

ROZDZIAŁ 5

Zagadnienie komiwojażera

Przypomnijmy, że w zagadnieniu komiwojażera handlowiec ma odwiedzić n miejscowości (klientów) startując ze swojej miejscowości i wracając tam po skończonej podróży, przy czym podane są odległości (czas, względnie koszty podróży) c_{ij} między miejscowościami i -tą i j -tą, $i, j = 1, 2, \dots, n$. Problem polega na wyznaczeniu trasy o minimalnej łącznej długości (względnie łącznych kosztach podróży) przy czym każda miejscowość ma być odwiedzona dokładnie jeden raz.

Przedstawimy *zagadnienie komiwojażera* (ang. *traveling salesman problem* – TSP) używając języka teorii grafów jako problem optymalizacji dyskretnej. Dany jest graf nieskierowany $G = (V, E)$ i funkcja odległości (wag, kosztów) $c : E \rightarrow \mathbb{R}$. Należy wyznaczyć *cykl Hamiltona* (czyli cykl prosty rozpinający graf G) o najmniejszej długości (wadze, koszcie). Aby zadanie to posiadało rozwiązanie należy założyć, że graf G jest hamiltonowski, czyli, że istnieje jego podgraf będący cyklem Hamiltona. Założenie to spełnia oczywiście graf pełny. Możemy bez straty ogólności rozważać od razu założyć, że G jest grafem pełnym. W razie potrzeby możemy dany graf uzupełnić do grafu pełnego przypisując dostatecznie duże wagi dołączonym krawędziom (czytelnikowi pozostawimy problem jak duże powinny one być).

W dalszym ciągu, jeśli nie będzie powiedziane inaczej, będziemy zakładać, że funkcja odległości spełnia tzw. warunek trójkąta:

$$c_{uw} \leq c_{uv} + c_{vw}$$

dla dowolnych $u, v, w \in V$. Ponieważ graf G jest nieskierowany, więc oczywiście $c_{uv} = c_{vu}$ dla dowolnych wierzchołków $u, v \in V$. Wówczas rozpatrywany problem nazywa się *symetrycznym zagadnieniem komiwojażera*. Zagadnienie komiwojażera można również rozpatrywać na grafach skierowanych, szczególnie w sytuacji, gdy $c_{uv} \neq c_{vu}$ dla pewnych $u, v \in V$. Mówimy wówczas o *niesymetrycznym zagadnieniu komiwojażera*.

Problem komiwojażera jest jednym z najbardziej znanych problemów optymalizacji dyskretnej. Jest to jednocześnie jeden z najtrudniejszych problemów tej optymalizacji. W istocie jest to tzw. problem \mathcal{NP} -trudny. W konsekwencji, jeśli $\mathcal{P} \neq \mathcal{NP}$ (co do tej pory nie zostało udowodnione), to nie istnieje metoda wyznaczająca dokładne rozwiązanie zagadnienia komiwojażera w czasie wielomianowym. Natomiast istnieją algorytmy, które w rozsądnym czasie rozwiązują w sposób przybliżony to zagadnienie dla dużych n z dokładnością do ułamka procenta.

Podchodząc do problemu naiwnie można powiedzieć, że wystarczy porównać długości wszystkich cykli Hamiltona grafu pełnego i wybrać najkrótszy z nich. Trudność polega jednak na tym, że cykli tych jest w sumie $(n-1)!/2$ i już dla $n = 20$ czas obliczeń wyniósłby blisko 2 lata zakładając optymistycznie, że w ciągu jednej sekundy jesteśmy w stanie przeszukać miliard

cykli Hamiltona. Tymczasem dla największego rozwiązanego do tej pory w sposób przybliżony problemu komiwojażera n jest większe niż 80 000.

Widzimy więc, że musimy znaleźć bardziej efektywne metody wyznaczania rozwiązań zagadnienia komiwojażera. Metody te dzielą się na heurystyczne (przybliżone) i metody podziału i ograniczeń (dokładne).

Metody heurystyczne nie wyznaczają wprawdzie rozwiązań optymalnych w ogólnym przypadku, mają jednak tę zaletę, że najczęściej mają złożoność wielomianową. Dają one ponadto możliwość oszacowania z góry długości najkrótszego cyklu Hamiltona. Natomiast metody podziału i ograniczeń wyznaczają wprawdzie rozwiązania optymalne redukując jednocześnie omawianą wyżej nieefektywność pełnego przeglądu, ich złożoność jest jednak wykładnicza. W konsekwencji musimy często przerwać ich działanie zanim wyznaczone zostanie rozwiązanie optymalne.

5.1 Rozwiązania heurystyczne

Przedstawimy teraz kilka algorytmów heurystycznych rozwiązania zagadnienia komiwojażera mających *własność zachłanności*.

5.1.1 Algorytm najbliższego sąsiada

Opis tego *algorytmu najbliższego sąsiada* jest bardzo prosty. Jako punkt startowy wybiera się jakikolwiek wierzchołek grafu V . W każdej iteracji przechodzi się z wybranego właśnie wierzchołka do najbliższego spośród wierzchołków do tej pory nie wybranych. Jeśli wybrane zostały już wszystkie wierzchołki grafu G należy wrócić do wierzchołka startowego. Wybrane w ten sposób kolejne wierzchołki utworzą oczywiście cykl Hamiltona.

Przykład 5.1.1 Dla podanej tabeli odległości wyznaczyć rozwiązanie zagadnienia komiwojażera dla 16 miast wojewódzkich w Polsce otrzymane metodą najbliższego sąsiada. Jaka jest długość tego cyklu?

