A DYNAMIC PROGRAMMING MODEL FOR NETWORK SERVICE SCHEDULING

Jesuk Ko
Department of Industrial and Information Engineering, Gwangju University
592-1 Jinwol-dong, Nam-gu, Gwangju 503-703, Korea

Keywords: Video on demand (VOD), Resource allocation, Dynamic programming.

Abstract: Video on Demand (VOD) is one of the most promising services in Broadband Integrated Services Digital Network (B-ISDN) for the next generation. VOD can be classified into two types of services: Near VOD (NVOD) and Interactive VOD (IVOD). For either service, some video servers should be installed at some nodes in the tree structured VOD network, so that each node with a video server stores video programs and distributes stored programs to customers. Given a tree-structured VOD network and the total number of programs being served in the network, the resource allocation problem in a VOD network providing a mixture of IVOD and NVOD services is to determine where to install video servers for IVOD service and both IVOD and NVOD services. In this study we develop an efficient dynamic programming algorithm for solving the problem. We also implement the algorithm based on a service policy assumed in this paper.

1 INTRODUCTION

The emergence of B-ISDN (Broadband-Integrated Service Digital Network) and the advance of several technologies such as ATM (Asynchronous Transfer Mode) technology, image compression / retrieval technology, and multimedia storage / transmission technology make it possible to provide customers with high bandwidth interactive services such as video on demand (VOD), home shopping, video conferencing, etc. VOD seems to be an especially attractive service for the next generation. These VOD services can be classified into two types: interactive VOD (IVOD) and near VOD (NVOD) (Petit et al, 1998; IGI Consulting, 2000).

IVOD is a real-time service that provides a customer with a requested program for the customer to control it. However, IVOD requires expensive and highly developed Video Server (VS)s and storage media to support the real-time service, and incurs a large amount of program storage and transmission costs due to point-to-point connections on demand. Consequently, NVOD service should be utilized from the economical VOD service point of view (Gelman and Halfin, 1999; Sincoskie, 1997).

NVOD distributes periodically some programs on several channels for each program so that customers can begin to watch their requested programs from scratch after waiting an acceptable amount of time. Customers who do not want to wait for the NVOD service can switch the request to IVOD service. NVOD service is not a real-time service and does not depend on customers' requests, but requires relatively cheaper VS and storage media than those for IVOD service. Moreover, NVOD service requires a relatively small amount of program storage and incurs lower transmission costs compared with IVOD service because one channel can be allocated to several customers simultaneously.

In this paper we consider the resource allocation problem in a VOD network providing a mixture of IVOD and NVOD services (RAPINVOD). The RAPINVOD problem is to determine where to install video servers for IVOD service and, by considering customers demands, which programs should be stored at each video server for both IVOD and NVOD services so as to minimize the sum of operating costs. There might be several costs related to the operation of the mixed IVOD and NVOD services, but we just consider three kinds of costs for each service: a program transmission cost, a program storage cost, and a VS installation cost.

To the best of our knowledge, the problem RAPINVOD has yet to be carefully analyzed by researchers. Hodge *et al*. (1994, 1998) and Ishihara *et al*. (1996) have proposed only a service policy for the mixture of IVOD and NVOD services such that some popular programs are distributed through

Ko J. (2006).
A DYNAMIC PROGRAMMING MODEL FOR NETWORK SERVICE SCHEDULING.
In *Proceedings of the Third International Conference on Informatics in Control, Automation and Robotics*, pages 47-53
DOI: 10.5220/0001199000470053

NVOD service since the total cost will be increased if all the programs are distributed only through IVOD service. In particular, Hodge *et al.* (1994) analyzed technologies and costs required for IVOD and NVOD services. Kim *et al.* (1996) proposed a dynamic programming algorithm for the resource allocation problem in a VOD network providing only IVOD service (RAPIVOD).

In this paper we propose a service policy for providing NVOD service and also develop a dynamic programming algorithm for solving RAPINVOD under this policy by extending the dynamic programming algorithm proposed earlier by Kim *et al.* (1996).

