Scientia Iranica, Vol. 9, No. 3, pp 289–296

SCIENTIA
IRANICA

Research Note

Optimal Control of Service Rates in Jackson Networks for Solving a Bicriteria Optimal Control Problem

A. Azaron[1] and F. Kianfar*

Each dynamic job shop can be represented as a network of queues, in which each service station indicates a machine or a production department. Now, assume that the service rates of these service stations can be controlled. In this paper, a new model is developed for bicriteria optimal control of service rates of all service stations in a class of Jackson networks, in which the expected value of the shortest path of the network and, also, the total operating costs of all service stations of the network, per period, are minimized. The expected value of the shortest path of such a network of queues is equal to the expected value of the time that the first product is completed. This is an important factor in production systems, because, related to design and manufacture of a new product, the first manufactured product often has the maximum flow time, in effect, which can be minimized through the model. The networks of queues analyzed in this paper, have all the specifications of Jackson networks, except for not containing $M/M/C$ queueing systems.

INTRODUCTION

One of the most important subjects in queueing theory is the network of queues, because of its applications and also due to the complexity of the subject. A network of queues contains several service stations and each customer should refer to some of them. Many problems in the area of production or services can be formulated in the form of queueing networks. Taking into account the absence of the subject of the shortest path in the research of queueing networks in the literature, the method referred to here can be useful for solving many problems in the areas of production systems, reliability modeling and computer networks.

In this paper, a model is developed for optimal control of service rates of service stations in a network of queues in the steady-state, in which the expected value of the shortest path of the network and, also, the total operating costs of all service stations of the network per period, are minimized. These networks of queues have all the specifications of Jackson networks except for not containing $M/M/C$ queueing systems. Therefore, they only contain $M/M/1$ and $M/M/\infty$ queueing systems. Practically, if it is necessary to wait for starting the service in a production department, this could be represented by a $M/M/1$ queueing system and, if enough servers exist in this production department and it is possible not to wait in queue, a $M/M/\infty$ queueing system can be considered.

It is also assumed that the demand for each product arrives at the source node according to a Poisson process and the finished products leave the system from the sink node.

If the service rates of the service stations or machines are increased, the expected value of the shortest path of the network will be reduced, but the total operating costs per period will be raised, which is undesirable. Therefore, a bicriteria problem should be solved, in which the first criterion is minimizing the expected value of the shortest path of the network and the second criterion is minimizing the total operating costs of all service stations of the network per period.

Finally, the weighted sum approach is used for obtaining the optimal values of this bicriteria optimal control problem, which is transformed into a bicriteria nonlinear programming after discretization.

The length of a path in each network of queues is the sum of the lengths of the nodes of the network, in which the length of each node is equal to the waiting time in the system. Although no papers

1. *Department of Industrial Engineering, Bu-Ali-Sina University, Hamadan, I.R. Iran.*

*. *Corresponding Author, Department of Industrial Engineering, Sharif University of Technology, Tehran, I.R. Iran.*

could be found corresponding to the shortest path in networks of queues, there are several papers available concerning the distribution function of the shortest path in stochastic networks. Martin [1] found the distribution function of the shortest path and, also, the expected value of the shortest path in stochastic networks, in which the arc lengths are independent random variables with polynomial distribution functions, in the form of multiple integrals. Frank [2] computed the probability that the time of the shortest path in the stochastic network is smaller than a specific value. He assumed that the arc lengths are continuous random variables. Mirchandani [3] presented another method for finding the distribution function of the shortest path in stochastic networks. It is not required to solve multiple integrals in this paper, but this method can only be used in special cases where the arc lengths are discrete random variables. Kulkarni [4] presented an algorithm for finding the distribution function of the shortest path in directed stochastic networks, in which the arc lengths are independent random variables with exponential distributions, based on continuous time Markov processes. The framework of this paper is used for finding the expected value of the shortest path of the network of queues, after transforming it to an equivalent stochastic network.

