The Generalized Bin Packing Problem under Uncertainty

ROBERTO TADEI, GUIDO PERBOLI
Department of Control and Computer Engineering, Politecnico di Torino
Corso Duca degli Abruzzi, 24, I-10129, Turin, ITALY
{roberto.tadei, guido.perboli}@polito.it

Abstract: In this paper we introduce the Generalized Bin Packing Problem under Uncertainty, a new packing problem where, given a set of items characterized by volume and stochastic revenue and a set of bins characterized by volume and cost, we want to select a subset of items to be loaded into a subset of bins which maximizes the total profit, given by the difference between the expected total revenue of the loaded items and the total cost of the used bins, by satisfying the volume and bin availability constraints. The item revenues are random variables with unknown probability distribution. By using the asymptotic theory of extreme values a nonlinear integer deterministic model is derived.

Key-Words: packing, stochastic revenue, asymptotic approximation, nonlinear integer deterministic model.

1 Introduction

Given a set of items characterized by volume and stochastic revenue and a set of bins characterized by volume and cost, the Generalized Bin Packing Problem under Uncertainty (GBPPu) chooses a subset of items to be loaded into a subset of bins which maximizes the total profit, given by the difference between the expected total revenue of the loaded items and the total cost of the used bins, by satisfying the volume and bin availability constraints. The item revenues are random variables with unknown probability distribution. They are composed by a deterministic revenue plus a random term, which represents the revenue oscillations due to the handling operations of the bins which are necessary for item loading and dispatching.

The GBPPu frequently arises in real-life applications, in particular in logistics, where the freight consolidation is essential to optimize the delivery process. In this case, a series of handling operations for the bins must be performed at the logistic platforms and these operations could highly affect the final total revenue of the loading [8].

In this paper we introduce a stochastic model for the GBPPu. In most papers dealing with uncertainty, the probability distribution of the random variables is given and their expected value can be calculated. This is not the case of the GBPPu, where the probability distribution of the stochastic item revenue is unknown, because it is unmeasurable in practice and any strong assumption on its shape would be arbitrary. We show that, by using some results of the asymptotic theory of extreme values, the probability distribution of the maximum item revenue becomes a Gumbel (or double exponential) distribution and the expected total revenue of the loaded items can be easily calculated.

2 Literature review

The GBPPu is the stochastic version of the Generalized Bin Packing Problem (GBPP), introduced by Baldi et al. [1] as a generalization of the Variable Cost and Size Bin Packing Problem [2]. The GBPPu literature does not exist yet, being this paper the first one on this problem. Also the literature on the GBBP is quite limited, due to its recent introduction. For this reason, in the following we will also consider some relevant literature on a similar problem, which is a special case of the GBPPu, i.e. the Stochastic Bin Packing Problem.
In the classical Bin Packing Problem we have a set of bins with given volume and a set of items to be compulsorily loaded. The problem aims to load all items in the minimum number of bins. In the Stochastic Bin Packing Problem the source of uncertainty is usually the item volume [3, 5, 6, 7]. These papers mainly deal with on-line algorithms for solving the Stochastic Bin Packing Problem where strong hypotheses on the probability distribution of the random terms are usually done.

3 The GBPPu model

Let us consider the following parameters and data

- $I$: set of items
- $T$: set of bin types
- $J_t$: set of bins of type $t \in T$
- $L_t$: set of handling operating scenarios for bins of type $t \in T$
- $w_i$: volume of item $i \in I$
- $r_{ij}^t$: revenue of loading item $i \in I$ into bin $j \in J_t$, $t \in T$
- $U_t$: maximum number of bins of type $t \in T$
- $C_t$: cost of a bin of type $t \in T$
- $W_t$: volume of a bin of type $t \in T$
- $U$: maximum number of bins of any type

and variables

- $x_{ij}^t$: item-to-bin assignment boolean variable which is equal to 1 if item $i \in I$ is loaded into bin $j \in J_t$, $t \in T$, 0 otherwise
- $y_{ij}^t$: bin selection boolean variable which is equal to 1 if bin $j \in J_t$, $t \in T$ is selected, 0 otherwise
- $\theta^{tl}$: random revenue of loading any item into a bin of type $t \in T$ under handling operating scenario $l \in L_t$.

