A probabilistic analysis of the next fit decreasing bin packing heuristic

Citation for published version (APA):

Document status and date:
Published: 01/01/1986

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 17. Nov. 2023
A PROBABILISTIC ANALYSIS OF THE NEXT FIT DECREASING BIN PACKING HEURISTIC

J. CSIRIK and G. GALAMBOS
Joszef 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. RINNOY 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 \( a_1, \ldots, a_n \) (\( 0 < a_i < 1 \)), 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 \( \text{NP} \)-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

\[
a_1 \geq a_2 \geq \ldots \geq a_n.
\]

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 accommodate 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 \( \text{OPT}(n) \) satisfies

\[
\lim_{n \to \infty} \frac{\text{E}(\text{OPT}(n))}{n/2} = 1.
\]

For the NFD heuristic, it will be shown below that

\[
\lim_{n \to \infty} \frac{\text{E}(\text{NFD}(n))}{n/2} = 2\left(\frac{\pi^2}{6} - 1\right) = 1.289\ldots,
\]

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—bat 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 m 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 $E(\text{NFD}(n))$, we approximate the performance of the NFD heuristic by that of the Sliced NFD heuristic with parameter $r$ (SNFD), 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 placed in groups of size $r$. Obviously, for any realization of the item sizes, $\text{SNFD}(n) \geq \text{NFD}(n)$ and $\lim_{n \to \infty} \text{SNFD}(n) = \text{NFD}(n)$

Let $k_i$ be the number of items whose size falls in the interval $(1/(i+1), 1/i]$, and let $K_r = k_1 + \cdots + k_r$. Then, clearly,

$$\text{SNFD}(n) = k_1 + \frac{k_2}{2} + \cdots + \frac{k_r}{r/r - 1} + \frac{1}{r} + r,$$

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

$$n \sum_{i=1}^{r-1} \frac{1}{i^2(i+1)} + \frac{n}{r^2} + r$$

and hence, from choosing $r$ appropriately as a function of $n$, the deviation probability is

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

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

$$\text{NFD}(n) \geq (k_1 - 1) + \left(\frac{k_2}{2} - 1\right)$$

so that, for any fixed $r$,

$$\lim_{n \to \infty} \frac{E(\text{NFD}(n))}{n/2} = 2 \sum_{i=1}^{r-1} \frac{1}{i^2} - 2 + \frac{2}{r}.$$ (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

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

3. Deviations from the expected value

In this section, we study the deviation probability of $|\text{NFD}(n) - E(\text{NFD}(n))| \geq m$.

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

$$\Pr\{\text{NFD}(n) - E(\text{NFD}(n)) \geq m\}$$

and

$$\Pr\{\text{NFD}(n) - E(\text{NFD}(n)) \leq -m\}.$$ (11)

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

$$\Pr\left\{\sum_{i=1}^{r-1} \frac{k_i}{i} + \frac{K_r}{r} + r - E(\text{NFD}(n)) > m\right\},$$ (12)

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

$$\Pr\left\{\sum_{i=1}^{r-1} \frac{k_i}{i} - (r - 1) - E(\text{NFD}(n)) < -m\right\}.$$ (13)

Now, the Laplace–Stieltjes transform

$$k(\exp((\Sigma_{i=1}^{r-1} k_i + \lambda, K_r)),$$

which is well known to equal

$$\exp(\lambda, K_r)/\lambda^{r+1} + \exp(\lambda/r)^r).$$

Hence, for every $\lambda > 0$, the devia- bility

$$E(\exp(\lambda(\Sigma_{i=1}^{r-1} (k_i + \lambda, K_r))) = \exp(\lambda, K_r)/\lambda^{r+1} + \exp(\lambda/r)^r).$$

Using this Laplace–Stieltjes transform, one now easily verifies that, for every $r$, $\Sigma_{i=1}^{r-1} (k_i + \lambda, K_r)$ is distributed as $\Sigma_n$ i.i.d. random variables, where the $y_{ij}$ are i.i.d. random variables, with

$$\Pr\{y_{ij} = 1/i\} = 1/i(i+1)$$

and

$$\Pr\{y_{ij} = 1/r\} = 1/r.$$ (14)

Similarly, one verifies that $\Sigma_{i=1}^{r-1} (k_i + \lambda, K_r)$ is distributed as $\Sigma_n$ with

$$\Pr\{y_{ij} = 1/i\} = 1/i(i+1)$$

and

$$\Pr\{y_{ij} = 0\} = 1/r.$$ (15)