There are several papers corresponding to the control of the parameters of the networks of queues. Jordan and Ku [5] considered admission policies to two multiserver loss queues in series with two types of traffic. The first type requires service at the first queue and enters the second queue with a positive probability. The second type requires service at only the second queue. They showed that, under appropriate conditions, the optimal admission policy that maximizes the expected total discounted reward over an infinite horizon, is given by a switching curve. Schechner and Yao [6] considered the control of the service rate at each node of a closed Jackson network. They assumed that for each node, there is a holding and an operating cost. It was also assumed that both costs are arbitrary functions of the number of jobs at the node. The objective is to minimize the time-average expected total costs. They showed that an optimal control, characterized by a set of thresholds (one for each node), exists such that it is optimal for each node to serve at zero rate, if the number of jobs is below the threshold and to serve at maximum allowed rate, when the number of jobs exceeds the threshold. Lazar [7] considered the special case of the following model: Two node cyclic networks with the service rate of one node being controllable for the objective of minimizing throughput. Lazar showed that there exists one optimal control of the threshold type. Tseng and Hsiao [8] analyzed the optimal control of arrival at a two-station network of queues for the objective of maximum system throughput under a time-delay constraint optimality system criterion. They showed that the optimality problem is formulated using dynamic programming with a convex cost function. Shioyama [9] developed an optimal control problem in a queueing network system. The system consists of the first stage with a server and the second stage with two servers. Two types of customers are first served at the first stage server and, subsequently, proceed to the queue at the server corresponding to their types in the second stage. When the first stage server completes a service, he determines the type of customer to be next served. The optimal control problem is to select the type of customer to be next served, in order to minimize the expected cost per hour. He formulated the problem as an undiscounted semi-Markov decision process.

In these papers, nothing is found regarding control of the queues' parameters for minimizing the expected value of the shortest path of the network of queues, which is an important factor in production systems. In the next section, a method is presented for transforming the network of queues to a stochastic network. Then, the framework of the bicriteria optimal control is described. Subsequently numerical example is solved, followed by the conclusion.

TRANSFORMING THE NETWORK OF QUEUES TO A STOCHASTIC NETWORK

The main steps for transforming the network of queues to a stochastic network are as follows:

Step 1: Compute the density function of the waiting time in system for each node, taking into account the relations of the queueing theory;

Step 2: Transform the network of queues to a stochastic network by transforming each node that contains a service station to a stochastic arc corresponding to the waiting time in system.

In this step, suppose that the arcs $b_1, b_2, \cdots, b_n$ end at the service station of the node k and the arcs $d_1, d_2, \cdots, d_m$ start from that node and the waiting time in system for this service station is equal to T_k. In the transformed network, node k is transformed to an arc (k', k'') with the length of T_k, in which node k' settles between the arcs b_i for $i = 1, \cdots, n$ and arc k and node k'' settles between arc k and the arcs d_j for $j = 1, \cdots, m$. Therefore, a network of queues is transformed to a stochastic network. The indicated process is the opposite of absorbing an edge in a graph (G.e) (see [10] for more details).

Now, Let $G = (V, A)$ be a directed network. Let V represent the set of nodes, A the set of arcs of the network and s, t the source node and the sink node of the network, respectively. Let $l(u, v)$ represent the

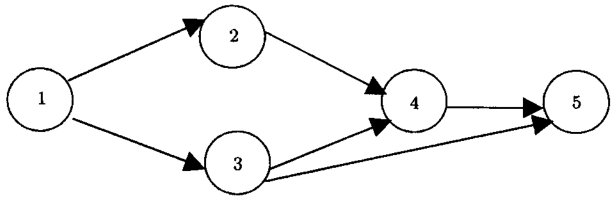

Figure 1. The network corresponding to the above example.

length of arc $(u, v) \in A$, which is an exponential random variable with the parameter $\lambda(u, v)$ (the expected length of arc (u, v) is equal to $1/\lambda(u, v)$).

For constructing the proper stochastic process, it is convenient to visualize the network as a communication network with the nodes as stations capable of receiving and transmitting messages and arcs as one-way communication links connecting pairs of nodes. As soon as a node receives a message over one of the incoming arcs, it transmits it along all the outgoing arcs and then disables itself. Now, let $X(t)$ be the set of all disable nodes at time t.

