsatisfied traffic demand.

Let $\pi(b, D)$ denote the random profit function by serving traffic demand in the CN with the linear penalty cost, namely,

$$\pi(b, D) = r(b \wedge D) - q(D - b)^+ - cb. \tag{1}$$

Let $\Pi(b, D)$ denote the mean profit function with the linear penalty cost as follows:

$$\Pi(b, D) = r \int_0^b x f(x) dx + rb \int_b^{+\infty} f(x) dx$$
$$-q \int_b^{+\infty} (x - b) f(x) dx - cb. \tag{2}$$

The objective function of the system is

$$\Pi^* = \max_{b>0} \{\Pi(b, D)\} \tag{3}$$

subject to

$$P(b \geq \delta D) \geq 1 - \epsilon \tag{4}$$

and

$$b \leq C_{max} \tag{5}$$

where Π^* is the optimal profit function. $P(b \geq \delta D) \geq 1 - \epsilon$ is the loss rate constraint with δ ($0 \leq \delta \leq 1$) and ϵ ($0 \leq \epsilon \leq 1$) as the parameters defined in [10]. $C_{max} > 0$ is the maximal capacity that can be allocated in the CN.

With the above formulation, we can derive the optimal capacity of bandwidth allocation. First, we analyze the property of the CN mean profit function $\Pi(b, D)$.

The first order derivative of $\Pi(b, D)$ with respect to b is given as follows:

$$\frac{d\Pi(b, D)}{db} = (r + q - c) - (r + q)F(b). \tag{6}$$

The second order derivative of $\Pi(b, D)$ with respect to b is given as follows:

$$\frac{d^2\Pi(b, D)}{db^2} = -(r + q)f(b). \tag{7}$$

With the assumptions in Section 2, we know that $f(b) \geq 0$, $r + q > 0$, hence,

$$\frac{d^2\Pi(b, D)}{db^2} \leq 0. \tag{8}$$

Therefore, we can say that $\Pi(b, D)$ is a concave function of b. So, the optimal bandwidth capacity without constraints is

$$F^{-1}\left(\frac{r + q - c}{r + q}\right) \tag{9}$$

where $F^{-1}(\cdot)$ is the inverse function of $F(\cdot)$.

Finally, if we consider the loss rate constraint and the maximal capacity constraint as in [10], the optimal bandwidth capacity for the CN is

$$\left[F^{-1}\left(\frac{r + q - c}{r + q}\right) \vee \delta F^{-1}(1 - \epsilon)\right] \wedge C_{max}. \tag{10}$$

where $\vee$ represents the choice of the larger value between the two components.

4   Risk Analysis in the CN with Penalty Cost

The mean-variance analysis, which was first introduced by Markowitz [11], had been a standard tool in financial risk management. It uses a parameter α ($0 \leq \alpha \leq 1$) to characterize the risk averseness, which is a quantitative balance between the mean profit and the risk of its shortfall [12]. When α increases from 0 to 1, it indicates the willingness to sacrifice the mean profit to avoid risk of its variance.

Due to the random arrival users, the profit is also uncertain and is dependent on the distribution of the demand. So, in many cases, the optimal profit can not been obtained as desired. Based on this, we define the risk as the deviation from the optimal profit in this paper.

The random profit function and the mean profit function of the CN are given by Eq. (1) and Eq. (2) presented in Section 3, respectively. By using the method of integral by parts, Eq. (2) becomes

$$\Pi(b, D) = -(r + q)\int_0^b F(x)dx - q\int_0^{+\infty} xf(x)dx$$
$$+(r + q - c)b. \tag{11}$$

The variance profit function can be obtained by the formula given as follows:

$$Var[\pi(b, D)] = E[(\pi(b, D))^2] - (\Pi(b, D))^2. \tag{12}$$

By using the mean-variance approach to investigate the risk of profit shortfall, the objective function, which is denoted by Φ^*, is given as follows:

$$\Phi^* = \max_{b>0} \left\{ \Pi(b, D) - \alpha Var[\pi(b, D)] \right\} \tag{13}$$

where α is the risk averseness parameter, $\pi(b, D)$ is the random profit function given by Eq. (1), $\Pi(b, D)$ is the mean profit function given by Eq. (11), and $Var[\pi(b, D)]$ is the variance function given by Eq. (12).