This paper is organized as follows. In Section 2, we first describe VOD network architecture and several assumptions and then introduce the concepts of program vision probability and mean service demand. We also define the rate of lost service request for an NVOD program. Section 3 addresses a dynamic programming algorithm for the RAPIVOD problem. In Section 4, we propose an extension of the dynamic programming algorithm, given in Section 3, so as to provide a solution for the RAPINVOD problem. Section 5 concludes the paper.

2 THE TARGET PROBLEM

In this study, we consider two kinds of directed and tree-structured VOD networks, one providing only IVOD service and the other providing a mixture of IVOD and NVOD services. We assume that these networks consist of N interconnected central offices (COs) represented by nodes which are labeled in the Breadth First Search (BFS) order. It is assumed that at most one VS for IVOD service can be installed for each CO and exactly one VS for NVOD service can be installed at the root node of the network. We assume that the program warehouse containing the programs to be provided is located at the root node of the network. The program warehouse provides some programs which are initially stored at the program storage of a video server (VS) in a CO on schedule whenever customers request those programs. We also assume that each customer is connected to exactly one leaf node (CO) in the network by a dedicated link so that the transmission cost from the leaf node to the customer can be ignored. Each CO corresponding to a non-leaf node not only transfers IVOD programs from the CO to the immediately linked COs (i.e., its successors), but also copies NVOD programs distributed from a VS for NVOD service and multi-broadcasts those to its successors.

Let $P[i,1]$ be the set of nodes on the path from node i to the root node 1, i.e., $P[i,1]=\{i, i_1 = PD_i, i_2 = PD_{i_1}, \cdots, 1\}$, where PD_n is the predecessor of node n for each $n = 2,3,\cdots,N$. Then it is assumed that a customer connected to a leaf node i can receive the requested IVOD program from a VS on the path $P[i,1]$. Therefore, all of the IVOD programs requested by customers connected to the leaf node i should be stored at some VS on the path $P[i,1]$. We assume that the unit storage cost for every program is identical for all COs and the link capacity between two consecutive COs is unlimited.

Let J be the total number of programs being provided in the network. It is assumed that all of the programs are sorted in decreasing order of customers' preference and an IVOD program with higher preference is stored at a VS closer to customers in order to reduce the transmission cost. Moreover it is assumed that a more highly preferred NVOD program has a higher priority to be stored at the root node since more customers will be served on each channel for an NVOD program.

NVOD program j is distributed on m_j channels from the root node where we assume that $m_j \geq m_i$ if $i< j$ so that the rate of lost service requests for NVOD programs can be reduced (Ishihara et al, 1996). The service provider then needs to determine the number of channels for each NVOD program.

2.1 Program Vision Probability and Mean Service Demand

We assume in this paper that the demand for each program is determined by customers preference which is sorted in a decreasing order, although it varies with several factors such as service time, service type (IVOD or NVOD service), and customers location, etc. Giovanni *et al.* (1994) defined the *vision probability* of program j as follows:

$$P_j = \frac{P_{j-1}}{D_{HP}},\ j = 2,3,...,J\ ,$$

$$P_1 = \frac{1-\left(1/D_{HP}\right)}{1-\left(1/D_{HP}\right)^J}\ ,\ \sum_{j=1}^{J} P_j = 1 \qquad (1)$$

where D_{HP} is the ratio between the (j-1)-th and j-th program vision probabilities.

Note that $P_1 \geq P_2 \geq \cdots \geq P_J$ and thus $D_{HP} \geq 1$. In this paper we also use equation (1) as the definition of the program vision probability. It is assumed that the same program requested by customers connected to all leaf nodes in the network has the same program vision probability.

We now define the *mean service demand* at node n to be the mean traffic volume occurring during one unit of time of the busiest period. The mean traffic volume is the product of three values: the number of customers connected to the node n, the probability that customers will request the service during the busiest period, and the mean service time. More precisely speaking, let $\vec{T} = (V, E)$ be a directed and tree structured VOD network and $\underline{T}(n) = \{q \in V \mid n \in P[q,1]\}$ be the complete subtree of $\vec{T}$ rooted at node n, where $V = \{1,2,\cdots,N\}$ and $E = \{(PD_i, i) \mid i = 2,3,\cdots,N\}$ are the set of nodes and the set of links (arcs), respectively. For convenience, we denote an arc (PD_i, i) as just arc i since there is point-to-point correspondence between E.

