An Asymptotic Approximation Scheme for the Concave-cost Bin Packing Problem

Joseph Y.-T. Leung

Department of Computer Science
New Jersey Institute of Technology
Newark, NJ 07102, U.S.A.
Email: leung@cis.njit.edu

Chung-Lun Li

Department of Logistics
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
Email: lgtclli@polyu.edu.hk

February 2007

Revised: August 2007

Abstract

We consider a generalized one-dimensional bin packing model in which the cost of a bin is a nondecreasing concave function of the utilization of the bin. We show that for any given positive constant ϵ, there exists a polynomial-time approximation algorithm with an asymptotic worst-case performance ratio of no more than $1 + \epsilon$.

Keywords: Bin packing; Concavity; Asymptotic worst-case analysis

1 Introduction

We consider a generalized one-dimensional bin packing problem where the objective is to minimize a concave cost function of the bin utilization. The problem can be defined as follows: Given a list of n items $L = (a_1, a_2, \ldots, a_n)$, where each item has size $a_i \in (0, 1]$, our goal is to pack these items into unit-capacity bins so as to minimize the total cost. The cost of each bin, $f(x)$, is a nondecreasing concave function of the total size x of the items assigned to that bin, where $f(0) = 0$. Our objective is to minimize the total cost of all the nonempty bins in the packing. We assume that $a_1, a_2, \ldots, a_n$ are rational numbers. This problem has applications in packing of goods for mixed truckload and less-than-truckload services, and in production scheduling with learning effects [8].

Li and Chen [8] have considered four heuristics for this problem: First Fit (FF), Best Fit (BF), First Fit Decreasing (FFD), and Best Fit Decreasing (BFD). They show that FF and BF obey an absolute worst-case bound of 2, while FFD and BFD obey an absolute worst-case bound of 1.5. Both bounds are the best possible. In addition, they conduct some computational experiments to test the performance of these heuristics. Their results show that FFD and BFD perform much better than FF and BF, BF performs slightly better than FF, while FFD and BFD perform almost identically.

In this note, we consider the *asymptotic* worst-case performance ratio, rather than the *absolute* worst-case performance ratio. We show that for any $\epsilon > 0$, there exists a polynomial-time algorithm A_ϵ with an asymptotic worst-case performance ratio of $1 + \epsilon$, that is, $Z(A_\epsilon) \leq (1 + \epsilon)Z^* + K_\epsilon$, where $Z(A_\epsilon)$ is the total cost of the solution generated by the algorithm, Z^* is the total cost of the optimal solution, and K_ϵ is a constant depending only on ϵ.

Our problem is a generalization of the classical bin packing problem. The classical bin packing problem has been studied extensively; see, for example, the survey paper by Coffman *et al.* [2] as well as the more recent developments in [1, 3, 6, 9]. Fernandez de la Vega and Lueker [4] have shown that classical bin packing can be solved asymptotically within $1 + \epsilon$ in linear time. As we shall see later, our result is an adaptation of the framework of Fernandez de la Vega and Lueker's approximation scheme, but with a new bin assignment method and a new lower bound on the optimal solution value.

2 The Main Result

We first state an important property of the concave-cost bin packing problem.

Property 1 *(i)* $f(x+\Delta)+f(y-\Delta) \leq f(x)+f(y)$, *for any* $\Delta > 0$ *and any* x, y *such that* $x \geq y \geq 0$. *(ii)* $Z^* \geq \sum_{i=1}^{n} a_i \cdot f(1)$.

Proof: See Lemmas 1 and 2 of [8]. ∎

Note that Property 1(i) can be interpreted as follows: Suppose Bin 1 currently has a fill-level of x and Bin 2 currently has a fill-level of y, where $x \geq y$. Then, increasing the fill-level of Bin 1 while simultaneously decreasing the fill-level of Bin 2 by the same amount will not increase the total cost of the solution. This implies the following: Suppose we need to add a new item to the existing packing. Then, it is optimal to assign it to the bin with the highest fill-level (among those bins into which it will fit).

