Raport Badawczy
Research Report

RB/21/2016

Losowe zadanie harmonogramowania (szeregowania) prac z terminami zakończenia: Wpływ przedziałów terminów zakończenia na zachowanie wartości rozwiązania optymalnego

K. Szkatuła

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Zakładu zgłaszający pracę:
Prof. dr hab. inż. Zbigniew Nahorski

Warszawa 2016

Losowe zadanie harmonogramowania (szeregowania) prac z terminami zakończenia: Wpływ przedziałów terminów zakończenia na zachowanie wartości rozwiązania optymalnego

Krzysztof SZKATUŁA
Instytut Badań Systemowych PAN
ul. Newelska 6, 01-447 Warszawa
Uniwersytet Humanistyczno-Przyrodniczy w Siedlach
ul. Konarskiego 2, 08-110 Siedlce
E-mail: Krzysztof.Szkatula@ibspan.waw.pl

10 grudnia, 2016

Streszczenie

W pracy rozważony został wpływ wzajemnych relacji terminów zakończenia prac (przedziałów terminów zakończenia, ang. deadlines intervals) na asymptotyczny wzrost wartości rozwiązań optymalnych losowego zadania szeregowania prac z terminami zakończenia (ang. Sequencing Jobs with Deadlines problem, SJD). Celem problemu optymalizacyjnego jest maksymalizacja łącznego zysku z prac wykonanych w terminie. Został zaproponowany asymptotycznie sub-optymalny algorytm do rozwiązywania tego zadania. W pracy przyjęto założenie, że współczynniki zadań są realizacjami wzajemnie niezależnych zmiennych losowych o rozkładzie jednostajnym na przedziale $[0, 1)$, $n \to \infty$, gdzie terminy zakończenia prac są wielkościami deterministycznymi.

Słowa kluczowe: Harmonogramowanie (Scheduling), zadanie załadunku (Knapsack Problem), analiza probabilistyczna (Probabilistic Analysis), algorytm przybliżony (Approximate Algorithm), zysk (Profit), kryterium kosztowe (Cost Criterion)

1 Wprowadzenie

Zadanie szeregowania prac z terminami zakończenia (ang. Sequencing Jobs with Deadlines problem, w skrócie SJD) polega na maksymalizacji ważonej ilości prac wykonanych przed terminem zakończenia (ang. deadlines). Terminy zakończenia są szczególnym przypadkiem okna czasowego (ang. due windows, due intervals), patrz Janiak i inni [5]. Każda praca j ($j = 1, ..., n$) musi być wykonana na pojedyńczej, tej samej, maszynie. Wykonanie pracy wymaga czasu t_j i jest zadany termin jej wykonania $d_j(n)$. W przypadku wykonania pracy przed upływem terminu jej wykonania otrzymywany jest zysk p_j. Naszym celem jest maksymalizacja całościowego zysku z prac wykonanych przed terminem, co może byc uznane za ekwiwalentne z minimalizacją pełnego kosztu wykonania prac.

Z formalnego punktu widzenia, zgodnie z klasyfikacją Deterministycznych Zadań Harmonogramowania, Zadanie Szeregowania Prac z Terminami Zakończenia należy do klasy Zadań Harmonogramowania na Pojedynczej Maszynie (Single Machine Scheduling, w skrócie SMS). Dokładniej Zadanie Szeregowania Prac z Terminami Zakończenia jest rozważane jako Zadanie Harmonogramowania z Kryteriami Optymalizacyjnymi z Terminami Zakończenia (Scheduling Problem with Optimisation Criteria involving Due Dates), sklasyfikowane jako $1 \mid\mid \Sigma w_j U_j$, patrz Błażewicz et al. [1]. Klasie Zadań Harmonogramowania na Pojedyńczej Maszynie poświęcono w literaturze dużo uwagi, zarówno z uwagi na ich samodzielne znaczenie badawcze jak również z uwagi na fakt, że często są one rozważane jako część bardziej ogólnych i złożonych zadań.

Przyjmując założenie, że

$$d_1(n) \leq d_2(n) \leq \cdots \leq d_n(n) \tag{1}$$

zadanie Szeregowania Prac z Terminami Zakończenia może zostać sformułowane w postaci problemu programowania binarnego, patrz Lawler i Moore [8]:

$$\begin{aligned} & z_{OPT}(n) = \max \sum_{j=1}^{n} p_j x_j \\ & \sum_{j=1}^{i} t_j x_j \leq d_i(n), \quad i = 1, \ldots, n \\ & \text{gdzie} \quad x_j = 0 \text{ lub } 1, \quad j = 1, \ldots, n \end{aligned} \tag{2}$$

gdzie $x_j = 1$ tylko w przypadku wykonania pracy j w terminie. Prace należy wykonywać zgodnie z kolejnością wyznaczoną przez (1), prace nie wykonane w terminie mogą być wykonane w dowolnej kolejności lub w ogóle nie wykonane. Bez utraty ogólności można przyjąć założenie, że

$$0 < t_j \leq d_j(n) \text{ oraz } p_j > 0,\ j = 1, \ldots, n$$

Zadanie Szeregowania Prac z Terminami Zakończenia jest NP-trudnym problemem Optymalizacji Dyskretnej, patrz Garey i Johnson [4], ale istnieją efektywne algorytmy przybliżone dla jego rozwiązywania. Sahni [11] zaproponował pseudowielomianowy algorytm oparty na idei programowania dynamicznego, w pracach Dudzińskiego i Szkatuły [2] zaproponowano proste i efektywne algorytmy heurystyczne. W literaturze rozpatrywana była uproszczona wersja tego zadania, gdzie zamiast indywidualnych czasów wykonania t_i, $i = 1, \ldots, n$, dla wszystkich prac została przyjęta jednakowa jednostka czasu, np. $t_i = c$, gdzie c jest pewną stałą. Dla uproszczonej wersji zadania szeregowania prac zostało zaproponowanych wiele efektywnych algorytmów typu zachłannego, patrz Puntambekar [10]. Ponadto algorytmy typu zachłannego dla uproszczonej wersji zadania szeregowania prac są często wykorzystywane w procesie dydaktycznym na uczelniach, patrz Kocur [7].

