JOURNAL OF ALGORITHMS 12, 189–203 (1991)

Probabilistic Analysis of Algorithms for Dual Bin Packing Problems

J. CSIRIK

Jószef Attila University, Szeged, Hungary

J. B. G. FRENK

Erasmus University Rotterdam, The Netherlands

G. GALAMBOS

Jószef Attila University, Szeged, Hungary

AND

A. H. G. RINNOOY KAN

Erasmus University Rotterdam, The Netherlands

Received April 30, 1987; revised May 25, 1989

In the dual bin packing problem, the objective is to assign items of given size to the largest possible number of bins, subject to the constraint that the total size of the items assigned to any bin is at least equal to 1. We carry out a probabilistic analysis of this problem under the assumption that the items are drawn independently from the uniform distribution on [0, 1] and reveal the connections between this problem and the classical bin packing problem as well as to renewal theory.
© 1991 Academic Press, Inc.

1. INTRODUCTION

Given *n items* of *size* $a_1, \ldots, a_n$ ($a_i \in (0,1)$, $i = 1, \ldots, n$), the classical *bin packing problem* is to assign the items to the *smallest possible* number of *bins*, subject to the constraint that the total size of the items assigned to any bin is *at most* equal to 1. The *dual bin packing problem*, which is the subject of this paper, is to assign the items to the *largest possible* number of bins, subject to the constraint that the total size of the items assigned to any bin is *at least* equal to 1. The problem could also appropriately be called the *bin covering problem* and was first studied by Assmann *et al.* [2]

0196-6774/91 $3.00

(For a recent survey, see Csirik et al. [13]). Their main concern was worst-case analysis of approximation algorithms. While superficially similar to its traditional counterpart, the problem poses a challenge of its own: as in the case of more general packing and covering problems, a result for one problem occasionally carries over immediately to the other, but generally the differences between them are as pronounced as their common traits.

We shall see examples of both phenomena as we carry out an exploration of the dual bin packing problem. We shall do so from a *probabilistic* point of view, i.e., we shall assume that the item sizes $a_1, a_2, \ldots$ are drawn independently from the uniform distribution on [0, 1]. Many results will in fact be seen to hold under more general assumptions, but the uniform distribution provides a traditional starting point for this type of enquiry.

In addition to the bin packing problem, we shall also consider a two-dimensional analogue, the *dual vector packing problem*. Here, we are given n pairs $(a_1, b_1), \ldots, (a_n, b_n)$ that have to be assigned to the largest possible number of bins subject to the constraint that both the sum of the a-coordinates and the sum of the b-coordinates of the pairs in any bin are at least equal to 1. Many of our results can in fact be extended to the obvious m-dimensional version of this problem, for any fixed m.

In Section 2, we consider the optimal solution value OPT(n) to the dual bin packing problem (i.e., the largest possible number of bins that can be covered by n items) and prove that

$$\limsup_{n \to \infty} \frac{\mathscr{E}(\mathrm{OPT}(n)) - n/2}{n^{1/2}} \le -(32\pi)^{-1/2}; \tag{1}$$

i.e., for n large enough, $\mathscr{E}(\mathrm{OPT}(n)) = n/2 - \Omega(n^{1/2})$. In Section 3, we demonstrate that this estimate is the best possible one up to a multiplicative constant by demonstrating that a simple heuristic, the *pairing heuristic*, produces a value $PA(n)$ satisfying

$$\mathscr{E}(PA(n)) \ge \frac{n}{2} - \left(\frac{n}{2\pi}\right)^{1/2} - \alpha \tag{2}$$

for some constant α. This heuristic can be adapted to show that the expected solution value of the dual vector packing problem is also asymptotic to $n/2$.

These two results have their counterparts in the classical bin packing problem, where an upper bound of $n/2 + O(n^{1/2})$ on the optimal solution value can be proved to be best possible in a similar fashion [10, 11]. (Actually, our technique yields an improvement on the best known upper bound on the multiplicative constant for this case.) The result in Section 4, where we analyze the expected performance of a suitably adapted version of the *next fit heuristic*, has a different flavor. Using techniques from renewal theory that do not carry over to the classical case, we establish the strong result that the solution value $NF(n)$ satisfies

$$\lim_{n\to\infty}\left(\mathscr{E}(NF(n)) - \frac{n}{e}\right) = \frac{2}{e} - 1 = -0.2642\ldots. \qquad (3)$$

This result improves $\lim_{n\uparrow\infty}(\mathscr{E}(NF(n))/n) = 1/e$, which was derived in Section 8 of [1] by modeling the algorithm as a Markov process with discrete time steps (the arrival of each item) and continuous state space (the level of the current bin).

