Wydział Matematyki i Nauk Informacyjnych
Politechniki Warszawskiej

Teoria Algorytmów i Obliczeń
Laboratorium - Etap 1

Specyfikacja wstępna

Adrian Bednarz,
Bartłomiej Dach,
Tymon Felski

Wersja 1.0

19.10.2017

Lista zmian:

Data	Autor	Opis	Wersja
15.10.2017	Tymon Felski	Stworzenie szablonu dokumentu	1.0

Spis treści

1 Opis problemu

Niniejszy rozdział poświęcony jest dokładnemu opisaniu podstawowej wersji zadanego problemu.

Dane są zbiory:

- E - **ekspertów** realizujących projekty,
- S - **umiejętności** posiadanych przez ekspertów,
- P - **projektów** do zrealizowania.

Każdemu ekspertowi przypisany jest wektor binarny opisujący posiadane przez niego umiejętności. Przykładowo, jeżeli ekspert posiada umiejętność i, to w wektorze umiejętności odpowiadającemu temu ekspertowi na i-tym miejscu znajduje się znak 1, w przeciwnym wypadku - 0.

`Załóżmy, że liczność zbioru umiejętności S jest równa 5. Wówczas ponumerujemy umiejętności rozważane w problemie liczbami z zakresu [1, 5].`

`Pewien ekspert ze zbioru E posiada umiejętności 2, 3 i 5. Wówczas wektor opisujący jego umiejętności to:`

`[0, 1, 1, 0, 1]`

Przykład 1: Wektor opisujący umiejętności eksperta

Każdemu projektowi przypisany jest wektor liczbowy opisujący zapotrzebowanie na ekspertów posiadających dane umiejętności. Przykładowo, jeżeli do realizacji projektu potrzeba trzech ekspertów posiadających umiejętność i, to w wektorze umiejętności odpowiadającemu temu projektowi na i-tym miejscu znajduje się liczba 3.

`Utrzymując założenie o liczności zbioru umiejętności z poprzedniego przykładu, rozważmy pewien projekt ze zbioru P. Niech jego zapotrzebowanie na ekspertów posiadających umiejętności 1 i 4 wynosi odpowiednio 4 i 3, a na pozostałe - 0. Wówczas wektor opisujący zapotrzebowanie tego projektu to:`

`[4, 0, 0, 3, 0]`

Przykład 2: Wektor opisujący zapotrzebowanie projektu

Wszystkie projekty realizowane są w tym samym oknie czasowym, tzn. prace nad każdym z nich rozpoczynają się w momencie t_0 i kończą w późniejszym momencie t_k. Oznacza to, że jeżeli dany ekspert zostanie zatrudniony do pracy nad projektem P_1, to nie będzie mógł brać udziału w równoległym projekcie P_2. Ponadto każdy ekspert podczas pracy nad projektem może wykorzystywać tylko jedną z posiadanych umiejętności i nie może jej zmienić w trakcie trwania prac.

Prace nad danym projektem zostaną zakończone nawet jeżeli nie zostanie mu przydzielona wymagana liczba ekspertów posiadających potrzebne umiejętności, określona przez wektor liczbowy odpowiadający temu projektowi. Będzie on natomiast zrealizowany gorzej niż w przypadku, gdyby przypisana została odpowiednia liczba ekspertów. Może się również zdarzyć, że najbardziej optymalne okaże się takie przypisanie ekspertów, że nad pewnym projektem nie będzie pracował nikt.

Wówczas mamy do czynienia z brakami. Poprzez braki rozumiemy różnicę pomiędzy zapotrzebowaniem projektu na ekspertów o danych umiejętnościach a rzeczywistym przydziałem. Aby wyznaczyć braki w danym projekcie, należy odjąć wektor zapotrzebowania projektu na ekspertów od wektora zawierającego informację o ekspertach przydzielonych do tego projektu i zsumować elementy uzyskanej różnicy. Dokładna definicja tego pojęcia znajduje się w rodziale zawięrającym dowód poprawności (definicja nr 8).

```
Niech wektorem opisującym zapotrzebowanie projektu na ekspertów będzie:
                            [4, 0, 0, 3, 0]
Załóżmy, że do tego projektu zostali przypisani eksperci opisani przez:
                            [1, 0, 1, 0, 1] (wykorzystuje 1)
                            [1, 0, 0, 0, 1] (wykorzystuje 1)
                            [1, 1, 0, 1, 0] (wykorzystuje 4)
                            [0, 0, 0, 1, 1] (wykorzystuje 4)
Stąd przydział ekspertów do tego projektu można opisać wektorem:
                            [2, 0, 0, 2, 0]
Braki w tym projekcie obliczymy następująco:
    sum([4, 0, 0, 3, 0] - [2, 0, 0, 2, 0]) = sum([2, 0, 0, 1, 0]) = 3
```

Przykład 3: Wektor opisujący zapotrzebowanie projektu

Naszym celem jest zminimalizowanie braków w obrębie wszystkich projektów (sumy wszystkich braków), czyli znalezienie optymalnego przydziału ekspertów do projektów.