Hence, (12) can be rewritten as

$$\Pr\left\{\sum_{j=1}^{n} y_{ij} - E(\text{NFD}(n)) \geq m - r\right\}.$$ (16)
which, in combination with (7), is easily shown to be bounded by
\[
\Pr\left( \sum_{j=1}^{n} (y_j - \mathbb{E}(y_j)) \geq n\left(t - \frac{1}{r^2}\right) - 2r + 1 \right).
\]
(15)

Similarly, (13) is bounded by
\[
\Pr\left( \sum_{j=1}^{n} (z_j - \mathbb{E}(z_j)) \leq -n\left(t - \frac{1}{r^2}\right) + 2r - 1 \right).
\]
(16)

Now, since both \(y_j\) and \(z_j\) are bounded by 1, a famous result from Hoeffding [4] implies that (15) and (16) are bounded by \(\exp(-2n(t - 1/r^2 - 2/r)\mathbb{E})\). Taking \(r = \left[n^{1/3}\right]\), we obtain the strong result that, for all \(t\),
\[
\Pr\left| \text{NFD}(n) - \mathbb{E}(\text{NFD}(n)) \right| > nt \leq 2 \exp\left(-2n\left(t - \frac{3}{n^{2/3}}\right)\right).
\]
(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 \(x\),
\[
\lim_{n \to \infty} \frac{\Pr\left( \text{NFD}(n) - \mathbb{E}(\text{NFD}(n)) \leq x \right)}{\sqrt{n} \sigma_n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-y^2/2\right) \, dy,
\]
(18)

where
\[
\sigma_n^2 = \lim_{r \to \infty} \text{var}(z_j).
\]
\[
= \lim_{r \to \infty} \sum_{i=1}^{r-1} \left( \frac{1}{i^2(i+1)} - \frac{1}{(i+1)^2} \right)^2
\]
\[
+ \left( \frac{1}{r^2} - \frac{1}{r^2(i+1)} \right)^2
\]
\[
= \zeta(3) - \left(\frac{\pi^2}{6} - 1\right) - \left(\frac{\pi^2}{6} - 1\right)^2
\]
\[
= \frac{\pi^2}{6} - \frac{\pi^4}{36} = 0.14118\ldots,
\]
(19)

where \(\zeta\) is the Riemann zeta function. One easily verifies that \(\lim_{n \to \infty} \text{var}(y_j)\) is also equal to (19).

To prove this central limit theorem, we first observe that
\[
\Pr\left( \text{NFD}(n) - n\left(\pi^2/3 - 1\right) \leq x \right) \leq \Pr\left( \sum_{i=1}^{r-1} (k_i/i) - n\left(\pi^2/3 - 1\right) \leq x \right)
\]
\[
= \Pr\left( \frac{\sum_{i=1}^{r-1} (z_j - \mathbb{E}(z_j))}{\sqrt{n} \text{var}(z_j)} \leq \frac{x\sigma_n}{\sqrt{\text{var}(z_j)}} \right)
\]
\[
= \Pr\left( \frac{\sum_{i=1}^{n} y_{ij}}{\sqrt{n} \text{var}(z_j)} \leq \frac{\sigma_n}{\sqrt{\text{var}(z_j)}} \right)
\]
\[
+ \frac{r - 1}{\sqrt{n} \text{var}(z_j)} \frac{\sqrt{n} \sum_{i=1}^{\infty} 1/i^2(i+1)}{\sqrt{\text{var}(z_j)}}
\]
(20)

Taking \(r = \left[n^{1/3}\right]\), 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) converges to \((1/\sqrt{2\pi}) \int_{-\infty}^{x} \exp(-y^2/2) \, dy\) as \(n \to \infty\).

The random variables \(y_j\) provide a lower bound on \(\Pr((\text{NFD}(n) - n\left(\pi^2/3 - 1\right))/\sqrt{n} \sigma_n \leq x)\) 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 \(\lim_{n \to \infty} \mathbb{E}(\text{NFD}(n))^k\) 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 \(\lim_{x \to 0} f(x) = c > 0\). These details are left to the reader; essentially, one redefine \(y_j\) and \(z_j\) by letting \(\Pr(y_j = 1/i) = \Pr(z_j = 1/i) = F(1/i) - F(1/i + 1)\) \((i = 1, \ldots, r - 1)\), \(\Pr(y_j = 1/r) = \Pr(z_j = 0) = F(1/r)\), 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

\[
\lim_{n \to \infty} \frac{E(RH(n))}{n/2} = 1.237\ldots,
\]

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