



Production Planning Model for a Flexible Manufacturing System

Pankaj Chandra

Faculty of Management, McGill University
1001 Sherbrooke St. W., Montreal, Quebec, Canada H3A 1G5

Abstract

In this paper we develop a production planning model which considers the problem of selecting optimal routes for manufacturing various part types in a Flexible Manufacturing System. We address the planning problem and its shop floor implications in a closed loop framework thereby making order release policies more effective. An hierarchical solution procedure is outlined. We illustrate the model with an example.

1. Introduction

Short product life cycles, mass customization, and short delivery" times are some of the important challenges that industries are facing today. Firms are looking for practices and technologies that will help them in addressing such needs. Recent advances in manufacturing technology, which cater to the growing needs of the industry, have strengthened the case for improving or introducing flexibility in manufacturing systems. Flexible manufacturing system (FMS) is such an example of this new technology. It comprises workcenters (consisting of machines) which are connected by a transportation system and are capable of performing a number of different tasks - all under computer control, thereby achieving improved productivity and reduction in the costs with losing sensitivity to product quality. The machines have automatic tool changing capabilities, which enable them to perform a set of operations while minimizing set up delays. Several different types of parts can be processed on a flexible manufacturing system. Parts, in general, have the flexibility to follow a variety of workcenter sequences and are made to demand (i.e., pull system). This reduces the lot size, sometimes to as low as one. The purpose of this paper is to provide an analytical methodology for production planning in a flexible manufacturing system.

FMSs can be analyzed using queuing network models for performance evaluation and optimization for system design and planning (Chatterjee, et. al., 1984) [5]. In their model, a workpiece goes through a series of operations at various workcenters before it is unloaded from the system. A path for a given workpiece is defined as a sequential set of operations. The flexibility of the system allows for various kinds of parts to be processed at the same time. The workcenter may consist of a single machine capable of performing a single job or multiple jobs (i.e., multi-tool machine with automated tool changing capabilities) or a group of machines. Hence, a workpiece which enters the system has the option of going to several different workstations. This is governed by the type of operation required on a job and gives rise to at least one path which the job may follow before being unloaded. Therefore, the network of workcenters and the transportation system (connecting workcenters and the load/unload center) allows many possible paths for a workpiece to move through the system. It has been found that variety and fluctuations in the demand of various workpieces or parts being routed influence the nature of the FMS routing procedure (Jaikumar, 1984) [6] and, consequently, the design of the planning system.

The pioneering work in the area of analytical modelling of FMS was done by Solberg (1978)[14] with the development of the CAN-Q model based on a closed queuing network. Kimemia and Gershwin (1978)[8] have presented an optimization model with workcenter level complexity (i.e., with all machines at a workcenter having the same process time for all part types) for an open-queuing network. Secco Suardo (1978)[13] coupled a non-linear programming formulation with a closed-queuing network model for a single class of jobs.

Stecke (1983)[15] has defined a set of production planning problems and attempts to solve the pooling and



the loading subproblems. It has been suggested that work load should remain unbalanced when the size of machine groupings (i.e., after maximum possible pooling) is unequal (Stecke, 1983)[15]. Wittrock (1985)[20] used heuristic algorithms to solve the part mix allocation problem, the objective being minimization of makespan or total completion time of each day's mix. He argues for studying setup times since, often, part types are grouped into families and a machine incurs a setup time while shifting from one family to another. Needless to say, there is a need to carry out capacity planning and operations scheduling before hand in order to guarantee a high usage rate of the expensive flexible manufacturing system (Schaluch, 1982)[12]. He observed that planning results would become better in measure as the degree of freedom for planning, such as alternative machines, alternative sequence of operations, and alternative operation variants are taken into consideration. Stecke and Kim (1988)[18] have presented a model which selects a subset of part types to be machined together and determines their aggregate production ratios over some upcoming period. The authors use CAN-Q for determining the relative ratios of workloads on machine types.