2 Problem znajdowania maksymalnego przepływu

Okazuje się, że problem opisany w sekcji 1 można uogólnić do znanego problemu znajdowania maksymalnego przepływu w sieciach. W tej sekcji zdefiniowane są podstawowe pojęcia potrzebne do opisu tego problemu.

Definicja 1. **Siecią** nazywamy czwórkę uporządkowaną $S = (G, c, s, t)$, gdzie:

- $G = (V, E)$ jest grafem skierowanym,
- $c : E \to \mathbb{N}$ to tzw. funkcja przepustowości,
- $s, t \in V, s \neq t$ są dwoma wyróżnionymi wierzchołkami grafu G — kolejno źródłem i ujściem sieci.

Definicja 2. **Przepływem** w sieci S nazywamy funkcję $f : E \to \mathbb{N}$ spełniającą następujące warunki:

1. $\forall_{e \in E} \quad 0 \leqslant f(e) \leqslant c(e)$,

2. $(\forall v \in V - \{s,t\}) \sum_{u\colon uv \in E} f(uv) = \sum_{u\colon vu \in E} f(vu)$ — tzw. prawo Kirchhoffa.

W ogólniejszym przypadku funkcje przepustowości i przepływu mogą mieć wartości nieujemne rzeczywiste, lecz założenie o całkowitości zapewnia, że algorytmy wyznaczające maksymalny przepływ zawsze zakończą działanie.

Prawo Kirchhoffa stanowi, że suma wartości przepływu na krawędziach wchodzących do danego wierzchołka musi być równa sumie wartości przepływu na krawędziach wychodzących z tego wierzchołka.

Definicja 3. Wartością przepływu f w sieci S nazywamy liczbę

$$W(f) = \sum_{u\colon su \in E} f(su) - \sum_{u\colon us \in E} f(us)$$

Powyższe definicje wystarczą, aby zdefiniować problem maksymalnego przepływu.

Definicja 4 (Problem maksymalnego przepływu)**.** Dana jest sieć $S = (G, c, s, t)$. Szukamy przepływu f o maksymalnej wartości $W(f)$, zwanego również **przepływem maksymalnym**.

Zagadnienie znajdowania maksymalnego przepływu jest rozwiązywalne przez wiele zachłannych algorytmów opartych na metodzie Forda-Fulkersona, polegającej na znajdowaniu ścieżek w tzw. sieci rezydualnej. Szczegółowy opis jednego z takich algorytmów znajduje się w następnej sekcji.

3 Algorytm

4 Dowód poprawności

W tej sekcji wykażemy związek między postawionym problemem a zagadnieniem wyznaczania przepływu maksymalnego oraz równoważność rozwiązań obu zadań.

Na początek zdefiniujmy w sposób formalny pojęcia użyte w oryginalnym zadaniu. Załóżmy, że dane są następujące zbiory:

- **zbiór ekspertów**, oznaczony E,
- **zbiór umiejętności**, oznaczony S,
- **zbiór projektów**, oznaczony P.

Definicja 5. Funkcją umiejętności nazywamy funkcję

$$\text{ability} : E \times S \to \{0, 1\}$$

gdzie dla eksperta $e \in E$ oraz umiejętności $s \in S$ zachodzi $\text{ability}(e, s) = 1$ wtedy i tylko wtedy, gdy ekspert e posiada umiejętność s, zaś 0 w przeciwnym przypadku.

Definicja 6. Zapotrzebowaniem projektu nazywamy funkcję

$$\text{need} : P \times S \to \mathbb{N}$$

gdzie dla projektu $p \in P$ i umiejętności $s \in S$ zachodzi $\text{need}(p, s) = n$ wtedy i tylko wtedy, gdy w projekcie p liczba potrzebnych ekspertów w dziedzinie umiejętności s wynosi n.

Definicja 7. Przyporządkowaniem eksperta nazywamy funkcję

$$\text{assign} : E \to P \times S$$

gdzie dla projektu $p \in P$ i umiejętności $s \in S$ zachodzi $\text{assign}(e) = (p, s)$ wtedy i tylko wtedy, gdy:

- ekspert e posiada umiejętność s (tj. $\text{ability}(e, s) = 1$),
- ekspert e został przyporządkowany do pracy w projekcie p w dziedzinie umiejętności s.

