Computers & Industrial Engineering 96 (2016) 16–30

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

A joint chance-constrained programming approach for call center workforce scheduling under uncertain call arrival forecasts

M. Excoffier [a], C. Gicquel [a,*], O. Jouini [b]

[a] Laboratoire de Recherche en Informatique, Université Paris Sud, Campus d'Orsay, bâtiment 650, 91405 Orsay Cedex, France
[b] Laboratoire Génie Industriel, Ecole Centrale Paris, Grande Voie des Vignes, 92290 Chatenay-Malabry, France

ARTICLE INFO

Article history:
Received 13 December 2014
Received in revised form 2 September 2015
Accepted 12 March 2016
Available online 18 March 2016

Keywords:
Personnel planning
Call center shift scheduling
Customer abandonment
Stochastic programming
Probabilistic constraints
Mixed-integer linear programming

ABSTRACT

We consider a workforce management problem arising in call centers, namely the shift-scheduling problem. It consists in determining the number of agents to be assigned to a set of predefined shifts so as to optimize the trade-off between manpower cost and customer quality of service. We focus on explicitly taking into account in the shift-scheduling problem the uncertainties in the future call arrival rates forecasts. We model them as independent random variables following a continuous probability distribution. The resulting stochastic optimization problem is handled as a joint chance-constrained program and is reformulated as an equivalent large-size mixed-integer linear program. One key point of the proposed solution approach is that this reformulation is achieved without resorting to a scenario generation procedure to discretize the continuous probability distributions. Our computational results show that the proposed approach can efficiently solve real-size instances of the problem, enabling us to draw some useful managerial insights on the underlying risk-cost trade-off.

1. Introduction

Call centers can be broadly defined as facilities designed to support the delivery of some interactive service via telephone communications (Gans, Koole, & Mandelbaum, 2003). Applications include among others telemarketing, customer service, help desk support and emergency dispatch. In most cases, the primary function of a call center is to receive phone calls that have been initiated by customers. Such operations, known as "inbound" call centers, are the topic of the present paper.

Personnel planning is a key issue in call center management. Namely, as reported in Aksin, Armony, and Mehrotra (2007), call centers are labor-intensive operations in which the cost of the staff members handling the phone calls (known as the agents) typically accounts for 60–80% of all the operating expenses. An efficient workforce management is thus crucial to achieve profitability in a call center.

Call center workforce management involves three main levels of decision-making (see e.g. Mehrotra, Ozluk, & Saltzman, 2010). Long-term planning decisions (6–12 months ahead) include the determination of how many agents to hire and train at what times based on aggregate long-term forecasts of demand for services. Short-term decisions (1–2 weeks ahead) involve the scheduling of an available pool of agents over an horizon typically spanning one week. These decisions are based on detailed short-term forecasts of agent requirements. Finally, real-time adjustment decisions, such as agent schedule updating and call routing, have to be made on a intra-day basis.

The present work is related to short-term workforce management decisions in call centers. These decisions usually involve two main steps. First, a range of possible shift patterns is defined and managers have to determine the number of agents to be assigned to each shift. Second, a rostering procedure combines shifts into rosters and assigns rosters to individual employees.

In the present paper, we focus on the first step of this process and consider the shift scheduling problem. We thus seek to determine the number of agents to be assigned to a set of predefined shifts so as to meet two potentially conflicting objectives, namely minimizing the manpower cost and delivering a high quality of service to the call center customers. To achieve this, we will look for shift schedules where the supply of agent resources is matched as closely as possible with the demand for services. Namely, understaffing would lead to customer dissatisfaction due to a poor quality of service while overstaffing would result in useless overservice to customers and higher than needed operating costs.

One of the major difficulties to be tackled while trying to match supply and demand in a call center is that the level of demand for

* Corresponding author. Tel.: +33 (0)1 69 15 42 26.
E-mail address: celine.gicquel@lri.fr (C. Gicquel).

http://dx.doi.org/10.1016/j.cie.2016.03.013
0360-8352/

services, i.e. the workload, is highly variable. This is mainly due to the fact that the call arrival rate (number of calls reaching the call center per unit of time) is subject to strong fluctuations over the course of a day or a week. This difficulty is usually handled in practice by dividing the scheduling horizon into a number of time periods of 15–60 min. In each period, the call arrival process is modeled as a Poisson process with a constant and deterministically known arrival rate and the call center is treated as an independent queuing system in stationary state (see e.g. Green, Kolesar, & Soares, 2001). This allows to determine the required staffing level for each period which is set to be the minimum number of agents for which the target quality of service (expressed e.g. as the maximum allowed proportion of customers hanging up before being answered) is reached. A deterministic optimization problem has then to be solved to find the minimum-cost shift schedule ensuring that the required number of agents is staffed in each period of the horizon.

However, as mentioned e.g. by Aksin et al. (2007), at the time when decision on shift schedules is made, call arrival rates are most often not deterministically known. We only have estimations obtained via a forecasting procedure whose outcome is a point forecast and some probabilistic representation of the forecasting errors. The input data of the shift scheduling problem are thus subject to uncertainty: not taking this into account while building the shift schedule might lead to significant discrepancies between the call center performance targeted at the time scheduling decisions are made and the one actually obtained in practice (see Gans, Shen, Zhou, & Korolev, 2015).

