16th European Symposium on Computer Aided Process Engineering
and 9th International Symposium on Process Systems Engineering
W. Marquardt, C. Pantelides (Editors)

Supply Chain Design, Management and Optimization

Dean Kassmann and Russell Allgor

Amazon.com; P.O. Box 81226; Seattle, WA 98108-1226

Abstract

Modeling and optimization are the traditional workhorses of supply chain management. The techniques have been used by many companies for planning, manufacturing, and logistics decision making. These techniques generally rely heavily on approaches grounded in operations research that excel at capturing stochastics and the discrete nature of these problems. Approaches fundamental to the process industries such as identification, dynamic simulation, model-based control, and more generally, operational-based decision making are often not understood or fully embraced by supply chain practitioners. This paper discusses the challenges and opportunities in using modeling and optimization techniques in supply chain management at Amazon. We will also discuss the application of control and feedback to supply chain systems, and discuss theoretical and practical challenges as well as opportunities in applying these ideas to real-world supply chain decision systems.

Keywords: optimization, modeling, control, supply chain management.

1. Introduction

Supply chain management refers to the decision technologies and business processes used to manage the logistics and operations of complex supply-demand networks. Rich sets of research and development opportunities associated with solving this class of problems exist. In this paper we touch on a subset of these problems, providing concrete examples from the online retail industry, specifically Amazon.com. Amazon sells millions of unique products to millions of customers world wide for billions of dollars in annual sales. The examples, although simple, represent some of the challenges and opportunities common to large scale supply-demand networks. Demand management issues, although important, will not be addressed here.

This paper is outlined as follows. Section 2 provides a high level overview of the Amazon supply-demand network with needed background and context required for the discussion. Section 3 covers several of the challenges in more detail including capacity planning, inventory planning and control, customer order assignment, and demand forecasting; we conclude in Section 4.

2. The Amazon Supply-Demand Network

The Amazon supply chain differs from traditional supply chains in several respects. First, it is shallower than many supply chains. Inventory is procured from suppliers, received in the fulfillment centers (FCs), and packed and shipped directly to customers. No large scale and complex manufacturing component exists in the supply chain. Likewise, no brick and mortar retail outlets are present, and each FC can serve each customer in a given marketplace. Second, the number of products available for sale is huge compared with most supply chains. Amazon offers wide selection of products, spanning dozens of product lines. At any point in time approximately ten million unique items may be up for sale on the Amazon website. Some of those items are sold by Amazon; some are sold by third party merchants via the Marketplace and Merchants@ programs[1]. Amazon holds a large number of these items in its fulfillment network at any given time. This contrasts many supply-demand networks that deal with hundreds of products or up to tens of thousands of products.

The basic order and inventory flows are relatively simple. Customer orders are placed on the website and enter the fulfillment network through an order assignment system that assigns the units in each order to FCs across the network in order to minimize the fulfillment costs. When the time arrives to fulfill the order, a request is sent to the appropriate fulfillment center. The items are picked by one or more associates who walk the fulfillment center floor, retrieving the physical items for incoming orders. The picked items are sorted according to customer order and placed in a box for packing. The box is then labeled and directed to an outbound dock for shipping to the customer.

The inventory flow is similar to the order flow. Using demand forecasts, inventory planning systems determine how much inventory to hold in each location for each product along with the frequency and quantity in which to purchase it. Once inventory arrives and is put away, it can be picked for a customer order as explained above.

The supply chain is a dynamic system consisting of many manipulated, feed forward, and control variables. The primary states in the supply chain are orders, inventory, and labor (e.g., staffing level) each of which evolve on different time scales. The orders typically remain in the system from hours to days, inventory from days to weeks, and labor from months to years. Applying the appropriate technology to design, manage, and optimize this dynamic system results in many research opportunities. The remainder of the paper discusses specifics about the above problems and relates the use of modeling, optimization, and control to each area.