For any given list Λ, let $S(\Lambda)$ denote the total size of all the items in the list. Let $\epsilon > 0$ be a given constant. Let $\epsilon_1 < \min\{\epsilon, 1\}$ be a positive constant with the exact value defined later. We divide the list L into two sublists, namely $L_1 = (b_1, b_2, \ldots, b_{n_1})$ and $L_2 = (c_1, c_2, \ldots, c_{n_2})$, where each item in L_1 has size less than ϵ_1 and each item in L_2 has size greater than or equal to ϵ_1. We consider two cases: "$S(L) \geq n_2$" and "$S(L) < n_2$."

We first consider Case 1: $S(L) \geq n_2$. In this case we first assign the items in L_2, with one item per bin. This will take n_2 bins. We then assign the items in L_1 in a first-fit manner, that is, we assign the items according to their indices, and we always assign an item to the first (lowest indexed) bin into which it will fit. Let α be the number of nonempty bins in the resulting packing. Note that $\alpha \geq n_2$. The total cost of this packing, denoted $Z(A_\epsilon)$, is at most $\alpha \cdot f(1)$. Furthermore, in this packing, every bin, except the last one, has a fill-level of at least $1 - \epsilon_1$. Thus, the total size of the items must be at least $(\alpha - 1)(1 - \epsilon_1)$. By Property 1(ii), $Z^* \geq (\alpha - 1)(1 - \epsilon_1)f(1)$. Therefore, if we select ϵ_1 in such a way that $\epsilon_1 \leq \epsilon/(1 + \epsilon)$, then

$$\begin{aligned} Z(A_\epsilon) &\leq (\alpha - 1)f(1) + f(1) \\ &\leq \frac{Z^*}{1 - \epsilon_1} + f(1) \\ &\leq (1 + \epsilon)Z^* + f(1). \end{aligned}$$

Next, we consider Case 2: $S(L) < n_2$. We first show that we can obtain a $(1 + \epsilon)$-approximation in polynomial time when L_2 contains only m distinct item sizes, where m is a constant. We then show that the original problem can be reduced to one in which L_2 contains m_ϵ distinct item sizes, where m_ϵ is a constant depending only on ϵ.

2.1 When L_2 contains only m distinct item sizes

In this subsection, we consider the situation where the sizes of the elements in L_2 have m distinct values. Since the smallest item in L_2 has size greater than or equal to ϵ_1, there are at most a constant number q of possible bin types (i.e., the multiset of items from L_2 that are assigned to a bin). Consequently, there are at most $O(n_2^q)$ possible packings of the elements in L_2 (see p. 65 of [2]). For each packing P, we sort the bins in descending order of their fill-levels. We then treat the items in L_1 as one big "breakable" item of size $S(L_1)$, and assign it to the bins as follows: We fill the level of the first bin with the big item until it is full, and cut the big item at the boundary. We then fill the level of the second bin with the big item until it is full, and again cut the big item at the boundary. We repeat this process until the entire big item has been assigned. At the end, we can evaluate the total cost of the packing. For example, suppose $L_1 = (0.10, 0.10, 0.20)$ and $L_2 = (0.36, 0.36, 0.42, 0.42)$, and we consider the packing P depicted in Figure 1(a). In this example, the big item has size $S(L_1) = 0.40$ and is assigned to bins 1 and 2 as shown in Figure 1(b). The total cost of this packing is $f(1) + f(0.96)$.

Among all the $O(n_2^q)$ possible packings, we select the one with the lowest total cost (with ties broken arbitrarily). Let $\hat{P}$ denote the packing selected. Let $\hat{Z}$ denote the total cost of packing $\hat{P}$, and let α be the number of nonempty bins utilized in $\hat{P}$. Obviously, $\hat{P}$ is not a legitimate packing in that the items of L_1 (which form the big item) are not packed legitimately. We now re-pack the items in L_1. This is done by first removing the big item, and then assigning the items in L_1 to the unoccupied space of the bins in a first-fit manner. Let $\bar{P}$ be the resulting packing, and let β be the number of nonempty bins utilized in $\bar{P}$. Clearly, $\beta \geq \alpha$. For example, suppose the packing depicted in Figure 1(b) is the selected packing $\hat{P}$, with $\alpha = 2$. Then, the resulting packing $\bar{P}$ is depicted in Figure 1(c), with $\beta = 3$.

