Omega 55 (2015) 10–23

Contents lists available at ScienceDirect

Omega

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

Logic-based Benders decomposition for an inventory-location problem with service constraints ☆

David Wheatley [1], Fatma Gzara *, Elizabeth Jewkes

Department of Management Sciences, University of Waterloo, 200 University Ave W., Waterloo, Ontario, Canada N2L 3G1

ARTICLE INFO

Article history:
Received 10 October 2013
Accepted 3 February 2015
Available online 10 February 2015

Keywords:
Integer programming
Inventory control
Location
Mathematical programming
Operational/OR
Optimization

ABSTRACT

We study an integrated inventory-location problem with service requirements faced by an aerospace company in designing its service parts logistics network. Customer demand is Poisson distributed and the service levels are time-based leading to highly non-linear, stochastic service constraints and a nonlinear, mixed-integer optimization problem. Unlike previous work in the literature, which propose approximations for the nonlinear constraints, we present an exact solution methodology using logic-based Benders decomposition. We decompose the problem to separate the location decisions in the master problem from the inventory decisions in the subproblem. We propose a new family of valid cuts and prove that the algorithm is guaranteed to converge to optimality. This is the first attempt to solve this type of problem exactly. Then, we present a new restrict-and-decompose scheme to further decompose the Benders master problem by part. We test on industry instances as well as random instances. Using the exact algorithm and restrict-and-decompose scheme we are able to solve industry instances with up to 60 parts within reasonable time, while the maximum number of parts attempted in the literature is 5.

1. Introduction

The sale of after-market service parts is a significant source of revenue in the aerospace industry, but it is a complex industry with many unique challenges. Customers who operate aircraft around the world discover failed parts during routine maintenance and may require replacement parts immediately. An aircraft on ground event is estimated [1] to cost between U.S.$5,000 and U.S.$150,000 per hour, depending on the operator and model of aircraft. It is therefore common practice in the aerospace after-market industry to have service agreements that require delivery of failed parts within a specified time window to ensure timely return to service of an aircraft on ground.

This paper is motivated by a problem faced by a major North American aerospace company that manufactures aircraft and provides after-market support to aircraft operators. The company operates a service parts logistics network to ensure that each customer has access to a service centre stocked with service parts within a given time window. Ideally, the service parts provider would like to deliver 100% of customer orders within the time window, but there are two challenges. First, demand for service parts is random and typically infrequent, so the company must consider the trade-off between carrying more inventory and achieving higher service levels. Second, operating a dense network of service centres that can deliver parts within a short time window may lead to high inventory levels. To address these challenges, the location of service centers and assignment of customers should be done in conjunction with inventory stocking decisions. The goal is to locate service centres around the world, stock service parts inventory, and assign customer demand to meet contracted customer service levels, while minimizing total system costs.

The contributions of the paper are as follows. To the best of our knowledge, we present the first exact methodology to solve the integrated inventory-location problem for service parts with a system-wide time-based part service level. Three of the papers cited in the literature review section [2–4] provide approximate algorithms that may find good quality solutions, but do not guarantee optimality. Our novel logic-based Benders decomposition algorithm provides a family of valid cuts and a certificate of optimality. The second contribution is a new procedure to construct tight valid cuts and an effective method to construct feasible solutions using partial solutions from the master problem. Another major contribution is a restrict-and-decompose scheme that

☆This manuscript was processed by Associate Editor Bish.

* Corresponding author.
E-mail addresses: dmwheatl@uwaterloo.ca (D. Wheatley), fgzara@uwaterloo.ca (F. Gzara), emjewkes@uwaterloo.ca (E. Jewkes).

[1] Supported by NSERC CRD Grant RPGIN-298767-2007.

http://dx.doi.org/10.1016/j.omega.2015.02.001
0305-0483/

improves on the logic-based Benders decomposition algorithm by reducing computation time and that efficiently solves instances with a large number of parts. Previous research on integrated models has attempted problems with a single part [2,3], or up to five parts [4]. The potential of our approach is illustrated by solving problems using data from a real aerospace company with up to 60 parts.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature on inventory-location problems for service parts and logic-based Benders decomposition. Section 3 describes the nonlinear and stochastic mixed-integer programming formulation for the inventory-location problem. Section 4 presents an exact solution methodology based on logic-based Benders decomposition, details the family of valid cuts, and proves the validity of the cuts and the complete algorithm. Section 5 describes the restrict-and-decompose scheme that is tailored to solve problem instances with a large number of parts. Section 6 provides computational results and managerial insights on real industry instances as well as random instances used in the literature. Section 7 concludes the paper.