Let L_n be the set of successors of node n(i.e., $L_n = \{q \in V \mid PD_q = n\}$). If n is a leaf node (i.e., $L_n = \varnothing$), then the mean service demand R_n at node n can be determined by the following value: the mean traffic volume at node n ÷ (the unit_service time). For nodes other than leaf nodes (i.e., n such that $L_n \neq \varnothing$), $R_n = \sum_{q \in W} R_q$ where $W = \{q \in T(n) \mid L_q = \varnothing\}$.

2.2 Rate of Lost Service Requests for an NVOD Program

NVOD service distributes programs on several channels periodically. For instance, if a program with the service duration of two hours is distributed on five channels, then the program can be distributed repeatedly through NVOD service per every 0.4 hours, i.e., 24 minutes. As mentioned earlier, customers may feel that the waiting time is too long and cancel the request.

We define the *rate of lost service requests* for an NVOD program to be the probability that a customer who requested the NVOD program cancels the request. An NVOD program with vision probability P_j is distributed on m_j number of channels. Then, if $V(m_j)$ is the time interval between the starts of two consecutive distributions of this NVOD program, i.e., the maximum amount of time that a customer should await the program can be obtained by

$$V(m_j) = \frac{\tau_j}{m_j}, j = 1, 2, 3, \ldots, J \tag{2}$$

where $0 < V(m_j) < \infty$ and $\tau_j > 0$ is the service duration of program.

Now, let the random variable T be the time that a customer waits for the requested NVOD program, with $f(t)$ its probability density function. Then the probability that a customer will wait for more than t hours is

$$P(T > t) = \int_t^\infty f(x)dx \tag{3}$$

Therefore, $P_f(V(m_j))$, the probability that a customer will wait the requested NVOD program with vision probability P_j, can be calculated by

$$P_f(V(m_j)) = \int_0^{V(m_j)} P(T > t)dt \Big/ V(m_j) \tag{4}$$

Consequently, the rate of lost service requests for an NVOD program with vision probability P_j, denoted by $\overline{P}_f(V(m_j))$, is

$$\overline{P}_f(V(m_j)) = 1 - P_f(V(m_j)) \tag{5}$$

For example, if T is exponentially distributed with parameter δ, i.e., if $f(t) = \delta \exp(-\delta t)$ with $0 < t < \infty$ and $\delta > 0$, then

$$P_f(V(m_j)) = \frac{m_j}{\tau_j \cdot \delta}\left(1 - \exp\left(-\frac{\tau_j \cdot \delta}{m_j}\right)\right) \tag{6}$$

Here, the parameter δ is the mean queuing rate that a customer will receive an NVOD service.

Since NVOD service is usually cheaper than IVOD service and each customer also makes a decision to wait or not to await the requested NVOD program, we assume the following: (i) if a program is distributed through NVOD service, then a customer who requested the program wants to receive NVOD service rather than IVOD service, (ii) a proportion of γ of the customers who request NVOD programs and choose not to wait for it request IVOD instead. Consequently, a proportion of $1-\gamma$ of such customers clear their requests and choose neither service.

3 DYNAMIC PROGRAMMING FOR RAPIVOD

This paper considers three kinds of costs for IVOD service: a program transmission cost, a program storage cost, and a VS installation cost. Then the resource allocation problem in a VOD network providing only IVOD service (RAPIVOD) is to decide where we should install VSs, which and how many programs should be stored at each VS, so that all the demands are satisfied with the minimum total cost. We propose a dynamic programming algorithm for solving RAPIVOD in this section.