In the present paper, we propose a stochastic programming based approach to explicitly take into account the uncertainty on the call arrival rates in the shift scheduling problem. Our contributions are threefold. First we model the forecasting error on the call arrival rate in each period as a random variable following a continuous probability distribution. This is in contrast with most previously published approaches which rely on discrete probability distributions through the use of a finite set of scenarios to represent the uncertainty. Second, we propose to model the stochastic shift scheduling problem as a one-stage stochastic program involving a joint chance constraint. Such a model is particularly relevant when the call center is evaluated based on its ability at reaching, on all periods of the scheduling horizon, the target quality of service and when the call center management focuses on the risk of not meeting this objective. Another advantage is that it does not require introducing a penalty cost to be incurred when the target quality of service is not reached, the value of which might be difficult to estimate in practice. Third we present an efficient solution approach based on the reformulation of the stochastic program as an equivalent deterministic large-size mixed-integer linear program. This approach relies on the assumption that the random variables modeling the forecasting errors are independent from one another and comprises two main steps: (1) the reformulation of the joint chance-constrained program into an equivalent deterministic program involving a series of non-linear terms and (2) the building of a numerical representation for these non-linear terms exploiting the fact that their underlying mathematical expressions involve non-increasing piecewise constant functions. Our computational experiments show that the proposed solution approach is capable of solving real-size instances of the problem within computation times compatible with an industrial use.

The remainder of the paper is organized as follows. We provide in Section 2 an overview of the related literature. We then introduce in Section 3 the proposed joint chance-constrained stochastic program. We explain how, under the assumption of independence between the forecasting errors, it can be reformulated as a stochastic program involving a set of individual chance constraints. Section 4 is devoted to the presentation of the proposed solution approach. Its main idea is to exploit the fact that, all other parameters being fixed, the minimum number of agents needed to handle the phone calls is a non-decreasing piecewise constant function of the call arrival rate. A small illustrative example is discussed in Section 5. The results of computational experiments carried out on real-size instances are provided in Section 6.

2. Literature review

Given the size of the call center industry and the complexity associated with its operations, call centers have emerged as a fertile ground for Operations Research. We refer the reader to Aksin et al. (2007) and Gans et al. (2003) for a general introduction to this field and focus in what follows on the recently emerged research stream on stochastic call center shift scheduling. We distinguish three main features to classify the related papers: the call center setting, the representation of the uncertainty and the risk management measures.

In terms of call center architecture, the simplest case consists in a setting where a single pool of homogeneous agents handles a single class of infinitely patient calls. This amounts to using an Erlang C model to represent the call center in each period of the scheduling horizon (see Liao, Koole, van Delft, & Jouini, 2012; Liao, van Delft, & Vial, 2012). However, the importance of modeling customer impatience and abandonment in call centers has been underlined in several papers such as Gans et al. (2003) and Mandelbaum and Zeltyn (2009b). Thus, similarly to Gans et al. (2015) and Robbins and Harrison (2010), we use in the present paper a representation of the call center as an Erlang A model. For both the Erlang C and the Erlang A models, the performance evaluation of the call center can be done by exploiting analytical results available in the queuing theory literature. A more complicated setting corresponds to skill-based routing call centers. In this case, the performance evaluation of the call center has to be made by relying either on simulation or on approximations under various asymptotic regimes. Stochastic shift scheduling for skill-based routing call centers has been studied by Bodur and Luedtke (2014), Gurvich, Luedtke, and Tezcan (2010), Helber and Henken (2010) and Ye et al. (2014).

Another important feature to be considered is the way the uncertainty is described in the stochastic optimization model. In the present case, the uncertainty comes from the difficulty in exactly forecasting the future call arrival rates. The forecasting procedure provides a point estimate of the call arrival rate for each scheduling period as well as a probabilistic description of the forecasting error, e.g. in terms of a normal distribution with a zero mean and a standard deviation reflecting the forecast quality. However, in most cases, explicitly handling continuous probability distributions in a stochastic optimization problem is very computationally challenging as it requires computing multi-dimensional integrals during the solution procedure. Moreover, in the specific case under study, one is faced with an additional difficulty: the need to translate the probabilistic description of the uncertainty on the call arrival rates into a probabilistic description of the uncertainty on the agent requirements. Even for the simplest call center settings, there is no known analytical expression to directly carry out such a translation. This implies that the agent requirement in a given scheduling period is a random variable, whose probability distribution cannot be described by an analytical mathematical expression. This might explain why, to the best of our knowledge, all previously published approaches for stochastic call center shift scheduling rely on the use of discrete probability distributions to represent the uncertainty on the call arrival rates and translate each corresponding call arrival rate scenario into an agent requirement scenario in a pre-optimization step. Thus, Liao, Koole et al. (2012) and Liao, van Delft et al. (2012) consider that the