We consider a fully distributed communication network, where the traffic demand offered to the whole CN forms a Poisson process with arrival rate $\lambda > 0$. The interarrival times are exponentially distributed with rate λ. Let X be a random variable representing the time between successive demand arrivals in the Poisson process, then we have the probability distribution function $F_X(x)$ and the probability density function $f_X(x)$ of X as follows:

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0, \end{cases} \tag{14}$$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0. \end{cases} \tag{15}$$

The mean and varianceof the exponential distribution are $1/\lambda$ and $1/\lambda^2$, respectively.

Based on the assumption of the traffic demand, Eq. (11) can be obtained as follows:

$$\Pi(b, D) = -\frac{r + q}{\lambda} e^{-\lambda b} + \frac{r}{\lambda} - cb. \tag{16}$$

By using the definition of expectation and method of integral by parts, we can obtain the first component in Eq. (12) as follows:

$$E\left[(\pi(b, D))^2\right] = e^{-\lambda b} \left(-\frac{2b}{\lambda} r^2 - \frac{2}{\lambda^2} r^2 + \frac{2}{\lambda^2} q^2 + \frac{2qcb}{\lambda} + \frac{2rcb}{\lambda} \right)$$
$$- \frac{2rqb}{\lambda} e^{-2\lambda b} + \frac{2r^2}{\lambda^2} - \frac{2rcb}{\lambda} + c^2 b^2. \tag{17}$$

With the similar method, we can obtain the second component in Eq. (12) as follows:

$$(\Pi(b, D))^2 = e^{-\lambda b} \left(-\frac{2r^2 - 2rq}{\lambda^2} + \frac{2rcb + 2qcb}{\lambda} \right) + \frac{r^2}{\lambda^2}$$
$$- \frac{r^2 + 2rq + q^2}{\lambda^2} e^{-2\lambda b} - \frac{2rcb}{\lambda} + c^2 b^2. \tag{18}$$

Substituting Eqs. (17) and (18) into Eq. (12), we can obtain that

$$Var[\pi(b,D)] = e^{-2\lambda b}\left(-\frac{r^2 + 2rq + q^2}{\lambda^2} - \frac{2qcb}{\lambda}\right) + \frac{r^2}{\lambda^2}$$
$$+ e^{-\lambda b}\left(\frac{2rq}{\lambda^2} + \frac{2q^2}{\lambda^2} - \frac{2br^2}{\lambda}\right). \tag{19}$$

5   Numerical Results

In this section, based on the assumption of traffic demand in a CN presented in Section 4, we give some numerical results to show the impact of the penalty cost on the network performance.

According to the engineering experience, we choose several different arrival rates to represent the different cases of traffic load in the CN as follows: $\lambda = 0.01, 0.1, 0.5, 0.9$. Where $\lambda = 0.01$ represents the case that the traffic load in the CN is low, $\lambda = 0.1$ and $\lambda = 0.5$ represent the cases that the traffic load in the CN is normal, and $\lambda = 0.9$ represents the case that the traffic load in the CN is heavy.
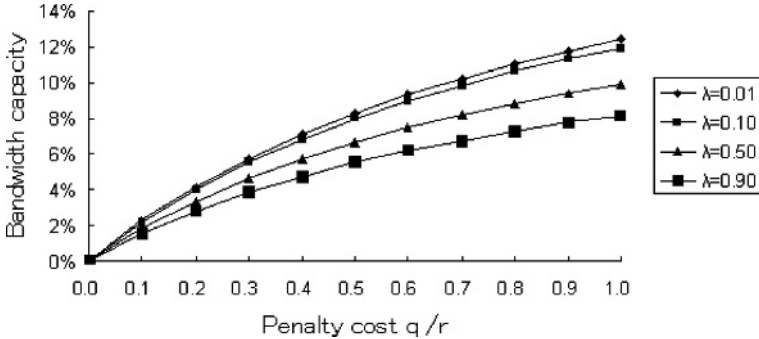

Fig. 1. Impact of penalty cost on bandwidth capacity of the CN

5.1   Impact on the Bandwidth Capacity

In this subsection, we study the impact of the penalty cost on the optimal bandwidth capacity of the CN. Note that the optimal bandwidth capacity without penalty cost presented in [10], is $F^{-1}\left(\frac{r-c}{r}\right)$. However, in this paper the optimal bandwidth capacity with the penalty cost is given by Eq. (9).

