Capacity Assignment and Routing for Interactive Multimedia Service Networks

Byung Ha Lim and June Sung Park

Abstract: **A binary linear integer program is formulated for the problem of expanding the capacity of a fiber optic network and routing the traffic to deliver new interactive multimedia services. A two-phase Lagrangian dual search procedure and a Lagrangian heuristic are developed. Computational results show superior performance of the two-phase subgradient optimization compared with the conventional one-phase approach.**

Index Terms: **Capacity assignment and routing, fiber optic network design, integer program, Lagrangian dual search.**

I. INTRODUCTION

Market demand and vendors supply of interactive multimedia services have started showing a noticeable growth. In 2005, \$11 million came from video on demand (VOD) out of \$211 million in revenue generated by IPTV. Informa Media and Telecoms forecasts that although about 63 million U.S. households participate in VOD services in 2006, 909 million households will have VOD and VOD revenue will reach \$10 billion in 2012. AT&T Interactive TV, Microsoft WebTV, and Time Warner Cable, for example, already deliver two-way interactive multimedia services to homes. TV set-top boxes, PCs and other specialized home devices will be increasingly used for a variety of multimedia communications including Web access, video on demand, distance learning, video telephony, home shopping, home banking, telemedicine, just to name a few.

Telecommunication carriers have been investing in fiber optic networks to increase bandwidth and offer broadband multimedia services to customers at lower operating and repair costs. In the US, most cable television carriers and telephone companies have adopted the hybrid fiber-coax (HFC) system-fiber all the way to a neighborhood and then a single coax cable routed to several hundred homes. Several telephone companies (such as Verizon and Southwestern Bell in U.S. and Hanaro Telecommunication and KT in Korea) have also invested in the more expensive fiber-to-the-curb (FTTC) solution, where the fiber is terminated in an optical network unit that serves about 30 homes.

This research was conducted in collaboration with a telecommunication company (a local exchange carrier) to solve the problem of expanding an existing fiber optic network so that the company introduces new broadband multimedia services to its customers. Known *a priori* are slack capacities of network links in the existing fiber optic network, the location of a level-1 gateway that supplies multimedia contents and the demand for multimedia services at each switching office serving a residential area (which we call a demand node). See Fig. 3 for a depiction of a multimedia service network.

The level-1 gateway node has asynchronous transfer mode (ATM) switches and database systems storing video, audio and data for interactive services. The upstream signal to the gateway typically needs a bandwidth of a few tens of kilobits per second (Kbps), but the downloading of a movie from the gateway, for example, requires a bandwidth of several megabits per second (Mbps). Audio and video streams from the gateway to demand nodes are transmitted using ATM services. All streams from the gateway to each demand node are routed along a single permanent virtual circuit (PVC).

We consider designing multimedia distribution networks with a centralized traffic pattern where all traffic flows from the single gateway to demand nodes. The decisions to be made are: 1) The amount of optical fiber to add to each link of the existing network in order to support the multimedia traffic and 2) a single routing path from the gateway to each demand node.

Extensive research has been conducted with respect to coding, compression, traffic shaping, bandwidth allocation, and congestion control in order to improve the performance of multimedia distribution over an ATM network (see for example Li *et al.* 1996, McManus and Ross (1996), Nussbaumer *et al.* (1995), and Rangan *et al.* (1992)). However, the design of ATM-based multimedia distribution networks, which have a centralized traffic pattern, has received little attention so far. Previous research on ATM network design (such as topology design, link capacity assignment and routing) has mostly focused on the case with distributed traffic patterns where traffic demands can exist between any pair of nodes (see Gerla *et al.* (1989), Medhi (1995), Medhi and Tipper (2000), and Sato and Sato (1991)).

We formulate the problem of designing ATM-based multimedia distribution networks as a binary integer linear program. We solve this problem using a two-phase Lagrangian dual search procedure, which we believe is a new approach to solving an integer program. This two-phase approach alleviates the problem of slow convergence when we solve a disaggregate formulation of the Lagrangian dual from scratch. We solve an aggregate formulation in the first phase to quickly find a dual feasible solution. This becomes the starting solution in phase two in which we conduct a more intensive dual search using a disaggregate formulation. Subgradient optimization is used for the dual search in both phases. We also develop a Lagrangian heuristic to recover primal feasibility from solutions of the Lagrangian relaxation, giving both lower and upper bounds for the primal. Our computation results show superior performance of the two-

Manuscript received May 6, 2008; approved for publication by Jeonghoon Mo, Division III Editor, February 9, 2009.

This Research was supported by the Chung-Ang University Research Grants.

B. H. Lim is with College of Business Administration, Chung-Ang University, Seoul, Korea, email: blim@cau.ac.kr.

J. S. Park is with Industrial Engineering, Computer Science and Knowledge Service Engineering, KAIST, Seoul, Korea, email: june.park@samsung.com.

phase approach compared with the conventional one-phase approach.

The paper is organized as follows. Section II presents the optimization model. Section III constructs two alternative Lagrangian relaxation problems and the two-phase, Lagrangian dual-based solution procedure. Section IV reports computational results, and Section V contains concluding remarks.

