Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

CHANCE-CONSTRAINED STAFFING WITH RECOURSE FOR MULTI-SKILL CALL CENTERS WITH ARRIVAL-RATE UNCERTAINTY

Wyean Chan
Thuy Anh Ta
Pierre L'Ecuyer
Fabian Bastin

Département d'informatique et de recherche opérationnelle
Université de Montréal, Canada

ABSTRACT

We consider a two-stage stochastic staffing problem for multi-skill call centers. The objective is to minimize the total cost of agents under a chance constraint, defined over the randomness of the arrival rates, to meet all the expected service level targets. First, we determine an initial staffing based on an imperfect forecast. Then, this staffing is corrected by applying recourse when the forecast becomes more accurate. We consider the recourse actions of adding or removing agents at the price of some penalty costs. We present a method that combines simulation with integer or linear programming and cut generation.

1 INTRODUCTION

In this study, we consider an inbound multi-skill call center, where each customer requests a type of service, or *call type*. The employees who handle the calls are called *agents*, and they are often trained to serve only a subset of call types (the *skill set*). These agents are divided into *groups*, based on their skill sets. A *skill-based* router determines the matching between calls and agents. The most popular measure of quality of service is the *service level* (SL), defined as the long-term ratio of calls whose waiting times do not exceed a constant threshold, called the *acceptable waiting time*.

The *staffing problem* consists of allocating agents to each group, at minimal cost, such that the SLs are no less than some defined targets. Staffing for multi-skill call centers has been studied by several authors, see Wallace and Whitt (2005) and Cezik and L'Ecuyer (2008) for example. Bassamboo et al. (2006) and Gurvich et al. (2010) include explicitly stochastic arrival rates into their models, and they approximate the abandonment and SL functions by a fluid model (large-scale approximation). Those authors do not consider the possibility of recourse actions on the staffing to address the arrival-rate uncertainty. On the other hand, Mehrotra et al. (2010) and Gans et al. (2012) optimize the work shifts with recourse and stochastic arrival rates, but only for a simplified model with a single call type.

In this study, we propose a methodology to solve the two-stage stochastic staffing problem with recourse for multi-skill call centers with arrival-rate uncertainty. We present the model for the special case with a single period, but our approach can be extended for multiple periods.

2 THE MODEL

We consider a call center with K call types and I agent groups. Let $\Lambda = (\lambda_1, \ldots, \lambda_K)$ be the random vector of arrival rates of the K call types. If Λ follows a continuous distribution, then we approximate the distribution of Λ by Quasi-Monte Carlo sampling. Now assume that Λ is defined over a finite number of M possible realizations, say $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_M$, where $\boldsymbol{\lambda}_m = (\lambda_{1,m}, \ldots, \lambda_{K,m})$ with corresponding probability mass $p_m \geq 0$ and $\sum_{m=1}^{M} p_m = 1$. In Stage 1, the manager determines the staffing $\boldsymbol{y} = (y_1, \ldots, y_I)$ with cost $c_i > 0$ for each agent in group i, some days or weeks in advance, knowing only the distribution of Λ.

978-1-4799-7486-3/14/$31.00 ©2014 IEEE

In Stage 2, suppose that at a date near the targeted day, the manager obtains the exact arrival rates (strong assumption). Then, he can apply recourse actions based on these accurate revisions, and no further modification will be needed later on. For the scenario m, we define $r_{i,m}^{+}$ as the number of agents to add in group i, and $r_{i,m}^{-}$ as the number of agents to remove. The actual number of agents to be staffed in group i is $y_i + r_{i,m}^{+} - r_{i,m}^{-}$.

