21

Statistical Resource Allocation and Pricing in Broadband Communication Networks

Hongbin Ji and Joseph Y. Hui*
Rutgers, The State University of New Jersey
Department of Electrical and Computer Engineering
Piscataway, New Jersey 08855-0909, USA

Abstract

This paper presents resource management for broadband communication networks from an economic point of view.† To be specific, we consider bandwidth allocation and buffer dimensioning. A pricing scheme is also proposed to be dependent on the amount of allocated bandwidth and buffer.

First, we exploit the utility function to represent satisfaction level of a user who requests a certain type of connection service from the network. Then we address the problem of how to allocate bandwidth and buffer capacity of one network component into virtual paths (VPs). On the one hand, the problem is formulated as a non-cooperative K-person game. Two cases, namely unconstrained and constrained, are considered. On the other hand, the problem may be formulated as a cooperative game. The objective is to find an optimal resource allocation such that the total utility of one network component will be maximized. Alternatively, the objective is to maximize the minimum utility among VPs. It is shown that these two game problems may have the same solution for unconstrained case.

Keywords

Asynchronous Transfer Mode (ATM) networks, utility function, resource management, pricing, Nash equilibrium.

1 INTRODUCTION

Broadband communication networks (e.g. ATM networks) are under development for multimedia applications such as data, audio, image and video transmission. Extended services and multiple quality-of-service (QoS) will be provided. The capability to reserve bandwidth and dynamically set up calls make ATM an ideal basis for the support of multimedia applications as pointed out by McDysan and Spohn (1994).

---

†The author is now at AT&T Bell Laboratories, Room 3H-309, Holmdel, NJ 07733-3030.

L. Mason et al. (eds.), *Broadband Communications*

Bandwidth and buffer are the major resources in broadband networks. When a certain type of connection request arrives, a virtual circuit (VC) is set up through the network. The VC carries the call from its origin to destination on one physical connection. Multiple VCs may be integrated to use a virtual path (VP). How to effectively allocate the resources of a physical ATM trunk among the VPs is crucial for managing QoS. A good resource allocation scheme may avoid congestion and entail efficient utilization of resources.

People usually approach this problem from an engineering point of view. For example, Hui (1988) addressed resource allocation for broadband networks through measuring congestion at different levels, i.e., the cell, burst and call levels. An algorithm was proposed for multilayer bandwidth allocation emulating some functions of virtual circuit setup, fast circuit switching, and cell switching. Dziong, Choquette, Liao and Mason (1990) investigated effective bandwidth allocation in an ATM link. The result was used to construct some simple admission control strategy. A routing scheme was given based on residual effective bandwidth. Resource management in broadband ISDN was considered by Burgin (1990) based on VPs. A cost-benefit analysis was presented to determine the conditions under which capacity should be reserved on VPs and the benefits obtained by dynamically updating this reservation. Eckberg, Luan and Lucantoni (1990) studied bandwidth management for congestion control in broadband packet networks. It was noted that bandwidth management can be thought of as applying a throughput-burstiness filter to packet flows entering the network. They focused on characterizing this filter in teletraffic terms.

Hui, Gursoy, Moayeri and Yates (1991) proposed a layered broadband switching architecture with physical or virtual path configurations. A graph framework was introduced to describe network layers of network design, path configurations, dynamic call routing, burst switching and ATM cell switching. These hierarchical layers of switching are performed at decreasing time scales. A layered notion of equivalent bandwidth for satisfying layered grade of service parameters was introduced for making connections at these time-scales.

Monteiro and Gerla (1994) presented bandwidth allocation in ATM networks at different levels and in different stages. At physical level, ATM topology can be dynamically reconfigured by adding or removing trunks between ATM switches. The bandwidth allocation in this level is made possible by the SONET synchronous transfer mode (STM) infrastructure equipped with digital cross connection systems (DCSs). At ATM level, bandwidth can be allocated to individual VCs and VPs. They found that it was convenient to organize the VPs in a connectionless overlay network. This introduced connectionless server bandwidth allocation. Gun, Kulkarni and Narayanan (1994) looked at the problem of bandwidth allocation and access control in high-speed networks. They formulated the design problem so as to minimize the allocated bandwidth subject to service guarantees and stability conditions for the input and output buffers.