II. MODEL FORMULATION

Fiber cables are available at discrete bandwidth levels such as Optical Carrier (OC) 6, 12, 24, 48, 96, 120 and 144. OC-1 signal is transmitted at 51.84 Mbps, and OC-X has X times the OC-1 speed. The multimedia service demand is measured in integral number of OC-1 circuits. The capacity of network links is also measured in number of OC-1 circuits. For engineering and economic reasons, the telephone company had the policy of adding a single fiber cable of certain capacity when it increases the capacity of an existing link. It is not allowed to create a new link; that is, the network topology remains unchanged.

Our problem is thus defined as follows: Given an existing fiber network with a known slack capacity on each link and a known service demand at each node, minimize the total cost of fiber installation and usage by finding the optimal path between the gateway and each demand node and the optimal capacity to add to each existing link, subject to the constraint that the traffic flow may not exceed the capacity of each link.

We use the following notation in presenting the model:

Model Parameters:

V	set of network nodes $\{0, 1, \cdots, m\}$ where 0 denotes the gateway node
$N \subseteq V$	set of nodes with multimedia service demands
E	set of undirected links (j, k), $j, k \in V$, on the network
$E_i \subseteq E$	set of links that are at least on one path from the gateway to node i
$N_{jk} \subseteq N$	set of demand nodes whose paths to the gateway include link (j, k)
T	set of fiber cable types with different capacities; $t \in \{0, 1, \cdots, \|T\|\}$ where 0 denotes no cables, and types $1, 2, \cdots$ denote OC-6, 12, $\cdots$, etc.
e_{jk}	slack capacity on link (j, k) measured in number of OC-1 circuits
r_t	capacity of a fiber cable of type t in number of OC-1 circuits
h_{jkt}	cost of installing a new fiber cable of type t on link (j, k)
c_{jk}	cost of using an OC-1 circuit on link (j, k)
d_i	multimedia service demand at node i in number of OC-1 circuits

Decision Variables:

x_{ijk}	1 if the PVC from node i to the gateway includes link (j, k); 0 otherwise
y_{jkt}	1 if a fiber cable of type t is installed on link (j, k); 0 otherwise.

Model [P]

$$\text{Minimize} \sum_{(j,k)\in E} \sum_{t\in T} h_{jkt} y_{jkt} + \sum_{(j,k)\in E} c_{jk} \sum_{i\in N_{jk}} d_i x_{ijk} \quad (1)$$

$$\text{s.t.} \sum_{t\in T} y_{jkt} = 1 \qquad \forall (j,k) \in E \quad (2)$$

$$\sum_{i\in N_{jk}} d_i x_{ijk} \le e_{jk} + \sum_{t\in T} r_t y_{jkt} \quad \forall (j,k) \in E \quad (3)$$

$$\sum_{j\in V} x_{ij0} - \sum_{j\in V} x_{i0j} = -1 \qquad \forall i \in N \quad (4)$$

$$\sum_{j\in V} x_{iji} - \sum_{j\in V} x_{iij} = 1 \qquad \forall i \in N \quad (5)$$

$$\sum_{j\in V} x_{ijk} - \sum_{j\in V} x_{ikj} = 0 \quad \forall k \in V, \forall i \in N. \quad (6)$$

Objective function (1) represents the total cost of fiber installation and usage. Constraints (2) require that only a single cable of chosen type or no cable is added to a link. Note that if the fiber cable of type 0 is selected, it means not adding any capacity to the link. Constraints (3) state that the total flow of multimedia traffic on a link must not exceed the total available capacity on the link which is the sum of the current slack capacity and the new added capacity. Constraints (4), (5), and (6) require that $\{x_{ijk},\ (j,k) \in E_i\}$ form a single path from the gateway 0 to each demand node i.

III. SOLUTION PROCEDURE

An overview of the solution procedure is shown in Fig.1. We first apply a simple primal heuristic to obtain a feasible solution of problem [P], which is used to calculate the initial step size of the subgradient move. Next, the two-phase subgradient search algorithm is applied. The first phase solves an aggregate formulation of the Lagrangian dual called [ALD]. The incumbent dual solution from the first phase is used as the initial dual solution for the second phase that solves a disaggregate formulation called [DLD]. The Lagrangian heuristic is applied to Lagrangian relaxation solutions generated in the second phase, producing primal feasible solutions. Readers not familiar with the subgradient procedure for solving a Lagrangian dual may refer to Fisher (1981).

A. Aggregate Lagrangian Dual

First, construct a variant of the primal, called [AP], that is equivalent to [P]. Model [AP] introduces new variable f_{jk} that aggregates all the flows of multimedia traffic on link (j, k), i.e., $f_{jk} = \sum_{i\in N_{jk}} d_i x_{ijk}$. [ALD] shown in Fig. 1 is a Lagrangian dual of [AP].

