Journal of the Korean
OR/MS Society
Vol.15, No.1, June, 1990

A Fixed Priority Queue Median with Jockeying on a Network

Kyung-Hee Jung*

Abstract

This paper is concerned with determining a minisum location with jockeying for a server on a probabilistic network in which each customer type enters the network system permitting with jockeying through a specified node and a nonpreemptive service policy is in effect. An algorithm to locate a single Fixed Priority Queue Median with Jockeying (FPQMJ) on acyclic networks is developed by using the Generalized Benders' Decomposition technique. The results are then extended to a general network.

1. Introduction

Minisum location problems incorporating queueing aspects were first by Berman, et. al. [2]. They studied the problem in which customers requiring service enter a network only through the nodes in a Poisson stream and called the optimal single server location, the Stochastic Queue Median (SQM). Each customer joins a nodal queue if the server is busy, and waiting customers are served using a First-Come-First-Served (FCFS) service rule. The objective is to minimize the mean response time for all customers. In this context, the mean response time refers to the sum of the travel time and the mean waiting time in the queue.

* Policy & Economic Analysis Dept. Korea Electrotechnology Research Institute

They develop an algorithm for the SQM on a general network. Chiu, et. al. [3] consider the SQM on a tree network and develop two specialized algorithms by utilizing the properties of convexity.

Chiu and Larson [4] consider an optimal location of p-servers on a general network with zero queue capacity, that is referred to a p-server-facility-loss-median (p-SFLM). They prove that p-SFLM corresponds to Hakimi's median if the costs of lost demands are nonnegative.

On a network congested with a queueing system, the servers must return to their base location. Otherwise, the service times for service requests are not independent and available results in queueing theory for an M/G/c system is no longer applicable. The First-Come-First-Served (FCFS) queue discipline is assumed in each case.

More recently, Jung, et. al. [11] developed algorithms for determining a Fixed Priority Queue Median (FPQM) on chain and tree graphs. A FPQM is a minisum location for a server on a probabilistic network in which each priority class of customers enters the network through a specified node and a nonpreemptive service policy is in effect.

In this paper, we consider the problem in which customers requiring service enter the system through priority nodes and are allowed to jockey with some penalties to higher priority nodes. The objective is to find the location of a single server and the switched customer rates which minimize the mean weighted waiting time in the queue, assuming that the network system permits the customer's jockeying behavior.

This paper is organized as follows. We begin with a detailed problem description and then discuss the structure of the objective function over a path by utilizing the properties of convexity. An algorithm to locate a single Fixed Priority Queue Median with Jockeying for queue position on a chain graph is developed by using the Generalized Benders' Decomposition. These results are then extended to tree graphs and general networks.

2. Problem Description

Consider a finite, connected and undirected stochastic network $G(N, L)$ where $N(|N|=n)$ and L denote the sets of nodes and links, respectively. The arrival pattern of demands through the nodes and the service scheme are probabilistic. Then lengths of links are nonnegative and there are no multiple links between any two nodes on G. Then the shortest path distance, $d(x, y)$, is well-defined between two points x, y on G.

Assume that the priority of each node is predetermined and remains constant in time, the higher the nodal number, the higher the priority. Type i customers requiring service enter the network system only through node i. Whenever a customer arrives for service and finds that a traveling server is not available, then it is possible for customers to react in various ways. A customer may decide to wait no matter how long the queue becomes. If the queue is too long, the customer may leave the network system without waiting. On the other hand, an impatient customer may react in some other ways after joining the queue. For example, balking, reneging, jockeying, or bribing or cheating for queue position. In this paper, we will assume the input process allows "jockeying for queue position" in which the objective is to minimize the mean weighted waiting time in the queue for all customers. Non-preemptive service policy independent of a customer's priority is in effect. Each nodal queue capacity is infinite, and as soon as a server is free, he chooses the highest priority node among the existing nodal queues and serves the nodal customer. Within a priority group, a First-Come-First-Served (FCFS) queue discipline is observed. Otherwise, a server stays at his home location on G.

In practice, however, it seems reasonable that customers jockeying for queue position are penalized in proportion to the traveling time of a jockeying customer to the switched node. We assume here that after switching to node i from node j, he has the mean nodal service time of node i. Such a customer will be called a pseudo-customer of node i in this paper.

Because of the penalty and the new mean nodal service time, switching to a higher priority node does not necessarily reduce the mean weighted waiting time in the queue. Therefore, the main concerns are what proportion of the arrival rates at nodes should be switched and where the server's home should be located to optimize the problem, assuming that the network system permits the customer's jockeying behavior.