3. Optimization and Modeling Opportunities

The management of the supply chain could be represented by an infinite horizon stochastic optimal control problem containing discrete and continuous decisions. We do not attempt to solve that problem directly. In practice, we break down the problem into smaller subproblems that are handled by individual systems. We highlight the following areas, which are critical to the operation of our supply chain in this section: capacity planning, inventory management, customer order assignment, and demand forecasting. We present an overview of the areas and some of the optimization and modeling opportunities. We then compare the techniques which have commonly been applied to these areas with those that are more prevalent in the process industries.

3.1. Capacity Planning

Capacity planning and network management determine appropriate capital investments and the best way in which to fulfill anticipated demand through the fulfillment network.

The objective is to determine what physical infrastructure in which to invest and how much of each product to fulfill from each facility. The goal is to minimize expected fulfillment costs subject to throughput, labor, and storage capacity constraints. The challenge is to balance capital investment with the variable costs of the fulfillment network subject to a highly seasonal pattern of customer demand. The primary decisions are the capital investments (e.g., new facilities, new storage modules, conveyor, etc.), where to fulfill the customer orders for various products at every instant in time, and how to manage staffing levels in each facility into the future. These plans must include commissioning and decommissioning of facilities, recommendations for construction projects, and deal with the highly seasonal nature of the retail business.

This problem is analogous to the process design problems faced in the process industries, yet the uncertainty and seasonality of the customer demand distinguishes it from problems more commonly addressed in that area. Thus, our design problem differs from the classical design of a continuous process. The problems can be classified as a mixed nonlinear stochastic combinatorial financial optimization problems. For most real world instances, a full stochastic formulation is not tractable. Instead, most practitioners, including us, cast the problem as multi-period LPs or MILPs; stochastic elements are ignored, and various scenarios are analyzed. For these multi-period (i.e., finite horizon) problems, the choice of horizon length and the manner in which final time constraints are imposed is important. The capacity planning problem that we face does not reach steady-state conditions because of the growth and seasonality of our business. We address this by considering a time horizon that looks far enough into the future to ensure that constraints are no longer active and the 'cost to go' at the end of the time horizon can be sufficiently discounted. For example, the labor plans must ensure that the horizon includes the staffing requirements through the labor ramp and decay of the peak season. This ensures that enough trained staff are available to handle peak production volumes and deal with the situation in which demand increases at a faster rate than staff can be hired and trained. In addition, the horizon should look far enough past the end of the peak season to ensure that the proper balance of temporary and permanent staffing is maintained. This ensures that too many permanent workers are not hired at peak. Similar considerations impact the horizon for capital investment.

In contrast to many of the process design and control problems that have been addressed, most approaches to planning do not include explicit disturbance models or state estimation techniques. Perfect state estimation is assumed, and little, if any, attention is given to stability of the closed loop system. Most research is focused on investment decisions and the minimization of operating costs under each investment. Detailed operational models and control strategies are not typically included in these formulations; instead, aggregate models of the operating costs are incorporated. This is similar to separating the process design and control problems.

3.2. Inventory Management

Inventory planning is a function of many things. It depends upon the statistics of future uncertain demand, capacity constraints, product lead times and other factors. The problem includes the optimization of both the purchasing and the reverse logistics. For each item, there are multiple vendors and sourcing channels for procurement into a given location. Each has different costs and dynamics. For example, bulk discounts may exist when purchasing a product from a vendor (e.g., the per unit cost is greater when ordering single units instead of case or pallet quantities), freight and payment terms will vary. Different channels to remove overstock product also exist. For example, disposition options will often include vendor returns, liquidation, and markdown strategies, and the available options may depend not only upon the product, but also on the terms under which it was purchased. The inventory management problem is the analog to the level control problem.

