Journal of the Korean
OR/MS Society
Vol. 14, No. 2, Dec., 1989

Reducing Congestion in General Queuing Networks†

Sung-Chul Kim*

일반 대기 네트웍에 있어서 시스템 설계에 관한 연구

김 성 철*

Abstract

We develope an algorithm to determine the optimal loading policy, which minimizes the congestion in general queueing networks with variable stations. Under more specified condition, majorization and arrangement ordering are established to compare, respectively, various loading and assignment policies. Implications of results are also discussed.

1. Introduction

Consider an open network of queues where the arrival processes need not be Poisson and the service-time distributions need not be exponential(general queueing network). There are M single-server stations in the network with the infinite capacity waiting rooms and the first-come-first-served queueing disciplines. The service times at station i(i=1,…,M) are independent and identically distributed with fixed service rate μ_i and finite squared coefficient of variation(SCV) Cs_i^2. Jobs arrive at a station following a renewal process, which is independent and identically distributed. For station i(i=1,…,M), let $\lambda_i(>0)$ be the effective arrival rate(which can be determined from the traffic equations) and Ca_i^2 be the SCV of the interarrival times. Then, $\rho_i=\lambda_i/\mu_i$ can be interpreted as the service intensity at station i, and we shall refer to ρ_i as the "loading" of station i, and $\underline{\rho}=(\rho_1,\cdots,\rho_M)$ as the loading vector. To ensure steady-state, we require $\rho_i<1$ for all i.

Let $\nu_i=\lambda_i/\lambda$, where λ is the total exter-

* Department of Business Administration, Duk Sung Women's University

† 이 연구는 1989년도 한국과학재단 연구비 지원에 의한 결과임
과제번호 : 891-0915-009-2

nal arrival rate, then ν_i can be interpreted as the expected number of visits to station i for each job and we shall refer to ν_i as the visit frequency to station i. Also let ω_i(respectively, T_i) be the mean waiting time before service(respectively, including service) that an arbitrary job spends at station i, and T be the mean sojourn time in the network. Then, using approximations based on the decomposition approach [Shanthikumar & Buzacott(1981), Whitt(1983), Bitran & Tirupati(1988)](where the nodes are analyzed as being stochastically independent), $T_i=\nu_i(1/\mu_i+\omega_i)$ and $T=T_1+\cdots+T_M$. Also let N_i be the mean queue length at station i including both in queue and in service, and N be the mean number of jobs in the network. Then, based on the familiar formular, $N_i=\lambda_i(1/\mu_i+\omega_i)$ [Heyman & Sobel(1982), Frankel & et. al(1981)], $N_i=\rho_i+\lambda_i\omega_i$ and $N=N_1+\cdots+N_M$.

Suppose, for any $\underline{\rho}$, we have $L=|\underline{\rho}|=\rho_1+\cdots+\rho_M$, then L can be interpreted as the average total work requirement of jobs within the network. Given a fixed L, the loading vector basically represents a distribution of L among the stations. For Jackson networks(with the exponential service times and the poisson arrival rates), balanced loading minimizes the mean number of jobs in the network(N)[Shanthikumar(1982), Shanthikumar & Stecke(1986)]. Yao(1985, 1987) extended the above conclusion by establishing a partial ordering, majorization, within the set of loading vectors with a total work requirement not exceeding L.

In this paper we study the above problem for the more general case of non-exponential service times and non-poisson arrival rates under the highly variable condition, i.e. $Ca_i^2\geq 1(i=1,\cdots,M)$. However, in the context of manufacturing systems, the upstream production systems could be quite unstable, diversified jobs are processed simultaneously in the system and machine breakdowns and operator unavailability may occur. Also, notice the relationship, $Cd_i^2=\rho_i^2\ Cs_i^2+(1-\rho_i^2)Ca_i^2$ for all i [Shanthikumar & Buzacott(1981), Whitt(1983)], where Cd_i^2 is the SCV of the departure process at station i which is in turn the arrival process to the other stations. If these factors are aggregated, it is very likely to have highly variable stations. And the objective is to achieve the minimum sojourn time in the network at a given overall loading, i.e., all the operation mix of job types, and determine which station will be able to perform each operation of job types within the associated technical constraints. Then by Little Law, the mean number of jobs in the network, N, is also minimized.

