FTTC/xDSL Network Planning under Uncertain Demand

J.S. Choi*, D.W. Tcha*, S.H. Lee*, and D.H. Ahn**

* Graduate School of Management, KAIST(Korea Advanced Institute of Science and Technology)
207–43, Cheongryangri-Dong, Dongdaemun-Gu, Seoul, 130–012, Korea
Phone: +82–2–958–3377, Fax: +82–2–958–3376
e-mails : jschoi@telmal.kasit.ac.kr, tchadw@sorak.kaist.ac.kr, iskra@kgsm.kaist.ac.kr
** Wise i tech co.
1702, Seoulventuretown, 679-5, Yuksam-dong Kangnam-gu, Seoul, 135-080, Korea
Phone: +82-2-6246-1400
e-mail : dhahn@wise.co.kr

ABSTRACT

Dramatic growth in broadband demand makes it inevitable for network operators to upgrade their access network, FTTC/xDSL is one of the most attractive technologies the telephone companies can take, since it utilizes the existing twisted copper pairs while providing enough bandwidth for the time being. Studies have been performed on the access network expansion planning problems, but they overlooked one of the most important factors, the demand uncertainty. Though considering uncertainty makes the problem more difficult, we cannot let it unchallenged when we think of the enormous amount of the investment to be made on the access networks. In this paper, we devise an algorithm to solve the problem of FTTC/xDSL planning under uncertain demand optimally, using the column generation technique and applying DP, and finally compare the results of this algorithm and the traditional sequential approach.

KEYWORDS

FTTC/xDSL, Residential Broadband, Access Network Planning

1. Introduction

With the explosively increasing demand for data transmission capacity caused by the spectacular growth in the multimedia service as well as the Internet traffic, it is inevitable for the network operators including telephone companies to upgrade their access networks. These networks, changed little since installed, have been until very recently regarded as the bottleneck of the Broadband era. [1] A number of solutions for broadband network access via various physical media, however, are currently being available. Digital Subscriber Line (DSL or xDSL) and Hybrid Fiber / Coax (HFC) technologies, which use the existing twisted copper pairs for telephone service and the coaxial cable for the traditional CATV service, respectively, are very attractive because of their maximum use of the existing media. Other technologies using optical fiber and wireless or satellite-based technologies are being developed for the places where very high data rates and mobility may be needed. The special attention, however, has to be called for making an investment in the access networks since these networks account for over 50% of the total investment in communication facilities. [2] One of the major reasons for the heavy investment is the dispersed nature of the customers.[3]

The length of phone line heavily affects the achievable bandwidth. FTTC(fiber to the curb)/xDSL brings the high-bandwidth capability to customers who are connected by twisted pairs and served by what is called an ONU (optical network unit). For example, if an ONU is located about 300m away from a customer, in such places as a street pedestal, controlled vault, or possibly a telephone pole, the bandwidth can be as large as 50Mbps.[1,4]

The methods to resolve demands by locating multiplexers (or concentrators) instead of installing huge amount of twisted-pairs in the associated area sections have been studied by Balakrishnan[2] and Shaw[5]. They devised very efficient algorithms to solve the local access telecommunication network expansion planning problem (LATN) exploiting the special structure of the underlying graph, tree.

The problem to construct FTTC/xDSL networks on the existing telephone networks are very similar to the LATN problem, in the sense that we locate ONUs instead of concentrators and decide groups of customers to be served by each of them. However, there are two marked differences between POTS (plain old telephone service) and broadband service: the difficulties in forecasting demand and the existence of several broadband service alternatives. Since broadband service has not a long history and the market is rapidly growing, enough data has not been accumulated for forecasting. Furthermore, the practice of forecasting is usually done in small sized units, enlarging the forecast variance. Fierce service competitions among operators make customers expect quick service installation. The shortage thus leads to the lost of sales directly and influences the future sales indirectly. The uncertainty in demand is one of key complicating factors that a network planner has to overcome. The traditional practice of the sequential approach starting from an arbitrarily given demand cannot be effectively applied. For example, let's suppose there are two adjacent demand nodes of which demands are not deterministic. Planners would determine the demand for each node a priori and then think of locating the ONUs. But if the sum of the preset demands in the first stage exceeds an ONU's capacity they have no choice but to locate two ONUs while locating an ONU may be more economic.