Można łatwo zauważyć, że Zadanie Szeregowania Prac z Terminami Zakończenia (2) stanowi szczególny przypadek wielowymiarowego zadania załadunku, powszechnie znanego problemu optymalizacyjnego, patrz Kellner i inni [6], w poniższym sformułowaniu:

$$z_{OPT}(n) = \max \sum\nolimits_{i=1}^{n} c_i \cdot x_i \text{ s.t. } \sum\nolimits_{i=1}^{n} a_{ji} \cdot x_i \leq b_j(n),\ x_i = 0 \text{ or } 1,\ i,\ j = 1, \ldots, n$$

gdzie $c_j = p_j$, $a_{ij} = t_j$, $1 \le i \le j$, $a_{ij} = 0$, $j < i \le n$, $b_j(n) = d_j(n)$, $j = 1, \ldots, n$. W przypadku kiedy wszystkie ograniczenia, poza ostatnim, zostaną usunięte wtedy zadanie szeregowania prac przyjmuje postać klasycznego (jedno-wymiarowego) zadania załadunku:

$$z_{OPT}(n) = \max \sum_{i=1}^{n} p_i \cdot x_i \text{ s.t. } \sum_{i=1}^{n} t_i \cdot x_i \le d_n(n), \; x_i = 0 \text{ or } 1, \; i = 1, \ldots, n \tag{3}$$

Wielowymiarowe zadanie załadunku jest dobrze znanym silnie NP-trudnym zadaniem Optymalizacji Dyskretnej, podczas gdy zadanie szeregowania prac z terminami zakończenia oraz jedno-wymiarowe zadanie załadunku są NP-trudnymi zadaniami optymalizacji dyskretnej, ale nie są zadaniami silnie NP-trudnymi, patrz Garey i Johnson [4].

Właściwości probabilistyczne losowej wersji wielowymiarowego zadania załadunku były rozpatrywane w pracach Frieze i Clarke [3], Mamer i Schilling [9], Schilling [12],[13] oraz Szkatuła [14],[15]. W związku z bardzo znacznymi różnicami pomiędzy ogólną postacią zadania załadunku oraz zadaniem szeregowania prac z terminami zakończenia powyżej przytoczone wyniki nie mogą zostać zastosowane w bezpośredni sposób do zadania szeregowania prac. W pracy Szkatuła [16] został przeanalizowany asymptotyczny wzrost, przy $n \to \infty$, wartości rozwiązania optymalnego $z_{OPT}(n)$ dla klasy losowych zadań szeregowania prac z terminami zakończenia.

Celem tej pracy jest zbadanie wpływu przedziałów terminów zakończenia na asymptotyczne zachowanie wartości rozwiązania optymalnego $z_{OPT}(n)$, przy $n \to \infty$, w przypadku losowej wersji zadania szeregowania prac z terminami zakończenia, gdzie przedziały terminów zakończenia są określone przez wzajemne relacje $d_j(n)$, $j = 1, \ldots, n$. Ponadto w pracy został zaproponowany prosty algorytm heurystyczny i została udowodniona jego asymptotyczna sub-optymalność w sensie analizy średniego przypadku (analizy probabilistycznej).

Przedstawione w pracy wyniki stanowią wkład do teorii harmonogramowania oraz do analizy probabilistycznej zadań optymalizacji kombinatorycznej. Wyniki te mogą być też przydatne przy budowie i analizie algorytmów przybliżonych dla zadania szeregowania prac z terminami zakończenia.

W pracy stosowane są następujące oznaczenia: $V_n \approx Y_n$, $n \to \infty$ oznacza:

- $Y_n \cdot (1 - o(1)) \le V_n \le Y_n \cdot (1 + o(1)$ jeśli V_n oraz Y_n są ciągami liczb;
- $\lim_{n\to\infty} P\{Y_n \cdot (1 - o(1)) \le V_n \le Y_n \cdot (1 + o(1))\} = 1$ jeśli V_n jest ciągiem zmiennych losowych oraz Y_n jest ciągiem liczb lub zmiennych losowych, gdzie $o(1) > 0$ oraz $\lim_{n\to\infty} o(1) = 0$ jak jest to powszechnie przyjęte.

W rozdziale 2 zostały przedstawione oszacowania (2), uzyskane przy pomocy teorii dualności. Zostały one wykorzystane w rozdziale 3 prezentującym wyniki analizy probabilistycznej zadania szeregowania prac z terminami zakończenia. W rozdziale 4 zostały przedstawione główne wyniki tej pracy dotyczące przedziałów terminów zakończenia oraz został zaproponowany algorytm przybliżony. Rozdział 5 jest poświęcony omówieniu uzyskanych wyników.

2 Oszacowania rozwiązania optymalnego

Rozważmy funkcję Lagrange'a dla (2), patrz Szkatuła [16]:

$$\begin{aligned}
F_n(x,\Lambda) &= \sum_{j=1}^{n} p_j x_j + \sum_{i=1}^{n} \lambda_i \cdot \left(d_i(n) - \sum\nolimits_{j=1}^{i} t_j x_j\right) = \\
&= \sum_{i=1}^{n} \lambda_i d_i(n) + \sum_{j=1}^{n} p_j x_j - \sum_{i=1}^{n} \lambda_i \cdot \left(\sum\nolimits_{j=1}^{i} t_j x_j\right) = \\
&= \sum_{i=1}^{n} \lambda_i d_i(n) + \sum_{j=1}^{n} \left(p_j - \left(\sum\nolimits_{i=j}^{n} \lambda_i\right) \cdot t_j\right) \cdot x_j = \\
&= \sum_{i=1}^{n} \lambda_i d_i(n) + \sum_{j=1}^{n} (p_j - \Lambda_j \cdot t_j) \cdot x_j
\end{aligned}$$

