Makespan estimations in flexible manufacturing systems

Citation for published version (APA):

Document status and date:
Published: 01/01/1994

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Makespan Estimations in Flexible Manufacturing Systems

Jan C. Fransoo, Ton G. de Kok, and Jan Paulli

Research Report TUE/BDK/LBS/94-14
December, 1994
Makespan Estimations in Flexible Manufacturing Systems

Jan C. Fransoo*, Ton G. de Kok* and Jan Paulli**

Research Report TUE/BDK/LBS/94-14
December, 1994

* Graduate School of Industrial Engineering and Management Science
  Eindhoven University of Technology
  P.O. Box 513, Paviljoen F16
  NL-5600 MB Eindhoven
  The Netherlands
  Phone: +31.40.472230
  E-mail: J.C.FRANSOO@BDK.TUE.NL and A.G.D.KOK@BDK.TUE.NL

** Department of Operations Research
  University of Aarhus
  Ny Munkegade, Bldg 530
  DK-8000 Aarhus C
  Denmark
  Phone: +45.89423542
  E-mail: JANP@MI.AAU.DK

This paper should not be quoted or referred to without the prior written permission of the authors.

Please address all correspondence to the first author

This working paper will also be published as Publication no. 95/6, Department of Operations Research, University of Aarhus, Aarhus, Denmark

The research of Dr. Fransoo has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences
Makespan Estimations in Flexible Manufacturing Systems

Jan C. Fransoo¹ & Ton G. de Kok
Graduate School of Industrial Engineering and Management Science, Eindhoven University of Technology, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands.
Jan Paulli²
Department of Operations Research, University of Aarhus, Bldg 530, Ny Munkegade, DK-8000 Aarhus C, Denmark.

Abstract

In the production control concepts of production departments it is usually necessary to estimate the makespan, since this is required as an input in the MRP-system. This paper addresses the makespan estimation problem if the department consists of a flexible manufacturing system (FMS). We compare a makespan estimation based on the analysis of a stochastic queueing network model of the FMS and a makespan minimizing algorithm based on a combinatorial algorithm. Results indicate that the queueing network model has some added value in estimating the makespan, given a limited availability of detailed information and instant solution requirements.

1. Introduction

The scheduling problem in flexible manufacturing systems (FMSs) has been addressed extensively in the literature. Usually, the scheduling decision is considered part of a decision hierarchy, in which batching and tooling decisions are included as well. Additionally, in many practical cases, the Material Requirements Planning (MRP) decision is also part of the decision hierarchy (see figure 1). Consider a production situation which consists of a fabrication stage and an assembly stage. A subset of the

---

¹ Author to whom correspondence should be addressed
² The research for this paper was performed while the author was a visitor at Eindhoven University of Technology.
intermediates is fabricated on a large FMS, which has been modelled as a separate production stage at the MRP level. At the MRP level, jobs are generated for the FMS stage to be completed within a time bucket, which we assume -- without loss of generality -- to be a week. It is important to know in advance whether this set of jobs is likely to be completed within a week. It is useful to have some insight into the bottleneck situation with this specific set of jobs, in order to be able to modify the requirements plan, the job routing, or any other characteristics which may be modified in the short run. Note that in FMS environments bottlenecks are shifting from one machine to another depending on the product mix. To evaluate a set of alternative solutions, it is important that these data are available without much PC-CPU time.

Basically, two lines of thought can be followed in order to properly evaluate the alternatives. The first line of thought is based on scheduling theory and generates detailed production schedules for a given set of jobs and production characteristics. It contains a lot of detail, and it is therefore difficult to model and sensitive to minor specification changes. The second line of thought uses queueing theory as a basis for the aggregate analysis. Generally, this requires less detailed modelling and the solution is more robust. However, not all details can be incorporated into the model, which may affect the quality of the decision.

The objective of this paper is to compare these two approaches in an FMS-setting. Both approaches will be compared regarding their use for the aggregate planning problem, at the tactical (MRP) decision level. In section 2, we will give a brief overview of the literature. Section 3 contains a description of the problem and the approach which is
proposed in this paper. Simulation experiments have been used for testing both approaches. These experiments have been described in Section 4. The paper is concluded by a discussion of the results obtained.

It should be noted that the purpose of this paper is primarily to introduce a line of research where different approaches addressing the same object of research are assessed. We expect this paper to lead to some discussion regarding the relevance and usefulness of research like this. It is not a complete analysis of a specific topic in this area, since sufficient insight into this material has is not yet been developed. On the other hand, it seriously assesses the value of a number of OR-methods which are generally assumed to be applicable to solve problems like the one introduced in this paper.