Although a complete enumeration of all the possible solutions might be used to find the optimal solution for the RAPIVOD problem, this would be very inefficient if no computationally infeasible when the number of COs and programs increase, since the size of the solution space grows exponentially. Moreover, cost functions are non-linear in general and thus it is necessary to find an efficient solution technique for this kind of problem. Let $TC_n(k)$ be the cost of the program transmission on arc $n\left(PD_n, n\right)$ when k programs are stored on $T(n)$. The transmission cost of the remaining (J-k) programs, which will have lower program vision probabilities, on arc n depends upon their mean service demands. Therefore, $TC_n(k)$ can be expressed as follows:

$$TC_n(k) = \begin{cases} g_1\left(C_t, D_n, \sum_{j=k+1}^{J}\left(R_n \times P_j\right)\right), & \text{if } n \neq 1 \\ 0, & \text{otherwise,} \end{cases}$$

where C_t is the unit transmission cost of an IVOD program and D_n is the distance between node n and its predecessor PD_n.

For example, if $L_n \neq \varnothing$ for $n \neq 1$ and $g_1(a,b,c)$ is defined by $(a \times b \times c)^{\phi_t}$ with $\phi_t > 0$, then $TC_n(k)$ is expressed by

$$\left(C_t \times D_n \times \sum_{j=k+1}^{J}\left(R_n \times P_j\right)\right)^{\phi_t},$$

where ϕ_t is the parameter of the transmission cost. The third quantity $\sum_{j=k+1}^{J}\left(R_n \times P_j\right)$ represents the total amount of mean service demands on node n which is equal to the total traffic volume on arc n during the busiest period of time when k kinds of programs are assumed to be stored on $T(n)$.

Let $SC(k, x_n)$ be the cost function of the program storage on node n when x_n kinds of programs out of k ones are stored at node n and the remaining $k - x_n$ kinds of programs are stored on $T(q)$ for all $q \in L_n$ (i.e., $k - x_n$ kinds of programs are stored at some nodes on the path $P[u, q]$ for each leaf node $u \in T(q)$ and all $q \in L_n$). Note that programs associated with the program vision probabilities from the $(k - x_n + 1)$-th through the k-th are stored at node n because of our program storage policy assumed in this paper. Here, we assume that the unit program storage cost is the same for all programs. Let $\lceil x \rceil$ be the smallest integer larger than or equal to x. Then $SC(k, x_n)$ can be expressed as follows:

$$SC(k, x_n) = \begin{cases} g_2\left(C_s, \sum_{j=k-x_n+1}^{k}\left\lceil \frac{R_n \times P_j}{h} \right\rceil\right), & \text{if } x_n \neq 0 \\ 0, & \text{otherwise,} \end{cases}$$

where C_s is the unit storage cost of an IVOD program and h is the number of multiple accesses for an IVOD program.

For example, if $x_n \neq 0$ and $g_2(a,b)$ is defined by $(a \times b)^{\phi_s}$ with $\phi_s > 0$, then $SC(k, x_n)$ is expressed by

$$\left(C_s \times \sum_{j=k-x_n+1}^{k}\left\lceil \frac{R_n \times P_j}{h} \right\rceil\right)^{\phi_s},$$

where ϕ_s is the parameter of the storage cost. The quantity $\left\lceil \frac{R_n \times P_j}{h} \right\rceil$ represents the number of programs with the j-th program vision probability stored at a VS located at node n.

If at least one program is stored at node n (i.e., if $x_n \neq 0$), then a VS should be installed in node n. Let $IC(k, x_n)$ be the cost function of the installation of a VS on node n under the same situation given for $SC(k, x_n)$. Then $IC(k, x_n)$ can be expressed as follows:

$$IC(k, x_n) = g_3(C_v, y(x_n)), \text{ where } y(x_n) = \begin{cases} 1, & \text{if } x_n \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

where C_v is the installation cost of a VS for IVOD service.

For example, if the function $g_3(a,b)$ is defined by $a \times b$, then $IC(k, x_n)$ is expressed by $C_v \times y(x_n)$.