Definition 1

For describing the evolution of the stochastic process $\{X(t), t \geq 0\}$, for each $X \subset V$, in which $s \in X$ and $t \in \bar{X} = V - X$. These sets are defined as follows:

1. $\bar{X}_1 \subset \bar{X}$ is the set of nodes that do not belong to X, in which each path that connects each node of this set to the sink node t, contains at least one member of X;
2. $S(X) = X \bigcup \bar{X}_1$.

Example

In the network shown in Figure 1, if one considers $X = \{1,2\}$, then $\bar{X}_1 = \phi$ and $S(X) = \{1,2\}$. Now, consider $X = \{1,4\}$. The only path that connects node $\{2\}$ to node $\{5\}$ passes through node $\{4\}$, but for node $\{3\}$ there is a path that does not include any nodes of X. Therefore, $\bar{X}_1 = \{2\}$ and $S(X) = \{1,2,4\}$.

Definition 2

$$\Omega = \{X \subset V / s \in X, t \in \bar{X}, X = S(X)\}, \tag{1}$$

$$\Omega^* = \Omega \cup V. \tag{2}$$

In the above example, $\Omega^* = \{(1), (1,2), (1,3), (1,2,3), (1,2,4), (1,2,3,4), (1,2,3,4,5)\}$.

Definition 3

If $X \subset V$ such that $s \in X$ and $t \in \bar{X}$, then a cut is defined as:

$$C(X, \bar{X}) = \{(u, v) \in A / u \in X, v \in \bar{X}\}. \tag{3}$$

There is a unique minimal cut contained in $C(X, \bar{X})$. Denote this cut by $C(X)$. If $X \in \Omega$ then $C(X, \bar{X}) = C(X)$.

It is shown that $\{X(t), t \geq 0\}$ is a continuous time Markov process with state space Ω^* and this infinitesimal generator matrix $Q = [q(X, Y)](X, Y \in \Omega^*)$ (see [4] for more details):

$$q(X,Y) = \begin{cases} \sum\limits_{(u,v) \in C(X)} \lambda(u,v) & \text{if } Y = S(X \bigcup \{v\}) \\ -\sum\limits_{(u,v) \in C(X)} \lambda(u,v) & \text{if } Y = X \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Let T represent the length of the shortest path in the network. If 1 is considered as the initial state and N as the final state of the stochastic process $\{X(t), t \geq 0\}$, it is clear that $T = \min\{t > 0 : X(t) = N / X(0) = 1\}$. Thus, the length of the shortest path in the network is equal to the time until $\{X(t), t \geq 0\}$ gets absorbed in the final state starting from state 1. It is required to compute $F(t) = P\{T \leq t\}$ or the distribution function of the shortest path in the stochastic network. The Chapman-Kolmogorov backward or forward equations can be applied for computing $F(t)$. Using the backward algorithm, the following is defined:

$$P_i(t) = P\{X(t) = N / X(0) = i\}, \quad 1 \leq i \leq N. \tag{5}$$

Therefore, $F(t) = P_1(t)$.

The system of differential equations for vector $P(t) = [P_1(t), P_2(t), \cdots, P_N(t)]^T$ is given by:

$$P'(t) = Q.P(t),$$

$$P(0) = [0, 0, \cdots, 1]^T, \tag{6}$$

where $P(t)$ represents the state vector of the system and Q is the infinitesimal generator matrix of the stochastic process $\{X(t), t \geq 0\}$. Owing to the upper triangular nature of Q, the above differential equations can be easily solved by using an analytical or a numerical method. Therefore, if the waiting time in system for each service station is an exponential random variable, after transforming the service stations to the arcs, the above results can be utilized for finding the distribution function of the shortest path in the network of queues.

Networks of Queues with $M/M/\infty$ Queueing Systems

The simplest case for finding the shortest path is a network with all of its nodes containing $M/M/\infty$ service stations. In this case the waiting time in system, is equal to the service time with exponential distribution because there is no queue. Therefore, each node that contains a service station with service rate μ can be transformed to an exponential arc with parameter μ, as mentioned previously. Then the previous results can be applied to obtain the differential equations corresponding to the state vector of the system. After solving these equations, the distribution function of the shortest path in the network is obtained.

