Project Scheduling Approach to Production Planning

A. Márkus[1] (2), J. Váncza[1], T. Kis[1], A. Kovács[2]
[1]Computer and Automation Research Institute, Hungarian Academy of Sciences
[2]Faculty of Electrical Engineering and Informatics,
Budapest University of Technology and Economics, Budapest, Hungary

Abstract
We suggest a novel approach to modelling and solving production planning (PP) problems in make-to-order production environments. This approach unifies the capacity and the material flow oriented aspects of PP. Orders are modelled as projects that compete for limited resources. The projects consist of networks of variable-intensity activities which may require several resources at the same time. The goal is to generate production plans that satisfy all the temporal and resource constraints and minimize additional or external resource usage over the planning horizon. The model is applicable at different aggregation levels of PP, as we show with two real-life case studies. Although the capacity and material flow aspects of PP are coupled, a special solver developed for the project model efficiently solves problems of real-life sizes.

Keywords:
Production, planning, project

1 INTRODUCTION

Production planning (PP) matches future production load and capacities by generating plans that determine the flow of materials and the use of resources over a given planning horizon.

Finding an appropriate match between load and capacities is a chicken-and-egg problem. On the one hand, the actual load – and the corresponding flow of materials – should fit to the available resource capacities. On the other hand, the level of capacities should be determined so as to meet the demand over time. The problem is usually settled by fixing either the load or the capacities [1].

However, this approach can hardly be taken nowadays when the traditional boundaries of companies are getting dissolved. In the era of supply chain management, virtual enterprises and production networks the tasks of PP may cross the organizational boundaries of the firms. Decisions on the use of resources should concern both internal and *external* capacities; the internal flow of materials should be synchronized with the incoming and outgoing flows [2]. All this makes the PP problem extremely hard to solve. Conversely, the complex situations call for efficient, robust decision support methods at each node of a production network [2,3]. Hence, there is a need of intuitive and flexible models *and* fast, reliable solution techniques that scale-up well also to large problem instances. An important practical requirement for any new method is that it should be able to work by using data stored in existing – so-called legacy – production information systems.

There are long-standing recipes to handle complexity. *Aggregation* removes details in the representation of products and orders, production processes, resource capacities, and time [1,4]. Similar problems formulated with more details are limited by shorter planning horizon. The solution is generated in a process where higher level solutions provide constraints to lower level problems. *Decomposition* separates PP problems into a resource and a load oriented subproblem. Decomposition is usually applied on several levels of aggregation. E.g., decomposed planning functions are traditionally termed as master production scheduling and material requirements planning on the side of the load, whereas rough-cut capacity planning and capacity requirements planning on the side of resources [1].

Below we suggest a novel approach to modelling and solving long- and medium-term capacity and production planning problems. Our method is based on a generalized version of the *resource-constrained project scheduling* problem [5] and unifies the resource and the material flow oriented aspects of PP. By introducing *activities*, we handle together resource and temporal constraints. Hence, planning can work without using lead time estimates that, in a turbulent environment, cannot represent individual orders any more [6]. The project-based approach captures the strong goal-oriented nature of *make-to-order* and *engineering-to-order* production.

The paper is structured in the following way: In Section 2 the project model is presented along with the main principles of the solution technique. Then we describe how to solve PP problems in two different production environments by taking the suggested project-based approach (Sections 3 and 4). Finally, conclusions of the two case studies are drawn in Section 5.

2 THE BASIC MODEL

2.1 Resource-constrained project scheduling

Resource-constrained project scheduling problems are concerned with scheduling a number of discrete activities, each requiring some resources. Constraints due to the limited capacities of resources and precedence relations between the activities are

prescribed. The classical model assumes fixed activity durations and a constant rate of resource usage during the entire processing of every activity [5,7]. However, in *aggregate* planning the above assumptions cannot be taken and there is also no need to generate detailed solutions for future periods that will certainly be different to what is anticipated. Hence, after studying real PP problems we extended the classical model by allowing (1) *preemption* of activity execution, (2) *variable-intensity* activities, and (3) *continuously* divisible resources.

