A PROBABILISTIC ANALYSIS OF THE NEXT FIT DECREASING BIN PACKING HEURISTIC

J. CSIRIK and G. GALAMBOS

József Attila University of Szeged, Szeged, Somogyi u.7, Hungary

J.B.G. FRENK

University of Technology, Eindhoven, P.O. Box 513, The Netherlands

A.M. FRIEZE

Queen Mary College, Mile End Road, London E1 4NS, UK

A.H.G. RINNOOY KAN

Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

Received April 1986
Revised September 1986

A probabilistic analysis is presented of the Next Fit Decreasing bin packing heuristic, in which bins are opened to accommodate the items in order of decreasing size.

bin packing * probabilistic analysis * Next Fit Decreasing heuristic

1. Introduction

Given a list of *n items* of size $a_1, \ldots, a_n$ $(0 \leqslant a_i \leqslant 1)$, the famous *bin packing problem* is to find the smallest number of *bins* in which these items can be packed, subject to the constraint that the total size of the items assigned to any bin cannot exceed 1. This problem is well known to be NP-hard, and the analysis of simple approximation methods for its solution represents a permanent challenge (see, e.g., Coffman et al. [2]).

The *Next Fit Decreasing* (NFD) heuristic is a good example of such a method. The items on the list are first reindexed so that

$$a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n. \tag{1}$$

They are then assigned to bins in this order; a new bin is opened whenever there is not enough room left in the one most recently opened to accomodate the current item. The number of bins opened according to this rule can be shown to exceed the minimum number possible by slightly less than 70 percent in the *worst case* (Baker and Coffman [1]).

We are interested, however, in a *probabilistic* analysis of this heuristic, carried out under the assumption that $a_1, \ldots, a_n$ are drawn independently from a uniform distribution on [0, 1]. It is well known that the optimal solution value OPT(n) satisfies

$$\lim_{n \to \infty} \frac{\mathrm{E(OPT}(n))}{n/2} = 1. \tag{2}$$

For the NFD heuristic, it will be shown below that

$$\lim_{n \to \infty} \frac{\mathrm{E(NFD}(n))}{n/2} = 2\left(\frac{\pi^2}{6} - 1\right) = 1.289\ldots, \tag{3}$$

so that the expected deviation from optimality is slightly less than 30 percent. This result and some powerful generalizations turn out to have been derived independently – but in a more complicated fashion – in Hofri and Kamhi [5].

In Section 2 of this note, we provide a short proof of (3). We shown in Section 3 that the probability that NFD(n) differs from its expected value by more than an amount tn decreases exponentially fast. In Section 4, we establish a central limit theorem for this random variable.

2. The expected solution value

To analyse the expected solution value E(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 to contain at most $r-1$ items and any remaining items are packed in groups of size r. Obviously, for any realization of the item sizes, $\mathrm{SNFD}_r(n) \geq \mathrm{NFD}(n)$ and $\lim_{r\to\infty} \mathrm{SNFD}_r(n) = \mathrm{NFD}(n)$

Let k_i be the number of items whose size falls in the interval $(1/(i+1), 1/i]$ $(i=1, 2,\ldots)$, and let $K_i = k_i + k_{i+1} + \ldots$ Then, clearly,

$$\mathrm{SNFD}_r(n) \leq k_1 + \frac{k_2}{2} + \ldots + \frac{k_{r-1}}{r-1} + \frac{K_r}{r} + r, \tag{4}$$

where the last term is induced to allow for rounding errors. Since $E(k_i) = n/(i(i+1))$ and $E(K_r) = n/r$, the expected value of the right hand side of (4) is equal to

$$n\sum_{i=1}^{r-1} \frac{1}{i^2(i+1)} + \frac{n}{r^2} + r$$
$$= n\sum_{i=1}^{r-1}\left(\frac{1}{i^2} - \frac{1}{i} + \frac{1}{i+1}\right) + \frac{n}{r^2} + r, \tag{5}$$

and hence, from choosing r appropriately as a function of n,

$$\limsup_{n\to\infty} \frac{E(\mathrm{NFD}(n))}{n/2} \leq 2\left(\frac{\pi^2}{6} - 1\right). \tag{6}$$

On the other hand, if the items are packed by the NFD rule and bins containing items from more than one interval $(1/(i+1), 1/i]$ as well as bins containing items smaller than $1/r$ are ignored, then we have that

$$\mathrm{NFD}(n) \geq (k_1 - 1) + \left(\frac{k_2}{2} - 1\right) + \ldots + \left(\frac{k_{r-1}}{r-1} - 1\right), \tag{7}$$

so that, for any fixed r,

$$\liminf_{n\to\infty} \frac{E(\mathrm{NFD}(n))}{n/2} \geq 2\sum_{i=1}^{r-1} \frac{1}{i^2} - 2 + \frac{2}{r}. \tag{8}$$

