根付き部分木ネットワークの利益最大化

MAXIMIZING THE PROFIT OF A ROOTED SUBTREE

安部友輔
Yusuke ABE
指導教員　五島洋行

法政大学大学院工学研究科システム工学専攻(経営系)修士課程

Given a connected undirected graph, one root vertex, non-negative income for each vertex including the root vertex, and non-negative cost for each edge, we aim to find a rooted tree for which the total profit is maximized. Such tree is known as not-necessarily-spanning. This problem can be applied to various real-life situations. We first prove that this problem is NP-hard in a strong sense. The key idea is a reduction from the uncapacitated facility location problem known as NP-hard in a strong sense. We then present three heuristic algorithms to solve the problem. Finally, we show several computational results using the three algorithms and compare the results in terms of both the objective function value and computation time. The algorithms perform well on approximation ratio.

Key Words *: combinatorial optimization, graph algorithm, heuristics, not-necessarily-spanning tree, NP-hard, mixed integer programming.*

１．はじめに

本研究では「最大利益根付き部分木問題(maximum-profit-rooted-subtree problem(MPRS))」の問題のクラスや，効率的な解法などについて検討する．具体的には，点の集合，枝の集合に加え，点に付された収入，枝に付されたコスト，根の指定が与えられる．それに対して，最大利益根付き部分木を求めるという問題である．ケーブル TV やガス, 水道事業など様々な事業がこの問題として定式化できる．

この重要な問題は古林ら[1]によってはじめて提案された．さらに[1]ではアルゴリズムを提案しており，そのアルゴリズムの優れた特徴と，しかし必ずしも最適解を求められないことを述べている. その後, 古林らは[2]にて，高速なアルゴリズムも提案しているものの，これも最適解を必ずしも求められない.

そこで本研究では，この問題の NP 困難性を証明した．さらに，混合整数計画問題として定式化をした上で，新たなアルゴリズムの提案と計算実験を行った．計算実験では古林らの二つのアルゴリズムとの比較も行った．

なお，「最大利益根付き部分木問題」は最適化問題の一つであり，図 1 のように位置付けられる．ネットワークとは，グラフ理論でいう“グラフ”に，重みを持たせたものである．「最大利益根付き部分木問題」はここに位置付けられる．他にここに位置付けられるものとして，有名な巡回セールスマン問題がある．

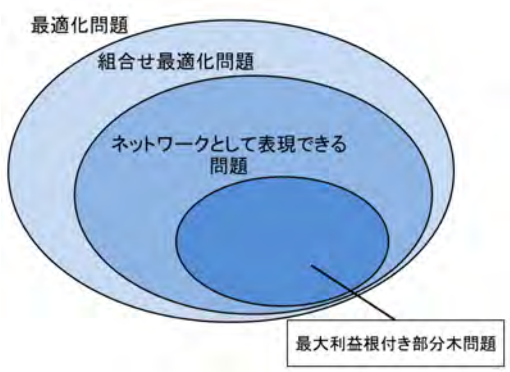

図１　最大利益根付き部分木問題の位置付け.

２．問題の説明

入力は連結な無向グラフ $G = (V, E)$，根の指定，各点に付された収入，各枝に付されたコストである．それに対して出力は，利益最大の根付き部分木である．なお，根を含み連結で無閉路な部分グラフのことを根付き部分木という．

$G = (V, E)$は各点に番号が付いているとする．根は番号 1 であり，他の点は $2, 3, \ldots, n\ (\geq 2)$と正の整数でそれぞれ異なる番号が付いているとする．また，この番号付けによって，Vのサイズは n となる．

各点 $i(\in V)$は，実数である非負収入 p_i を持つ．なお，出力は必ず点 r を含むため，$p_1 = 0$ とする．各枝$\{i, j\}(\in E)$は，実数である非負コスト c_{ij} を持つ．G は無向グラフで

あるため，$c_{ij} = c_{ji}$ となる．G における，点 r を根とする部分木を $T = (N, A)$とする．利益である profit(T)を，

$$\mathrm{profit}(T) = \sum_{i \in N} p_i - \sum_{\{i,j\} \in A} c_{ij} . \tag{1}$$