	Warszawa	Kraków	Wrocław	Poznań	Gdańsk	Szczecin	Łódź	Lublin	Białystok	Katowice	Rzeszów	Kielce	Olsztyn	Bydgoszcz	Opole	Zielona Góra
Warszawa	-	295	343	310	342	573	135	170	195	285	297	177	214	286	350	452
Kraków	295	-	268	393	587	632	224	272	476	79	162	118	475	454	190	422
Wrocław	343	268	-	164	486	371	220	460	504	197	370	243	406	290	87	189
Poznań	310	393	164	-	308	265	211	413	497	338	463	305	279	130	237	182
Gdańsk	342	587	486	308	-	364	334	482	421	502	574	426	171	161	491	443
Szczecin	573	632	371	265	364	-	453	675	685	548	777	617	386	242	459	198
Łódź	135	224	220	211	334	453	-	250	325	199	334	139	257	200	217	328
Lublin	170	272	460	413	482	675	250	-	235	265	163	159	348	454	414	516
Białystok	195	476	504	497	421	685	325	235	-	401	467	347	239	434	512	603
Katowice	285	79	197	338	502	548	199	265	401	-	184	144	441	394	147	357
Rzeszów	297	162	370	463	574	777	334	163	467	184	-	169	506	545	332	483
Kielce	177	118	243	305	426	617	139	159	347	144	169	-	338	353	226	420
Olsztyn	214	475	406	279	171	386	257	348	239	441	506	338	-	213	378	402
Bydgoszcz	286	454	290	130	161	242	200	454	434	394	545	353	213	-	270	302
Opole	350	190	87	237	491	459	217	414	512	147	332	226	378	270	-	263
Zielona Góra	452	422	189	182	443	198	328	516	603	357	483	420	402	302	263	-



Trasa wyznaczona tą metodą, to Warszawa \rightarrow Łódź \rightarrow Kielce \rightarrow Katowice \rightarrow Kraków \rightarrow Rzeszów \rightarrow Opole \rightarrow Wrocław \rightarrow Poznań \rightarrow Bydgoszcz \rightarrow Gdańsk \rightarrow Olsztyn \rightarrow Białystok \rightarrow Lublin \rightarrow Zielona Góra \rightarrow Szczecin \rightarrow Warszawa, a jej długość wynosi 3576 km.

Widzimy, że opisana procedura posiada – podobnie jak algorytm Kruskala czy Prima – własność zachłanności, tzn. jest lokalnie optymalna. W przeciwieństwie jednak do tych dwóch algorytmów nie wyznacza ona rozwiązania optymalnego. Przyczynę tego stanu rzeczy nietrudno wyjaśnić. Wybierając najbliższego sąsiada omijamy wierzchołki, które leżą również niedaleko

wierzchołka właśnie wybranego, a które musimy odwiedzić później przy znacznie wyższym koszcie. Można pokazać, że długość cyklu Hamiltona wyznaczonego przez opisany algorytm nie przekracza $(\frac{1}{2}\lceil \log_2 n \rceil + \frac{1}{2})c^*$, gdzie c^* jest wartością optymalną (długością najkrótszego cyklu Hamiltona). Ponadto efektywności tej nie da się poprawić w sposób istotny. Istnieje bowiem przykład, że długość cyklu Hamiltona wyznaczonego przez algorytm najbliższego sąsiada wynosi $\lceil \frac{1}{3} \log_2(n+1) + \frac{4}{9} \rceil c^*$. Ta niekorzystna własność algorytmu powoduje, że ma on niewielkie zastosowanie.

5.1.2 Algorytmy dołączania

W *algorytmach dołączania* w k -tej iteracji wyznacza się cykl złożony z $k+1$ krawędzi grafu G poprzez dołączenie do bieżącego cyklu pewnego nowego wierzchołka i zamianę jednej krawędzi bieżącego cyklu na dwie nowe. Z bieżącego cyklu usuwa się krawędź, dla której zamiana na dwie nowe prowadzi do najmniejszego wydłużenia powstałego cyklu. Jako punkt startowy wybiera się dowolny wierzchołek grafu G i cykl pusty (bez krawędzi) – w konsekwencji pierwsza iteracja różni się tym od pozostałych, że nie usuwamy krawędzi początkowego cyklu a jedynie dołączamy dwie nowe. Algorytmy dołączania różnią się kryteriami wyboru dołączanego wierzchołka.

Algorytm swobodnego dołączania

W *algorytmie swobodnego dołączania* nowy wierzchołek wybiera się dowolnie, na przykład losowo.

Algorytm najbliższego dołączania

W *algorytmie najbliższego dołączania* nowy wierzchołek jest wierzchołkiem leżącym najbliższej bieżącego cyklu. Najczęściej jako początkowe dwa wierzchołki bierze się wierzchołki sąsiednie połączone łukiem którego waga jest najmniejsza. Dołączany wierzchołek dołącza się do bieżącego cyklu w miejsce dające najmniejszy przyrost długości.

Ćwiczenie 5.1.2 Wyznaczyć rozwiązanie zagadnienia komiwojażera dla 16 miast wojewódzkich w Polsce otrzymane metodą najbliższego dołączania. Jaka jest długość tego cyklu?

Najkrótsza trasa otrzymana tą metodą to Kraków → Kielce → Lublin → Białystok → Warszawa → Łódź → Bydgoszcz → Gdańsk → Olsztyn → Poznań → Szczecin → Zielona Góra → Wrocław → Opole → Rzeszów → Katowice → Kraków o długości 2748 km

Algorytm najdalszego dołączania

W *algorytmie najdalszego dołączania* nowy wierzchołek jest wierzchołkiem leżącym najdalej od bieżącego cyklu. Najczęściej jako początkowe dwa wierzchołki bierze się wierzchołki sąsiednie połączone łukiem którego waga jest największa. Podobnie jak w przypadku algorytmu najbliższego dołączania, nowy wierzchołek dołącza się do bieżącego cyklu w miejsce dające najmniejszy przyrost długości.

W praktyce okazuje się, że wśród algorytmów dołączania algorytm ten prowadzi najczęściej do najlepszych rezultatów (choć nie ma takiej gwarancji) – długość otrzymanego cyklu Hamiltona najczęściej najlepiej ogranicza z góry długość najkrótszego cyklu Hamiltona wśród wszystkich algorytmów dołączania.

Ćwiczenie 5.1.3 Wyznaczyć rozwiązanie zagadnienia komiwojażera dla 16 miast wojewódzkich w Polsce otrzymane metodą najdalszego dołączania. Jaka jest długość tego cyklu?

Pierwsze 5 iteracji tego algorytmu daje cykl Szczecin \rightarrow Białystok \rightarrow Gdańsk \rightarrow Wrocław \rightarrow Rzeszów \rightarrow Szczecin. W każdej z kolejnych iteracji dołączamy miasto leżące najdalej od bieżącego cyklu. Ostateczna trasa (metodą najdalszego dołączania): Sz \rightarrow Bi \rightarrow Lu \rightarrow W \rightarrow Ki \rightarrow K \rightarrow Ld \rightarrow Gd \rightarrow By \rightarrow P \rightarrow Zg \rightarrow Ol \rightarrow Wr \rightarrow Op \rightarrow Ka \rightarrow Rz \rightarrow Sz. Długość trasy: około 4026 km.