2.2 Activities and resources

An instance of the problem is given by a set $N = \{1,...,n\}$ of *activities*, a set $R = \{1,...,r\}$ of continuously divisible and renewable *resources*, and a directed acyclic graph $D = (N, A)$ representing *precedence constraints* among the activities. Each activity $i \in N$ must entirely be processed within its *time window*: between its earliest starting time e^i and deadline d^i.

Each activity may require the simultaneous use of some *resources*. The entire processing of activity i requires a total of r_k^i units of resource k, for each $k \in R$. The *intensity* of each activity may vary over time, and the resource usage is proportional to the intensity. If x_t^i is the intensity of activity i in time period t, then it requires $r_k^i * x_t^i$ units of resource k in that period. However, the intensity of executing an activity is limited: in any time period $t \in [e^i, d^i]$ at most $a^i \leq 1$ fraction of activity i may be completed.

The *capacity* of each resource $k \in R$ is fixed period by period over the horizon. In each time period t, a certain internal capacity of each resource k is available. Internal resource capacities can be used free of charge. Additional external capacities are also available, but at the expense of some cost per resource units.

2.3 Generalized precedence relations

The most typical precedence constraint between a pair of activities i, j prescribes that activity i must finish before activity j may start. However, in practical problems a precedence relation between a pair of activities may have the following form: start activity j only if 25% of activity i has been completed, or the last 30% of activity j can be done only after 60% of activity i was completed.

Our model supports the above kind of generalized precedence relations. Note that there could be several relations between a pair of activities i, j.

2.4 Optimization criteria

In our basic model the *cost of using external resources* is to be minimized. This optimization criterion is motivated by practical applications that will be described in Sections 3 and 4.

We note here that classical optimization criteria, like project *duration*, *maximum tardiness* or *weighted tardiness* fit also in the proposed framework.

2.5 Problem statement

The problem consists of determining for each activity i an intensity x_t^i in each time period $t \in [e^i, d^i]$ such that $0 \leq x_t^i \leq a^i$, $\sum x_t^i = 1$, all the precedence constraints among the activities are fulfilled, the resource demands do not exceed the resource availabilities in any time period, and the total cost of using external capacity is minimized.

For a simple example, see Fig. 1 with data of two projects. Activity 1 and 2 are linked by a precedence relation. The resources have a unit capacity.

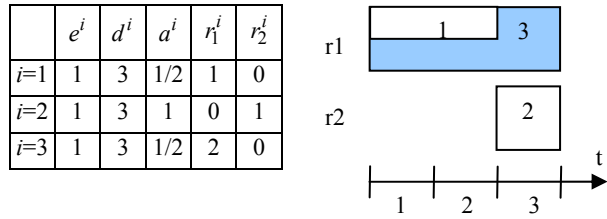|	e^i	d^i	a^i	r_1^i	r_2^i
$i=1$	1	3	1/2	1	0
$i=2$	1	3	1	0	1
$i=3$	1	3	1/2	2	0

Figure 1: Data of two sample projects and the solution.

The above problem has been formalized as a mixed integer-linear program. A detailed analysis of this mathematical program has shown that the problem is NP-complete in the strong sense. However, the analysis resulted also in a linear programming re-formulation with some cutting planes. The solution method uses them in a branch-and-cut algorithm that finds optimal solutions. Our extensive numerical tests confirmed that branch-and-cut is a viable approach for solving even very large problem instances. For details of the analysis and solution method, see [8].

3 APPLICATION TO PRODUCTION AND CAPACITY PLANNING

3.1 Background

In Case Study 1, we consider a factory with a long tradition in manufacturing special, one-of-a-kind equipment for producing a mass product for household use. The factory handles accepted orders as *projects* that must be completed between their release dates and deadlines.

