Optimum QoS Classes in Interworking of Next Generation Networks

Behrouz Khoshnevis and Babak H. Khalaj

Abstract: **In this paper, we consider the problem of optimum selection of quality-of-service (QoS) classes in interworking between the networks in a next-generation-network (NGN) environment. After introducing the delay-cost and loss-cost characteristics, we discuss the time-invariant (TI) and time-variant (TV) scenarios. For the TI case, we show that under nearly lossless transmission condition, each network can make its own optimization regardless of other networks. For the TV case, we present sufficient conditions under which the optimum QoS class of each network can be considered fixed with respect to time without considerable degradation in the optimization target. Therefore, under the conditions presented in this paper, the QoS of a flow in each network can be determined solely by considering the characteristics of that network and this QoS class can be held fixed during the flow period.**

Index Terms: **Delay, interworking, loss, next-generation-network (NGN), quality-of-service (QoS).**

I. INTRODUCTION

The next generation network (NGN) concept has been given diverse definitions in literature [1]–[5]. From our point of view, an NGN environment is a set of not necessarily similar networks or domains which need to interwork to send and receive data packets between the source-destination pairs which are not located in the same network or domain. Means of interworking between heterogenous networks in an NGN structure has been a hot research field in literature, however, it is not our major concern in this paper. We will assume that these networks are capable of forwarding data packets to their adjacent networks on a predetermined flow path. An important problem which arises in this scenario is the selection of optimum quality-of-service (QoS) classes for the flow in such networks. This problem has been partially mentioned in [6]–[8], where the authors introduce a generic framework for interworking purposes. However, to the best of our knowledge, this problem has not been discussed and formulated in a mathematical framework in earlier works.

The optimization target that we will use in this paper is the "net benefit" and is widely used in literature (see [9]–[12] and references therein). The net benefit X is defined as utility function U minus the cost C, $X = U - C$, where the "utility function" demonstrates the user (source-destination) satisfaction with the flow transmission quality and the "cost" is the actual cost (per unit time) that the user is paying for transmission of the flow.

The utility U depends on the delay and loss that the flow has experienced and should be a decreasing function with respect to these quantities. We will use the following expression to model this dependency:

$$U(L, D) = -\alpha_D D - \alpha_L L \tag{1}$$

where $\alpha_D > 0$ and $\alpha_L > 0$ demonstrate the sensitivity of the utility function to delay and loss. This linear dependence simplifies the analysis and has been verified, for some specific cases, by the experimental results of [13], which shows that the quality experienced by the users in an interactive audio application decreases almost linearly in delay and loss of the flow.

The flow cost C, on the other hand, depends of the flow rate and the "QoS class" used for the flow. The higher the rate or the QoS class is (the better service), the more cost will be imposed on the flow. For a flow with QoS class c and data rate of r, the cost will be modeled as $C = rc$.

Before describing the problem setup, we will need an explicit definition of the term "QoS class," denoted by c. The communication networks usually classify the data flows to some priority groups, which are normally referred to as the QoS classes. The flows are categorized into these classes based on their QoS requirements, e.g., tolerable delay and loss. Normally, there is a direct relation between a QoS class and the price (e.g., $\frac{\$}{\text{bit}}$) that the network imposes on the unit of the traffic, which uses that specific QoS class. By considering the one-to-one relation between the QoS classes and the corresponding prices, we refer to these prices (per unit traffic) as the "QoS classes". The problem, considered in this paper, is to determine which QoS classes should be used for a single flow, as it passes through a heterogenous set of networks, to maximize to total net benefit. We have made two major assumptions to take the first step in solving this problem:

- This problem is basically a discrete optimization problem, since the QoS classes are of discrete nature. However, we will consider a continuous nonnegative variable c as the QoS class, to reach a continuous optimization problem, which will provide closed form expressions and interesting intuitive results. This continuity assumption can also be justified by assuming high QoS granularity in the network, as it is expected from NGNs.
- The QoS constraints of the flow will be relaxed to acquire a more general picture of the problem. We will consider the optimization problem subject to the QoS constraints in our future work.

The problem setup has been depicted in Fig. 1. The source in network 1 wants to send a flow (or a flow fragment) of duration T and rate r to the destination in network n. The networks

Manuscript received May 09, 2007.

B. Khoshnevis was with the Advanced Communication Research Institute, Sharif University of Technology, Tehran, Iran, during this research, email: b_khoshnevis@alum.sharif.edu.

B. H. Khalaj is with the Electrical Engineering faculty of Sharif University of Technology, Tehran, Iran, email: khalaj@sharif.edu.

This work is funded by Iran National Science Foundation (INSF) under grant number 84094/14.

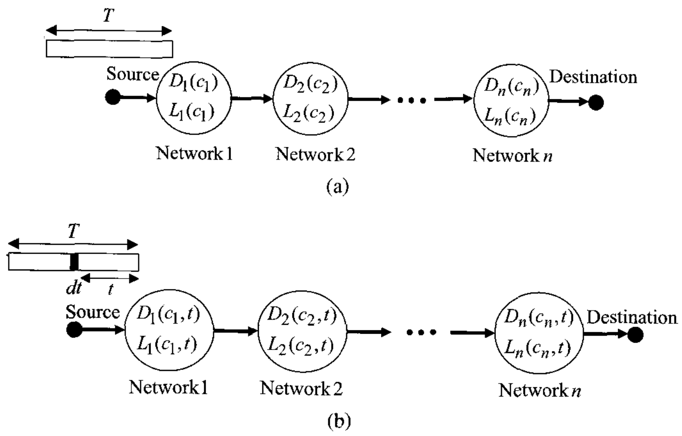


Fig. 1. Series of networks in a NGN environment: (a) Time-invariant characteristics and (b) time-variant characteristics.

on the path have different delay-cost and loss-cost characteristics, which will be defined later. A similar setup of Fig. 1 has also been considered in [12], where the authors allocate different service levels to a flow, as it passes through the routers in a single network, in order to maximize the net benefit. Our problem model is different from [12] in two main ways. 1) The problem in Fig. 1, deals with a higher level problem with multiple heterogenous networks (having different characteristics), while the routers in [12] belong to a single network and have similar behaviors. 2) The utility and cost models in [12] are more simplistic. For example, the dependency of the cost on the flow rate has been ignored.

