European Journal of Operational Research 224 (2013) 497–506

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing and Logistics

Inventory sharing in integrated network design and inventory optimization with low-demand parts

Ilyas Mohamed Iyoob, Erhan Kutanoglu *

The University of Texas at Austin, Austin, TX 78712, United States

ARTICLE INFO

Article history:
Received 11 November 2011
Accepted 19 September 2012
Available online 27 September 2012

Keywords:
Inventory
Logistics
Service parts logistics
Inventory sharing
Network design

ABSTRACT

Service Parts Logistics (SPL) problems induce strong interaction between network design and inventory stocking due to high costs and low demands of parts and response time based service requirements. These pressures motivate the inventory sharing practice among stocking facilities. We incorporate inventory sharing effects within a simplified version of the integrated SPL problem, capturing the sharing fill rates in 2-facility inventory sharing pools. The problem decides which facilities in which pools should be stocked and how the demand should be allocated to stocked facilities, given full inventory sharing between the facilities within each pool so as to minimize the total facility, inventory and transportation costs subject to a time-based service level constraint. Our analysis for the single pool problem leads us to model this otherwise non-linear integer optimization problem as a modified version of the binary knapsack problem. Our numerical results show that a greedy heuristic for a network of 100 facilities is on average within 0.12% of the optimal solution. Furthermore, we observe that a greater degree of sharing occurs when a large amount of customer demands are located in the area overlapping the time windows of both facilities in 2-facility pools.

1. Introduction

Service Parts Logistics (SPL) is about providing spare parts used in the post-sales maintenance and repair of high-capital equipment, such as heavy machinery, manufacturing equipment, and medical devices. These products are expensive, and are made of expensive parts, motivating the careful maintenance and repair of failed parts, rather than replacing the whole product. Due to the high reliability of products, part failures are typically very infrequent, leading to demands for service (spare) parts that are very low. However, the need for making the products in use functioning again quickly is nevertheless paramount, mainly because most of these products support the critical missions of the organization using the product.

For example, a manufacturing plant may rely on an expensive machine that might be a bottleneck, hence a part failure causing down time should be addressed as quickly as possible. This motivates the time-based service level requirements seen in SPL systems. Thus, in SPL, two issues, strategically locating service parts stocking facilities to serve geographically dispersed customers needing service, and deciding the part stock levels to be used for maintenance and repair, become extremely important. The former is typically called logistics network design optimization and can be viewed as an extended version of the facility location and demand allocation problem. The latter is called inventory stocking problem, and is solved to obtain the optimal inventory levels across all facility locations such that the response time based target service levels are achieved.

The interdependency between the network design and inventory stocking decisions and the tradeoffs among the cost components (facility installation, transportation, and inventory) and service levels motivate the simultaneous optimization of these decisions which are modeled and analyzed recently (Candas and Kutanoglu, 2007; Jeet et al., 2009). If a facility's stock does not have a part needed by a customer, then these models assume that the demand from the customer is passed up to a central warehouse which ships the part directly to the customer's site in an emergency mode. Although this assumption is accurate in some SPL systems, in others, practicing field managers take advantage of inventory available at nearby facilities by requesting direct shipment from a neighbor facility stocked with the needed part to the customer, and using the shipment from the central warehouse as a last resort (i.e., if the part is not available in all the facilities considered for inventory sharing). This type of inventory sharing among nearby facilities can be especially useful to increase the service levels for a given system-wide stock, mainly because there is a chance that the shipment from the neighboring facility is within the time window due to

* Corresponding author.
E-mail address: erhank@mail.utexas.edu (E. Kutanoglu).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.09.033

close proximity of the locations involved. In this paper, we extend the integrated network design and inventory models to explicitly include this type of inventory sharing practice explicitly.

Lee (1987) was the first that modeled inventory sharing through lateral transshipments. After dividing facilities among a limited number of pooling groups, he assumed that all the facilities are identical in each pooling group. He also developed an algorithm to solve for optimal stock levels while minimizing holding costs subject to service level constraints. Dropping the assumption of identical facilities in each pooling group, Axsäter (1990) developed a methodology for better results than Lee (1987) assuming exponentially distributed replenishment times. Dada (1992) further extended the inventory sharing model to allow items in transit to be used for filling orders as well. This was the first model of its kind. Finally Alfredsson and Verrijdt (1999) combined the works of Dada (1992) and Axsäter (1990) to develop an analytic model that calculates the fraction of demand of a customer satisfied by different sources such as the regular order fills from the assigned facility, lateral transshipments from other facilities, and direct shipments from the central warehouse. Alfredsson and Verrijdt (1999) concluded that the distribution of shipment times has a negligible effect on the overall model. More recent work further advanced the inventory sharing analysis into different types of sharing policies (partial or full) by Kranenburg and van Houtum (2009). The iterative procedure of obtaining sharing-based fill rates in Alfredsson and Verrijdt (1999) was also used to gain insights into effects of inventory sharing in service parts logistics systems with given stock levels (Kutanoglu, 2008) and as a basis for an implicit enumeration algorithm for given demand assignments across facilities (Kutanoglu and Mahajan, 2009). A rather extensive and recent review on inventory sharing and lateral transshipments is by Paterson et al. (2011).