Note, by Assmann's result, that the expected relative error given by $(\mathscr{E}(\mathrm{OPT}(n)) - \mathscr{E}(NF(n)))/\mathscr{E}(\mathrm{OPT}(n))$ converges to $1 - 2/e$. A similar strong result is obtained for an appropriately modified version of *next fit*, applied to the dual vector packing problem. Both results can be easily extended to distributions other than the uniform one. In Section 5, we present a probabilistic analysis of the *next fit decreasing* heuristic, which can again be easily adapted to our model. Surprisingly, its performance is inferior to that of *next fit*, in remarkable contrast to their behavior on the classical bin packing model. Some concluding remarks are contained in Section 6.

2. The Expected Optimal Solution Value

In deriving an upper bound on the optimal solution value to the dual bin packing problem $\mathrm{OPT}(n)$, we shall find it convenient to assume that n is even or, equivalently, to focus on $\mathrm{OPT}(2n)$.

To obtain an upper bound on the expected value of this random variable, we start by defining b_{2n} to be the number of *big* items (i.e., those with size greater than or equal to $\frac{1}{2}$). Since, with probability 1, each bin must contain at least two items in any feasible solution, we almost always have that $\mathrm{OPT}(2n) \le n$. If, however, we know that $b_{2n} < n$, then the best that we can hope for is to pair each big item with a small item to cover a bin, and to divide the remaining small items in groups of three of which each covers an additional bin. Hence, in this case $\mathrm{OPT}(2n) \le b_{2n} + (2n - 2b_{2n})/3 = 2n/3 + b_{2n}/3$.

Since obviously $\mathrm{OPT}(2n) \le \sum_{i=1}^{2n} a_i$, we have that

$$\mathscr{E}(\mathrm{OPT}(2n)) \le \sum_{k=n}^{2n} \mathscr{E}\left(\min\left(\sum_{i=1}^{2n} a_i, n\right) \Big| b_{2n} = k\right)\binom{2n}{k}2^{-2n}$$
$$+ \sum_{k=0}^{n-1} \mathscr{E}\left(\min\left(\sum_{i=1}^{2n} a_i, \frac{2n}{3} + \frac{k}{3}\right) \Big| b_{2n} = k\right)\binom{2n}{k}2^{-2n}. \qquad (4)$$

The first term in (4) is clearly bounded from above by

$$n \sum_{k=n}^{2n} \binom{2n}{k} 2^{-2n}, \tag{5}$$

and since (cf. [12, p. 34])

$$2 \sum_{k=n}^{2n} \binom{2n}{k} = \sum_{k=0}^{2n} \binom{2n}{k} + \binom{2n}{n} = 2^{2n} + \binom{2n}{n},$$

this is equal to

$$n 2^{-2n-1}\left(2^{2n} + \binom{2n}{n}\right) = n/2 + n 2^{-2n-1}\binom{2n}{n}. \tag{6}$$

If we define

$$d_i = \begin{cases} a_i & \left(0 \le a_i < \frac{1}{2}\right) \\ 1 - a_i & \left(\frac{1}{2} \le a_i \le 1\right) \end{cases}, \tag{7}$$

we may observe by the exchangeability of $(a_1, \ldots, a_n)$ and the independence of b_{2n} and $\{d_i\}_{i=1}^{2n}$ (cf. [8]) that for every k,

$$\begin{aligned} &\mathscr{E}\left(\min\left(\sum_{i=1}^{2n} a_i, \frac{2n}{3} + \frac{k}{3}\right) \middle| b_{2n} = k\right) \\ &\quad = \mathscr{E}\left(\min\left(\sum_{i=1}^{k} (1 - d_i) + \sum_{i=k+1}^{2n} d_i, \frac{2n}{3} + \frac{k}{3}\right)\right). \end{aligned} \tag{8}$$

Hence, the second term in (4) equals

$$\sum_{k=0}^{n-1} k \binom{2n}{k} 2^{-2n} + \sum_{k=0}^{n-1} \mathscr{E}\left(\min\left(\sum_{i=k+1}^{2n} d_i - \sum_{i=1}^{k} d_i, \frac{2}{3}(n - k)\right)\right)\binom{2n}{k} 2^{-2n}. \tag{9}$$

The first term of (9) is equal to $n/2 - 2n 2^{-2n-1}\binom{2n}{n}$ (cf. [12, p. 34]). We bound the minimum in the second term of (9) by $(2n - 2k)\mathscr{E} d_i =$