2. Literature review

A full coverage of the literature dealing with scheduling of FMSs would be too extensive, so we will only give a few examples of some of the different approaches found in the literature. For a review of FMS scheduling literature (as well as other FMS related topics) see e.g. Kaighobadi and Venkatesh (1994), Gunasekaran et al. (1993) and Tempelmeier and Kuhn (1992).

Quite often the FMS scheduling problem is seen as part of a larger (hierarchical) planning system, taking care of e.g. the part types to be processed at the same time, the set-up of machines and the scheduling. An example of such an approach is the model presented in Sawik (1990), in which the objective of the scheduling is to minimize the maximal lateness of processing the jobs.

Another approach is to consider the FMS scheduling as a separate problem, taking the batches and the set-up of the system as given. Examples of such a model include the integer programming model presented in Sherali et al. (1990) and the model presented in Paulli (1995). In both of these papers the objective is to minimize the maximal completion time of the jobs. Finally, one of the often used scheduling approaches is to use dispatching rules. Here a dispatching/priority rule is used to choose the course of action every time the status of the system changes. Montazeri and Van Wassenhove (1990) give an extensive simulation study of numerous dispatching rules under several objectives.

As a conclusion to this brief overview of literature dealing with FMS scheduling one can
say that what is to be understood as the scheduling problem in an FMS is not clearly defined and consequently it is very hard to assess the quality of the different approaches.

Hierarchical models for production planning usually discriminate between an aggregate capacity planning decision on the tactical level and a detailed scheduling decision on the operational level. This approach is suited to fit the organizational structure (Meal, 1978). Planning decisions may be taken at the Master Production Scheduling level and usually assume fixed lead times at the operational level (Vollmann et al., 1979). This lead time is then controlled by some form of input-output or workload control (Bertrand, 1985). In order to keep track of the lead times which will be realized, adequate aggregate models of the detailed production situation need to be available. In job shops, stochastic queueing network models are used to investigate these issues (Bertrand and Van Ooijen(1991)).

A lot of authors model production systems, and in particular FMS's, as open or closed queueing networks. An overview of this modelling approach can be found in Suri et al. (1993) and Buzacott and Shantikumar (1992). In particular, we built on the results of Whitt (1983). It should be noted that we initially used a standard Mean Value Analysis approximation scheme for non-exponential service times along the lines suggested by Reiser (1979). However, we found that the analysis yielded utilizations of machines exceeding 1. Therefore we discarded this method, in spite of the fact that it is widely used in many publications (cf. Suri et al., 1993). It should be noted that in most publications on FMSs modelled as closed queueing networks no explicit validation is given of this modelling approach.

3. Makespan estimation models

A makespan estimation is useful at the tactical planning level. At this level decisions regarding the allocation of capacity, short term workforce (overtime work), order acceptance, etc. are made. In order to properly make these decisions, a model which shows the rough consequences of various alternatives might provide useful support. This support can be of various levels, depending upon the quality of the model.

First, the model can give ordinal information regarding the proposed alternatives. This means that each of the alternatives can be sorted in terms of ascending makespan by the aggregate model. The ordinal information means that the sequence is correct as if it were
done in detail, but it does not indicate the absolute value (time) of the makespan.
Second, the model can give relative information regarding the proposed alternatives. This
means that it does not only provide differences ("The makespan of set A is shorter than
the makespan of set B"), but also the quantity of the difference ("The makespan of A is
x% shorter than the makespan of B"). Third, the model can give absolute information
regarding the proposed alternatives. This means that the makespan which the model
generates also provides absolute information whether the products can be manufactured
within a certain time constraint ("The makespan of A equals y").

It is not an obvious thought to estimate the makespan by using a queueing network. This
is especially true because a queueing model is based on average flows through the
system, while a deterministic scheduling algorithm uses specific problem information to
determine sequences on each of the machines. Furthermore a queueing model assumes
that we process an infinite number of jobs to ensure stationarity, while a scheduling
model assumes a finite set of jobs to be processed. Two problems, which may seem
similar at the aggregate level, may have different characteristics making the scheduling
problem fairly easy in one case and very difficult in the other. The SQN-model is
expected to capture a part of this specificity at the aggregate level. Which part it actually
does cover, is, as of yet, unknown.

