〈研究論文〉

Journal of the Korean Institute
of Industrial Engineers
Vol. 15, No. 1, June. 1989

A Study on Information Flow and Routing in a Class of Flexible Manufacturing Systems†

Kim, Sung Chul*

Abstract

The routing flexibility in FMSs can be achieved through two levels : the parallel operations and the alternative stations. To accommodate this pertaining flexibility for part routing in FMSs, a hierarchical routing principle is developed which considers both the routing flexibility and the performance effectiveness, which can also be easily implemented. For the proper implementation of this routing algorithm, the required information flow is identified which turns out to be very simple.

1. INTRODUCTION

In this work, we study a dynamic control of part-routing in a class of manufacturing systems. The practical system studied is a network of part-manufacturing machines which is usually termed as a flexible manufacturing system(FMS)(Groover 1980).

A FMS is essentially an automated jobshop with a set of versatile work stations(machines) and automated material handling device. Under the computer control and numerical control techniques. this system allows considerable flexibility in assigning operations along with their associated tools among machines and permits very general flow patterns of the jobs within the system. And a part may follow several different routes in terms of sequence of operations and/or machine combinations to perform these operations.

Due to the complexity of the proper implementation of FMSs. it is not likely that job routing follows a fixed pattern which is determined at the pre-production planning level independent of the operational state of the on-line system(Buzacott and Shanthikumar 1980), and if these decisions are made at the pre-production planning level. routing flexibility is simply deteriorated.

*Dept. of Business Administration, Duk Sung Women's University

†이 논문은 1988년도 문교부지원 학술진흥재단의 자유공모 과제 학술연구 조성비에 의하여 연구되었음.

To make use of this routing flexibility, it requires the detailed information about the status of the machines and jobs and the capacity of collecting, retrieving and utilizing on-line information, and thereby making adaptive decisions. Therefore, in this new manufacturing systems, the invisible information flow becomes more important in addition to the physically visible material flows.

A formal and systematic approach to dynamic part routing in FMSs is relatively few. Buzacott(1982) studied a set of dynamic routing rules which utilize simple information of system states and provide some insights into system performance. Yao(1987) developed a probabilistic shortest queue routing which routes a part with the highest probability to the station with the largest empty spaces. Extentions of the probabilistic shortest queue routing to multiple job types and closed queueing network with arbitrary configuration can be found in Yao and Buzacott(1985). Though these studies all utilize, to some extent, the flexibility pertaining to part routing, the concept of routing flexibility was first quantified in Yao(1985). A concept, routing entropy, is formulated as a measure for routing flexibility, and a routing principle depending on the on-line information of the operational status is proposed.

The purpose of this paper is to develop a simple and easily implementable dynamic part routing rule in a class of FMSs with single server stations which utilize the routing flexibility pertaining to parts routing but also accomplish the performance effectiveness, and demonstrate its applicability.

2. Sturcture of Routing Flexibility.

Yao(1985) identifies four different levels of the production task, (i)part types, (ii) sequential subsets, (iii) parallel operations, and (iv)alternative stations, and has shown that routing flexibility can be achieved through the last two levels, i.e. through parallel operations and through alternative stations. In this section, we explore the details of the underlying structure of routing flexibility.

The routing flexibility through the alternative stations comes from the capability of versatile machines which can perform many different operations and is closely related with the production planning problem of FMS. To explain, suppose we have product types, $t=1, \cdots, T$, to be manufactured and their relative product quantities, $Q(t)$, $t=1, \cdots, T$, then we have the number of each operation type, ω_j, $j=1, \cdots, J$, from the operation mix of part types. Let ν_i be the number of operations assigned to station i, $i=1, \cdots, M$, and δ_{ij}, $i=1, \cdots, M$, and $j=1, \cdots, J$, be a variable which represents the number of operation type j assigned to station i. Then, the number of type j operations, ω_j, and the number of operations assigned to station i, ν_i, can be written respectively as :

$$\omega_j=\sum_{i=1}^{M} \delta_{ij}, \quad j=1, \cdots, J,$$
$$\nu_i=\sum_{j=1}^{J}\delta_{ij}, \quad i=1, \cdots, M. \tag{2.1}$$

Let s_j be the processing time of the type j operation, $j=1, \cdots, J$, and μ_i be the service rate of machine at station i, $i=1, \cdots, M$, then the mean processing time of machine at station i, denoted by $1/\mu_i$. is :

$$1/\mu_i=\sum_{j=1}^{J}s_j\delta_{ij}/\sum_{j=1}^{J}\delta_{ij}, \quad i=1, \cdots, M. \tag{2.2}$$

The average work requirement assigned to station i, denoted by ρ_i, $i=1, \cdots, M$, is equal to the product of the mean processing time of a machine at station i and the number of operations assigned to station i, that is :

$$\rho_i=\nu_i(1/\mu_i)=\sum_{j=1}^{J}s_j\delta_{ij}, \quad i=1, \cdots, M, \tag{2.3}$$

and can be used as a measure of the relative work-load assigned to station i.