The *internal* resources are well-organized and stable. At subcontracting partners, there are *external* capacities for all resources, but for a given unit costs. The crucial problem is to determine the timing and resource assignments of the activities of all the projects so as to satisfy the temporal and resource constraints, and to minimize the cost of external resource usage. Production and capacity planning should be supported in an integrated way, at two levels of aggregation:

- On the long term, with a 1-1.5 year horizon, by considering the various departments (like mechanical design, components machining, mechanical assembly, electric design, electric assembly, installation, etc.) as resources.

- On the medium term, with a quarter horizon, by considering the groups of machine and labour resources of components machining.

The planning problems should be solved on both levels with a rolling horizon, since the actual states of the projects are reported week by week, and the planners have to make new plans that take into account the deviations and new requests as well.

3.2 The planning problem

In terms of our model presented above, each resource has an inside capacity (given in work-hours per week) in each week of the horizon. Inside capacities can be used free of charge. Further on, additional capacities of subcontractors can be used at some extra cost. However, these parameters may vary from week to week. The actual values of capacities and prices are influenced by several factors such as maintenance, holidays, urgency, market position etc.

Each project consists of several *activities*. Each activity may require a number of different resources. Each

activity can be executed with time-varying intensity and can be broken by assigning zero intensity for (some) week(s), with non-zero intensity before and after. Activities and their relations are modelled in the same way on both planning levels. Hence, below we refer to the specific level only if necessary.

Activities are described by the *amount of work* to be done by the specific resources. On the higher level, each activity requires a single resource, and any resource is used at most once in a project. On the detailed level of planning, activities require several resources at the same time, typically machine and labour.

Each activity should be performed within its *time window*. Initially, the window of an activity is defined by the earliest start date and deadline of the project it belongs to. By considering the precedence relations between activities of a project, time windows can be narrowed. Planning at the higher aggregation level determines the time windows of activities on the detailed level.

In each week, any resource may be partially assigned to several projects in parallel: e.g., supposing a 40 hours week, 5 hours may go to Project1, 15 hours to Project2 and 18 hours to Project3, the remaining 2 hours left free.

The volumes of work may be fractional numbers. E.g., the execution of 10 units of work in four consecutive weeks can be uniform as 2.5 + 2.5 + 2.5 + 2.5 or varying as 1.0 + 0.0 + 2.5 + 6.5. The question is what should be the *intensity* of each activity in each period of its time window. To better approach engineering reality, the maximal allowable intensity of each activity is also represented. E.g., even if the factory has capacity of 200 units per week for assembling components, an activity that needs 200 units of assembly can hardly be performed in a single week. Instead, it should be spread over the planning horizon. As this example shows, sharing resources among several projects may be not just a planning trick but a matter of engineering convenience.

The activities of a project are connected by *precedence relations*. The projects are independent, hence there are no precedences between the activities of different projects. However, we need generalized precedences to capture overlapping activities like below:

- Machining of some parts in the project may start while some other ones are still being designed.

- As soon as there is a large amount of parts in rough-cut state (such as 25% of the total volume required in the project), they may be forwarded to fine-cutting.

- A small percentage of electric design is postponed for final corrections after assembly.

These types of precedences fit well into the mathematical model outlined above. The variable-intensity, breakable activity model, together with the generalized precedence relations, can describe complex temporal relations – including partial overlaps – between two activities.

Summing up, the objective of production planning in Case Study 1 is to minimize the total cost of external resources while keeping all the project deadlines and the constraints discussed here. Solving the problem means determining the intensity of each activity in each time period of its time window.

3.3 Solution and experiments

We have implemented an environment for testing the representation power of the model and the efficiency of the solution method. The model was adequate with the expectations of both the managers and the engineers who should use the planner system.

Further requirements concerning the profile of the intensities of activities have been satisfied by post-processing the results.