Getting back to the problem in Fig. 1, the total delay and loss values for the networks in Fig. 1 will be equal to:

$$D = \sum_{i=1}^{n} D_i, \tag{2}$$

$$L = 1 - \prod_{i=1}^{n} (1 - L_i) \tag{3}$$

where $n > 1$ is the number of the networks. D_i and L_i are the delay and loss values imposed on the flow by network i and depend solely on the QoS class $c_i > 0$ chosen for the flow in network i. The $D_i - c_i$ and $L_i - c_i$ dependence will be called delay-cost and loss-cost characteristics of network i.

For the series of networks in Fig. 1, the total cost will be:

$$C = \sum_{i=0}^{n} r_i c_i$$

where, r_i is the flow entrance rate to network i:

$$r_i = r_1 \prod_{j=1}^{i-1} (1 - L_j) \tag{4}$$

and r_1 is initial flow rate. The total net benefit can be formulated as:

$$X = -\alpha_D D - \alpha_L L - \sum_{i=1}^{n} r_i c_i. \tag{5}$$

In the following sections, we will continue the problem in two cases of time-invariant (TI) and time-variant (TV) network characteristics. In Section II, we will present some general assumptions about the network characteristics based on which we will show that the optimization problem has a solution. Subsequently, we will show that under nearly lossless transmission condition, each network can make its own optimization without considering the characteristics of other networks. In Section III, we will assume that the network characteristics change with time. After formulating the optimization problem, we will present sufficient conditions so that fixed sub-optimum QoS class can be applied to the whole flow with controlled performance degradation with respect to the optimum time-dependent QoS class scenario. Finally, Section IV summarizes the results.

II. TI NETWORK CHARACTERISTICS

Throughout this paper, we will assume that the network characteristics of each network $1 \leq i \leq n$ satisfies the following requirements:

The delay function $D_i(c_i)$ and loss function $L_i(c_i)$ are non-increasing, positive, convex with continuous derivative on $\mathbb{R}_+$, and,

$$\begin{cases} \lim_{c_i \to 0} D_i(c_i) = \infty, \\ \lim_{c_i \to \infty} D_i(c_i) = 0, \end{cases} \tag{6}$$

$$\begin{cases} \lim_{c_i \to 0} L_i(c_i) = 1, \\ \lim_{c_i \to \infty} L_i(c_i) = 0, \\ \lim_{c_i \to 0} |\frac{\partial L_i}{\partial c_i}| < \infty. \end{cases} \tag{7}$$

The first assumptions in (6) and (7) are justified by noting that no connection can be established (i.e., infinite delay and 100% loss) with the zero QoS class, as no cost has been charged against the flow. On the other hand, the decreasing property of the characteristics and the asymptotic assumptions for $c \to \infty$ assure that the delay and the loss will decrease and converge to zero as the QoS class improves. The final assumption in (7) and the convexity assumptions are crucial in proving that the optimization problem does have a solution.

Unfortunately, we are not aware of any experimental results for describing the delay-cost and loss-cost characteristics, and therefore, we will apply our analytical and simulation results to some hypothetical examples of such characteristics. A simple example for delay-cost and loss-cost characteristics satisfying the above requirements is:

$$L_i(c_i) = \frac{\beta_i}{c_i + \beta_i}, \quad \beta_i > 0, \tag{8}$$

$$D_i(c_i) = \frac{\gamma_i}{c_i}, \quad \gamma_i > 0 \tag{9}$$

where β_i and γ_i control the expense of a specific level of service. The larger these values are, the higher will be the cost to be paid in order to achieve a fixed level of service. Fig. 2 shows some examples of this characteristics, where we have used **U** as a unit for c.[1]

[1]We will assume 1 $\mathbf{U} = 10^{-9} \frac{\$}{\text{bit}}$ in the simulations.