With these three cost functions, we now present an efficient dynamic programming for solving RAPIVOD. For a given node n, we assume that k kinds of programs are stored on $T(n)$ for k = 0, 1, 2, ..., J. Let $f(n, k)$ be the minimum total cost related to storing k kinds of programs on $T(n)$. Suppose that we have found $f(q, k)$ for all $q \in L_n$ and k = 0, 1, 2, ..., J. Then $f(n, k)$ can be determined by the following recursive formula:

$$f(n,k) = \begin{cases} \min\limits_{0 \le x_n \le k}\left\{SC(k,x_n) + IC(k,x_n) + \sum_{q \in L_n} f(q, k-x_n)\right\} + TC_n(k), & \text{if } L_n \neq \varnothing \\ SC(k,k) + IC(k,k) + TC_n(k), & \text{otherwise.} \end{cases}$$

It is important to notice that all the nodes in the network are labeled in BFS order and our dynamic programming incorporates a bottom-up approach which solves the restricted resource allocation problem on $T(n)$ by the node n in the reverse of BFS order.

We now summarize the main idea of our dynamic programming approach. We begin with the leaf node N. If k kinds of programs are stored on $T(N)=\{N\}$, then all of those programs should be stored at node N itself and the (J-k) number of programs with the lower program vision probabilities than the k-th (i.e., programs with the program vision probability from (k+1)-th to the J-th) should be stored at some VSs on the path $P[PD_N,1]$. Therefore, to find the minimum total cost $f(N,k)$ for all k = 0, 1, 2, ..., J, the cost of storing k kinds of programs at node N, the video server installation cost at node N, and the transmission cost of the J-k number of programs on arc N should be evaluated by considering the service demand for each program at node N. Consequently, $f(N,k)$ can be obtained by the sum of those costs, i.e., $SC(k,k)+IC(k,k)+TC_N(k)$, for all k = 0, 1, 2, ..., J. Now, we consider the complete subtree $T(N$-1) of $\bar{T}$ rooted at node N-1. If the node N-1 is a leaf node, then $T(N\text{-}1)=\{N\text{-}1\}$ and thus $f(N\text{-}1,k)$ can be obtained by the same argument for $f(N,k)$, which is equal to $SC(k,k)+IC(k,k)+TC_{N\text{-}1}(k)$, for all k = 0, 1, 2, ..., J. Otherwise (i.e., if $L_{N-1} \neq \varnothing$), $T(N$-1) consists of $\{N\text{-}1, N\}$ where $PD_N = N-1$. Suppose that k kinds of programs are stored on $T(N$-1) and x_{N-1} kinds of programs are stored at node N-1. Then, to find $f(N\text{-}1,k)$, it is enough to evaluate the cost of storing x_{N-1} kinds of programs at node N-1, the video server installation cost at node N-1, and the transmission cost of the (J-k) number of programs on arc N-1 for each $x_{N-1} = 0,1,\cdots,k$, since $k - x_{N-1}$ kinds of programs are stored at node N and we have already found the minimum total cost $f(N,\ k - x_{N-1}\ )$. Therefore, $f(N\text{-}1,k)$ can be obtained by

$$\min_{0 \le x_{N-1} \le k} \{SC(k,x_{N-1}) + IC(k,x_{N-1}) + f(N,k-x_{N-1})\} + TC_{N-1}(k)$$

We continue the above procedure by visiting nodes in the reverse of BFS order until we arrive at the root node 1 of $\bar{T}$, and finally find the optimal value $f(1,J)$ of RAPIVOD.

To find the optimal solution x_n^* for n=1, 2, ..., N, we first define the following value for each n=1, 2, ..., N and k=0, 1, 2, ..., J:

$$x(n,k) = \begin{cases} \operatorname*{argmin}_{0 \le x_n \le k} \left\{ SC(k,x_n) + IC(k,x_n) + \sum_{q \in L_n} f(q,k-x_n) \right\}, & \text{if } L_n \neq \varnothing \\ k, & \text{otherwise} \end{cases}$$

Then the optimal solution can be obtained by $x_n^* = x(n, J - w)$ in the BFS order for all n = 2, 3, ..., N, where $w = \sum_{i \in P[PD_n,1]} x_i^*$ and $x_1^* = x(1,J)$.