Based on the above preparation, we show the numerical results. We choose the unit revenue r as the benchmark of the linear penalty cost q. The horizontal axis (q/r) of Fig. 1 corresponds to the increase of the linear penalty cost. The ordinate axis (b/b^*) of Fig. 1 corresponds to the percentage difference of the optimal bandwidth capacity from the benchmark b^*, where b^* is the optimal
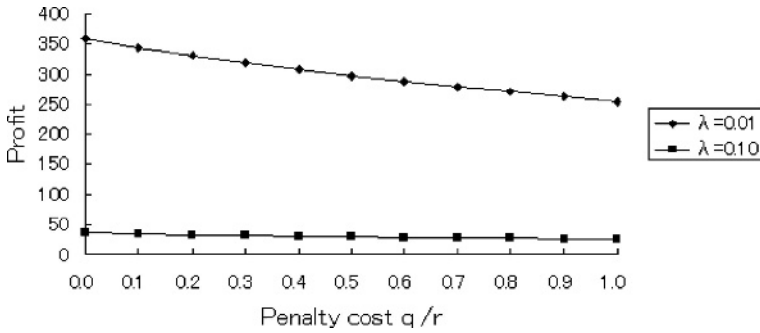
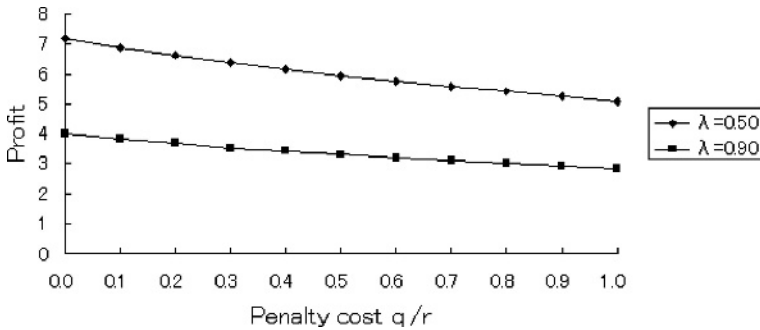**Fig. 2.** Impact of penalty cost on mean profit function of the CN

Fig. 3. Impact of penalty cost on mean profit function of the CN

bandwidth capacity without the penalty cost, and b is the optimal bandwidth capacity with the linear penalty cost. Our numerical results include the optimal bandwidth capacity obtained without penalty cost presented in [10], which is one point in the ordinate axis corresponding to $q/r = 0.0$ in Fig. 1.

For comparing with the model presented in [10], we choose the CN system parameters as follows: the unit revenue $r = 7.5$, the unit cost $c = 1.5$. Let the percentage difference of the penalty cost increase from 0.0 to 1.0 by 0.1 each step with all other parameters unchanged.

From the numerical results shown in Fig. 1, we can conclude that:

(1) In all curves, the impact of the penalty cost on the bandwidth capacity increases as the penalty cost increases.
(2) The curve with a smaller arrival rate has a quicker increasing speed than the curve with a larger arrival rate.
(3) With the same penalty cost, the heavier the traffic load in the CN is, the less the impact of the penalty cost on the bandwidth capacity will be.

Comparing with the results presented in [10] without linear penalty cost, the numerical results in our paper reveal a distinct impact of linear penalty cost on the network bandwidth capacity. It implies that if we consider the penalty cost,

the CN needs to be allocated more bandwidth capacity to guarantee the network performance.

5.2   Impact on the Mean Profit Function

In this subsection, we study the impact of the penalty cost on the mean profit function. Note that the mean profit function $\Pi(b, D)$ without penalty cost, which is presented in [10], is given as follows:

$$\Pi(b, D) = r \int_0^b x f(x) dx + rb \int_b^{+\infty} f(x) dx - cb.$$

However, in this paper the mean profit function with the penalty cost is given by Eq. (2).