Several authors (including Stecke and Solberg, 1981[15]; Ammons et. al., 1984[1]; Chakravarty and Shtub, 1984[3]; Rajagopalan, 1986 [11]) have addressed various issues related to loading and control of FMSs. None, however, incorporate the state of the shop in their formulations. Few attempts have been made to solve the production planning and scheduling problems simultaneously. Kusiak (1986)[9] , Stecke (1986)[17] and Nelson (1986)[10] propose hierarchical production planning frameworks which involve solving the planning and scheduling problems in a top down fashion. To our best knowledge, this remains as the most feasible approach available. The results of Karmarkar (1987)[7] show the need for incorporating the shop floor status in determining order release policies and consequently planning decisions. Kusiak (1986)[9] outlines the necessity of exploiting the "interaction between FMS and its environment". One way of making production planning more reactive, in this hierarchical framework, is to address the planning problem and the shop floor implications in a closed-loop framework. This would render order releases to a scheduling module more relevant and effective.

Given a fixed mix of parts (as defined by its demand) and knowing the workcenters that can perform each operation, we want to determine the optimal number of parts to be processed on each feasible route and subsequently the workcenters on which to process them such that the overall production rate of the system is maximized or the total variable costs are minimized. Such a planning exercise becomes necessary in light of the above discussion, especially, since it forms the basis for production scheduling.

In the next section we develop a mathematical programming model which maximizes the throughput rate of the system while the build up of queues within the system is modelled as a constraint. An illustrative example is provided in Section 3 and conclusions are presented in Section 4. Details of the solution procedures are given in the Appendix.

2 Optimization Model for a FMS and the Solution Algorithm

In this section we develop a production planning model for a flexible manufacturing system. We also describe a solution algorithm for solving the above problem. The manufacturing system comprises a set of workcenters which is a collection of machines with a dedicated inter-machine transport facility. Each machine can perform one or more operations. One or more operations is required to process any given part type. A route is defined as a sequence of workcenters (and machines) which a part type visits before it becomes an end product.

Given a product mix, we seek to determine the mix of routes which maximize the total production rate for all the different part types in the manufacturing system. The parameters and variables that describe the problem are:

Parameters

p = number of part types
 d = number of operations



n = number of workcenters

b = number of machine types

r_i = number of feasible routes for producing part type i

S_{ml} = setup time for operation m on machine type l

O_{imlj} = processing time for part type i for operation m on machine l at workcenter j

d_i = demand for part type i

C_l = available capacity at machine l

a_i = fraction of part type i in the total production

I = maximum allowable work-in-process inventory

$t_{(j,k)}$ = travel time between workcenters j and k

Variables

X_{ijr} = production rate of part type i on machine type l at workcenter j on route r

PR_i = production rate of part type i over all routes

$q_j(X_{1j}, \dots, X_{pj})$ = average queue length at workcenter j comprising different part types

The optimization problem pp which maximizes the overall throughput rate of the system subject to the constraint imposed on the average level of work-in-process inventory can be formulated as follows:

$$\text{PP : Maximize } \sum_{i=1}^p PR_i$$

subject to:

$$PR_i = \sum_{r=1}^{r_i} X_{i|j^*r}, \quad \forall j^* \in M(j^*), \forall l^* \in M(l^*j^*) \quad (1)$$

$$PR_i = a_i \sum_{i=1}^p PR_i \quad \forall i \quad (2)$$

$$\sum_{j=1}^n q_j(X_{1j}, \dots, X_{pj}) + \sum_{i=1}^p \sum_{r=1}^{r_i} \sum_{l=1}^b (X_{ijr} \cdot t_{(j,k)}) \leq I \quad \forall (j,k) \in A \quad (3)$$

$$\sum_r X_{i|j^*r} \geq d_i \quad \forall i \quad (4)$$

$$\sum_i \sum_m \sum_r X_{ijr} \cdot O_{imlj} \leq C_l \quad \forall j, l \quad (5)$$

$$X_{ijr} \text{ and } PR_i \geq 0 \quad \forall i, j, l, r \quad (6)$$

Constraint 1 establishes the relationship between the production rate for each part type and the production rates at bottleneck stations on each route for every part type. The production rate on any route, for a given part type, is given by the production rate of the bottleneck machine on that route. Since it is possible that a machine can be on more than one route, the workload at any given machine comprises amount produced on different routes and, may be, consisting of different part types. It can also be appreciated that the bottleneck workcenter may be different for each route. l^* is the bottleneck machine on route r for part type i and is determined using the algorithm outlined in the Appendix. Similarly, j^* is the workcenter which includes the bottleneck machine l^* on the route r . $M(j^*)$ is the set of bottleneck workcenters and $M(l^*j^*)$ is the set of