Algorytm najtańszego dołączania

W algorytmie najtańszego dołączania nowy wierzchołek wybierany jest w ten sposób, aby opisana zamiana krawędzi prowadziła do jak najmniejszego wydłużenia powstałego cyklu. Algorytm ten, mimo że pojedyncza iteracja jest najkosztowniejsza wśród opisanych algorytmów dołączania – prowadzi w praktyce do dłuższych cykli Hamiltona aniżeli algorytm najdalszego dołączania.

5.1.3 Algorytmy poprawy

W algorytmach poprawy punktem startowym jest dowolny cykl Hamiltona. W każdej iteracji dokonuje się wymiany pewnych krawędzi na nowe w taki sposób aby powstały graf był nadal cyklem i aby jego długość uległa skróceniu w stosunku do długości bieżącego cyklu Hamiltona.

Algorytm 2-optymalny

W algorytmie 2-optymalnym usuwa się z bieżącego cyklu Hamiltona T dwie krawędzie i w ich miejsce wprowadza się dwie nowe dbając jednocześnie o to, aby powstały graf był nadal cyklem. Aby to było możliwe usuwane krawędzie nie mogą być związane ze wspólnym wierzchołkiem. Istnieje $n(n-3)/2$ różnych sposobów usunięcia takich dwóch krawędzi. Zauważmy jednak, że istnieje tylko jeden sposób wprowadzenia dwóch nowych krawędzi w miejsce usuniętych prowadzący do nowego cyklu Hamiltona T' . Jeśli wymiana ta prowadzi do skrócenia powstałego cyklu T' , to zastępuje się bieżący cykl T nowym cyklem T' . Opisana procedura zatrzymuje się, jeśli dla bieżącego cyklu T żadna z $n(n-3)/2$ wymian nie prowadzi do skrócenia cyklu Hamiltona. Powstały w ten sposób cykl Hamiltona nazywa się *2-optymalny*. Nie musi być on jednak optymalny. Ponadto algorytm ten ma jedną wadę – ma złożoność wykładniczą.

Algorytm 3-optymalny

W algorytmie 3-optymalnym postępuje się podobnie jak w algorytmie 2-optymalnym, przy czym z bieżącego cyklu Hamiltona usuwa się i w ich miejsce wprowadza 3 nowe krawędzie. Okazuje się, że stosowanie tego algorytmu prowadzi w praktyce do znacznie krótszych cykli Hamiltona w stosunku do cykli generowanych przez algorytm 2-optymalny.

Opisaną procedurę można rozszerzyć w sposób naturalny do tzw. algorytmu k -optimalnego. Dla $k > 3$ jest on jednak rzadko stosowany z uwagi na rosnącą bardzo szybko ze wzrostem k złożoność obliczeniową pojedynczej iteracji i ze względu na (najczęściej) małe w stosunku do algorytmu 3-optimalnego zmniejszenie długości cyklu Hamiltona.

5.2 Metody podziału i ograniczeń dla zagadnienia komiwojażera

Metody podziału i ograniczeń przedstawione w ustępie 2.4 dotyczyły zadania maksymalizacji, natomiast zagadnienie komiwojażera jest zadaniem minimalizacji. Stosując je musimy więc zamienić odpowiednio ograniczenie górne na ograniczenie dolne i odwrotnie. W dalszej części zdefiniujemy dokładniej rodzinę, o której była mowa w ustępie 2.4. Teraz powiemy tylko, że będzie ona pewną rodziną \mathcal{D} , której elementami są pewne zbiory $D_{A,B}$ podgrafów grafu G (dokładniej cykli Hamiltona), dla których narzucono dodatkowe ograniczenia postaci: łuki z pewnego zbioru $A \subset E$ należą do każdego grafu ze zbioru $D_{A,B}$ zaś łuki z pewnego zbioru $B \subset E$ nie należą do żadnego grafu z tego zbioru. Oczywiście zbiory A i B muszą być rozłączne. W trakcie realizacji jakiegokolwiek metody podziału i ograniczeń zbiory A i B będą odpowiednio powiększane tak, aby coraz bardziej zawęzić rodzinę przeszukiwanych grafów (aż dojdzie się do sytuacji, w której $A \cup B = E$) i tak, aby w końcu doprowadzić do rozwiązania optymalnego. Po drodze wyznaczane będą ograniczenia górne $\bar{\varphi}$, które pozwolą usunąć pewne rodziny grafów z rodziny przeszukiwanych grafów. Ponieważ należy się liczyć z tym, że rozwiązanie optymalne nie będzie wyznaczone w czasie wielomianowym, trzeba odpowiednio wybrać $\varepsilon > 0$, tak aby rozwiązanie otrzymane po zatrzymaniu algorytmu było zadowalające. Należy również narzucić ograniczenie górne na liczbę wykonanych iteracji algorytmu.

Przejdziemy teraz do dokładniejszego przedstawienia metody podziału i ograniczeń dla zagadnienia komiwojażera.

Oznaczmy

$$\mathcal{D} = \{D_{A,B} : A, B \subset E, A \cap B = \emptyset\}.$$

W przedstawionej poniżej metodzie podziału i ograniczeń dla zagadnienia komiwojażera nie podajemy konkretnej postaci funkcji ograniczenia dolnego $\varphi : \mathcal{D} \rightarrow \mathbb{R}$. Zakładamy jedynie, że funkcja ta spełnia warunki:

$$1^0 G_1 \subseteq G_2 \in \mathcal{D} \implies \varphi(G_1) \geq \varphi(G_2),$$

$$2^0 \text{ Jeśli } H \text{ jest cyklem Hamiltona, to } \varphi(\{H\}) = c(H).$$

Zauważmy, że zbiór złożony z jednego cyklu Hamiltona H jest elementem rodziny zbiorów \mathcal{D} . Wystarczy za A wziąć zbiór wszystkich krawędzi grafu H , zaś za B – zbiór pusty.

Metoda podziału i ograniczeń dla TSPna grafie skierowanym

Wejście. Graf skierowany $G = (V, E)$ posiadający przynajmniej jeden cykl Hamiltona, funkcja wag $c : E \rightarrow \mathbb{R}_+$, tolerancja optymalności $\varepsilon \geq 0$

Wyjście. Rozwiązanie ε -optymalne zagadnienia komiwojażera

Krok 0 (*inicjalizacja*)

(a) Położyć $\mathcal{G} = \{D_{\emptyset, \emptyset}\}$.