Ponadto, dla każdego projektu p i umiejętności s musi zachodzić

$$\left|\text{assign}^{-1}[p, s]\right| \leq \text{need}(p, s)$$

gdzie $\text{assign}^{-1}[p, s]$ jest przeciwobrazem funkcji assign dla argumentów p, s.

Definicja 8. Liczbą braków w projekcie p dla danego przyporządkowania assign nazywamy liczbę

$$\text{missing}(p, \text{assign}) = \sum_{s \in S} \left(\text{need}(p, s) - \left|\text{assign}^{-1}[p, s]\right|\right)$$

Definicja 9. Całkowitą liczbą braków dla danego przyporządkowania assign nazywamy liczbę

$$M(\text{assign}) = \sum_{p \in P} \text{missing}(p, \text{assign})$$

Widoczne jest, że M jest parametrem minimalizowanym w postawionym problemie, zależnym od końcowego przyporządkowania.

Na podstawie powyższych definicji skonstruujemy teraz sieć, której użyjemy do wyznaczenia rozwiązań problemu.

Definicja 10. Siecią przydziałów nazwiemy sieć $N = (G, c, s_G, t_G)$, gdzie:

- $G = (V_G, E_G)$ jest grafem skierowanym takim, że:
  - $V_G = E \cup S \cup P \cup \{s_G, t_G\}$,
  - $E_G = \{(e, s) : \text{ability}(e, s) = 1, e \in E, s \in S\} \cup \{(s, p) : \text{need}(s, p) > 0, s \in S, p \in P\}$, tj. krawędziami połączeni są eksperci z ich opanowanymi umiejętnościami, oraz projekty z potrzebnymi do ich realizacji umiejętnościami.
- $c : E_G \to \mathbb{N}$ jest funkcją pojemności zdefiniowaną następująco:
  - jeżeli $e = s_G u, u \in E$, to $c(e) = 1$,

– jeżeli $e = es, e \in E, s \in S$, to $c(e) = \text{ability}(e, s) = 1$,

– jeżeli $e = sp, s \in S, p \in P$, to $c(e) = \text{need}(p, s)$,

– jeżeli $e = pt_G, p \in P$, to

$$c(e) = \sum_{sp \in E_G} c(up)$$

(tj. pojemność tej krawędzi jest równa sumie pojemności krawędzi wchodzących do wierzchołka p).

- s_G, t_G są wyróżnionymi wierzchołkami z V_G — kolejno źródłem i ujściem.

Lemat 1. *Dla dowolnego przyporządkowania* assign *mamy*

$$M(\text{assign}) \geq \left(\sum_{p \in P} \sum_{s \in S} \text{need}(p, s) \right) - |E|$$

Dowód. Mamy

$$\begin{aligned}
M(\text{assign}) &= \sum_{p \in P} \text{missing}(p, \text{assign}) = \\
&= \sum_{p \in P} \left(\sum_{s \in S} \text{need}(p, s) - \left|\text{assign}^{-1}[p, s]\right| \right) = \\
&= \left(\sum_{p \in P} \sum_{s \in S} \text{need}(p, s) \right) - \left(\sum_{p \in P} \sum_{s \in S} \left|\text{assign}^{-1}[p, s]\right| \right) = \\
&\geq \left(\sum_{p \in P} \sum_{s \in S} \text{need}(p, s) \right) - |E|
\end{aligned}$$

na mocy rozłączności przeciwobrazów funkcji assign. □

Twierdzenie 1. *Przepływ maksymalny w sieci przydziałów wyznacza przyporządkowanie o minimalnej możliwej wartości parametru M.*

Dowód. Aby dowieść to twierdzenie, wykażemy kolejno, że:

1. Dowolny przepływ w sieci przydziałów wyznacza poprawne (niekoniecznie maksymalne) przyporządkowanie ekspertów do zadań.

   Niech dany będzie pewien przepływ f w sieci przydziałów N. Przyporządkowanie ekspertów do projektów wyznaczamy w następujący sposób:

   (a) Pewnego eksperta $e \in E$ przypisujemy do umiejętności $s \in S$, jeżeli $f(e, s) = 1$.

   (b) Niech dana będzie pewna umiejętność $s \in S$. Oznaczmy zbiór ekspertów przypisanych do tej umiejętności w punkcie (a) jako E_s.
   Zbiór E_s dzielimy na rozłączne podzbiory $E_{s,p}$ takie, że $|E_{s,p}| = f(s, p)$.

   (c) Dla każdego z uzyskanych podzbiorów $E_{s,p}$, gdzie $s \in S, p \in P$ określamy wartość przyporządkowania assign_f jako

   $$(\forall e \in E_{s,p})\, \text{assign}_f(e) = (s, p)$$

Zauważmy następujące fakty:

- Rozważmy wierzchołek $e \in E$ odpowiadający pewnemu ekspertowi.