Networks of Queues with $M/M/1$ Queueing Systems

In this case, the density function of the waiting time, in system, in each $M/M/1$ service station is:

$$w(t) = (\mu - \lambda)e^{-(\mu-\lambda)t}, \quad t > 0, \tag{7}$$

where λ and μ represent the arrival and service rate of the service station, respectively. Therefore, the distribution of the waiting time in system, is exponential with parameter $(\mu - \lambda)$. Each node that contains a $M/M/1$ service station can, therefore, be transformed to an exponential arc with parameter $(\mu - \lambda)$ and the distribution function of the shortest path in the network found, according to the proposed method.

FRAMEWORK OF THE BICRITERIA OPTIMAL CONTROL PROBLEM

Each dynamic job shop system can be represented as a network of queues, in which each service station indicates a machine or a production department. Now, assume that the service rates of these service stations can be controlled. It is also assumed that the demand for each product arrives at the source node according to a Poisson process and the finished products leave the system from the sink node. Assume that the number of products which are produced by the system is equal to m. It is clear that each product spends a time equal to the waiting time in system, in each machine or service station. Now, assume that there are N_i production processes for producing product i (each production process corresponds to one path of the network of queues). Therefore, path i_j of the network of queues ($i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, N_i$) indicates jth production process for producing product i. For example, consider a dynamic job shop system which is represented as the network of queues shown in Figure 2. Assume that 2 products are produced by this production system. Path $1_1 = 1 - 2 - 4 - 6$ indicates the unique production process for producing product 1 and path $2_1 = 1 - 3 - 4 - 6$ and, also, path $2_2 = 1 - 3 - 5 - 6$ indicate two production processes for producing product 2.

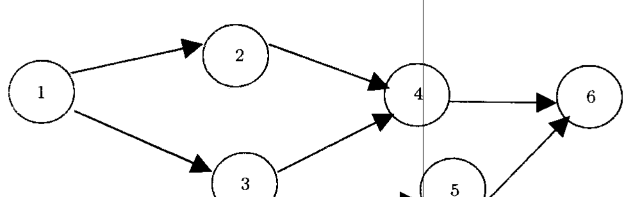

Figure 2. The dynamic job shop system.

Representing T_{ij} as the completion time of product i through production process j, C_i as the completion time of the first finished product i and T as the time of the shortest path of the network, this relation is obtained for each C_i:

$$C_i = \min\{T_{ij}\}, \quad j = 1, 2, ..., N_i. \tag{8}$$

Therefore, it is concluded that:

$$T = \min\{C_i\}, \quad i = 1, 2, \cdots, m. \tag{9}$$

Therefore, $E(T)$ or the expected value of the shortest path of such network of queues is equal to the expected value of the time that the first product is completed. $E(T)$ is an important factor in production systems because, related to design and manufacture of a new product, the first manufactured product has often the maximum flow time. Therefore, if $E(T)$ is minimized, the maximum flow time or $F_{\max}$ will be minimized, which is an important performance measure in production systems. Also, if there should be a process type of layout, in which this production system would produce other products except the new products, the first new manufactured product would have to wait for service in common facilities.

If the service rates of the service stations or machines were increased, this factor would be reduced, but C or the total operating costs per period would be raised, which is undesirable. Therefore, a bicriteria problem should be solved in which the first criterion is minimizing the expected value of the shortest path of the network and the second criterion is minimizing the total operating costs of all service stations of the network, per period. Finally, the weighted sum approach is utilized for obtaining the optimal values of this bicriteria optimal control problem, which is transformed to a bicriteria nonlinear programming after discretization.

When the service rates of the service stations are constant then, after transforming the network of queues to the corresponding stochastic network, the expected value of the shortest path of the network or $E(T)$ can be obtained from the following equation:

$$E(T) = \int_0^\infty (1 - P_1(t))dt, \tag{10}$$

where $P_1(t)$ is computed from Equation 6, as described before.