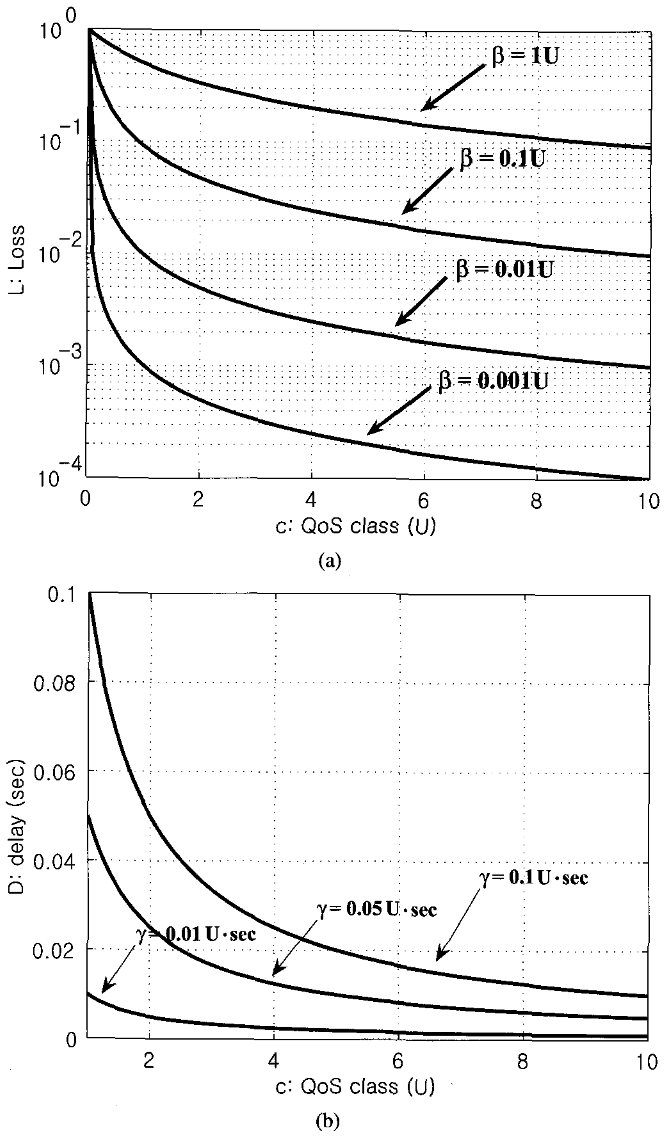


Fig. 2. Examples of network characteristics: (a) Loss-cost characteristics and (b) delay-cost characteristics.

The goal is to find optimum QoS classes c_i^*, $1 \le i \le n$, for which the net benefit is maximized:

$$\underline{c}^* = \arg\max_{\underline{c} \in \mathbb{R}_+^n} \{X(\underline{c})\} \tag{10}$$

where, by $\arg\max_{\underline{c} \in \mathbb{R}_+^n} \{X(\underline{c})\}$ we mean the value of the vector $\underline{c} = [c_1, c_2, \cdots, c_n]$ which maximizes the net benefit X, defined in (5). The optimal values of $\{c_i\}_{i=1}^n$ can be obtained by setting the partial derivatives of X to be equal to zero.

$$\frac{\partial X(\underline{c})}{\partial c_i} = 0, \quad \forall\, 1 \le i \le n \tag{11}$$

which will be of the following form after considering (2)–(5):

$$-\frac{\alpha_L (1-L)}{1-L_i}\frac{\partial L_i}{\partial c_i} - \alpha_D \frac{\partial D_i}{\partial c_i} + \frac{\partial L_i/\partial c_i}{1-L_i} \sum_{j=i+1}^{n} r_j c_j - r_i = 0, \tag{12}$$

for $c_i = c_i^*$ and $1 \le i \le n$.

Based on the requirements imposed on $D_i(c_i)$ and $L_i(c_i)$, we have shown in Appendix I that:

$$\lim_{c_i \to \infty} \frac{\partial D_i}{\partial c_i} = \lim_{c_i \to \infty} \frac{\partial L_i}{\partial c_i} = 0, \tag{13}$$

$$\lim_{c_i \to 0} \frac{\partial D_i}{\partial c_i} = -\infty. \tag{14}$$

Combining these with (12), we will have:

$$\lim_{c_i \to \infty} \frac{\partial X}{\partial c_i} < 0, \tag{15}$$

$$\lim_{c_i \to 0} \frac{\partial X}{\partial c_i} > 0 \tag{16}$$

which by considering the continuity of the derivatives, assures that the system of equations of (11) has a solution in $\mathbb{R}_+^n$.

It is evident from (12) that in order to find optimum QoS class of each network we need to know the status of all networks. However, in the following we will show that under nearly lossless transmission (NLT) condition, each network can choose its own optimum QoS class without the knowledge of the characteristics of other networks.

Assuming the NLT condition, we have $L_i \ll 1$ for $1 \le i \le n$, and subsequently $r_i \approx r_1$ and $L \approx \sum_{i=1}^n L_i$. Therefore, from (5), we will have

$$X(\underline{c}) \approx \sum_{i=1}^{n} X_i(c_i) = \sum_{i=1}^{n} (-\alpha_D D_i - \alpha_L L_i - r_1 c_i). \tag{17}$$

Consequently, the optimum QoS classes can be calculated after solving the following system of equations

$$-\alpha_L \frac{\partial L_i}{\partial c_i} - \alpha_D \frac{\partial D_i}{\partial c_i} - r_1 = 0, \tag{18}$$

for $1 \le i \le n$, which can be easily shown to have a solution in $\mathbb{R}_+$. Therefore, in order to find the optimum QoS class of network i, the knowledge of characteristics of other networks is not required.

To verify this result, we have simulated a scenario with three networks ($n = 3$) and varied the loss-cost characteristic of one of the networks, while keeping all other characteristics fixed. The simulation values have been set to: $r_1 = 10^6 \frac{\text{bit}}{\text{sec}}$, $\alpha_D = 0.01 \frac{\$}{\text{sec}^2}$, $\alpha_L = 0.01 \frac{\$}{\text{sec}}$, $\beta_1 = \beta_2 = 0.1$ U, and $\gamma_1 = \gamma_2 = \gamma_3 = 0.01$ U $\cdot$ sec. Fig. 3 compares the net benefit achieved by the cooperative optimization in (12) and independent optimization in (18), for different values of β_3. As expected from the analysis, for small values of β_3, these two optimizations perform very close. As β_3 increases, the NLT assumption is no longer correct, and we observe considerable degradation in the benefit achieved by independent optimization, compared to the optimal net benefit.

In the next section, we will generalize the problem to the case of time-variant network characteristics.