And, the average work requirement of jobs within he system, $\sum\rho_i$, is :

$$\sum_{i=1}^{M}\rho_i=\sum_{i=1}^{M}\sum_{j=1}^{J}s_j\delta_{ij}=\sum_{j=1}^{J}s_j\omega_j. \tag{2.4}$$

The different loading policies have the different average work requirement of jobs among stations. Then, how to distribute the average work requirement of jobs among stations becomes an important production planning problem in FMSs. In this case, the total number of the distinguished states of possible allocations is equal to $\prod_{j=1}^{J} \frac{M+\omega_j-1}{M-1})/M!$. For details, we refer to Stecke(1983), and Kim(1988).

A major estimation of performance can be measured through the evaluation of the Jackson queueing network(Jackson 1963, Gorden and Newell 1967) which has been extensively applied to study FMSs(Buzacott and Yao 1986). In Stecke(1983), various other objective functions are considered regarding material movement, tool magazine utilization, and operation priorities.

The alternative stations for each operation can be identified through the solutions of the above production planning problem, i.e. δ_{ij}^*'s, the required production quantities of operation type j at station i, $i=1, \cdots, M$, and $j=1, \cdots, J$. And it is important to point out that, for queueing network models of FMSs, the values of δ_{ij}^*'s can also be used directly to determine visit ratios and these can replace the routing matrix.

However, consider a route of a part type, that is, a path of successive station visits of a part type to complete all it's required operations. Then the required operations may have technological restrictions on the order in which operations can be performed, and a route consists of a set of operations which can be grouped into a series of sequence constrained subsets, $\ell(t)=1, \cdots, L(t)$. That is, any operation in subset $\ell+1(t)$ can not be initiated before all the operations in subset $\ell(t)$ are completed. And, in each subset $\ell(t)$, there exist operations $n_{\ell(t)}=1, \cdots, N_{\ell(t)}$, which are sequence independent and then can be performed in any order. These sequence-independent operations will be referred to, in the sequel, as paralled operations.

Therefore, the parallel operations and the alternative stations completely characterize the routing flexibility of a part. And it is quite clear that, while a part flows through the system, the parallel operations and the alternative stations of a part keep changing, which in turn provides a basis for characterizing its routing flexibility.

3. Hierachical Routing Principle

Based on the analysis in the previous section, any routing machanism should consist of the following two levels : (i) choose the next operation among the immediate parallel operations, (ii) choose one station among several alternative stations.

We first study the way of choosing the next operation to be performed among immediate parallel operations. The concept of routing entropy introduced by Yao(1985) can be used to measure the routing flexibility. To explain, we focus on the ℓth sequential subset of a part(to facilitate presentation, we don't differentiate the part types here), with $N_\ell(n_\ell = 1, \cdots, N_\ell)$ parallel operations in the subset, and with $M_n(\ell)(m_n(\ell)=1, \cdots, M_n(\ell))$ alternative stations for each operation n_ℓ, $n_\ell=1, \cdots, N_\ell$. Then, the total number of possible routings in this sequential subset ℓ is equal to ψ_ℓ, where $\psi_\ell = N_\ell! \cdot \prod_{n\ell=1}^{N\ell} M_n(\ell)$. The routing entropy of an operation n_ℓ, $E_n(\ell)$, the routing entropy of a part in the sequential subset ℓ, $E(\ell)$, and the routing entropy of a part type t over all it's sequential subsets, $E_\Sigma(t)$, can be summarized as follows :

$$E_n(\ell) = -\sum_{m_n(\ell)=1}^{M_n(\ell)} \{(1/M_n(\ell)) \ln(1/M_n(\ell))\}$$
$$= -\ln\{1/M_n(\ell)\}, \tag{3.1}$$
$$E(\ell) = -\sum \{1/N_\ell! \prod_{n\ell=1}^{N\ell} M_n(\ell)\} \ln\{1/N_\ell! \prod_{n\ell=1}^{N\ell} M_n(\ell)\},$$

where the range of summation covers all of the possible ψ_ℓ routings in the sequential subset ℓ,

