1 Introduction

Scheduling is a decision-making process used in many manufacturing and services (incl. logistics services like transportation and distribution) industries. “It deals with the allocation of resources to tasks over given time periods and its goal is to optimize one or more objectives” (Pinedo 2008).

The resources and tasks in different industries can have different meanings: the resources may be machines in a factory, workers in a distribution centre, container cranes in a port, and so on. The tasks may be operations in a production process, customer orders to be delivered in a distribution centre, containers to be transported in a port, and so on. The objectives can also take many different forms. One objective may be the minimization of the completion time of all tasks and another may be the minimization of the number of tasks completed after their respective due dates.

This thesis focuses on scheduling in production processes and presents a new progress in this field. Therefore, from now on, the resources are called machines and the tasks are called jobs in the thesis. However, the results of this thesis are not limited to production processes but also applicable to logistics and other services processes, which can be modelled by the same scheduling problem that this thesis considers.

1.1 The role of production scheduling

In the supply chain (SC) planning matrix from Meyr et al. (2005) (shown in Figure 1-1), they classify the production scheduling as a short-term planning module used for *machine scheduling* and *shop floor control* in a SC planning process. It closely cooperates with three other modules: *Material Requirements Planning* (MRP), *Production Planning* and *Master Planning*. The history of using production scheduling can be tracked back to the 1910s. Gantt (1916) explicitly discusses scheduling, especially in the job shop environment, and Gantt (1919) develops a line of charts (known as Gantt charts) to help production planer to assign resources to tasks and control the production progress (Herrmann 2006). In comparison to scheduling, the other modules came into existence much later, with the advent of computers in the field of production planning. For instance, MRP was developed in the 1960s and Production Planning and the Master Planning were developed as parts of the Enterprise Resource Planning (ERP) system still later.

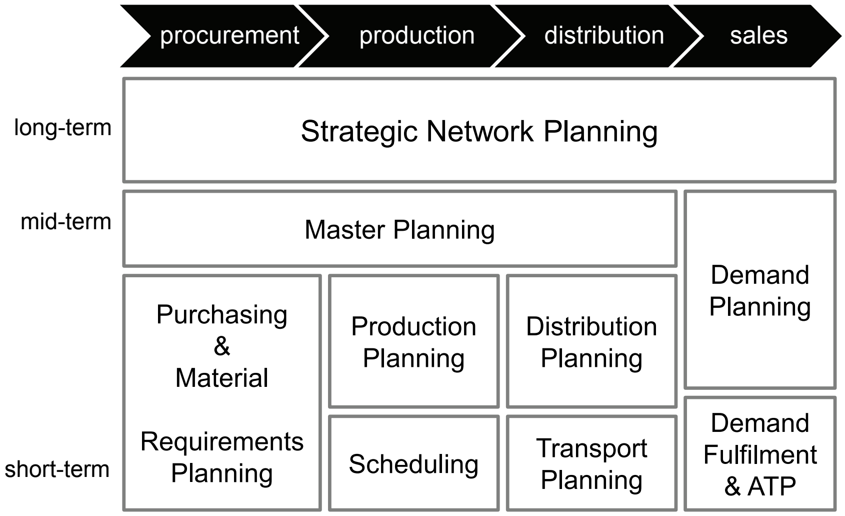

Figure 1-1. Software modules covering the supply chain planning matrix (Meyr et al. 2005, p. 105)

Figure 1-2 demonstrates the interaction between the scheduling module and the other modules in the SC planning matrix. The *Production Planning* (PP) module receives the information about orders or demand forecasts from the *Demand Planning* module via the *Master Planning*. Afterwards, the PP determines jointly with the MRP the quantity of jobs (shop orders) and their release dates and due dates for the next scheduling period. With the inputs from the PP, the *Scheduling* module generates a production schedule. An additional *dispatching* module closely connected to the shopfloor management system (now called Manufacturing Execution Systems (MES)) dispatches the job order for a machine according to the schedule and informs which job should be loaded onto the shopfloor next. There are generally three different implementations of the scheduling module:

Manual: The production schedulers make schedules by means of a chart and their own experiments. There are several problems of this implementation. Firstly, the schedulers cannot consider the information from the shopfloor and the other systems in real-time. Secondly, most scheduling problems are too complicated to be solved by a human. Such a “production scheduling system is a broken collection of independent plans that are frequently ignored, periodic meetings where unreliable information is shared, expediters who run from one crisis to another, and ad-hoc decisions made by persons who cannot see the entire system” (Herrmann 2006).