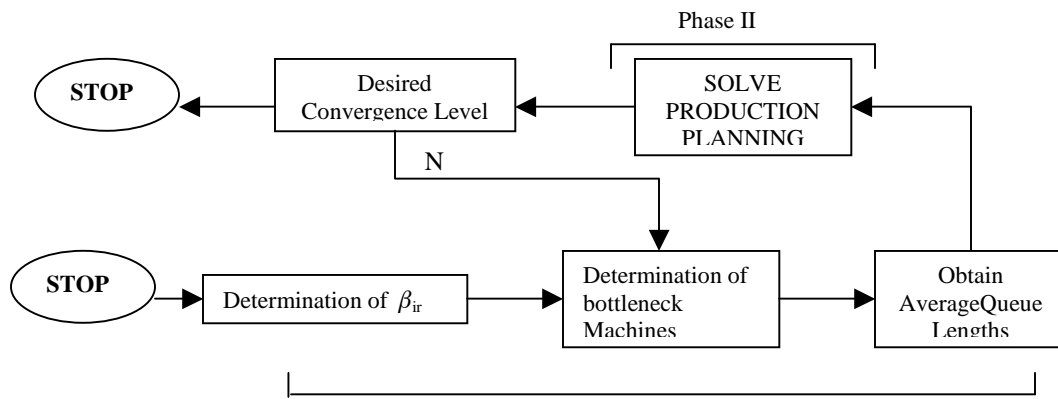
bottleneck machines in each workcenter, j^* . Constraint 2 is the ratio requirement for each part type i such that $E_i = 1$. It can capture the relationship between part types as defined by the bill of material. Constraint 3 ensures that the average WIP does not exceed the prescribed level. The WIP level is the sum of average queue lengths at all workcenters and the pipeline stock (i.e., stock under transportation between workcenters). A is the set of all workcenter pairs which are connected by the material handling system. The average queue length at any workcenter includes arrivals from other workcenters (or other machines in the same workcenter) as well as external arrivals. However, if $j = 1$ is the load station and $j = n$ is the unload station, then $q_1 = \text{external queue}$ while all other workcenters have no external queues. Constraint 4 ensures that the production rate over all routes, for each part type, meets the demand for that part type. Constraint 5 ensures that the amount produced on any given machine, over all possible routes, does not exceed the capacity of the machine. Constraint 6 are the non-negativity constraints.

The above mentioned production planning model embodies a spectrum of issues ranging from the determination of routes for each part type to the determination of average queue lengths. We have developed an hierarchical solution algorithm (which is shown in Figure 1) for solving the problem PP. The heuristic procedure involves solving an LP along with a queuing subproblem PP. The heuristic procedure involves solving an LP along with a queuing subproblem -these two phases are tied together in a closed-loop framework. As shown in Figure 1, Phase I determines the proportion of a given part type that should be manufactured on each of the feasible routes (i.e., the starting solution); this subsequently forms the input to queue-length-determination problem which provides us with the average queue length at each workcenter. Phase II uses the average queue values from Phase I to form the production planning problem. It determines the production quantity for each part type on every feasible route. Each phase provides input to the other -the cycling process, thereby, updates and improves the optimal part-route mix at the end of each iteration until it converges. The hierarchical solution procedure requires enumeration of feasible routes for each part type in terms of operation and machine visits. Any standard route generation algorithm can be used for obtaining the routes (e.g., route generator of Chatterjee et. al., 1984[5] or one can be developed using the constructs in Chandra, 1993 [4]). The initialization of the solution procedure is done via an aggregate determination of the quantity of given part types that should be manufactured on a given feasible route (i.e., β_{ir}). This has been dealt as SUBPROBLEM 2 in the Appendix. SUBPROBLEM 1 in the Appendix outlines a procedure for determining the bottleneck machine on each feasible route and discusses related issues. The queuing subproblem (i.e., the determination of the average queue length at each workcenter) is solved using the following heuristic which modifies Suri and Hilderbrant's (1984)[19] variant of Schweitzer-Bard algorithm for the above problem: "

$$W_{ij} = \left\{ \left[\frac{\sum_{r=1}^R \beta_{ir} - 1}{\sum_{r=1}^R \beta_{ir}} \right] \cdot Q_{ij} \cdot O_{imij} \right\} + \left\{ \sum_{v=1, v \neq i}^P Q_{vj} \cdot O_{vmij} \right\}$$

W_{ij} is the average waiting time for part type i at workcenter j and Q_{ij} is the average queue length at workcenter j as part type i enters the workcenter. All other variables

Figure 1: Hierarchical Solution Procedure



have their defined meaning. The above heuristic is employed to determine $g_i(X_1, \dots, X_n)$ using 'Mean Value Analysis' (Suri and Hilderbrant, 1984)[19] for the multiple product case. The heuristic converges in finite number of steps (convergence results are shown in Section 3).

