Int. J. Production Economics 133 (2011) 439–450

Contents lists available at ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

An iterative approach to item-level tactical production and inventory planning

Feng Tian [a], Sean P. Willems [b,*], Karl G. Kempf [c]

[a] College of Business & Public Administration, Governors State University, University Park, IL 60484, USA
[b] School of Management, Boston University, Boston, MA 02215, USA
[c] Decision Technologies Group, Intel Corporation, Chandler, AZ 85044, USA

ARTICLE INFO

Available online 21 August 2010

Keywords:
Inventory
Production
Hierarchical planning
Supply chain management

ABSTRACT

In this paper, we propose an iterative approach to jointly solve the problems of tactical safety stock placement and tactical production planning. These problems have traditionally been solved in isolation, even though both problems operate in the same decision making space and the outputs of one naturally serve as the inputs to the other. For simple supply chain network structures, two stages and one or many products, we provide sufficient conditions to guarantee the iteration algorithm's termination. Through examples, we show how the algorithm works and prove its applicability on a realistic industrial-scale problem.

1. Introduction

The strategic-tactical-operational framework developed by Anthony (1965) is ingrained in the operations-management lexicon. In a classic manifestation of this framework, determining how much production capacity to have and where to have it are strategic decisions, determining how to allocate production capacity to product families is a tactical decision, and producing an item-level production schedule is an operational decision. Not only do these decisions operate at different frequencies (i.e., a company does not evaluate its capacity acquisition strategy on a weekly basis), they also operate with different levels of scope and granularity. For example, setting the production schedule for the next day requires a precise statement of every item at each location while a biannual capacity acquisition study aggregates items beyond the product family to product types that represent major market segments by manufacturing origin.

The literature that addresses supply chain aspects of Anthony's (1965) framework is vast. Even with attention limited in scope to production–inventory problems, researchers must make hard choices to limit the scope and granularity of their models. We will restrict our attention to the large subset of the literature that models the interaction of tactical production planning with a number of other production–inventory problems. This subset can be divided into approaches that break the problems into a hierarchy of decisions and approaches that solve a monolithic unified model.

Hax and Meal (1975) propose a hierarchical solution procedure that spans capacity planning through detailed scheduling. The hierarchical planning approach relies on aggregating data for higher-level decisions and having the optimal decisions from each higher-level model serve as a constraint for the next-lower model in the hierarchy.

Bitran et al. (1981) solves the production allocation and item-level scheduling problems for a multiple-item single-echelon system. Family and item disaggregation subsystems are both represented by means of knapsack problems. Bitran et al. (1982) expands this approach to a two-echelon system. While the application of the framework to the two-echelon setting is conceptually straightforward, problem-specific knowledge must be exercised to determine the appropriate aggregation structure. Specifically, the solution to the aggregate top-level model does not ensure the existence of a feasible disaggregation for the item-level problem. To ensure feasibility, it is necessary to either add sufficient conditions at the aggregated planning level (Gfreer and Zapfel, 1995), or apply an iterative scheme in the hierarchical structure (Jornsten and Leisten, 1995).

Billington et al. (1983) and Bradley and Arntzen (1999) are representative of monolithic approaches. Billington et al. (1983) simultaneously determine the stage lead-times and the item-level production plan. To ease the computational burden, product structure compression is employed to collapse stages that do not influence the resulting solution. Compression works well in cases where only a few resources are constrained. Bradley and Artzen develop a monolithic mathematical program to address strategic capacity acquisition, tactical production planning, and operational

* Corresponding author. Tel.: +1 617 353 2287; fax: +1 617 353 4098.
E-mail addresses: f-tian@govst.edu (F. Tian), willems@bu.edu (S.P. Willems), karl.g.kempf@intel.com (K.G. Kempf).

0925-5273/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpe.2010.07.011

scheduling. The decision variables are capacity investments, raw material purchases, and the production schedule. The objective function maximizes return on assets. For this class of modeling, demand is deterministic so inventory represents time-phased imbalances between production and demand. Neither model explicitly considers setting safety stock levels, although exogenously determined safety stocks could be incorporated as constraints. Spitter et al. (2005) and Fandel and Stammen-Hegene (2006) are indicative of the advances in this line of modeling. Spitter et al. (2005) is similar in spirit to Billington et al. (1983) but allows capacity for an order to be allocated any time during the leadtime. Fandel and Stammen-Hegene (2006) improve the production plan by considering general lot sizing and scheduling across multiple machines.

Byrne and Bakir (1999) adopts a hybrid simulation-analytical approach to protect a production plan against operational sources of variability. A linear program generates an optimal production plan and then a simulation verifies the feasibility of the production levels. The solution procedure adjusts capacity between successive iterations until capacity constraints are satisfied. Kim and Kim (2001) propose an extended linear programming model and include more information during iterations for a similar hybrid approach. Numerical analysis shows that their approach can find a better solution in fewer iterations than Byrne and Bakir (1999).