(b) Wyznaczyć ograniczenie dolne $\varphi(D_{\emptyset, \emptyset})$ lub położyć $\varphi(D_{\emptyset, \emptyset}) = 0$.

(c) Wybrać jakikolwiek cykl Hamiltona $\bar{H} \in D_{\emptyset, \emptyset}$ i położyć $\bar{\varphi} = c(\bar{H})$ lub położyć $\bar{\varphi} = +\infty$.

Krok 1 (*ograniczenie*)

- (a) Wyznaczyć $\underline{\varphi} = \min_{D_{A,B} \in \mathcal{G}} \varphi(D_{A,B})$.
- (b) Jeśli stwierdzi się, że $\varphi(D_{A,B}) = c(H)$ dla pewnego $H \in D_{A,B}$ (np. $D_{A,B}$ składa się z jednego cyklu Hamiltona), to usunąć $D_{A,B}$ z rodziny \mathcal{G} (w zbiorze $D_{A,B}$ nie ma cyklu Hamiltona krótszego niż długość cyklu H). Jeśli dodatkowo $c(H) < \bar{\varphi}$, to położyć $\bar{\varphi} = c(H)$ i $\bar{H} = H$ ($c(H)$ jest nowym ograniczeniem górnym i H jest kandydatem na rozwiązanie optymalne).
- (c) Usunąć z rodziny \mathcal{G} te podzbiory $D_{A,B}$, dla których $\varphi(D_{A,B}) \geq \bar{\varphi}$ (podzbiory te nie będą więcej rozważane, gdyż nie zawierają one rozwiązania optymalnego).
- (d) Jeśli $\mathcal{G} = \emptyset$, to \bar{H} jest rozwiązaniem optymalnym zadania TSP – algorytm zatrzymuje się.
- (e) Jeśli $\bar{\varphi} - \underline{\varphi} \leq \varepsilon$, to \bar{H} jest rozwiązaniem ε -optymalnym zadania TSP – algorytm zatrzymuje się.

Krok 2 (podział)

- (a) Wybrać podzbiór $D_{A,B} \in \mathcal{G}$ oraz łuk $e \notin A \cup B$ i rozłożyć ten podzbiór na dwa rozłączne podzbiory $D_{A \cup \{e\}, B}$ i $D_{A, B \cup \{e\}}$.
- (b) Położyć $\mathcal{G} = \mathcal{G} \cup \{D_{A \cup \{e\}, B} \cup D_{A, B \cup \{e\}}\} \setminus \{D_{A,B}\}$.
- (c) Wyznaczyć $\varphi(D_{A \cup \{e\}, B})$ i $\varphi(D_{A, B \cup \{e\}})$.
- (d) Przejsć do kroku 1.

W dalszej części podamy jak można zdefiniować funkcję ograniczenia dolnego w metodzie podziału i ograniczeń dla TSP. Bez szkody dla ogólności rozważań przyjmujemy, że graf G jest pełny.

Niech (V, E, c) będzie siecią skierowaną pełną i niech T^* będzie najkrótszym cyklem Hamiltona w tej sieci. Oznaczmy jego długość w tej sieci przez c^* . Metody podziału i ograniczeń, które przedstawione będą w tym ustępie korzystają z własności wyrażonej w następującym lemacie.

Lemat 5.2.1 T^* jest najkrótszym cyklem Hamiltona w sieci (V, E, \bar{c}) , gdzie $\bar{c}_{ij} = c_{ij} - y_i - w_j$, $y = (y_1, \dots, y_n)$, $w = (w_1, \dots, w_n) \in \mathbb{R}^n$. Ponadto długość cyklu T^* w tej sieci wynosi $\bar{c}^* = c^* - e^\top y - e^\top w$, gdzie $e = (1, \dots, 1)$.

Dowód. W zmodyfikowanej sieci (V, E, \bar{c}) od długości poszczególnych łuków c_{ij} odjęto stałe wielkości y_i dla początków i w_j dla końców tych łuków. Zauważmy, że dla dowolnego cyklu Hamiltona T każdy węzeł jest jednocześnie początkiem dokładnie jednego łuku i końcem dokładnie jednego łuku. Wobec tego po modyfikacji długość dowolnego cyklu Hamiltona zmniejszy się o stałą wielkość $\sum_{i=1}^n y_i + \sum_{j=1}^n w_j = e^\top y + e^\top w$. Nowa funkcja celu (długość cyklu Hamiltona w zmodyfikowanej sieci) będzie się więc różnić od wyjściowej funkcji celu (dla wyjściowej sieci) o stałą wielkość. Jest oczywiste, że taka zmiana nie wpłynie na dowolne rozwiązanie optymalne (najkrótszy cykl Hamiltona).

Uwaga 5.2.2 Lemat 5.2.1 jest prawdziwy również dla sieci nieskierowanej (V, E, c) . W tym przypadku zmodyfikowaną funkcję odległości $\bar{c} : E \rightarrow \mathbb{R}$ określa się następująco: $\bar{c}_{ij} = c_{ij} - y_i - y_j$, $i, j = 1, \dots, n$, gdzie $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. Wówczas długość cyklu T^* w sieci (V, E, \bar{c}) wynosi $\bar{c}^* = c^* - 2e^\top y$.

Uwaga 5.2.3 Zmodyfikowana funkcja odległości \bar{c} nie musi spełniać warunku trójkąta, nawet jeśli wyjściowa funkcja c go spełnia.

5.2.1 Metoda redukcji

Przedstawimy teraz pewną metodę podziału i ograniczeń, zwaną *metodą redukcji*, dla zagadnienia komiwojażera na sieci skierowanej (V, E, c) . Skorzystamy przy tym z własności wyrażonej w lemacie 5.2.1 dobierając odpowiednio wektory $y, w \in \mathbb{R}^n$. Zauważmy, że jeśli wektory $y, w \in \mathbb{R}^n$ zostały dobrane tak, aby $\bar{c}_{ij} = c_{ij} - y_i - w_j \geq 0, i, j = 1, \dots, n$, to oczywiście długość cyklu T^* w sieci (V, E, \bar{c}) jest nieujemna. Wobec tego, dla takiego doboru $y, w \in \mathbb{R}^n$, z lematu 5.2.1 otrzymujemy dolne ograniczenie długości cyklu T^* w wyjściowej sieci (V, E, c)

$$c^* \geq e^\top y + e^\top w.$$

Możemy spodziewać się dobrego ograniczenia dolnego wartości c^* przyjmując $y_i = \min_j c_{ij}, i = 1, \dots, n$ i $w_j = \min_i (c_{ij} - y_i), j = 1, \dots, n$, albo $w_j = \min_i c_{ij}, j = 1, \dots, n$ i $y_i = \min_j (c_{ij} - w_j), i = 1, \dots, n$, gdyż wówczas w każdym wierszu i w każdej kolumnie zmodyfikowanej macierzy kosztów \bar{c} znajdzie się przynajmniej jedno zero. Opisana metoda oszacowania z dołu długości najkrótszego cyklu Hamiltona pozwoli nam skonstruować funkcję ograniczenia dolnego zdefiniowaną na określonej wyżej rodzinie \mathcal{D} .