Service requests through nodes are generated independently in homogeneous (stationary) Poisson streams, each with mean rate λ_i and the nodal service times vary with priority groups. The service time has two compnents – a nodal service time and a traveling time on links. These times are independent, and indentically distributed with a general distribution.

The independence of service times restricts a server to make a round-trip between a home location and each node on G requiring service. Otherwise, the service times for the different group customers will depend on the previous customer types. In this case, available results in queueing theory for an M/G/c system are no longer applicable. Therefore, it is assumed that whenever available, a server travels to the demanding node with the highest priority, serves a customer with the appropriate given nodal service time, then returns to his home location on a network. If there is still any other node having a queue at any node, he starts to serve. If not, he remains at his home location until the network system becomes busy.

If we represent the actual service time of each customer type, it is the sum of twice the travel time to each node demanding service and its nodal service time. Denote the mean service time of type i customers by $E(S_i(x))$ where a server is located at x on G. Then

$$E(S_i(x)) = 2t(i,x) + E(Y_i),$$

$$t(i,x) = d(i,x)/v,$$

where $E(Y_i)$ is the average nodal service time at node i,
$d(i,x)$ is the shortest path distance between node i and a server's home location, x,
v is the traveling velocity of a server over a path, and
$t(i,x)$ is the shortest traveling time between node i and x.

As defined earlier, node n has the highest priority and node 1 the lowest on $G(N,L)$.

By allowing jockeying for priority customers on the network system, the new average arrival rates at node $i(i=1,\cdots,n)$, λ_i', are :

$$\lambda_i' = \lambda_i - \sum_{k=i+1}^{n} \gamma_{ik} + \sum_{k=1}^{i-1} \gamma_{ki} \tag{1}$$

where γ_{ij} is the rate of pseudo-customers from node i to j.

Now the objective of this paper is to find an optimal location of a single server on G and optimal pseudo-customer rates which minimizes the weighted sum of the mean waiting time in the queue of each priority class. This model is basically a minisum location (median) problem congested with a fixed priority queueing system with jockeying. This study seeks an optimal server location stated as above which is called a Fixed Priority Queue Median with Jockeying (FPQMJ).

3. A Single on a Probabilistic Network Incorporating an M/G/1 Priority Queue

The mean waiting times in the queue of each priority class, W_i, for an M/G/1 nonpreemptive, fixed priority queueing system with an infinite queue capacity, are found by Cobham[5] :

$$W_i = \begin{cases} \dfrac{\sum_{j=1}^{n}\lambda_j E(S_j^2)/2}{(1-\sum_{j=i}^{n}\lambda_j E(S_j))(1-\sum_{j=i+1}^{n}\lambda_j E(S_j))} & \text{if } \sum_{j=i}^{n}\lambda_j E(S_j) < 1 \\ \infty & \text{otherwise} \end{cases}$$

where λ_i is the constant mean arrival rate of type i customers,

$E(S_i)$ is the mean service time for type i customers, and

$E(S_i^2)$ is the second moment of service time for type i customers.

When the network system is operating on an M/G/1 fixed priority queue discipline, the mean waiting time in the queue of each customer group desiring service can also be formulated in terms of a server location, x, on G and a new arrival rate, $\lambda_i'(\tilde{\gamma})$ by utilizing Cobham's solution with constant service time and arrival rate for each class.

Therefore, the mean waiting time in the queue of the i^{th} priority class, $W_i(\tilde{\gamma}, x)$, at a server location, $x/\varepsilon/G$, is :

$$W_i(\tilde{\gamma}, x) = \begin{cases} \dfrac{\sum_{j=1}^{n}\lambda_j'(\tilde{\gamma})E(S_j^2(x))/2}{(1-\sum_{j=i}^{n}\lambda_j'(\tilde{\gamma})E(S_j(x)))(1-\sum_{j=i+1}^{n}\lambda_j'(\tilde{\gamma})E(S_j(x)))} + t(i, x) & \text{if } \sum_{j=i}^{n}\lambda_j'(\tilde{\gamma})E(S_j(x)) < 1 \\ \infty & \text{otherwise} \end{cases} \quad (2)$$

where

$$E(S_i(x)) = 2t(i, x) + E(Y_i) = \frac{2}{v} d(i, x) + E(Y_i), \quad (3)$$

$$E(S_i^2(x)) = \frac{4}{v^2} d^2(i, x) + \frac{4}{v} E(Y_i) d(i, x) + E(Y_i^2), \quad (4)$$

$E(Y_i)$: the mean nodal service time of customer type i,

$E(Y_i^2)$: the second of the nodal service time of customer type i,

$E(S_i(x))$: the mean service time for type i customers at a server location, x,

$E(S_i^2(x))$: the second moment of service time for type i customers at x.