In this paper, we consider resource allocation from an economic point of view. A utility function is defined for each class of connections. The utility depends on allocated resources and the number of connections of the same class. As more bandwidth and buffer are allocated, QoS will be improved. Thus users will be more satisfied. When QoS is upgraded to a certain level, users will become indifferent to any further QoS improvement. Thus, we assume that utility is an increasing concave function of allocated bandwidth and buffer.

MacKie-Mason and Varian (1994) defined utility based on allocated bandwidth and the network congestion level. Ji, Hui and Karasan (1996) developed an economic model in a more general way, i.e. users' utility is based on the QoS. Cocchi, Shenker, Estrin and Zhang (1993) specified utility as a linear function of delay and cell loss for different applications such as ftp, mail, telnet, etc. They also studied the role of pricing policies in multiple service class networks. By simulations, they found that it was possible to set the prices so that users of every application type were more satisfied with the combined cost and performance of a network with service-class sensitive prices.

In this paper, we also study pricing scheme. The price is charged to each VP (instead of each connection) according to the allocated resources. This pricing scheme is so simple that it will not cause much extra cost when it is implemented in practical ATM networks. Ji, Hui and Karasan (1996) developed a pricing scheme which charges each connection according to its externalities on the other connections. Pricing a network service in a competitive environment was studied by Liau, Lutton, and Kouatchon (1994). Murphy, Murphy and Posner (1994) presented distributed pricing for embedded ATM networks. A tariff structure, which encourages the cooperative sharing of information between users and the network, was proposed by Kelly (1994) for high speed multiservice networks. Sairamesh, Ferguson and Yemini (1995) discussed the economy and formulated economic based problems for allocating resources.

The rest of this paper is organized as follows. In Section 2, we develop system and economic model for pursuing resource allocation. Some reasonable assumption is also made. Then in Section 3, we address non-cooperative resource allocation. There exists Nash equilibrium under the assumption. An algorithm is proposed to find the non-cooperative optimal resource allocation. In Section 4, we present cooperative resource allocation. The problem is formulated as maximization of total utility in one ATM component or maximization of the minimum utility of the VPs in the ATM component. An algorithm is also proposed to search for cooperative optimal resource allocation. Last section concludes the paper and gives direction for future work.

2 SYSTEM AND ECONOMIC MODELS

Since a broadband network trunk has very large amount of bandwidth and buffer capacity, it is possible to support hundreds of connections simultaneously. Each connection has distinct characteristics and QoS requirements. Without loss of generality, we make the following assumption.

Assumption 1 *The carried connections are generically categorized into K classes according to their characteristic and QoS requirement.*

It is suggested in ATM Forum that traffic is classified into constant bit rate (CBR), variable bit rate (VBR), available bit rate (ABR) or unspecified bit rate (UBR). CBR traffic may include real-time video applications, VBR traffic consists of real-time and non-real-time video and data services, ABR traffic is for data service and UBR for any kind of non-real-time services.

Furthermore, we make another assumption:

Assumption 2 *A physical trunk is correspondingly decomposed into K virtual paths (VPs), and each VP just supports one class of connections exclusively.*

Assumption 2 was also made by Bolla, et al, (1993) and Mishra and Tripathi (1993), Ji (1995), Ji, Hui and Karasan (1996). The major advantage of such decomposition lies in twofold. First, interferences among different classes of traffic are avoided. Thus, QOS management becomes flexible and easy. For example, bursty data traffic would cause significant cell delay variance (CDV) for real-time video traffic if they were multiplexed into the same VP. Second, the benefits of statistical multiplexing are still maintained under the decomposition since the VCs which carry the same class connections are multiplexed into the corresponding VP.

When a call arrives, a virtual circuit (VC) will be established over its correspondingly VP. Each virtual path has dedicated bandwidth and buffer, which are dynamically assigned by the network. It is associated with one type of QoS. So, multiple grades of services are supported through the K virtual paths.

Different VPs might implement distinct queueing disciplines like FCFS, LCFS, etc. The discipline implemented at one VP should be appropriate for the carried class of traffic. And each class of traffic is separated into its corresponding VP.

Let C_i and B_i denote the bandwidth and buffer capacity allocated for the i-th VP, respectively, and N_i denote the number of connections currently established in the i-th VP. These N_i connections are statistically multiplexed into the VP. And each connection has an associated virtual circuit identifier (VCI).