$$= -\ln\{1/\psi_\ell\}$$
$$= \sum_{n\ell=1}^{N\ell} E_n(\ell) + \ln N_\ell!, \tag{3.2}$$

and

$$E_\Sigma(t) = \sum_{\ell=1}^{L} E(\ell). \tag{3.3}$$

Let E_Σ be the total routing entropy of the operation mix of production task, then

$$E_\Sigma = \sum_{t=1}^{T} Q(t) E_\Sigma(t). \tag{3.4}$$

The concept of routing entropy indicates the routing flexibility by analogy with the entropy of information theory, physics, and thermodynamics which takes a similar mathematical form. Occasionally the terms 'Commentropy' or 'Negentropy' are also used.

Also we note the following points which are of special interest to our study : As a part goes through it's operations, (i) both the total amount of work and routing entropy decrease, (ii) within a sequential subset ℓ, the total entropy in the sequential subset, $E(\ell)$, keeps decreasing until all the operations in subste ℓ are accomplished, and the nth operation performed reduces $E(\ell)$ by the following amount.

$$\Delta_n E(\ell) = E_n(\ell) + \ln(N_\ell - n + 1). \tag{3.5}$$

Under the first-come-first-served queueing discipline, the term $\ln(N_\ell - n + 1)$ has no bearing on the type of operation and depends only on the order that they are actually performed.

Therefore, the following routing rule is immediate : if a part has several parallel operations available, always perform next the operation which has the smallest routing entropy, $E_n(\ell)$, i.e. the smallest number of alternative stations, and retain the flexible operations to cope with potential future disturbances. As a direct consequence of the result, the best routing rule here can be identified as the one that delivers the operations in the increasing order of their routing entropy.

Now, we have an operation to perform next, then we need a method to assign this operation to one of the alternative stations. Conventional routing mechanism discussed in the open literature is to assign an operation randomly to stations according to their routing probabilities(Kelly 1979). However, stochastic analysis is primarily important for the system design purposes and there is no known implementation of probabilistic routing in any existing FMS. Therefore, what is required is some means of assigning operations to stations.

The probabilistic routing is essentially a randomized version of the deterministic routing. Suppose a deterministic routing, scheme which assigns an operation type n_ℓ to alternative stations, i.e. $1, \cdots, M_n(\ell)$, more orderly, while keeping the other things being fixed, then this certainly contributes to a better system performance. The formal approach for this line in the store-and-forward computer network can be found in Yum(1981).

To explain, suppose a part has operation n_ℓ to be performed next and it belongs to operation type j, and the operation has $M_n(\ell)$ alternative stations which are equivalent to the alternative stations of operation type j. The required production quantity of operation type j at station $m_n(\ell)$, $m_n(\ell)=1, \cdots, M_n(\ell)$, is $\delta_{m_n^*(\ell)j}$. The probabilistic routing assumes to assign an operation type j to station $m_n(\ell)$ with probability $p(m_n(\ell))$, $m_n(\ell)=1, \cdots, M_n(\ell)$, which is determined by $\delta_{m_n^*(\ell)j}/\omega_j$. On the other hand, the deterministic routing with $P(m_n(\ell))$s assigns the operation type j to stations in a sequential manner which minimizes the sum of squares of the deviations from rational $P(m_n(\ell))$s.

Let the first m assignment decisions consist of the sequence$\{\epsilon_j(1), \cdots, \epsilon_j(m)\}$ where $\epsilon_j(i)$ is the station of the ith decision, i.e. $\epsilon_j(1)$ is the station from which the first operation type j get service, $\epsilon_j(2)$ is the station from which the second operation type j get service, and so forth. Let $\Phi(m_n(\ell)|m)$, $m_n(\ell)=1, \cdots, M_n(\ell)$, be the total number of assignments at station $m_n(\ell)$ among these m decisions, then the sequence in the set can be determined as follows:

step 1 : m=1

step 2 : compute for all i, $i=1, \cdots, M_n(\ell)$,

$$\tau(i)=\sum_{m_n(\ell)\neq i}\{(\Phi(m_n(\ell)|m-1)/m)-P(m_n(\ell))\}^2+\{(\Phi(i|m-1)+1)/m)-P(i)\}^2 \qquad (3.6)$$

step 3 : $\epsilon_j(m)=\arg\min\{\tau(i),\ i=1, \cdots, M_n(\ell)\}$ (3.7)

step 4 : m=m+1 ; Go To step 2.