We penalize the recourse actions: adding an agent to group i increases the cost by $q_i^{+} > c_i$, whereas removing an agent reduces the cost by $q_i^{-} < c_i$. Let $g_k(\boldsymbol{\lambda},\boldsymbol{y})$ and $g(\boldsymbol{\lambda},\boldsymbol{y})$ be the SL functions for call type k and the aggregated SL over all calls. The corresponding minimal SL thresholds are l_k and l. If the manager wants to satisfy all the SL constraints for a fraction $\Phi \in [0,1]$ of the M scenarios, then the chance constraint is $\mathbb{P}_{\boldsymbol{\lambda}}\left[g(\boldsymbol{\lambda}_m,\boldsymbol{y}+\boldsymbol{r}_m^{+}-\boldsymbol{r}_m^{-}) \geq l \text{ and } g_k(\boldsymbol{\lambda}_m,\boldsymbol{y}+\boldsymbol{r}_m^{+}-\boldsymbol{r}_m^{-}) \geq l_k, \forall k\right] \geq \Phi$, where $\mathbb{P}_{\boldsymbol{\lambda}}$ is the probability measure on λ. Let ϕ_m be a binary variable that controls the SL feasibility of scenario m, where $\phi_m = 1$ if scenario m must be feasible, else $\phi_m = 0$. Let $\boldsymbol{r}_m^{+} = (r_{1,m}^{+},\ldots,r_{I,m}^{+})$ and $\boldsymbol{r}_m^{-} = (r_{1,m}^{-},\ldots,r_{I,m}^{-})$. The problem is:

$$
\begin{aligned}
\min \quad & \sum_{m=1}^{M} p_m \phi_m \left[\sum_{i=1}^{I} (c_i y_i + q_i^{+} r_{i,m}^{+} - q_i^{-} r_{i,m}^{-})\right] \\
\text{subject to:} \quad & \\
& \sum_{m=1}^{M} p_m \phi_m \geq \Phi, \\
& g(\boldsymbol{\lambda}_m,\boldsymbol{y}+\boldsymbol{r}_m^{+}-\boldsymbol{r}_m^{-}) \geq l\phi_m, && \forall m, \\
& g_k(\boldsymbol{\lambda}_m,\boldsymbol{y}+\boldsymbol{r}_m^{+}-\boldsymbol{r}_m^{-}) \geq l_k\phi_m, && \forall m, \forall k, \\
& y_i + r_{i,m}^{+} \geq r_{i,m}^{-}, && \forall i, \forall m, \\
& \phi_m \in \{0,1\}, && \forall m, \\
& y_i, r_{i,m}^{+}, r_{i,m}^{-} \geq 0, \text{ and integer }, && \forall i, \forall m.
\end{aligned}
\tag{P1}
$$

3 METHODOLOGY

To solve (P1), we replace the constraints on g and g_k by a set of linear cuts. We add new linear cuts iteratively using the subgradients of the SL functions, estimated by simulation, until we obtain a feasible solution. We execute a simulation-based local search to refine the final solution and reduce the error from simulation noise. We propose an extension to the simulation-based algorithm using linear programming of Ceẓik and L'Ecuyer (2008) by integrating the recourse variables $r_{i,m}^{+}$ and $r_{i,m}^{-}$, and the binary variables ϕ_m for the chance constraint. In this extension, our method optimizes multiple scenarios simultaneously. Numerical results show our method performs much better than a classic two-step method (first, find the staffing requirement of each scenario, then optimize with recourse) as the penalty costs q_i^{+} and q_i^{-} increase.

REFERENCES

Bassamboo, A., J. M. Harrison, and A. Zeevi. 2006. "Design and Control of a Large Call Center: Asymptotic Analysis of an LP-Based Method". *Operations Research* 54 (3): 419–435.

Ceẓik, M. T., and P. L'Ecuyer. 2008. "Staffing Multiskill Call Centers via Linear Programming and Simulation". *Management Science* 54 (2): 310–323.

Gans, N., H. Shen, Y.-P. Zhou, N. Korolev, A. McCord, and H. Ristock. 2012. "Parametric Stochastic Programming Models for Call-Center Workforce Scheduling". Working paper.

Gurvich, I., J. Luedtke, and T. Tezcan. 2010. "Staffing Call Centers with Uncertain Demand Forecasts: A Chance-Constrained Optimization Approach". *Management Science* 56 (7): 1093–1115.

Mehrotra, V., O. Ozlük, and R. Saltzman. 2010. "Intelligent Procedures for Intra-day Updating of Call Center Agent Schedules". *Production and Operations Management* 19 (3): 353–367.

Wallace, R. B., and W. Whitt. 2005. "A staffing algorithm for call centers with skill-based routing". *Manufacturing and Service Operations Management* 7 (4): 276–294.