III. TIME-VARIANT NETWORK CHARACTERISTICS

In this section, we will consider the delay-cost and loss-cost characteristics which change over time. We will use the following assumptions throughout the analysis:

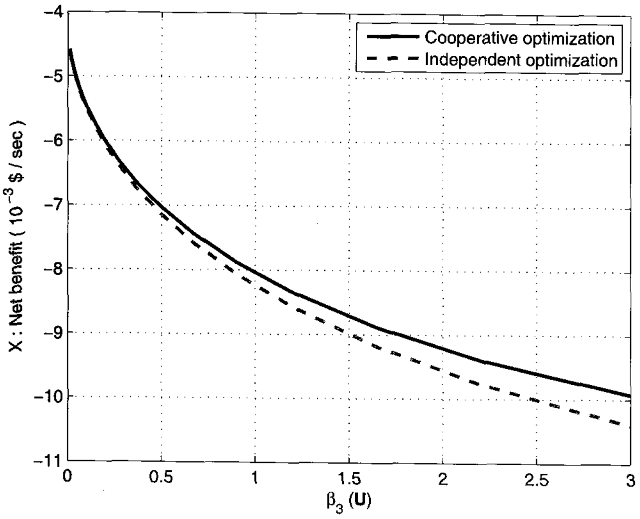


Fig. 3. Comparison of cooperative and independent optimizations for three networks. All the characteristics are fixed, except the loss-cost characteristic of the third network, which is controlled by β_3. For the QoS class unit, we have used: $1\ \mathbf{U} = 10^{-9}$\$/bit.

1) The networks perform under the NLT condition.
2) The characteristics of the networks do not depend on the flow arrival time.

It will be explicitly mentioned, when we use each of these assumptions.

The problem model has been shown in Fig. 1(b). The time-dependent characteristics of the networks will be denoted by $D_i(c_i, t)$ and $L_i(c_i, t)$ for $1 \le i \le n$. We assume that $D_i(c_i, t)$ and $L_i(c_i, t)$ satisfy the requirements described in the beginning of Section II, for all $t > 0$. Consider the $(t, t + dt)$ section of the flow with QoS class equal to $c_1(t)$. This QoS class results in delay $D_1(c_1(t), t)$ and this time-dependent delay will cause different delays for different sections of the flow. Therefore, the flow will experience a time dispersion after passing through the network and will not keep its initial duration and rate. However, if we assume that:

$$\left|\frac{\partial}{\partial t} D(c_i(t), t)\right| < \frac{\epsilon}{n}, \ \forall\, 1 \le i \le n, \tag{19}$$

it can be easily shown that, the difference between the delay values of any two sections of the flow will be less that $\epsilon T/n$ after passing through each of the networks. Therefore, the total difference between the delay values will be less than ϵT which can be ignored when compared with the the flow duration, if $\epsilon \ll 1$. Similarly, the variation in the flow rate will be less than $\epsilon r_1 \ll r_1$. Therefore, if (19) holds, we can assume that the flow keeps its initial duration and rate after passing through the networks.

For every $1 \le i \le n$, the flow will pass through network i toward next network with an average delay of $D_{av,i}$, where:

$$D_{av,i} = \frac{1}{T} \int_{T_i}^{T_i+T} D_i(c_i(t), t)dt \tag{20}$$

where T_i is the flow arrival time at the boundary of network i:

$$\begin{aligned} T_i &= T_{i-1} + D_{av,i-1}, \ \forall\, 1 < i \le n, \\ T_1 &= 0. \end{aligned} \tag{21}$$

It is evident that the average delay and also average loss and cost of network i will depend not only on its own characteristics but also on the total delay imposed by the preceding networks. However, as mentioned in assumption 2 (in the beginning of this section), we will assume that the networks operate under their normal condition, which means that the average behavior of the networks can be assumed to be independent of the flow arrival time. Therefore, the average delay, loss and cost of each network can be expressed as:

$$D_{av,i} = \frac{1}{T} \int_0^T D_i(c_i(t), t)dt, \tag{22}$$

$$L_{av,i} = \frac{1}{T} \int_0^T L_i(c_i(t), t)dt, \tag{23}$$

$$C_{av,i} = \frac{r_i}{T} \int_0^T c_i(t)dt \tag{24}$$

where

$$r_i = r_1 \prod_{j=1}^{i-1} (1 - L_{av,j}). \tag{25}$$

Also the total average delay, loss and cost will be given by:

$$D_{av} = \sum_{i=1}^{n} D_{av,i}, \tag{26}$$

$$L_{av} = 1 - \prod_{i=1}^{n} (1 - L_{av,i}) \approx \sum_{i=1}^{n} L_{av,i}, \tag{27}$$

$$C_{av} = \sum_{i=1}^{n} C_{av,i} \tag{28}$$

where (27) is based on the NLT assumption. Therefore, the average net benefit will be equal to:

$$X_{av} = -\alpha_D D_{av} - \alpha_L L_{av} - C_{av} \tag{29}$$

$$\approx \sum_{i=1}^{n} \frac{1}{T} \int_0^T X_i(c_i(t), t)dt \tag{30}$$

where

$$X_i(c_i(t), t) = -Y_i(c_i(t), t) - r_i c_i(t) \tag{31}$$

$$\approx -Y_i(c_i(t), t) - r_1 c_i(t) \tag{32}$$

where $Y_i(c_i(t), t)$ holds the terms related to the network characteristics:

$$Y_i(c_i(t), t) = \alpha_D D_i(c_i(t), t) + \alpha_L L_i(c_i(t), t). \tag{33}$$

The approximation in (32) results from $r_i \approx r_1$, which is a result of NLT assumption.