Przykład 5.2.4 Dla funkcji (macierzy) odległości

$$c = \begin{bmatrix} \infty & 3 & 5 & 22 & 16 \\ 13 & \infty & 9 & 11 & 5 \\ 6 & 8 & \infty & 3 & 12 \\ 8 & 9 & 14 & \infty & 6 \\ 12 & 4 & 7 & 9 & \infty \end{bmatrix}$$

i dla $y_i = \min_j c_{ij}, i = 1, \dots, 5$ oraz $w_j = \min_i (c_{ij} - y_i), j = 1, \dots, 5$ zmodyfikowaną funkcję (macierz) odległości otrzymujemy z powyższej macierzy poprzez odjęcie od kolejnych wierszy wielkości 3, 5, 3, 6, 4, a następnie odjęcie od kolejnych kolumn wielkości 2, 0, 2, 0, 0. Ma ona postać

$$\bar{c} = \begin{bmatrix} \infty & 0 & 0 & 19 & 13 \\ 6 & \infty & 2 & 6 & 0 \\ 1 & 5 & \infty & 0 & 9 \\ 0 & 3 & 6 & \infty & 0 \\ 6 & 0 & 1 & 5 & \infty \end{bmatrix}$$

Widzimy, że dla funkcji \bar{c} dowolny cykl Hamiltona ma nieujemną długość, czyli, zgodnie z tym co było powiedziane wyżej, mamy

$$c^* \geq 3 + 5 + 3 + 6 + 4 + 2 + 2 = 25.$$

Opisana wyżej metoda wyznaczenia ograniczenia dolnego długości najkrótszego cyklu Hamiltona ma jeszcze jedną zaletę, którą wykorzystamy przy konstrukcji metody redukcji. Pojawienie się zer w wierszach i kolumnach macierzy \bar{c} pozwoli nam na wykonanie odpowiedniego podziału zbioru rozwiązań dopuszczalnych (cykli Hamiltona) na dwa podzbiory, z których do pierwszego należą wszystkie cykle Hamiltona zawierające określony łuk, a do drugiego – nie zawierające go. Łuk ten będzie określony przez odpowiedni element zerowy macierzy \bar{c} . Pozwoli to zredukować wymiar rozpatrywanego zagadnienia o jeden, gdyż od tego momentu będziemy mogli traktować

końce tego łuku jako jeden wierzchołek. Łuk ten należy bowiem do wszystkich cykli Hamiltona należących do pierwszego podzbioru, a długość tego łuku w zmodyfikowanej sieci jest zerowa. Ponieważ elementów zerowych macierzy \bar{c} jest co najmniej tyle ile jest wierzchołków grafu, więc mamy swobodę przy wyborze łuku, względem którego dokonamy podziału. Wybierzemy ten łuk, którego wybór pozwoli na największy wzrost dolnego ograniczenia długości cykli Hamiltona nie zawierających wybranego łuku. Przy takim wyborze będzie większa szansa na ewentualne usunięcie rozwiązań nie zawierających tego łuku, jeśli wśród rozwiązań łuk ten zawierających znajdziemy najkrótszy cykl Hamiltona.