とする．出力は profit(T)を最大化する T である．この問題を最大利益根付き部分木問題 (maximum-profit-rooted-subtree problem(MPRS))と呼ぶ．例を図 2 に示す．

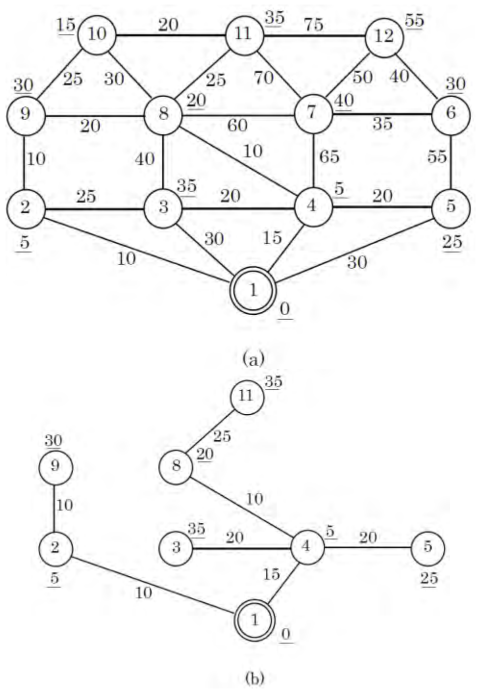

図 2　(a)は MPRS のインスタンス例. 二重丸が根を表している．(b)は(a)の最適解であり，最適値は 45.

3．混合整数計画問題

MPRS を混合整数計画問題として定式化する．

まずは，MPRS のインスタンスから，双方向である有向グラフ $D = (N_D, A_D)$を次のようにつくる．例を図 3 に示す．

- $N_D := V$.
- $A_D := \{(i,j) | \{i,j\} \in E,\quad j \neq 1\}$.
- 各枝$(i,j) \in A_D$ は重み $w_{ij} := p_j - c_{ij}$ を持つ．

D 上で最大重み根付き部分有向木を見つけることは，元問題において最適解を見つけることと等価である．

次に，各枝(i,j)に対応した 0-1 変数 x_{ij} と，各点 i に対応した実数変数 u_i を定義する．

$$x_{ij} = \begin{cases} 1, & \text{枝}(i,j)\text{を解に含む場合}, \\ 0, & \text{それ以外}, \end{cases}$$

$$u_1 = 1,$$

$$2 \leq u_i \leq n, \quad i = 2, 3, \ldots, n.$$

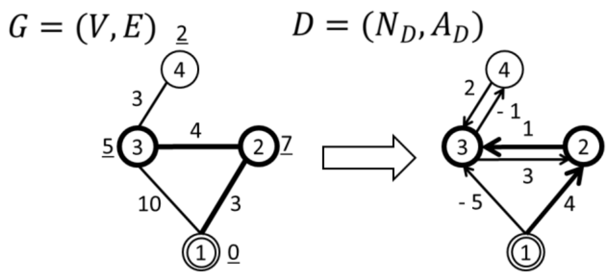

図 3　グラフ Dへの変換例．二重丸が根を表している．太線が最適解であり，最適値は 5.

MPRS を次のように定式化する．

$$\text{maximize} \qquad \sum_{(i,j) \in A_D} w_{ij} x_{ij} \tag{2}$$

$$\text{subject to} \qquad \sum_{(i,j) \in A_D} x_{ij} \leq 1, \quad \text{for all } j \in N_D \setminus \{1\}, \tag{3}$$

$$x_{ij} \leq \sum_{(k,i) \in A_D \setminus \{(i,j)\}} x_{ki} \quad , \quad \text{for all } (i,j) \in A_D \setminus \mathrm{OUT}(1), \tag{4}$$

$$u_i + 1 - (n-1)(1 - x_{ij}) + (n-3) x_{ji} \leq u_j, \quad \text{for all } (i,j) \in A_D \setminus \mathrm{OUT}(1). \tag{5}$$

式(3)–(5)は解が根付き部分有向木になるための条件を表す．式(3)は解において，選択された点に入ってくる枝が高々一つであることを示している．このため，式(3)を入次数制約と呼ぶ．