2. Literature review

Service parts are often characterized by having low rates of demand with unpredictable but high priority demand. The seminal works in service parts inventory management are by Sherbrooke [5] and Graves [6], while the book by Muckstadt [7] is now considered a primary reference. A base stock inventory policy is typically chosen for low-demand service parts, see [8] for justification. Works such as Roni, Jin and Eksioglu [9] study the management of surge demand for commodities in emergency situations, but the high demand models are not appropriate for most service parts.

This paper builds on previous studies [2–4] of integrated inventory-location problems with time-based service level constraints. Candas and Kutanoglu [2] and Jeet et al. [3] study a problem where fill rates at individual locations can vary (i.e., *variable fill rates*), provided the system-wide fill rate meets the contracted service level. The fill rates are calculated by determining the long-term probability that a part is available at the assigned service centre immediately when it is ordered by a customer. Gzara et al. [4] consider two different service level aggregations, one of which is the system-wide part service level. The system-wide part service level is also the focus of this paper, motivated by the need for global aircraft operators to have parts available to customers around the world within a specific time window.

The time-based service levels used in [2–4] and in this paper do not consider the possibility that a replenishment part may be in transit to a service centre and still be delivered to a customer on time. This possibility is considered by Caggiano et al. [10], who study generalized time-based service levels and determine base stock levels for multiple parts in a multi-location, multi-echelon service parts distribution network where the locations are known *a priori*. Solving an integrated facility location and inventory problem with this generalized time-based service level is an area for future research.

This paper also does not consider the possibility that a part that is not immediately available at the assigned service centre may be delivered from another location within the time window. This strategy is referred to as Emergency Lateral Transshipments (ELT), and earlier approaches to solving these problems are surveyed by [11]. These two assumptions reflect the logistics strategies used by the industry partner.

As a result, the time-based service level in [4] and this paper is effectively a distance-based restriction that does not allow customer demand to be assigned to a location outside of the time window. The models in [2,3] allow customers to be assigned to service centres that lie outside of the time window, and thus cannot possibly receive the service part on time. This requires providing a higher level of service to customers that are within the time window to achieve the desired target service level. Finally, the models in both [2] and [4] assume that unfilled demand is backordered, while the model in [3] assumes that unfilled demand is treated as lost sales. The two assumptions require subtly different computations to evaluate the achieved service level, and should be considered as distinct problems. This paper assumes unfilled demand is backordered, as many large or expensive aerospace parts are only available from a single supplier.

In summary, the problem considered in this paper is the same as that studied in [4] and relates but is not directly comparable to the problems studied in [2,3]. Direct comparisons are made in Section 5 between the results of the exact algorithm presented in this paper and the approximate method found in [4], while direct comparisons of the approximation methods of [2–4] are found in [4].

2.1. *Approximate and exact methodologies for various service level aggregations*

The calculation of the fill rate depends on how aggregated the service level is and different service level aggregations may require different solution methodologies. When service is measured at the service centre-part level, [4] develops a reformulation that finds an exact solution. The reformulation uses pre-calculated threshold demand quantities to meet the service level for any given level of inventory. However, this technique cannot be directly applied to other service level aggregations [4].

When the service is measured at the part level, no efficient exact methodologies are known to exist. [2] creates piece-wise linear step functions to linearize the fill rate constraints which may either overestimate or underestimate inventory resulting in suboptimal or infeasible solutions. [3] explicitly use fill rate decision variables and write a convex model that they linearize using an outer approximation scheme. Then lower bounds are found by solving a separate lower bounding problem, while upper bounds are found by solving a relaxed model with higher service levels than required. Both approaches seem to be tailored specifically to the problem at hand, making their suitability for use on other problems questionable. The work in [4] linearizes the backorder fill rates in a novel way, allowing a solution to choose an inventory stocking level for a part only if the assigned demand for that part falls within a precomputed interval. The approach uses a large number of discretization points that does not increase the number of constraints, which appears to be useful in solving smaller problems quickly. However, the work does not include any attempts at finding lower bounds, so the quality of the solutions cannot be measured.

All of the literature reviewed in this section has used approximation techniques or heuristics to solve the integrated inventory-location problem with service levels at the part level. In contrast, this paper presents an exact methodology to solve the same problem, and provides a measure of the quality of the solutions found by approximations. The use of an exact method is desirable for two reasons. The first is providing a certificate of optimality: proving whether or not a solution found by an approximate method is optimal. The second is the guaranteed convergence to optimality, which allows an exact method to always find an optimal solution when approximate methods may fail. Practitioners are expected to make these inventory-location decisions