Let us assume, as usual, that $\theta^{tl}$ are independent and identically distributed (i.i.d.) random variables with a common probability distribution

$$\Pr\{\theta^{tl} \leq x\} = F(x)$$

(1)

The main feature of our approach consists in considering such probability distribution as unknown as the random revenues are extremely difficult to be measured in practice.

Let $\tilde{r}_{ij}^{tl}(\theta^{tl})$ be the random revenue of loading item $i$ into a bin $j$ of type $t$ under handling operating scenario $l$ given by

$$\tilde{r}_{ij}^{tl}(\theta^{tl}) = r_{ij}^t + \theta^{tl}, \ i \in I, j \in J_t, t \in T, l \in L_t$$

(2)

Let us define with $\overline{\theta}^t$ the maximum of the random revenues under the alternative handling operating scenarios $l \in L_t$

$$\overline{\theta}^t = \max_{l \in L_t} \theta^{tl}, \ t \in T$$

(3)

$\overline{\theta}^t$ is still of course a random variable with unknown probability distribution (because $F(x)$ is not known) given by

$$B_t(x) = \Pr\{\overline{\theta}^t \leq x\}$$

(4)
As $\theta^l \leq x \iff \theta^l \leq x$, $l \in L_t$ and $\theta^l$ are independent, using (1) one gets
\[ B_l(x) = \prod_{l \in L_t} \Pr\{\theta^l \leq x\} = \prod_{l \in L_t} F(x) = [F(x)]^{L_t} \]  
(5)

We assume that the bin loading policies are efficiency-based so that, among the alternative handling operating scenarios $l \in L_t$, the one which maximizes the random revenues $\tilde{r}^l_{ij}(\theta^l)$ will be selected, then
\[ \tilde{r}^l_{ij}(\theta^l) = \max_{l \in L_t} \tilde{r}^l_{ij}(\theta^l) = r^l_{ij} + \max_{l \in L_t} \theta^l = r^l_{ij} + \theta^l, \quad i \in I, j \in J_t, t \in T \]  
(6)

The GBPPu is then formulated as follows

\[
\begin{align*}
\text{Maximize}_{y} & \quad \left\{ \mathbb{E}_{\theta^l} \left[ \max_{x} \sum_{i \in I} \sum_{t \in T} \sum_{j \in J_t} r^l_{ij}(\theta^l)x^t_{ij} \right] - \sum_{t \in T} \sum_{j \in J_t} C_t y^t_j \right\} \\
\text{Subject to} & \quad \sum_{i \in I} w_i x^t_{ij} \leq W_t y^t_j \quad j \in J_t, t \in T \\
& \quad \sum_{t \in T} \sum_{j \in J_t} x^t_{ij} \leq 1 \quad i \in I \\
& \quad \sum_{j \in J_t} y^t_j \leq U_t \quad t \in T \\
& \quad \sum_{t \in T} \sum_{j \in J_t} y^t_j \leq U \\
& \quad y^t_j \in \{0, 1\} \quad j \in J_t, t \in T \\
& \quad x^t_{ij} \in \{0, 1\} \quad i \in I, j \in J_t, t \in T 
\end{align*}
\]  
(7)-(13)

The objective function (7) maximizes the total profit, given by the expected maximum total revenue yielded by the loaded items minus the total cost of the used bins. Constraints (8) have the double effect of linking the usage of the bins to the accommodation of items into them and to limit the capacity of each used bin. Constraints (9) ensure that each item, if loaded, is loaded into no more than one bin. Constraints (10) enforce the maximum number of bins of each type that can be used, while constraint (11) limits the total number of selected bins. Finally, (12)-(13) are the integrality constraints.