Extensive tests have been carried out with problem instances of real-life size and project mix. On the higher aggregation level typically 5 to 10 resources had to be dealt with and were required by some 150-250 activities. Problems on the lower level are of the same complexity since they have more resources, but due to the shorter planning horizon, less activities. On the average, the algorithm could solve problems of such size within 30 sec CPU time on a Pentium 4 (1.6 GHz) personal computer. The branch-and-cut search algorithm explored typically 0-100 nodes (i.e., solved some problem instances *without search*). Hence, thanks to its efficiency, the solver is applicable as the engine of an interactive, decision support PP system. Fig. 2 gives an example of how activities of various projects share a particular resource over a year's planning horizon.
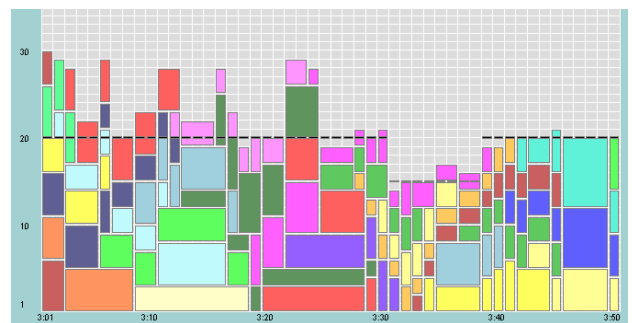

Figure 2: Planned load of a resource, with external needs above the internal capacity limit (dashed line).

4 APPLICATION TO MATERIAL AND CAPACITY REQUIREMENTS PLANNING

4.1 Background and problem statement

The factory of Case Study 2 manufactures mechanical products of high value by using machining and welding centers, assembly and inspection stations and some highly specialized machines. The products have tree-structured *Bill of Materials* (BOM). The root of the tree is the end product, while the leaves are purchased items. Each internal node corresponds to a make item, to which a sequence of manufacturing operations is assigned. These operations are described in the *routings*. Though products of different families differ also in their production technology, the problem cannot be decomposed: the sets of resources used for producing different families are not disjoint. Production is performed in a *make-to-order* manner, where deadline observance is an absolute must, even for unpredicted orders. Since quality assurance is a key issue, tests may result in extra adjustment operations. Hence, production technology contains a factor of uncertainty in itself. The objectives in Case Study 2 have been as follows:

- the development of a medium-term capacity and production planner;

- the creation of a short-term job shop scheduler system that takes into account all the specific features of the production;

- the development of a simulation model of the factory that represents the non-deterministic events, tests and validates the results of the scheduler system.

Here we focus on *medium-term capacity and production planning*. At this level of aggregation, the PP problem is stated as follows:

- given (1) the orders to be completed within the horizon (including open shop orders) (2) the BOMs of products, (3) routings that specify the sequence, processing time and the resource requirements of operations, and (4) the available resource (machine and worker) capacities;

- determine for each unit of the planning horizon (1) the operations that should be performed that time, and (2) the eventual extra capacities needed.

Below we show how this problem, that integrates *capacity* and *material requirements planning*, can be solved by taking the project-based approach.

4.2 Application of the project model

We consider machines and workers as *resources*. Both machine and worker capacities vary in time but are known in advance throughout the planning horizon: this *resource calendar* is given by the higher-level capacity planning. Hence, the resource parameters of the model presented in Sect. 2 are known for each resource.

Each product order whose time window falls (even partially) within the planning horizon is considered as a *project*. The definition of activities is, however, not as straightforward as it was in the previous case. The description of activities should be based on data available on the detailed structure and technology of the products. Hence, the project model is built on the BOMs and routings. The one-to-one mapping of operations to activities is not convenient, because the problem instances would contain too many activities. Instead, we *merge operations* into activities and – respecting the ordering in the BOMs and routings – define precedence relations between some of the activities. The operations within an activity may call for different resources, hence an aggregated activity may require several resources. Principles of this transformation are as follows:

- The total resource demand of an activity should not exceed the internal capacity limit per time unit.

- The number of activities in a project should be as small as possible.

- The depth of the project's precedence tree should be minimal (so that it contain as many parallel branches as possible).

