Polska
Akademia
Nauk
Instytut
Badań
Systemowych

Methodology and applications of decision support systems

Proceedings of the 3-rd
Polish-Finnish Symposium
Gdańsk-Sobieszewo, September 26-29, 1988

edited by
Roman Kulikowski

Polska Akademia Nauk Instytut Badań Systemowych

Methodology and applications of decision support systems

Proceedings of the 3-rd Polish-Finnish Symposium
Gdańsk-Sobieszewo, September 26-29, 1988

edited by
Roman Kulikowski

Warszawa 1989

Secretary of the Conference
dr. Andrzej Stachurski

Wykonano z gotowych oryginałów tekstowych
dostarczonych przez autorów

ISBN 83-00-02543-X

A GREEDY-LIKE APPROXIMATE ALGORITHM FOR THE SEQUENCING JOBS WITH DEADLINES PROBLEM : AN AVERAGE CASE APPROACH

Krzysztof Szkatuła
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6, 01-447 Warszawa, Poland

ABSTRACT

In this paper, scheduling jobs with deadlines problem is considerd. A threshold algorithm for solving it is proposed. It is shown, in contrast to the worst case, that the threshold algorithm is asymptotically optimal in the average case.

KEY WORDS:

scheduling jobs with deadlines problem, threshold algorithm, probabilistic analysis

1. INTRODUCTION

In this paper we are concerned with the scheduling jobs with deadlines (SJD) problem. It can be formulated as follows:

n jobs have to be processed on one machine. Each job j has a processing time t_j and a deadline d_j. If job j is completed before its dedline then profit p_j is earned. The problem consists in finding a schedule of jobs which maximizes the total profit i.e. to find a permutation ϕ^* of $\langle 1,\dots,n\rangle$ ($\phi \in \Phi$, Φ set of all permutations of $\langle 1,\dots,n\rangle$) such that

$$\sum_{j=1}^{n} p_{\phi^*(j)} \cdot \operatorname{sgn}^+(d_{\phi^*(j)} - \sum_{i=1}^{j} t_{\phi^*(i)}) =$$

$$= \max_{\phi \in \Phi} \sum_{j=1}^{n} p_{\phi(j)} \cdot \operatorname{sgn}^+(d_{\phi(j)} - \sum_{i=1}^{j} t_{\phi(i)})$$

where

$$\operatorname{sgn}^+(a) = \begin{cases} 1 & \text{if } a \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

This is an important optimization model. Without loss of generality we may assume that

$$0 < p_j, \ t_j \leq 1 \ , \ j=1,\ldots,n \ ,$$

and jobs are sorted in order of nondecreasing deadlines. The deadline of job j is then denoted by $d_j(n)$, $d_j(n) \leq d_{j+1}(n)$, $j=1,\ldots,n-1$.

Now the SJD problem can be formulated as the special binary programming problem [Lawler and Moore (1969)] :

$$z_{OPT}(n) = \max \sum_{i=1}^{n} p_i \cdot x_i$$

$$\sum_{i=1}^{j} t_i \cdot x_i \leq d_j(n) \qquad (1.1)$$

$$x_i = 0 \text{ or } 1 \quad i,j=1,\ldots,n$$

where x_j is equal to 1 if job j is processed on time and 0 if it is tardy.

The SJD problem is known to be an NP-hard combinatorial optimization problem [Garey and Johnson (1979)]. Thus for solving practically large instances of the problem within a limited amount of time (as it often happens in practice) it is necessary to use approximate algorithms.

A simple heuristic algorithm for solving (1.1) is proposed in this paper. It is shown that in the so called worst case no asymptotical accuracy of this algorithm is guaranted. On the other hand in the so called average case this algorithm is asymtptotically optimal (has 0 asymptotical error) for the described class of probabilistic SJD problems.