The structure of the general LP optimization problem (i.e., PP) is such that a solution procedure can be developed by application of the Dantzig- Wolfe price directive decomposition procedure for solving the LP problem for the entire network (especially for large size problems as described in Chandra (1993) [4]). For small size problems one can use any standard simplex code (e.g., LINDO).

3 Numerical Results for a Four Station System

Consider the system depicted in Figure 2. There are two workcenters in the system and the load and unload stations. There are two types of parts being manufactured. Each workcenter comprises two machine types. Each workcenter is connected by a dedicated conveyor. Three different kinds of operations can be performed by the four machines. Operation 1 needs to be performed on part type 1 while operation 2 and 3 need to be performed on part type 2.

$$d = 3 \text{ } \forall [m = 1, 2, 3]$$

$$n = 4 \text{ } \forall [7 = L, 1, 2, UL] \text{ } p = 2 \text{ } \forall [i = 1, 2,]$$

$$b = 4 \text{ } \forall [I = L1, L2, L3, L4]$$

- Load and Unload operations take 1 minute each

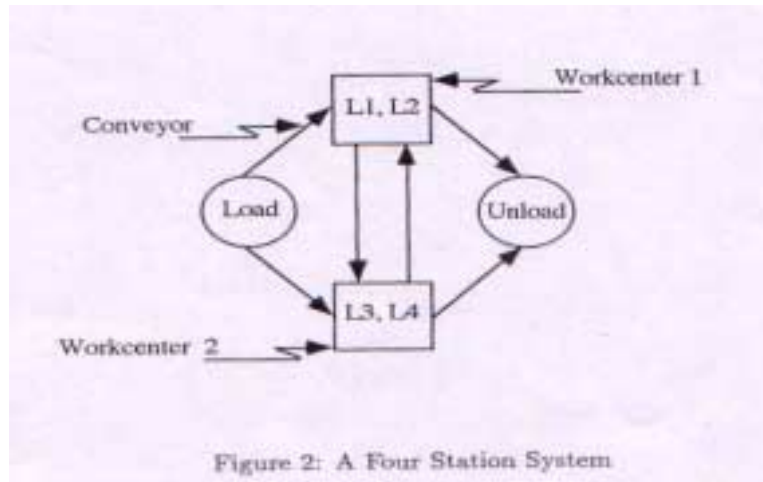


Figure 2: A Four Station System

- Average travel time on each arc = 1 minute, i.e., $t_{(j,k)} = 1 \text{ min.}$ for all $(j,k) \in A$
- $a_1 = a_2 = 0.5$
- The average operation times (in minutes) on each machine type are given as follows:

Operation	Machine Type			
	1	2	3	4
1	12	-	-	11
2	8	7	-	-
3	-	-	13	-

Routings for the Two part Types

Part type 1

ROUTE 1 : L – W1(L1) – UL
 ROUTE 2 : L – W2(L4) – UL

Part type 2

ROUTE 3: L – W1 (L1) – W2(L3) –UL
 ROUTE 4: L – W1(L2) – W2(L3) –UL
 ROUTE 5: L – W2(L3) – W1(L1) –UL
 ROUTE 6: L – W2(L3) – W1(L2) –UL

where W1 and W2 are workcenters 1 and 2 and L1, L2, L3 and L4 are the machines used at those workcenters. Since this design is simple, routes can be enumerated by inspection. Alternatively, for a complex design, the route generation algorithm of Chatterjee et. al. (1984)[5] can be employed for this task using the constructs and matrices that have been developed In Chandra (1993)[4].

The starting solutions (.*dir*) for the above routes were determined using Subproblem 2. From the analysis of bottleneck machines (Subproblem 1) for the above routes for part type 1, the bottleneck machines are L1 for Route 1 and L4 for Route 2. For part type 2, the bottleneck machine is L3 over all routes. These form the starting solution to the overall problem as well as inputs to the queuing subproblem.