The fundamental operational question in inventory management is how many units of a given product should be held at a given location based on current and forecasted demand subject to the dynamics of the supply-demand network. The academic literature on inventory management is large. See [5,8,9] for a collection of papers and reviews. One focus in the literature is on finding optimal inventory control policies under different assumptions. The control policy is generally a closed form analytic formula that relates the current demand and inventory level to a purchase quantity. Different assumptions on demand (deterministic vs. stochastic; stationary vs. nonstationary, constant vs. time-varying), the number of products (single vs. many), and number of locations and structure of the supply-demand network all yield different models. The classical economic order quantity[11], dynamic lot size model[11], (Q,r) and (S,s) models[4,6], and other model variants are the result. The (S,s) policy is the rough analog to LQR theory and the Kalman gain for unconstrained single product systems with stochastic time varying demand. The inventory literature assumes that the models are correct and that the only variation comes in through the demand and lead time variation. A common theme in these problems is to simultaneously determine the optimal inventory levels along with the dynamic control policy. This contrasts the process industries where the problems are normally separate. The problems tend to be fairly simple because closed form solutions are sought for these inventory policies. Most problems deal with only a single product. Constraints are rarely considered. Even though these problems don't deal with much of the real world complexity required to solve practical problems, they provide useful insight into the general dynamics of supply chains. These insights lead to rules of thumb that can be used by practitioners and analysts for policy making. The main theoretical tool for most of the above analysis is dynamic programming (DP); the Hamilton-Jacobi-Bellman (HJB) equation provides the foundation for provably optimal policies[3].

At the other extreme is an attempt to fully formulate the detailed inventory control problem and solve the resulting stochastic optimization problem. Quite often much of this work is application centric; relying on solution techniques rather than an overarching theory to provide commonality between application instances. Here again the workhorse is DP or approximate dynamic programming (ADP); however, in this instance it is used as a computational tool rather than a theoretical tool. Closed loop stability of the control policies is never really considered as the DP policy will be nominally stable as long the HBJ equation admits a solution. Robustness is not generally discussed. Model predictive control, although suboptimal compared to DP, has not been actively investigated.

Most of the difficulty with these problems stems from the combinatorial size of the problem, the stochastics and non-stationary nature of demand, variable lead times, and the complex, often discontinuous, cost structures. Practical problems contain economies of scale and constraints that involve multiple products, such as inbound freight costs and vendor PO minimums, that are not often addressed by the inventory management literature. Finally, one time purchase opportunities may arise. As a result of these difficulties, there are many insights and opportunities that can be gained from taking a broader perspective to the control and the inventory management space. To date, no research provides a comprehensive approach to address the complexity of the problems faced in practice.

3.3. Order Assignment

Based on the customer expectations set during the checkout process[2], the units in a customer order must be assigned to fulfillment centers across the network for processing. Each unit in an order has the potential to be fulfilled from a number of different physical locations based on inventory levels and the ship and delivery estimates provided to the customer at checkout. In its simplest form this is an exact set cover problem in which the units in the order are partitioned into sets, and each set of units is assigned to a feasible location whose cost can be calculated. The goal is to minimize the expected fulfillment cost of a feasible assignment. These problems are NP-hard for which no provably good approximation algorithms are available. The problem can also be modeled as a fixed cost multi-commodity network flow problem, in which nonlinear concave costs are easily handled. Although the shipping cost structure faced by Amazon is typically non-concave, we have found that it is reasonably approximated by concave functions in most cases. In addition, we have found that a large fraction of orders equate to problem instances that are small enough to employ exact solution techniques, and that variants of some approximation algorithms[7] perform well in practice.