gdzie $x = \{x_1, \ldots, x_n\}$, $\Lambda = \{\lambda_1, \ldots \lambda_n\}$, $\Lambda_j = \sum_{i=j}^{n} \lambda_i$. Niech dla każdego Λ, $\lambda_j \geq 0$, $j = 1, \ldots, n$

$$\begin{aligned}
\varphi_n(\Lambda) &= \max_{x\in\{0,1\}^n} F_n(x,\Lambda) = \sum_{i=1}^{n} \lambda_i d_i(n) + \sum_{j=1}^{n} (p_j - \Lambda_j \cdot t_j) \cdot x_j(\Lambda_j) = \\
&= \sum_{i=1}^{n} \lambda_i d_i(n) + \sum_{j=1}^{n} (p_j(\Lambda_j) - \Lambda_j \cdot t_j(\Lambda_j)) = \\
&= \sum_{j=1}^{n} p_j(\Lambda_j) + \sum_{i=1}^{n} \lambda_i \cdot \left(d_i(n) - \sum\nolimits_{j=1}^{i} t_j(\Lambda_j)\right)
\end{aligned}$$

gdzie

$$x_j(\Lambda_j) = \begin{cases} 1 & \text{jeżeli } p_j - \Lambda_j t_j > 0 \\ 0 & \text{inaczej} \end{cases} \tag{4}$$

$$\begin{aligned}
p_j(\Lambda_j) &= \begin{cases} p_j & \text{jeżeli } p_j - \Lambda_j t_j > 0 \\ 0 & \text{inaczej} \end{cases}; \\
t_j(\Lambda_j) &= \begin{cases} t_j & \text{jeżeli } p_j - \Lambda_j t_j > 0 \\ 0 & \text{inaczej} \end{cases}
\end{aligned}$$

Oznaczmy:

$$\begin{aligned}
z_n(\Lambda) &= \sum_{j=1}^{n} p_j(\Lambda_j); \; s_i(\Lambda) = \sum_{j=1}^{i} t_j(\Lambda_j); \\
\hat{p}(\Lambda_j) &= \begin{cases} p_j(\Lambda_j) & \text{jeżeli } s_j(\Lambda) \leq d_j(n) \\ 0 & \text{inaczej} \end{cases}; \; \hat{z}_n(\Lambda) = \sum_{j=1}^{n} \hat{p}(\Lambda_j); \\
D_n(\Lambda) &= \sum_{i=1}^{n} \lambda_i \cdot d_i(n); \; S_n(\Lambda) = \sum_{i=1}^{n} \lambda_i \cdot s_i(\Lambda) = \\
&= \sum_{j=1}^{n} \Lambda_j \cdot t_j(\Lambda_j); \; \varphi_n(\Lambda) = z_n(\Lambda) + D_n(\Lambda) - S_n(\Lambda)
\end{aligned}$$

The problem dual to SJD (2) is then as follows:

$$\Phi_n^* = \min_{\Lambda \geq 0} \varphi_n(\Lambda)$$

Z właściwości $z_n(\Lambda)$, $\hat{z}_n$ (Λ), $S_n(\Lambda)$, $D_n(\Lambda)$, $\varphi_n(\Lambda)$ oraz $\Phi_n^*(\Lambda)$ mamy dla każdego $\Lambda \geq 0$:

$$z_n(\Lambda) \geq S_n(\Lambda)$$

oraz

$$\hat{z}_n(\Lambda) \leq z_{OPT}(n) \leq \Phi_n^* \leq \varphi_n(\Lambda) = z_n(\Lambda) + D_n(\Lambda) - S_n(\Lambda) \tag{5}$$

3 Analiza probabilistyczna

W dalszej części pracy będzie rozważany następujący model losowy zadania szeregowania prac z terminami zakończenia (2), patrz Szkatuła [16]:

- $n \geq 1$, n jest dodatnią liczbą całkowitą, $n \to \infty$, i, $j = 1, \ldots, n$;
- t_j, p_j, są realizacjami wzajemnie niezależnych zmiennych losowych o rozkładzie równomiernym w przedziale $(0, 1]$;
- $0 \leq d_1(n) \leq d_2(n) \leq \cdots \leq d_n(n)$ oraz $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$ są wielkościami deterministycznymi, $d_j(n)$ są funkcjami n.

Dystrybuanty i wartości oczekiwane zmiennych losowych $t_i(\Lambda_i)$, $p_i(\Lambda_i)$, w przypadku asymptotycznym, jako funkcje n, Λ_i, kiedy $n \longrightarrow \infty$, przyjmują następującą postać:

$$\begin{aligned} G_i(\Lambda_i, x) &= P\{t_i(\Lambda_i) < x\} = P\{t_i < x \cup t_i > x \cap p_i < \Lambda_i \cdot t_i\} = \\ &= \begin{cases} \left.\begin{matrix} \frac{\Lambda_i}{2} + x\left(1 - \frac{x \cdot \Lambda_i}{2}\right) & \Lambda_i \leq 1 \\ 1 - \frac{1}{2\Lambda_i} + x \cdot \left(1 - \frac{x \cdot \Lambda_i}{2}\right) & \Lambda_i > 1 \end{matrix}\right\} & 0 < x \leq \min\left\{1, \frac{1}{\Lambda_i}\right\} \\ 1 & x > \min\left\{1, \frac{1}{\Lambda_i}\right\} \end{cases} \end{aligned}$$

$$\begin{aligned} H_i(\Lambda_i, x) &= P\{H_i(\Lambda_i) < x\} = P\{p_i < x \cup t_i > x \cap p_i < \Lambda_i \cdot t_i\} = \\ &= \begin{cases} 1 - \frac{1 - x^2}{2 \cdot \Lambda_i} & \text{jeżeli} \quad \Lambda_i \geq 1 \\ \frac{\Lambda_i}{2} + \frac{x^2}{2 \cdot \Lambda_i} & \text{jeżeli} \quad x \leq \Lambda_i \leq 1 \\ x \quad \text{jeżeli} \quad \Lambda_i \leq x \leq 1 \\ 1 \quad \text{jeżeli} \quad x \geq 1 \end{cases} \end{aligned}$$

