European Journal of Operational Research 204 (2010) 316–327

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Optimum service capacity and demand management with price incentives

Özgür Özlük [a], Abdelghani A. Elimam [b,*,1], Eduardo Interaminense [c]

[a] Decision Sciences Department, College of Business, San Francisco State University, San Francisco, CA 94132, USA
[b] Mechanical Engineering Department, School of Sciences and Engineering, American University in Cairo, Egypt
[c] Business Development, MCI, San Francisco, CA 94105, USA

ARTICLE INFO

Article history:
Received 2 July 2008
Accepted 13 October 2009
Available online 30 October 2009

Keywords:
Mixed integer programming
Workload smoothing
Delivery dates
Heuristics
Incentives

ABSTRACT

Service firms periodically face fluctuating demand levels. They incur high costs to handle peak demand and pay for under-utilized capacity during low demand periods. In this paper, we develop a mixed integer programming (MIP) model based on the real life experience of a Brazilian telecommunications firm. The model determines the optimum staffing requirements with different seniority levels for employees, as well as the distribution and balancing of workload utilizing flexibility of some customers in their service completion day. The proposed MIP uses monetary incentives to smooth the workload by redistributing some of the peak demand, thereby increasing capacity utilization. Due to the intractable nature of optimizing the proposed MIP model, we present a heuristic solution approach. The MIP model is applied to the case of the examined Brazilian Telecommunications firm. The computational work on this base case and its extensions shows that the proposed MIP model is of merit, leading to approximately seventeen percent reduction in the base case operating costs. Extensive computational work demonstrates that our heuristic provides quality solutions in very short computational times. The model can also be used to select new customers based on the workload, the revenue potential of these new customers and their flexibility in accepting alternate service completion dates. The generic structure of the proposed approach allows for its application to a wide variety of service organizations facing similar capacity and demand management challenges. Such wide applicability enhances the value of our work and its expected benefits.

1. Introduction

The capacity of a service company may refer to the number of customer care representatives at a call center, the number of tellers in a bank or other types of staff in various service organizations. In any period where demand is less than capacity, the unused capacity at the end of that period cannot be carried to the next period. Tellers sitting idle today cannot start processing tomorrow's customers who have not yet arrived. In addition, Thompson and Strickland (2001) argue that unused capacity imposes a significant cost because idle capacity translates into fewer units carrying the fixed cost burden. On the other hand, if demand in a given period exceeds capacity, this may force the company to either schedule employees for overtime, hire temporary employees or let the excess demand go unmet. Not only overtime costs more but it could also have an adverse effect on employee morale and consequently customer satisfaction. Unmet demand may frustrate customers and may trigger the loss of current and future business.

The balancing act between the low and peak demand periods make capacity decisions and demand management central to the successful operation of a service company. This paper focuses on the problem of deciding on the optimum staffing levels and the redistribution of uneven demand in order to reduce the cost of overtime, temporary staffing and idle capacity. We develop and apply a mixed integer programming model to optimize the staffing and scheduling decisions of a service company with deterministic demand. A heuristic approach is also presented to provide good solutions in very short CPU time.

The paper is organized as follows. The problem definition and the motivation are given in Section 2. Section 3 summarizes the literature search. Section 4 provides the MIP model and its characterization. In Section 5, we present a heuristic approach to solve the problem. Section 6 includes the computational work as well as the treatment of special cases of the problem. The paper ends with the concluding remarks including the proposed future research.

2. Problem definition

2.1. Motivation: BST informatica

The problem studied in this paper originated in a small Brazilian telecommunications (BST) company – the name of the company is changed in compliance with the anonymity request of

* Corresponding author. Tel.: +20 2 7975356; fax: +20 2 795765.
E-mail address: aelimam@aucegypt.edu (A.A. Elimam).

[1] Part of this work was ongoing while the author was with San Francisco State University.

0377-2217/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2009.10.005

the organization. BST provides its customers with solutions for the cost-effective management of overall telecommunications expenses. BST operates two divisions: software and services. The software division is responsible for producing software used to streamline an organization's telecommunications tracking and billing systems. The services division provides on-site and off-site consulting services. These services include tracking and analyzing data collected by the BST software products, and generating recommendations on ways to reduce telecommunication expenses. The services division already accounts for nearly 50% of the annual revenues for BST. Its contribution to the company's profit is even larger because typical service contracts provide higher gross margins than that of the software products. In addition, the services division is more likely to continuously generate revenues through annual contracts with monthly payments. For these reasons, it would benefit BST to optimize its current services division operations.