A wide range of applications with various preferences are envisioned to use ATM networks. Users' preferences are represented by utility function, which is defined as a function of allocated resources and the number of connections. Formally,

Definition 1 *User's utility function, denoted by $u_i(C_i, B_i, N_i)$, is a function of the allocated resources, C_i, B_i and the number of connections N_i, for $i = 1, 2, ..., K$.*

We observe that users will be more satisfied when QoS is improved as more resources are allocated. When QoS is upgraded to a certain level, users will become indifferent to any further QoS improvement. According to this observation, we make the following reasonable assumptions.

Assumption 3 *Utility u_i is differentiable and monotonically increasing concave function of C_i, B_i, for $i = 1, 2, ..., K$.*

The above assumption will guarantee the existence of optimal resource allocation. It should be noted that the number of connections may change once a connection is admitted or terminated. The time scale of changing N_i may be a few milliseconds. However, resource allocation should be implemented in a larger time scales, for instance, one second. The reason is that change of resource allocation is costly. It might also cause instability and chaos if resources were re-allocated frequently.

3 NON-COOPERATIVE RESOURCE ALLOCATION

Without loss of generality, we consider one ATM component. The idea generated in this section can be extended into the whole ATM networks.

By collecting historical data, we may obtain statistics on the number of i-th class connections N_i for $i = 1, 2, ..., K$. Let $P_{N_i}(l_i)$ denote the probability when the number of i-th class connections is equal to l_i. The problem is how to allocate the available resources of one ATM component among the K VPs such that the resources will be efficiently utilized.

We know that the total utility of the i-th VP is given by $N_i u_i(C_i, B_i, N_i)$, for $i = 1, 2, ..., K$. The expected value is given by $E[N_i u_i(C_i, B_i, N_i)]$. If u_i is assumed to be a concave function of N_i (Ji, Hui and Karasan, 1996), from Jensen's inequality, we have:

$$E[N_i u_i(C_i, B_i, N_i)] \leq E[N_i] u_i(C_i, B_i, E[N_i]), \ \forall i \tag{1}$$

A simple resource allocation may try to optimize the upper bound in the above equation (Ji, 1995; Ji, Hui and Karasan, 1996). In this paper, our objective is to maximize the expected total utility.

3.1 Unconstrained Resource Allocation

Before multimedia applications become popular, resources in the broadband ATM networks seem relatively plentiful. In this case, the constraints of available resources can be neglected.

Problem Statement

Let p_j and q_j denote the usage cost for unit bandwidth and unit buffer, respectively, for $j = 1, 2, ..., K$. Note that the prices p_j and q_j are specified by network service providers. They may be the same for all j if the providers don't want to distinguish services cost. In general, p_j and $q_j, \forall j$ are different due to distinct services cost of different classes of connections. How to appropriately specify the prices will be addressed in Section 3.1.

The expected total utility of the j-th VP is given by $E[N_j u_j(C_j, B_j, N_j)]$ while the cost charged by network service provider is given by $p_j C_j + q_j B_j$. The revenues for the j-th class of users, denoted by J_j, is given by the difference between expected utility and service charges, i.e. $J_j = E[N_j u_j(C_j, B_j, N_j)] - p_j C_j - q_j B_j$. Each class of users request an optimal amount of resources such that their expected revenues are maximized, i.e.

$$\max_{\{C_j; B_j\}} J_j = E[N_j u_j(C_j, B_j, N_j)] - p_j C_j - q_j B_j \tag{2}$$

for $j = 1, 2, ..., K$. Since $P_{N_j}(l_j)$ represents the probability when the number of j-th class of connections is l_j, straightforwardly we have:

$$J_j = \sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j u_j(C_j, B_j, l_j) - p_j C_j - q_j B_j \tag{3}$$

where L_j denotes the maximum number of connections which can be supported on j-th VP.

From Assumptions 3, we know that J_j is a concave function of C_j, B_j. Therefore, there exists a optimal solution, which is called non-cooperative (or Nash) equilibrium (Nash, 1950). The optimal solution, denoted by C_j^*, B_j^*, for $j = 1, 2, ..., K$, should satisfy:

$$\sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j \left. \frac{\partial u_j(C_j, B_j, l_j)}{\partial C_j} \right|_{C_j^*, B_j^*} - p_j = 0$$
$$\sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j \left. \frac{\partial u_j(C_j, B_j, l_j)}{\partial B_j} \right|_{C_j^*, B_j^*} - q_j = 0 \quad (4)$$

For each VP, we have two unknowns and two equations. Then we may represent the solution by:

$$C_j^* = C_j^*(p_j, q_j)$$
$$B_j^* = B_j^*(p_j, q_j) \quad (5)$$

for $j = 1, 2, ..., K$. It should be noted that there is no conflict among different classes of users. This is because the available resources are assumed to be plentiful (or unlimited). Also note that Nash equilibrium is unique in this case.