Recently, there arise considerable interests in the design of the general queueing network. One of such problem is the capacity allocation problem, which is closely related to the loading problem. A given total flexible capacity is to be allocated among M service stations such that the service capacity at station i, $\mu_i(i=1,\cdots,M)$, is at least greater than λ_i(equivalently $\rho_i<1$) and the mean queue length in the network is minimized. Bitran and Tirupati(1989) developed a first derivative algorithm based on the parametric decomposition approach. In Wein(1989), the capacity allocation is achieved in proportion to the square root of variability using a Brownian approximation of the general queueing network[Harrison & Williams(1987)].

In section 2, the decomposition approximation is briefly outlined, which will be the building block for the analysis. In section 3, we start with problem formulation and present the solution algorithm. Further results are also discussed concerning the properties of the optimal solution. A numerical example is illustrated in section 4, and section 5 concludes this paper.

2. The Decomposition Approximation

For queueing networks, exact results are available for only a limited class of product-form networks[Jackson(1957, 1963), Gorden & Newell(1967), Baskett et al.(1975), and Kelly(1979)]. To evaluate the general queueing networks without product-form solutions, the decomposition approach was reported to provide good estimates of the performance measures. In the decomposition approach, each station is decomposed into individual queues and each individual queue is analyzed in isolation.

Brief explanation for decomposition approximation is as follows. The first step of this approach is characterized by:

(i) the fundamental traffic equations which evaluate the effective arrival rates,

$$\lambda_i = a_i + \Sigma_{j=1}^{M} \lambda_j \gamma_{ji},\ i=1,\cdots,M \qquad \cdots (2.1)$$

and (ii) a set of following equations which compute the SCV's of the interarrival times approximately[Bitran & Tirupati(1988)]:

$$\lambda_i Ca_i^2 - \Sigma_{j=1}^{M} \{\lambda_j(1-\rho_j^2)\gamma_{ji}^2\, Ca_j^2\} = a_i Ca_i^0 + \Sigma_{j=1}^{M} \{\lambda_j \gamma_{ji}\, (\rho_j^2\, r_{ji}\, Cs_j^2 + 1 - \gamma_{ji})\},\quad i=1,\cdots,M \qquad \cdots (2.2)$$

where

γ_{ij}: the routing probability from station i to station j,

a_i: the external arrival rate to station i,

and

Ca_i^0: the SCV of the external arrival process to station i.

The second step is to approximate the performance of the decomposed nodes as GI/G/1 queues based on the first two moments, i.e. mean and SCV, of the interarrival and service time distributions computed in step 1. Among others, we use the combined approximation of Kraemer and Langenbach-Belz(1976) and Whitt(1983) for GI/G/1 queue, and the estimate for the mean waiting time at station i, T_i, is given by:

$$T_i(\rho_i, Ca_i^2, Cs_i^2)$$
$$= \nu_i\{1/\mu_i + \omega_i(\rho_i, Ca_i^2, Cs_i^2)\}$$
$$= (1/\lambda)\{\rho_i + \rho_i^2(Ca_i^2+Cs_i^2)\ h(\rho_i, Ca_i^2, Cs_i^2)/2(1-\rho_i)\} \quad \cdots\cdots (2.3)$$

where

$$h(\rho_i, Ca_i^2, Cs_i^2) = \begin{cases} \exp\{-2(1-Ca_i^2)(1-\rho_i)/3(Ca_i^2+Cs_i^2)\rho_i\}, & Ca_i^2<1 \\ 1, & Ca_i^2 \geq 1. \end{cases}$$

The last step is to synthesize the performance of the isolated stations to estimate the performance of the network. The mean sojourn time in the network is given by:

$$T = \Sigma_{i=1}^{M} T_i \quad \cdots\cdots (2.4)$$

This decomposition approach will be used as a building-block for the analysis in section 3.

3. The Loading Problem

The loading problem to be considered is to allocate the average total work requirement of jobs within the network to the M stations such that the mean sojourn time is minimized. Based on the decomposition approach(using the approximation of Whitt(1983) with $Ca_i^2 \geqq 1$ for all the single-server queues), the problem at hand is to minimize

$$T = \Sigma_{i=1}^{M} T_i = (1/\lambda)\ \Sigma_{i=1}^{M}\{\rho_i + \rho_i^2(Ca_i^2+Cs_i^2)/2(1-\rho_i)\} \quad \cdots\cdots (3.1)$$

subject to the constraints

$$\Sigma_{i=1}^{M} \rho_i = L \quad \cdots\cdots (3.2)$$

and equation(2.2) $\cdots\cdots$ (3.3)

Lemma 1: If the coefficients Ca_i^2 and Cs_i^2 $(i=1,\cdots,M)$ are assumed to be independent of ρ_i(equivalently, if constrant(3.3) is relaxed), the mean waiting time at a decomposed node i, $T_i(i=1,\cdots,M)$, is a non-decreasing convex function of ρ_i.