Model [AP]

$$\text{Minimize} \sum_{(j,k)\in E} \sum_{t\in T} h_{jkt} y_{jkt} + \sum_{(j,k)\in E} c_{jk} f_{jk} \quad (7)$$

s.t. $(2), (4), (5), (6)$, and

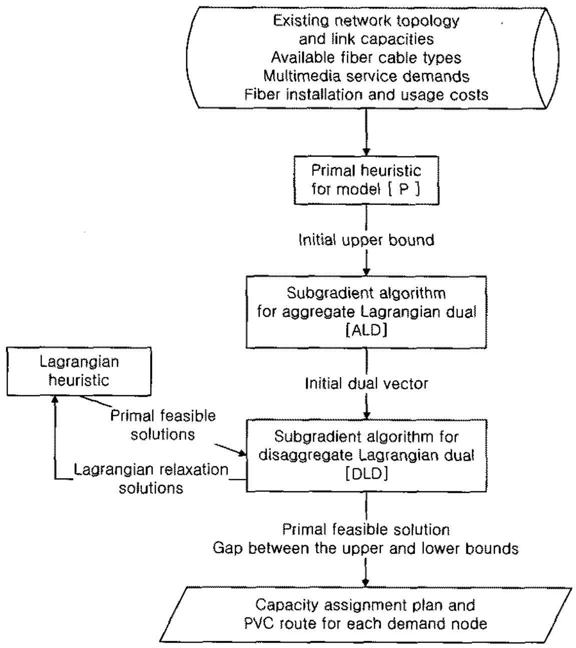

Fig. 1. Solution procedure.

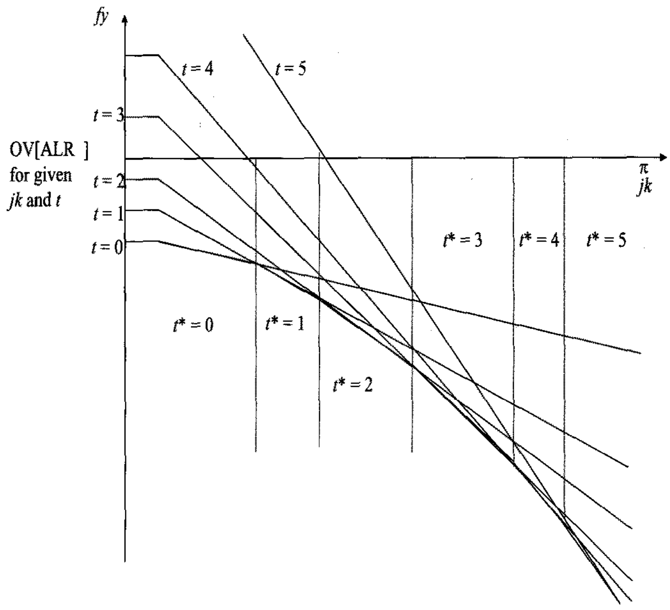

Fig. 2. OV[ALR$^{jk}_{fy}(\pi)|y_{jk\tilde{t}}$=1] curve for different t.

$$f_{jk} \le e_{jk} + \sum_{t\in T} r_t y_{jkt} \qquad \forall (j,k)\in E \quad (8)$$

$$f_{jk} \ge \sum_{i\in N_{jk}} d_i x_{ijk} \qquad \forall (j,k)\in E \quad (9)$$

$$f_{jk} \in \left\{ \sum_{i\in N_{jk}} d_i z_{ijk},\, z_{ijk}\in\{0,1\}\ \forall i\in N \right\} \qquad \forall (j,k)\in E. \quad (10)$$

Constraints (10) are surrogate constraints requiring f_{jk} to be a subset sum of $\{d_i, i\in N\}$. These constraints are added to improve the Lagrangian bound.

An aggregate formulation of Lagrangian relaxation [ALR(π)] is constructed from [AP] by dualizing (9) using multipliers $\pi = [\pi_{jk},(j,k)\in E]\ge 0$.

Model [ALR(π)]

$$\text{Minimize} \sum_{(j,k)\in E}\sum_{t\in T} h_{jkt}y_{jkt} + \sum_{(j,k)\in E}(c_{jk}-\pi_{jk})f_{jk} + \sum_{(j,k)\in E}\pi_{jk}\sum_{i\in N_{jk}} d_i x_{ijk} \quad (11)$$

s.t. (2), (4), (5), (6), (8), and (10).

The introduction of f allows decomposing [ALR(π)] into two independent subproblems-[ALR$_{fy}(\pi)$] involving only f, y and constraints (2), (8), and (10), and [ALR$_x(\pi)$] involving only x and constraints (4), (5), and (6).

[ALR$_{fy}(\pi)$] can in turn be decomposed into $|E|$ independent subproblems [ALR$^{jk}_{fy}(\pi)$] over the set of network links. Each [ALR$^{jk}_{fy}(\pi)$] can be further decomposed over set T of fiber cable types. Each of these subproblems, denoted by [ALR$^{jk}_{fy}(\pi)|y_{jk\tilde{t}}$=1] for given link (j,k) and cable type $\tilde{t}$, is formulated as:

$$\begin{aligned} \text{Minimize}\quad & h_{jk\tilde{t}} + (c_{jk}-\pi_{jk})f_{jk} \\ \text{s.t.}\quad & f_{jk}\le e_{jk}+r_{\tilde{t}} \\ & f_{jk}\in\left\{\sum_{i\in N} d_i z_{ijk}, z_{ijk}\in\{0,1\}, \forall i\in N_{jk}\right\}. \end{aligned}$$

The latter is equivalent to:

$$\begin{aligned} \text{Minimize}\quad & h_{jk\tilde{t}} + (c_{jk}-\pi_{jk})\sum_{i\in N_{jk}} d_i z_{ijk} \\ \text{s.t.}\quad & \sum_{i\in N_{jk}} d_i z_{ijk} \le e_{jk}+r_{\tilde{t}} \\ & z_{ijk}\in\{0,1\} \qquad \forall i\in N_{jk} \end{aligned}$$

which is a $0/1$ knapsack problem where the bang-for-buck ratios are identical for all items.