Below, we will first describe the aggregate model we used for the estimation of the
makespan. Then, we will describe the scheduling algorithm we used for the calculation
of the makespan.

3.1 SQN and modelling

The aggregate model is based on a queueing network model. Yet we start with a formal
description of the FMS we consider. It should be noted that in spite of the fact that we
consider the same FMS in both the queueing model and scheduling model, the model
description is slightly different because of the underlying modelling approach. This
underlines the main theme of our paper, which is to apply different modelling approaches
to the same situation.

We consider an FMS consisting of a number of work cells, each consisting of a number
of work stations. Jobs are routed along the work stations using a pallet on a material
handling system. There is a finite number of pallets available. There is only one type of
pallets, which is able to carry any job. Each work station has its own queue. We consider
a finite set of jobs that needs to be processed by the FMS. Each job belongs to a particular family for which a fixed sequence of operations is given. For each operation, we have a number of (parallel) work stations that can process the job. After an operation has been finished at a particular work station, the next work station is selected from the set of feasible parallel work stations according to some probability and the job joins the queue in front of the selected work station. Processing times are considered to be random variables determined by their mean and variance. In the numerical experiments we assumed that the processing time of a job at a work station is deterministic, which is reasonable from a practical point of view. It should be noted, however, that stochastic processing times can be modelled analogously. Processing times are dependent on the job and the processing phase, but not on the work station selected. Hence we assume that parallel work stations are identical. Note that in this definition the concept of work cell is quite flexible, since we only consider logical work cells, dependent on the job routing.

Especially the selection of the next work station differs from the scheduling approach discussed in section 3.2. The scheduling approach attempts to select the machine in such a way that an optimal schedule results. In the SQN-model, however, the "selection" is captured in the transitions probabilities. We restrict to giving an outline of the analysis. The analysis is analogous to the one in Whitt (1983) for open queueing networks with deterministic routing.

Based on the operational characteristics of the families, we can determine for each work station the mean and variance of its weighted processing time. These weights are based on the routings of each job and the number of jobs to be processed in each family. The number of jobs to be processed is considered to be the external arrival rate of jobs per period for a particular family. Furthermore we can derive a routing matrix describing the probability that a job leaving a specific work station visits another or the same specific work station next.

Using the external arrival rates, the processing time characteristics and the routing matrix we may apply standard approximate MVA schemes. As indicated above we found utilization degrees exceeding 1. The standard approximate MVA scheme used is based on substituting Pollaczek-Khintchine's formula for the M/G/1-queue in the Mean-Value equation to correct for non-exponential service times. Now mean throughput times and mean throughput, or equivalently utilization degrees, are tightly linked in closed queueing networks as follows from Little's formula (cf. Tijms (1986) amongst others). The approximation error in the throughput time apparently introduces utilization degrees
exceeding 1. Therefore we used an alternative method which is related to the decomposition approach discussed in Suri et al. (1993) for closed-loop systems, which does not suffer from infeasible utilization degrees and gives similar results as the MVA schemes in case of exponential processing times. An extensive numerical comparison of approximate MVA schemes and the approach sketched below is currently under investigation.

The basis of our analysis is the analysis of an open network of GI/G/1 queues. Instead of applying the QNA approach of Whitt (1983), we use more accurate approximations for average waiting times and second moments of interdeparture times (cf. Tijms,1986). We close the open network by adding a node from which all jobs enter the open system and to which all jobs leave the open system. For given values of the relative arrival rates from the artificial node we compute iteratively by substitution the associated coefficients of variation of these arrival processes. We find that this iteration converges quite fast. As of yet, we cannot give a formal proof for this convergence. After convergence we compute, using the GI/G/1 analysis for each work station, the overall throughput times and average number of jobs in the system. Since this number should be equal to the number of available pallets, we compare the two numbers. If the number found is smaller than the number of available pallets, then we multiply the relative arrival rates by some factor greater than 1, otherwise we multiply the relative arrival rates by some factor smaller than 1. Basically, we apply a bisection scheme by varying one of the external arrival rates, while keeping the arrival rates in the same proportion to each other.

This analysis finally results in the external arrival rates, i.e. the number of jobs that can be processed per time unit (i.e. per week) with the given number of pallets in the FMS. The next step in the analysis is to compute a makespan estimate for the set of jobs to be processed each period. The makespan estimate is computed as follows.

\[ M = \frac{\lambda \times TPT}{N} \]

where

- \( M \) = estimated makespan
- \( \lambda \) = total number of jobs to be processed
- \( TPT \) = average overall throughput time per job
- \( N \) = number of pallets

In section 3.2 we describe the algorithm applied to the problem of scheduling the FMS.
In section 4 we give a numerical comparison of the two modelling approaches, thereby testing the validity of application of queueing models.