It is assumed that each $E(Y_i)$ and $E(Y_i^2)$ are known for $i=1, 2, \cdots, n$, and v is a constant on G.

In the finite case of (2), the first term in the right side is the mean waiting time of type i customers caused by customers of higher priority or the same priority that entered earlier. To get the actual waiting time, the traveling time of the server, $t(i, x)$, should be added to the first term of (2).

Notice that the value of the first term in (2) decreases with increasing priority. However, this is not necessarily true if the second term, $t(i, x)$, is added. Therefore, $W_i(\tilde{\gamma}, x)$ and $E(S_i(x))$ for each priority type i vary with a server location, $x/\varepsilon/G$, for a constant $\tilde{\gamma}$.

Our objective is to minimize the mean waiting time in the queue for all nodal demands, regardless of their priority types. That is, the model is considered from the viewpoint of the server. Then the average wait in the queue over all customers is :

$$\sum_{i=1}^{n} \frac{\lambda_i'(\tilde{\gamma})}{\lambda'(\tilde{\gamma})} W_i(\tilde{\gamma}, x),$$

where $W_i(\tilde{\gamma}, x)$ is the mean waiting time in queue of customer type i,

$\lambda'(\tilde{\gamma})$ is the sum of the mean arrival rates of type i customer, $\lambda_i'(\tilde{\gamma})$, for $i=1, 2, \cdots, n$, i.e. $\lambda'(\tilde{\gamma}) = \sum_{i=1}^{n} \lambda_i'(\tilde{\gamma})$, and x is a server's home location on G.

Other than fixed nodal priorities, relative nodal weights due to service times depending on a server's home place are considerable. This is reflected in the measures by which the mean waiting times are evaluated. We shall assume that values, $E(S_i(x))$, describe relative importance for priority group i for $i=1, 2, \cdots, n$, that is, these represent normalizing factors in the performance measure.

Therefore, the mean weighted time in the queue is given by :

$$\sum_{i=1}^{n} E(S_i(x)) \frac{\lambda_i'(\tilde{\gamma})}{\lambda'(\tilde{\gamma})} W_i(\tilde{\gamma}, x), \text{ for } x \varepsilon G.$$

Since the sum of all arrival rates, $\lambda'(\tilde{\gamma}) = \sum_{i=1}^{n} \lambda_i'(\tilde{\gamma})$, is a constant, the above is equivalent to :

$\sum_{i=1}^{n}\lambda_i'(\tilde{\gamma})E(S_i(x))W_i(\tilde{\gamma},x)$ for $x\varepsilon G$.

In this paper, we assume that customers jockeying for queue position are penalized by the traveling time of a pseudo-customer to the switched node. Then the total waiting time of a pseudo-customer is the sum of the waiting time at his new nodal queue and the traveling time to the switched node from his previous node. That is, the mean waiting time of jockeyed customers at node i, $W_i'(\tilde{\gamma},x)$, is :

$$W_i'(\tilde{\gamma},x)=\sum_{k=1}^{i-1}\gamma_{ki}\{W_i(\tilde{\gamma},x)+\frac{1}{v'}\ d(k,i)\}$$

where v′ is the traveling velocity of pseudo-customers on a network.

Note that traveling velocities of pseudo-customers are equal, regardless of their priority types and pseudo customers at node i are from lower priority types, types 1 through i-1. Then, a mathematical statement of the problem is as follows :

$$\min_{\substack{x\,\varepsilon\,G(N,L)\\ |N|=n}} AWJ(\tilde{\gamma},x) \tag{5}$$

subject to $\gamma_{ij}\geq 0$

$\lambda_i'(\tilde{\gamma})\geq 0$ for $i, j=1,2,\cdots,n$

where $AWJ(\tilde{\gamma},x)$ denotes :

$$AWJ(\tilde{\gamma},x)=\sum_{i=1}^{n}\lambda_i'(\tilde{\gamma})E(S_i(x))W_i(\tilde{\gamma},x)+\frac{1}{v'}\sum_{i}E(S_i(x))\{\sum_{k=1}^{i-1}\gamma_{ki}d(k,i)\}.$$

Constraints in (5) result from the restrictions that the rates of flow-in and flow-out are nonnegative, and the new nodal rates after switching should be also nonnegative.