A popular approach of solving a complex problem is to divide it into some small subproblems and to make them solved separately and sequentially, with the solution of one subproblem used as data for the next subproblem. There are two main drawbacks associated to this multi-phase approach: 1) Without any knowledge on the optimum, it is difficult to set the parameters for some of the subproblems which appear in the earlier phases, and 2) wrong decisions taken at one of the earlier phases are "passed" to the subsequent phases. [6] In this paper, we integrate the two subproblems, the inventory problem and the facility location problem to a single comprehensive model and develop an algorithm to solve this model optimally and compare its performance with that of the sequential approach.

2. Problem Description

FTTC/xDSL networks are planned based on the tree-structured copper loop. The real world access network can be represented by an undirected tree $G=(V,E)$, where $V=\{1,2,...,n\}$. The root of the tree, G, node 1 in this paper, is the central office (CO) and a node, either a demand node or an ONU candidate node, corresponds to a distribution point (DP) or a group of DPs where an ONU can be located. For each node $i \in V$, the demand, X_i, has independent normal distribution with mean μ_i and variance σ_i^2. We assume that nodes of the tree G are labeled in a depth-first-search (DFS) order, starting with the root node 1. For the notation simplicity, we denote p_i^j be the relative predecessor of node i when the node j is assumed to be the root of the tree G. Let $T(j)$ be the subtree made of the successors of node j and itself, and $Last(j)$ be the last node of $T(j)$ in DFS order. Then, the goal of the problem in this paper is to find the most economic set of ONU locations and their capacities, and set of node disjoint cluster of demand nodes under uncertain demand (OLUD).

Problem OLUD

$$\min \sum_{j\in V} f_j(y_j)+\sum_{j\in V}\sum_{i\in V} c_{ij}\, x_{ij}+p\cdot\sum_{j\in V} E\Big\{\sum_{\{i|x_{ij}>0\}} X_i - y_j\Big\}^+ \quad (1)$$

$$\text{s.t.} \quad x_{ij} \le x_{p_i^j j} \quad \text{for all } i\in V,\ j\in V \quad (2)$$

$$\sum_j x_{ij} \le 1 \quad \text{for all } i\in V \quad (3)$$

$$y_j \ge 0 \text{ and integer for all } j\in V$$

$$x_{ij}\in\{0,1\} \text{ for all } i\in N,\ j\in V$$

The cost of locating ONU at node j, $f_j(.)$, is composed of the location-dependent fixed cost and ONU itself, and its installation costs varying with the capacity. To be consistent with reality, we assume this function has the staircase shape. The cost of serving demand of node i by the ONU located at node j, $c_{ij}(.)$, depends on the demand distribution of node i and the cable installation or rewiring amount. Most customer premises are directly connected to CO through copper or fiber cable and therefore cable expansions are almost not needed. So the most influential costs are the shortage and the rewiring cost. Rewiring might be required for the link between the nodes and their serving ONU if they are not directly connected. The shortage may be occurred in the DP level [3]. The physical connections, however, are considered to be always available and it is easy to allocate the bandwidth (or the capacity) of an ONU to demand nodes that the DP level shortage does not occur unless the realized demand of the group served by an ONU exceeds the ONU capacity (ONU level shortage). The third term of the objective function, one of the most distinguishing part from other studies, is the expected penalty cost caused by the demand uncertainty, the shortage of ONU capacity. The expected shortage cost is determined by the capacity of an ONU and the demand distribution of the cluster served by the ONU. Note that even for a given

ONU location and capacity, the expected shortage cost can be varied according to the members of the cluster. Constraints (2) force each cluster to be contiguous, i.e., if a node i is served by the ONU at node j then all nodes on the path between the node i and node j should be served by the ONU. In real world, this rule is kept for the sake of easier OAM. Constraints (3), together with the integrality of the variables x_{ij}, are for the non-bifurcation of demand requirements. Since we allow the ONU level shortage in this model, we don't have to provide all demand nodes and relax the requirement to connect each demand node to one of ONUs.

This problem can be treated as one of optimal subtree packing problems with its objective function non-linear. The subtree packing problem can be solved by such methods as Primal column generation algorithm, Lagrangian relaxation, and Dynamic programming. The third term of the objective function suggests the use of the column generation algorithm, and the special-structure of the underlying graph the application of dynamic programming.

3. Reformulation

In this section, the problem is reformulated into a column generation problem, and the related theories are presented. We first reformulate the ONU location under uncertainty problem using subtrees

For a given tree $G=(V,E)$, there are a number of subtrees T_1, T_2, ..., with each $T_j \subseteq V$ for $j=1,2,\ldots$.