式(4)は解において，ある枝が選択されるためには，その先行の枝が少なくとも一つ選択されなければならないことを示している．この制約式により解は連結なグラフとなるため，式(4)を連結制約と呼ぶ．また，この制約式により，選択された点は根から到達可となる．式(3), (4)は文献[3]で提案されたものである．

閉路の制約式は様々なものがある[4]．本研究では，有名な閉路制約に対して持ち上げ操作を適用した式(5)を採用する．

4．NP 困難性の証明

MPRS が NP 困難である容量制約なし施設配置問題 (Uncapacitated Facility Location Problem(UFLP))に多項式時間で帰着できることにより，NP 困難性を証明する．UFLP は次のような問題である：

入力は，顧客集合 D と施設集合 F，$i \in F$ の開設コスト $o_i \in \mathbb{R}^+$，顧客 $j \in D$ が施設 $i \in F$ を利用するときの利用コスト $s_{ij} \in \mathbb{R}^+$である．出力は，開設コストと利用コストの総和が最小となるような，施設の部分集合 X と施設と，顧客の割り当て $\sigma : D \to X$ である[5].

定理 1. MPRS は NP 困難である．

証明.

UFLP のインスタンスからグラフ $G = (V, E)$をつくる．点集合を $V = F \cup D \cup \{r\}$とし，枝集合を $E = E_F \cup E_s$とする．ここで，　$E_F = \{\{r, i\} \mid i \in F\}$，$E_s = \{\{i, j\} \mid i \in F, j \in D\}$とする．

次に，点と枝の重み関数を定義する．M_tを UFLP のすべてのコストの総和とし，M_oをすべての開設コストの総和とする．すなわち，$M_t = \sum_{i \in F} o_i + \sum_{i \in F, j \in D} s_{ij}$，$M_o = \sum_{i \in F} o_i$となる．点 $v \in F \cup \{r\}$の重みを $p_v = 0$ とし，点 $v \in D$ の重みを $p_v = M_t + M_o$ とする．また，枝$\{r, i\}(i \in F)$の重みを $c_{ri} = o_i$，枝$\{i, j\}(i \in F, j \in D)$の重みを $c_{ij} = M_o + s_{ij}$とする．この変換の様子を図 4 に示す．この変換は明らかに多項式時間で行える．

UFLP のインスタンスがコスト K の最適解(X, σ)を持つことの必要十分条件が，MPRS において$|D|M_t - K$ となる最適な部分木を持つことであることを示せばよい．UFLP の最適解(X, σ)から，MPRS における部分木 $T = (N, A)$を次のように構成する．

点集合を $N = \{r\} \cup X \cup D$，枝集合 $A = \{\{r, i\} \mid i \in X\} \cup \{\{\sigma(j), j\} \mid j \in D\}$とする．このとき部分木 T の利益は，

$$
\begin{aligned}
&\sum_{j \in D} p_j - \Big(\sum_{i \in X} c_{ri} + \sum_{j \in D} c_{\sigma(j)j}\Big) \\
&= |D|(M_t + M_o) - \Big(\sum_{i \in X} o_i + |D| M_o + \sum_{j \in D} s_{\sigma(j)j}\Big) \qquad (6) \\
&= |D| M_t - K.
\end{aligned}
$$

となり，UFLP のインスタンスの最適解のコストが K のとき，MPRS において利益$|D|M_t - K$ となる部分木 T が存在する．

逆を証明する．グラフ G における MPRS の最適解を $T = (N, A)$とする．もし，点集合 N が D を必ず含んでいて，D の各点の次数が 1 であるならば，部分木 T は UFLP のインスタンスの最適解(X, σ)に対応する．この議論を次に示す．