Also suppose ω_j, $\delta_{1j}{}^*, \cdots, \delta_{M_n^*(\ell)j}$ be relatively prime, then the generated sequences are all recurrent with recurrent period ω_j(Yum 1981). That is, the overall sequence S(j) takes the form of {(·)}, where (·) means that the sequence inside is to be repeated. As a numerical example, consider operation type j with $M_n(\ell)=2$. Let $P(1)=2/3$ and $P(2)=1/3$, then $S(j)=\{(1, 2, 1)\}$ with period 3 according to the above algorithm, and it represents the sequence of stations that assigns the first j operation which requires service to station 1, the second j operation to station 2, the third j operation to station 1, the fourth j operation to station 1, and so forth. This property considerably simplifies the routing mechanism, and the sequence inside the bracket completely determines the order of assigning operation type j to their alternative stations.

Therefore, the overall routing scheme can be summarized at two different levels : (i)if a part has several immediate parallel operations, perform the operation next which has the smallest routing entropy, (ii)among the several alternative stations, route the operation to the next station in the sequence $S(n(\ell))$.

An example is in order. Suppose at a certain point in time, a part enters it's ℓth sequential subset, which carries two parallel operations : turning(two alternative stations : T_1 and T_2 with $S(T)=\{(T_1, T_2, T_1)\}$) and boring(three alternative stations : B_1, B_2, and B_3 with $S(B)=\{(B_1, B_3, B_1, B_2, B_1, B_3, B_1)\}$). The last turning operation up to now was performed at station T_2, the second element in the sequence, and the last boring operation at station B_1, the third element in the sequence. The first decision is whether to go for turning first followed by boring, or the other way around. Since $E_T(\ell)$ value for turning operation is ln 2, while $E_B(\ell)=\ln 3$ for boring operation, the turning operation should by performed first. Then the next decision is whether to route the turning operation to station T_1 or station T_2, and since the next element in the sequence S(T) is T_1, it should be delivered to station T_1.

It is perhaps of interest to note the followings : (i) the more precise definition of the routing entropy could have been obtained such that it takes into account the work-load distribution, i.e. visit ratio, that is

$$E_n(\ell)=-\sum_{m_n(\ell)=1}^{M_n(\ell)} P(m_n(\ell)) \ln P(m_n(\ell)),$$

and, hence, $E(\ell)$, $E_\Sigma(t)$, and E_Σ accordingly. However, the routing entropy has been formulated to take simpler form with it's own meaning, (ii) to implement this routing mechanism, each part circulating in the system maintains and keeps track of only a list of it's immediate parallel operation(s) and the corresponding alternative station. And this information is updated whenever a part finishes an operation and hence constitutes an information flow, (iii) instead of the deterministic routing modelled here, we can consider the production-dependent probabilistic routing, that is, parts are routed to the latest station in production schedule with the highest probability. In general, such a state dependent routing doesn't have a simple analytical solution. Even though reversibility, key to the tractability, is assumed, the transition rates of the queue length process can not be factorized into two functions, one depending on the total number of parts at the source station and the other depending on the destination station, and, hence, it can not be modelled to have product-form solution (Kelly 1979.)

4. Application

To make proper use of the proposed routing principle by accomodating the service interruptions such as machine breakdowns, maintenance, and so on, the service rates of machines and the measure of flexibility, i.e. the routing entropy can be reformulated as follows. Suppose the proportion of time that the station i is operational, that is, the reliability of station i, is r(i), then the effective processing time of the machine at station i, $1/u_i$, can be redefined as the actual service time plus the duration of interruptions during the service, i.e. $1/u_i \rightarrow 1/(u_i r(i))$(Vinod and Altiok 1986). Therefore, the long-term optimal loading of station i, ρ_i^*, of the production planning problem in Kim(1988) should be adjusted to $\rho_i^* r(i)$ through equation (2.3), and hence, the required production quantity of operation type j at station i, δ_{ij}^*, i=1, ⋯, M and j=1, ⋯, J, in the short/medium term perspective, and, therefore, the generated sequence of operation type j, S(j), j=1, ⋯, J, can be obtained accordingly. In addition, the measure, routing entropy, can be reinterpreted to accomodate the reliability of the station, i.e.,

$$E_n(\ell) = -\sum_{m_n(\ell)=1}^{M_n(\ell)} \{r(m_n(\ell))/\sum_{m_n(\ell)=1}^{M_n(\ell)} r(m_n(\ell))\} \cdot \ln\{r(m_n(\ell))/\sum_{m_n(\ell)=1}^{M_n(\ell)} r(m_n(\ell))\}, \tag{4.1}$$

and $E(\ell)$, $E_\Sigma(t)$, and E_Σ accordingly.