Each subtree T_j has an associated value c_{T_j}. Let $\mathbf{T}$ denote the collection of all subtrees in the underlying network $G=(V,E)$. Note that we are not given the subtrees and their values explicitly. Let A be the node-subtree incident matrix, i.e., $a_{ij}=1$ if node i is in subtree T_j, and $a_{ij}=0$ otherwise. Each subtree is composed of an ONU location node and a set of contiguous demand nodes served by this ONU. Each decision variable w_{T_i} denotes the existence of the corresponding subtree in the solution. For a given subtree T_j, there are a number of combinations of ONU location and its capacity. Note that the solution of the reformulated problem is the set of disjoint subtrees that constitute the whole tree. The subtree with the lowest cost dominates the subtrees consisted of the same nodes. So we consider subtrees with the lowest cost only. Let $c_{T_j} = \min_k \{\sum_i c_{ik}\, x_{ik} + f_k(y_k)\}$ denote the cost of subtree T_j.

Main problem

$$\min \sum_{T_j \in \mathbf{T}} c_{T_j} w_{T_j} \tag{4}$$

$$\text{s.t.} \quad A w_{T_j} = 1 \tag{5}$$

$$w_{T_j} \in \{0,1\} \text{ for all } T_j \in \mathbf{T}$$

Let u_i s be the duals corresponding to the constrains (5). Then the subproblem is to find the minimum cost subtree rooted at node j with the available nodes confined to the nodes in $T(j)$. For simplicity of explanation we assume that the ONU is located at node j.

Subproblem TKUD(j)

$$\text{Min } f_j(y_j) + \sum_{i \in T(j)} c_{ij} x_{ij} - \sum_{i \in T(j)} u_i x_{ij} + p \cdot E\left\{\sum_{\{i|x_{ij}>0\}} \beta_i - y_j\right\}^+$$

$$\text{s.t. } \quad x_{ij} \le x_{p_i j}$$

$$x_{ij} \in \{0,1\} \text{ , } y_j \in \{0, a_{j1}, a_{j2}, \ldots\}$$

Solution methods of this reformulated problem are composed of two main parts. One is how to find the columns of minimum costs and the other is how to solve the main problem efficiently. Fortunately, we have the structure of the main problem same as the LATN problem. Therefore if we find the minimum cost subtree rooted at k for $k = 1,\ldots,n$ in the backward order of DFS, then the main problem can be solved in n iterations. [5]

The states and stages of the knapsack problem, which Shaw used for his LATN algorithm, cannot be applied to the subproblem TKUD. The available capacities, the states defined in solving the knapsack problem by dynamic programming, cannot be defined in this problem since the shortage is allowed. Moreover, the available capacities (or the capacities used by the group of demand nodes) are not the only information in deciding whether to include a node in the subtree. Let's suppose that we are given tree G as shown in the figure 1. Assume that the set of node 1 and node 2, $\{1,2\}$, and the set of node 1 and node 3, $\{1,3\}$, have the same cost and mean demand. The DP for the knapsack does not differentiate the set $\{1,2\}$ and the set $\{1,3\}$ when we are at the stage of 4. But the set $\{1,2\}$ and the set $\{1,3\}$ can have different variances, say, 0 and non-zero, respectively. In that case, we cannot say that two sets are indifferent because the values of the set $\{1,2,4\}$ and the set $\{1,3,4\}$ may be different in the next stage.

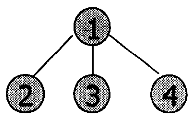


Figure 1. Subtree Example

4. Application of Dynamic Programming

To solve the TKUD problem, we have to consider all contiguous subtrees to compare their values. We devised an algorithm based on the knapsack DP while considering the contiguity property that affects the dependency of the nodes in TKUD. The state of the system at any point of time contains information on the mean and variance of the nodes in the subtree. First, we assume that we are dealing with the subproblem j, i.e., the problem confined to the subtree $T(j)$. For each $k = j$ to $Last(j)$, the states in stage k will be denoted by (k,a,b) where $a \ge 0, b \ge 0$. In any state (k,a,b) in stage k, there are two tasks; (i) to find the subtrees that include the parent node of node k and to calculate the value function; and (ii) to remove (k,a,b) which is uneconomic afterward. Now we define the value function, in state (k,a,b) to be

$$\begin{aligned} f(k,a,b) &= h(k,a,b) + g(a,b) \\ &= h(i, a-\mu_k, b-\sigma_k^2) + c_{kj} - u_k + g(a,b) \end{aligned} \qquad \text{for } j \le i \le k-1$$