Various relaxations and estimations of (1.1) are given in Section 2. A threshold algorithm for solving (1.1) is proposed in Section 3. Probabilistic analysis of the threshold solutions and the SJD problem is performed in Section 4. In Section 5 case of uniform distribution is considered.

In this paper the following notation is used:

For the infinite sequences u_n, v_n, $n \to \infty$, we will write:

$$u_n \approx v_n \quad \text{if} \quad \lim_{n\to\infty} \frac{u_n}{v_n} = 1$$

$$u_n = o(v_n) \quad \text{if} \quad \lim_{n\to\infty} \frac{u_n}{v_n} = 0$$

$u_n = O(v_n)$ if there exist constant c such that $u_n \le c\cdot v_n$

$u_n = \Theta(v_n)$ if there exist constants b and c such that

$$b\cdot v_n \le u_n \le c\cdot v_n$$

$\int_a^b f(x)dg(x)$ denotes the Lebesgue-Stiltjes integral on $(a,b]$

$P(\cdot)$ denotes the probability of an event $(\cdot)$

For the random variable X, $E(X)$ denotes its expected value and $Var(X)$ its variance

We will say that the sequence of random variables X_n converges in probability to X [Loeve (1977)] (we will write $X_n \xrightarrow{P} X$) if for every $\varepsilon>0$

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

For two sequences of random variables X_n, Y_n we write

$$X_n \overset{P}{\approx} Y_n \qquad \text{if} \qquad \frac{X_n}{Y_n} \xrightarrow{P} 1$$

2. RELAXATIONS AND ESTIMATIONS

Let us consider the following relaxation of (1.1)

$$z_{KR}(n) = \sum_{i=1}^{n} p_i \cdot x_i$$

$$\sum_{i=1}^{n} t_i \cdot x_i \leq d_n(n)$$

$$0 \leq x_i \leq 1, \; i=1,\ldots,n$$

Introducing

$$p_i(\lambda) = \begin{cases} p_i & \text{if } \frac{p_i}{t_i} > \lambda \\ 0 & \text{otherwise} \end{cases} \qquad t_i(\lambda) = \begin{cases} t_i & \text{if } \frac{p_i}{t_i} > \lambda \\ 0 & \text{otherwise} \end{cases}$$

$$z_n(\lambda) = \sum_{i=1}^{n} p_i(\lambda) \quad , \quad s_n(\lambda) = \sum_{i=1}^{n} t_i(\lambda)$$

we can state the dual problem to it as follows

$$\phi_R(n) = \min_{\lambda\geq 0} \{z_n(\lambda) + \lambda\cdot(d_n(n) - s_n(\lambda))\}$$

For arbitrary $\lambda \geq 0$ we obtain

$$z_{OPT}(n)\leq z_{kr}(n)\leq \phi_R(n)\leq z_n(\lambda)+\lambda(d_n(n)-s_n(\lambda)) \qquad (2.1)$$

Let us consider the arbitrary vector x ,

$$x = \{x_1,\dots,x_n | x_i = 0 \text{ or } 1,\ i=1,\dots,n\}$$

$$\text{with} \quad z_n = \sum_{i=1}^{n} p_i\cdot x_i \quad , \quad s_j = \sum_{i=1}^{j} t_i\cdot x_i \quad , \quad j = 1,\dots,n\ .$$

The solution of (1.1) given by it can be infeasible i.e. there is at least one j, $1 \leq j \leq n$, such that $s_j > d_j(n)$.

If $s_{j-1} \leq d_{j-1}(n)$ and $s_j = s_{j-1} + t_j\cdot x_j > d_j(n)$, $j \in \langle 2,\dots,n \rangle$, then setting $x_j = 0$ we can obtain $s_j = s_{j-1} \leq d_{j-1}(n) \leq d_j(n)$.