Now, assume that the service rates of the service stations of the network can be controlled and the operating cost of each service station is a linear function of its service rate. Let b_1 represent the goal for the expected value of the shortest path of the network and b_2 represent the goal for the total operating costs of all service stations of the network per period. E_1 and E_2 are free variables, which represent the deviations from the first and the second goals. Let w_1 and w_2 represent the weights corresponding to the deviations from the first and the second goals. Let $A(i)$ represent the set of adjacent nodes of node i. Let r_{ij} represent the probability that the product, whose service was finished in the service station settled in node i, goes to the service station settled in node $j \in A(i)$. Let λ_i represent the rate of arrival process to the service station settled in node i, μ_i represent the service rate of the service station settled in node i and c_i represent the cost of increasing the service rate by one unit. The values of μ_i are the controllers of the system and the values of c_i are the coefficients of μ_i in the second criterion. The other assumptions are as follows:

1. The total number of $M/M/1$ and $M/M/\infty$ service stations settled in the network of queues is equal to n;
2. The demands for all products arrive at the source node according to the Poisson process with the rate λ, in which r_i percent of this rate $i = 1, 2, \cdots, m$ would be the rate of demand for product i;
3. The values of r_{ij} for all i and $j \in A(i)$ are independent from the state of the system and are determined by the manager of the production system, taking into account the limitations of the system and the demands of the products. It should be noted that each λ_i is computed from the following equation:

$$\lambda_i = \sum_{j=1}^{n} r_{ji}\lambda_j, \quad i = 2, 3, \cdots, n. \tag{11}$$

Taking into account the above assumptions, the infinitesimal generator matrix Q is not constant and is a function of the control vector $\mu = [\mu_1, \mu_2, \cdots, \mu_n]^T$. Therefore, the system of differential equations for the vector $P(t) = [P_1(t), P_2(t), \cdots, P_N(t)]^T$ is given by:

$$\begin{aligned} &P'(t) = Q(\mu).P(t), \\ &P_i(0) = 0 \quad i = 1, 2, \cdots, N-1, \\ &P_N(t) = 1. \end{aligned} \tag{12}$$

Representing A as the set of nodes which contain $M/M/1$ service stations and B as the set of nodes which contain $M/M/\infty$ service stations, the following relations should be satisfied to exist the response in the steady-state:

$$\begin{aligned} &\mu_i > \lambda_i \quad i \in A, \\ &\mu_i > 0 \quad i \in B. \end{aligned} \tag{13}$$

There are no such constraints in mathematical programming. Therefore the following constraints are used instead of ones in the final optimal control problem:

$$\begin{aligned} &\mu_i \geq \lambda_i + \varepsilon \quad i \in A, \\ &\mu_i \geq \varepsilon \quad i \in B. \end{aligned} \tag{14}$$

Taking into account the above notations and assumptions, the appropriate model to satisfy the indicated goals is:

$$\min Z = w_1 E_1 + w_2 E_2$$

s.t.:

$$\begin{aligned} &\int_0^{\infty} (1 - P_1(t))dt - E_1 = b_1, \\ &\sum_{i=1}^{n} c_i\mu_i - E_2 = b_2, \\ &P'(t) = Q(\mu).P(t), \\ &P_i(0) = 0 \quad i = 1, 2, \cdots, N-1, \\ &P_N(t) = 1, \\ &\mu_i \geq \lambda_i + \varepsilon \quad i \in A, \\ &\mu_i \geq \varepsilon \quad i \in B, \\ &P(t) \geq 0, \\ &E_1, E_2 \quad \text{free variables.} \end{aligned} \tag{15}$$

This optimal control problem can be made discrete and transformed into a nonlinear programming. For this purpose, the differential equations should be transformed into the equivalent difference equations and, also, the integral term should be transformed into the equivalent summation term. Therefore, the continuous-time system $P'(t) = Q(\mu).P(t)$ is transformed into the following discrete-time system by dividing the time interval into K equal portions with length of Δt (see [11] for more details). If Δt is sufficiently small, it can be assumed that $P(t)$ varies only in times $0, \Delta t, \cdots, (K-1)\Delta t$. Therefore, if $P(k\Delta t)$ or kth value of P is considered as $P(k)$, the related discrete-time system would be:

$$P(k+1) = P(k) + Q(\mu).P(k)\Delta t, \quad k = 0, 1, \cdots, K-1. \tag{16}$$

If this procedure is continued, the constraint corresponding to the first goal is transformed into the following constraint:

$$\sum_{k=0}^{K} (1 - P_1(k))\Delta t - E_1 = b_1. \tag{17}$$

Finally, the optimal value of the control vector $\mu =$

$[\mu_1, \mu_2, \cdots, \mu_n]^T$ is obtained through solving the following nonlinear programming:

$$\min Z = w_1 E_1 + w_2 E_2$$

s.t.:

$$\sum_{k=0}^{K} (1 - P_1(k))\Delta t - E_1 = b_1,$$

$$\sum_{i=1}^{n} c_i \mu_i - E_2 = b_2,$$

$$P(k+1) = P(k) + Q(\mu).P(k)\Delta t, \quad k = 0, 1, \cdots, K-1,$$

$$P_i(0) = 0, \quad i = 1, 2, \cdots, N-1,$$

$$P_N(k) = 1, \quad k = 0, 1, \cdots, K,$$

$$\mu_i \geq \lambda_i + \varepsilon, \quad i \in A,$$

$$\mu_i \geq \varepsilon, \quad i \in B,$$

$$P_i(k) \geq 0, \quad i = 1, 2, \cdots, N-1, \quad k = 0, 1, \cdots, K,$$

$$E_1, E_2 \quad \text{free variables.} \tag{18}$$

It is understood that each $P_i(t)$, for $i = 1, 2, \cdots, N$, in the continuous-time system is a distribution function. Therefore, the following relation should be satisfied:

$$lim \; P_i(t) = 1, \quad i = 1, 2, \cdots, N, \quad t \to \infty. \tag{19}$$

Consequently, each $P_i(k)$, for $i = 1, 2, \cdots, N$, in the discrete-time system should also posses this property. This means that the following relation for $i = 1, 2, \cdots, N$ should also be satisfied:

$$P_i(K\Delta t) \cong 1. \tag{20}$$

It is clear that when $K\Delta t \to \infty$, Relation 20 is satisfied. On the other hand, Δt should be sufficiently small, because of the accuracy of transforming the differential equations to the equivalent difference equations. Therefore, for increasing the accuracy of the discrete-time model, K should be large.

If the following constraints are combined with the other constraints, nonlinear programming (Relations 18) would have $K(N-1) + n + 2$ constraints and $K(N-1) + 2n + 2$ variables,

$$P_i(0) = 0, \quad i = 1, 2, \cdots, N-1,$$

$$P_N(k) = 1, \quad k = 0, 1, \cdots, K. \tag{21}$$

NUMERICAL EXAMPLE

Consider the network of queues shown in Figure 3 and the characteristics of the service stations in Table 1. There are no service stations in node 5. The other assumptions are as follows:

1. This production system produces only one product and the demand for this product reaches to the source node, according to Poisson process with the rate $\lambda = 2$;
2. Each product whose service is finished in the service station settled in node 1 goes to one of the service stations settled in node 2 or 3 with equal probabilities. This means that $r_{12} = r_{13} = 0.5$. Therefore, taking into account the Relation 11, the rate of arrival process to each service station can be easily computed.

Now, each node that contains a service station is transformed into one arc whose length is equal to the waiting time in system, for this service station. Therefore, the network of queues is transformed into a stochastic network shown in Figure 4.

In the above network, arc 1 indicates the waiting time in system, in $M/M/\infty$ service station settled in node 1 of the network of queues, which has exponential distribution with parameter μ_1. Arc 2 indicates the waiting time in system, in $M/M/1$ service station settled in node 2 of the network of queues, which has exponential distribution with parameter $(\mu_2 - 1)$. Arc 3 indicates the waiting time in system, in $M/M/1$ service station settled in node 3 of the network of queues which has exponential distribution with parameter $(\mu_3 - 1)$. Arc 4 indicates the waiting time in system, in $M/M/1$ service station

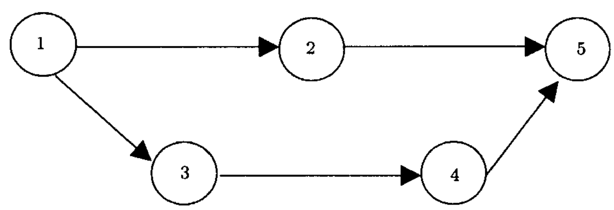

Figure 3. The network of queues corresponding to the numerical example.

Table 1. Characteristics of the service stations.