Optimal value OV[ALR$^{jk}_{fy}(\pi)|y_{jk\tilde{t}}$=1] is $h_{jk\tilde{t}}$ if $c_{jk}\ge\pi_{jk}$, and $h_{jk\tilde{t}}+(c_{jk}-\pi_{jk})f^*_{jk}$ otherwise, where $f^*_{jk} = max\left\{\sum_{i\in N_{jk}} d_i z_{ijk} : \sum_{i\in N_{jk}} d_i z_{ijk} \le e_{jk}+r_{\tilde{t}},\ z_{ijk}\in\{0,1\}\ \forall i\in N_{jk}\right\}$. Quantity f^*_{jk} is uniquely determined given $\tilde{t}$ (i.e., y_{jkt} solutions). It follows that OV[ALR$^{jk}_{fy}(\pi)|y_{jk\tilde{t}}$=1] is a linear function of π_{jk} in the range $[c_{jk},\infty]$ of π_{jk}. Fig. 2 shows, for each value of $\tilde{t}$, the piecewise-linear curve of OV[ALR$^{jk}_{fy}(\pi)|y_{jk\tilde{t}}$=1] over the (j,k)-coordinate of π. Since OV[ALR$^{jk}_{fy}(\pi)$]= $\min_{\tilde{t}\in T}$\{OV[ALR$^{jk}_{fy}(\pi)|y_{jk\tilde{t}}=1$]\}, it is the lower envelop of the curves of OV[ALR$^{jk}_{fy}(\pi)|y_{jk\tilde{t}}$=1] for $\tilde{t} = 0,1,\cdots,|T|$ as shown in thick lines in Fig. 2. We can a priori determine optimal f and y solutions of [ALR$^{jk}_{fy}(\pi)$] for different ranges of π_{jk} by evaluating the lower envelop curve for each link (j,k). We do this before starting the Lagrangian dual search and save the result in a data structure for later lookups. Consequently we do not need to re-solve subproblems [$\mathrm{ALR}^{jk}_{fy}(\pi)$] as π changes in Lagrangian dual search iterations.

Turning to the other subproblem [$\mathrm{ALR}_x(\pi)$], this can be decomposed over the set of demand nodes into $|N|$ independent shortest path problems denoted by [$\mathrm{ALR}^i_x(\pi)$]. Each of them is given by: Minimize $d_i \sum_{(j,k)\in E_i} \pi_{jk}x_{ijk}$ subject to the condition that $\{x_{ijk}, (j,k) \in E_i\}$ form a path from node i to the gateway.

B. Disaggregate Lagrangian Dual

Another variant of the primal, called [DP], that is equivalent to [P] is constructed. Model [DP] deploys copy variables z_{ijk} of x_{ijk}. [DLD] shown in Fig. 1 is a Lagrangian dual of [DP].

Model [DP]

$$\text{Minimize} \sum_{(j,k)\in E}\sum_{t\in T} h_{jkt}y_{jkt} + \sum_{(j,k)\in E} c_{jk} \sum_{i\in N_{jk}} d_i z_{ijk} \quad (12)$$

s.t. $(2), (4), (5), (6)$, and

$$\sum_{i\in N_{jk}} d_i z_{ijk} \le e_{jk} + \sum_{t\in T} r_t y_{jkt} \quad \forall (j,k)\in E \quad (13)$$

$$z_{ijk} \ge x_{ijk} \quad \forall i \in N,\ \forall (j,k)\in E. \quad (14)$$

A disaggregate formulation of Lagrangian relaxation [DLR] is constructed from [DP]. Dualizing constraints (14) using multipliers $\lambda = [\lambda_{ijk},\ i\in N_{jk},\ (j,k)\in E] \ge 0$, we obtain:

Model [DLR(λ)]

$$\text{Minimize} \sum_{(j,k)\in E}\sum_{t\in T} h_{jkt}y_{jkt} + \sum_{(j,k)\in E}\sum_{i\in N_{jk}} (d_i c_{jk} - \lambda_{ijk}) z_{ijk} + \sum_{(j,k)\in E}\sum_{i\in N_{jk}} \lambda_{ijk}x_{ijk} \quad (15)$$

s.t. (2), (4), (5), (6), and (13).

Similar to [ALR], [DLR] is decomposable into $|E| \times |T|$ 0/1 knapsack problems and $|N|$ shortest path problems. The 0/1 knapsack problem [$\mathrm{DLR}^{jk}_{fy}(\lambda)|y_{jk\bar{t}}{=}1$] is given by: *Minimize* $h_{jk\bar{t}} + \sum_{i\in N_{jk}} (d_i c_{jk} - \lambda_{ijk}) z_{ijk}$ s.t. $\sum_{i\in N_{jk}} d_i z_{ijk} \le e_{jk} + r_{\bar{t}}$. The shortest path problem [$\mathrm{DLR}^i_x(\lambda)$] is given by: Minimize $\sum_{(j,k)\in E_i} \lambda_{ijk}x_{ijk}$ subject to the condition that $\{x_{ijk},\ (j,k) \in E_i\}$ form a path from node i to the gateway.