$$\begin{aligned} E(t_i(\Lambda_i)) &= \int_0^1 x \cdot dG_i(\Lambda_i, x) = \int_0^{\min\{1/\Lambda_i\}} x \cdot (1 - \Lambda_i \cdot x) dx = \\ &= \begin{cases} \frac{1}{6 \cdot \Lambda_i^2} & \text{if } \Lambda_i \geq 1 \\ \frac{1}{2} - \frac{\Lambda_i}{3} & \text{otherwise } 0 \leq \Lambda_i \leq 1 \end{cases} \end{aligned}$$

$$E(p_i(\Lambda_i)) = \int_0^1 x \cdot dH_i(\Lambda_i, x) = \int_0^1 x \cdot G_i(\Lambda_i, x/\Lambda_i)\, dx =$$
$$= \begin{cases} \frac{1}{3 \cdot \Lambda_i} & \text{jeżeli } \Lambda_i \geq 1 \\ \frac{1}{2} - \frac{\Lambda_i^2}{6} & \text{inaczej } 0 \leq \Lambda_i \leq 1 \end{cases}$$

Szukamy takich wartości $\lambda_1, \ldots, \lambda_n \geq 0$, że zachodzi

$$E(s_i(\Lambda_i(n))) \leq d_i(n), \text{ dla wszystkich } i = 1, \ldots, n \tag{6}$$

Mamy $\Lambda_j = \sum_{i=j}^{n} \lambda_i$, $\lambda_j \geq 0$, $j = 1, \ldots, n$ oraz

$$\Lambda_1 \geq \Lambda_2 \geq \cdots \geq \Lambda_n \text{ i dlatego } E(t_1(\Lambda_1)) \leq E(t_2(\Lambda_2)) \leq \cdots \leq E(t_n(\Lambda_n)). \tag{7}$$

Zauważmy, że jeśli dla pewnych wartości i, $1 \leq i \leq n$

$$E(s_i(\Lambda)) = d_i(n) \text{ oraz } d_{i+1}(n) - d_i(n) < E(t_{i+1}(\Lambda_{i+1})) \tag{8}$$

wtedy

$$E(s_{i+1}(\Lambda) > d_{i+1}(n)$$

co oznacza, że jeśli dla pewnych terminów zakończenia $d_1(n), d_2(n), \cdots, d_n(n)$, jest wyznaczony wektor Λ, taki że zachodzi (8) to wtedy (6) nie będzie spełnione dla wszystkich $i = 1, \ldots, n$. Jest to spowodowane monotonicznością Λ_j oraz $E(t_j(\Lambda_j)$, patrz (7). Dlatego w pracy Szkatuła [16] został zaproponowany algorytm rekurencyjny wyznaczający wartości $\Lambda(n)$, $\lambda_i(n) \geq 0$, $i = 1, \ldots, n$, takie że dla wszystkich $0 \leq d_1(n) \leq d_2(n) \leq \cdots \leq d_n(n)$ będzie spełnione (6.

Algorytm 1 *Procedura wyznaczenia $\delta_j(n)$ oraz $\Lambda_j(n)$, $j = 1, \ldots, n$*

Krok Inicjalizujący: Let $l \leftarrow 0$, $d_0(n) \leftarrow 0$
Główny Krok Rekurencyjny: Niech

$$j^* = \max_{l < m \leq n} \left\{ m \,\middle|\, \frac{d_m(n) - d_l(n)}{m - l} = \min_{l < j \leq n} \frac{d_j(n) - d_l(n)}{j - l} \right\} \tag{9}$$

oraz

$$\delta_k(n) = \min \left\{ \frac{d_{j^*}(n) - d_l(n)}{j^* - l}, \frac{1}{2} \right\};\ \Lambda_k(n) = \arg\{E(t_k(\Lambda_k)) = \delta_k(n)\} \tag{10}$$

dla $k = l + 1, \ldots, j^*$.
Krok Sprawdzający: jeśli $j^* = n$ to wtedy procedura jest zakończona. W przeciwnym przypadku $l \leftarrow j^*$ i **Główny Krok Rekurencyjny** jest powtarzany do momentu kiedy $j^* = n$.

Po zakończeniu pracy Algorytmu 1 wartości $\delta_1(n), \delta_2(n), \ldots, \delta_n(n)$ oraz $\Lambda_1(n)$, $\Lambda_2(n), \ldots, \Lambda_n(n)$ są wyznaczone. Poniżej są przedstawione ich wybrane właściwości:

- $\delta_1(n) \leq \delta_2(n) \leq \ldots \leq \delta_n(n)$;
- $\sum_{j=1}^{i} \delta_j(n) \leq d_i(n)$. Jeśli $\delta_i(n) < \delta_{i+1}(n)$ to wtedy $\sum_{j=1}^{i} \delta_j(n) = d_i(n)$;

- $\Lambda_j(n) = \begin{cases} \sqrt{\frac{1}{6\cdot\delta_j(n)}} & \text{jeśli } 0 < \delta_j(n) \le \frac{1}{6} \\ \frac{3}{2} - 3\cdot\delta_j(n) & \text{jeślif } \frac{1}{6} < \delta_j(n) \le \frac{1}{2} \end{cases}, j = 1, \ldots, n;$

- $\Lambda_1(n) \ge \Lambda_2(n) \ge \ldots \ge \Lambda_n(n);$

- Jeśli $\delta_j(n) = \delta_{j+1}(n)$ to wtedy $\Lambda_j(n) = \Lambda_{j+1}(n)$, $\lambda_j(n) = 0$, $E(s_j(\Lambda(n)) \le d_j(n)$;