$(n-k)/2$ to obtain (cf. [12, p. 34])

$$\begin{aligned}\frac{1}{2}\sum_{k=0}^{n-1}(n-k)\binom{2n}{k}2^{-2n} &= \tfrac{1}{2}n\sum_{k=0}^{n}\binom{2n}{k}2^{-2n} - \frac{1}{2}\sum_{k=0}^{n}k\binom{2n}{k}2^{-2n} \\ &= \tfrac{1}{2}n\left(\tfrac{1}{2} + 2^{-2n-1}\binom{2n}{n}\right) - \frac{n}{4} \\ &= \tfrac{1}{2}n2^{-2n-1}\binom{2n}{n}. \qquad (10)\end{aligned}$$

Summing up the various components in (6), (9), and (10), we conclude from (4) that

$$\mathscr{E}(\mathrm{OPT}(2n)) \le n - \tfrac{1}{2}n2^{-2n-1}\binom{2n}{n}. \qquad (11)$$

Since, for large n, $\binom{2n}{n}$ is asymptotic to $(\pi n)^{-1/2}2^{2n}$, we obtain the desired result (1):

$$\limsup_{n\to\infty}\frac{\mathscr{E}(\mathrm{OPT}(2n)) - n}{(2n)^{1/2}} \le -(32\pi)^{-1/2}. \qquad (12)$$

In the above proof it is only necessary to require that the distribution $F(x)$ of the item sizes a_i satisfies $F(x) = 1 - F(1-x)$ for every $0 \le x \le \frac{1}{2}$ and F is not degenerate at $\frac{1}{2}$. This means that the inequality (12) with $(32\pi)^{1/2}$ replaced by some other constant depending on $\mathscr{E}d_i$ also holds for this larger class of distributions.

For the optimal solution value to the classical bin packing problem, the above technique yields an asymptotic lower bound equal to $n/2 + (32\pi)^{-1/2}n^{1/2}$ which is a slight improvement over the result in [11].

3. THE PAIRING HEURISTIC

In this section, we demonstrate that the upper bound (12) is sharp by showing that a certain heuristic for the dual bin packing problem produces a solution value that is equal to $n/2 - O(n^{1/2})$ in expectation.

For this purpose, we adapt the binary pairing heuristic for the classical bin packing problem [11, 10] to obtain a pairing heuristic (PA) for dual bin packing. In this heuristic, the largest unassigned item is selected and combined with the smallest unassigned item such that together they can cover a bin (i.e., such that the sum of their sizes exceeds 1). If no such item exists, all items then remaining are added to the bin most recently opened, and the algorithm terminates.

We analyze this heuristic along the lines of [9], using the random variables d_i defined in (7). If we label d_i by "$+1$" and call it *big* if $a_i \geq \frac{1}{2}$, and label it by "-1" and call it *small* if $a_i < \frac{1}{2}$, and consider the labeled sequence d_i in $[0, \frac{1}{2}]$ in increasing order, then the PA heuristic amounts to matching each successive "$+1$" to the unassigned "-1" that is closest to its right. If there are no unassigned -1's to its right, match it the rightmost $+1$. If such a $+1$ also does not exist, then put it in the bin most recently opened. If u_n is the number of unmatched small d_i, then one can verify that

$$PA(n) \geq \frac{n - u_n}{2} - \frac{1}{2}$$

and hence,

$$\mathscr{E}(PA(n)) \geq \frac{n}{2} - \frac{1}{2}\mathscr{E}u_n - \frac{1}{2}.$$

To compute $\mathscr{E}u_n$, we first observe that the sequence of $+1$'s and -1's can be viewed as a realization of a *Bernoulli process* [4], defined by a sequence e_j $(j = 1, 2, \ldots)$ consisting of i.i.d. random variables with $\Pr\{e_j = +1\} = \Pr\{e_j = -1\} = \frac{1}{2}$. (This is a nontrivial statement; we leave the proof to the reader [8].) We have that $u_n = \max_{0 \leq k \leq n}\{-s_k\}$, with $s_k = \sum_{j=1}^{k} e_j$, $k \geq 1$ and $s_0 = 0$. Actually, $s_k \stackrel{d}{=} -s_k$ (i.e., s_k has the same distribution as $-s_k$), so that it suffices to compute the expectation of $\max_{0 \leq k \leq n}\{s_k\}$.

According to the theory of *fluctuations* (cf. [3, p. 287]), we know that (assuming n is even)

$$\begin{aligned} \mathscr{E}u_n &= \sum_{k=1}^{n} \frac{1}{k}\mathscr{E}s_k^+ \\ &= \sum_{k=1}^{n/2} \frac{1}{2k}\mathscr{E}s_{2k}^+ + \sum_{k=1}^{n/2} \frac{1}{2k-1}\mathscr{E}s_{2k-1}^+, \end{aligned} \tag{13}$$

where generally $x^+ = \max(x, 0)$.