Queuing Subproblem

The computer implementation for determining the average queue length (using the modified Schwitzer-Bard heuristic given in the earlier section; this is used within the mean value analysis framework) was done in Fortran 77. For the base total demand of 120 units for the entire planning period, the results obtained for the four workcenter system, were as follows:

Average Queue Length at Node 1: 4
 Node 2: 4
 Node 3: 4
 Node 4: 4

The heuristic was tested for convergence and it was found to converge in a finite number of steps. The following table exhibits the number of iterations taken to converge, the cpu time used on DEC 10, and the average queue lengths obtained for a ten workcenter FMS while varying the part types from 2 to 10:

Table 1: Performance of the queueing heuristic.

No. of Part types	No. of workcentres	No. of Interactions	CPU Secs.	Total No. of parts in Queue
2	10	6	0.47	20
4	10	7	0.91	40
6	10	7	1.17	50
8	10	7	1.45	70
10	10	7	2.01	92

The efficacy of this implementation vis-a- vis other methods (which needless to mention are also approximations) still remains an exercise.

Now that all the subproblems have been solved the original production planning problem (PP) is formulated and solution obtained. The value of *I* is the limit on work-in-process inventory in the system for a given demand for each part type. Total number of parts in queue in the system is 16. Results were obtained for the current demand level (i.e., 60 units of each part type). Subsequently, the demand level was



varied and the whole exercise repeated. Movement of parts from one workcenter to another with the increase in demand levels was studied.

At the base demand levels (60 units of each part type per planning period) all of type 1 parts followed route 2, i.e., L – W2(L4) – UL. The proportion of part type 1 that use workcenter 2 is shown in Figure 3 as a function of the work-in-process inventory.

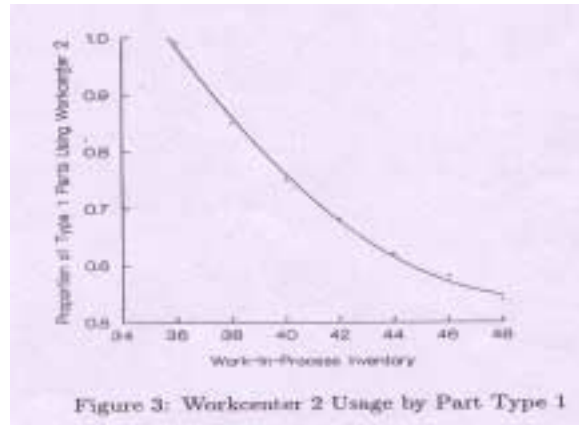


Figure 3: Workcenter 2 Usage by Part Type 1

When the in-process inventory is low the proportion of part type 1 using workcenter 2 is high since it is the faster workcenter (operation time for operation 1 is the lowest at machine L4). As the number of parts in the system increases, more type 1 parts start utilizing workcenter 1. Similar results were reported by Secco-Suardo (1978) [13] while optimizing a closed network of queues. However, in his study the proportion of parts visiting the faster station depended on the ready number of pallets. Figure 4 shows the variation of the production rate of type 1 with I and was found to be directly proportional to the work in process inventory. Nevertheless, maximum production rate stabilizes at 10.45 parts/hr. as I tend to infinity. At that production rate both are machines (L1 and L2), which can perform, operation 1 on part type 1 are utilized to the fullest.

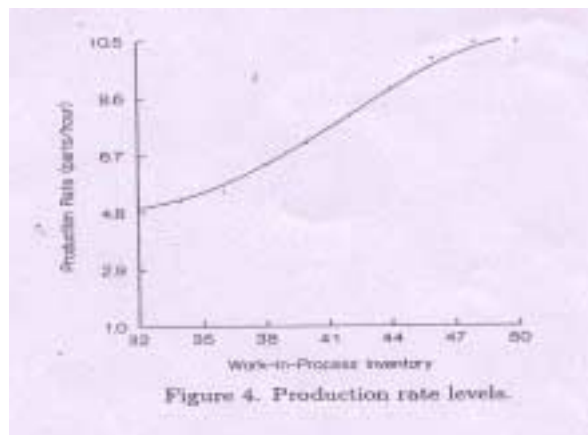


Figure 4: Production rate levels.



Type 2 parts, on the other, hand, always follow route 4: L-W1(L2) W2(L3) UL They go to workcenter, 1 first and then to workcenter 2. Workcenter 2 (machine L3) is the bottleneck hence using route 6(L W3 (L3)_W1(L2) UL) increases the work-in-process inventory correspondingly.