Proof: A simple algebraic manipulation leads to $dT_i/d\rho_i>0$ and $d^2T_i/d\rho_i^2<0$ for all i.

Remark: In GI/G/1 queue, the mean waiting time(respectively, mean number of jobs) is known to be a non-increasing convex function of the service rate(μ)[Weber(1983) (respectively, Hung-Yuan & Kumin(1983))].

The independence of the coefficients, Ca_i^2 and $Cs_i^2(i=1,\cdots,M)$, of ρ_i is generally an adopted assumption in the open literature [Kleinrock(1964), Shanthikumar & Yao(1988), Bitran & Tirupati(1989), and Wein(1989)]. If there are diversified jobs to be manufactured in the network, they are indeed incensitive to the changes of ρ_i [Bitran & Tirupati(1989)].

Since the objective is convex function and constraints form a convex set, the problem reduces to a convex problem. To solve this problem, we form the Lagrangian function

$$\mathcal{L}=(1/\lambda)\ \Sigma_{i=1}^{M}\{\rho_i+\rho_i^2(Ca_i^2+Cs_i^2)/2(1-\rho_i)\}-\pi\ (\Sigma_{i=1}^{M}\rho_i-\mathrm{L}) \quad \cdots\cdots (3.4)$$

and solve the M+1 equations,

$$d\mathcal{L}/d\rho_i=(1/\lambda)\{1+\rho_i(2-\rho_i)(Ca_i^2+Cs_i^2)/2(1-\rho_i)^2\}-\pi=0 \quad \cdots\cdots (3.5)$$

$$d\mathcal{L}/d\pi=-(\Sigma\rho_i-\mathrm{L})=0 \quad \cdots\cdots (3.6)$$

resulting in

$$\rho_i=1\pm1/\{2(\lambda\ \pi-1)/(Ca_i^2+Cs_i^2)+1\}^{1/2} \quad \cdots\cdots (3.7)$$

$$\Sigma_{i=1}^{M}\ \rho_i=\mathrm{L}, \quad \cdots\cdots (3.8)$$

where π is the Lagrangian multiplier. Since $\rho^*_i<1$, the right hand side of (3.7) should be

$$\rho^*_i=1-1/\{2(\lambda\pi-1)/\ (Ca_i^2+Cs_i^2)+1\}^{1/2} \quad \cdots\cdots (3.9)$$

and(3, 8) reduces to

$$\Sigma_{i=1}^{M}1/\{2(\lambda\ \pi-1)/(Ca_i^2+Cs_i^2)+1\}^{1/2}=\mathrm{M}-\mathrm{L} \quad \cdots\cdots (3.10)$$

To obtain the solution at this point is to determine value for π^* that satisfies equation(3.10). But, it is not difficult to see that the left hand side of (3.10) is non-increasing covex function of $\pi(\pi\geqq0)$ and this property considerably facilitates the procedure to determine the optimal π^*, and it turns to be an elementary nonlinear problem.

With exponential service times and poisson arrival rates for all stations, our assignment(3.8) reduces to Wein(1989)'s capacity assignment. However his assignment scheme requires not only the balanced heavy loading condition but also a certain skew-symmetry variability condition, and its applicability is quite limited.

Theorem 1: Suppose the M stations are numbered such that $Ca_1^2+Cs_1^2\leqq Ca_2^2+Cs_2^2\leqq\cdots\leqq Ca_M^2+Cs_M^2$, then the optimal loading has the following property; $\rho^*_1\geqq\rho^*_2\geqq\cdots\geqq\rho^*_M$.

Proof: It is a direct consequence of equation(3.9).

The above theorem can be intuitively explained as follows: T can be reduced by assigning smaller ρ_i(by assigning either smaller λ_i or larger μ_i) for more variable station.

Now, suppose there is one more condition that the variability parameter $Ca_i^2+Cs_i^2$ is the same for all stations(i=1,…,M), then we can establish the further conculsions. In the sequel, we refer this condition as the equally variable condition.

Theorem 2: Under the equally variable condition, the balanced loading(ρ^*_i=L/M for i=1,…,M), minimizes T.