Now, using the identity $\sum_{p=1}^{k} p\binom{2k}{k-p} = \frac{1}{2}k\binom{2k}{k}$ (cf. [12, p. 34]), we find that

$$\begin{aligned}
\mathscr{E}s_{2k}^{+} &= 2\sum_{p=1}^{k} \Pr\{s_{2k} = 2p\}p \\
&= 2\sum_{p=1}^{k} p\binom{2k}{k-p}2^{-2k} \\
&= 2^{-2k+1}\sum_{p=1}^{k}\binom{2k}{k-p}p \\
&= 2^{-2k}k\binom{2k}{k}.
\end{aligned} \tag{14}$$

Similarly, using the identity

$$\sum_{p=1}^{k}(p-1)\binom{2k-1}{k-p} = 2^{2k-3} - 2^{2k-2} + \tfrac{1}{2}(2k-1)\binom{2k-2}{k-1}$$

(cf. [12, p. 34]), we obtain

$$\begin{aligned}
\mathscr{E}s_{2k-1}^{+} &= 2^{-2k+1}\sum_{p=1}^{k}(2p-1)\binom{2k-1}{k-p} \\
&= 2^{-2k+2}\sum_{p=1}^{k}(p-1)\binom{2k-1}{k-p} + 2^{-2k+1}\sum_{p=1}^{k}\binom{2k-1}{k-p} \\
&= 2^{-2k+2}\left(2^{2k-3} - 2^{2k-2} + \tfrac{1}{2}(2k-1)\binom{2k-2}{k-1}\right) + \tfrac{1}{2} \\
&= 2^{-2k+1}(2k-1)\binom{2k-2}{k-1}.
\end{aligned} \tag{15}$$

Now, $2^{-2k}\binom{2k}{k} = (-1)^k\binom{-1/2}{k}$ (cf. [5, p. 63]), with $\binom{-1/2}{k}$ defined as $(-\frac{1}{2})(-\frac{1}{2}-1)\ldots(-\frac{1}{2}-k+1)/k!$, so that (cf. [5, p. 64])

$$\begin{aligned}
\sum_{k=1}^{n/2}\frac{1}{2k}\mathscr{E}s_{2k}^{+} &= \frac{1}{2}\sum_{k=0}^{n/2}(-1)^k\binom{-1/2}{k} - \frac{1}{2} \\
&= \frac{1}{2}(-1)^{n/2}\binom{-3/2}{n/2} - \frac{1}{2}.
\end{aligned} \tag{16}$$

A similar manipulation with respect to $\sum_{k=1}^{n/2}(\mathscr{E}s_{2k-1}^{+}/(2k-1))$ yields as a final exact result:

$$\mathscr{E}u_n = \frac{1}{2}(-1)^{n/2}\binom{-3/2}{n/2} + \frac{1}{2}(-1)^{n/2-1}\binom{-3/2}{n/2-1} - \frac{1}{2}.$$

Using the identities $\binom{-3/2}{n/2} = -(n+2)\binom{-1/2}{n/2+1}$ and $\binom{-3/2}{n/2-1} = (-n/(n+1))\binom{-3/2}{n/2}$ the above expression breaks down to

$$\begin{aligned}\mathscr{E}u_n &= \frac{(n+2)(2n+1)}{2n+2}(-1)^{n/2+1}\binom{-1/2}{n/2+1} - \frac{1}{2}\\ &= \frac{(n+2)(2n+1)}{2n+2}\binom{n+2}{n/2+1}2^{-(n+2)} - \frac{1}{2}. \qquad (17)\end{aligned}$$

A refinement of Stirling's formula (cf. [5, p. 54]) then produces as an approximation that

$$|\mathscr{E}u_n - (2n/\pi)^{1/2}| \leq \alpha \qquad (18)$$

for some constant α. Hence,

$$\begin{aligned}\mathscr{E}(PA(n)) &\geq n/2 - \mathscr{E}(u_n/2) - \tfrac{1}{2}\\ &\geq \frac{n}{2} - \left(\frac{n}{2\pi}\right)^{1/2} - \alpha, \qquad (19)\end{aligned}$$

as was to be proved (cf. (2)).

We observe that the analysis leading to (19) is valid under much more general conditions than imposed here. Rather than *independence*, all that turns out to be needed is *exchangeability* and a *symmetry* condition on the joint distribution of the item sizes. We do not pursue this generalization in detail here.