Property 2 *$\hat{Z} \leq Z^*$.*

Proof: Let δ be the greatest common divisor of $a_1, a_2, \ldots, a_n, 1$. Suppose we split each of the items in L_1 into multiple items of size δ (that is, an item of size a_i is replaced by a_i/δ items of size δ). Then the total cost of optimal solution of the resulting "relaxation problem" must be a lower bound of the optimal solution value of the original problem. One way to solve this relaxation problem is to try all possible packings of the items of L_2, and then in each such possible packing, we pack the remaining δ-sized items optimally. By Property 1(i), the optimal way to pack the remaining items is to assign them to those bins with the highest fill-levels. Since an item of size δ can fit into any empty bin space, this method of assigning the δ-sized items will result in the same packing as $\hat{P}$. Therefore, $\hat{Z}$ must be a lower bound of Z^*. ∎

If $\beta > \alpha$, then every bin in $\bar{P}$, except the last one, is filled to a level of at least $1 - \epsilon_1$, and we can resort to the same argument as in Case 1 to show that the total cost of $\bar{P}$ is at most $\frac{Z^*}{1-\epsilon_1} + f(1)$. Thus, we focus on the case in which $\beta = \alpha$. Consider the packing $\bar{P}$. Let B_1 be the set of bins that either contain some items from L_1 or have a fill-level of more than $1 - \epsilon_1$. Let B_2 be the set of bins that contain only items from L_2 and have a fill-level no more than $1 - \epsilon_1$. We have $\alpha = \beta = |B_1| + |B_2|$. Let us compare the total cost of the bins in the two packings $\bar{P}$ and $\hat{P}$.

The total cost of the bins in B_2 under the packing $\bar{P}$, denoted Z_{B_2}, is the same as that under the packing $\hat{P}$, because these bins do not contain any items from L_1 in either of the two packings. On the other hand, every bin in B_1, except the last one, is filled to a level of at least $1-\epsilon_1$ under the packing $\bar{P}$. This implies that the total size of the items in those bins is at least $(1-\epsilon_1)(|B_1|-1)$. Applying Property 1(ii) to these items, we know that in the packing $\hat{P}$, the total cost of the bins in B_1 must be at least $(1-\epsilon_1)(|B_1|-1)f(1)$. Hence, the total cost of the packing $\bar{P}$ is at most $Z_{B_2}+|B_1|f(1) \leq \frac{1}{1-\epsilon_1}\big[Z_{B_2}+(1-\epsilon_1)(|B_1|-1)f(1)\big]+f(1) \leq \frac{1}{1-\epsilon_1}\hat{Z}+f(1) \leq \frac{1}{1-\epsilon_1}Z^*+f(1)$, where the last inequality follows from Property 2.

2.2 Reducing the original problem to the one where L_2 has m_ϵ distinct item sizes

We now turn our attention to the reduction of our original problem to the one in which L_2 has m_ϵ distinct item sizes. Recall that $L_2=(c_1,c_2,\ldots,c_{n_2})$. We reindex the items in such a way that $c_1 \leq c_2 \leq \cdots \leq c_{n_2}$. Let $m=\lceil \epsilon_1^{-2}\rceil+1$. Let $n_2=mh+r$, where $h=\lfloor n_2/m\rfloor$ and $0 \leq r \leq m-1$. Consider the list

$$L_2'=(c_1,\ldots,c_1;c_{h+1},\ldots,c_{h+1};c_{2h+1},\ldots,c_{2h+1};\ldots;c_{(m-1)h+1},\ldots,c_{(m-1)h+1};c_{mh+1},c_{mh+2},\ldots,c_{n_2}),$$

which consists of h copies of c_1, followed by h copies of c_{h+1}, followed by h copies of c_{2h+1}, and so on. Finally, it has h copies of $c_{(m-1)h+1}$, followed by the last r elements of L_2. We also consider the list

$$L_2''=(c_{h+1},\ldots,c_{h+1};c_{2h+1},\ldots,c_{2h+1};\ldots;c_{(m-1)h+1},\ldots,c_{(m-1)h+1};1,\ldots,1;c_{mh+1},c_{mh+2},\ldots,c_{n_2}),$$