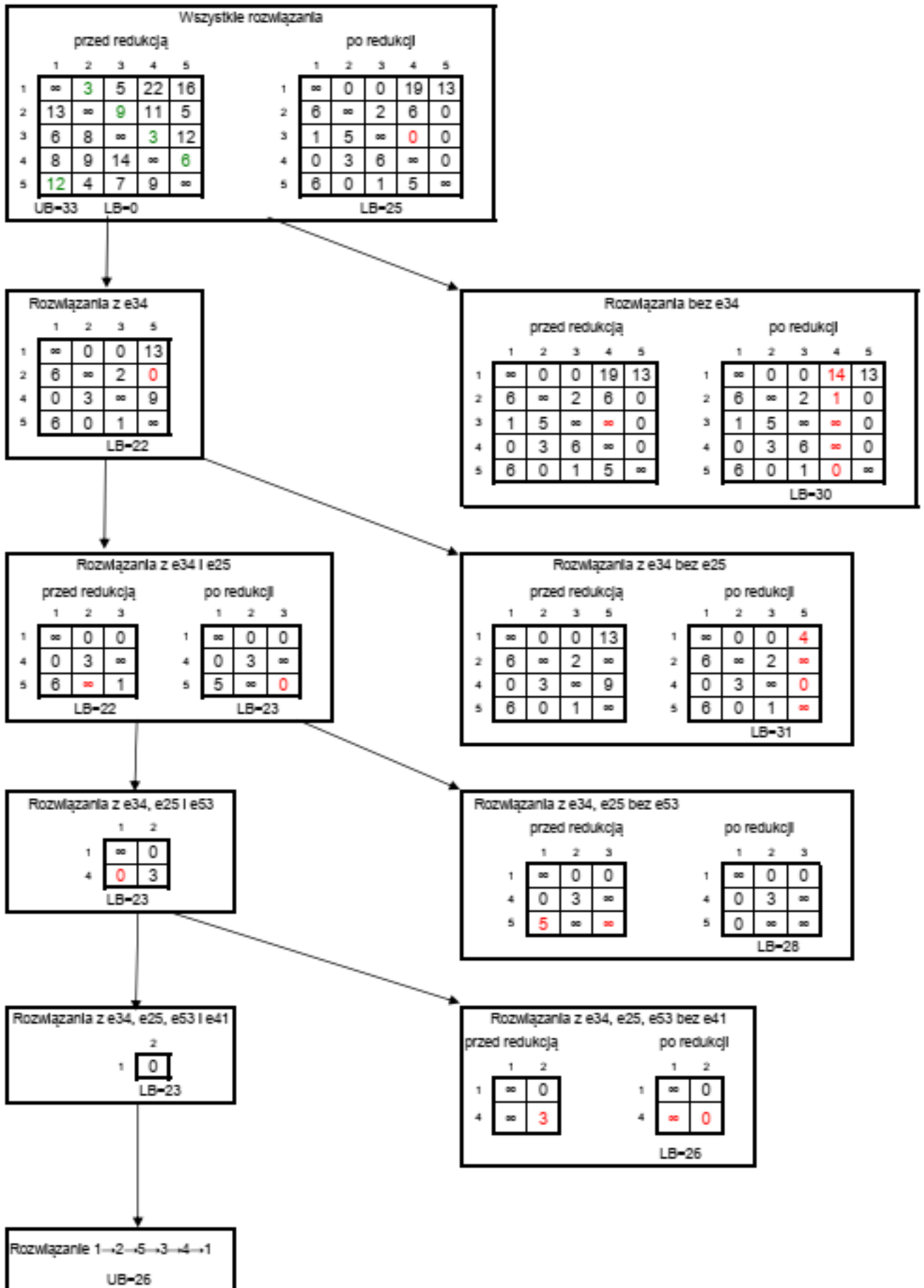
Przykład 5.2.5 Wróćmy do przedstawionego wyżej przykładu, na którym przedstawimy omawianą metodę podziału i ograniczeń łącznie z metodą redukcji. Zauważmy, że wybór w macierzy \bar{c} elementu $\bar{c}_{34} = 0$ (lub równoważnie łuku e_{34}) pozwoli na kolejny wzrost o 5 ograniczenia dolnego długości wszystkich cykli Hamiltona nie zawierających łuku e_{34} (w macierzy \bar{c} możemy wstawić $\bar{c}_{34} = \infty$ przy rozpatrywaniu cykli Hamiltona nie zawierających łuku e_{34}). Przy wyborze dowolnej innej krawędzi wzrost ten będzie mniejszy. Wiemy więc, że dowolny cykl Hamiltona nie zawierający łuku e_{34} ma długość co najmniej $25 + 5 = 30$. Z drugiej strony rozpatrywanie wszystkich rozwiązań zawierających łuk e_{34} jest równoważne rozpatrywaniu macierzy \bar{c} , w której usunięto 3. wiersz i 4. kolumnę i wstawiono ∞ w miejsce \bar{c}_{43} (możemy to uczynić, gdyż żaden cykl Hamiltona zawierający łuk e_{34} nie może zawierać łuku e_{43}).

W podobny sposób postępujemy ze zbiorem cykli Hamiltona zawierającymi łuk e_{34} . W tym zbiorze rozwiązań nie da się dokonać redukcji, ponieważ w każdym wierszu i w każdej kolumnie macierzy kosztów jest przynajmniej jedno zero. Natomiast zbiorze cykli Hamiltona nie zawierających łuku e_{34} możemy dokonać redukcji zwiększając jednocześnie o 5 ograniczenie dolne, bowiem w kolumnie 4. najmniejszy element wynosi 5.

Następnie zbiór cykli Hamiltona zawierających łuk e_{34} rozkładamy na dwa zbiory: cykle Hamiltona zawierające łuk e_{25} i na cykle Hamiltona nie zawierające tego łuku.

Dalej postępujemy podobnie jak opisaliśmy to wyżej. Proces ten możemy przedstawić w poniższym diagramie. Po lewej stronie otrzymaliśmy w końcu zbiór złożony z jednego cyklu Hamiltona $1 \rightarrow 2 \rightarrow 5 \rightarrow 3 \rightarrow 4 \rightarrow 1$ o długości 26. Wielkość tę możemy wobec tego traktować jako nowe ograniczenie górne długości najkrótszego cyklu Hamiltona. Teraz możemy usunąć wszystkie zbiory cykli Hamiltona, dla których ograniczenie dolne jest większe od 26, bowiem tam nie będzie cykli Hamiltona krótszych niż 26. Został więc nam do zbadania jeden zbiór cykli Hamiltona z e_{34} i e_{25} i e_{53} ale bez e_{41} , dla którego ograniczenie dolne wynosi 26. Ten zbiór też można właściwie usunąć, bo możemy w nim ewentualnie znaleźć cykl Hamiltona o długości 26, ale cykl o tej długości już przecież znaleźliśmy. Widzimy też zaletę wyboru łuku, według którego dokonujemy podziału. Wybór ten powoduje największy możliwy wzrost dolnego ograniczenia w zbiorze cykli Hamiltona nie zawierających tego łuku (po prawej stronie poniższego diagramu). Dzięki temu, możemy usunąć wiele zbiorów cykli Hamiltona o ograniczeniu dolnym większym niż wyznaczone ograniczenie górne najkrótszego cyklu Hamiltona po lewej stronie diagramu. W naszym przykładzie po znalezieniu cyklu Hamiltona o długości 26 mogliśmy usunąć wszystkie

zbiory rozwiązań po prawej stronie.



Ostatecznym rozwiązaniem zagadnienia komiwojażera otrzymanym metodą redukcji jest cykl Hamiltona $1 \rightarrow 2 \rightarrow 5 \rightarrow 3 \rightarrow 4 \rightarrow 1$ o długości 26.

5.2.2 1-drzewa ograniczające

Podamy teraz inną metodę wyznaczania ograniczenia dolnego długości najkrótszego cyklu Hamiltona w grafie nieskierowanym (pełnym) $G = (V, E)$ z określoną funkcją długości $c : E \rightarrow \mathbb{R}$.

Definicja 5.2.6 Niech $v \in V$, $H = (V \setminus \{v\}, F)$ będzie drzewem rozpinającym graf $G \setminus \{v\}$ i niech f, g będą dwiema różnymi krawędziami związanymi z wierzchołkiem v . Graf $H_1 = (V, F \cup \{f, g\})$ nazywamy 1-drzewem.