For the classical bin packing problem, the expected solution value of the binary pairing heuristic is given by $n/2 + \mathscr{E}(u_n/2)$. Thus, (17) provides an *exact* expression for this value, improving on the asymptotic estimates that have appeared in the literature.

A variation on the binary pairing heuristic can be used to analyze the optimal solution OPTV(n) of the dual vector packing problem, under the assumption that a_i and b_i are independently uniformly distributed on [0, 1].

To describe this heuristic, divide $[0,1] \times [0,1]$ into the regions A_k and B_k ($k = 0, 1, \ldots, m-1$), where

$$A_k = \left[0, \frac{1}{2}\right] \times \left[\frac{k}{m}, \frac{k+1}{m}\right] \qquad \text{and} \qquad B_k = \left[\frac{1}{2}, 1\right] \times \left[\frac{k}{m}, \frac{k+1}{m}\right]$$

(m arbitrary, but fixed) and label the stochastic pairs (a_i, b_i) ($i = 1, \ldots, n$)

$$\begin{aligned} (i) &= k \qquad \text{if } (a_i, b_i) \in A_k \cup B_{m-k} \qquad (k = 1, \ldots, m-1) \\ (i) &= 0 \qquad \text{if } (a_i, b_i) \in A_0 \cup B_0. \end{aligned} \tag{20}$$

Now it is easy to check that, conditional on $(i) = k$, a_i is still uniformly distributed for $k = 1, \ldots, m - 1$. Consider now all the pairs (a_i, b_i) with (i) equal to k ($k = 1, \ldots, m - 1$), and apply the pairing heuristic for the one-dimensional dual binpacking problem to their first coordinates.

The number of filled bins then equals at least $1/2(w_k - u_{k,n})$, where w_k is the number of items (a_i, b_i) with $(i) = k$ and $u_{k,n}$ is the number of unmatched small items among the elements with label $(i) = k$.

Hence the total number of filled bins PAV(n) satisfies

$$\begin{aligned} \mathscr{E}(\mathrm{PAV}(n)) &\geq \mathscr{E}\left(\sum_{k=1}^{m-1} (w_k - u_{k,n})/2\right) \\ &= \frac{n}{2}\frac{m-1}{m} - \sum_{k=1}^{m-1} \mathscr{E}(u_{k,n}/2), \end{aligned} \tag{21}$$

where we use the fact that w_k is binomially distributed with parameters n and $1/m$ so that $\mathscr{E}(w_k) = n/m$.

We know that (cf. (18)) $\mathscr{E}(u_{k,n} | w_k = p) \leq C\sqrt{p}$ for some constant C, and hence by Jensen's inequality

$$\mathscr{E}(u_{k,n}) \leq C\mathscr{E}\left(\sqrt{w_k}\right) \leq C\sqrt{\mathscr{E}w_k} = C\sqrt{n/m}\,. \tag{22}$$

Thus, $\mathscr{E}(\mathrm{PAV}(n)) \geq n(m-1)/2m - \frac{1}{2}C(m-1)\sqrt{n/m}$, or

$$\liminf_{n\to\infty} \frac{\mathscr{E}(\mathrm{PAV}(n))}{n} \geq \frac{m-1}{2m}, \tag{23}$$

so that

$$\liminf_{n\to\infty} \frac{\mathscr{E}(\mathrm{OPTV}(n))}{n} \geq \frac{1}{2}. \tag{24}$$

Since it is obvious that $\limsup_{n\to\infty}(\mathscr{E}(\mathrm{OPTV}(n))/n) \leq 1/2$, we obtain

$$\lim_{n\to\infty} \frac{\mathscr{E}(\mathrm{OPTV}(n))}{n} = \frac{1}{2}. \tag{25}$$

4. The Next Fit Heuristic

A simple and natural solution method for the dual bin packing problem is given by an adaption of the well-known next fit heuristic for classical bin packing. We continue to assume that item sizes are uniformly distributed on $[0, 1]$.

In a next fit heuristic (NF) for dual bin packing, one opens a bin and assigns items in arbitrary order until the sum of their sizes exceeds 1 and the bin is covered. The process then repeats itself. This algorithm is also discussed in [2].

The number of items v_1 assigned to the first bin is equal to $\inf\{k \geq 1 \mid \sum_{i=1}^{k} a_i \geq 1\}$. The NF heuristic is such that the same distribution applies to the number of items v_j assigned to the jth bin, for any j. Since the sequence $a_i, i \geq 1$, consists of independent and uniformly distributed random variables we obtain that the random variables $v_j, j = 1, 2, \ldots$ are also independent and identically distributed.