Proof: In this case, let $C=Ca_i^2+Cs_i^2$ for all i. Then (3.10) reduces to

$$\mathrm{M}/\{2(\lambda\ \pi-1)/C+1\}^{1/2}=\mathrm{M}-\mathrm{L}. \quad \cdots (3.11)$$

Hence,

$$\pi^*=\mathrm{M}^2\ C/2\lambda(\mathrm{M}-\mathrm{L})^2-C/2\lambda+1/\lambda. \quad \cdots\cdots (3.12)$$

Substituting the value of π^* into equation(3.9) yields:

$$\rho^*_i = 1-(M-L)/M = L/M \quad \cdots\cdots\cdots\cdots (3.13)$$

Theorem 3: Under the equally variable condition, the mean sojourn time as a function of $\underline{\rho}$, $T(\underline{\rho})$, is a Schur-convex function(see Appendix) of $\underline{\rho}$.

Proof: The symmetry of $T(\underline{\rho})$ with respect to $\underline{\rho}$ is obvious, and it suffices to show that $\rho_1 \geqq \rho_2$ implies$(d/d\rho_1)T(\underline{\rho})-(d/d\rho_2)T(\underline{\rho}) \geqq 0$. Then, we immediately have

$$(d/d\rho_1)T(\underline{\rho})-(d/d\rho_2)T(\underline{\rho})$$
$$=(\rho_1-\rho_2)\ \{2-(\rho_1+\rho_2)\} \geqq 0. \quad \cdots\cdots\cdots (3.14)$$

Notice that, from Theorem 3, given $\Sigma_{i=1}^{M} \rho_i = L$, where L is constant, the minimum mean sojourn time is obtained under the equal elements of $\underline{\rho}$, i.e. $\rho_i = L/M$ for all i, and hence, Theorem 2 is a direct consequence of Theorem 3.

Corollary 1: If $\underline{\rho}^1 \leqq_{wm} \underline{\rho}^2$, then $T(\underline{\rho}^1) \leqq T(\underline{\rho}^2)$. Theorem 3 can be extended to arrangement increasing function.

Theorem 4: Under the equally variable condition, the mean sojourn time as a function of two assignment vector$(\underline{\lambda}, \underline{\mu})$, $T(\underline{\lambda}, \underline{\mu})$ is an arrangement increasing function(see Appendix) of$(\underline{\lambda}, \underline{\mu})$.

Proof: The same line of proof is given in Yao(1985).

As a direct consequence of Theorem 4:

Corollary 2: The assignment vector$(\underline{\lambda}\downarrow, \underline{\mu}\downarrow)$ minimizes T and $(\underline{\lambda}\downarrow, \underline{\mu}\uparrow)$ maximizes T.

In corollary 2, ↓ and ↑ denote, respectively, the decreasing and increasing arrangements of a vector's components. Therefore to reduce the congestion in the network, we should match larger μ_i to larger λ_i.

4. Numerical Examples

In this section, we illustrate the problem examined through several examples. The parameters used in the examples are summarized in Table 1, together with the optimal solutions.

Table 1. Summary of the Prameters and Optimal Solutions

M	L	λ	Ca_1	Cs_j^2	$Ca_i^2+Cs_i^2$	π^*	ρ*	T	
3	2	5	1.0	1.6	2.6	2.279994	0.666666	1.439997	balanced
			1.2	1.4	2.6		0.666666		loading
			2.5	0.1	2.6		0.666666		
5	3	8	1.0	0.8	1.8	1.097839	0.6780463	1.190946	

M	L	λ	Ca_i	Cs_j^2	$Ca_i^2+Cs_i^2$	π*	ρ*	T	
			1.2	1.0	2.2		0.6480966		
			2.0	0.8	2.8		0.6095383		
			1.5	1.5	3.0		0.5980166		
			4.0	2.2	6.2		0.4662816		
9	6	6	1.0	1.5	2.5	1.833333	0.6666666	3.499999	balanced
			1.1	1.4	2.5		0.6666666		loading
			1.2	1.3	2.5		0.6666666		
			1.4	1.1	2.5		0.6666666		
			1.5	1.0	2.5		0.6666666		
			1.6	0.9	2.5		0.6666666		
			2.0	0.5	2.5		0.6666666		
			2.4	0.1	2.5		0.6666666		
			2.5	0.0	2.5		0.6666666		
12	11	15	1.0	0.3	1.3	14.30529	0.9458566	13.06049	
			1.4	0.0	1.4		0.9438191		
			1.0	0.8	1.8		0.9363256		
			1.5	0.3	1.8		0.9363256		
			1.4	0.6	2.0		0.9328964		
			1.9	0.5	2.4		0.9265248		
			1.2	2.0	3.2		0.9152343		
			1.7	2.0	3.7		0.9089034		
			3.0	1.5	4.5		0.8996267		
			1.4	4.0	5.4		0.8901571		
			2.0	3.5	5.5		0.8891571		
			4.0	3.0	7.0		0.8751613		