Notice the following differences between [ALR(π)] and [DLR(λ)]:

1) Dimension of the dual space is $|N \times E|$ versus $|E|$, and thus it is more time consuming to solve [DLR(λ)] than [ALR(π)], but the observable dual value is likely to be higher with [DLR(λ)]
2) Bang-for-buck ratios of knapsack problems which are [$\mathrm{DLR}^{jk}_{zy}(\lambda)|y_{jk\bar{t}}{=}1$] are no longer identical for all demands $i \in N_{jk}$, and consequently we have to re-solve subproblems [$\mathrm{DLR}^{jk}_{zy}(\lambda)$] in every iteration of the dual search.

C. Lagrangian Dual Search

The Lagrangian dual [ALD] (resp. [DLD]) is to maximize OV[ALR(π)] s.t. $\pi \in \mathcal{R}^{|E|}_+$ (resp. OV[DLR(λ)] s.t. $\lambda \in \mathcal{R}^{|N\times E|}_+$). The two duals are both solved approximately using subgradient search algorithms. Implementation details that were found particularly effective for our subgradient search are discussed below.

- The dual vector is initialized such that $\pi_{jk} = c_{jk}\ \forall (j,k) \in E$ in the first phase of solving [ALD]. At start of the second phase to solve [DLD], we set $\lambda_{ijk} = d_i\pi^*_{jk}\ \forall i \in N_{jk}, (j,k) \in E$, where π^*_{jk} are the incumbent dual solution obtained from the first phase.
- The scalar value used for step size adjustment is initialized at 2 in the first phase, and 1 in the second phase.
- In both phases, the scalar is halved if the dual value does not improve for 30 consecutive iterations. In this case, the dual vector is reset to the one that has yielded the best dual value up to that point. (This is the same search strategy as used in Park *et al.* (1998).)
- The Lagrangian heuristic is not invoked in the first phase. In the second phase, it is invoked only when the current Lagrangian solution is sufficiently different from the last five Lagrangian solutions to which the heuristic was applied. The difference is measured in number of x_{ijk} and y_{jkt} values that are different in two solutions. If this number is below 20% of $|N|$ comparing with any of the five previous solutions, then the Lagrangian heuristic is skipped.
- In both phases, the subgradient search terminates if either the step size becomes less than 0.00001, or the gap between the upper and lower bounds becomes smaller than 1% of the lower bound.

D. Primal and Lagrangian Heuristics

D.1 *Primal* heuristic module

a) Sort demand nodes in decreasing order of their demand volume.
b) With each demand node i in the sorted list
   i) For every link (j,k) in E, set Cost(j,k) = cost of the demand flow d_i+ cost of adding a fiber cable if necessary to support the demand flow. Set Cost(j,k) to ∞ if even the largest-capacity cable cannot be added to support d_i.
   ii) Find the shortest path from node i to the gateway using Cost(j,k) as the link distance. If the cost of the shortest path is ∞, stop (in which case no feasible solution is found). Otherwise, fix the added fiber capacity and the demand flow on each link along the shortest path. Update the slack capacity in each link.

This primal heuristic takes O($|N| \times |V|^2$). When enough fibers can be added, the primal heuristic easily find a feasible solution. However, this primal heuristic does not guarantee a feasible solution when only a limited number of new fibers can be added. If a feasible solution is not found, proceed to dual search procedure without an incumbent solution. Note that the feasible solution found using the primal heuristic is used only to provide an initial upper bound for the dual search procedure.

The Lagrangian heuristic has two successive modules. *Build* module constructs a primal feasible solution from an optimal

Table 1. Performance of solution procedures.

Problem category	Two-phase dual search (TDS)		Disaggregate dual search (DDS)		Aggregate dual search (ADS)	
	Gap(%)	CPU(sec)	Gap(%)	CPU(sec)	Gap(%)	CPU(sec)
(25,50,7)	0.8	5	0.9	7	17.1	10
(25,50,12)	0.9	24	1.2	52	11.3	15
(25,100,7)	0.9	8	1.2	31	5.8	9
(25,100,12)	0.8	13	1.0	23	9.7	19
Average	0.9	12.5	1.1	28.3	11.0	13.3
(50,100,15)	2.1	436	3.1	919	15.5	76
(50,100,25)	1.0	678	1.3	1673	13.3	106
(50,200,15)	1.3	251	2.5	508	9.2	83
(50,200,25)	1.2	1053	3.3	2321	11.6	124
Average	1.4	604.5	2.6	1355.3	12.4	97.3
(100,200,30)	2.1	Forced to stop after 1.5 hours	4.2	Forced to stop after 1.5 hours	19.7	600
(100,200,50)	13.8	Forced to stop after 1.5 hours	16.9	Forced to stop after 1.5 hours	22.2	830
(100,400,30)	1.7	4173	3.0	Forced to stop after 1.5 hours	13.4	603
(100,400,50)	8.6	Forced to stop after 1.5 hours	23.5	Forced to stop after 1.5 hours	15.1	949
Average	6.6	Forced to stop after 1.5 hours	11.9	Forced to stop after 1.5 hours	17.6	745.5

solution of [DLR(λ)]. Then, *Improve* module, beginning with the feasible solution obtained by *Build*, tries to reduce the total cost while maintaining the feasibility. We provide a pseudocode description of the two modules below.