$f(k,a,b)$ is the minimum cost that can be achieved if the mean and the variance of the subtree are a and b, respectively. It consists of the stage dependent linear term and the independent non-linear term. $h(a,b)$ is defined as the minimum cost (or maximum revenue) of the contiguous node set of which sums of means and variances are a and b, respectively. $g(a,b)$ is the expected shortage cost for fixed ONU capacity, say H, when means and variances are a and b, respectively. Since the number of subtrees exponentially increases with the number of nodes, it would be practically impossible to enumerate all subtrees by this algorithm. There are some properties on $g(a,b)$ that would be applied in the algorithm and decrease the number of calculations.

Property 1. *For any variance* $b \geq 0$, *if* $a \leq c$, $g(a,b) \leq g(c,b)$ *and* $g(a+\Delta,b) \leq g(c+\Delta,b)$ *hold for* $\Delta \geq 0$.

<Proof> Note that first- and second-partial derivatives of $g(a,b)$, $\dfrac{\partial g(a,b)}{\partial a}$ and $\dfrac{\partial^2 g(a,b)}{\partial a^2}$ respectively, are non-negative for any a.

Property 2. *For any mean demand* $a \geq 0$, *if* $b \leq d$, $g(a,b) \leq g(a,d)$ *and* $g(a,b+\Delta) \leq g(a,d+\Delta)$ *hold for* $\Delta \geq 0$.

<Proof> Note that the first- and second-partial derivatives of $g(a,b)$, $\dfrac{\partial g(a,b)}{\partial b}$ and $\dfrac{\partial^2 g(a,b)}{\partial b^2}$, are non-negative for any b.

Corollary 3. *If* $a \leq c$ *and* $b \leq d$, *then* $g(a,b) \leq g(c,d)$ *and* $g(a+\Delta_1,b+\Delta_2) \leq g(c+\Delta_1,d+\Delta_2)$ *hold for* $\Delta_1 \geq 0$ *and* $\Delta_2 \geq 0$.

Definition. We define that $f(k,a,b)$ dominates $f(k,c,d)$ if $f(k,a+\Delta_1,b+\Delta_2) \leq f(k,c+\Delta_1,d+\Delta_2)$ hold for $\Delta_1 \geq 0$ and $\Delta_2 \geq 0$.

Theorem 4. *When the location and capacity of the facility (or ONU) is fixed,* $f(k,a,b)$ *dominates* $f(k,c,d)$ *if* $f(k,a,b) \leq f(k,c,d)$ *and* $a \leq c$ *and* $b \leq d$ *are satisfied.*

<Proof> The value function $f(k,a,b)$ is composed of the node dependent cost and the shortage cost. From corollary 3, if $a \leq c$ and $b \leq d$ then $g(a,b) \leq g(c,d)$ and $g(a+\Delta_1,b+\Delta_2) \leq g(c+\Delta_1,d+\Delta_2)$. Addition of a node brings the same effect of adding to both of $h(k,a,b)$ and $h(k,c,d)$.

By concretizing the idea of Theorem 4 in the algorithm, we can reduce the number of calculations. There are other dominance relationships to further improve the efficiency of the algorithm, which is reserved for the reader.

Algorithm TKUD;

```
begin
   /* To assign an ONU at each node and calculate the cost for each node */
   for k = 1 to N
      dual(k) = y(k) + g((H-mean(k))/sqrt(variance(k))) * sqrt(variance(k))
   /* To find the minimumcost subtrees rooted at node k */
   for k = N to 1 step -1
```