5. Conclusion and Discussions

We considered the loading problem in the general queueing network and presented a solution method. The optimal loading was shown to be negatively dependent [Block, Savits, & Shaked(1982)] on the square root of variabity of station and a better loading policy was identified to be the one which assigns the more loadings to less variable stations. In the problem examined, we assumed that all the stations are highly variable and this assumption is generally acceptable in the context of a certain manufacturing systems[Yao & Buzacott(1985), Whitt(1982), Bitran & Tirupati(1989)]. However, for $Ca_i^2<1(i=1, \cdots, M)$, there is so far no obvious method to present the basic structure of the problem except the greedy-like heuristics[Bitran & Tirupati(1989)].

Under more specified condition(the equally variable condition), majorization and arrangement orderings are established which not only identify the optimal policy but also compare various loading policies as far as the majorization relation is relevant. These results can also be easily extended to the low variability condition($Ca_i^2 \leqq 1$, $i=1,\cdots,M$) and can be proved as in Theorem 3. They can also be shown the other way using the fact that the objective function is convex and symmetric[Marshall & Olkin(1979)]. However, if these conditions are voilated(i.e. some stations are more variable, $Ca_i^2 \geqq 1$, and the others are not, $Ca_i^2 \leqq 1$), the majorization and arrangement orderings do not apply simply because the objective function is not symmetric any more.

The majorization and arrangement results in Jackson queueing network with single-server stations are well known results[Yao(1985, 1987)]. Here we extended these orderings to a more general case, the queueing network with the equally variable condition where the interarrival and service time distributions need not be exponential. In case of the general Queueing Network, the results are considered to be new.

Finally, the independence of the coefficients Ca_i^2 and Cs_i^2 of ρ_i can be relaxed and the effects of the coefficients can be directly included in the optimal loading by recomputing the coefficients in an iterative manner. However, it is worthwhile to notice that it is more important to have the initial network configuration appropriately since the penalties resulting from equation(3.9) without an iterative scheme are quite small and well within the errors of approximation.

Appendix

We present here a summary of results on majorization and arrangement orderings used in this article. In the sequel, M-O denotes the reference Marshall and Olkin(1979).

1. $\underline{x}$ and $\underline{y}$ are two n-dimensional vectors(with real-valued components). Let $x_{[i]}$ and $y_{[i]}$ be their ith largest components, i.e., $x_{[1]} \geqq \cdots \geqq x_{[n]}$ and $y_{[1]} \geqq \cdots \geqq y_{[n]}$. The majzorization ordering, $\underline{x} \leqq_m \underline{y}$, is defined as

$$\Sigma_{i=1}^{K} x_{[i]} \leqq \Sigma_{i=1}^{K} y_{[i]} \quad (k=1,\cdots,n-1) \text{ and}$$
$$\Sigma_{i=1}^{n} x_{[i]} = \Sigma_{i=1}^{n} y_{[i]} \quad \cdots\cdots\cdots\cdots \text{(A.1)}$$

2. If the last equation in (A.1) changes to $\leqq$, the ordering is defined as week(sub-) majorization, denoted as $\underline{x} \leqq_{wm} \underline{y}$.(§ 1A of M-O.)

3. If $\underline{x} \leqq_{wm} \underline{y}$, then there exists $\underline{z}$ such that

$\underline{x} \leqq \underline{z}$(componentwise) and $\underline{z} \leqq_m \underline{y}$(§ 5A of M-O.)

4. A real-valued function f(·) is defined as a Schur-convex(concave) function, if $\underline{x} \leqq_m \underline{y} \rightarrow f(\underline{x}) \leqq (\geqq) f(\underline{y})$. If $f(\underline{x})$ is continuously differentiable, then it is Schur-convex(concave) if and only if for all i, j

$$x_i \leqq x_j \rightarrow (d/dx_i)f(\underline{x}) \leqq (\geqq) (d/dx_j)f(\underline{x}), \text{ (§ 3A of M-O.)} \quad \cdots\cdots \text{(A-2)}$$

5. The arrangement ordering of two assignment vectors, $(\underline{\lambda}, \underline{\mu}^2) \leqq_a (\underline{\lambda}, \underline{\mu}^1)$, is defined as follows: $\underline{\mu}^1$ can be reached $\underline{\mu}^2$ through successive pairwise interchanges of components with each interchange correcting one inversion of the decreasing order of components. Consider the following as an example: suppose $\mu_1 \geqq \mu_2 \geqq \mu_3 \geqq \cdots \geqq \mu_M$, and $\underline{\mu}^2 = (\mu_3, \mu_1, \mu_2, \cdots, \mu_M)$, $\underline{\mu}^1 = (\mu_1, \mu_3, \mu_2, \cdots, \mu_M)$.(§ 6F of M-O.)