Before proceeding further, we need to introduce the notation for partial differentiation. For a bivariate function $f(x_1, x_2)$ and $1 \le i \le 2$:

$$f^{(i)}(x_1, x_2) = \frac{\partial f(x_1, x_2)}{\partial x_i}.$$

In order to maximize the average net benefit in (30), we just need to maximize $X_i(c_i(t), t)$. Therefore, the optimum QoS class $c_i^*(t)$ for each network $1 \leq i \leq n$, will be the solution of the following equation:

$$X_i^{(1)}(c_i^*(t), t) = 0. \tag{34}$$

And the corresponding optimum net benefit will be given by:

$$X_{av}^* \approx \sum_{i=1}^{n} \frac{1}{T} \int_0^T X_i(c_i^*(t), t)dt. \tag{35}$$

It should be noted the networks will be able to choose their optimum QoS classes independently as a result of the NLT condition, however, these QoS classes will depend on time, as it should be evident from (34). Therefore, the network or domain administrator needs to monitor the network characteristics continuously in order to solve the equation in (34). This result not only is subject to practical limitations, but also violates the DiffServ [14] approach which recommends use of a single QoS class for a flow in a network or domain. Therefore, in the following, we will look for sufficient conditions under which, it is possible to use a single QoS class for the flow and perform desirably close to the optimum performance.

Assume that a single QoS of $c_i^*(0)$ is used for the whole flow in network i. Then, the total average net benefit will be given by:

$$X_{av}^{\circ} \approx \sum_{i=1}^{n} \frac{1}{T} \int_0^T X_i(c_i^*(0), t)dt. \tag{36}$$

In theorem 1, we provide sufficient conditions, to ensure that X_{av}° can be close enough to X_{av}^*.

Theorem 1: Assume that NLT condition holds and we have the following for all $1 \leq i \leq n$:

$$n\left|\frac{\partial}{\partial t} D_i(c_i^*(t), t)\right| \ll 1, \tag{37}$$

$$n\left|\frac{\partial}{\partial t} D_i(c_i^*(0), t)\right| \ll 1. \tag{38}$$

Define:

$$Z_i(c, t) = \left|\frac{Y_i^{(2)}(c, t)}{Y_i(c, t) + r_1 c}\right| \tag{39}$$

where $Y_i(c, t)$ is defined in (33). If

$$Z_i(c_i^*(t), t) < \frac{\epsilon_1}{T}, \tag{40}$$

$$Z_i(c_i^*(0), t) < \frac{\epsilon_2}{T}, \tag{41}$$

for all $1 \leq i \leq n$ and $\epsilon_1 \ll 1$, $\epsilon_2 \ll 1$, then we will have:

$$\frac{|X_{av}^* - X_{av}^{\circ}|}{|X_{av}^{\circ}|} < \epsilon_1 + \epsilon_2. \tag{42}$$

The proof of this theorem has been given in Appendix II. The conditions (37) and (38) assure that the flow keeps its shape and rate while passing through the networks, as discussed earlier in this section. The conditions (40) and (41), as will be shown in the following example, will control how fast network characteristics change during the flow time. Finally, the result (42) assures the closeness of X_{av}° and X_{av}^*.

We will apply Theorem 1 to an example to find out how its conditions may be satisfied in a typical scenario. For this example we will use time-dependent version of the examples in (8) and (9):

$$L_i(c_i, t) = \frac{\beta_i(t)}{c_i + \beta_i(t)}, \tag{43}$$

$$D_i(c_i, t) = \frac{\gamma_i(t)}{c_i} \tag{44}$$

where $\beta_i(t)$ and $\gamma_i(t)$ are positive bounded functions of $0 \leq t \leq T$ for all $1 \leq i \leq n$. From (33) we will have:

$$Y_i(c_i, t) = \alpha_D \frac{\gamma_i(t)}{c_i} + \alpha_L \frac{\beta_i(t)}{c_i + \beta_i(t)} \approx \frac{f_i(t)}{c_i} \tag{45}$$

where:

$$f_i(t) = \alpha_D \gamma_i(t) + \alpha_L \beta_i(t), \tag{46}$$

and the approximation in (45) results from NLT assumption, $c_i \gg \beta_i(t)$. From (32) and (34) we will have:

$$c_i^*(t) = \sqrt{f_i(t)/r_1}. \tag{47}$$

Substituting this in (45) and using (39), the condition (40) will be equivalent to:

$$\left|\frac{f_i'(t)}{f_i(t)}\right| < \frac{2\epsilon_1}{T}. \tag{48}$$

The sufficient conditions for (48) can be:

$$\left|\frac{\beta_i'(t)}{\beta_i(t)}\right| < \frac{2\epsilon_1}{T}, \quad \left|\frac{\gamma_i'(t)}{\gamma_i(t)}\right| < \frac{2\epsilon_1}{T}. \tag{49}$$

Doing similar manipulations for $c_i^*(0)$, it can be shown that the condition (40) will be equivalent to:

$$\left|\frac{f_i'(t)}{f_i(t) + f_i(0)}\right| < \frac{\epsilon_2}{T}. \tag{50}$$

Assuming $\epsilon_2 = 2\epsilon_1$ and noting that $f_i(0)$ is a positive quantity, (48) can result (50) and (50) will be redundant. Therefore, the conditions (40) and (41) in Theorem 1 can be satisfied by (48) or its sufficient conditions (49), which as mentioned before, control the characteristics variation speed with time.