Uwaga 5.2.7

- (a) Dowolny cykl Hamiltona jest 1-drzewem.
- (b) Długość dowolnego, a więc i najkrótszego cyklu Hamiltona jest ograniczona z dołu przez długość najkrótszego 1-drzewa.
- (c) Niech $v \in V$. Długość najkrótszego 1-drzewa można obliczyć wyznaczając najpierw długość najkrótszego drzewa rozpinającego graf $G \setminus \{v\}$ (na przykład metodą Kruskala lub metodą Prima) i dodając do niej długości dwóch najkrótszych krawędzi $f, g \in E$ związanych z wierzchołkiem v . Dokładniej, niech $T^* = (V \setminus \{v\}, F)$ będzie najkrótszym drzewem rozpinającym graf $G \setminus \{v\}$ i niech f, g będą dwiema (różnymi) najkrótszymi krawędziami, związanymi z wierzchołkiem v :

$$f \in \underset{e \in \delta(\{v\})}{\text{Argmin}} c(e), \quad g \in \underset{e \in \delta(\{v\}) \setminus \{f\}}{\text{Argmin}} c(e).$$

Wówczas $(V, F \cup \{f, g\})$ jest najkrótszym 1-drzewem. W konsekwencji wielkość $c(F \cup \{f, g\}) = c(F) + c(f) + c(g)$ jest dolnym ograniczeniem długości najkrótszego cyklu Hamiltona w sieci (V, E, c) czyli

$$c^* \geq c(F) + c(f) + c(g),$$

gdzie c^* oznacza długość najkrótszego cyklu Hamiltona w tej sieci.

Zgodnie z powyższą uwagą zastosowanie 1-drzew pozwala wykorzystanie ich długości jako dolnego ograniczenia w metodzie podziału i ograniczeń dla zagadnienia komiwojażera.

5.2.3 Metoda Helda–Karpa

1-drzewa nie muszą stanowić wystarczająco dobrego ograniczenia dolnego długości cyklu Hamiltona. Zilustrujemy to następującym przykładem

Przykład 5.2.8 Rozważmy graf nieskierowany (V, E) dla $V = \{v_1, \dots, v_5\}$, $E = \{e_{12}, e_{13}, e_{15}, e_{23}, e_{24}, e_{25}, e_{34}, e_{35}, e_{45}\}$ i funkcję odległości określoną przez macierz

$$c = \begin{bmatrix} \infty & 0 & 0 & \infty & 1 \\ 0 & \infty & 10 & 0 & 10 \\ 0 & 10 & \infty & 0 & 10 \\ \infty & 0 & 0 & \infty & 0 \\ 1 & 10 & 10 & 0 & \infty \end{bmatrix}.$$

Jeśli wybrać $v = v_1$, to najkrótsze 1-drzewo ma długość 0. Natomiast najkrótszy cykl Hamiltona ma długość 10.

W celu wyznaczenia lepszego ograniczenia dolnego długości najkrótszego cyklu Hamiltona w sieci nieskierowanej (V, E, c) możemy skorzystać z lematu 5.2.1 i z uwagi 5.2.2. Jeśli w powyższym przykładzie długości wszystkich krawędzi związanych z wierzchołkiem v_4 zwiększymy o 10, to zgodnie z uwagą 5.2.2 najkrótszy cykl Hamiltona jest taki sam jak w wyjściowym grafie, lecz jego długość w zmodyfikowanej sieci wynosi 30. Zwróćmy jednak uwagę na to, że najkrótsze 1-drzewo ma w zmodyfikowanej sieci również długość 30. Oznacza to, że w wyjściowej sieci nie ma cykli Hamiltona o długości mniejszej niż 10. Otrzymaliśmy w ten sposób lepsze niż poprzednio ograniczenie dolne długości najkrótszego cyklu Hamiltona.

W ogólnym przypadku rozumowanie przedstawione w uwadze 5.2.7 możemy powtórzyć dla sieci nieskierowanej (V, E, \bar{c}) dla ustalonego wektora $y \in \mathbb{R}^n$. Zgodnie z uwagą 5.2.2 otrzymamy wówczas

$$c^* - 2e^\top y = \bar{c}^* \geq \bar{c}(F) + \bar{c}(f) + \bar{c}(g),$$

gdzie \bar{c}^* oznacza długość najkrótszego cyklu Hamiltona w sieci (V, E, \bar{c}) , zaś c^* oznacza długość najkrótszego cyklu Hamiltona w sieci (V, E, c) . Zauważmy, że $y \in \mathbb{R}^n$ jest dowolnym wektorem. Potrzebny nam będzie jeszcze związek między długościami najkrótszego 1-drzewa w sieciach (V, E, c) i (V, E, \bar{c}) podobny do związku między długościami cyklu Hamiltona w tych sieciach, wyrażonego w uwadze 5.2.2. Aby przedstawić ten związek, wprowadzimy pewne oznaczenia. Niech $H = (W, F)$ będzie pewnym podgrafem grafu $G = (V, E)$. Stopniem wierzchołka $v \in W$ w grafie H nazywamy liczbę krawędzi $e \in F$ związanych z wierzchołkiem $v \in W$. Stopień wierzchołka oznaczamy symbolem $d_H(v)$. Oczywiście dla 1-drzewa T mamy $d_T(v) \geq 1$ dla dowolnego $v \in V$, dla cyklu Hamiltona T mamy $d_T(v) = 2$ dla dowolnego $v \in V$.

Lemat 5.2.9 *Niech T będzie podgrafem rozpinającym w sieci nieskierowanej (V, E, c) . Niech $y \in \mathbb{R}^n$ i niech $\bar{c}_{ij} = c_{ij} - y_i - y_j$, $i, j = 1, \dots, n$, będzie zmodyfikowaną funkcją odległości w tej sieci. Wówczas*

$$\bar{c}(T) = c(T) - \sum_{i=1}^n d_T(v_i) y_i.$$

Dowód.

Zbiór wszystkich 1-drzew oznaczmy symbolem \mathcal{T} .

Wniosek 5.2.10 *Niech $y \in \mathbb{R}^n$. Długość najkrótszego 1-drzewa w sieci (V, E, c) wynosi*

$$\min_{T \in \mathcal{T}} [c(T) - \sum_{j=1}^n (d_T(v_j) - 2) y_j]$$

Niech c^* oznacza długość najkrótszego cyklu Hamiltona w sieci (V, E, c) zaś c_T – długość 1-drzewa T w tej samej sieci.

Wniosek 5.2.11 *Dla długości c^* najkrótszego cyklu Hamiltona w sieci (V, E, c) słuszne jest oszacowanie*

$$c^* \geq \min_{T \in \mathcal{T}} [c_T - \sum_{j=1}^n (d_T(v_j) - 2) y_j],$$

gdzie $y \in \mathbb{R}^n$ jest dowolnym wektorem.