Algorithm

We develop an algorithm to find the optimal resource allocation. It should be noted that the algorithm is applicable to each VP.

Algorithm 1 *Find a non-cooperative optimal resource allocation:*
Step 1. Set an appropriate initial resources C_j^0, B_j^0, the desired precision of convergence ϵ, and iteration number $t = 0$.
Step 2. Find C^{t+1}, B^{t+1} such that

$$\sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j \left. \frac{\partial u_j(C_j, B_j, l_j)}{\partial C_j} \right|_{C_j^{t+1}, B_j^t} = p_j \quad (6)$$

and

$$\sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j \left. \frac{\partial u_j(C_j, B_j, l_j)}{\partial B_j} \right|_{C_j^t, B_j^{t+1}} = q_j \quad (7)$$

Step 3. If $|C_j^{t+1} - C_j^t| < \epsilon$ and $|B_j^{t+1} - B_j^t| < \epsilon$, stop. Otherwise, let $t = t + 1$ and go back to step 2.

Pricing Schemes

Roughly speaking, if the prices p_j, q_j are set too small, then the j-th class of users can always increase their revenues by choosing larger amount of resources. It implies that users will choose infinite amount of resources. In other words, resources will be abused if the prices are too low. On the other hand, if the prices are set too high, users can not get any revenues no matter how much resources they may request. In other words, users can not afford service charges. In both of these cases, there will be no solution to the problem of (2). Therefore, prices should be appropriately set such that resources will be efficiently utilized.

We can see that necessary conditions for good price lie on:

$$\lim_{C_i \to \infty} \sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial C_i} < p_i < \lim_{C_i \to 0} \sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial C_i}$$
$$\lim_{B_i \to \infty} \sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial B_i} < q_i < \lim_{B_i \to 0} \sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial B_i} \tag{8}$$

If the prices do not satisfy the left hand inequalities, users will abuse resources. On the other hand, if the prices do not satisfy the right hand inequalities, users will not request any resources because the prices seem too expensive.

Note that users' utility may saturate as resources are expanded continuously. It implies that:

$$\lim_{C_i \to \infty} \sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial C_i} = 0$$
$$\lim_{B_i \to \infty} \sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial B_i} = 0 \tag{9}$$

Then the left hand inequalities of equation (8) become trivial since the prices must be positive.

From network service providers' point of view, their objective is to find optimal prices such that their profits will be maximized, i.e.

$$\max_{\{p_i; q_i\}} \sum_{j=1}^{K} p_j C_j^*(p_j, q_j) + q_j B_j^*(p_j, q_j) \tag{10}$$

Then a necessary condition for optimal pricing is:

$$C_i^* + p_i^* \left.\frac{\partial C_i^*}{\partial p_i}\right|_{p_i^*, q_i^*} + q_i^* \left.\frac{\partial B_i^*}{\partial p_i}\right|_{p_i^*, q_i^*} = 0$$
$$B_i^* + p_i^* \left.\frac{\partial C_i^*}{\partial q_i}\right|_{p_i^*, q_i^*} + q_i^* \left.\frac{\partial B_i^*}{\partial q_i}\right|_{p_i^*, q_i^*} = 0 \tag{11}$$

for $i = 1, 2, ..., K$.

3.2 Constrained Resource Allocation

When multimedia applications become widespread, the resources in broadband ATM networks will become scarce. In this case, we may formulate the resource allocation as an K-person game problem, i.e.

$$\max_{\{C_j; B_j\}} J_j = \sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j u_j(C_j, B_j, l_j) - p_j C_j - q_j B_j, \quad \forall j \tag{12}$$

subject to the constraint of available bandwidth in the ATM component, denoted by C:

$$\sum_{j=1}^{K} C_j \leq C \tag{13}$$

and to the constraint of available buffer in the ATM component, denoted by B:

$$\sum_{j=1}^{K} B_j \leq B \tag{14}$$

If the resources are so plentiful that the following condition is satisfied

$$\begin{aligned} &\sum_{j=1}^{K} C_j^* \leq C \\ &\sum_{j=1}^{K} B_j^* \leq B \end{aligned} \tag{15}$$

where C_j^*, B_j^* are solutions to equation (4), then the constraints (13) and (14) can be relaxed. Thus, the solution of constrained resource allocation is the same as that of unconstrained case.