$\{\{r, i\} \in A \mid i \in X\}$，$\{\{i, j\} \in A \mid \sigma(j) = i, i \in F, j \in D\}$であるとすると，部分木 T の利益は$|D|M_t - K$ となる．式(6)から UFLP のインスタンスの最適解(X, σ)のコストは K である．よって，最適な部分木 T はすべての D を含んでいて，かつ，D の各点の次数は 1 である．D の各点 j は正の重み $p_j = M_t + M_o$を持つ．点 r から，D の各点 j に対して，点 $i \in F$ と 2 本の枝を含んだパスができる．$c_{ri} + c_{ij} = o_i + M_o + s_{ij} < M_t < M_t < M_t + M_o = p_j$であるため，点 $j \in D$ において，2 本の枝$\{r, i\}$と$\{i, j\}$との和は p_j より小さい．よって，点 $j \in D$ は部分木 T に含まれている．

次に部分木 T における，D の各点の次数が 1 であることを示す．

まず，点 $j \in D$ の次数が高々2 であると仮定する．そうすると，部分木 T においての，根 r から点 $i \in F$ までのパス P_i に含まれている枝$\{i, j\}$の点 i が存在する．ここで，枝集合 $A_{T'} = (A_T \cup \{r, i\}) \setminus \{i, j\}$とすると，$|A_T| = |A_{T'}|$となるため，$T' = (N, A_{T'})$は部分木であり，　T は r に繋がる点 i と点 j の二つの点を持つ．さらに，$o_i = c_{ri} < c_{ij} = M_o + s_{ij}$から，$T'$の利益は T の利益より大きくなり，T が最適解であることに矛盾する．よって，部分木 T における各点 $j \in D$ の次数は 1 である．□

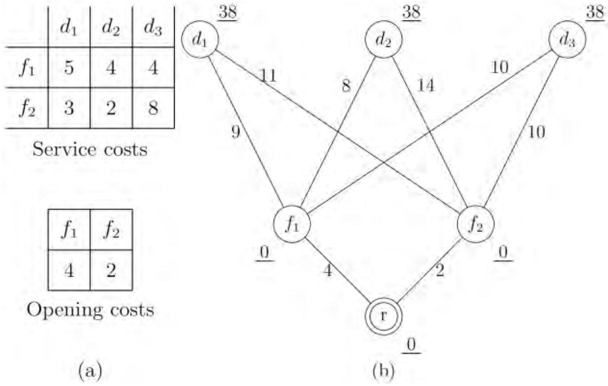

図 4　(a)は UFLP のインスタンス例．(b)は(a)を MPRS のインスタンスに変換したもの．

５．ヒューリスティックアルゴリズム

3 つの異なるアルゴリズムの概要を示す．

（１）剪定法

本研究で提案するアルゴリズム．

このアルゴリズムでは，まず入力グラフ G に対して，3 章で述べた双方向の有向グラフ D をつくる．次に D に対して，最大重み根付き全点有向木を求める．なお，最大重み根付き全点有向木の多項式時間アルゴリズムは既存である[5]．

最後に，求めた全点有向木から余分な枝を削除する．削除される枝が一つもない場合は，最適解となる．

（２）付随木連結法

文献[1]で提案されたアルゴリズム．

各点のそれぞれに，それを根とする部分木を持たせる．最初の段階では各点は自分自身のみを部分木として持っている．部分木はそれぞれの利益が大きくなるように拡大していく．

最終的には点 1 の部分木に含まれる点に対して，元グラフ上で最小全点木を張り，解を得る．

（３）最大重み経路法

文献[2]で提案されたアルゴリズム．

このアルゴリズムの主なアイデアは，最長経路を繰り返し求めることである．まずは，最大重み経路が存在するように，すなわち，正の重みの閉路が存在しないように，点を併合したグラフをつくる．そして，そのグラフ上で最大重み経路を求め，それらをつなぐことによって部分木を得る．次に，剪定法と同じ方法で余分な枝を削除する．

最後に，求めた点集合に対して，元グラフ上で最小全点木を張り，解とする．

６．計算実験

入力グラフは，網グラフとランダムグラフの 2 種類を用いた．網グラフは格子グラフに斜めの枝を加えたものである(図 5). ランダムグラフは LEDA(Library for Efficient Data types and Algorithms)の関数で発生させた．大きさは点数 9 個, 枝数 20 個～点数 2,500 個, 枝数 9,702 個とし, それぞれの大きさで 100 回ずつ実験を行った. 網グラフ, ランダムグラフともに, 点重み 1～1,000, 枝重み 1～2,200 の整数とし，一様分布で発生させた．