Dowód.

Uwaga 5.2.12 Zgodnie z wnioskiem 5.2.10 wielkość $\min_{T \in \mathcal{T}} [c(T) - \sum_{j=1}^n d_T(v_j)y_j]$ jest długością najkrótszego 1-drzewa w sieci (V, E, \bar{c}) . Może ona zatem być wyznaczona w czasie wielomianowym, na przykład metodą Kruskala lub metodą Prima (patrz uwaga 5.2.7(c)).

Problemem jest, jak dobrać wektor $y \in \mathbb{R}^n$ zapewniający największe ograniczenie dolne. Nietrudno zauważyć, że wektor taki jest rozwiązaniem zadania

$$\begin{array}{ll} \text{maksymalizować} & f(y) = \min_{T \in \mathcal{T}} (c(T) - g_T^\top y) \\ \text{względem} & y \in \mathbb{R}^n, \end{array}$$

gdzie $g_T = (d_T(v_1) - 2, \dots, d_T(v_n) - 2)^\top$. Jest to zadanie maksymalizacji funkcji wklęsłej (stanowi ona minimum funkcji liniowych). Funkcja ta nie jest jednak różniczkowalna. Mimo to istnieją efektywne metody rozwiązania tego typu zadań. Zgodnie z uwagą 5.2.12 wyznaczenie wartości minimalizowanej funkcji można wykonać w czasie wielomianowym. Znając najkrótsze 1-drzewo T' (czyli 1-drzewo realizujące to minimum) można wyznaczyć gradient funkcji liniowej $c_{T'} - g_{T'}^\top y$ realizującej $\min_{T \in \mathcal{T}} (c(T) - g_T^\top y)$. Gradient ten wynosi $-g_{T'}$ i jest jednocześnie tzw. subgradientem funkcji f w punkcie y . Najstarsza znana metoda rozwiązania powyższego zadania maksymalizacji funkcji wklęsłej pochodzi od Poljaka i ma postać procedury generującej ciąg kolejnych przybliżeń rozwiązania zadanej w postaci rekurencyjnej

$$y^{(k+1)} = y^{(k)} - t_k g_{T'}$$

(przypominającej nieco postać metody najszybszego spadku). W powyższej równości ciąg (t_k) jest wybrany zgodnie z odpowiednimi procedurami minimalizacji kierunkowej (line search). W najprostszej postaci podanej przez Poljaka w 1969 roku i wykorzystanej przez Helda i Karpa w roku 1971

$$t_k = \lambda \frac{\bar{f} - f(y^{(k)})}{\|g_{T'}\|^2}$$

gdzie $\lambda \in (0, 2)$, \bar{f} oznacza górne ograniczenie wartości optymalnej f^* funkcji celu f . Zauważmy, że w przypadku rozpatrywanego zadania zachodzi równość

$$\bar{f} - f(y^{(k)}) = \bar{c} - c^*(y^{(k)}),$$

gdzie \bar{c} jest górnym ograniczeniem długości najkrótszego cyklu Hamiltona, zaś $c^*(y^{(k)}) = \min_{T \in \mathcal{T}} [c(T) - \sum_{j=1}^n d_T(v_j)y_j^{(k)}]$ jest długością najkrótszego 1-drzewa w sieci $(V, E, c(y^{(k)}))$. Procedura zaproponowana przez Poljaka jest zbieżna do rozwiązania, jeśli za \bar{c} przyjąć c^* , czyli długość najkrótszego cyklu Hamiltona w sieci (V, E, c) , która jest na ogół nieznaną. Dla $\bar{c} > c^*$ można oczekiwać, że wielkości $f(y^{(k)})$ będą się początkowo zwiększać. Nie można natomiast oczekiwać, że będą one zbieżne do rozwiązania. Bardziej nowoczesne procedury maksymalizacji funkcji wklęsłych (lub minimalizacji funkcji wypukłych) wykorzystujące tzw. sterowanie poziomem i selekcje subgradientów są zbieżne do rozwiązania oraz – co jest równie ważne – dają dość szybko rozwiązania ε -optymalne.

5.2.4 Sformułowanie w postaci zadania programowania liniowego

5.2.5 Metoda relaksacji Lagrange'a

Literatura

- [Ceg02] A. Cegielski, *Programowanie matematyczne. Część 1. Programowanie liniowe*, Oficyna Wydawnicza Uniwersytetu Zielonogórskiego, Zielona Góra, 2002.
- [Ceg24] A. Cegielski, Podstawy optymalizacji (skrypt) 2024.
- [CCPS98] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, A. Schrijver, *Combinatorial Optimization*, John Wiley&Sons, New York 1998.
- [DD95] W. Domschke, A. Drexl, *Einführung in Operations Research*, Springer-Verlag, Berlin, 1995.
- [GN86] Z. Galas, I. Nykowski (red.), *Zbiór zadań z programowania matematycznego, część I*, PWN, Warszawa, 1986.
- [Gra80] W. Grabowski, *Programowanie matematyczne*, PWE, Warszawa, 1980.
- [HK71] M. Held, R.M. Karp, The traveling-salesman problem and minimum spanning trees: Part II, *Mathematical Programming*, 1 (1971), 6–25.
- [Hir-Lem] J.-B. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms, Vol I*, Springer-Verlag, Berlin, 1993.
- [H-UL93] J.-B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of Convex Analysis*, Springer-Verlag, Berlin, 2001.
- [Kal76] P. Kall, *Mathematische Methoden des Operations Research*, Teubner, Stuttgart, 1976.
- [KF65] A. Kaufmann, R. Faure, *Invitation a la recherche opérationelle*, Dunod, Paris, 1966. (Tłumaczenie na j. polski: *Badania operacyjne na co dzień*, PWN, Warszawa, 1968)
- [Ste89] H. Steinhaus, *Kalejdoskop matematyczny*, WSiP, Warszawa 1989.
- [SDK93] M. Sysło, N. Deo, J. S. Kowalik, *Algorytmy optymalizacji dyskretnej*, PWN, Warszawa, 1993.
- [Szw66] W. Szwarc, The time transportation problem, *Zastosowania Matematyki*, 8 (1966) 231-242.
- [Wal91] S. Walukiewicz, *Programowanie dyskretne*, PWN, Warszawa, 1986. (tłumaczenie angielskie: *Integer Programming*, PWN, Warszawa, 1991.)