At present, the services division has thirty analysts who collect technical and administrative data from customers. Then they analyze this data in order to prepare reports advising customers on potential practices to reduce telecommunication expenses. With all the tasks they need to perform, the analysts need to put in excessive amount of overtime. This has become an increasing burden for BST month after month. Fig. 1 shows the monthly working hours for six of the thirty analysts. For illustrative purposes, in the rest of the paper we will focus on these six analysts – four senior analysts and two junior analysts. The seniority level of an analyst shows the experience level of the analyst and affects the pay scale: naturally, senior analysts get paid higher than junior ones. These six analysts manage the accounts of 42 BST customers. Fig. 1 represents the current distribution of the workload based on the closing days for these 42 customers.

The main reason for the excessive overtime hours observed in Fig. 1 is the uneven distribution of, what is called, customer closing dates. A portion of the monthly workload for each customer has to be completed on a single day, which is referred to as the *closing day*. The closing day for a customer is the day of the month in which the customer's monthly account must be closed and numerous end-of-month reports must be generated.

Fig. 1 shows a major peak in the analysts' workload coinciding with the beginning of each month. This recurring peak in workload may lead to delays in deliverables to customers as well as incurring extra operational costs due to the need for overtime. A redistribution of the closing days would help BST reduce potential delays and overtime. However, not all the customers have the flexibility to select closing dates.

Current BST customers are classified into two classes based on the revenue they generate: Class 1 and Class 2 customers. Class 1 customers generate relatively lower revenue for BST. These customers have comparatively less complicated telecommunications infrastructure. As a consequence, they require smaller number of monthly activities with shorter durations. On the other hand, Class 2 customers require more detailed service including large number of tasks with longer durations. Typically, an analyst may spend 5–10 h a month for a Class 1 customer compared to 30–50 h for a Class 2 customer.

Class 2 customers usually have centralized cost accounting. For these customers, the telecommunications cost data must be available on a given day of the month. As a result, they lack the flexibility to choose their closing dates. However, Class 1 customers are smaller, more price sensitive organizations and might be willing to change their closing date given the right incentives.

This situation presented BST with two interesting questions:

- What is the optimum number of analysts of each seniority level?
- How to redistribute the workload in order to reduce overtime or temporary staffing costs while improving service level by capitalizing on the flexibility of Class 1 customers?

In an attempt to answer BST's questions, we developed a more generalized problem definition that will address the capacity and demand management issues of many service companies.

2.2. Generic problem definition

We examine companies such as BST that provide billing services to their customers. These billing services may include generating an invoice with a list of services offered, charges for these services and actions to help reduce customers' bills.

Analysts in such companies are classified into multiple levels, based on their skill, experience, and seniority. To illustrate, a company could have junior and senior analysts. Junior analysts represent the lowest tier of employees with regard to skills and experience. They are able to complete only the most basic activities for a customer. In contrast, a senior analyst is capable of processing all operational complexities for a customer.

Analysts work during normal hours, and they can also work overtime as needed. They are compensated with an additional premium for overtime work – the overtime pay to analysts being commensurate with their level. It is crucial for a service company to determine the capacity requirements expressed in the number of

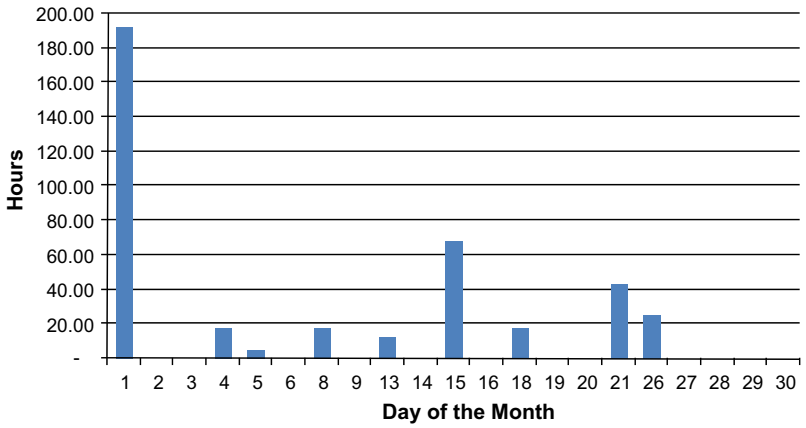

Fig. 1. Distribution of monthly workload.