### 4 Solving the problem

In this section we derive the form of the objective function (7) by showing that it depends on the unknown probability distribution $F(x)$. Let us assume that a bin selection $\{y^t_j\}$ is already given. It is obvious that the optimal values of the assignment variables $\{x^t_{ij}\}$ of problem (7)-(13) are
\[
\begin{align*}
  x^t_{ij} = \begin{cases} 
    1, & \text{if } (j, t) = \arg \max_{s \in T, k < J_t} y^t_k = 1 \tilde{r}^s_{ik}(\theta^l) \\
    0, & \text{otherwise}
  \end{cases}
\end{align*}
\]  
(14)

Because of (6), the maximum random revenue $\tilde{r}^l_{ij}(\tilde{\theta})$ for any item $i$ among the alternatives $(j, t)$ becomes
\[
\tilde{r}^l_{ij}(\tilde{\theta}) = \max_{t \in T, j \in J_t} r^l_{ij}(\tilde{\theta}) = \max_{t \in T, j \in J_t} \left( r^l_{ij} + \theta^l \right) 
\]  
(15)

Due to (14) and (15), the objective function (7) becomes
\[
\begin{align*}
\text{Maximize}_{y} & \quad \left\{ \sum_{i \in I} \mathbb{E}_{\theta^l} \left[ \tilde{r}^l_i(\tilde{\theta}) \right] - \sum_{t \in T} \sum_{j \in J_t} C_t y^t_j \right\}
\end{align*}
\]  
(16)
The calculation of the expected value in (16) would require the knowledge of the probability distribution of $\tau_i(\theta')$, i.e.,

$$G_i(x) = Pr\left\{\tau_i(\theta') \leq x\right\} = Pr\left\{\max_{t \in T, j \in J_t: y^j_i = 1} \left(r_{ij}^t + \theta'\right) \leq x\right\}$$  \hspace{1cm} (17)

As $\max_{t \in T, j \in J_t: y^j_i = 1} \left(r_{ij}^t + \theta'\right) \leq x \iff r_{ij}^t + \theta' \leq x, \ t \in T, j \in J_t : y^j_i = 1$, and the random variables $\theta'$ are independent (because $\theta^t_i$ are independent), due to (5), (17) becomes

$$G_i(x) = \prod_{t \in T} \prod_{j \in J_t: y^j_i = 1} Pr\left\{r_{ij}^t + \theta' \leq x\right\} = \prod_{t \in T} \prod_{j \in J_t: y^j_i = 1} Pr\left\{\theta' \leq x - r_{ij}^t\right\} =$$

$$= \prod_{t \in T} \prod_{j \in J_t: y^j_i = 1} B_t \left(x - r_{ij}^t\right) = \prod_{t \in T} \prod_{j \in J_t: y^j_i = 1} \left[F \left(x - r_{ij}^t\right)\right]^{\left|L_t\right|}$$  \hspace{1cm} (18)

Unfortunately, the unknown probability distribution $F(x)$ prevents the computation of $G_i(x)$ in (18). An efficient way to cope with this problem is introduced in the next section.

5 The asymptotic approximation of $G_i(x)$

In order to get an explicit form for $G_i(x)$ we consider its asymptotic approximation. The method we use is based on the following observation. Under a mild assumption on the shape of the unknown probability distribution $F(x)$, the probability distribution $G_i(x)$ tends towards a specific functional form as the number of alternative handling operating scenarios $\left|L_t\right|, t \in T$ becomes large.

Following Galambos [4], one can prove that the only condition requested for $F(x)$ is that it is asymptotically exponential in its right tail, i.e. there is a constant $\beta > 0$ such that

$$\lim_{y \to +\infty} \frac{1 - F(x + y)}{1 - F(y)} = e^{-\beta x}, \ \forall x \in R$$  \hspace{1cm} (19)

This is a very mild condition for the probability distribution $F(x)$ as we observe that many probability distributions show such behavior, among them the widely used distributions Gamma, Gumbel, Laplace, and Logistic.