近似比=最適値/近似値と定義した．近似比が 1 のとき, 近似解は最適解となる. 近似比が 1 に近ければ近いほど, 良い値となる．3 章で，混合整数計画問題として定式化したため, ソルバーCPLEX を用いて最適値を求めた. なお, 点 225 枝 812 のサイズ以上だと，最適値を求めるのに 7 日間以上かかるため値を求めていない．

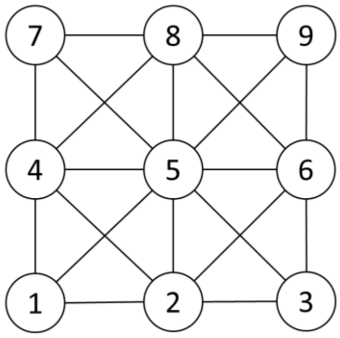

図 5　点 9 枝 20 の網グラフ.

表 1　網グラフにおける目的関数の平均値.

点	枝	剪定法	付随木	重み経路
9	20	1,013	1,034	1,040
25	72	3,929	3,993	3,879
49	156	8,773	8,841	8,458
225	812	46,098	46,487	44,770
625	2,352	136,690	137,247	132,735
841	3,192	183,627	184,626	178,741
1,600	6,162	357,121	359,217	347,860
2,500	9,702	561,821	565,497	547,137

表 2　網グラフにおける計算時間の平均値.

点	枝	剪定法	付随木	重み経路
9	20	0.001	0.003	0.002
25	72	0.008	0.004	0.002
49	156	0.037	0.006	0.002
225	812	0.876	0.155	0.006
625	2,352	8.584	3.100	0.018
841	3,192	19.329	7.650	0.031
1,600	6,162	92.591	58.171	0.111
2,500	9,702	296.394	225.976	0.265

表 3　網グラフにおける近似比の平均値.

点	枝	剪定法	付随木	重み経路
9	20	1.06	1.02	1.83
25	72	1.04	1.01	1.06
49	156	1.03	1.02	1.07

表 4　網グラフにおける近似比の標準偏差.

点	枝	剪定法	付随木	重み経路
9	20	0.379	0.064	8.249
25	72	0.115	0.029	0.124
49	156	0.053	0.029	0.087

表 1，3 から，付随木連結法，剪定法，最大重み経路法の順番で，目的関数値が良いことがわかる．表 3 から，最大重み経路法が最も速いといえる．次に，付随木連結法，剪定法の順に速い．

表 2 の近似比は 3 つのアルゴリズムとも良好な値である．さらに，表 4 からグラフサイズが大きくなるにつれて，バラつきが小さくなっていることがわかる．紙面の都合上，ランダムグラフの結果は載せていないが，網グラフとほとんど同じ結果が得られた．

７．おわりに

MPRS は様々な事業を含んでいる重要な問題である．MPRS に対するアルゴリズムはいくつか提案されていたものの，最適解を必ずしも求められなかった．

NP 困難性の証明を行ったことが本研究の一つの貢献である．また，MPRS を混合整数計画問題として定式化した．そのため，ソルバーCPLEX を使うことができた．さらに，計算実験では，3 つの手法が異なる特徴を持つことと，良好な近似比を示すことができた．

参考文献

1) 古林隆，福馬敏子，榊原英行，露木真理子：最大利益根付き問題のアルゴリズム，日本オペレーションズ・リサーチ学会 2003 年秋季研究発表会アブストラクト集, pp. 166-167，2003

2) 古林隆，福馬敏子：最大利益根付木問題のアルゴリズム II -最大重み経路法-，日本オペレーションズ・リサーチ学会 2005 年春季研究発表会アブストラクト集, pp. 116-117，2005

3) V. V. Rao and R. Sridharan : Minimum-weight rooted not-necessarily-spanning arborescence problem, *Networks*, vol. 39, no. 2, pp. 77–87, 2002

4) 久保幹雄：ロジスティクスの数理，第 6 章，共立出版，2007

5) B. Korte and J. Vygen : Combinatorial Optimization, Fourth Edition, Chapters 6 and 22, Springer-Verlag, 2008

