European Journal of Operational Research 255 (2016) 729–746

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Discrete Optimization

Multi-period capacitated facility location under delayed demand satisfaction

Isabel Correia[a,b], Teresa Melo[c,*]

[a] Centro de Matemática e Aplicações (CMA), Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), P 2829-516, Caparica, Portugal
[b] Departamento de Matemática, Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), P 2829-516, Caparica, Portugal
[c] Business School, Saarland University of Applied Sciences, D 66123, Saarbrücken, Germany

ARTICLE INFO

Article history:
Received 4 August 2015
Accepted 16 June 2016
Available online 20 June 2016

Keywords:
Location
Multi-period
Capacity choice
Delivery lateness
MILP models

ABSTRACT

We address an extension of the classical multi-period facility location problem in which customers are sensitive to delivery lead times. Accordingly, two customer segments are considered. The first segment comprises customers that require timely demand satisfaction, whereas customers accepting delayed deliveries make up the second segment. Each customer belonging to the latter segment specifies a maximum delivery time. A tardiness penalty is incurred to each unit of demand that is not satisfied on time. In the problem that we study, a network is already in place with a number of facilities being operated at fixed locations. The network can be expanded by establishing new facilities at a finite set of potential sites and selecting their capacity levels from a set of available discrete sizes. In addition, existing facilities may be closed over the time horizon. Two mixed-integer linear programming formulations are proposed to re-design the network at minimum cost and a theoretical comparison of their linear relaxations is provided. We also extend the mathematical models to the case in which each customer accepting delayed demand satisfaction requires late shipments to occur at most once over the delivery lead time. To gain insight into how challenging these problems are to solve, a computational study is performed with randomly generated instances and using a general-purpose solver. Useful insights are derived from analyzing the impact of different delivery lead time restrictions on the network structure and cost.

1. Introduction

In this paper, we consider a company that operates a set of facilities at fixed locations to fulfil customer demand requirements for a given product. Customers are classified on the basis of their sensitivity to delivery lead times. In particular, two customer segments are considered. The first segment comprises customers requiring timely demand satisfaction, whereas customers accepting delayed deliveries make up the second segment. Each customer belonging to the latter segment specifies a desired maximum delivery lead time. Changing market and business conditions, frequently in conjunction with increased cost pressure and service requirements, compel the company to restructure its network of facilities. To this end, the network configuration can gradually change over a multi-period horizon through opening new facilities and closing initially existing facilities. Locations for new facilities are chosen from a finite set of candidate sites. Moreover, at each potential site, different capacity levels are also available for selection. The objective is to determine the optimal network configuration over the planning horizon so as to minimize the sum of fixed and variable costs. The former are associated with strategic decisions and comprise fixed costs for facility siting and operation, for capacity acquisition and for closing existing facilities. Variable costs for processing and distributing the product to the customers along with tardiness penalty costs for delayed demand satisfaction are also incurred. These costs are associated with tactical decisions.

The problem that we study is motivated by the online retail and manufacturing industries (e.g. computer and mobile telephone manufacturers) who often adopt different pricing policies for their products, thereby offering price incentives to customers in exchange for longer order lead times (Agatz, Fleischmann, & van Nunen, 2008). Multiple customer segments that are differentiated on the basis of delivery lead times and price options can also be encountered in other business environments (see e.g. Hung, Chew, Lee, & Liu, 2012). For example, Wang, Cohen, and Zheng (2002) report the case of a semiconductor equipment manufacturer that provides a two-class service policy. Customers with emergency demand for repairable service parts form the first class. These customers pay a premium price to have their returned defective parts repaired immediately. Non-emergency service is provided to the

* Corresponding author.
E-mail address: teresa.melo@htwsaar.de (T. Melo).

http://dx.doi.org/10.1016/j.ejor.2016.06.039

second class of customers who pay a lower price for a longer repair time. Cheong, Bhatnagar, and Graves (2005) describe a similar case in the textile dye industry where textile mills with small capacities favor shorter delivery lead times from their suppliers as opposed to large textile mills that can keep higher stock levels and, therefore, can cope with longer order fulfilment lead times. Wu and Wu (2015) describe this type of strategy as "demand postponement" because the company decides upon the actual delivery time for orders committed by customers who are less sensitive to lead times.

On the one hand, lower unit prices may result in decreased total revenue for the company but, on the other hand, the company will have to invest in increasing capacity to guarantee a shorter delivery time to those customers who are willing to pay a premium price. Hence, one of the main research objectives of our work is to investigate the interplay between location and capacity investment (or disinvestment) decisions and delivery time decisions in a facility system operated by a company. The coordination of location, capacity acquisition, distribution and demand fulfilment decisions has the potential to improve system efficiencies. In particular, retail and manufacturing environments can benefit from the insights provided by our research work.