Substituting $W_i(\tilde{\gamma},x)$ of Eq. (2) into Equation (5), then $AWJ(\tilde{\gamma},x)$ becomes :

$$AWJ(\tilde{\gamma},x)=\frac{\{\sum_{i=1}^{n}\lambda_i'(\tilde{\gamma})E(S_i^2(x))\}\{\sum_{i=1}^{n}\lambda_i'(\tilde{\gamma})E(S_i(x))\}}{2\{1-\sum_{i=1}^{n}\lambda_i'(\tilde{\gamma})E(S_i(x))\}}$$
$$+\frac{1}{v}\sum_{i=1}^{n}\lambda_i'(\tilde{\gamma})E(S_i(x))d(i,x)+\frac{1}{v'}\sum_{i=1}^{n}E(S_i(x))\sum_{k=1}^{i-1}\gamma_{ki}d(k,i).$$

Note that $\lambda_i'(\tilde{\gamma})E(S_i(x))$ is the utilization factor (or traffic intensity) of the i^{th} priority class at a server location x. This is the ratio of the rate at which customer enters the queueing system to the maximum rate at which the system can serve this customer. In particular, we would rather not locate a server for which the sum of the utilization factors is greater than or equal to one, whereby the mean waiting time of any customer type is infinite. Therefore, the objective value is apparently infinite.

We will now consider only the finite case of each mean waiting time in the queue whose condition is $\sum_{j=1}^{n}\lambda_j'(\tilde{\gamma})E(S_j(x)) < 1$ for type i customers.

Hakimi[8, 9] has shown the nodal optimality of minisum locations for both single facility and multi-facility on a general network. However, the nonlinearity of $AWJ(\tilde{\gamma}, x)$ on $x \varepsilon G$ does not limit the possible server location(s) to the nodes of G only. A server might be located anywhere, even at a point on a link of G. Therefore, existing algorithms for the minisum location model are not applicable here.

In the following section, we consider a single FPQMJ on a chain graph. Then in Sections 5 and 6, we extend the model to tree and general networks, respectively.

4. A Single FPQMJ on a Chain Graph

Consider a chain graph C with n distinct nodes designated 1, 2, ⋯, n. We will assume that the chain graph is laid out as an interval on the real line and l_j is the distance of node l_j from node 1. Note that $l_1 = 0$.

As far as γ_{ij} are known for i, j = 1, 2, ⋯, n, then the objective in (5) is only to find a server's location minimizing the mean weighted waiting time in the queue. $AWJ(\tilde{o}, x)$. In the following sections, AW(x) will denote the objective function with constant jockeyed rates. Then this problem can be formulated as :

$$\underset{\substack{x \in G(N, L) \\ |N| = n}}{\text{Min}} \; AW(x)$$

where

$$AW(x) = \frac{\{\sum_{i=1}^{n}\lambda_i E(S_i^2(x))\}\{\sum_{i=1}^{n}\lambda_i E(S_i(x))\}}{2\{1 - \sum_{i=1}^{n}\lambda_i E(S_i(x))\}} + \frac{1}{v}\sum_{i=1}^{n}\lambda_i E(S_i(x))\, d(i, x). \tag{6}$$

The following theorem follows directly from the properties of convex functions and for brevity is presented without proof.

Theorem 1 : Let AW : $C \to E_1$ be the mean weighted waiting time as defined in (6). Suppose that AW(x) is finite and twice differential on a chain C except at the two end nodes. Then AW(x) is convex on C, and furthermore, AW(x) is strictly convex if $\lambda = \sum_{i=1}^{n} \lambda_i$ is positive.

Since the FPQMJ model is a nonconvex program with two sets of variables, few procedures are available for its solution. By using the result of Theorem 1, the Generalized Benders' Decomposition (GBD) technique by Geoffrion[6] decomposes the problem of (5) into an iteration between two subproblems : a nonlinear program with a vector of nonnegative continuous variables, γ_{ij}, and a convex program with a server's location, x. Here, we utilize the concept of a complicating variable originally defined by J. F. Benders[1]. Benders created an approach for exploiting the structure of mathematical program problems with complicating variables which are temporarily fixed. In this case the problem reduces an ordinary linear program by fixing the values of the complicating variables, and Geoffrion generalized Benders' limit to a nonlinearity. With a fixed value of pseudo-customer rates, $\tilde{\gamma}$, the convex subproblem can be formulated as (6).