Suppose the network resources are so scarce that the condition (15) is violated. In this case, the optimal solution, denoted by C_j^c, B_j^c, to problem of (12) becomes:

$$\begin{aligned} C_j^c &= \min[C_j^*, C - \sum_{i=1, i\neq j}^{K} C_i^c] \\ B_j^c &= \min[B_j^*, B - \sum_{i=1, i\neq j}^{K} B_i^c] \end{aligned} \tag{16}$$

for $j = 1, 2, ..., K$. Note that the set $\{C_1^c, C_2^c, ..., C_K^c; B_1^c, B_2^c, ..., B_K^c\}$ is a non-cooperative equilibrium (or Nash equilibrium) (Nash, 1950). It should also be noted that the Nash equilibrium is not unique in this case. In next section, we shall consider resource allocation in cooperative way.

4 COOPERATIVE RESOURCE ALLOCATION

In this section, we consider resource allocation cooperatively. First, we formulate resource allocation as a problem of maximizing total utility of the ATM component. Second, we formulate it as max-min problem which may guarantee fairness among different kinds of applications.

4.1 Total Utility Maximization

Let us formulate the problem of maximizing total utility of one ATM component. Then we give an algorithm to find the solution.

Problem Statement

The problem is formulated as

$$\max_{\{C_i,B_i\}} \sum_{j=1}^{K} \left[\sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j u_j(C_j, B_j, l_j) - p_j C_j - q_j B_j \right] \tag{17}$$

subject to the constraints (13) and (14).

If the resources are plentiful, i.e., inequality (15) is satisfied, then the solution of above maximization problem, which is called cooperative equilibrium, is the same as non-cooperative equilibrium.

If inequality (15) is not met, then the optimal solution to (17) must be at the boundary of the convex set specified by (13) and (14). Using Lagrange multiplier method, we have

$$\begin{aligned} &\sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial C_i} - p_i - \lambda_1 = 0 \\ &\sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial B_i} - q_i - \lambda_2 = 0 \end{aligned} \tag{18}$$

for $i = 1, 2, ..., K$ and where λ_1, λ_2 are Lagrange multipliers. It should be noted that λ_1, λ_2 may represent investment cost of bandwidth and buffer, respectively (Ji, 1995, Ji, Hui and Karasan, 1996).

Algorithm

From equation (18), we have

$$\begin{aligned} &\sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial C_i} - p_i = \sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j \frac{\partial u_j(C_j, B_j, l_j)}{\partial C_j} - p_j \\ &\sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i \frac{\partial u_i(C_i, B_i, l_i)}{\partial B_i} - q_i = \sum_{l_j=1}^{L_i} P_{N_j}(l_j) l_j \frac{\partial u_j(C_j, B_j, l_j)}{\partial B_j} - q_j \end{aligned} \tag{19}$$

for $i, j = 1, 2, ..., K$.

We propose the following algorithm to obtain the cooperative optimal resource allocation.

Algorithm 2 *Searching Cooperative Optimal Resource Allocation*
Step 1. *Set an appropriate initial resource allocation for the first VP,* C_1^0, B_1^0*, a positive initial scaling factors* $\alpha^0 < 1, \beta^0 < 1$*, and the desired precision of convergence* ϵ*. Set iteration index* $t = 0$*.*
Step 2. *Measure the partial derivatives* $\partial u_1/\partial C_1$ *and* $\partial u_1/\partial B_1$ *at the point* (C_1^t, B_1^t, N_1)*. Then calculate* $\partial u_i/\partial C_i, \partial u_i/\partial B_i$ *according to the equation (19) for* $i = 2, 3, ..., K$*. Then obtain* C_i^{t+1}, B_i^{t+1} *for* $i = 2, 3, ..., K$*.*
Step 3. *If* $|\sum_{i=1}^K C_i^{t+1} - C| \leq \epsilon$ *and* $|\sum_{i=1}^K B_i^{t+1} - B| \leq \epsilon$*, then stop. Otherwise,*