References

[1] Baskett, F., Chandy, K.M., Muntz, R. R. and Palacios, F.G., 1975, Open, closed and mixed networks of queues with different classes of customers. Journal of Associated Computing Machinery, 22, 248-260.

[2] Bitran, G.R. and Tirupati, D., 1988, Multiproduct queueing networks with deterministic routing: Deomposition approach and the notion of interface. Management Science, 34, 75-100.

[3] Bitran, G.R. and Tirupati, D., 1989, Trade-off curves, targeting and balancing in queueing networks. Operations Researcheck, 37, 547-564.

[4] Block, H.W., Savits, T.H. and Shaked, M., 1982, Some concepts of negative dependence. Annals of Probability, 10, 765-772.

[5] Franlel, P., König, D., Arndt, U. and Schmidt, V., 1981, Queues and point processes(Berlin: Akademie-Verlag).

[6] Gorden, W.J. and Newell, G.F., 1967, Closed queueing networks with exponential servers. Operations Research, 15, 252-267.

[7] Harrison, J.M. and Williams, R., 1987, Brownian models of open queueing networks with homogenous customer populations. Stochastics, 22, 77-115.

[8] Heyman, D.P. and Sobel, M.J., 1982, Stochastic models in operations research, Vol. 1(New York: MacGraw-Hill).

[9] Jackson, J.R., 1957, Networks of waiting lines. Operations Research, 5, 518-521.

[10] Jackson, J.R., 1963, Jobshop like queueing systems. Management Science, 10, 131-142.

[11] Kelly, F., 1979, Reversibility and Stochastic Networks(New York: Wiley)

[12] Kraemer, W. and Langenbach-Belz, M., 1976, Approximate formulae for the delay in the queueing system GI/G/1. In Proceedings of 8th International Teletraffic Congress, Melbourne, 23, 1-8.

[13] Kleinrock, L., 1964, Communication Nets: Stochastic Message Flow and

Delay(New York: Dover Publications, Inc.)

[14] Marshall, A.W. and Olkin, I., 1979, Inequalities: Theory of Majorization and It's Applications(New York: Academic Press)

[15] Shanthikumar, J.G., 1982, On the superiority of balanced load in flexible manufacturing systems. Syracuse University, U.S.A.

[16] Shanthikumar, J.G. and Buzacott, J. A., 1981, Open queueing network models of dynamic job shops. Interational Journal of Production Research, 19, 255-266.

[17] Shanthikumar, J.G. and Stecke, K. E., 1986, Reducing work-in-process inventory in certain classes of flexible manufacturing systems. European Journal of Operations Research, 26, 266-271.

[18] Shanthikumar, J.G. and Yao, D.D., 1988, On server allocation in multiple center manufaturing system. Operations Research, 36, 333-342.

[19] Tu, H.Y. and Kumin, H., 1983, A convexity result for a class of GI/G/1 queueing systems. Operations Research, 31, 948-950.

[20] Weber, R.R., 1983, A note on waiting times in single server queues. Operations Research, 31, 950-951.

[21] Wein, L.M., 1989, Capacity allocation in generalized Jackson networks. Operations Research Letters, 8, 143-146.

[22] Whitt, W., 1982, Approximating a point processes by a renewal process, I: Two basic methods. Operations Research, 30, 125-147.

[23] Whitt, W., 1983, The queueing network analyzer. The Bell System Technical Journal, 62, 2779-2815.

[24] Yao, D.D., 1985, Some properties of the throughput function of closed networks of queues. Operations Research Letters, 3, 313-317.

[25] Yao, D.D, 1987, Majorization and arrangement orderings in open networks of queues. Annals of Operations Research, 9, 531-543.

[26] Yao, D.D. and Buzacott, J.A., 1985, Queueing models for a flexible machining stations, I: Diffusion approximations; II: The method of coxian phases. European Journal of Operations Research, 19, 233-248.