Note that the optimal solution holds the information about the video server location and the kinds of programs stored in the video server. In fact, if $x_n^* \neq 0$ for some n, then a video server should be installed at node n. Moreover, programs with the program vision probability from $(J - w - x_n^* + 1)$ -th through $(J - w)$ -th should be stored at node n, since it is assumed that the program with the lower program vision probability is stored at the farther node from the customer. For example, if $V=\{1,2\}$, J=7, $x_1^* = 3$, and $x_2^* = 4$, then programs with 7-th, 6-th, and 5-th program vision probabilities and programs with 4-th, 3-rd, 2-nd, and 1-st program vision probabilities should be stored at node 1 and node 2, respectively.

4 EXTENSION TO THE MIXED SERVICE OF IVOD AND NVOD

In this study we also consider three kinds of costs for for both IVOD and NVOD services: a program transmission cost, a program storage cost, and a VS installation cost. The storage allocation problem in a VOD network providing mixed IVOD and NVOD (RAPINVOD) service is to decide where we should install VSs for IVOD service, and which and how many programs should be stored at each VS for both IVOD and NVOD services, so that all the demands are satisfied with the minimum total cost. Note that a VS for NVOD service is assumed to be installed only at the root node of the given network.

In this section, we propose a dynamic programming for solving RAPINVOD. For mixed IVOD and NVOD service, we first need to determine an efficient rule for determining the number of channels assigned to each NVOD program. Note that we have assumed that impatient customers unwilling to await the NVOD service will receive IVOD service in the ratio of γ , $0 \le \gamma \le 1$. For this case, we might consider several possible rules for determining the number of channels of each NVOD program. Instead, we propose a rule which assigns the number such that the expected number of customers who cancel their NVOD service requests do not exceed L, where L is a fixed number. In fact, let m_j be the number of channels for an NVOD program with the j-th program vision probability. Then m_j is determined by the minimum number of channels satisfying $\overline{P}_f\left(V(m_j)\right)\times\left(R_1\times P_j\right)\le L$, where $\overline{P}_f\left(V(m_j)\right)$ is given in equation (5) and $R_1\times P_j$ represents the expected service demand for an NVOD program with the *j*-th program vision probability, since a VS for NVOD service can be installed only at node 1. Note that the mean service demand for an IVOD program with the *j*-th program vision probability is $\overline{P}_f\left(V(m_j)\right)\times\left(R_1\times P_j\right)\times\gamma$. The procedure for finding m_j can be described as follows.

Procedure Find_m_j
Step 1. (Initialization) $\overline{m}_j \leftarrow 0$;
Step 2. $\overline{m}_j \leftarrow \overline{m}_j + 1$;
Step 3. $\overline{P}_f\left(V(\overline{m}_j)\right)\leftarrow 1 - P_f\left(V(\overline{m}_j)\right)$;
Step 4. If $\overline{P}_f\left(V(\overline{m}_j)\right)\times\left(R_1\times P_j\right) > L$, then go to Step 2.
Step 5. $m_j \leftarrow \overline{m}_j$; Stop

Once we obtain the number of channels, m_j, for all $j = 1,2,\cdots,J$, we are able to decide which and how many programs should be served through IVOD and through NVOD, so as to minimize the sum of the operating costs of IVOD and NVOD services. Before we formulate the problem RAPINVOD, we first introduce three kinds of cost for NVOD service the transmission cost, the storage cost, and the video server installation cost.

Let NTC_s be the transmission cost for *s* kinds of NVOD programs stored at a VS on node 1 to all leaf nodes connected to customers by using m_j number of channels. For convenience, we set $m_0 = 0$. Then, NTC_s can be expressed as follows:

$$NTC_s = \sum_{n=2}^{N}\left(\sum_{j=0}^{s}\left(nct\times D_n\times m_j\right)\right)^{\phi_t'} \qquad (7)$$

where *nct* is the transmission cost for an NVOD program per unit distance and ϕ_t is the parameter for the transmission cost with $\phi_t \ge 0$.