Thus, the random solution value $\mathrm{NF}(n)$ is related to the *renewal process* R_n, associated with the sequence v_j and defined by $R_n = \sup\{m \geq 0 \mid \sum_{j=0}^{m} v_j \leq n, v_0 = 0\}$, in that $\mathrm{NF}(n) = R_n$. To compute $\mathscr{E}(\mathrm{NF}(n))$, it suffices to compute the *discrete renewal function* $\mathscr{E}R_n$.

We first observe that $\mathscr{E}v_j = \sum_{k=0}^{\infty} 1/k! = e$, that $\mathscr{E}v_j^2 = 2\sum_{k=1}^{\infty} k \Pr\{v_j \geq k\} - e = 3e < \infty$, and that the distribution of v_j satisfies the property that g.c.d. $\{n \mid n > 0, \Pr\{v_j = n\} > 0\} = 1$. Hence, the *weak renewal theorem* (cf. [5, p. 330] immediately yields that

$$\lim_{n \to \infty} \frac{\mathscr{E}(\mathrm{NF}(n))}{n} = \frac{1}{e}. \tag{26}$$

We obtain a much stronger result by considering $\lim_{n \to \infty}(\mathscr{E}(\mathrm{NF}(n)) - n/e)$. The *strong renewal theorem* yields (cf. [5, p. 341])

$$\lim_{n \to \infty} \left(\mathscr{E}(\mathrm{NF}(n)) - \frac{n}{e} \right) = \frac{2}{e} - 1. \tag{27}$$

In fact, convergence in (27) can be shown to be exponentially fast (cf. [7, p. 72]. In view of the result from Section 3, (27) implies that the expected relative error of the NF heuristic converges to $1 - {}^{2}/e$.

The weaker result (26) can be generalized to the case in which the item sizes are distributed uniformly over the interval $[0, u]$ ($u \in (0, 1)$). In that case, the right-hand side of (26) has to be replaced by $1/\mu$, with

$$\mu = \sum_{l=0}^{\bar{k}} (-1)^l \frac{1}{l!} \left(\frac{1}{u} - l \right)^l \exp\left(\frac{1}{u} - l \right), \tag{28}$$

$$\bar{k} = \lfloor 1/u \rfloor. \tag{29}$$

The derivation of this result is based on the result (cf. [6, p. 27]) that in this case

$$\Pr\left\{\sum_{i=1}^{n} a_i \le x\right\} = \frac{1}{u^n n!}\sum_{l=0}^{n}(-1)^l\binom{n}{l}(\max\{x - lu, 0\})^n \tag{30}$$

and will not be presented here in full detail.

Again, the analysis leading to the result in (27) can be generalized to every distribution on [0, 1] for which $\mathscr{E}v_1^2 < \infty$ and g.c.d.$\{n \mid n > 0, \Pr\{v_1 = n\} > 0\} = 1$. In view of the general applicability of renewal theory, this should not come as a surprise.

The obvious extension of the next fit heuristic to the dual vector packing problem can be analyzed similarly. Let us assume again that a_i and b_i are independently uniformly distributed on [0, 1]. We now have two random variables

$$t_a = \inf\left\{k \ge 1 \mid \sum_{i=1}^{k} a_i \ge 1\right\},$$

$$t_b = \inf\left\{k \ge 1 \mid \sum_{i=1}^{k} b_i \ge 1\right\} \tag{31}$$

and the number of items packed in an arbitrary bin equals $\max\{t_a, t_b\}$. Note that by the independence of t_a and t_b we have

$$\begin{aligned}\Pr\{\max\{t_a, t_b\} \le t\} &= \Pr\{t_a \le t\}\Pr\{t_b \le t\}\\ &= P\left\{\sum_{i=1}^{t} a_i \ge 1\right\}\Pr\left\{\sum_{i=1}^{t} b_i \ge 1\right\} = \left(1 - \frac{1}{t!}\right)^2. \end{aligned}\tag{32}$$

Hence

$$\sum_{t=0}^{\infty} P\{\max\{t_a, t_b\} \le t\} z^t = \frac{1}{1-z} - 2e^z + \sum_{t=0}^{\infty}\frac{z^t}{(t!)^2} \tag{33}$$

and this implies, with $\hat{F}(z) = \sum_{t=0}^{\infty}\Pr\{\max\{t_a, t_b\} = t\}z^t$, that

$$\frac{1 - \hat{F}(z)}{1 - z} = 2e^z - \sum_{t=0}^{\infty}\frac{z^t}{(t!)^2}. \tag{34}$$