which consists of h copies of c_{h+1}, followed by h copies of c_{2h+1}, and so on. Finally, it has h copies of $c_{(m-1)h+1}$, followed by h unit-sized items, plus the last r elements of L_2. Clearly, the i-th item of L_2' is no larger than the i-th item of L_2, which in turn is no larger than the i-th item of L_2'', for $i=1,2,\ldots,n_2$. Thus, the optimal packing of $L_1 \cup L_2'$ has total cost no higher than that of $L_1 \cup L_2$, and the optimal packing of $L_1 \cup L_2$ has total cost no higher than that of $L_1 \cup L_2''$. In other words,

$$Z' \leq Z^* \leq Z'', \tag{1}$$

where Z' and Z'' denote the total costs in the optimal packings of $L_1 \cup L_2'$ and $L_1 \cup L_2''$, respectively.

For example, suppose $m = 4$ and

$$L_2 = (0.11, 0.13, 0.17, 0.21, 0.25, 0.28, 0.32, 0.36, 0.45, 0.52, 0.63, 0.67, 0.72, 0.77, 0.81),$$

where $n_2 = 15$. Then $h = \lfloor n_2/m \rfloor = 3$. The list

$$L_2' = (0.11, 0.11, 0.11, 0.21, 0.21, 0.21, 0.32, 0.32, 0.32, 0.52, 0.52, 0.52, 0.72, 0.77, 0.81)$$

consists of three copies of $c_1 = 0.11$, three copies of $c_4 = 0.21$, three copies of $c_7 = 0.32$, three copies of $c_{10} = 0.52$, followed by $c_{13} = 0.72$, $c_{14} = 0.77$, and $c_{15} = 0.81$. The list

$$L_2'' = (0.21, 0.21, 0.21, 0.32, 0.32, 0.32, 0.52, 0.52, 0.52, 1, 1, 1, 0.72, 0.77, 0.81)$$

consists of three copies of $c_4 = 0.21$, three copies of $c_7 = 0.32$, three copies of $c_{10} = 0.52$, three copies of 1, followed by $c_{13} = 0.72$, $c_{14} = 0.77$, and $c_{15} = 0.81$.

Note that the two lists L_2' and L_2'' differ only in that, when going from L_2' to L_2'', we change h items of size c_1 to h items of size 1. Hence, given a packing of $L_1 \cup L_2'$, we can easily obtain a packing of $L_1 \cup L_2''$ by using no more than h additional bins. This implies that $Z'' \leq Z' + h \cdot f(1)$. Because the total size of the items in L_2' is at least $mh\epsilon_1$, we have $Z' \geq mh\epsilon_1 f(1) \geq h \cdot f(1)/\epsilon_1$. Thus,

$$Z'' \leq (1 + \epsilon_1) Z'. \tag{2}$$

Now, we can solve the original problem as follows: Solve the problem using the item sizes given by $L_1 \cup L_2''$, and restore the sizes of the items in L_2'' with their original item sizes. The total cost of this solution, $Z(A_\epsilon)$, satisfies the following:

$$\begin{aligned}
Z(A_\epsilon) &\leq \frac{1}{1 - \epsilon_1} Z'' + f(1) \\
&\leq \frac{1 + \epsilon_1}{1 - \epsilon_1} Z' + f(1) \qquad \text{(by (2))} \\
&\leq \frac{1 + \epsilon_1}{1 - \epsilon_1} Z^* + f(1) \qquad \text{(by (1))}.
\end{aligned}$$

Hence, if we select ϵ_1 in such a way that $\epsilon_1 \leq \frac{\epsilon}{2+\epsilon}$, then we get $Z(A_\epsilon) \leq (1 + \epsilon) Z^* + f(1)$.

2.3 Summary

Summarizing the above analysis, we conclude that if we select $\epsilon_1 = \min\left\{\frac{\epsilon}{2+\epsilon}, 1\right\}$, then in both Cases 1 and 2, we have $Z(A_\epsilon) \leq (1+\epsilon)Z^* + f(1)$. In Case 2 of the above procedure, evaluating the total cost of each possible packing P requires $O(n \log n)$ time. List L_2'' contains items with only m_ϵ distinct values and each bin can fill at most $1/\epsilon_1$ items from L_2''. Thus, there are $O(m_\epsilon^{1/\epsilon_1})$ possible ways a bin can fill those items from L_2''. Hence, there are at most $O\big(n^{m_\epsilon^{1/\epsilon_1}}\big)$ possible packing P. Therefore, the running time of our algorithm is polynomial in n.