4 Conclusions

We have presented an optimization model for production planning in a Flexible Manufacturing System. Our model takes into consideration the state of the shop in making decision, regarding the quantity of different products that are to be produced on various, routes in the manufacturing system. We develop an hierarchical solution procedure, that solves the planning and the queuing subproblems in an iterative fashion to obtain the, result. An example was solved to exhibit the, feasibility of the proposed system result, leading to the convergence of the queuing heuristic were reported. We feel that implication of the result, that were obtained. We feel that this closed one framework at an early stage of production planning make, the plan more reactive and effective when implemented via scheduling process as it considers the impact of loading on shop floor queues while making the production planning decisions. This exercise of determining the optimal proportion of each type to be manufactured on the, various feasible routes is important from the point of view of the scheduling problem. Development of a scheduling heuristic, for this, system and a study of the performance of the hierarchical planning and scheduling problem would be, the next step in this Research.

Acknowledgment

This, research has been partially supported by NSERC Grant OGPO04250 Valuable comments, by Professor, Morris Cohen and Christopher Jones and an anonymous referee are gratefully acknowledge.

References

- [1] Ammons JC., CB Lofgren, and LF M, Gioni, (1984) "A Large,-Scale, Workstation Loading Problem" in KE Stecke, and R Suri (eds.), *Proceedings, of the First, ORSA/TIMS conference on Flexible, Manufacturing System*, Ann Arbor, PP 249-255
- [2] Buzacott, JA and DD Yao, (1986) "Flexible Manufacturing System: A Review of Models" *Management Science* 32, pp 890-905
- [3] Chakravarty, AK and A Shtub, (1984) "Selecting Parts and Loading Flexible Manufacturing Systems in KE Stecke, and R Suri (eds.) *Proceedings, of the, First, ORSA/TIMS Conference on Flexible, Manufacturing Systems* Ann Arbor, PP 284-289
- [4] Chandra., P" (1993) "Production Planning Model for, Flexible Manufacturing System", Working Paper, Faculty of Management, McGill University, Montreal
- 151 Chatterjee, A" MA Cohen, and W. Maxwell, (1984) "Manufacturing Flexibility Models, and Measurements," in K.E. Stecke and R. Suri (eds.) *Proceeding, of the, First, ORSA/TIMS Conference on Flexible, Manufacturing systems*", Ann Arbor, pp 49-64
- [6] JaikumM, R (1984) "Flexible Manufacturing System, A Management Perspective", Working Paper # 1-784-078, Harvard business School
- [7] Karmarkar, US" (1987) "Lot Sizes", Lead Time and In-Process Inventories," *Management Science*, 33, pp 409-418
- [8] Kimemia JG and SB Gec,hwin, (1978) "Network Flow Optimization in a Flexible Manufacturing System", *IEEE Conference on Decision Control*, pp 633- 639
- [9] Kusiak, A" (1986) " Application of Operational Reserch Model, and Technique, in Flexible Manufacturing System," *European Journal of Operational Reserch*, 14, pp 336-345
- [10] Nelson CA" (1986) " A Mathematic Proramming Formulation of Element, of Manufacturing Strategy FMS Application, in KE Stecke and R Suri (eds.) *Proceeding, of the, Second ORSA/TIMS Conference*

on Flexible Manufacturing System, Elsevier Science Publishers, B.V. Amsterdam, pp 31-42



- [11] Rajagopalan, S (1986) "Formulation and Heuristic Solution, for Part Grouping and Tool Loading in Flexible Manufacturing System.", in KE Stecke and R Suri (eds) *Proceedings, of second ORSA/TIMS Conference on flexible Manufacturing System*, Elsevier Science Publishers, BV Amsterdam, pp 311-320
- [12] Schlauch, R" (1982) "Production Scheduling in FMS", Technical Report, Fraunhofer Institute for Manufacturing Engineering & Automation, Stuttgart, W Germany
- [13] Secco Suardo, G (1978) "Optimization of a Closed Network of Queue" *LIDS MITReport No. ESL FR834-3*, MIT., Cambridge
- [14] Solberg, JJ (1978) "Analytic Performance Evaluation for the Design of Flexible Manufacturing Systems *Proceeding, of the, IEEE Decision and Control conference*, San Diego, CA, pp 640-646
- [15] Stecke, KE and JJ Solberg, (1981) "Loading and Control Policies, for Flexible Manufacturing System", *International Journal of production Research*, 19, pp 481-490.