Probabilistic analysis for stochastic Rectangle Packing Problem algorithm

Adam Kurpisz[1]

Abstract: *In this paper we consider a modified Rectangle Packing Problem where for each rectangle (module) its width and height come from uniform probability distribution between 0 and 1 - U(0,1). We provide a probabilistic analysis of a simple polynomial-time algorithm for Rectangle Packing Problem. At the end a comparison to the computer simulations is presented.*

Keywords: *production engineering, optimization, Rectangle Packing Problem*

1. Introduction

In this paper we consider a Rectangle Packing Problem (RPP) which is a part of production process in many branches of industry, i.e. VLSI circuit design, stone-cutting etc. Due to the cost optimization, planning a production process is a challenging task. In many branches of industry one can face a problem of minimizing the waste of material used in production process, which can be modeled with Rectangle Packing Problem. The RPP is classified as a discrete optimization problem, which could be solved (not always to optimality) by variety of methods developed in the past four decades [1]. Let us define a Rectangle Packing Problem [2]. We are given a set $B = \{1, \dots, n\}$ of n rectangles. Each rectangle $i \in B$ is characterized by its width x_i, height y_i. The rotation (by 90°) of the rectangles is allowed. The solution of the Rectangle Packing Problem (RPP) is such a packing, i.e., placement coordinates of a bottom-left corner of each block $i \in B$, that any two blocks do not overlap and the area of minimum enclosing rectangle of the packing is minimized. We denote as b_x, b_y the width and height of the enclosing rectangle, respectively. The measure of the solution is $VoFP$ (Value of Filling Percentage).

$$VoFP = \frac{\sum_{i \in B} x_i \cdot y_i}{b_x \cdot b_y}. \tag{1}$$

In this paper we consider stochastic RPP. Let X_i, Y_i for $i \in B$ be the independent random variables from continuous uniform distribution $U(0,1)$ where 0 and 1 are the minimum and the maximum achievable values, respectively. The values of X_i, Y_i describes the width and height of the i'th rectangle. We consider continuous uniform distribution, thus all the values from $[0,1]$ are equiprobable. Now, in a similar way B_x, B_y are some random variables which values describe the width and height of the enclosing rectangle. We consider the expected value of $VoFP$ as a measure of effectiveness of an algorithm.

$$E[VoFP] = \frac{E[\sum_{i \in B} X_i \cdot Y_i]}{E[B_x \cdot B_y]}. \tag{2}$$

[1] Institute of Mathematics and Computer Science, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, adam.kurpisz@pwr.wroc.pl,

2. Algorithm

In this section we analyze the algorithm presented in [3]. First we sort all the rectangles with respect to their height. Then we put one rectangle next to the other. Since the procedure allows the schedule to be multilayer, at particular point we start the next layer, i.e. we put the rectangle on the top of the first rectangle from the previous layer. The number of layers we choose as $\sqrt{n}$. Thus the number of rectangles in each layer is also $\sqrt{n}$. We do not allow any rotation of rectangles. Without loss of generality the number of rectangles in input instance is assumed to satisfy $n = k^2 for\ some\ k \in N$. The formal description of the algorithm can be seen below.

Algorithm
sort rectangles with respect to their y_i value
from 1 **to** $\sqrt{n}$ **do:**
from 1 **to** $\sqrt{n}$ **do:**
take i'th rectangle and put inline close to the previous rectangle
start a new layer
calculate the width b_x and height b_y of enclosing rectangle
return $b_x \cdot b_y$

An example of resulting packing of the algorithm can be seen in Figure 1.

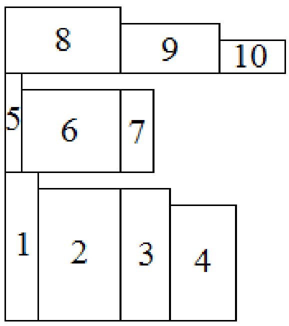

Fig. 1. Visualization of packing obtained with the algorithm.

First we recall the probability density function (PDF) and cumulative density function (CDF) for continuous uniform distribution $U(0,1)$ denoted as $f(x)$ and $F(x)$ respectively:

$$f(x) = 1, \tag{3}$$

$$F(x) = x\ . \tag{4}$$

We calculate the probability density function of the highest rectangle over all n given rectangles. Following the calculations from [3] the PDF of the highest rectangle can be expressed as the n order statistic from the set of n random variables. The general formula of PDF for k ordered statistic can be formed as:

$$f_{k:n}(x) = n\binom{n-1}{k-1} f(x)F(x)^{k-1}(1-F(x))^{n-k}. \tag{5}$$

Now we would like to calculate a total height of the schedule. The height of the first layer is equal to the maximum height over n rectangles, i.e. the n ordered statistic. The height of the second layer is the $n-\sqrt{n}$ ordered statistic, etc. The expected value of the k ordered statistics takes the form:

$$E[X_{k:n}] = \int_0^1 xn\binom{n-1}{k-1} x^{k-1}(1-x)^{n-k}dx = \frac{n-k\sqrt{n}}{n+1}. \tag{6}$$