Let NSC_s be the storage cost for *s* kinds of NVOD programs at a VS on node 1. Then NSC_s can be expressed as follows:

$$NSC_s = \left(\sum_{j=0}^{s}\left(ncs\times\left\lceil\frac{m_j}{H}\right\rceil\right)\right)^{\phi_s'} \qquad (8)$$

where *ncs* is the unit storage cost for an NVOD program and ϕ_s' is the parameter for the storage cost with $\phi_s' \ge 0$. The quantity $\left\lceil\frac{m_j}{H}\right\rceil$ in (8) represents the number of programs with the *j*-th program vision probability stored at VS on node 1 for NVOD service.

Let NFC_s be the installation cost of a VS for NVOD service on node 1 when *s* kinds of programs are served by NVOD. Then, since a VS for NVOD service should be installed at node 1 if at least one program is served by NVOD, NFC_s can be expressed as follows:

$$NFC_s = ncv\times y_s, \qquad (9)$$

where *ncv* is the installation cost of a VS for NVOD service and y_s is defined by

$$y_s = \begin{cases} 1, & \text{if } s\ne 0 \\ 0, & \text{otherwise.}\end{cases}$$

With the above three costs related to NVOD service, the problem RAPINVOD can be formulated as follows:

$$\min_{0\le s\le J}\left\{NTC_s + NSC_s + NFC_s + f_s(1,J)\right\} \quad (10)$$

where $f_s(1,J)$ is the minimum total cost for providing IVOD programs with the program vision probabilities rearranged by considering the rate of lost service requests for *s* kinds of NVOD programs and can be obtained by the dynamic programming approach.

We now summarize the main idea of our dynamic programming procedure for solving RAPINVOD. Initially, the number of channels for each NVOD program is obtained by applying the procedure 'Find_m_j'. We first evaluate the IVOD operating cost corresponding to providing only IVOD service by applying the dynamic programming technique and then begin by allocating programs to the VS for NVOD service in the decreasing order of program vision probabilities and finding the total cost, i.e., the sum of the NVOD and IVOD operating costs. In finding the IVOD operating cost, all the programs for IVOD service should be rearranged in decreasing order of program vision probabilities because all the customers who cancel the requested NVOD programs receive IVOD service in the ratio of γ and thus the program vision probabilities of programs for IVOD service that are also allocated for NVOD service should be changed. Once we determine the maximum number, s^*, of kinds of programs which should be stored at the VS for NVOD service, the locations of VSs for IVOD service and the kind and number of IVOD programs stored at each VS can be found simultaneously when $f_{s^*}(1,J)$ is evaluated using the dynamic programming method. We now describe our dynamic programming procedure for solving equation (10) as follows:

Procedure Solve_RAPINVOD
Step 1. (Initialization)
Find_m_j for all $j = 1,2,\cdots,J$;
$TCOST \leftarrow \infty$; $s^* \leftarrow 0$; $s \leftarrow 0$;

Step 2. $NCOST \leftarrow NTC_s + NSC_s + NFC_s$;
Step 3. If $s = 0$, then go to Step 6;
Step 4. $P_s \leftarrow P_s \times \overline{P}(V(m_s)) \times \gamma$;
Step 5. Sort all the programs in descending order of their updated program vision probabilities;
Step 6. Evaluate $f_s(1, J)$;
Step 7. If $NCOST + f_s(1, J) < TCOST$, then $TCOST \leftarrow NCOST + f_s(1, J)$; $s^* \leftarrow s$;
Step 8. If $s < J$, then $s \leftarrow s + 1$ and go to Step 2; Otherwise, stop;

Note that the returned values $TCOST$ and s^* from the procedure 'Solve_RAPINVOD' are the optimal value and the optimal number of kinds of programs for NVOD service, respectively. Moreover, the total number of programs stored at a VS on node 1 for NVOD service is $\sum_{j=0}^{s^*} \left\lceil \frac{m_j}{H} \right\rceil$.