For condition (37), we will have:

$$\begin{aligned} n\left|\frac{\partial}{\partial t} D_i(c_i^*(t), t)\right| &= n\left|\frac{\partial}{\partial t}\frac{\gamma_i(t)}{c_i^*(t)}\right| \\ &= n\sqrt{r_1}\left|\frac{2\gamma_i'(t)f_i(t) - \gamma_i(t)f_i'(t)}{2f_i(t)\sqrt{f_i(t)}}\right| \\ &< n\sqrt{r_1}\frac{2|\gamma_i'(t)|f_i(t) + \gamma_i(t)|f_i'(t)|}{2f_i(t)\sqrt{f_i(t)}}. \end{aligned} \tag{51}$$

Incorporating this with (48) and (49), we will have:

$$n\left|\frac{\partial}{\partial t} D_i(c_i^*(t), t)\right| < \frac{3n\sqrt{r_1}\epsilon_1\gamma_i(t)}{T\sqrt{\alpha_D\gamma_i(t) + \alpha_L\beta_i(t)}} \tag{52}$$

$$< \frac{3n\epsilon_1}{T}\sqrt{\frac{r_1\Gamma}{\alpha_D}} \tag{53}$$

where Γ is an upper bound for $\gamma_i(t)$ for all $0 \leq t \leq T$ and $1 \leq i \leq n$. Therefore, in order to satisfy (37) we need:

$$\frac{3n\epsilon_1}{T}\sqrt{\frac{r_1\Gamma}{\alpha_D}} \ll 1. \tag{54}$$

Finally, for condition (38), we will have:

$$n\left|\frac{\partial}{\partial t}D_i(c_i^*(0),t)\right| = n\left|\frac{\partial}{\partial t}\frac{\gamma_i(t)}{c_i^*(0)}\right| = \frac{n\sqrt{r_1}\gamma_i'(t)}{\sqrt{\alpha_D\gamma_i(0)+\alpha_L\beta_i(0)}} < n\sqrt{\frac{r_1}{\alpha_D}}\frac{|\gamma_i'(t)|}{\sqrt{\gamma_i(0)}} < \frac{2n\epsilon_1}{T}\sqrt{\frac{r_1}{\alpha_D}}\frac{\gamma_i(t)}{\sqrt{\gamma_i(0)}}. \tag{55}$$

Applying Lemma 1 presented in Appendix II to (49), it can be shown that, $\gamma_i(t)/\gamma_i(0) < 1 + 2\epsilon_1$. Combining this with (55), we will have:

$$n\left|\frac{\partial}{\partial t}D_i(c_i^*(0),t)\right| < 2n\frac{\epsilon_1}{T}(1+2\epsilon_1)\sqrt{\frac{r_1\Gamma}{\alpha_D}} < \frac{3n\epsilon_1}{T}\sqrt{\frac{r_1\Gamma}{\alpha_D}}$$

where the last inequality has been resulted from $\epsilon_1 \ll 1$. Therefore, (54) will also satisfy condition (41). The following theorem summarizes the presented example.

Theorem 2: Assume that NLT holds and (43) and (44) describe the network characteristics for $1 \leq i \leq n$. If:

$$\max\left\{\left|\frac{\gamma_i'(t)}{\gamma_i(t)}\right|, \left|\frac{\beta_i'(t)}{\beta_i(t)}\right|\right\} < \frac{2\epsilon}{3T}, \tag{56}$$

$$\epsilon \ll \min\left\{\frac{T}{n}\sqrt{\frac{\alpha_D}{r_1\Gamma}}, 3\right\} \tag{57}$$

where Γ is an upper bound for $\gamma_i(t)$ for all $0 \leq t \leq T$ and $1 \leq i \leq n$, we will have:

$$\frac{|X_{av}^* - X_{av}^\circ|}{|X_{av}^\circ|} < \epsilon. \tag{58}$$

According to the lemma presented in Appendix II, the condition (56) will result in:

$$\frac{|\beta_i(t) - \beta_i(0)|}{\beta_i(0)} < \frac{2\epsilon}{3}, \tag{59}$$

and similar expression for $\gamma_i(t)$. Therefore, if ϵ is small enough to satisfy (57), Theorem 2 states that we can operate over $(1-\epsilon)\%$ times the optimal performance provided that the network characteristics change is limited to $\pm(2\epsilon/3)\%$ during the flow duration.

Fig. 4 shows the simulation result for the time-variant scenario. Three networks with time-dependent characteristics have been considered. For ease of simulation, we have assumed that controlling variables β_i and γ_i decrease in a linear fashion during the simulation time $0 \leq t \leq T$:

$$\beta_i(t) = \beta_i(0) - \zeta t,$$
$$\gamma_i(t) = \gamma_i(0) - \eta t$$

where $\zeta > 0$ is such that:

$$\max_{0\leq t\leq T}\left|\frac{\beta_i'(t)}{\beta_i(t)}\right| = \left|\frac{\zeta}{\beta_i(T)}\right| = \frac{2\epsilon}{3T},$$

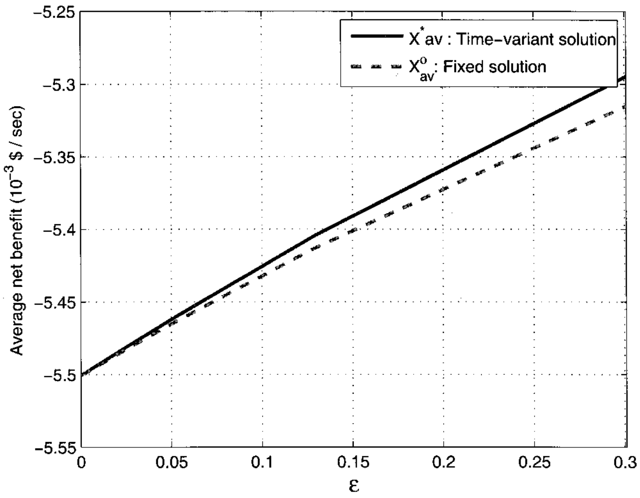


Fig. 4. Comparison of the average net benefits achieved by optimal time-dependent solutions $c_i^*(t)$ in (34) and fixed solutions $c_i^*(0)$.

and similarly for η and γ_i, for a given ϵ value.