And thus the expected value of the total height takes the form:

$$E[B_y] = \sum_{k=0}^{\sqrt{n}-1} E[X_{n-k\sqrt{n}:n}] = \sum_{k=0}^{\sqrt{n}-1} \frac{n-k\sqrt{n}}{n+1} = \frac{n(\sqrt{n}+1)}{2(n+1)}. \tag{7}$$

The next step is to approximate the expected value of the width of the schedule. Every layer composes of $\sqrt{n}$ blocks. The total width of the schedule is the maximum width over $\sqrt{n}$ layers. To approximate the density function of the width of any layer we use the Central Limit Theorem:

$$\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - E[X_i]\right) \xrightarrow{d} N(0,\sigma^2), \tag{8}$$

where $\sigma^2 = \frac{1}{12}$ is the variance of the uniform distribution $U(0,1)$. Now let B_x' be the random variable describing the width of any layer. From Central Limit Theorem:

$$E[B_x'] \cong E\left[\sqrt{n}\sigma N(0,1) + nE[X_1]\right]. \tag{9}$$

Now using (6) and the cumulative density function of standard normal distribution we have:

$$E[B_x] = E\left[(B_x')_{\sqrt{n}:\sqrt{n}}\right] \cong \sqrt[4]{n}\sigma \int_{-\infty}^{\infty} x\sqrt{n}\frac{1}{\sqrt{2\pi}} Exp\left[-\frac{x^2}{2}\right]\left(\frac{1}{2}\left(1+Erf\left[\frac{x}{\sqrt{2}}\right]\right)\right)^{\sqrt{n}-1} dx + \frac{\sqrt{n}}{2}, \tag{10}$$

where *Erf()* is a Gauss error function.

The last thing is to calculate the expected value of an area of a particular rectangle. Since for each $i \in B$, X_i and Y_i are independent, then

$$E[X_i \cdot Y_i] = E[X_i] \cdot E[Y_i] = \frac{1}{4}. \tag{11}$$

Now we can calculate the expected value of $VoFP$ for sample value of number of rectangles. Let us set $n = 10000$. The $E[VoFP]$ is equal to:

$$E[VoFP] = \frac{\frac{n}{4}}{E[B_y]E[B_x]} \cong \frac{\frac{n}{4}}{\frac{n(\sqrt{n}+1)}{2(n+1)}\left(\sqrt[4]{n}\sqrt{\frac{1}{12}}2{,}50759 + \frac{\sqrt{n}}{2}\right)} \cong \frac{2500}{2890} \cong 0{,}865. \quad (12)$$

2.1. Computer simulations

In this subsection we present a comparison between theoretic and simulated results.

Table 1. Expected value of VoFP.

	Theoretic $E[VoFP]$	Simulated $E[VoFP]$
n=100	0,710	0,620
n=10000	0,865	0,878
n=40000	0,894	0,897

3. Conclusions

In this paper we analyze a simple algorithm for a stochastic Rectangle Packing Problem. We show that the theoretic results are close to the simulated, and both are close to the optimum. The approximation of the total width of the schedule can be improved using the Berry–Esseen theorem which describes the rate of convergence in Central Limit Theorem.

Acknowledgements

Adam Kurpisz work is co-financed by the European Union as part of the European Social Fund.

Bibliography

[1] C. Blum, A. Roli Metaheuristics in combinatorial optimization: Overview and conceptual comparison. ACM Computing Surveys, 35, (3), 2003, 268-308.
[2] B. S. Baker, E. G. Coifman, R. L. Rivest Orthogonal packings in two dimensions. SIAM J. Comput., 9, (4), 1980, 846–855.
[3] A. Kurpisz Simple Algorithms for Rectangle Packing Problem – from Theory to Experiments, Innovations in Management and Production Engineering, 2013, 207-214.

ANALIZA PROBABILISTYCZNA ALGORYTMU DLA STOCHASTYCZNEGO PROBLEMU PAKOWANIA PROSTOKĄTÓW.

W tej pracy rozważamy Problem Pakowania Bloków, który występuje w wielu gałęziach przemysłu, np. w projektowaniu układów scalonych, cięciu bloków kamiennych, przemyśle tekstylnym. Rozważamy zmodyfikowany problem, w którym wysokości i szerokości bloków są zadane jednostajnym, ciągłym rozkładem prawdopodobieństwa *U(0,1)*. W tej pracy prezentujemy analizę probabilistyczną skuteczności algorytmu przestawionego w pracy [3]. Prezentujemy również wyniki symulacji komputerowych i porównujmy je z wynikami teoretycznymi.