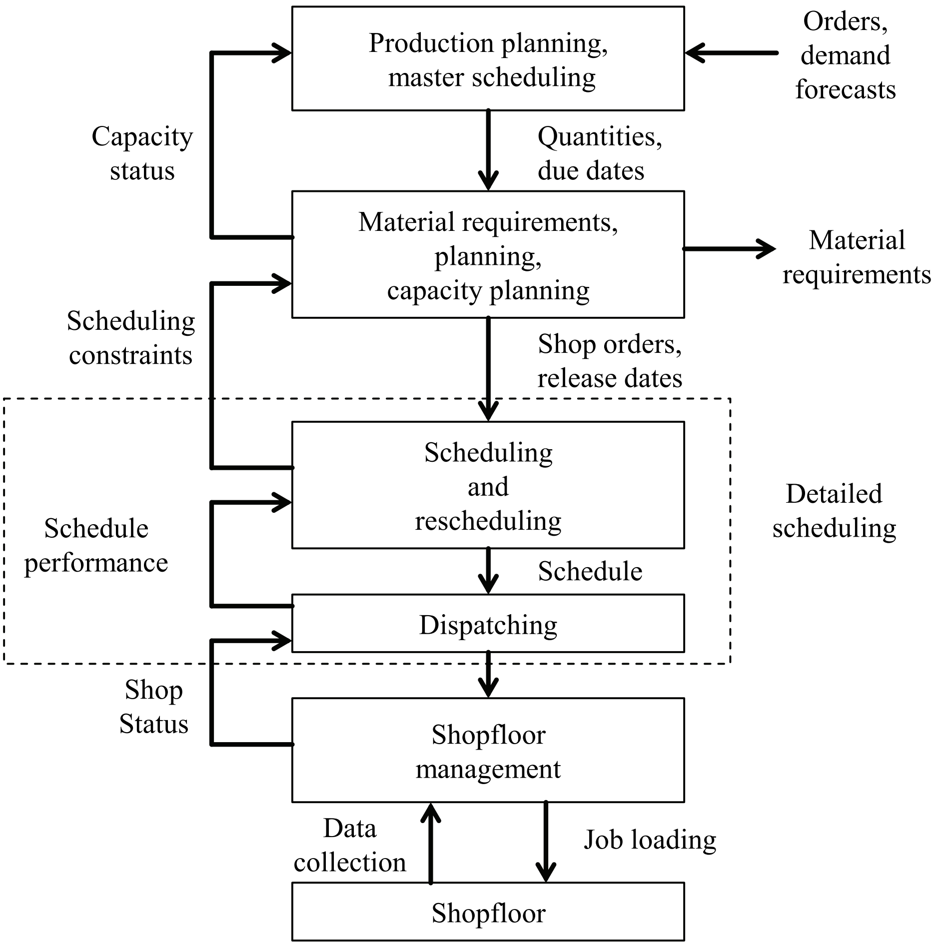

Figure 1-2. Information flow diagram in a manufacturing system (Pinedo 2008, p. 5)

Dispatching rules: (also called priority rules) are rules to sequence the jobs waiting at a machine. “Such rules are based on attributes of each job and may use simple sorting or a series of logical rules that separate jobs into different priority classes” (Herrmann 2006). The job with the highest priority will be processed next, when the machine becomes available. The advantages of the dispatching rules are aplenty. Firstly, they are easy to implement. Secondly, they are very efficient and can be scaled well to large problem instances. Thirdly, they can find reasonably good solutions. Due to these advantages, dispatching rules are widely used in industry. However, the main disadvantage of these rules is that they do not guarantee an optimal solution, because a dispatching rule only considers limited information from several jobs and only for one machine. The consideration of the whole production systems is missing, especially for multi resource systems like job shops and flow shops.

Scheduling algorithms: Scheduling algorithms not only consider the local optimum for one machine (like dispatching rules) but also concentrate on finding the global optimum for the whole production system. In the last decade and a half, rapid advancement in information technology has made the development of complex algorithms, from mathematical programming to artificial intelligence, possible. Thus, development of better scheduling algorithms has become more important.

1.2 Challenges for implementing scheduling algorithms in industry

Currently, researchers and manufactures pay more and more attention to develop scheduling algorithms. However, there are still challenges they are facing when implementing scheduling algorithms in industry:

NP-hard: Most scheduling problems are NP-hard. It means, algorithms that can find optimal solutions to these hard problems in a reasonable amount of time are unlikely to exist (Herrmann 2006). Even developing algorithms which can reach the near-optimum is complicated.

Problem size in practice: The size of a scheduling problem instance (i.e. the number of jobs and machines) in practice is usually much larger than the instances discussed in the literature. Hence, a scheduling algorithm must have an appropriate run-time and has to be scaled well to larger problem instances.

Dynamics and uncertainties: A production process is dynamic process. It contains uncertainties (like disturbances and changes of job information), which lead to deviations between the schedule and the shopfloor progress. Therefore, the schedule has to be updated regularly. This process is called *rescheduling*. In the rescheduling situation, a scheduling algorithm has to consider additional dynamic aspects like release dates of the jobs for the next period, the *work in process* (WIP) and the machine available dates.

Limited time for scheduling: As a short-term planning process, manufactures usually regenerate the schedule once each day or once each working shift. The available time for a scheduling is the transition time between two shifts.

Flexibility and extensionality: When in the future the production environment of the firm has been changed, e.g. through new machine layouts or by purchase of machines with new features, the scheduling algorithm should be flexible in such a way that it can be simply adapted and extended to fulfil these changes.

1.3 Expectations

Regarding the challenges mentioned above, the objective of this thesis is to present a scheduling algorithm, which solves the dynamic job shop scheduling problem and fulfils the following requirements:

Effective: It should solve this NP-hard scheduling problem optimally or near-optimally.

Problem size in practice: It should scale well to larger problem instances of a realistic industrial size.

Dynamics: It should be able to solve job shop problems with release dates for jobs and machines and consider WIP.

Efficient: It should have a low run-time complexity and consider the balance between the performance (solution quality) and the computational time.

Flexible and Extendable: It should be easily adjustable in order to meet different situations of the considered production environment. Furthermore, it should be simply extended to meet changes or more complex situations of the considered scheduling problems.

1.4 Motivation and goals

The job shop scheduling problem is widely used for modelling manufacturing processes, for instance for the engineer-to-order or make-to-order processes in the machinery and plant engineering industry, aeroplane and aerospace industry, as well as semiconductor industry. Therefore, it has received great attention from scheduling researchers.

Due to the increasing market competition, manufacturers currently pay more attention to due date related production objectives in order to fulfil the customers' demand on punctuality (Pinedo 2008; Gordon et al. 2012). Consequently, the job shop scheduling problem with optimality criteria related to due date performance has become increasingly important.