The simulation parameters are set to: $r_1 = 10^6 \frac{\text{bit}}{\text{sec}}$, $\alpha_D = 0.01\frac{\$}{\text{sec}^2}$, $\alpha_L = 0.01\frac{\$}{\text{sec}}$, $\beta_1(0) = \beta_2(0) = \beta_3(0) = 0.1$ **U**, and $\gamma_1(0) = \gamma_2(0) = \gamma_3(0) = 0.01$ **U** $\cdot$ sec. Fig. 4 compares the average net benefit resulted from the optimal time-dependent solutions $c_i^*(t)$ in (34) and fixed solutions $c_i^*(0)$. As we expect from Theorem 2, for small enough values of ϵ, the two schemes show very close performance.[2]

Although the result in Theorem 2 is based on specific characteristics of (43) and (44), we believe that it presents a good insight to the problem and if the general mathematical format of the delay-cost and loss-cost characteristics are known, they can be directly applied to Theorem 1 in order to derive similar results of Theorem 2. One possible application of Theorem 2 can be determining the proper flow fragment sizes. The flow fragments should be small enough so that the network characteristics do not make considerable change during each fragment duration and our performance can be kept as close to the optimum performance, as desired.

IV. CONCLUSION

In this paper, we discussed the problem of choosing optimum QoS classes for a single flow in interworking between a set of networks or domains, which we referred to as a NGN environment. First, the scenario of TI network characteristics was considered and we concluded that if NLT condition holds, the networks can choose their optimum QoS class without knowing the status of other networks. Next, the time-variant scenario was discussed and we presented sufficient conditions under which a time-independent QoS class can be chosen in each network while the performance can be kept close enough to optimum case, where the optimum QoS class depends on time and would change during flow duration.

[2]The curves in Fig. 4 are increasing because as we increase ϵ, the corresponding ζ and η increase and the network characteristics improve, i.e., the services get less expensive. Therefore, the average net benefit will increase with ϵ.

APPENDIX I

In this section, we intend to show that,

$$\lim_{c_i\to\infty}\frac{\partial D_i}{\partial c_i} = \lim_{c_i\to\infty}\frac{\partial L_i}{\partial c_i} = 0,$$
$$\lim_{c_i\to 0}\frac{\partial D_i}{\partial c_i} = -\infty$$

given that $L_i(c_i)$ and $D_i(c_i)$ satisfy the requirements presented in the beginning of Section II. Due to similarity of the proof for the above equalities, we will only prove $\lim_{c_i\to\infty}\frac{\partial L_i}{\partial c_i}=0$. We will first show that, this limit exists. For this purpose, we will use the following theorem in analysis.

Theorem 3: If $f(x)$ is a non-decreasing (non-increasing) function on $\mathbf{R}$ and is upper (lower) bounded, then $\lim_{x\to+\infty} f(x)$ exists.

According to the properties of $L_i(c_i)$ mentioned in the beginning of Section II, $L_i'(c_i)$ is non-decreasing ($L_i(c_i)$ is convex and $L_i''(c_i)>0$) and upper-bounded ($L_i(c_i)$ is non-increasing and $L_i'(c_i)<0$), therefore, according to the above theorem, $l=\lim_{c_i\to\infty} L_i'(c_i)$ exists and is obviously non-positive. Now assume that $l\neq 0$. According to the definition of limits, for any given $0<\epsilon<-l$, there exists $\mathbf{c}(\epsilon)$, such that:

$$l-\epsilon < L_i'(c_i) < l+\epsilon, \ \forall\, c_i > \mathbf{c}(\epsilon).$$

After integrating on the right side of the inequality, we will have:

$$L_i(c_i) - L_i(\mathbf{c}(\epsilon)) < (1+\epsilon)(\mathbf{c_i} - \mathbf{c}(\epsilon)).$$

By tending c_i to positive infinity and by noting that $\lim_{c_i\to+\infty} L_i(c_i)=0$ and $l+\epsilon<0$, the left side of the inequality will tend to a negative number while the right side will tend to negative infinity, which is a contradiction. Therefore, $l=\lim_{c_i\to\infty} L_i'(c_i)=0$.

APPENDIX II

In this section, we will provide the proof for Theorem 1, presented in Section III of the paper.

Proof: Conditions (37) and (38) are basic assumptions we have made in order to assure that the flow keeps its duration and rate while passing through networks.

Starting with (40):

$$\frac{\partial}{\partial t}X_i(c_i^*(t),t) = X_i^{(1)}(c_i^*(t),t)\frac{\partial c_i^*(t)}{\partial t} + X_i^{(2)}(c_i^*(t),t). \quad (60)$$

From the definition of $X_i(c,t)=-Y_i(c,t)-r_1c$ in (32), we will have $X_i^{(2)}(c,t) = -Y_i^{(2)}(c,t)$. Combining this with (60) and (34), we will have, $\partial/\partial t X_i(c_i^*(t),t) = -Y_i^{(2)}(c_i^*(t),t)$. Similarly, it can be shown that, $\partial/\partial t X_i(c_i^*(0),t) = -Y_i^{(2)}(c_i^*(0),t)$. Therefore, conditions (40) and (41) in Theorem 2 will be equivalent to:

$$\left|\frac{\frac{\partial}{\partial t}X_i(c_i^*(t),t)}{X_i(c_i^*(t),t)}\right| < \frac{\epsilon_1}{T}, \quad (61)$$

$$\left|\frac{\frac{\partial}{\partial t}X_i(c_i^*(0),t)}{X_i(c_i^*(0),t)}\right| < \frac{\epsilon_2}{T}. \quad (62)$$

To continue the proof, we will need the following lemma which has also been referred to in Section III.