- Jeśli $\delta_j(n) < \delta_{j+1}(n)$ to wtedy $\Lambda_j(n) < \Lambda_{j+1}(n)$, $\lambda_j(n) > 0$, $E(s_j(\Lambda(n))) = d_j(n)$;

- $\sum_{i=1}^{n} \lambda_i(n)\cdot\left(\sum_{j=1}^{i} \delta_j(n)\right) = \sum_{i=1}^{n} \lambda_i(n)\cdot E(s_i(\Lambda(n))) = \sum_{i=1}^{n} \lambda_i(n)\cdot d_i(n);$

- $E(p_j(\Lambda_j(n))) = \begin{cases} \sqrt{\frac{2}{3}}\cdot\delta_j(n) & \text{jeśli } 0 < \delta_j(n) \le \frac{1}{6} \\ \frac{1}{8} + \frac{3}{2}\cdot\delta_j(n)\cdot(1-\delta_j(n)) & \text{jeśli } \frac{1}{6} \le \delta_j(n) \le \frac{1}{2} \end{cases}.$

Z powyższego wynika, że:

$$\begin{aligned} E(Z_n(\Lambda(n))) &= \sum_{j=1}^{n} E(p(\Lambda_j(n))) = \\ &= \sum_{j=1}^{n} \left(\tau_j(n)\left(\frac{1}{8} + \frac{3}{2}\delta_j(n)(1-\delta_j(n))\right) + \bar{\tau}_j(n)\sqrt{\frac{2\cdot\delta_j(n)}{3}}\right) \end{aligned}$$

gdzie $\tau_j(n) = \begin{cases} 1 & \text{jeśli } \frac{1}{6} < \delta_j(n) \le \frac{1}{2} \\ 0 & \text{inaczej} \end{cases}$ oraz $\bar{\tau}_j(n) = 1 - \tau_j(n)$, $j = 1, \ldots, n$.

Główny wynik pracy Szkatuła [16] jest przedstawiony w poniższym twierdzeniu.

Twierdzenie 1 *Niech p_j, t_j, $j = 1, \ldots, n$, będą raelizacjami wzajemnie niezależnych zmiennych losowych o rozkładzie równomiernym w przedziale $(0, 1]$, $d_1(n) \le d_2(n) \le \ldots \le d_n(n)$, gdzie $d_j(n)$ są funkcjami deterministycznymi oraz $\delta_j(n)$ są określone przez Agorytm 1. Jeśli*

$$\frac{\ln(n)}{n\cdot\delta_1(n)} \approx 0$$

to wtedy:

$$z_{OPT}(n) \approx \sum_{j=1}^{n} \left(\tau_j(n)\left(\frac{1}{8} + \frac{3\cdot\delta_j(n)}{2}(1-\delta_j(n))\right) + \bar{\tau}_j(n)\sqrt{\frac{2\cdot\delta_j(n)}{3}}\right) \quad (11)$$

Głowna idea dowodu Twierdzenia 1 jest oparta na wykazaniu, że:

$$\hat{z}_n(\Lambda(n)) \approx E(z_n(\Lambda(n))) \approx \varphi_n(\Lambda(n)) \text{ oraz } E(z_n(\Lambda(n))) \approx z_n(\Lambda(n))$$

oraz zastosowaniu (5). Więcej szczegółów znajduje się w pracy Szkatuła [16].

4 Przedziały terminów zakończenia oraz algorytm przybliżony

Opisany w rozdziale 3 sposób wyznaczenia przez Algorytm 1 wartości $\delta_j(n)$, $j = 1, \ldots, n$, oznacza, że wartości $\delta_j(n)$ oraz ich właściwości w istotnym stopniu zależą od wzajemnych relacji pomiędzy $d_j(n)$, $j = 1, \ldots, n$. It is assumed that $d_j(n)$ are monotonic cf.

W (1) przyjęto założenie, że $d_j(n)$ są monotoniczne. Wzór (9) w Algorytmie 1 może zostać przedstawiony w poniższym sformułowaniu

$$j^* = \max_{l<j\leq n} \left\{ j \,\middle|\, d_{j-1}(n) \geq \frac{j-1}{j} \cdot d_j(n) \right\} \tag{12}$$

gdzie $l \leftarrow 0$ w kroku inicjalizującym oraz w kroku rekurencyjnym jeśli $j^* < n$ to wtedy $l \leftarrow j^*$. Algorytm 1 kończy pracę kiedy $j^* = n$. Wzór (12) pokazuje, że wartości $\delta_1(n), \delta_2(n), \ldots, \delta_n(n)$ zależą od

$$d_1(n), d_2(n) - d_1(n), \ldots, d_n(n) - d_{n-1}(n).$$

W szczególności z właściwości wynika $\delta_j(n)$, $\Lambda_j(n) = \sum_{i=j}^{n} \lambda_i(n)$, $j = 1, \ldots, n$, że:

$$\begin{aligned} \text{jeśli } \delta_j(n) &= \delta_{j+1}(n) \text{ to } \Lambda_j(n) = \Lambda_{j+1}(n),\ \lambda_j(n) = 0 \text{ oraz } E(s_j(\Lambda(n))) \leq d_j(n); \\ \text{jeśli } \delta_j(n) &< \delta_{j+1}(n) \text{ to } \Lambda_j(n) > \Lambda_{j+1}(n),\ \lambda_j(n) > 0 \text{ oraz } E(s_j(\Lambda(n))) = d_j(n). \end{aligned}$$

Z teorii dualności wynika, że jeśli $\lambda_i(n) = 0$ to odpowiednie ograniczenie w (2) jest „nieaktywne". Oznacza to, że jest ono spełnione przez pozostałe „aktywne" ograniczenia dla których $\lambda_j(n) > 0$, $i, j \in \{1, \ldots, n\}$. Wzór (12) pozwala wyodrębnić 3 różne klasy $\delta_j(n)$, $\Lambda_j(n)$, $\lambda_i(n)$, $i, j = 1, \ldots, n$.