Step 3.1. *If* $|\sum_{i=1}^K C_i^{t+1} - C| > \epsilon$*,*
If $\sum_{i=1}^K C^{t+1} - C$ *are positive*
If $\sum_{i=1}^K C^{t+1} - C$ *has the same sign as* $\sum_{i=1}^K C^t - C$*,*
double the scaling factor, i.e. $\alpha^{t+1} = 2\alpha^t$*;*
set $C_1^{t+1} = \alpha^{t+1} C_1^t$*, t=t+1 and go to step 3.2.*
Otherwise, reduce the scaling factoring by 2, i.e. $\alpha^{t+1} = \alpha^t/2$*;*
set $C_1^{t+1} = \alpha^{t+1} C_1^t$*, t=t+1 and go to step 3.2.*
Otherwise,
If $\sum_{i=1}^K C^{t+1} - C$ *has the same sign as* $\sum_{i=1}^K C^t - C$*,*
double the scaling factor, i.e. $\alpha^{t+1} = 2\alpha^t$*;*
set $C_1^{t+1} = \frac{C_1^t}{\alpha^{t+1}}$*, t=t+1 and go to step 3.2.*
Otherwise, reduce the scaling factoring by 2, i.e. $\alpha^{t+1} = \alpha^t/2$*;*
set $C_1^{t+1} = \frac{C_1^t}{\alpha^{t+1}}$*, t=t+1 and go to step 3.2.*
Step 3.2. *If* $|\sum_{i=1}^K B_i^{t+1} - B| > \epsilon$*,*
If $\sum_{i=1}^K B^{t+1} - B$ *are positive*
If $\sum_{i=1}^K B^{t+1} - B$ *has the same sign as* $\sum_{i=1}^K B^t - B$*,*
double the scaling factor, i.e. $\beta^{t+1} = 2\beta^t$*;*
set $B_1^{t+1} = \alpha^{t+1} B_1^t$*, t=t+1 and go to step 2.*
Otherwise, reduce the scaling factoring by 2, i.e. $\beta^{t+1} = \beta^t/2$*;*
set $B_1^{t+1} = \beta^{t+1} B_1^t$*, t=t+1 and go to step 2.*
Otherwise,
If $\sum_{i=1}^K B^{t+1} - B$ *has the same sign as* $\sum_{i=1}^K B^t - B$*,*
double the scaling factor, i.e. $\beta^{t+1} = 2\beta^t$*;*
set $B_1^{t+1} = \frac{B_1^t}{\beta^{t+1}}$*, t=t+1 and go to step 2.*
Otherwise, reduce the scaling factoring by 2, i.e. $\beta^{t+1} = \beta^t/2$*;*
set $B_1^{t+1} = \frac{B_1^t}{\beta^{t+1}}$*, t=t+1 and go to step 2.*

In Step 2, since u_i is a monotonically increasing function of C_i, B_i, we can easily obtain C_i, B_i once the information of the partial derivatives $\partial u_i/\partial C_i, \partial u_i/\partial B_i$ is available.

4.2 Max-Min Optimization

Alternatively, we may formulate the cooperative resource allocation as a max-min optimization problem:

$$\max_{\{C_i;B_i\}} \min_j \sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j u_j(C_j, B_j, l_j) \tag{20}$$

subject to the constraints (13) and (14).

If the resources are plentiful, i.e., inequality (15) is satisfied, then solution of above problem is the same as the unconstrained case.

If the resources are scarce, i.e., inequality (15) is not met, then solution of max-min optimization, denoted by C_i^m, B_i^m, will occur at the boundary of the convex set specified by the constraints (13) and (14). The optimal solution will satisfy:

$$\begin{aligned} C_j^m &= \min[C_j^*, C - \sum_{i=1, i\neq j}^{K} C_i^m] \\ B_j^m &= \min[B_j^*, B - \sum_{i=1, i\neq j}^{K} B_i^m] \end{aligned} \tag{21}$$

for $j = 1, 2, ..., K$. Note that above form is similar to equation (16). However, in this max-min optimization, fairness is achieved for those VPs whose $C_i^m \neq C_i^*, B_i^m \neq B_i^*; C_j^m \neq C_j^*, B_j^m \neq B_j^*$ where $i, j = 1, 2, ..., K$, i.e.

$$\sum_{l_j=1}^{L_j} P_{N_j}(l_j) l_j u_j(C_j^m, B_j^m, l_j) - p_j C_j^m - q_j B_j^m = \sum_{l_i=1}^{L_i} P_{N_i}(l_i) l_i u_i(C_i^m, B_i^m, l_i) - p_i C_i^m - q_i B_i^m \tag{22}$$

It is important to maintain fairness among different classes of multimedia applications.