3.2 The scheduling algorithm

For scheduling the FMS we used an approach suggested in Paulli (1995). Here the assumptions are:

- a number of jobs consisting of several consecutive operations has to be processed on a collection of machines.
- the system is set-up such that the operations can be performed on a subset of the machines.
- once the processing of an operation has started on a machine it can not be interrupted.
- the processing times do not depend on the machines.
- set-up times between operations and transit times between machines are negligible.
- a maximum of \( P \) pallets are allowed in the system at the same time (e.g. due to limitations of the material handling system).

The objective of the scheduling algorithm is to minimize the maximum completion time of the jobs (the makespan).

A special case of the above scheduling problem is the classical job-shop scheduling problem (each operation can be performed on one and only one machine and the number of pallets is equal to the number of machines). Consequently, the FMS-scheduling problem is NP-hard and a heuristic problem for solving the problem is necessary. The algorithm suggested in Paulli (1995) is based on the similarities between the FMS scheduling problem outlined above and the classical job-shop scheduling problem. The idea is to make a starting schedule using rather simple dispatching rules and then try to improve the schedule using the similarities with the job-shop scheduling problem. For details on the improvement step we refer to Paulli (1995), but the starting solution procedure is described below.

The starting schedule:

1. The jobs are released to the system in a fixed sequence. The sequence is determined using the following two rules:
   - mix part types as much as possible.
in case of ties, sequence the job with the longest total processing time first.

2. Release the first P jobs to the system.
3. Make a partial schedule for the jobs currently in the system until one of the jobs is finished.
4. If not all jobs have been released to the system then release the next job in the sequence. If all jobs have been scheduled then stop, otherwise goto step 3.

The partial schedule in step 3 is determined using the following procedure:

i. Each time an operation is finished the job is put in a central queue, waiting for processing.

ii. Each time a machine becomes free we choose according to a FIFO rule (First-In-First-Out) from the set of jobs that wait for processing and have a next operation that can be performed on this machine.

4. Simulation Experiments

In the simulation experiments, we have modelled a specific FMS, similar to the one modelled in Montazeri and Van Wassenhove (1990). The FMS has the following characteristics.

There are four types of machines: A, B, C and D (e.g. drilling and milling machines, grinders and lathes). There are three machines of types A and B (denoted A1, A2, A3, B1, B2 and B3) and two machines of types C and D (denoted C1, C2, D1 and D2). The machines of a given type may or may not be identical. In the identical case any operation that can be performed on one of the machines of a specific type (e.g. A1) can also be performed on the other machines of this type (i.e. A2 and A3). In the non-identical case some operations may be performed on e.g. A1 alone, while others may be performed also on A2 and/or A3. However, it is impossible that an operation can be performed on machines of the different types (e.g. A1 and B1). This flexibility is determined by the product routing specification. If a lot of operations can be performed on multiple machines, we consider the system to have a high routing flexibility.

The material handling system is assumed to be sufficiently fast, so that the parts in the system never have to wait for the material handling system. Therefore, we do not
consider the scheduling of the material handling system; it will never impose constraints upon the use of the system. However, the limitation of the material handling system is represented by the limited number of pallets it is able to handle at the same time.

The pallets and fixtures are assumed to be universal (i.e. not dedicated to a specific part type). However, a combination of the price of the pallets and fixtures and limitations in the material handling system limits the maximal number of pallets in the system at the same time to a fixed number. In-process buffers are assumed to have enough capacity, so that having too many parts in front of a single machine need not to be considered. The loading/unloading station is assumed to be sufficiently fast, so that a new pallet (with a part mounted on it) can enter the system right after a pallet with a finished part leaves the system. Finally, the system is assumed to available 40 hours a week. We will not consider the problem of splitting the jobs into shifts.

For a variety of problems and pallet numbers, the following performance measures will be determined:
1. the lower bound of the makespan (LB) using the procedure explained in Appendix 1.
2. the SQN-based makespan estimation (QM), using the SQN-procedure described in section 3.1.
3. the best makespan (BM), using the scheduling procedure described in section 3.2.