We now use the example to demonstrate our algorithm. All the assumptions for IVOD service are assumed to be the same as those used in Section 2, except that the cost functions of the program transmission and the program storage for IVOD service as well as NVOD service are assumed to be linear(i.e., $\phi_t, \phi_t', \phi_s, \phi_s' = 1$) here. We assume that the program transmission cost, the program storage cost, and the video server installation cost for IVOD service are more expensive than those for NVOD service. We also assume that the unit transmission cost of an NVOD program is more expensive than the unit storage cost of the program. Moreover, it is assumed that all the customers unwilling to await the NVOD service will receive IVOD service (i.e., $\gamma = 1$) and the mean queuing time that a customer will await the requested NVOD program is 20 minutes (i.e., $\delta = \frac{1}{20} = 0.05$). It is also assumed that $L = 710$ and the service duration for all programs is identically equal to 120 minutes (i.e., $\tau_j = 120$ for all $j = 1, 2, ..., J$).

5 CONCLUSIONS

In this paper we have first introduced a dynamic programming algorithm for optimally providing only IVOD service (RAPIVOD). Then we have proposed a procedure for determining the number of channels assigned to each NVOD program under the assumption that the mean number of customers who cancel their requests for NVOD service is given. Finally we have proposed an efficient dynamic programming algorithm for optimally providing a mix of IVOD and NVOD services (RAPINVOD) by extending the key idea of the earlier dynamic programming algorithm for solving RAPIVOD.

It is expected that our algorithms can be applied to several optimization problems which arise in resource allocation problems in networks that provide various types of multimedia services.

REFERENCES

Brubeck, D.W., 1996. Hierarchical storage management in distributed VOD system. *IEEE Multimedia*. 3, 37-47.

Fu, X., Sun, J., and Cai, A., 2000. Scheduling of hierarchical storage in distributed VOD systems. *Journal-China Institute of Communications*. 21, 64-69.

Gelman, A.D. and Halfin, S., 1999. Analysis of resource sharing in information providing services. IEEE GLOBECOM'99, 312-316.

Giovanni, L.De., Langellotti, A.M., Patitucci, L.M., and Petrini, L., 1994. Dimensioning of hierarchical storage for video on demand services. IEEE ICC'94, 1739-1743.

Hessenmuller, H., 1994. The use of large CATV networks as an infrastructure for interactive video services. Australian Telecommunication Networks & Applications Conference, 69-74.

Hodge, W., Mabon, S., and Jone T. Powers, Jr., 1998. Video on demand: architecture, systems, and applications. *SMPTE Journal*, 791-803.

Hodge, W. and Milligan, C., 1994. True video on demand vs. near video on demand, statistical modeling, cost & performance trade-offs. NCTA Technical Paper, 157-172.

IGI Consulting, 2000. Video dialtone & video on demand. VDT 2000, 3.

Ishihara, T., Tanaka, J., Nakajima, I., Okuda, M., and Yamashita, H., 1996. Modeling and evaluation of broadband access networks for multimedia services. IEEE GLOBECOM'96, 117-125.

Jun, S.B. and Lee, W.S., 1998. Video allocation methods in a multi-level server for large-scale VOD services. *IEEE Transactions on Consumer Electronics*. 44, 1309-1318.

Kim, Y.G., Cho, M.R., and Kim, J.Y., 1996. Storage allocation in multi-level VOD network using dynamic programming. *IE Interfaces*. 9(3), 202-213.

Schaffa, F. and Nussbaumer, J-P., 1995. On bandwidth and storage tradeoffs in multimedia distribution networks. IEEE INFOCOM'95, 1020-1026.

Petit, G.H., Deloddere, D., and Verbiest, W., 1998. Bandwidth resource optimization in video-on-demand network architectures. IEEE GLOBECOM'98, 91-97.

Sinkoskie, W.D., 1991. System architecture for a large scale video on demand service. *Computer_Networks and ISDN Systems*. 22, 155-162.

Sincoskie, W.D., 1997. Video on demand: is it feasible?. IEEE GLOBECOM'97, 201-205.