3 Concluding Remarks

We have presented an asymptotic polynomial-time approximation scheme for the bin packing problem with concave costs of bin utilization. However, the time bound of our approximation scheme is a "high order polynomial." Fernandez de la Vega and Lueker [4] have developed a linear time asymptotic approximation scheme for the classical bin packing problem. It is an interesting open question whether a similar "low order polynomial" approximation scheme can be developed for the bin packing problem with concave costs of bin utilization. Karmarkar and Karp [7] have presented an asymptotic fully polynomial-time approximation scheme for the classical bin packing problem (see also [5]). It is also a challenging open question whether an asymptotic fully polynomial-time approximation scheme exists for our problem.

Acknowledgments

This research was supported in part by the Research Grants Council of Hong Kong under grant PolyU5312/04E. The work of the first author was supported in part by the National Science Foundation under grants DMI-0300156 and DMI-0556010.

References

[1] C. Alves, J.M. Valerio de Carvalho, Accelerating column generation for variable sized bin-packing problems, European Journal of Operational Research, 183 (2007) 1333–1352.

[2] E.G. Coffman, Jr., M.R. Garey, D.S. Johnson, Approximation algorithms for bin packing: A survey, in D.S. Hochbaum (ed.), Approximation Algorithms for NP-hard Problems, PWS Publishing, Boston, MA, 1997, pp. 46–93.

[3] T.G. Crainic, G. Perboli, M. Pezzuto, R. Tadei, Computing the asymptotic worst-case of bin packing lower bounds, European Journal of Operational Research, 183 (2007) 1295–1303.

[4] W. Fernandez de la Vega, G.S. Lueker, Bin packing can be solved within $1 + \epsilon$ in linear time, Combinatorica, 1 (1981) 349–355.

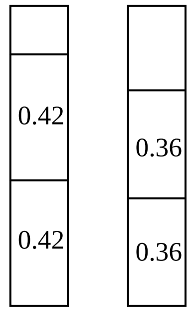
[5] D.S. Hochbaum, Various notions of approximations: Good, better, best, and more, in D.S. Hochbaum (ed.), Approximation Algorithms for NP-hard Problems, PWS Publishing, Boston, MA, 1997, pp. 346–398.

[6] J. Kang, S. Park, Algorithms for the variable sized bin packing problem, European Journal of Operational Research, 147 (2003) 365–372.

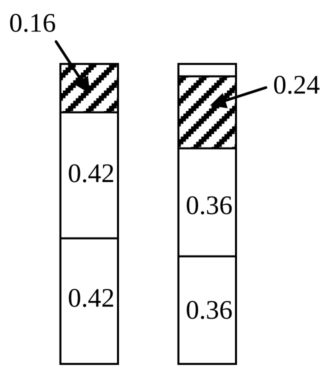
[7] N. Karmarkar, R.M. Karp, An efficient approximation scheme for the one-dimensional bin-packing problem, Proceedings of the 23rd Annual Symposium on Foundations of Computer Science, 1982, pp. 312–320.

[8] C.-L. Li, Z.-L. Chen, Bin-packing problem with concave costs of bin utilization, Naval Research Logistics, 53 (2006) 298–308.

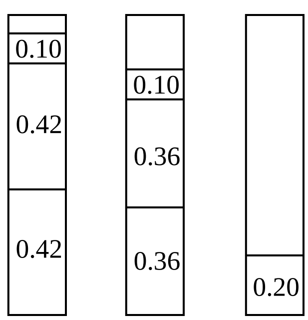
[9] J. Yang, J.Y.-T. Leung, The ordered open-end bin packing problem, Operations Research, 51 (2003) 759–770.



(a) Packing P of items in list L_2



(b) Packing P with the assignment of the big item (i.e., packing $\hat{P}$)



(c) Packing P after the re-packing of the big item (i.e., packing $\overline{P}$)

Figure 1. A numerical example