The reader should realize that the present study is a preliminary study to identify the problem and perform an initial analysis. The experiment, results and analysis are subject to a number of limitations. First, the problem setting is limited. Although a variety of demand characteristics, routings and pallets numbers are tested, the physical system (primary process) is identical in all cases. Consequently, the differences in the hardness of the problems is entirely determined by the job characteristics and not by the characteristics of the physical system. Second, only one scheduling algorithm is used. The performance of this particular algorithm is assessed in Paulli (1995). Third, only one lower bound method is used. The benefits of this specific lower bound are discussed in Paulli (1995).

The discussion section will further address the limitations and benefits of this study and address avenues of further research in this area.

Given the system described above, 50 different problems have been generated (see
Appendix 2). For each problem, 10 different pallet numbers (16 to 25 pallets) have been considered. Then, for each of these 500 problems, the three performance measures have been determined (with the lower bound obviously independent of the number of pallets).

The hypothesis is, that the ratio of QM and LB determines the hardness of the problem and consequently the ratio between BM and LB. If the problem is fairly easy (QM/LB is small), then BM is expected to be close to the lower bound. On the other land, if QM/LB is large, then BM is expected to be relatively far off the lower bound. Consequently, QM and LB can be used to estimate BM. Note that QM and LB are easy and fast to calculate, while BM requires substantially more computing time.

Figure 2 shows 50 observations, with a fixed number of pallets (16) for each problem. On the horizontal axis represents the ratio QM/LB. The solid line shows the observations of BM/LB. A slight trend can be recognized in the curve. Two problems show a considerable deviation from the trend curve, namely the problems identified as numbers 7 (considerably more difficult than expected) and problem number 2 (considerably easier than expected). An explanation for these exceptions has not yet been found. The dashed lines are two linear approximations of the observations. Line (2) is the regression curve. As can be observed from the graph, the correlation coefficient $r^2$ is not high (0.27). Line (1) is the lower bound plus a fixed percentage (1.065*LB) as an estimator of BM, where 1.065 is the average of all observations of BM/LB. This estimator will be called LB+. LB+ does not use any information regarding the complexity of the problem. The added value of LB+ can be determined by comparing the prediction accuracy of lines (2) and (1). These are shown in table 1.
Table 1. Prediction accuracy (50 observations, 16 pallets)

<table>
<thead>
<tr>
<th></th>
<th>MSE</th>
<th>MAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$BM/LB = 0.138$</td>
<td>0.00085</td>
<td>0.0229</td>
</tr>
<tr>
<td>$+ 0.864$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BM/LB = 1.065$</td>
<td>0.00116</td>
<td>0.0260</td>
</tr>
</tbody>
</table>

The added value of the queueing information over the use of $LB+$ is about 13.5% in MAD (Mean Absolute Deviation) and 36.5% in MSE (Mean Square Error). The variation is however still quite large and the performance of the linear estimator is not yet satisfactory. Apparently, more factors influence the distance of the best makespan to the lower bound.

Figure 2 shows the results from using two different linear approximations for the makespan for 16 pallets in the system considered. What happens with the relationship between the two estimation methods as the number of pallets increases? The improvement in MSE decreases for cases with more pallets (see figure 3). The improvement ratio decreases fast to about 8% for 19 pallets and then gradually further to about 4% for 25 pallets. It is noteworthy that the ratio $BM/LB$ decreases also when the number of pallets increases, as does $r^2$. Apparently, the algorithm can find better solutions, on average, as the number of pallets increases. These
solutions are then however less correlated with the difficulty of the problem expressed by QM/LB than the cases with fewer pallets. Obviously, when the number of pallets is larger, a greater number of combinatorial solutions exist and better solutions can be found for specific cases. The SQN-model does not cover this modelling aspect.