As will be shown in Section 2, our study is the first to embed customer segments having distinct sensitivity to delivery lead times in a multi-period facility location setting. This new feature is combined with different time scales for strategic and tactical decisions. This means that tactical decisions regarding the commodity flow from operating facilities to customers may take place in any time period, whereas strategic location decisions can only be made over a subset of the time periods of the planning horizon. Furthermore, the decision space is extended with strategic facility sizing decisions, an aspect that is not often encountered in the literature. In fact, most facility location models consider capacity as an exogenous factor. However, from an application point of view, capacity is often purchased in the form of equipment which is only available at a few discrete sizes. Capacity choices incur specific fixed installation costs that are subject to economies of scale. Hence, a further research objective of this study is to add three new dimensions (customer segments with distinct service requirements, different decision time scales and multiple capacity choices) to the classical multi-period facility location problem which is known to be an NP-hard problem. From a computational standpoint, this results in a challenging problem for which the possibility of solving large-scale instances to optimality within acceptable time is rather limited. In such cases, one often resorts to heuristic methods to obtain feasible solutions. However, to be able to measure the quality of such solutions it is of paramount importance to have (good) lower bounds for the problem. Therefore, the third research objective of our work is to develop different mathematical formulations and to compare them in terms of the LP-relaxation bound they provide. Additional inequalities are also proposed in an attempt to strengthen these bounds. Finally, we conduct an extensive computational study to obtain managerial insights that illustrate the far-reaching implications of delivery lead time restrictions on the network structure and its cost. Without the support of the models developed in this paper it would otherwise be difficult to obtain most of these insights. Given the typically high investment costs and the limited reversibility of strategic decisions, it is essential for stakeholders to perceive the impact of such decisions on overall system performance.

The remainder of this article is organized as follows. Section 2 summarizes the relevant literature. In Section 3, we develop two mixed-integer linear programming (MILP) formulations for the problem under study and present a theoretical comparison of their linear relaxations. In particular, we also consider the special case in which customers accepting late deliveries wish to receive single shipments even if they arrive with some delay. In other words, partial, late deliveries are not allowed for such customers. In Section 4, additional inequalities are proposed to enhance the original formulations. Section 5 reports and discusses the results of an extensive computational study using general-purpose optimization software. Finally, in Section 6, conclusions are provided and directions for future research are identified.

2. Literature review

Discrete facility location models are typically concerned with determining the number, location and capacities of facilities that should be established to serve the demands of a set of spatially distributed customers with least total cost. This field of location analysis has been an active and rich research area over the past decades. A wide variety of applications have emerged in many contexts such as strategic logistics planning (see e.g. Alumur, Kara, & Melo, 2015, chap. 16 and Melo, Nickel, & Saldanha da Gama, 2009) and telecommunications (see e.g. Fortz, 2015, chap. 20), just to name a few.

Most discrete location models ensure the satisfaction of customer demands by imposing distance and/or time limits as service level requirements. In contrast, location problems with flexibility regarding demand fulfilment have received much less attention. The case of unfilled demand can be treated either with the lost sales assumption or with the backorder assumption. The former situation applies to contexts in which satisfying all customer demands may not be economically attractive due to high investment costs on establishing new facilities with appropriate capacities. In a static setting, Alumur et al. (2015) describe a generic model for a facility location problem arising in logistics network design that includes this feature, while Correia, Melo, and Saldanha-da-Gama (2013) address this issue in the design of a two-echelon production–distribution network over multiple time periods. The models developed by Badri, Bashiri, and Hejazi (2013), Bashiri, Badri, and Talebi (2012), Canel and Khumawala (1996) and Sousa, Shah, and Papageorgiou (2008) also allow lost sales over a dynamic horizon. In the previous studies (Badri et al., 2013; Bashiri et al., 2012; Canel and Khumawala, 1996; Correia et al., 2013; Sousa et al., 2008), strategic location and tactical logistics decisions are made under a profit maximization objective. In addition, Correia et al. (2013) also investigate their problem from a cost minimization perspective with additional constraints enforcing a minimum rate for demand fulfilment. For a number of test instances with small and moderate sizes, the MILP formulations proposed in Alumur et al. (2015); Bashiri et al. (2012); Canel and Khumawala (1996); Correia et al. (2013); Sousa et al. (2008) could be solved to optimality by a commercial MILP solver within acceptable time. Badri et al. (2013) developed a Lagrangian-based heuristic through dualizing a set of constraints that limit the expenditures for opening new facilities and expanding the capacity at existing locations.

The lost sales assumption is also present in the problem addressed by Altiparmak, Gen, Lin, and Paksoy (2006) through the maximization of the overall fraction of demand that is delivered to customers. This objective is integrated with the minimization of the total cost of designing and operating a multi-stage network and the maximization of the capacity utilization of facilities. These three objectives are combined into a single-objective function by building a weighted sum and feasible solutions are determined with a genetic algorithm. The latter solution methodology was also adopted by Lieckens and Vandaele (2007) for a facility location problem arising in reverse logistics with stochastic lead times for processing and moving used products. In this case, a fraction of the returned products may not be collected and demand for reused products may be only partially met. Cheong et al. (2005) follow a different approach to deal with lost sales in an uncapacitated two-echelon distribution network. To this end, each customer is