Lemat 1 *Jeśli dla wszystkich* $j = 2, \ldots, n$

$$d_{j-1}(n) \geq \frac{j-1}{j} \cdot d_j(n) \tag{13}$$

to wtedy

$$\delta_1(n) = \delta_2(n) = \ldots = \delta_n(n); \Lambda_i(n) = \Lambda_{i+1}(n) = \lambda_n(n), \lambda_i(n) = 0, i = 1, \ldots, n-1. \tag{14}$$

Dowód. Jeśli (13) zachodzi to w Algorytmie 1, zgodnie z (9) oraz (12), Główny Krok Rekurencyjny jest wykonywany jednokrotnie dla $j^* = n$; (14) wynika wprost z (10). ∎

W przypadku rozważanym w Lemacie 1 tylko ostatnie ograniczenie jest aktywne i w tym przypadku zadanie szeregowania prac z terminami zakończenia przyjmuje postać jednowymiarowego zadania załadunku (3).

Lemat 2 *Jeśli dla wszystkich* $j = 2, \ldots, n$

$$d_{j-1}(n) < \frac{j-1}{j} \cdot d_j(n) \tag{15}$$

to wtedy

$$\delta_1(n) < \delta_2(n) < \ldots < \delta_n(n);\ \Lambda_1(n) > \Lambda_2(n) > \ldots > \Lambda_n(n),\ \lambda_i(n) > 0, \tag{16}$$

dla $i = 1, \ldots, n$.

Dowód. Jeśli (15) zachodzi to w Algorytmie 1 , zgodnie z (9) oraz (12), Główny Krok Rekurencyjny jest wykonywany n razy dla $j^* = 1, 2, \ldots, n$; (16) wynika wprost z (10). ■

W przypadku rozważanym w Lemacie 2 wszystkie ograniczenia są aktywne.

Lemat 3 *Jeśli istnieje $j\prime$, $1 < j\prime < n-1$ takie, że*

$$d_{j-1}(n) \geq \frac{j-1}{j} \cdot d_j(n),\ j = 2, \ldots, j\prime;\ \textit{and}\ d_{j\prime+1}(n) < \frac{j\prime+1}{j\prime+2} \cdot d_{j\prime+2}(n) \quad (17)$$

to wtedy

$$\delta_1(n) = \ldots = \delta_{j\prime}(n) < \delta_{j\prime+1}(n);\ \Lambda_1(n) = \ldots = \Lambda_{j\prime}(n) > \Lambda_{j\prime+1}(n),\ \lambda_i(n) = 0, \quad (18)$$

gdzie $i = 1, \ldots, j\prime$.

Dowód. W tym przypadku kiedy (17) zachodzi to w Algorytmie 1, zgodnie z (9) oraz (12), pierwsze wykonanie Głównego Kroku Rekurencyjnego daje $j^* = j\prime$. Następnie rozpoczynając od $l = j\prime + 1$ Główny Krok Rekurencyjny zostanie wykonany co najmniej raz; (18) wynika wprost z (10). ■

W Lemacie 3 został rozważony przypadek kiedy co najmniej $j\prime$ pierwszych ograniczeń jest nieaktywnych oraz pewne ograniczenia, przynajmniej $j\prime + 1$, są aktywne. Może też się zdarzyć, że ta sytuacja będzie się powtarzać. Na przykład mogą istnieć $j\prime\prime$ oraz $j\prime\prime\prime$, $j\prime < j\prime\prime < j\prime\prime\prime < n$, takie, że ograniczenia $j\prime+1, \ldots, j\prime\prime$ są aktywne, ograniczenia $j\prime\prime + 1, \ldots, j\prime\prime\prime$ są nieaktywne i tak dalej.

Dla dowolnego zbioru $d_1(n), d_2(n), \ldots, d_n(n)$ Lematy 1, 2 i 3 przedstawiają wszystkie możliwe relacje pomiędzy terminami zakończenia prac oraz wynikający z nich status „aktywności" ograniczeń.

Trzy powyżej zaprezentowane lematy pozwalają zdefiniować przedziały terminów zakończenia odpowiadające trzem rozważonym przypadkom. Poniższe twierdzenie prezentuje główny wynik pracy.

Twierdzenie 2 *Jeżeli dla wszystkich $j = 2, \ldots, n-1$ zachodzi*

$$d_j(n) \in \left[\frac{j}{j+1} \cdot d_{j+1}(n), \frac{j}{j-1} \cdot d_{j-1}(n)\right] \quad (19)$$

to wtedy również zachodzi Lemat 1 i (14) opisuje wzajemne relacje pomiędzy $\delta_j(n)$, $\Lambda_j(n)$, $\lambda_i(n)$, $i, j = 1, \ldots, n$. Jeżeli dla wszystkichl $j = 2, \ldots, n-1$ zachodzi

$$d_j(n) \in \left(\frac{j}{j-1} \cdot d_{j-1}(n), \frac{j}{j+1} \cdot d_{j+1}(n)\right) \quad (20)$$

to wtedy również zachodzi Lemat 2 i (16) opisuje wzajemne relacje pomiędzy $\delta_j(n)$, $\Lambda_j(n)$, $\lambda_i(n)$, $i, j = 1, \ldots, n$.
Jeżeli istnieje $j\prime$, $2 < j\prime < n-1$ takie, że dla wszystkich $j = 2, \ldots, j\prime - 1$ zachodzi

$$d_j(n) \in \left[\frac{j}{j+1} \cdot d_{j+1}(n), \frac{j}{j-1} \cdot d_{j-1}(n)\right]\ \textit{oraz}\ d_{j\prime+1}(n) < \frac{j\prime+1}{j\prime+2} \cdot d_{j\prime+2}(n) \quad (21)$$

to wtedy również zachodzi Lemat 3 i (18) opisuje wzajemne relacje pomiędzy $\delta_j(n)$, $\Lambda_j(n)$, $\lambda_i(n)$, $i, j = 1, \ldots, n$.