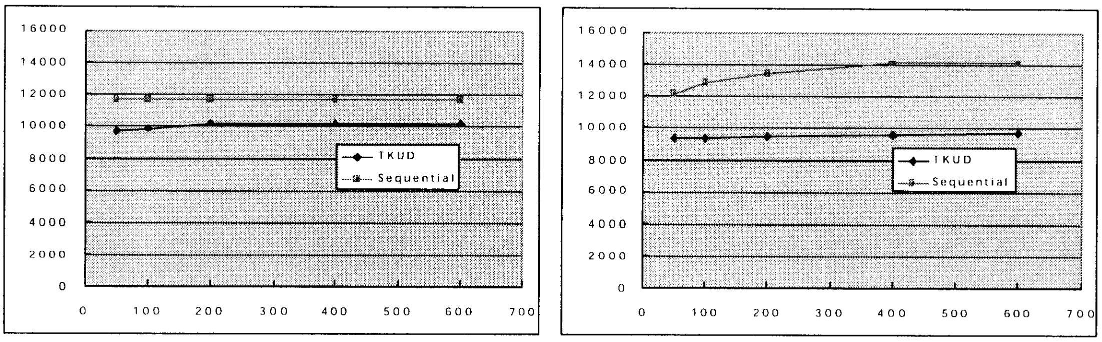
```
/* To start DP with a linked list with the root f(k, mean(k), variance(k)) */
 for each j, mean, variance in the list
     delete f(j, mean, variance)
i = k
f(k,mean(k),variance(k)) = 0
/* To search all nodes to find the least cost subtree rooted at k */
while i <> LastNode(k)
    begin
        i = i + 1
        for each j, mean, variance in the list such that Parent(i) ≤ j < i
            f(i, mean+mean(i), variance+variance(i))
                = ONUcost(k) + g((H- (mean+mean(i)))/sqrt(variance+variance(i)))
                    * sqrt(variance+variance(i)) + sum of duals in the subtree
        /* if i is the leaf node then remove f(j,mean,variance) dominated */
        if i = LastNode(i) then
            for each j, mean, variance such that LastNode(j) = i
            delete f(i, mean, variance) from the list if dominated
    end
    find the least valued f(i, mean, variance) in the list
    and store the subtree in the solution
next k
end
```

In this algorithm, we assume that an ONU is located at the root of subtree $T(j)$ and the capacity of the ONU is fixed at H, the generalization of the algorithm is left for readers. Note that $g(a,b)$ is the function of the normal distribution that can be transformed to the standard normal distribution. Instead of calculating $g(a,b)$, we can use the equation $g(a,b) = \sqrt{b} \cdot g((a-H)/\sqrt{b},1)$. The last note for the algorithm is that for the time-consuming job of finding dominated subtrees, suggested is the use of a threshold value, and mean and variance such as those of the root node.

5. Computational Results

To test the above algorithm, we used two sets of problems, differing in the variance scale. Each set of problems has its own ratio of the shortage costs to the ONU costs. We coded the algorithm in Visual Basic Application language of the Excel and ran on a Celleron processor PC. It took on the average around 1 minute to exactly solve problems with 40 nodes.

We compare the TKUD algorithm with the sequential method, presetting the demand for each node and solving the facility location problem. As expected, the TKUD algorithm outperformed the sequential method averagely by 20%. And we found that the larger the variance, the more difference in the objective values of two methods. The effects of the ratio of the shortage cost to the facility cost are found in one of two ways, either increasing or decreasing of the objective values. This phenomenon can be interpreted as follows: (i) the increased ratio forces each node to have more deterministic demand in the sequential method (ii) if the ratio is low, an ONU will serve more demand nodes in TKUD algorithm since the shortage is relatively more favorable than installing more ONUs.



(a) low variance (b) high variance

Figure 2. Comparison of TKUD and sequential method

6. Conclusion

In this paper, we formulated the ONU location problem under uncertain demand and solve it optimally by column generation technique and application of knapsack DP. We tested this algorithm and found that it outperformed the sequential approach. Contributions of this paper are that we considered the demand uncertainty that is one of the most important factors in the real world broadband access network planning and design. Exploiting the properties of the shortage function and we made an efficient algorithm to solve the unified and complex problem optimally while not so much computation increased. This increase may be the reasonable cost for the uncertainty consideration.

REFERENCES

[1] Modiano E. and Olcer S., (1999) "Broadband Access Networks", Computer Networks, Vol. 31, 279-281

[2] Balakrishnan A. and Matragi T. L. and Wong R. T., (1995) "A Decomposition Algorithm for Local Access Telecommunications Network Expansion Planning", Operations Research, Vol. 43, No. 1, 58-76

[3] Tcha D. W. and Choi J. S., (1999) "Comparative Economic Analysis between Direct and Indirect Wiring in the Copper-based Local Loop", Vol. 70, 531-535

[4] Cioffi J., Siverman P. and Starr T., (1999), "Digital Subscriber Lines", Computer Networks, Vol. 31, 283-311

[5] Shaw D. X. and Cho G., (1994) "Limited Column Generation for Local Access Telecommunication Network Expansion and Its Extension", Technical Report, School of IE, Purdue University, West Lafayette, Indiana

[6] Gouveia L. and Lopes M. J., (1997) "Using generalized capacitated trees for designing the totpology of local access networks", Telecommunication Systems, Vol. 7, 315-337