D.2 *Build* heuristic module

a) Initialize the capacity of each network link based on the y solution of the Lagrangian relaxation.
b) Sort demand nodes in decreasing order of their demand volume.
c) With each demand node i in the sorted list
   i) For every link (j, k) in E, set Cost(j, k) = cost of the demand flow d_i+ cost of adding a fiber cable if necessary to support the demand flow. The cost of demand flow is set to 0 if link (j, k) was used in routing the demand d_i in x-solution of the Lagrangian relaxation, and to $c_{jk}d_i$ otherwise. Set Cost(j, k) to ∞ if even the largest-capacity cable cannot be added to support d_i.
   ii) Same as Step b) of *Primal* heuristic module.

D.3 *Improve* heuristic module

a) Initialize the primal solution to the feasible solution obtained by *Build* module.
b) For each link with an added fiber, if there is an excessive slack capacity on the link, replace the added fiber by a lower-capacity fiber (the one with the lowest, yet sufficient capacity).
c) With each demand node i
   i) Clear the network of the flow of demand d_i, and update slack capacities along the path of d_i. For each link with an added fiber, replace the fiber by a lower-capacity one if possible (as in Step b).
   ii) Same as Step c.i) of *Build* module except that the Lagrangian x solution is not applied to the calculation of Cost(j, k).
   iii) Same as Step b) of *Primal* heuristic module except that a feasible solution is always found here.

IV. COMPUTATIONAL RESULT

The telecommunication company that worked with us prepared problem parameters so that they closely described the real-world situation. Nodes are uniformly distributed in a 100×130 rectangle. 20% of all nodes are designated as hub nodes. The slack capacity e_{jk} is U[30,120] if link (j, k) is between two hub nodes, and U[0,90] otherwise. U[a,b] stands for a random integer uniformly distributed over [a,b]. Demands d_i follow a normal distribution N(50,5). The usage cost c_{jk} and installation cost h_{jkt} are respectively set to 0.1 and 1.2 per unit distance per OC-1. A fiber cable of type t has 6, 12, 24, 48, 96, 120 and 144 OC-1 circuits for $t = 1, \cdots, 7$, respectively.

Based on these parameter distributions, we generate 12 different categories of test problems by varying the following attributes: network size $|V|$ = 25, 50 or 100; network density $|E|/|V|$ (i.e., the average node degree) = 2 or 4; and demand density $|E|/|V|$ = 0.3 or 0.5. We set the largest values of the three attributes rather unrealistically high so as to perform a stress testing with the solution procedures. We randomly generate 5 problem instances in each category, making a total of 60 test problems. Each problem category is labeled by the triple $(|V|, |E|, |N|)$ as listed in Table 1.

We apply three different solution procedures to test problems: 1) Two-phase dual search (Fig. 1), 2) disaggregate dual search solving [DLD], and 3) aggregate dual search solving [ALD]. The subgradient optimization, the *Primal* heuristic and the Lagrangian heuristic discussed in Section III are applied in all three procedures. Only variation is that in both single-phase procedures 2 and 3 we set the limit on non-improving consecutive iterations to 50 instead of 30 to allow more extensive searches. The CPU run time is limited to 1.5 hours for all three procedures. All the algorithms were coded in Fortran and run on an