The right hand side of (8) is monotonically increasing in r and converges to $2(\pi^2/6 - 1)$. Hence, in combination with (6), we conclude that

$$\lim_{n\to\infty} \frac{E(\mathrm{NFD}(n))}{n/2} = 2\left(\frac{\pi^2}{6} - 1\right). \tag{9}$$

3. Deviations from the expected value

In this section, we study the deviation probability,

$$\Pr\{\,|\mathrm{NFD}(n) - E(\mathrm{NFD}(n))| \geq nt\}. \tag{10}$$

It is, of course, bounded (from above) by

$$\Pr\{\mathrm{NFD}(n) - E(\mathrm{NFD}(n)) \geq nt\}$$
$$+ \Pr\{\mathrm{NFD}(n) - E(\mathrm{NFD}(n)) \leq -nt\}. \tag{11}$$

The first probability in (11) is bounded by (cf. (4))

$$\Pr\left\{\sum_{i=1}^{r-1} \frac{k_i}{i} + \frac{K_r}{r} + r - E(\mathrm{NFD}(n)) \geq nt\right\}, \tag{12}$$

the second one is bounded by (cf. (7))

$$\Pr\left\{\sum_{i=1}^{r-1} \frac{k_i}{i} - (r-1) - E(\mathrm{NFD}(n)) \leq -nt\right\}. \tag{13}$$

Now, the Laplace–Stieltjes transform $E(\exp(\sum_{i=1}^{r-1}\lambda_i k_i + \lambda_r K_r))$ $(\lambda_i \geq 0,\ i=1,\ldots, r)$ is well known to equal $(\sum_{i=1}^{r-1} \exp(\lambda_i)/(i(i+1)) + \exp(\lambda_r)/r)^n$. Hence, for every $\lambda > 0$, $E(\exp(\lambda(\sum_{i=1}^{r-1}(k_i/i) + K_r/r))) = (\sum_{i=1}^{r-1} \exp(\lambda/i)/(i(i+1)) + \exp(\lambda/r)/r)^n$.

Using this Laplace–Stieltjes transform, one now easily verifies that, for every r, $\sum_{i=1}^{r-1}(k_i/i) + K_r/r$ is distributed as $\sum_{j=1}^{n} y_{jr}$, where the y_{jr} are i.i.d. random variables with $\Pr\{y_{jr} = 1/i\} = 1/i(i+1)$ $(i=1,\ldots, r-1)$ and $\Pr\{y_{jr} = 1/r\} = 1/r$.

Similarly, one verifies that $\sum_{i=1}^{r-1}(k_i/i)$ is distributed as $\sum_{j=1}^{n} z_{jr}$ with $\Pr\{z_{jr} = 1/i\} = 1/i(i+1)$ $(i=1,\ldots, r-1)$, $\Pr\{z_{jr} = 0\} = 1/r$.

Hence, (12) can be rewritten as

$$\Pr\left\{\sum_{j=1}^{n} y_{jr} - E(\mathrm{NFD}(n)) \geq nt - r\right\}, \tag{14}$$

which, in combination with (7), is easily shown to be bounded by

$$\Pr\left\{ \sum_{j=1}^{n} \left(y_{jr} - \mathrm{E}(y_{jr}) \right) \geqslant n\left(t - \frac{1}{r^2} \right) - 2r + 1 \right\}. \tag{15}$$

Similarly, (13) is bounded by

$$\Pr\left\{ \sum_{j=1}^{n} \left(z_{jr} - \mathrm{E}(z_{jr}) \right) \leqslant -n\left(t - \frac{1}{r^2} \right) + 2r - 1 \right\}. \tag{16}$$

Now, since both y_{jr} and z_{jr} are bounded by 1, a famous result from Hoeffding [4] implies that (15) and (16) are bounded by $\exp(-2n(t - 1/r^2 - 2r/n)^2)$. Taking $r = [n^{1/3}]$, we obtain the strong result that, for all t,

$$\Pr\{ |\mathrm{NFD}(n) - \mathrm{E}(\mathrm{NFD}(n))| \geqslant nt \}$$

$$\leqslant 2 \exp\left(-2n\left(t - \frac{3}{n^{2/3}} \right) \right). \tag{17}$$

We refer to Rhee and Talagrand [7] for similar results obtained for other bin packing heuristics by quite different techniques.