The convexity of AW(x) at $x \varepsilon C$ guarantees the existence of a local minimum, x*, which is also a global minimum. Although any standard line search procedure could be used to find x*, we will describe an efficient procedure for first localizing and then determining x*. This solution procedure will then be extended to tree graphs. Toward this end, let us rewrite AW(x) as follows :

$$AW(x) = \frac{1}{2}\left(\sum_{i=1}^{n} \lambda_i' E(Y_i^2)\right) g_1(x) + \frac{1}{v}\left(1 + \sum_{i=1}^{n} \lambda_i' E(Y_i)\right) g_2(x) + \frac{2}{v^2} g_3(x) + \frac{2}{v^2} g_4(x)$$

where $g_i(x)$ are defined :

$$g_1(x) = \frac{\sum_{i=1}^{n} \lambda_i' E(Y_i) + \frac{2}{v}\sum_{i=1}^{n} \lambda_i' d(i, x)}{1 - \sum_{i=1}^{n} \lambda_i' E(Y_i) - \frac{2}{v}\sum_{i=1}^{n} \lambda_i' d(i, x)},$$

$$g_2(x) = \frac{\sum_{i=1}^{n} \lambda_i' E(Y_i) d(i, x)}{1 - \sum_{i=1}^{n} \lambda_i' E(Y_i) - \frac{2}{v}\sum_{i=1}^{n} \lambda_i' d(i, x)},$$

$$g_3(x) = \frac{\left(\sum_{i=1}^{n} \lambda_i' d(i, x)\right)\left(\sum_{i=1}^{n} \lambda_i' E(Y_i) d(i, x)\right)}{1 - \sum_{i=1}^{n} \lambda_i' E(Y_i) - \frac{2}{v}\sum_{i=1}^{n} \lambda_i' d(i, x)},$$

$$g_4(x)=\frac{\sum_{i=1}^{n}\lambda_i' d^2(i,x)}{1-\sum_{i=1}^{n}\lambda_i' E(Y_i)-\frac{2}{v}\sum_{i=1}^{n}\lambda_i' d(i,x)}.$$

It can be readily verified that the functions $g_i(x)$, $i=1,2,3,4$, are convex on C. To initiate the localization procedure, we present two basic results shown in Jung, et. al. [11].

Theorem 2 : let $u, v : S \rightarrow E_1$ be convex functions over a convex set $S \subseteq E_1$. Suppose that $\hat{x}, \bar{x} \varepsilon S$ minimize $u(\cdot)$ and $v(\cdot)$, respectively. Then the optimal solution x^* that minimizes the function $(u+v)(\cdot)$ is located on the interval joining $\hat{x}$ and $\bar{x}$.

Theorem 3 : let $u, v : S \rightarrow E_1$ be nonnegative functions where $S \subseteq E_1$. Suppose that $\hat{x}$, $\bar{x} \varepsilon S$ minimize $u(\cdot)$ and $v(\cdot)$, respectively. If x^* minimizes the product of two functions, $u \cdot v(\cdot)$, then x^* is located on the interval joining $\hat{x}$ and $\bar{x}$.

The localization process will consist of applying the above theorems to the optimal solutions to the following subproblems :

(P1) $\min_{x \in C} \sum_{i=1}^{n} \lambda_i d(i,x)$

(P2) $\min_{x \in C} \sum_{i=1}^{n} \lambda_i E(Y_i) d(i,x)$, and

(P3) $\min_{x \in C} \sum_{i=1}^{n} \lambda_i d^2(i,x)$.

Note that solutions of subproblems (P1) and (P2) can be easily found by Goldman's [7] algorithm since these two subproblems are simply minisum location problems on a chain (tree). Also the solution of (P3) could be found using the necessary and sufficient conditions provided by Shier and Dearing [12] for nonlinear single-facility median type probems. However, we will now develop an efficient procedure for a chain graph.

Theorem 4 : $x_3^* = \frac{\sum_{i=1}^{n} \lambda_i l_i}{\sum_{i=1}^{n} \lambda_i}$ solves subproblem (P3).

Note that x_3^* is simply the ratio of the weighted sum of all distances from node 1 and the sum of all nodal weights (i.e. arrival rates). Thus the solutions of subproblems (P1), (P2), and (P3), x_1^*, x_2^*, and x_3^*, respectively, can be obtained quite easily via either Goldman's algorithm or Theorem 4.

Recall that $AW(x)$ and $g_i(x)$, $i=1,\cdots,4$ are convex. Therefore, by Theorem 2, a minimum of $AW(x)$ lies on the interval defined by the minimal points of $g_i(x)$, $i=1,\cdots,4$. Thus, minimizing $g_i(x)$, $i=1,\cdots,4$ would localize the initial interval of uncertainty on C. Firstly, the minimization of $g_1(x)$ is equivalent to solving subproblem (P1). Next, the minimum of $g_2(x)$ is located at a point between x_1^*, and x_2^* by Theorem 3, as is the minimum of $g_3(x)$. Similarly, the minimum of $g_4(x)$ is located between x_1^*, x_3^*. And since $AW(x)$ is a nonnegative combination of $g_i(x)$, $i=1,\cdots,4$, the solution, x^*, which minimizes $AW(x)$ will also lie in the interval defined by the solutions to subproblems (P1), (P2), and (P3).