Since the BOMs and the sequential routings define a tree-shaped production process, the above transformation was formalized as a tree-partitioning problem. The last two requirements are in conflict, hence a trade-off must be found. Pareto optimal solutions are generated by a polynomial-time, dynamic programming method.

The approach was validated and tested with real-life data. First, projects were generated from existing routing tables and BOMs, then, using the resource calendars, the planning problem was solved on a 15 weeks horizon, with a time unit of one week. Then, the first two weeks' production plan was passed to a constraint-based finite job-shop scheduler that worked with a 0.1 hour time unit. Since the activity model aggregates operations, there is no guarantee that a solution of the project model leads to a feasible (i.e., executable) discrete schedule on the finer level. A solution with an adequate performance on both levels is established iteratively, by the cooperation of the project-based planner and the discrete job-shop scheduler. We built also a discrete-event model of the factory to test schedules against uncertainties.

Typical problem instances had up to 500 orders (projects). The projects consisted of 2 to 10 activities that were merged from 30 to 430 operations. Altogether cc. 150 machines and worker resources were considered. Generating the project model was a matter of seconds, and using the same solver as in Case Study 1, its solution took never more than one minute.

5 CONCLUSIONS

In this paper we have presented a new model for project scheduling and applied it to integrated capacity and production planning problems. The method is able to support planning even if no detailed information on production technology is available. Conversely, it is applicable also for detailed planning when all product and production technology information are to be exploited. The model can handle design, engineering, as well as traditional production processes.

Having introduced variable-intensity tasks and continuous resources, the model can be solved by customized mathematical programming methods very efficiently. Hence, it can be applied in a dynamic setting when re-planning is initiated by unexpected changes. The approach is an alternative to control-theoretic planning methods that, remarkably, integrate the load and capacity oriented aspects of PP by working on continuous models [9,10]. The fast response time provides opportunity to analyze various scenarios and to drive the engine of a decision support tool that can be used in an exploratory mode of operation. Our current goal is to develop a complete workflow model [11] whose key components are driven by the project-based planner.

6 ACKNOWLEDGMENTS

7 REFERENCES

[1] Graves, S.C., Rinnooy Kan, A.H.G, Zipkin, P.H. (eds.), 1993, Logistics of Production and Inventory, Elsevier.

[2] Wiendahl, H.-P., Lutz, S., 2002, Production in Networks, Annals of the CIRP, 51/2:1-14.

[3] Váncza, J., Márkus, A., 2000, An Agent Model for Incentive-Based Production Scheduling, Computers in Industry, 43:173-187.

[4] Maropoulos, P.G., McKay, K.R., Bramall, D.G., 2002, Resource-Aware Aggregate Planning for the Distributed Manufacturing Enterprise, Annals of the CIRP, 51/1:363-366.

[5] Demeulemeester, E.L., Herroelen, W.S., 2002, Project Scheduling: A Research Handbook, Kluwer.

[6] Wiendahl, H.-H., Roth, N., Westkämper, E., 2002, Logical Positioning in a Turbulent Environment, Annals of the CIRP, 51/1:383-386.

[7] Kumara, S.R.T, Lee, Y.-H., Chatterjee, K., 2002, Distributed Multiproject Resource Control: A Market-Based Approach, Annals of the CIRP, 51/1:367-370.

[8] Kis, T., 2002, A Branch-and-Cut Algorithm for Scheduling Projects with Variable-Intensity Activities. Submitted to Mathematical Programming.

[9] Wiendahl, H.-P., Breithaupt, J.-W., 2001, Backlog-Oriented Automatic Production Control, Annals of the CIRP, 50/1:331-334.

[10] Scholz-Reiter, B., Freitag, M., Schmieder, A., 2002, Modelling and Control of Production Systems Based on Nonlinear Dynamics Theory, Annals of the CIRP, 51/1:375-378.

[11] Lutters, D., Mentink, R.J., van Houten, F.J.A.M, Kals, H.J.J, 2001, Workflow Management Based on Information Management, Annals of the CIRP, 50/1:309-312.