Thus, starting with $x'_1 = x_1\cdot \text{sgn}^+(d_1(n) - t_1)$, $s'_1 = t_1\cdot x'_1$ and recursively setting $x'_j = x_j\cdot \text{sgn}^+(d_j(n) - s'_{j-1} - t_j)$, $s'_j = s'_{j-1} + t_j\cdot x'_j$, $j = 2,\dots,n$, feasibile solution of (1.1) $x' = (x'_1,\dots,x'_n)$ can be produced . The value of the goal function of it is equal to $z'_n = \sum_{i=1}^{n} p_i \cdot x'_i$.

Obviously $z'_n \le z_{OPT}(n)$. Let:

$$y_i = \begin{cases} 1 & \text{if } s_i > d_i(n) \\ 0 & \text{otherwise} \end{cases} \qquad i=1,\ldots,n$$

and

$$r_n = \sum_{i=1}^{n} p_i \cdot y_i \cdot x_i$$

Because $s'_j \le s_j$, $j=1,\ldots,n$, for arbitrary $\lambda \ge 0$ and $x=\{x_1,\ldots,x_n | x_i=0 \text{ or } 1,\ i=1,\ldots,n\}$ we have from (2.1)

$$z_n - r_n \le z'_n \le z_{OPT}(n) \le z_n(\lambda) + \lambda\cdot(d_n(n) - s_n(\lambda)) \qquad (2.2)$$

3. THRESHOLD ALGORITHM

When practically large NP-hard problems are considered they are usually solved using approximate algorithms.

Let us consider arbitrary approximate algorithm A solving instance ρ of the given problem P where:

n - size of the instance ρ , $z_A(n)$ - value of the approximate solution produced by A , $z_{OPT}(n)$ value of the optimal solution of ρ, P_n set of all instances of the problem P of the size n.

From the point of view of asymptotical accuracy, approximate algorithms could be classified as follows :

1) Algorithm A is asymptotically optimal for problem P if for every $\varepsilon > 0$ there exist $n_0 \ge 1$ such that

$$\left| \frac{z_{OPT}(n)}{z_A(n)} - 1 \right| \le \varepsilon \quad \text{for every } n \ge n_0 \text{ and } \rho \in P_n$$

2) A has ε ($\varepsilon \ge 0$) asymptotical relative error if for

every δ , γ , n_1 , $\delta > \varepsilon > \gamma$, $n_1 \geq 1$, there exist n_0 , $n_0 \geq 1$ such that

$$\left| \frac{z_{OPT}(n)}{z_A(n)} - 1 \right| \leq \delta \text{ for every } n \geq n_0 \text{ and } \rho \in P_n$$

and there exist $n \geq n_1$ and $\rho \in P_n$ such that

$$\left| \frac{z_{OPT}(n)}{z_A(n)} - 1 \right| \geq \gamma$$

3) A has infinite relative error for P if for every $\varepsilon \geq 0$, $n_0 \geq 1$, there exist $n \geq n_0$, $\rho \in P_n$ such that

$$\left| \frac{z_{OPT}(n)}{z_A(n)} - 1 \right| \geq \varepsilon$$

It is easy to observe that :

(i) Every optimal algorithm ($z_A(n) = z_{OPT}(n)$ for every $n \geq 1$ and $\rho \in P_n$) is also asymptotically optimal.

(ii) An asymtocically optimal algorithm is an algorithm with 0 relative error .

We propose a simple-greedy like algorithm for the SJD problem. It does not even need explicit sorting (i.e. it is a so called on line algorithm) , which is usually the case.

The efficiency of every decision variable i ($1 \leq i \leq n$) of (1.1) is equal to $\frac{p_i}{t_i}$. The larger it is the more promising is the corresponding decision variable.

We will consider the so called threshold value λ. Let:

$$x_i(\lambda) = \begin{cases} 1 & \text{if } \frac{p_i}{t_i} > \lambda \\ 0 & \text{otherwise} \end{cases} \qquad i=1,\dots,n$$

We have produced vector $(x_1(\lambda),\dots,x_n(\lambda) \mid x_i(\lambda)=0 \text{ or } 1, i=1,\dots,n)$

$x_i(\lambda)$ is equal to 1 only when the efficiency of the decision variable i $(1 \le i \le n)$ is greater than the given threshold value λ (i.e. we are considering only the best, according to the threshold value λ, decision variables).