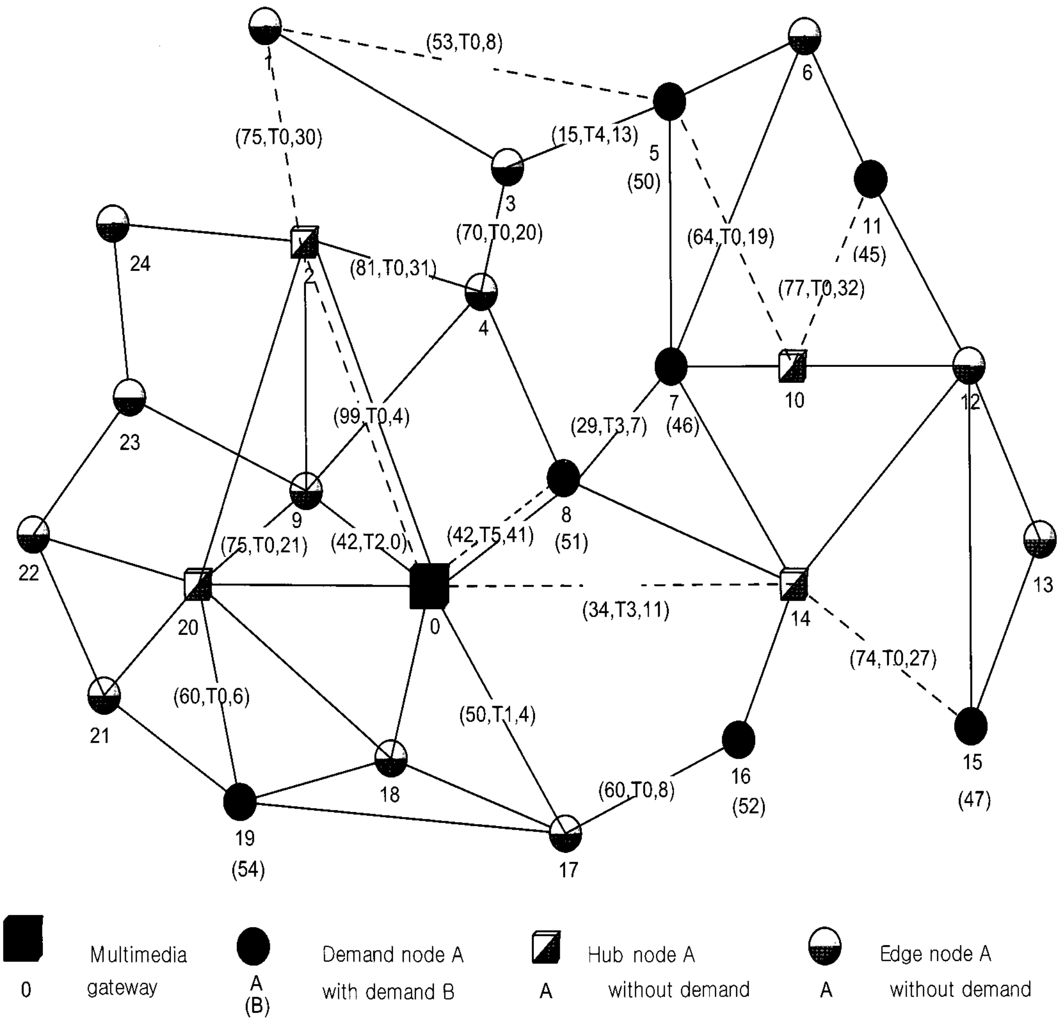

Fig. 3. A feasible capacity assignment and routing for multimedia network: Problem (25,50,7).

IBM 3090 model 400J.

Table 1 reports the gap between the upper and lower bounds on OV[P] as percentage of the lower bound, and the CPU run time, for each problem category and for each solution procedure. The figures in the table are the average across five random problem instances. We summarize the findings from Table 1.

- The disaggregate dual search (DDS) produces a much smaller gap than the aggregate dual search (ADS) for most problem categories, although DDS consumes more CPU time as expected.
- It is interesting, however, to observe that ADS produces a better solution in less time than DDS for the largest problem $(100, 400, 50)$.
- The two-phase dual search (TDS) outperforms the disaggregate dual search (DDS) and the aggregate dual search (ADS) for all problem categories. In 25-nodes problem categories, the average gap of TDS is only 0.85% with 12.5 seconds while DDS's is 1.1% with 28.3 seconds and ADS's 11.0% with 13.3 seconds. In 50-nodes problem categories, TDS consumes more CPU times (604.5 sec.) than ADS (97.3 sec.) and less than DDS (1355.3 sec.) as expected. However its gap is much smaller (1.4%) than ADS's (12.4%) and DDS's (2.55%). In 100-nodes problem categories, even though it consumes more time than ADS and less time than DDS, TDS produces a lot better solution (gap 6.6 %) than ADS (17.6%) and DDS (11.9%).
- The two-phase dual search (TDS) produces a gap of at most 2% for all problem categories except two-$(100, 200, 50)$ and $(100, 400, 50)$. It is unusual to have 100 switching offices in a region served by a single multimedia gateway. It is also quite unusual for a fiber optic network to have an average node degree greater than 2. Therefore, the two problem categories with large gaps are rather unrealistic cases that we generated for the purpose of stress testing.

Fig. 3 shows a feasible multimedia network found for the problem category $(25, 50, 7)$. The PVC path for each demand node is shown in thick lines (some solid and others dotted). The triplet shown on each link indicates the initial slack capacity e_{jk}, the type t of fiber cable added, and the final slack capacity $e_{jk} + r_t - f_{jk}$. For example, the triplet of the link between node 5 and node 3 is $(15, T4, 13)$. It means that the initial slack capacity of the link $(5, 3)$ is 15, a fiber cable of type 4 is added, and the final slack capacity is 13.

The demand node 11 with demand of 45 is routed along a long path $(11, 10, 5, 1, 2, 0)$, but consuming only initial slack capacities. And, the demand node 5 which has demand of 50 uses the path $(5,3,4,2,0)$. In this case, a fiber cable of type 4 (which is 48 OC-1) is added to link $(5,3)$ so that the capacity of link $(5,3)$ will be 63 (initial capacity (15) + new cable added (48)) which is enough to cover the demand node 5 with demand of 50. And then, the slack capacity of link $(5,3)$ is $13(=63-50)$. No cable is added to subsequent link $(3,4)$, $(4,2)$, and $(2,0)$ because the demand of node 5 (which is 50) can be routed using initial slack capacities on links $(3,4)$, $(4,2)$, and $(2,0)$. Link $(0,8)$ needs a type 5 cable and is used to support both demand node 7 with demand of 46 and demand node 8 with demand of 51. These demands (demand node 7 and demand node 8), as well as demand node 15 and demand node 16, are all routed on a minimum-hop path. As such, it seems impossible to determine capacity assignments and routing paths without a formal model and a model-specific solution procedure.