In this paper, we extend the works of Lee (1987) and Alfredsson and Verrijdt (1999) to include inventory sharing within the integrated network design (including demand allocation) and inventory optimization problem. Given the already complex nature of the integrated models mentioned above, we make certain simplifying assumptions when we consider inventory sharing to make the models more amenable to further analysis. Even with these assumptions, limiting the stock levels and the number of facilities sharing inventory (i.e. pool size), the inventory sharing models considered in the paper are challenging and informative, giving insights to when inventory sharing is beneficial, and how it affects the network design and inventory stocking decisions. (Note that similar assumptions limiting the stock levels to 0 or 1 are made before in the service parts logistics literature, see, e.g., Cohn and Barnhart (2006).) In this case, network design boils down to which facilities to open (if there is significant facility costs) and stock (if there is significant holding cost).

The rest of the paper is organized as follows: Section 2 provides a list of assumptions and Section 3 presents the challenges of estimating fill rates when inventory is shared across multiple facilities. Section 4 formulates and solves the single pool problem, followed by the multiple pool problem formulation in Section 5. Sections 4 and 5 also include algorithms and heuristics as solution techniques. In Section 6 we compare the sharing vs no sharing case to identify cases in which sharing is most beneficial as well as cases when sharing might actually hurt the system. Computational results are shown in Section 7, followed by conclusions and future research in Section 8.

2. Assumptions

As mentioned earlier, a "pool" is defined as a set of stocking facilities that are allowed to share inventory. Demand for a service part requested from a facility is satisfied in the following order:

- If the facility has stock on-hand, this on-hand stock is used to satisfy the demand.
- If the facility is out of stock and another facility in the pool has stock on-hand, then demand is satisfied by the other facility.
- If all the facilities in the pool are out of stock, then demand is satisfied by a direct shipment from the central warehouse which has infinite supply and is far from all customers and facilities.

Note that all the facilities are replenished from the central warehouse according to the base stock policy. Moreover, customers have to be assigned to some facility and cannot be satisfied directly from the central warehouse.

For computational reasons we assume that each pool has two facilities. Moreover, due to the low demand nature of spare parts, we assume that a stock level of 1 is sufficient to satisfy all the demand at a facility. For a lead time of 1 period, this corresponds to a maximum demand rate of 0.01 per facility per time period for practically 100% service (Fig. 1). When we consider inventory sharing between facilities, estimating fill rates becomes a challenge. Alfredsson and Verrijdt (1999) estimated these probabilities iteratively using Markov chains. In this paper however, we use the low demand nature of spare parts to simplify the Markov chain based iterative process into a set of simultaneous equations. This enables us to incorporate the inventory sharing formulae within the integrated network design and inventory optimization problem with time based service level constraints.

3. Fill rate estimation

As a first step in formulating the inventory optimization problem, we modify the iterative methodology of Alfredsson and Verrijdt (1999) in estimating the local and lateral fill rates (due to inventory sharing) for a single pool with two facilities. We then formulate equations to estimate the fill rates, thus eliminating the need of the iterative procedure.

Suppose $I = \{1,2\}$ is the set of facilities in a single pool (Fig. 2). At each facility $i \in I$, let S_i denote the stock level and let λ_i be the (local) demand rate (this is the demand originally assigned to facility i). The replenishment rate from the central warehouse is denoted by μ, which is the reciprocal of the lead time. We also define:

β_i – fraction of demand (of facility i) satisfied by on-hand stock at facility i,
γ_i – fraction of demand of facility i satisfied by lateral transshipment from another facility in the pool, and
θ – fraction of demand satisfied by a direct shipment from the central warehouse (CW).

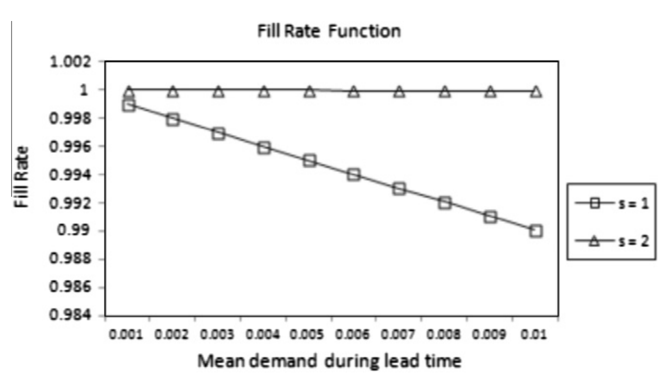

Fig. 1. Fill rate as a function of mean demand during lead time.