To ensure feasibility of the threshold solution to (1.1) we may use the simple procedure proposed at the end of the previous section .

Combination of these two procedures provides the threshold algorithm :

Threshold algorithm

1° $z'(\lambda):=0$, $s'(\lambda):=0$;

$x_i'(n):=0$, $i=1,\dots,n$;

2° **for** $i = 1$ **to** n **do**

begin

if $s'(\lambda) + t_i > d_i(n)$ **or** $\frac{p_i}{t_i} \le \lambda$ **then go to** 3°

$x_i'(\lambda):=1$

$z'(\lambda):=z'(\lambda) + p_i$

$s'(\lambda):=s'(\lambda) + t_i$

3° **end**

$z_{THR}(n,\lambda):= z'(\lambda)$

STOP

Observing that $p_i(\lambda) = p_i \cdot x_i(\lambda)$, $t_i(\lambda) = t_i \cdot x_i(\lambda)$ and introducing

$$y_i(\lambda) = \begin{cases} 1 & \text{if } s_i(\lambda) > d_i(n) \\ 0 & \text{otherwise} \end{cases} \qquad i=1,\dots,n$$

$$r_n(\lambda) = \sum_{i=1}^{n} p_i(\lambda) \cdot y_i(\lambda)$$

we can obtain from (2.2) for arbitrary $\lambda \ge 0$:

$$1 - \frac{r_n(\lambda)}{z_n(\lambda)} \le \frac{z_{THR}(n,\lambda)}{z_n(\lambda)} \le \frac{z_{OPT}(n)}{z_n(\lambda)} \le \max\left\{1 \,,\, \frac{d_n(n)}{s_n(\lambda)}\right\} \qquad (3.1)$$

For every given threshold value $\lambda \ge 0$ threshold algorithm has extremely low computational complexity - $\Theta(n)$.

Let us consider the following SJD problem :

$$z_{OPT}(n) = \max \left(\sum_{i=1}^{n-1} \alpha\cdot\gamma\cdot x_i + \beta\cdot x_n \right)$$

$$\sum_{i=1}^{j} \alpha\cdot x_i \le \frac{j}{n-1}\cdot\alpha, \qquad j = 1,\dots,n-1$$

$$\sum_{i=1}^{n-1} \alpha\cdot x_i + \beta\cdot x_n \le \beta$$

$$x_i = 0 \text{ or } 1 \qquad i = 1,\dots,n$$

where $\alpha > 0$, $\gamma > 1$, $\alpha\cdot\gamma < \beta \le 1$

Then :

$$z_{OPT}(n) = \beta$$

$$z_{THR}(n,\lambda) = \begin{cases} 0 & \text{if } \lambda > \gamma \\ \alpha\cdot\gamma & \text{otherwise} \end{cases}$$

So $\dfrac{z_{THR}(n,\lambda)}{z_{OPT}(n)} \le \dfrac{\alpha\cdot\gamma}{\beta}$ and taking apropriate values of α ,γ and β these ratios could be arbitrarily close to zero for every $\lambda \ge 0$.

This example shows that in the so called worst case (for all instances of the SJD problem) threshold algorithm has infinite relative error .

4. PROBABILISTIC ANALYSIS OF THE THRESHOLD ALGORITHM AND THE SJD PROBLEM

The goal of this section is to show that in the so called average case (i.e. with probability approaching 1 as n tends to infinity), the threshold algorithm is asymptotically optimal for a rather wide class of random SJD problems.