De Kok and Fransoo (2003) present a problem to coordinate the release of materials and resources across a multi-echelon network. They refer to this as the supply chain operations problem (SCOP) and present two solution methods. One approach formulates a linear program (LP) that assumes starting inventories are zero. They then conduct a simulation to determine the appropriate safety stock levels to support the plan and then rerun the linear program. The second approach assumes synchronized base stock (SBS) policies and analytically computes the resulting base stock levels. For a set of test problems, the SBS approach outperforms the LP approach.

While Byrne and Bakir (1999), Kim and Kim (2001), and De Kok and Fransoo (2003) are notable exceptions, the majority of the literature does not focus on the determination of safety stock inventory. Hax and Candea (1984) is indicative of the more standard approach where tactical production planning problems and detailed operational scheduling are clearly laid out with established linkages but safety stock is determined exogenously and at best serves as a constraint to production planning and scheduling models. Maxwell et al. (1983) explicitly recognize this problem and propose a three-phase modeling framework to recognize the relationship between lead time, capacity, lot sizes, and safety stock. They propose phase one as creation of the master production plan, phase two as planning for uncertainty, and phase three as real time resource allocation. Safety stock setting is the key problem in phase two since it provides protection for the created master production plan.

Our work takes a different philosophical perspective, conceptually outlined in Kempf (2004). In effect, this research approach is iteratively solving the phase one and phase two problems outlined by Maxwell et al. (1983). We propose a procedure to iteratively solve two optimization problems: the tactical problem of production planning and the tactical problem of safety stock placement. The value of this approach is it integrates two well-developed research streams, allowing the joint solution to overcome the limitations of each individual approach while simultaneously preserving the optimality, within constraints, of each individual solution.

Each research stream has made significant advancements in isolation. The area of tactical safety stock optimization is summarized in Graves and Willems (2003). In brief, tactical safety stock optimization seeks to optimize inventory levels across the multi-echelon supply chain. To accomplish this, these approaches must make additional assumptions and settle for heuristic solutions relative to the exact solutions that can be derived when the problem scope is limited to single-stage inventory problems. On the positive side, papers including Billington et al. (2004) and Bossert and Willems (2007) document that these models have been successfully applied in practice.

The area of tactical production planning has a vast associated literature. Beyond the articles referenced earlier, summary overviews are provided by Shapiro (1993) and Fleischmann and Meyr (2003). For our purposes, we are concerned with linear-programming based approaches that minimize the sum of production cost, inventory cost, and penalty cost over a tactical horizon that often measures 12–24 weeks. A specific example of a relevant formulation is presented in Bean et al. (2005).

The rest of the paper is arranged as follows: Section 2 describes the iteration algorithm. Section 3 establishes termination criteria for a two-stage single-echelon network producing either one or N products. Section 4 shows the implementation of the algorithm for a realistic industrial-scale planning problem. Section 5 concludes and describes future research.

2. An iterative algorithm

The supply chain is modeled at the SKU-location level as a graph with node set N and arc set A. Every stage corresponds to a processing function. Examples include transportation from one location to another, manufacturing, and placement in a warehouse to satisfy demand. Arcs denote the precedence relationship between stages. We will find it useful to partition N into three disjoint sets: N_S, N_I, N_D. The set of supply stages, N_S, have no incoming arcs and demand stages, N_D, have no outgoing arcs. The set of intermediate stages, N_I, each have at least one incoming arc and one outgoing arc. Inventory will only be held at the end of stages in N_I, after their processing activity has completed. N_S and N_D can be thought of as dummy stages which are required to populate data for the model.

We model a production system with stationary demand operating under a periodic review policy. Demand must be filled in the period it arrives, otherwise it is lost. The ending inventory for any internal stage j in period t is found by the balance equation

$$I_{j,t} = I_{j,t-1} + P_{j,t-T_j} - \sum_{k:(j,k)\in A} S_{j,k,t} \tag{1}$$

where $P_{j,t}$ is the quantity started at stage j in period t, $S_{j,k,t}$ is the quantity shipped from stage j to stage k in period t. T_j is stage j's processing lead time. The ending inventory for stage j is the starting inventory plus the units that started at stage j. T_j time periods ago minus the units stage j ships out.

Production minimums are introduced to enforce a practical policy that if a stage is designed to produce a certain product, it is always required to make at least a minimum amount of this product every period. This is a common occurrence in many industries ensuring the stage maintains its capability to produce a product according to the designed tolerances. Planners usually impose a minimum production amount for each product assigned to each plant (Intel, 2005 and Intel, 2006).

The tactical production planning problem is formulated as a linear program **P1**

$$\text{P1}\quad \max \sum_{t=1}^{T}\left[\sum_{j\in N_D} r_j \sum_{i:(i,j)\in A} S_{i,j,t} - \sum_{j\in N_I}(c_j P_{j,t} + h_j I_{j,t} + e_j o_{j,t} + e_j u_{j,t})\right] \tag{2a}$$