Now that an initial interval of uncertainty has been established, any standard line search procedure can be utilized to obtain x^*. However, as an expedient, it can be seen that if at the k^{th} iteration, the interval of uncertainty, $[a_k, b_k]$, is a subset of a single link $[j, j+1]$, then $AW(x)$ can be rewritten as follows and x^* can be found directly from (7).

$$AW(x) = \frac{Sx^2+Tx+R}{1-c_1-c_2x},$$

where the constants are

$$S = \frac{c_2c_3}{v} + \frac{2\lambda}{v^2},$$

$$= \frac{c_2}{2}\sum_{i=1}^{n}\lambda_i E(Y_i^2) + \frac{c_3}{v} + \frac{c_1c_3+c_2c_4}{v} + \frac{2c_5}{v^2},$$

$$R = \frac{c_1}{2}\sum_{i=1}^{n}\lambda_i E(Y_i^2) + \frac{c_4(1+c_1)}{v} + \frac{2c_6}{v^2},$$

$$c_1 = \sum_{i=1}^{n}\lambda_i E(Y_i) - \frac{2}{v}\left(\sum_{i=1}^{j}\lambda_i l_i - \sum_{i=j+1}^{n}\lambda_i l_i\right),$$

$$c_2 = \frac{2}{v}\left(\sum_{i=1}^{j}\lambda_i - \sum_{i=j+1}^{n}\lambda_i\right),$$

$$c_3 = \sum_{i=1}^{j}\lambda_i E(Y_i) - \sum_{i=j+1}^{n}\lambda_i E(Y_i),$$

$$c_4 = \sum_{i=j+1}^{n}\lambda_i E(Y_i)\, l_i - \sum_{i=1}^{j}\lambda_i E(Y_i)\, l_i,$$

$$c_5 = -2\sum_{i=1}^{n}\lambda_i l_i, \text{ and,}$$

$$c_6 = \sum_{i=1}^{n}\lambda_i l_i^2.$$

The optimal solution is then given by

$$x^* = \frac{S(1-c_1) \pm \sqrt{S^2(1-c_1)^2 + Sc_2[T(1-c_1) + Rc_2]}}{Sc_2}. \quad (7)$$

A formal statement of the algorithm follows.

1-FPQM Chain Algorithm

Step 0] If a chain C has only a single node, stop.

Step 1] For the previous three subproblems, find optimal x_i^*, $i = 1, 2, 3$. Then reindex these as $x^*_{[1]} \le x^*_{[2]} \le x^*_{[3]}$.

If $\nabla AW(x^*_{[2]}) = 0$, then stop and $x^*_{[2]}$ is the optimum. Otherwise, let $a_1 = x^*_{[1]}$, $b_1 = x^*_{[2]}$ if $\nabla AW(x^*_{[2]}) > 0$; let $a_1 = x^*_{[2]}$, $b_1 = x^*_{[3]}$ if $\nabla AW(x^*_{[2]}) < 0$.

Step 2] If $\nabla AW(a_1) = 0$ or $\nabla AW(b_1) = 0$, then stop. Otherwise, let $k=1$ and go to Step 3.

Step 3] If a_k, b_k ε a link (p, q), find x^* from (7). If $x^* \varepsilon [a_k, b_k]$, then stop. Otherwise, choose x^* such that $AW(x^*) = \min\{AW(a_k), AW(b_k)\}$. If $a_k, b_k \notin$ a link (p, q), then go to step 4.

Step 4] Find the midpoint $\beta_k = (a_k + b_k)/2$ and determine $\nabla AW(\beta_k)$. If $\nabla AW(\beta_k) = 0$, then stop with $x^* = \beta_k$. Else, go to Step 5 if $\nabla AW(\beta_k) > 0$; go to Step 6 if $\nabla AW(\beta_k) < 0$.

Step 5] Let $a_{k+1} = a_k$ and $b_{k+1} = \beta_k$, replace k by k+1, and repeat Step 3.

Step 6] Let $a_{k+1} = \beta_k$ and $b_{k+1} = b_k$, replace k by k+1, and repeat Step 3.

This localization and solution procedure proved to be computationally several times faster than bisection alone.