To perform probabilistic analysis of the problem and of the algorithm, one needs a probabilistic model of the problem. To define the class of probabilistic SJD problems we will assume that p_i, (t_i), $i=1,\dots,n$, are realizations of identically distributed random variables (i.d.r.v.) P_i (T_i). This leads to the fact that all previously introduced quantities (but n, λ, $d_1(n),\dots,d_n(n)$) such as $p_i(\lambda), t_i(\lambda), y_i(\lambda)$, $i=1,\dots,n$, $z_{OPT}(n)$, $s_i(\lambda)$, $z_n(\lambda)$, $r_n(\lambda)$, $z_{THR}(n,\lambda)$, are also realizations of corresponding random variables $P_i(\lambda)$, $T_i(\lambda)$, $Y_i(\lambda)$, $i=1,\dots,n$, $Z_{OPT}(n)$, $S_i(\lambda)$, $Z_n(\lambda)$, $R_n(\lambda)$, $Z_{THR}(n,\lambda)$.

Let $H(x)$ $(G(x))$ be the cumulative distribution function (c.d.f.) of i.d.r.v. P_i (T_i), $i=1,\dots,n$.

Moreover it is assumed that P_i, T_i, $i=1,\dots,n$, are mutually independent and i.d.r.v. concentrated on the interval $[0,1]$ ($[0,1]$ m.i.i.d.r.v.]).

Theorem [Szkatuła (1988)]

If for every $n \geq 1$ and $d_1(n),\dots,d_n(n)$, P_i, T_i, $i=1,\dots,n$, are $[0,1]$ m.i.i.d.r.v. and there exist β, λ_n, ψ_n, $0 < \beta < \frac{1}{2}$, $\lambda_n > 0$, $0 \leq \psi_n \leq n$, such that:

(i) $$E(S_n(\lambda_n)) \approx d_n(n)$$

(ii) $$\lim_{n\to\infty} \frac{\ln(n)}{\lambda_n \cdot d_n(n)} = 0 \quad , \quad \alpha_n = \left[\frac{\ln(n)}{\lambda_n \cdot d_n(n)} \right]^{\beta}$$

(iii) At least ψ_n , $\psi_n \approx n$, items are fulfilling inequality:

$$d_j(n) > (1 + \alpha_n)\cdot E(S_j(\lambda_n)) \qquad 1 \le j \le n$$

then

$$Z_{OPT}(n) \overset{P}{\approx} Z_{THR}(n,\lambda) \overset{P}{\approx} Z_n(\lambda_n) \overset{P}{\approx} E(Z_n(\lambda_n)) \qquad (4.1)$$

■

An even more general result holds.

Corollary [Szkatuła (1988)]

If all assumptions of Theorem but (ii) are fulfilled and instead of (ii) the following hold:

(ii)' $$\lim_{n\to\infty} \frac{\ln(n)}{\gamma(\lambda_n)\cdot d_n(n)} = 0, \qquad \alpha_n = \left[\frac{\ln(n)}{\gamma(\lambda_n)\cdot d_n(n)} \right]^\beta$$

then

$$Z_{OPT}(n) \overset{P}{\approx} Z_{THR}(n,\lambda_n) \overset{P}{\approx} Z_n(\lambda_n) \qquad (4.2)$$

■

5. THE CASE OF THE UNIFORM DISTRIBUTIONS

The theorem and corollary have rather general formulations. It seemed to be of some interest to consider a specific instance and to interpret the assumptions and results of the Theorem and corollary for it.

As an apropriate instance we take uniform distribution of P_i , T_i , $1 \le i \le n$, which is often used in such a context [Burkard and Fincke (1985), Frieze and Clarke (1984, Szkatuła and Libura (1987)].

Let

$$H_i(x) = G_i(x) = x \qquad 0 \le x \le 1 \quad , \quad i=1,\dots,n$$

This means that realizatons of r.v. P_i , T_i, could be equal to every value from (0,1] with the same probability.