Obviously, in practice, processing times are not deterministic. Deterministic schedules are however often used in practical situations. Usually, schedules are not revised within a fixed period of time, e.g. one week. Consequently, jobs are shifted forward if they cannot be started or completed on time. Similar reactions occur when machines break down. This issue is addressed in an excellent article by Leon et al. (1994). We have briefly investigated the influence of stochastic processing times on the makespan, if the deterministic schedule is applied in a right-shift manner. If we call the right-shift makespan with Erlang-distributed processing times EM, figure 4 displays two ratios of EM and LB, viz. for Erlang(3) and Erlang (30) distributions, in order to show the effect of variance. The Erlang-distribution is characterized by the parameters \( n \) and \( \lambda \), where \( n \) is the number of exponential distributions with parameter \( \lambda \). The mean of an Erlang distribution is equal to \( n/\lambda \). If we assume that the deterministic processing times used in the scheduling algorithms are the mean of an Erlang-distributed processing time with parameter \( n \) and mean equal to the processing times, we are able to adjust \( \lambda \) to find the correct Erlang-distribution for each processing time. Since the variance of an Erlang-distribution is equal to \( n/\lambda^2 \), the variance of our processing time equals \( p^2/n \), where \( p \) is the deterministic processing time used in the scheduling algorithm. This implies, that as \( n \) increases, the variance of the processing times decreases. The
deterministic pattern can still be recognized in the EM/LB graphs. Consequently, a lightly ordinal relation exists between EM and BM, with a light trend of the QM/LB ratio as parameter in the trend. Again, the patterns are very weak and full of noise, such that conclusions cannot be based upon these observations. Apparently not only the algorithm determines the increase in makespan due to the introduction of stochastics (Leon et al., 1994), but so does the specific problem.

5. Discussion

The initial study described above discusses possibilities to improve aggregate planning of flexible manufacturing systems. We compared the results of an SQN-based model with the results of a deterministic scheduling algorithm.

For both aggregate estimators (QM and LB+) we need detailed data such as the routes and demand levels of the products. The QM and LB+ estimates are however very easy to compute as compared to the detailed near-optimal schedule. This is mostly due to the fact that the scheduling algorithm also determines the sequence and allocation of the jobs on the machines. This information is not necessary to make the aggregate decision.

As discussed above, this paper does not intend to fully review the use of queueing models in makespan estimation. It does however introduce a new area in operations research, since queueing models are widely used for loading decisions and detailed scheduling models are increasingly used for operational scheduling decisions. Comparisons have not been made before.

Since it is an initial study, it has some limitations, which should be taken into account when assessing the results of this study. First, only one FMS is considered. As of yet, it is unclear whether a different FMS structure will lead to different results. Second, although some relations have been found between the investigated parameters, the correlation is still unsatisfactorily low. Finally, the influence of the scheduling algorithm itself has not been investigated. In the comparison with the right-shift schedule, it is known to be a relevant factor (Leon et al., 1994).

The analysis of the limitations of this study and expansion of its scope opens new areas of research in operations scheduling and control. Although the results are not entirely positive, there are however some indications which look promising. Since there is a need
in practice for better aggregate models, we should continue research in this area.

6. References


Sawik, T., "Modelling and Scheduling of a Flexible Manufacturing System", European


Tempelmeier, H., and H. Kuhn, Flexibele Fertigungssysteme, Berlin, Germany: Springer-Verlag, 1992 (in German).


Appendix 1 Lower Bound Determination

\[
\begin{align*}
LB & = \max\{LB_{A}, \ldots, LB_{D}, LB_{A1}, \ldots, LB_{D2}\} \\
LB_{i} & = \frac{\text{the sum of the processing times of the operations, that have to be performed on one of the machines in machine group } i}{\text{no. of machines in machine group } i} \\
i & \in \{A, \ldots, D\} \\
LB_{j} & = \text{the sum of the processing times of the operations, that can only be performed on machine } j \\
j & \in \{A1, A2, A3, B1, B2, B3, C1, C2, D1, D2\}
\end{align*}
\]

So if the routing flexibility is high, \(LB_{j}\) are likely to be small and then one of \(LB_{i}\) will be the lower bound \(LB\). If the routing flexibility is low, one the machines is likely to be the bottleneck (or at least have to operate more, a priori) and then one of \(LB_{j}\) will be the
lower bound LB.

Appendix 2 Generation of Test Problems
The test problem were generated in the following way.

- Choose the no. of product type between 6 and 10.
- Choose the no. of operations for a product of a given product type between 4 and 8.
- Choose the demand for the product types between certain limits, defined such that the load on the system is on average 80%.
- Chose a target load level for the 4 machines groups between 70% and 90%.
- Generate processing times between 5 and 50 and choose a machine group on which the operations must be performed.
- Adjust all the processing times so that the target levels are (almost) met, i.e. multiply by target load divided by actual load for the machine group on which the operation is performed and then round off.
- Choose one of the machines in the machine group for performing the operation.
- With a given probability let the other machines in the machine group be able to perform the operation. This probability is doubled if the processing time of the operation multiplied by the demand for the product type is larger than the average processing time (27.5) multiplied by the average demand.

We believe that this is a rather realistic way of setting up the machines - if an operation is relatively time-consuming (either because it is long or because it is frequent) one will try to make it possible to perform it on several machines.)