Now the number of bins used for n items is given by $\mathrm{NFV}(n) = \sup\{m \ge 0 \mid \sum_{j=0}^{m} v_j \le n\}$, where v_j $(j \ge 1)$ are independent and identically distributed random variables ($v_0 = 0$), with $v_1 = \max\{t_a, t_b\}$. Hence, by the weak renewal theorem,

$$\lim_{n\to\infty} \frac{\mathscr{E}(\mathrm{NFV}(n))}{n} = \frac{1}{\mathscr{E}(\max\{t_a, t_b\})}$$
$$= \lim_{z \uparrow 1} \left(1/2e^z - \sum_{t=0}^{\infty} \frac{z^t}{(t!)^2} \right)$$
$$= \left(1/2e - \sum_{t=0}^{\infty} \frac{1}{(t!)^2} \right). \tag{35}$$

Because the distribution function of $\max\{t_a, t_b\}$ is lattice with the span equal to 1, the strong renewal theorem yields that

$$\lim_{n\to\infty} \left(\mathscr{E}(\mathrm{NFV}(n)) - \frac{n}{E_1} \right) = \frac{E_2 + E_1}{2E_1^2} - 1 \tag{36}$$

with $E_1 = \mathscr{E}(\max\{t_a, t_b\}) = 2e - \sum_{t=0}^{\infty} 1/(t!)^2 = 3.1567\ldots$ and $E_2 = \mathscr{E}((\max\{t_a, t_b\})^2)$.

Note that from (34)

$$\sum_{t=0}^{\infty} tP\{\max\{t_a, t_b\} > t\}$$
$$= \lim_{z \uparrow 1} \frac{d}{dz} \left(\frac{1 - \hat{F}(z)}{1 - z} \right)$$
$$= \lim_{z \uparrow 1} \left(2e^z - \sum_{t=1}^{\infty} \frac{z^{t-1}}{t!(t-1)!} \right)$$
$$= 2e - \sum_{t=0}^{\infty} \frac{1}{(t+1)!t!}. \tag{37}$$

It is also easy to prove that

$$\sum_{t=0}^{\infty} tP\{\max\{t_a, t_b\} > t\} = \tfrac{1}{2}E_2 - \tfrac{1}{2}E_1 \tag{38}$$

and, hence,

$$E_2 = 6e - \sum_{t=0}^{\infty} \frac{1}{(t!)^2} - 2\sum_{t=0}^{\infty} \frac{1}{(t+1)!t!} = 10.8488\ldots. \tag{39}$$

Thus, $\lim_{n\to\infty}(\mathscr{E}(\mathrm{NFV}(n)) - (0.3168\ldots)n) = -0.2974\ldots$.

5. The Next Fit Decreasing Heuristic

In this section we adapt and analyze the next fit decreasing heuristic (NFD) to our model. Given a list of n items of size $a_1, a_2, \ldots, a_n$ $(0 \le a_i \le 1)$, the NFD heuristic for the dual bin packing problem first reindexes the elements in decreasing order and then applies the NF heuristic to this new list. To analyze the behavior of the expected solution value $\mathscr{E}(\text{NFD}(n))$, we approximate the performance of the NFD heuristic by that of the *sliced NFD heuristic with parameter r* (SNFD_r), in which first items larger than $1/r$ are packed according to the NFD heuristic, the last opened bin is completed by adding elements of decreasing size smaller than $1/r$ and any remaining items are packed in groups of size $r + 1$ (possibly at the expense of feasibility; but in the limit, this will not hurt).

The number of bins used by this heuristic on n items is denoted by $\text{SNFD}_r(n)$. It is clear that

$$\text{SNFD}_r(n) \ge \text{NFD}(n) \qquad (r > 1) \tag{40}$$

and

$$\lim_{r \to \infty} \text{SNFD}_r(n) = \text{NFD}(n) \qquad \text{(a.s.)}. \tag{41}$$

Let k_i be the number of items whose size falls in the interval $(1/(i + 1), 1/i]$ $(i \ge 1)$ and let $K_i = k_i + k_{i+1} + \ldots$. Then, for any $r > 1$,

$$\begin{aligned} \text{SNFD}_r(n) &\le \left\lceil \frac{k_1}{2} \right\rceil + \left\lceil \frac{k_2}{3} \right\rceil + \cdots + \left\lceil \frac{k_{r-1}}{2} \right\rceil + \left\lceil \frac{K_r}{r+1} \right\rceil \\ &\le \frac{k_1}{2} + \frac{k_2}{3} + \cdots + \frac{k_{r-1}}{r} + \frac{K_r}{r+1} + r, \end{aligned} \tag{42}$$

where the last term is included to allow for the rounding errors.