Asumption (i) holds if

$$d_n(n) \;\le\; E\left(\sum_{i=1}^{n} T_i\right) \;+\; a(n) \qquad (5.1)$$

where $\quad E\left(\sum_{i=1}^{n} T_i\right) \;=\; \frac{1}{2}\cdot n, \qquad a(n) \;=\; o(n).$

Let $d'_n(n) = d_n(n) - b(n)$ where $b(n) = o(d_n(n))$.

If $d_n(n)$ is such that there exist $d'_n(n) \le \frac{1}{2}\cdot n$ for every $n \ge 1$ then (5.1) hold.

λ_n is choosen as follows :

$$\lambda_n \;=\; \arg(\,E(S(\lambda_n)) = d'_n(n)\,) \qquad (5.2)$$

Such a choice of λ_n fulfils (i)

1) If $\quad 0 < d'_n(n) \le \frac{1}{6}\cdot n \quad$ then

$$\lambda_n \;=\; \sqrt{\frac{n}{6\cdot d'_n(n)}}$$

$$E(Z_n(\lambda_n)) \;=\; \sqrt{\frac{2}{3}\cdot n\cdot d'_n(n)} \;\approx\; \sqrt{\frac{2}{3}\cdot n\cdot d_n(n)}$$

But the case:

$$d_n(n) \;=\; O\left(\frac{\ln^2(n)}{n}\right)$$

assumptions (ii) and (ii)' hold.

2) If $\quad \frac{1}{6}\cdot n \le d'_n(n) \le \frac{1}{2}\cdot n \quad$ then

$$\lambda_n \;=\; \frac{3}{2} \;-\; \frac{3\cdot d'_n(n)}{n}$$

$$E(Z_n(\lambda_n)) \;=\; \frac{1}{8}\cdot n \;+\; \frac{3}{2}\cdot d'_n(n)\cdot\left(1 - \frac{d'_n(n)}{n}\right) \;\approx$$

$$\approx\; \frac{1}{8}\cdot n \;+\; \frac{3}{2}\cdot d_n(n)\cdot\left(1 - \frac{d_n(n)}{n}\right)$$

If $d_n(n) \le \frac{1}{2}\cdot n$ then assumptions (ii) and (ii)' hold. If $d_n(n) > \frac{1}{2}\cdot n$ then only (ii)' hold and (ii) do not hold.

(iii) is equivalent in the case of the uniform distribution to the following condition :

At least ψ_n, $\psi_n \approx n$, items fulfill inequality :

$$d_j(n) > \frac{j}{n}\cdot(d_n(n) - c(n))$$

where

$$c(n) = b(n) + \alpha_n\cdot(b(n) - d_n(n)) = o(d_n(n)) .$$

RRFERENCES

1. R.E. Burkard and U. Fincke, Probabilistic asymptotic properties of some combinatorial optimization problems, *Discrete Appl. Math. 12 (1985), 21-29.*

2. A.M. Frieze and M.R.B. Clarke, Approximation algorithms for the m-dimensional knapsack problem: worst case and probabilistic analyses, *European Journal of Operational Research 15 (1984), 100-109.*

3. M.R. Garey and D.S. Johnson, Computers and Intractability : a guide to the theory of NP-completeness *(Freeman, San Francisco, 1979).*

4. E.L. Lawler and J.M. Moore, A functional equation and its application to resource allocation and sequencing problems, *Management Science 16 (1969) 77-84.*

5. M.Loeve, Probability Theory I, *(Springer Verlag, New York, 1977).*

6. K.Szkatuła and M.Libura, On Probabilistic Properties of Greedy-like Algorithms for the Binary Knapsack Problem,

in: G. *Andreatta, F. Mason and P. Serafini eds., Advanced School on Stochastics in Combinatorial optimization, Udine, Italy, 22-25 September 1986 (World Scientific, Singapore, 1987) 233-254.*

7. K. Szkatuła, Simple approximate algorithm for the sequencing jobs with deadlines problem, *Working paper ZPM-39 -A1526/CPBP 02.15/1.3.6/88.*