V. CONCLUSIONS

We have developed a binary integer linear programming model and a two-phase Lagrangian dual search procedure to determine a close-optimal solution for the problem of expanding link capacities and routing multimedia service demands on a fiber optic network. The two-phase subgradient search procedure first solves an aggregate formulation of the Lagrangian dual, quickly producing a dual feasible solution. Then, it solves a disaggregate formulation starting from the incumbent solution found in the first phase.

This two-phase approach generates a better solution than the conventional one-phase approach to solving either formulation. For realistic problems (with network nodes up to 50 and the average node degree up to 2), it finds a primal solution whose value is within 2% from the Lagrangian bound within about 10 minutes of CPU time.

This research can be extended to address the situations where the location of a level-1 gateway is determined together with link capacity assignments and routing; where multiple level-1 gateways supply multimedia services to a known set of demands on a fiber optic network; where multiple fiber cables can be added to a network link; and where installing new fiber links is allowed to meet a demand growth in addition to expanding existing links. We hope this work enhances insights into the problems of designing multimedia service networks and is useful for future extensions.

REFERENCES

[1] M. L. Fisher, "The Lagrangian relaxation method for solving integer programming problems," *Management Science*, vol. 27, no. 1, pp. 1–18, 1981.

[2] M. Gerla, J. A. S. Monteiro, and R. Pazos, "Topology design and bandwidth allocation in ATM nets," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 8, pp. 1253–1262, 1989.

[3] Informa Media & Telecoms, *On-Demand TV & PVRs*, 2008.

[4] V. O. K. Li, W. Liao, X. Qiu, and E. Wong, "Performance model of interactive video-on-demand systems," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 6, pp. 1099–1109, 1996.

[5] J. M. McManus and K. W. Ross, "Video-on-demand over ATM: Constant-rate transmission and transport," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 6, pp. 1087–1098, 1996.

[6] D. Medhi, "Multi-hour, multi-traffic class network design for virtual path-based dynamically reconfigurable wide-area ATM networks," *IEEE/ACM Trans. Netw.*, vol. 3, no. 6, pp. 809–818, 1995.

[7] D. Medhi and D. Tipper, "Some approaches to solving a multihour broadband network capacity design problem with single-path routing," *Telecommun. Syst.*, vol. 13, no. 2–4, pp. 269–291, 2000.

[8] J. P. Nussbaumer, B. V. Patel, F. Schaffa, and J. P. G. Sterbenz, "Networking requirements for interactive video on demand," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 5, pp. 779–787, 1995.

[9] J. S. Park, B. H. Lim, and Y. Lee, "A Lagrangian dual-based branch-and-bound algorithm for the generalized multi-assignment problem," *Management Science*, vol. 44, no. 12, pp. 271–282, 1998.

[10] P. V. Rangan, H. M. Vin, and S. Ramanathan, "Designing an on-demand multimedia service," *IEEE Commun. Mag.*, pp. 56–64, July 1992.

[11] Y. Sato, and K. Sato, "Virtual path and link capacity design for ATM networks," *IEEE J. Sel. Areas Commun.*, vol. 9, no. 1, pp. 104–111, 1991.

Byung Ha Lim received B.Sc. degree in Engineering from Seoul National University, and M.Sc. degree in Business Computing Science from Texas A&M University and Ph.D. degree in Business Administration majoring Management Information Systems from University of Iowa in 1997. From 1997 to 2002, he has been Professor at University of San Francisco. He is currently Professor at Chung-Ang University, Korea. His research interests include the analysis and modeling data communication systems, high-performance networking, e-business strategy, technology and management. He has published many papers in Management Science, Telecommunication Systems, Information Technology and Management, Asia Pacific Journal of Information Systems, and so forth.

June Sung Park received Interdisciplinary Ph.D. in Computer Science and Industrial Engineering from Ohio State University in 1988, and served as Professor of Information Systems at Louisiana State University and University of Iowa from 1987 to 2000. From 2001 to 2009, he served as Executive Vice President and Chief Technology Officer for Samsung SDS, an IT service provider with 10 thousand employees and US $4 billion revenue. Currently, he is Professor in Korea Advanced Institute of Science and Technology with research and teaching focus on high-technology service engineering and management. He is also President of Korea Software Technology Training Institution funded by Korean government. His areas of expertise include business process engineering, software engineering, database systems, telecommunication networks and operations research. He was Chair of the Technical Section on Telecommunications in the Institute for Operations Research and Management Science in the U.S. during 1998–2000, and has published numerous papers in international journals such as European Journal of Operations Research, IEEE Transactions on Knowledge and Data Engineering, Information Systems, Informs Journal on Computing, International Journal of Technology Management, Journal of Heuristics, Management Science, Telecommunication Systems, and so forth. He is Associate Editor for two journals: Telecommunication Systems and Information Technology and Management.