Consequently, the routing principle can be adopted to take into account the machine state by simply skipping the station with service interruption in the sequence S(j) modified by 1/r(i) whenever service interruption occurs at that station until the station becomes operational again. By these simple adjustments, decisions on the choice of station to do the next operation can be adaptively made in response to disruptions of the operation of the system caused by the service interruptions.

5. Conclusion

In this study, we developed an hierarchical routing principle in FMS which does take into account both parallel operations and alternative stations. It is known that not only state dependent probabilistic routing(Buzacott 1982, Towsley 1980, and Yao 1987) but also deterministic routing(Yum 1981) is better than fixed probabilistic routing in performance. However it appears to be difficult to characterize the optimal routing rule except for the several special cases(Buzacott 1982, Seidmann and Schweitzer 1982, Maimon and Gershwin 1988). On the other hand, as the routing flexibility increases while keeping the other things being the same, for instance the visit ratios, the system performance improves(Yum 1978). Therefore, the routing scheme proposed here does have merit in considering both the performance effectiveness and the routing flexibility, as well as its simplicity and easy implementation.

Reference

1. Buzacott, J.A., "'Optimal' Operating Rules for Automated Manufacturing Systems", *IEEE Trans. Automatic Control*, AC-27, pp. 80~86, Feb. 1982.
2. Buzacott, J.A., and J.G. Shanthikumar, "Models for Understanding Flexible Manufacturing Systems", *AIIE Transactions*, 12, pp. 339~350, Dec. 1980.
3. Buzacott, J.A., and D.D. Yao, "On Queuein Network Models of Flexible Manufacturing Systems", *Queueing Systems, Theory and Applications*, Vol. 1, No. 1, pp. 5~28, 1986.
4. Gorden, W.J. and G.F. Newell, "Closed Queueing Network with Exponential Servers, *Operations Research*, Vol. 15, No. 2, pp. 252~267, 1967.
5. Groover, M.P., *Automation, Production Systems, and Computer-Aided Manufacturing*, Prentice Hall, Inc., Englewood Cliffs, N.J., 1980.
6. Jackson, J.R., "Jobshop-Like Queueing Systems,", *Manaement Science*, Vol. 10, No. 1 , pp. 131~142, Oct. 1963.
7. Kelly, F.P., *Reverrsibility and Stochastic Networks*, John Wiley & Sons, New York, 19 79.
8. Kim, S.C., "The Work-Load Distribution Problems in a Class of Flexible Manufacturing Systems", 대한산업공학회지, 제15권 1호, pp. 65~76, 1989.
9. Maimon, O.Z. and S.B. Gershwin, "Dynamic Scheduling and Routing for Flexible Manufacturing Systems That Have Unreliable Machines", *Operations Research*, Vol. 3 6, No. 2, pp. 279~292, 1988.
10. Seidmann, A. and P.J. Schweitzer, "Real-Time On-Line Control of an FMS Cell", working paper series No, QM 8217, Graduate School of Management, University of Rochester, Rochester, New York, 1982.
11. Stecke, K.E., "Formulation and Solution of Nonlinear Integer Production Planning Problems for Flexible Manufacturing Systems", *Management Science*, Vol. 29, No. 3, pp. 273~288, March 1983.
12. Towsley, D., "Queueing Network Models with State-Dependent Routing", *J. Ass. Comp. Mach.*, Vol. 27, No. 2, pp. 323~337, April 1980.
13. Vinod, B. and T. Altiok, "Approximating Unreliable Queueing Networks under the Assumption of Exponentiality", *J. Opnl. Res. Soc.*, 37, pp. 309-316, 1986.
14. Yao, D.D., "Material and Information Flows in Flexible Manufacturing Systems", *Meterial Flow*, 2, pp. 143~149, 1985.
15. Yao, D.D., "Modeling a Class of Flexible Manufacturing Systems with Reversible Routing", *Operations Research*, Vol. 35, No. 1, pp. 87~93, 1987.
16. Yao, D.D., and J.A. Buzacott, "Modeling a Class of State-Dependent Routing in Flexible Manufacturing Systems", *Annals of Operations Research*, Vol. 3, pp. 153~167, 1 985.
17. Yum, T., *Analysis of Adaptive Routing Rule in Computer Communication Networks*, Ph.D. dissertation, Columbia University, 1978.
18. Yum, T., "The Design and Analysis of a Semidynamic Deterministic Routing Rules", *IEEE Trans. Communications*, Com-29, No. 4, pp. 498-504, April 1981.