Node	Type of Service Station	Service Rate
1	$M/M/\infty$	μ_1
2	$M/M/1$	μ_2
3	$M/M/1$	μ_3
4	$M/M/1$	μ_4

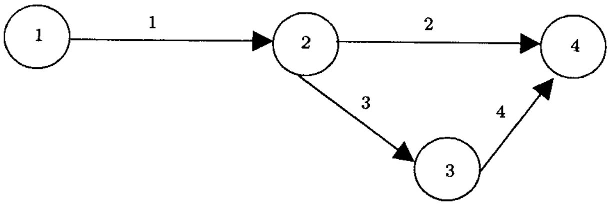

Figure 4. The stochastic network corresponding to the numerical example.

settled in node 4 of the network of queues, which has exponential distribution with parameter $(\mu_4 - 1)$.

The indicated stochastic process $\{X(t), t \geq 0\}$ has 4 states in the order of $\Omega^* = \{(1), (1,2), (1,2,3), (1,2,3,4)\}$. Table 2 shows the infinitesimal generator matrix $Q(\mu)$.

The appropriate nonlinear model to find the optimal value of the control vector $\mu = [\mu_1, \mu_2, \mu_3, \mu_4]^T$, after combining the constraints $P_4(k) = 1$ for $k = 0, 1, \cdots, K$ with the other constraints, would be:

$$\min Z = w_1 E_1 + w_2 E_2$$

s.t.:

$$\sum_{k=0}^{K}(1 - P_1(k))\Delta t - E_1 = b_1,$$

$$\sum_{i=1}^{4} c_i \mu_i - E_2 = b_2,$$

$$P_1(k+1) = P_1(k) - \mu_1 P_1(k)\Delta t + \mu_1 P_2(k)\Delta t,$$

$$k = 0, 1, \cdots, K-1,$$

$$\begin{aligned} P_2(k+1) = &P_2(k) - \mu_2 P_2(k)\Delta t - \mu_3 P_2(k)\Delta t \\ &+ 2P_2(k)\Delta t + \mu_3 P_3(k)\Delta t - P_3(k)\Delta t \\ &+ \mu_2 \Delta t - \Delta t, \end{aligned}$$

$$k = 0, 1, \cdots, K-1,$$

$$\begin{aligned} P_3(k+1) = &P_3(k) - \mu_2 P_3(k)\Delta t - \mu_4 P_3(k)\Delta t \\ &+ 2P_3(k)\Delta t + \mu_2 \Delta t + \mu_4 \Delta t - 2\Delta t, \end{aligned}$$

$$k = 0, 1, \cdots, K-1,$$

$$P_i(0) = 0 \quad i = 1, 2, 3,$$

$$\mu_1 \geq \varepsilon,$$

$$\mu_i \geq 1 + \varepsilon, \quad i = 2, 3, 4,$$

$$P_i(k) \geq 0 \quad i = 1, 2, 3, \quad k = 0, 1, \cdots, K,$$

$$E_1, E_2 \quad \text{free variables} \tag{22}$$

The values of c_i are:

$$c_1 = 20, \; c_2 = 12, \; c_3 = 10, \; c_4 = 15, \tag{23}$$

It is also assumed that b_1 or the determined goal for the expected value of the shortest path of the network is equal to 4 and b_2, or the determined goal for the total operating costs of all service stations of the network per hour, is equal to 20. The values of other parameters are:

$$w_1 = 0.5, \; w_2 = 0.5,$$

$$K = 8, \; \Delta t = 3, \; \varepsilon = 0.05. \tag{24}$$

GAMS is used to solve Nonlinear Programming 22. Table 3 shows the optimal values of $P_i(k)$ for $i = 1, 2, 3$ and $k = 0, 1, \cdots, 8$.

Table 4 shows the optimal values of μ_i for $i = 1, 2, 3, 4$ and also the optimal values of E_1 and E_2.

The optimal values of $E(T)$, or the expected value of the shortest path of the network of queues and C, or the total operating costs of all service stations of the network of queues per hour, would be:

$$E(T) = E_1 + b_1 = 8.01106,$$

$$C = E_2 + b_2 = 46.1495. \tag{25}$$

CONCLUSION

In this paper, a model is developed for optimal control of service rates of the service stations, in which the expected value of the shortest path of the network and,

Table 2. Matrix $Q(\mu)$ corresponding to the numerical example.