In our business, though, order assignment is even more complex. Network level operational goals must be satisfied in addition to the minimum cost order assignment. Every assignment of a unit to a facility consumes resources at that facility, such as labor capacity to pick, pack, and ship the items and processing capacity to move items within the facility. Poor assignment can starve or swamp a facility with work, so active control of the workload assigned to a facility is important. Many of the operational objectives of our fulfillment network can be recast as set points, zone constraints, and limits in a control-oriented formulation. Control research provides a formalism in which to represent these objectives in an online system. Practically, this is still very difficult. The challenge is to cast the network level objectives into a form that can be implemented in the decision logic of a single- or multi-order optimization algorithm. Our approach has been to solve the online problem using optimization and control techniques; we revisit these decisions to take advantage of new information and opportunities. The multi-order problems can become very large (over one hundred million variables) which provides challenges for both the solution algorithms and the software architecture [10]. This leads to research opportunities to provide better models of the business problem and improved solution techniques that can provide real-time answers within acceptable levels of accuracy.

3.4. Demand Forecasting

Most of the above methods rely on predictions or forecasts of demand. It is a key feed forward component in many of the decision algorithms. The forecasts are produced by forecasting systems for different time ranges in the future (daily, weekly, etc.) Forecasts are created for every single product and can be represented at national and local levels. Much of the forecasting infrastructure relies on standard time series modeling techniques. ARX, ARIMA, and exponentially weighted moving average models are often used to model demand. Product lifecycle affects the model choice. New release items, not yet published items, and well established items all have markedly different demand dynamics. A single model can have difficulty remaining accurate for the entire life of a product. Practical approaches like switching techniques and mixture modeling are used to overcome these problems.

State space formulations are not standard practice. Thus, it is much less common to see the use of Kalman filters and other estimation techniques in demand forecasting. Model identification for a small set of products can be accomplished manually. However, automatic identification and subsequent prediction for over a million products poses algorithmic and software challenges. External causal events, such as promotions, availability changes, competitor pricing, etc., also impact the forecasts. Capturing the multitude of these data signals, then screening and cleaning the data for problems is a technical challenge on its own.

4. Conclusion

Effective planning and operation of a complex supply-demand network is a difficult and rewarding problem. There are significant opportunities for the classic operations research techniques as well as the optimization, simulation and control techniques to add value. Each discipline touches on different aspects of the problem, and the existing body of research does not address the exact problems commonly faced by industrial practitioners. Many outstanding challenges lie in the modeling of these problems, the development of efficient algorithms to solve the models, and the architecting of software systems to implement these solutions and manage data and workflow. Addressing these problems and questions holds enormous potential for bringing tangible economic benefit and competitive advantage to a business.

5. References

1. Amazon.com, Inc., 2004, *Annual Report 2004,* http://www.amazon.com.
2. M. Ball, C.Y. Chen and Z.Y. Zhao, 2004, *Available to Promise*, in Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era, Simchi-Levi, D., D. Wu, and M. Shen, eds., Kluwer Academic Publishers, Boston, pp 447-484.
3. D. Bertsekas, 2000, *Dynamic Programming and Optimal Control, 2nd ed.*
4. A. Clark and H. Scarf, 1960, *Optimal policies for a mulit-echelon inventory problem*, Management Science, 6, pp 475-490.
5. S. Graves, Rinnooy Kan and P. Zipkin, 1993, *Handbooks in Operations Research and Management Sciences: Logistics of Production and Inventory*.
6. G. Hadley and T.M. Whitin, 1963, *Analysis of Inventory Systems,* Prentice-Hall, Englewood Cliffs.
7. K. Jain and V. V. Vazirani, 2001, Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation*, Journal of the ACM 48*, *pp. 274*--296.
8. T. de Kok and S. Graves, 2003, *Handbooks in Operations Research and Management Sciences: Supply Chain Management, Design, Coordination and Operations*.
9. S. Tayur, R Ganeshan, and M. Magazine, eds. 1998, *Quantitative Models for Supply Chain Management*, Kluwer Acad. Pub., Boston.
10. P. J. Xu, R. Allgor, and S.C. Graves, 2006 (forthcoming), The Benefits of Re-Evaluating Real-Time Fulfillment Decisions, Manufacturing and Service Operations Management. 8(1).
11. P. Zipkin, 2000, Foundations of Inventory Management, McGraw-Hill, Boston.