5 CONCLUSION

In this paper, we study resource allocation in broadband communication networks. Each connection is associated with a utility function which depends on the allocated bandwidth, buffer capacity and the number of same class connections. Then the resource allocation is formulated as a non-cooperative and/or cooperative game problem. Both unconstrained and constrained cases are considered.

A simple pricing scheme is proposed and discussed. Future work may be concerned with pursuit of dynamic resource allocation. Time may be divided into relatively large intervals. Resources are optimally allocated at the start of an interval according to current number of connections. The network service providers may give a price at the start of an interval according to current usage of the resources. This price is fixed during that interval. As time goes on and the number of users is changing, resources will correspondingly be re-allocated. The producer will also change the pricing scheme such that the resources are efficiently utilized.

REFERENCES

Bolla, R., Danovaro, F., Davoli, F. and Marchese, M. (1993) An integrated dynamic resource allocation scheme for ATM networks, *Proceedings of IEEE INFOCOM'93*, San Francisco, CA.

Burgin, J. (1990) Broadband ISDN resource management. *Computer Networks and ISDN Systems*, **20**,323–331.

Cocchi, R. and Shenker, S., Estrin, D. and Zhang, L. (1993) Pricing in computer networks: motivation, formulation, and example. *IEEE/ACM Trans. on Networking*,**1**, no. 6.

Dziong, Z., Choquette, J., Liao, K. Q. and Mason, L. (1990) Admission control and routing in ATM networks. *Computer Networks and ISDN Systems*, **20**, 189–196.

Eckberg, A. E., Luan, D. T. and Lucantoni, D. M. (1990) Bandwidth management: a congestion control strategy for broadband packet networks–characterizing the throughput-burstiness filter. *Computer networks and ISDN Systems*, **20**, 415–423.

Gun, L., Kulkarni, V. G. and Narayanan, A. (1994) Bandwidth allocation and access control in high-speed networks, *Annals of Operations Research*, **49**, 161–183.

Hui, J.Y. (1988) Resource allocation for broadband networks. *IEEE J. Selected Areas in Commun.*, **6**.

Hui, J. Y., Gursoy, M. B., Moayeri, N. and Yates, R. (1991) A layered broadband switching architecture with physical or virtual path configurations. *IEEE Journal on Selected Areas of Communication*, **9**, no. 9.

Ji, H. (1995), Resource allocation, admission control, pricing and routing in multimedia ATM networks. *Proceedings of IEEE Conference on Asia Pacific Communication Conference (APCC)*, Jakarta, Indonesia, Nov. 27-29.

Ji, H., Hui, J. Y. and Karasan, E. GoS-based pricing and resource allocation for multimedia broadband networks. *Proceedings of IEEE INFOCOM'96*, San Francisco, CA, March 24-28.

Kelly, F.P. (1994) On tariffs, policing and admission control for multiservice networks *Operations Research Letters.*, **15**, no. 9.

Liau, B., Lutton, J. L. and Kouatchou, J. (1994) Pricing a network service in a competitive environment. *Proc. of 14 ITC,* Antibes Juan-les-Pins, France, June.

MacKie-Mason, J. K. and H. Varian (1994), Pricing the Internet, *Public Access to the Internet,* Kahin, B. and Keller, J. (Eds.), Prentice-Hall, Englewood Cliffs, New Jersey.

McDysan, D. E. and Spohn, D. L. (1994) ATM: Theory and Application, McGraw-Hill, New York.

Mishra, P. P. and Tripathi, S. K. (1993) Dynamic bandwidth allocation in high speed integrated service networks, *Proceedings of IEEE INFOCOM'93*, San Francisco, CA.

Monteiro, J. A. S. and Gerla, M. (1994) Bandwidth allocation in ATM networks. *Annals of Operations Research*, **49**, 25–50.

Murphy, J., Murphy, L. and Posner, E. C. (1994) Distributed pricing for embedded ATM Networks, *Proc. of 14 ITC*, Antibes Juan-les-Pins, France, June.

Nash, J. (1950) Equilibrium points in n-person games, *Proc. Nat. Acad. Sci. U.S.A.*, **36**.

Sairamesh, J., Ferguson, D. F. and Yemini, Y. (1995) An approach to pricing, optimal allocation and quality of service provisioning in high-speed packet networks, *Proceedings of IEEE INFOCOM'95*, Boston, MA, April.