Lemma 1: For any real function $f(t)$, if, $|f'(t)/f(t)|<\epsilon/T$ for all $0<t<T$, and $\epsilon\ll 1$, Then,

$$\frac{||f(t)|-|f(0)||}{|f(0)|} < \frac{\epsilon t}{T}.$$

By applying Lemma 1 to (61), we will have

$$(1-\frac{\epsilon_1 t}{T})\,|X_i(c_i^*(0),0)| < |X_i(c_i^*(t),t)| < (1+\frac{\epsilon_1 t}{T})\,|X_i(c_i^*(0),0)|. \quad (63)$$

Also, for (62), we will have

$$(1-\frac{\epsilon_2 t}{T})\,|X_i(c_i^*(0),0)| < |X_i(c_i^*(0),t)| < (1+\frac{\epsilon_2 t}{T})\,|X_i(c_i^*(0),0)|$$

which yields

$$\frac{1}{1+\frac{\epsilon_2 t}{T}}\,|X_i(c_i^*(0),t)| < |X_i(c_i^*(0),0)| < \frac{1}{1-\frac{\epsilon_2 t}{T}}\,|X_i(c_i^*(0),t)|. \quad (64)$$

Combining (63) and (64) and noting that

$$\frac{1+\frac{\epsilon_1 t}{T}}{1-\frac{\epsilon_2 t}{T}} \approx 1+(\epsilon_1+\epsilon_2)\frac{t}{T},$$
$$\frac{1-\frac{\epsilon_1 t}{T}}{1+\frac{\epsilon_2 t}{T}} \approx 1-(\epsilon_1+\epsilon_2)\frac{t}{T}$$

and also by noting the $X_i(c,t)$ is generally a negative quantity, we will have

$$|X_i(c_i^*(t),t) - X_i(c_i^*(0),t)| < -(\epsilon_1+\epsilon_2)\frac{t}{T}X_i(c_i^*(0),t) < -(\epsilon_1+\epsilon_2)X_i(c_i^*(0),t).$$

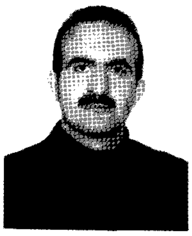
After taking integral of both sides and by summation from $i=1$ to n and noting the definitions in (35) and (36), the proof will be complete. □

REFERENCES

[1] *General Overview of NGN*, ITU-T Rec. Y.2001, Dec. 2004.
[2] *General Principles and General Reference Model for Next Generation Networks*, ITU-T Rec. Y.2011, Oct. 2004.
[3] *IEEE Commun. Mag.*, vol. 38, no. 10, Oct. 2000.
[4] *IEEE Commun. Mag.*, vol. 43, no. 10, Oct. 2005.
[5] *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, Feb. 2005.
[6] V. Fineberg, "A practical architecture for implementing end-to-end QoS in an IP network," *IEEE Commun. Mag.*, vol. 40, no. 1, pp. 122–130, Jan. 2002.
[7] S. I. Maniatis, E. G. Nikolouzou, and I. S. Venieris, "QoS issues in the converged 3G wireless and wired networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 44–53, Aug. 2002.
[8] S. I. Maniatis, E. G. Nikolouzou, and I. S. Venieris, "End-to-end QoS specification issues in the converged all-IP wired and wireless environment," *IEEE Commun. Mag.*, vol. 42, no. 6, pp. 80–86, June 2004.
[9] X. Wang and H. Schulzrinne, "Pricing network resources for adaptive applications in a differentiated services network," in *Proc. IEEE INFOCOM 2001*, Apr. 2001, pp. 943–952.
[10] P. Marbach, "Priority service and max-min fairness," in *Proc. IEEE INFOCOM 2002*, June 2002, pp. 266–275.
[11] J. Hou, J. Yang, and S. Papavassiliou, "Integration of pricing with call admission control to meet QoS requirements in cellular networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 13, no. 9, pp. 898–910, Sept. 2002.
[12] J. Yang, J. Ye, S. Papavassiliou, and N. Ansari, "A flexible and distributed architecture for adaptive end-to-end QoS provisioning in next-generation

networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 321–333, Feb. 2005.
[13] J. Janssen, D. D. Vleeschauwer, and G. H. Petit, "Delay and distortion bounds for packetized voice calls of traditional PSTN quality," in *Proc. IPTel 2000*, Berlin, Germany, Apr. 2000. GMD Report 95, pp. 105–110.
[14] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated services," RFC 2475, 1998.

Behrouz Khoshnevis received B.Sc. and M.Sc. degrees from Sharif University of Technology, Tehran, Iran, in 2003 and 2005, respectively, both in Electrical Engineering. From 2003 to 2006, he was with the Advanced Communications Research Institute, Sharif University of Technology, Tehran, Iran. He is currently pursuing the Ph.D. degree in Electrical and Computer Engineering at the University of Toronto, Toronto, Ontario, Canada. He is a member of IEEE Communications Society.



Babak H. Khalaj received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, in 1989 and the M.Sc. and Ph.D. degrees from Stanford University, Stanford, CA, in 1992 and 1995, respectively, all in electrical engineering. He joined KLA-Tencor in 1995 as a senior algorithm designer, working on advanced processing techniques for signal estimation. From 1996 to 1999, he was with Advanced Fiber Communications and Ikanos Communications. Since then, he has been a senior consultant in the area of data communications. He was the co-editor of the Special Compatibility Standard Draft for ANSI T1E1 group from 1998 to 1999, and he is the author of two U.S. patents and many papers in the signal processing and digital communications area.