Supposing that an optimal server location of AW(x) in the GBD subproblem, $\hat{x}^*$, is known, the objective becomes to find the optimal rates, $\tilde{\gamma}^*$, which minimize the mean weighted waiting time in the queue, $AWJ(\tilde{\gamma}, \hat{x}^*)$. By using vector notation, the GBD master program with jockeying at a server location x is given by :

$$\text{Min } AWJ(\tilde{\gamma}, \hat{x}^*) = \frac{\tilde{\gamma}^T A B^T \tilde{\gamma}}{2(1 - B^T \tilde{\gamma})} + C^T \tilde{\gamma} \quad (8)$$

subject to $\tilde{\gamma} \ge 0$

$D_i \tilde{\gamma} \ge 0$ for $i = 1, 2, \cdots, n$

where the vectors A, B, C, and D are well-defined.

An advantage of the GBD technique is that the GBD subproblem and the master program may be solved by any appropriate algorithm. There are many algorithms to solve the continuous nonlinear master problem in (8), containing extremely many variables. Gradient methods, implemented with Newton's method and sparse matrix techniques are useful.

In the convex subproblem we assume that the pseudo-customer rates are known. Their values are provided by an initial guess at the first iteration and by the solution to the master problem at subsequent iterations. However, we suggest to choose the zero jockeyed rates as a starting value which satisfies the nonnegativity. Then the iteration procedure to find a FPQMJ is summarized in the following.

1-FPQMJ Chain Algorithm

Step 0] Select an initial value of pseudo-customer rates, that is, $\gamma_{ij}=0$.

Step 1] Solve the GBD subproblem of (6) obtaining $\hat{x}^*$, by using the 1-FPQM Chain Algorithm.

Step 2] With a constant value, $\hat{x}^*$, set up and solve the master program (8) with a set of variables, $\tilde{\gamma}$, by using the appropriate nonlinear programming method.

Step 3] Find a new arrival rate at each node λ_i', by using Eq. (1), and then repeat Step 1 until no improvement in the objective value is obtained. The optimal objective value is $AWJ(\tilde{\gamma}^*, x^*)$ where, $\tilde{\gamma}^*$ is the set of optimal pseudo-rates and x^* is the optimal server location.

5. A Single FPQMJ on a Tree Graph

In the previous section, we developed a solution procedure by using the GBD algorithm for a single FPQMJ on a chain graph. We now extend these results to a tree graph, $T(N, L)$.

As previously mentioned, the GBD algorithm is based on an iterative interaction of master and subproblems. The problem in (5) is decomposed into two parts : the determination of pseudo-customer rates (master problem), and the determination of an optimal server location on a tree (subproblem). The only difference with Section 4 is the domain of subproblem which is extended to a tree graph. Since the 1-FPQMJ tree algorithm will be the same as the chain case developed in the previous section, in this section we will only show how to find an optimal server location x* on a tree, which is the GBD subproblem.

Again, consider subproblems (P1), (P2), and (P3) where the underlying graph is now a tree T(N, L) :

(P1) $\min_{x \in T} \sum_{i=1}^{n} \lambda_i' d(i, x)$,

(P2) $\min_{x \in T} \sum_{i=1}^{n} \lambda_i' E(Y_i) d(i, x)$, and

(P3) $\min_{x \in T} \sum_{i=1}^{n} \lambda_i' d^2(i, x)$.

Subproblems (P1) and (P2) are again readily solvable by Goldman's algorithm. The results of Shier and Dearing [12] are useful in solving subproblem (P3) and the following theorem follows directly from their results for general networks.

Theorem 5 : Let t denote ay node of T(N, L) and let $T_k(N_k, L_k)$ be a subtree of T obtained by disconnecting T at t. Treat t as an end node of T_k. x_3^* lies on T_k if and only if $(\sum_{i \in N_k} \lambda_i d(i, t)) / \sum_{i \notin N_k} \lambda_i d(i, t)$

Now, utilizing Theorem 5 and a Goldman-like procedure, one can easily isolate the location of x_3^* to a chain on even a single link. The optimal location, x_3^*, then follows from Theorem 4. For example, in the case of a single link (p, q), x_3^* is located at a distance $(\sum_{i \in N} \lambda_i d(i, p)) / \sum_{i \in N} \lambda_i$ from node p.

Now consider the problem $\min_{x \in T} AW(x)$. Recall from Section 3 that AW(x) is convex on a chain graph and since there is a unique chain between any two points on a tree T, it follows that AW(x) is also convex on a tree T. The following theorem is useful in localizing x* to a chain graph which is a subset of T.