Dowód. W celu udowodnienia Twierdzenia 2 wystarczy zauważyć, że Lemat 1 dowodzi prawdziwości (19), Lemat 2 dowodzi prawdziwości (20) oraz Lemat 3 dowodzi prawdziwości (21). ■

Wzory (19), (20) oraz (21) definiują w sposób rekurencyjny przedziały terminów zakończenia. Jeżeli terminy zakończenia należą do przedziałów zdefiniowanych odpowiednio przez wzory (19), (20) oraz (21) to ten fakt gwarantuje, że $\delta_j(n)$, $\Lambda_j(n)$, $\lambda_i(n)$, $i, j = 1, \dots, n$, będą należeć do jednej z 3 różnych klas zdefiniowanych przez Lematy 1-3 oraz Twierdzenie 2. To z kolei oznacza, że w ramach tych przedziałów wartości terminów zakończenia mogą się zmieniać, co daje pewną elastyczność w ustalaniu ich wartości. Z powodu rekurencyjnego sposobu definiowania przedziałów terminów zakończenia szczególny wpływ na postać tych przedziałów odgrywa pierwszy i ostatni z nich. Poniżej zostały zaprezentowane 3 przykłady ilustrujące zdefiniowane klasy $\delta_j(n)$, $\Lambda_j(n)$, $\lambda_i(n)$, $i, j = 1, \dots, n$.

Przykład 1

Niech

$$d_j(n) = \frac{j}{2} \text{ i wtedy } \delta_j(n) = \frac{1}{2},\ j = 1, \dots, n;\ d_n(n) = \frac{n}{2}.$$

W tym przypadku założenia (13) Lematu 1 oraz Twierdzenia 1 są spełnione i zgodnie z (11) mamy:

$$z_{OPT}(n) \approx \frac{1}{2} \cdot n \tag{22}$$

co oznacza, że w tym przypadku tylko ostatnie ograniczenie jest aktywne, rozwiązanie optymalne uzyskuje maksymalną wartość, wszystkie prace zostaną wykonane przed terminami zakończenia w przypadku asymptotycznym dla losowego modelu zadania szeregowania prac z terminami zakończenia.

Przykład 2

Niech

$$d_j(n) = \frac{j^2}{2n} \text{ i wtedy } \delta_j(n) = \frac{2j-1}{2n},\ j = 1, \dots, n;\ d_n(n) = \frac{n}{2}.$$

W tym przypadku założenia (15) Lematu 2 oraz Twierdzenia 1 są spełnione. Wszystkie ograniczenia są aktywne, wszystkie prace zostaną wykonane przed terminami zakończenia, zachodzi (22).

Przykład 3

Niech

$$\begin{aligned} d_j(n) &= \frac{j}{4} \text{ i wtedy } \delta_j(n) = \frac{1}{4},\ j = 1, \dots, n^*,\ n^* = \left\lceil \frac{n}{2} \right\rceil, \\ d_j(n) &= \frac{j^2}{2n} \text{ i wtedy } \delta_j(n) = \frac{2j-1}{2n},\ j = n^* + 1, \dots, n;\ d_n(n) = \frac{n}{2}. \end{aligned}$$

W tym przypadku założenia (17) Lematu 3, gdzie $j' = n^*$ oraz Twierdzenia 1 są spełnione. Zgodnie z (10) $\delta_j(n) = \frac{1}{2}$, $j = n^* + 1, \dots, n$, z (11) uzyskujemy:

$$z_{OPT}(n) \approx \frac{1}{2} \cdot n - \frac{3}{32} \cdot n^* - \frac{1}{2},$$

co oznacza że w tym przypadku tylko część ograniczeń będzie aktywna, a część nieaktywna, wartość rozwiązania optymalnego będzie mniejsza od możliwej maksymalnej wartości. Cześć prac nie zostanie wykonana przed terminami zakończenia (to znaczy nie zostanie uzyskany związany z nimi zysk) w przypadku asymptotycznym dla losowego modelu zadania szeregowania prac z terminami zakończenia.

Poniżej jest zaprezentowany heurystyczny algorytm typu zachłannego przeznaczony dla zadania szeregowania prac z terminami zakończenia (2) w ogólnym - deterministycznym przypadku. Ten algorytm wykorzystuje procedurę analogiczną do zastosowanej a Algorytmie 1.

Algorytm 2

Krok Inicjalizujący: Niech $l \leftarrow 0$, $d_0(n) \leftarrow 0$
Główny Krok Rekurencyjny: Niech

$$j^* = \max_{l<m\leq n} \left\{ m \;\middle|\; \frac{d_m(n) - d_l(n)}{m-l} = \min_{l<j\leq n} \frac{d_j(n) - d_l(n)}{j-l} \right\}$$

oraz

$$\Lambda_{j^*} = \min_{\Lambda \in \Phi_{j^*}} \left\{ \Lambda \;\middle|\; \sum_{j=l+1}^{j^*} t_j(\Lambda) \leq d_{j^*}(n) \right\}; \Phi_{j^*} = \left\{ \frac{p_{l+1}}{t_{l+1}}, \ldots, \frac{p_{j^*}}{t_{j^*}} \right\} \tag{23}$$

$$x_j \leftarrow x_j(\Lambda_{j^*}),\; j = l+1, \ldots, j^*, \text{ refer to (4).}$$

Krok Sprawdzający: jeśli $j^* = n$ to wtedy Algorytm 2 i kończy pracę. W przeciwnym przypadku $l \leftarrow j^*$ i **Główny Krok Rekurencyjny** jest powtarzany do momentu kiedy $j^* = n$

Algorytm 2 ma wyjątkowo niską złożoność obliczeniową rzędu $O(n)$. Algorytm ten nie wymaga sortowania elementów, jak na ogół ma to miejsce w przypadku algorytmów zachłannych. Z obliczeniowego punktu widzenia najbardziej kosztownymi operacjami są max oraz min, których złożoność obliczeniowa jest rzędu $O(k)$, gdzie k ilość elementów odpowiedniego zbioru. W przypadku kiedy operacje max oraz min są powtarzane k razy w zbiorze k elementów to złożoność obliczeniowa jest rzędu $O(k^2)$. Zgodnie z Lematami 1 - 3 ta sytuacja nie może się wydarzyć, gdyż albo **Główny Krok Rekurencyjny** jest wykonywany jednokrotnie (Lemat 1) albo zostanie on wykonany kilka razy (n w przypadku Lematu 2), ale liczba elementów odpowiednich zbiorów będzie niewielka, to znaczy rzędu $\frac{n}{m}$, gdzie n jest liczbą prac do wykonania (inaczej rozmiarem zadania), m liczbą koniecznych powtórzeń **Głównego Kroku Rekurencyjnego**. To oznacza, że ogólna złożoność obliczeniowa Algorytmu jest rzędu $O(n)$.