4. A central limit theorem

In this final section, we shall prove that for every x,

$$\lim_{n \to \infty} \Pr\left\{ \frac{\mathrm{NFD}(n) - n(\pi^2/6 - 1)}{\sqrt{n}\sigma_\infty} \leqslant x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-y^2/2)\, \mathrm{d}y, \tag{18}$$

where

$$\sigma_\infty^2 = \lim_{r \to \infty} \mathrm{var}(z_{jr})$$

$$= \lim_{r \to \infty} \sum_{i=1}^{r-1} \frac{1}{i^3(i+1)} - \left(\sum_{i=1}^{r-1} \frac{1}{i^2(i+1)} \right)^2$$

$$= \lim_{r \to \infty} \sum_{i=1}^{r-1} \left(\frac{1}{i^3} - \frac{1}{i^2(i+1)} \right) - \left(\sum_{i=1}^{r-1} \frac{1}{i^2(i+1)} \right)^2$$

$$= \zeta(3) - \left(\frac{\pi^2}{6} - 1 \right) - \left(\frac{\pi^2}{6} - 1 \right)^2$$

$$= \zeta(3) + \frac{\pi^2}{6} - \frac{\pi^4}{36} = 0.14118\ldots, \tag{19}$$

where ζ is the Riemann zeta function. One easily verifies that $\lim_{r \to \infty} \mathrm{var}(y_{jr})$ is also equal to (19).

To prove this central limit theorem, we first observe that

$$\Pr\left\{ \frac{\mathrm{NFD}(n) - n(\pi^2/6 - 1)}{\sqrt{n}\sigma_\infty} \leqslant x \right\}$$

$$\leqslant \Pr\left\{ \frac{\sum_{i=1}^{r-1} (k_i/i) - (r-1) - n(\pi^2/6 - 1)}{\sqrt{n}\,\sigma_\infty} \leqslant x \right\}$$

$$= \Pr\left\{ \frac{\sum_{j=1}^{n} \left(z_{jr} - \mathrm{E}(z_{jr}) \right)}{\sqrt{n\ \mathrm{var}(z_{jr})}} \leqslant \frac{x\sigma_\infty}{\sqrt{\mathrm{var}(z_{jr})}} + \frac{r-1}{\sqrt{n\ \mathrm{var}(z_{jr})}} + \frac{\sqrt{n} \sum_{i=1}^{\infty} 1/i^2(i+1)}{\sqrt{\mathrm{var}(z_{jr})}} \right\}. \tag{20}$$

Taking $r = [n^{1/3}]$, we find that the right hand side of the final inequality converges to x as $n \to \infty$. But then we can apply Theorem 7.1.2 in Chung [3] to conclude that (20) convergers to $(1/\sqrt{2\pi}) \int_{-\infty}^{x} \exp(-y^2/2)\, \mathrm{d}y$ as $n \to \infty$.

The random variables y_{jr} provide a lower bound on $\Pr\{(\mathrm{NFD}(n) - n(\pi^2/6 - 1))/\sqrt{n}\sigma_\infty \leqslant x\}$ in an exactly similar fashion. Together, the lower and the upper bound yield the desired result (18).

It turns out that the above result can be used to compute $\lim_{n \to \infty} \mathrm{E}((\mathrm{NFD}(n))^k)$ for any $k > 0$. We also observe that the results in this note can all be extended to a larger class of distribution functions F than the uniform one. E.g., the results in the last two sections are essentially also valid if the item sizes are generated from any distribution whose density function f satisfies $\lim_{x \downarrow 0} f(x) = c > 0$. These details are left to the reader; essentially, one redefines y_{jr} and z_{jr} by letting $\Pr\{y_{jr} = 1/i\} = \Pr\{z_{jr} = 1/i\} = F(1/i) - F(1/i + 1)$ $(i = 1, \ldots, r-1)$, $\Pr\{y_{jr} = 1/r\} = \Pr\{z_{jr} = 0\} = F(1/r)$, and uses the information on f to bound the latter right hand side.

As a final note, we observe that our results are also valid for the *Harmonic* heuristics introduced in Lee and Lee [6]. They can be easily adapted to show that the *Revised Harmonic* (RH) heuristic introduced in the same paper satisfies

$$\lim_{n\to\infty} \frac{E(RH(n))}{n/2} = 1.237\ldots, \tag{21}$$

i.e., slightly better than NFD, but still surprisingly poor for the heuristic that from a worst case point of view is the best on-line heuristic currently known.

References

[1] B.S. Baker and E.G. Coffman, Jr., "A tight asymptotic bound for Next Fit Decreasing bin packing", *SIAM J. Alg. Discr. Math.* **2**, 147–152 (1981).

[2] E.G. Coffman, Jr., G.S. Lueker and A.H.G. Rinnooy Kan, "An introduction to the probabilistic analysis of sequencing and packing heuristics", Technical Report, Econometric Institute, Erasmus University, Rotterdam, 1985.

[3] K.L. Chung, *A Course in Probability Theory*, Academic Press, New York, 1974.

[4] W. Hoeffding, "Probability inequalities for sums of bounded random variables", *Journ. Am. Stat. Assoc.* **58**, 13–30 (1963).

[5] M. Hofri and S. Kamhi, "A stochastic analysis of the NFD bin packing algorithm", *Journal on Algorithms*, forthcoming.

[6] C.C. Lee and D.T. Lee, "A simple on-line packing algorithm", *Journal of the ACM* **32**, 562–572 (1985).

[7] W.T. Rhee and M. Talagrand, "Martingale inequalities and NP-complete problems", *Math. of Op. Res.*, forthcoming.