Since a_i are uniformly distributed and independent, we obtain $\mathscr{E}k_i = n/(i(i + 1))$ and $\mathscr{E}K_i = n/i$. Hence,

$$\mathscr{E}(\text{SNFD}_r(n)) \le n \sum_{i=1}^{r-1} \frac{1}{i(i+1)^2} + \frac{n}{r(r+1)} + r \tag{43}$$

and this implies that, if r is suitably chosen as a function of n, then

$$\limsup_{n \to \infty} \frac{\mathscr{E}(\text{NFD}(n))}{n} \le \sum_{i=1}^{\infty} \frac{1}{i(i+1)^2}. \tag{44}$$

Moreover,

$$\text{NFD}(n) \ge \left(\frac{k_1}{2} - 1 \right) + \left(\frac{k_2}{3} - 1 \right) + \cdots + \left(\frac{k_{r-1}}{r} - 1 \right) \tag{45}$$

and we find that

$$\liminf_{n\to\infty} \frac{\mathscr{E}(\mathrm{NFD}(n))}{n} \geq \sum_{i=1}^{\infty} \frac{1}{i(i+1)^2}. \tag{46}$$

Since

$$\sum_{i=1}^{\infty} \frac{1}{i(i+1)^2} = \sum_{i=1}^{\infty} \left(\frac{1}{i} - \frac{1}{i+1} - \frac{1}{(i+1)^2} \right) = 2 - \sum_{i=1}^{\infty} \frac{1}{i^2} = 2 - \frac{\pi^2}{6},$$

we obtain from (45) and (46) that

$$\lim_{n\to\infty} \frac{\mathscr{E}(\mathrm{NFD}(n))}{n} = 2 - \frac{\pi^2}{6} = 0.3551\ldots. \tag{47}$$

We note that $\lim_{n\to\infty} \mathscr{E}(\mathrm{NF}(n))/n = 1/e = 0.3679\ldots$, so that the expected performance of the NF heuristic is better than the (expected) performance of the NFD heuristic. For the classical binpacking problem, exactly the reverse is true! We have no satisfying intuitive explanation for this phenomenon.

6. Concluding Remarks

The probabilistic analysis of the dual bin packing problem, carried out in the preceding sections, reveals its connections to the classical bin packing problem and, surprisingly, to renewal theory. It also leaves several open questions of interest. Perhaps most prominent among these would be the challenge to find an on-line heuristic for this problem with better expected relative error than the NF heuristic in Section 4.

Acknowledgments

The authors thank Professor Lueker and an anonymous referee for their constructive remarks.

References

1. S. F. Assmann, "Problems in Discrete Applied Mathematics," Ph.D thesis, Mathematics Department, MIT, Cambridge, MA, 1983.
2. S. F. Assmann, D. S. Johnson, D. J. Kleitman, and J. Y.-T., Leung, On a dual version of the one-dimensional binpacking problem, *J. Algorithms* **5** (1984), 502–525.
3. K. L. Chung, "A Course in Probability Theory," Wiley, New York, 1968.

4. E. CINLAR, "Introduction to Stochastic Processes," Prentice-Hall, Englewood Cliffs, NJ, 1975.
5. W. FELLER, "An Introduction to Probability Theory and Its Applications," Vol. 1, Wiley, New York, 1970.
6. W. FELLER, "An Introduction to Probability Theory and Its Applications," Vol. 2, Wiley, New York, 1971.
7. J. B. G. FRENK, "On Banach Algebras, Renewal Measures and Regenerative Processes," CWI Tract No. 38, CWI, Amsterdam, 1987.
8. J. B. G. FRENK, "On a Pairing Heuristic in Binpacking," unpublished Technical Report COSOR 86-13, Department of Mathematics and Computing Science, TU, Eindhoven, 1986.
9. R. M. KARP, Lecture notes, 1984, unpublished.
10. W. KNÖDEL, A bin packing algorithm with complexity $O(n \log n)$ and performance 1 in the stochastic limit, *in* "Mathematical Foundations of Computer Science 1981," (J. Gruska & M. Chytel, Eds.), Springer-Verlag, Berlin, 1981.
11. G. S. LUEKER, "An Average Case Analysis of Bin Packing with Uniformly Distributed Item Sizes," Report 181, University of California, Irvine, 1982.
12. J. RIORDAN, "Combinatorial Identities," Wiley, New York, 1968.
13. J. CSIRIK, J. B. G. FRENK, A dual version of binpacking, *Algorithms Review* **1** (2), 87–95, 1990.