Let us consider the following theorem, which provides the desired approximation for $G_i(x)$ (proof omitted).

**Theorem 1** Under assumption (19), the unknown probability distribution $G_i(x)$ becomes

$$G_i(x) = \exp \left(-A_i e^{-\beta x}\right)$$  \hspace{1cm} (20)

where

$$A_i = \sum_{t \in T} \sum_{j \in J_t: y^j_i = 1} e^{\beta r_{ij}} = \sum_{t \in T} \sum_{j \in J_t} y^j_i e^{\beta r_{ij}}$$  \hspace{1cm} (21)

is the "accessibility" of item $i$ to the overall set of bins $j$ of type $t$.

6 The deterministic approximation of the GBPPu

Having now an explicit form for $G_i(x)$, we can calculate $E_{\theta'} \left[\tau_i(\theta')\right]$ in (16) as follows

$$\tilde{r}_i = E_{\theta'} \left[\tau_i(\theta')\right] = \int_{-\infty}^{+\infty} x dG_i(x) = -\int_{-\infty}^{+\infty} x \exp \left(-A_i e^{-\beta x}\right) A_i e^{-\beta x} \beta dx$$  \hspace{1cm} (22)
Substituting for $t = A_i e^{-\beta x}$ one gets
\[
\tilde{r}_i = -1/\beta \int_0^{+\infty} \ln(t/A_i) e^{-t} dt = -1/\beta \int_0^{+\infty} e^{-t} \ln t dt + 1/\beta \ln A_i \int_0^{+\infty} e^{-t} dt = \\
\gamma/\beta + 1/\beta \ln A_i = 1/\beta (\ln A_i + \gamma)
\tag{23}
\]
where $\gamma = -\int_0^{+\infty} e^{-t} \ln t dt \simeq 0.5772$ is the Euler constant.

By substituting (23) in (16), the deterministic approximation of the GBPPu becomes the following nonlinear integer deterministic problem

Maximize \[y \left( \frac{1}{\beta} \sum_{i \in I} \ln A_i - \sum_{t \in T} \sum_{j \in J_t} C_{tij} y_{tj}^l \right) \tag{24}\]
Subject to \[\sum_{i \in I} w_i x_{tij}^l \leq W_{tij}^l \quad j \in J_t, t \in T \tag{25}\]
\[\sum_{t \in T} \sum_{j \in J_t} x_{tij}^l \leq 1 \quad i \in I \tag{26}\]
\[\sum_{j \in J_t} y_{tj}^l \leq U_t \quad t \in T \tag{27}\]
\[\sum_{t \in T} \sum_{j \in J_t} y_{tj}^l \leq U \tag{28}\]
\[y_{tj}^l \in \{0, 1\} \quad j \in J_t, t \in T \tag{29}\]
\[x_{tij}^l \in \{0, 1\} \quad i \in I, j \in J_t, t \in T \tag{30}\]

Problem (24)-(30) can be solved by some exact methods (e.g. Branch and Bound). In alternative, for large size instances approximate solutions can be obtained by using some heuristics.

Let us notice that the above problem needs to know a proper value of the positive constant $\beta$ in (24), which can be obtained by calibration.

7 Conclusion

In this paper we have addressed a new packing problem where, given a set of items characterized by volume and stochastic revenue and a set of bins characterized by volume and cost, we want to select a subset of items to be loaded into a subset of bins which maximizes the total profit, given by the difference between the expected total revenue yielded by the loaded items and the total cost of the used bins, by satisfying the volume and bin availability constraints. To the authors’ knowledge, this paper is the first one which introduces stochasticity in the Generalized Bin Packing Problem. The paper shows that, under a mild assumption on the shape of probability distribution of the stochastic item revenue, the unknown probability distribution of the maximum stochastic item revenue converges to a Gumbel distribution and a nonlinear deterministic approximation of the problem can be derived.

References