Theorem 6 : Let x_i^*, i=1, 2, 3 be the optimal locations on T corresponding to subproblems (P1), (P2), and (P3). If one of the x_i^* lies on a chain C connecting the remaining two, then x* also lies on C. Otherwise, let node k be the node at the intersection of the chains joining the x_i^*, i=1, 2, 3. Then there exists exactly one improving direction of AW(x) at node k. Say that this improving direction is toward x_j^*. Then x* lies on the chain joining node k and x_j^*.

Finally, we present an algorithm for determining x* that solves $\min_{x \in T} AW(x)$. This algorithm utilizes the localization properties of Theorem 6 and the solutions of subproblems (P1), (P2), and (P3) which can all be solved quickly by Goldman-type algorithms.

1-FROM Tree Algorithm

Step 1] Solve subproblems (P1), (P2), and (P3) to find x_i^*, i=1, 2, 3.

Step 2] If x_i^*, i=1, 2, 3 lie on a unique chain, then find x* by utilizing the Chain Algorithm. Otherwise, compute the directional derivatives of AW(x) at the intersection node k in the directions toward x_i^*, i=1, 2, 3.

Step 3] Let the improving direction be in the direction toward x_j^*. Use the Chain Algorithm to find x* on the chain from node k to x_j^*.

6. A Single FPQMJ on a General Network

While there exists a unique shortest path between any two points on a tree and the distance function d(x, y) is convex on a tree, these properties do not necessarily hold on a general network. To attempt to overcome this difficulty, one tries to find some interval on a general network in which the objective function is convex. One can then find a local optimum relative to this interval.

Regions of convexity on a graph are found by determining the set of breakpoints, or points of path indifference on each link. (See Berman, et. al. [2]). For a network with n nodes, there are at most (n-2) breakpoints on each link. For a point x on the interval between two adjacent breakpoints, the shortest path between x and each node is consistent, and thus AW(x) is convex on this interval.

Therefore, to determine a global minimum of AW(x) on a general network G, one could determine all the local minimums which exist between adjacent breakpoints and choose the best of these. Obviously, this can be a tedious process for a large network. However, Hooker[10] has recently improved this procedure by developing a lower bound technique and a path elimination technique which are both applicable to a single FPQMJ on a general network.

By utilizing the Hooker's improvement, we can also solve the GBD subproblem which obtains an optimal server location on a general network. The GBD solution procedure of

(5) on a general network is similar to the acyclic case and is therefore omitted the statement of a search procedure.

7. Conclusions

In this paper, algorithms for finding a Fixed Priority Queue Median with Jockeying (FPQMJ) on a chain and tree were developed. A FPQMJ is a minisum location on a probabilistic network in which each customer type enters the network system permitting jockeying through a specified node and nonpreemptive service policy is in effect.

References

[1] Benders, J. F. 1962. Partitioning Procedures for Solving Mixed-Variables Programming Problems. Numerische Mathematik, Vol. 4, pp. 238~252.

[2] Berman, O., R. C. Larson, and S. S. Chiu. 1985. Optimal Server Location on a Network Operating as an M/G/1 Queue. *Operations Research, 33*, pp. 746~771.

[3] Chiu, S. S., O. Berman, and R. C. Larson. 1985. Locating a Mobile Server Queueing Facility on a Tree Network. *Management Science, 31*, pp. 764~772.

[4] Chiu, S. S. and R. C. Larson, 1985. Locating an n-Server Facility in a Stochastic Environment. *Computers & Operations Research, 12*, pp. 509~516.

[5] Cobham, A. 1954. Priority Assignment in Waiting Line Problems. *Operations Research, 2*, pp. 70~76.

[6] Geoffrion, A. M. 1972. Generalized Benders Decomposition. *Journal of Optimization Theory and Applications, 10*, pp. 237~260.

[7] Goldman, A. J. 1971. Optimal Center Location in Simple Networks. *Transportation Science, 5*, pp. 212~221.

[8] Hakimi, S. L. 1964. Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph. *Operations Research, 12*, pp. 450~459.

[9] Hakimi, S. L. 1965. Optimum Distribution of Switching Centers in a Communication Network and Some Related Graph Theoretic Problems. *Operations Research, 13*, pp. 462~475.

〔10〕 Hooker, J. 1985. Solving Nonlinear Single Facility Network Location Problems. *Operations Research*, *34*, pp. 732~743.

〔11〕 Jung, K. H., T. M. Cavalier, and M. J. Chandra 1987. A Fixed Priority Queue Median on a Network. IMSE Working Paper #87-143, Pennsylvania State University.

〔12〕 Shier, D. R. and P. M. Dearing 1983. Optimal Locations for a Class of Nonlinear, Single-facility Location Problems on a Network. *Operations Research*, *31*, pp. 292~303.