  Z definicji zbioru krawędzi sieci i funkcji przepustowości, do wierzchołka tego wchodzi dokładnie jedna krawędź o pojemności 1, a wychodzi z niego co najwyżej $|S|$ krawędzi o pojemności 1.

  Stąd w przepływie f tylko jedna z krawędzi wychodzących może mieć przepływ 1, a więc każdy ekspert może być przyporządkowany do co najwyżej jednej umiejętności.

- Rozważmy dowolny wierzchołek $s \in S$ odpowiadający pewnej umiejętności.

  Z własności przepływu mamy

  $$\sum_{us \in E_G} f(us) = \sum_{sv \in E_G} f(sv)$$

  Wiedząc, że wszystkie krawędzie wchodzące do s wychodzą ze zbioru E, oraz że wszystkie krawędzie wychodzące z s wchodzą do zbioru P, mamy

  $$\sum_{e \in E} f(es) = \sum_{p \in P} f(sp)$$

  Krawędzie o niezerowym przepływie wchodzące do e reprezentują ekspertów przydzielonych do danej umiejętności, zaś krawędzie o niezerowym przepływie wychodzące z e reprezentują zapotrzebowanie projektów na ekspertów z umiejętnością s.

  Ponieważ suma przepływów krawędzi wchodzących i wychodzących jest taka sama, każdego eksperta w dziedzinie umiejętności s można przypisać do dokładnie jednego projektu, a więc można wykonać punkt (b) konstrukcji.

- Rozważmy dowolne dwa wierzchołki $s \in S, p \in P$ takie, że $sp \in E_G$. Z definicji sieci mamy $c(s,p) = \mathrm{need}(s,p)$, a z konstrukcji rozwiązania wynika, że $f(s,p) = \left|\mathrm{assign}_f^{-1}[s,p]\right|$. Stąd na mocy definicji przepływu mamy

  $$\left|\mathrm{assign}_f^{-1}[s,p]\right| = f(s,p) \leq c(s,p) = \mathrm{need}(s,p)$$

- Rozważmy dowolny wierzchołek $p \in P$. Z definicji funkcji pojemności, jeśli wszystkie krawędzie wchodzące do p będą wysycone przepływem (tj. $f(e) = c(e)$), to przepływ ten można przekazać w całości do ujścia krawędzią pt_G, bo

  $$c(pt_G) = \sum_{up \in E_G} c(up)$$

  Stąd pojemność krawędzi pt_G nie ogranicza wartości maksymalnego przepływu.

Wyznaczone przyporządkowanie assign_f spełnia więc wszystkie warunki prawidłowego przyporządkowania ekspertów do projektów.

2. Skonstruowane z przepływu maksymalnego przyporządkowanie wyznacza minimalną wartość parametru M.

   Dowód nie wprost: Załóżmy, że przyporządkowanie $\mathrm{assign}_{\max}$ skonstruowane na podstawie przepływu maksymalnego $f_{\max}$ w sposób opisany w punkcie 1. można rozszerzyć tak, aby zmniejszyć wartość parametru M.

Na mocy lematu mamy

$$\left(\sum_{p\in P}\sum_{s\in S}\mathrm{need}(p,s)\right)-\left(\sum_{p\in P}\sum_{s\in S}\left|\mathrm{assign}_{\max}^{-1}[p,s]\right|\right)=M(\mathrm{assign}_{\max})>$$
$$>\left(\sum_{p\in P}\sum_{s\in S}\mathrm{need}(p,s)\right)-|E|$$

a więc istnieje co najmniej jeden ekspert $e \in E$, który nie jest przyporządkowany do jednego z projektów.

(a) Jeżeli $(\forall s \in S)\,\mathrm{ability}(e,s) = 0$, to eksperta nie można przyporządkować do żadnego projektu.

(b) Załóżmy, że istnieje umiejętność $s \in S$ taka, że $\mathrm{ability}(e,s) = 1$.

   i. Jeżeli $(\forall p \in P)\,\mathrm{need}(p,s) = \left|\mathrm{assign}_{\max}^{-1}[p,s]\right|$, to eksperta nie można przyporządkować do żadnego projektu.

   ii. Załóżmy, że istnieje projekt p taki, że $\mathrm{need}(p,s) > \left|\mathrm{assign}_{\max}^{-1}[p,s]\right|$. Wówczas w sieci rezydualnej dla przepływu $f_{\max}$ istnieje ścieżka rozszerzająca postaci $s_G espt_G$, wzdłuż której można zwiększyć przepływ o 1. Sprzeczność z maksymalnością przepływu $f_{\max}$.

$\square$