We choose the same system parameters as those given in Subsection 5.1. Let the percentage difference of the penalty cost increase from 0.0 to 1.0 by 0.1 each step with all other parameters unchanged.

We choose the unit revenue r as the benchmark of the penalty cost q. The horizontal axes of Figs. 2 and 3 correspond to the increase of the penalty cost q/r. The ordinate axes of Figs. 2 and 3 correspond to the mean profit $\Pi(b, D)$ presented in Eq. (16). The unit of the ordinate axes of Figs. 2 and 3 corresponds to a unit price of the mean profit. Our numerical results include the mean profit obtained without penalty cost presented in [10], which are the points in the ordinate axes corresponding to $q/r = 0.0$ in Figs. 2 and 3.

From the numerical results shown in Figs. 2 and 3, we can conclude that:

(1) In all curves, the impact of the penalty cost on the mean profit function increases as the penalty cost increases.
(2) The curve with a smaller arrival rate has a quicker decreasing speed than the curve with a larger arrival rate.
(3) With the same penalty cost, the heavier the traffic load in the CN is, the less the mean profit will be.

Comparing with the model without the penalty cost presented in [10], the numerical results in our paper reveal a distinct impact of the penalty cost on the network optimal profit. Moreover, the numerical results with different arrival rates almost have the same increasing speed and impact on the mean profit function.

6   Conclusions

In this paper, we presented a stochastic model for optimizing bandwidth allocation in Communication Networks with the linear penalty cost. The model is proposed for offline traffic engineering optimization taking a centralized view of bandwidth allocation, performance control, and risk of profit shortfall. We have derived the optimal bandwidth allocation capacity with the linear penalty cost. We have analyzed the risk averseness in the CNs in the mean-variance framework. We have given numerical results to compare our model with the previous

model presented in [10] and shown the impact of the linear penalty cost on the network performance. We can conclude that the linear penalty cost has distinct impact on the network performance. The implications presented in this paper have good insights for traffic engineering design and planning.

References

[1] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja and X. Xiao: Overview and Principles of Internet Traffic Engineering, RFC 3272, IETF (2002)

[2] X. Xiao, A. Hannan, B. Bailey and L. M. Ni: Traffic Engineering with MPLS in the Internet. IEEE Network **14** (2000) 28-33

[3] P. Aukia et al.: RATES: A Server for MPLS Traffic Engineering. IEEE Network **14** (2000) 34-41

[4] A. Elwalid, C. Jin, S. Low, and I. Widjaja: Mate: MPLS Adaptive Traffic Engineering. Proc. of IEEE INFOCOM (2001) 1300-1309

[5] D. Mitra and K. G. Ramakrishnan: A Case Study of Multiservice Multipriority Traffic Engineering Design for Data Networks. Proc. of IEEE GLOBECOM (1999) 1077-1083

[6] S. Suri, M. Waldvogel, D. Bauer and P. R. Warkhede: Profile-based Routing and Traffic Engineering. J. Computer Communications **26** (2003) 351-365

[7] D. Mitra and Q. Wang: Stochastic Traffic Engineering, with Applications to Network Revenue Management. Proc. of IEEE INFOCOM (2003)

[8] D. Mitra and Q. Wang: Risk-aware Network Profit Management in A Two-tier Market. Proc. of 18th International Telegraffic Congress (2003)

[9] D. Mitra and Q. Wang: Stochastic Traffic Engineering for Demand Uncertainty and Risk-aware Network Revenue Management. ACM SIGMETRICS Performance Evaluation Review **32** (2004) 1-1

[10] J. Wu, W. Yue and S. Wang: Traffic Engineering Design and Optimization for Multimedia Communication Networks. IEICE Technical Report **104** (2005) 19-24

[11] H. M. Markowitz: Portfolio Selection, Efficient Diversification of Investments. Yale University Press, New Haven (1959)

[12] S. Wang and Y. Xia: Portfolio Selection and Asset Pricing. Springer-Verlag, Berlin (2002)