2012年度　　修士論文

論文題名　根付き部分木ネットワークの利益最大化

-Maximizing the profit of a rooted subtree-

指導教員　五島洋行

大学院工学研究科

システム工学専攻修士課程

11R6202

アベ　ユウスケ

氏名　安部友輔

Abstract In this thesis, we address *maximum-profit-rooted-subtree problem* (*MPRS*). In the problem, given a connected undirected graph, a root vertex, each vertex's non-negative income, and each edge's non-negative cost, we aim to find a rooted tree in which the total profit is maximized. The problem can be applied to various real-life situations such as cable television service and supply of energy (gas, water and electricity etc.). This important problem had been proposed by " Kobayashis(2003) " for the first time.

The main contributions of this research are to prove that the $MPRS$ is $NP-hard$, formulate the $MPRS$ as a mixed integer programming problem, and propose a heuristic algorithm.

The $MPRS$ is classified as an optimization problem. In optimization problems, an objective function and constraint equations are given. Under the constraints, we find a maximum or minimum value of the objective function. Combinatorial optimization problem is also classified as an optimization problem, and the solutions are represented by combination or permutation. Moreover, a part of combinatorial optimization problems can be formulated as a network. Our problem is positioned here. Using the ideas and terminologies in graph theory, this problem is referred to as *maximum-profit-rooted-subtree problem*, and we call this $MPRS$ for short.

Chapter 1 is an introduction, and we outline the background and organization of this thesis. Next, in chapter 2, we definite terminologies. In chapter 3, we formulate the $MPRS$ as a combinatorial optimization problem. Moreover, practical usages and previous researches are overviewed. In chapter 4, we formulate the $MPRS$ as a mixed integer programming problem. In chapter 5, we prove that the $MPRS$ is $NP-hard$. In chapter 6, we introduce three heuristic algorithms, one of which has been proposed by us. Though the other algorithms are existing, we rewrite these in detail. In chapter 7, we show the results of computational experiments and perform considerations. Finally, in chapter 8, we conclude our research giving future works.

論文要旨　本研究では「最大利益根付き部分木問題 (*maximum-profit-rooted-subtree problem* (*MPRS*))」の問題のクラスや，効率的な解法などについて検討する．具体的には，点の集合，枝の集合に加え，点に付された収入，枝に付されたコスト，根の指定が与えられる．それに対して，最大利益根付き部分木を求めるという問題である．ケーブル TV やガス，水道事業など様々な事業がこの問題として定式化できる．この重要な問題は「古林ら（2003）」によってはじめて提案された．

本研究の主な貢献は，NP 困難性の証明，混合整数計画問題としての定式化，新たなアルゴリズムの提案を行ったことである．

なお，「最大利益根付き部分木問題」は最適化問題の一つである．最適化問題とは，ある制約条件と目的関数が与えられる．そして，その制約条件の下で，目的関数を最大化，もしくは最小化する解を求める問題である．最適化問題の中でも組合せ最適化問題は，解が順列，組合せで表される問題であり，さらに組合せ最適化問題の中にはネットワークとして表現することができる問題がある．ネットワークとは，グラフ理論でいう“ グラフ ”に，重みを持たせたものである．「最大利益根付き部分木問題」はここに位置付けられる．

第 1 章では，研究背景と論文構成を述べる．次に第 2 章で用語の定義を行う．なお，グラフ理論における用語の定義は，文献によって異なることが多い．本論文では第 2 章での定義を使用することとする．それを用いて第 3 章で $MPRS$ を組合せ最適化問題として定式化する．さらに，応用と先行研究について述べる．第 4 章では $MPRS$ を混合整数計画問題として定式化する．第 5 章では NP 困難性の証明をする．なお，NP 困難である問題は，多項式時間で最適解を求めることが難しい問題である．次の第 6 章では近似解を求めるための提案手法を三つ紹介する．一つは我々の提案したものであり，二つは既存のアルゴリズムである．本論文では二つの既存アルゴリズムを詳しく書き直している．そして，第 7 章で計算実験の結果と考察を示す．最後に第 8 章で結論と今後の課題を述べる．