State	1	2	3	4
1	$-\mu_1$	μ_1	0	0
2	0	$-\mu_2 - \mu_3 + 2$	$\mu_3 - 1$	$\mu_2 - 1$
3	0	0	$-\mu_2 - \mu_4 + 2$	$\mu_2 + \mu_4 - 2$
4	0	0	0	0

Table 3. The optimal values of $P_i(k)$ for $i = 1, 2, 3$ and $k = 0, 1, \cdots, 8$.

k	0	1	2	3	4	5	6	7	8
$P_1(k)$	0	0	0.55232	0.85073	0.95065	0.98370	0.99462	0.99822	0.99941
$P_2(k)$	0	0.82463	0.99786	0.99991	0.99999	0.99999	1	1	1
$P_3(k)$	0	0.97463	0.99936	0.99998	0.99999	0.99999	1	1	1

Table 4. The optimal values of μ_i for $i = 1, 2, 3, 4$ and E_i for $i = 1, 2$.

μ_1	μ_2	μ_3	μ_4	E_1	E_2
0.22326	1.27488	1.06358	1.05	4.01106	26.1495

also, the total operating costs of the service stations of the network, per period, are minimized.

This model can be used for solving several problems in the fields of transportation networks, computer networks and production systems. In production systems, the expected value of the shortest path of the network of queues would be equal to the expected value of the time that the first product is completed. This factor is important in production systems, because, related to design and manufacture of a new product, the first manufactured product has often the maximum flow time, which can be minimized through the model-presented here.

This optimal control model was transformed into a nonlinear programming and, finally, the optimal values of the system controllers were obtained by a multi-objective optimization technique.

The limitation of this model is that the number of constraints of Nonlinear Programming 18 can grow exponentially with the network size. In the worst case scenario, for a complete transformed directed network with l nodes and $l(l-1)$ arcs, the size of the state space would be $2^{l-2}+1$ and, consequently, the number of constraints of the mentioned model would be $K(2^{l-2})+n+2$. Therefore, the number of constraints grows exponentially with l.

The model can be extended in the following directions:

1. It could be considered that the values of r_{ij} are not constant values, rather they could be system controllers, like the service rates and they could be optimally controlled. In this case, the distribution of the demand for each product could be optimally controlled in the related production processes.
2. The arc lengths among the service stations could be considered as independent random variables with exponential distributions. These lengths would be the transportation times among the departments or machines.
3. The model could be extended to the networks of queues with non-Markovian queueing systems, such as $M/G/1, G/M/1$ and $M/G/\infty$.

REFERENCES

1. Martin, J. "Distribution of time through a directed acyclic network", *Operations Res.*, **13**, pp 46-66 (1965).
2. Frank, H. "Shortest paths in probabilistic graphs", *Operations Res.*, **17**, pp 583-599 (1969).
3. Mirchandani, P. "Shortest distance and reliability of probabilistic networks", *Comput. Operations Res.*, **3**, pp 347-355 (1976).
4. Kulkarni, V. "Shortest paths in networks with exponentially distributed arc lengths", *Networks*, **16**, pp 255-274 (1986).
5. Jordan, S. and Ku, C. "Access control to two multiserver loss queue in series", *IEEE Trans. on Automatic Control*, **42**, pp 1017-1023 (1997).
6. Schechner, Z. and Yao, D. "Decentralized control of service rates in a closed Jackson network", *IEEE Trans. on Automatic Control*, **34**, pp 236-240 (1989).
7. Lazar, A. "Optimal flow control of queueing networks in equilibrium", *IEEE Trans. on Automatic Control*, **28**, pp 1001-1007 (1983).
8. Tseng, K. and Hsiao, M. "Optimal control of arrivals to token ring networks with exhaustive service discipline", *Operations Res.*, **43**, pp 89-101 (1995).
9. Shioyama, T. "Optimal control of a queueing network system with two types of customers", *European Journal of Operational Res.*, **52**, pp 367-372 (1991).
10. Bondy, J. and Murty, U., *Graph Theory with Applications*, Elsevier (1976).
11. Kirk, D., *Optimal Control Theory: An Introduction*, Prentice-Hall (1970).