W sensie analizy najgorszego przypadku ten algorytm zawsze uzyskuje rozwiązanie dopuszczalne zadania szeregowania prac z terminami zakończenia (2) co jest zagwarantowane przez (23), ale nie ma on żadnych gwarancji jakości (dokładności) uzyskanego rozwiązania. Powyższe oznacza, że w sensie analizy najgorszego przypadku jest to algorytm heurystyczny. Natomiast dla rozważanego w pracy modelu losowego zadania szeregowania (2) Algorytm 2 jest asymptotycznie suboptymalny. W tym przypadku Algorytm 2 zachowuje się identycznie z procedurą wykorzystywaną w Algorytmie 1. Asymptotyczna sub-optymalność Algorytmu wynika z Twierdzenia 1.

5 Uwagi końcowe

W pracy został rozważony jeden z klasycznych problemów teorii harmonogramowania - zadanie szeregowania prac z terminami zakończenia. Na podstawie poprzednich wyników autora, patrz Szkatuła [16] została przeprowadzona analiza probabilistyczna wpływu wzajemnych relacji przedziałów terminów zakończenia oraz funkcjonalnych właściwości w przypadku losowego modelu zadania. W wyniku przeprowadzonej analizy trzy szczególne kategorie wzajemnych relacji przedziałów terminów zakończenia zostały określone. Następnie na podstawie wyników uzyskanych Lematach 1, 2 oraz 3 zostało sformułowane Twierdzenie 2,w którym zdefiniowane zostały rekurencyjne przedziały terminów zakończenia dla rozważanego w pracy losowego modelu zadania szeregowania prac z terminami zakończenia. W świetle uzyskanych wyników rola pierwszego i ostatniego ograniczenia, $d_1(n)$ oraz $d_n(n)$ odpowiednio, jest zasadnicza. Przedziały terminów zakończenia mogą dać znaczącą elastyczność w formułowaniu zadań szeregowania prac z terminami zakończenia, gdyż wzajemne relacje pomiędzy terminami zakończenia mogą być analizowane w znacznie wygodniejszy i bardziej przekonywujący sposób.

Innym ciekawym wynikiem pracy jest prosty algorytm heurystyczny o bardzo niskiej złożoności obliczeniowej, który w średnim przypadku, to znaczy dla rozważanego w pracy modelu losowego zadania jest asymptotycznie sub-optymalny.

Uzyskane wyniki wzbogacają bazę wiedzy problemów harmonogramowania, w szczególności dla zadania szeregowania prac z terminami zakończenia, zarówno z punktu widzenia teorii jak również możliwości rozwiązywania praktycznych problemów.

Podziękowania

Praca była częściowo wsparta przez Narodowe Centrum Nauki (NCN), decyzja nr.: DEC-2011/03/B/ST6/00364.

Literatura

[1] J. Błażewicz, K. Ecker, E. Pesch, G. Schmidt, and J. Weglarz. *Scheduling Computer and Manufacturing Processes.* Springer Verlag, 1996.

[2] K. Dudziński and K. Szkatuła. A note on sequencing jobs with deadlines problem. *European Journal of Operational Research*, 59:333–336, 1992.

[3] A. Frieze and M. Clarke. Approximation algorithms for the m-dimensional 0-1 knapsack problem: Worst case and probabilistic analysis. *European Journal of Operational Research*, 15:100–109, 1984.

[4] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, San Francisco, 1979.

[5] A. Janiak, W. A. Janiak, T. Krysiak, and T. Kwiatkowski. A survey on scheduling problems with due windows. *European Journal of Operational Research*, 242:347–357, 2015.

[6] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems.* Springer Verlag, 2004.

[7] G. Kocur. *Courses on Computer Algorithms in Systems Engineering, Greedy algorithms (lecture 10): knapsack, job sequence.* Massachusetts Institute of Technology: MIT OpenCourseWare, 2010.

[8] E. Lawler and J. Moore. A functional equation and its application to resource alocation and sequencing problems. *Management Science*, 16:77–84, 1969.

[9] J. Mamer and K. Schilling. On the growth of random knapsacks. *Discrete Applied Mathematics*, 28:223–230, 1990.

[10] A. Puntambekar. *Analysis of Algorithm and Design.* Technical Publications, 2009.

[11] S. Sahni. Algorithms for scheduling independent jobs. *Journal of ACM*, 23:116–127, 1976.

[12] K. Schilling. The growth of m-constraint random knapsacks. *European Journal of Operational Research*, 46:109–112, 1990.

[13] K. Schilling. Random knapsacks with many constraints. *Discrete Applied Mathematics*, 48:163–174, 1994.

[14] K. Szkatuła. On the growth of multi-constraint random knapsacks with various right-hand sides of the constraints. *European Journal of Operational Reserch*, 73:199–204, 1994.

[15] K. Szkatuła. The growth of multi-constraint random knapsacks with large right-hand sides of the constraints. *Operations Research Letters*, 21:25–30, 1997.

[16] K. Szkatuła. Random sequencing jobs with deadlines problem: growth of the optimal solutions values. *European Journal of Operational Research*, 109:160–169, 1998.