1999年度日本オペレーションズ・リサーチ学会
秋季研究発表会

最大リーフ全域木問題に対する厳密解法

02501520 神戸商科大学 藤江 哲也 FUJIE Tetsuya

1 はじめに

連結無向グラフ $G=(V,E)$ が与えられたとき，G の全域木 (spanning tree) の中でリーフ (次数 1 の頂点) 数が最大のものを求める問題を最大リーフ全域木問題 (Maximum Leaves Spanning Tree Problem : MLSTP) という．MLSTP は NP 困難であり [3]，いくつかの近似解法が存在する [6, 8]．現時点で知られている最良の近似比は 2 である．一方，MLSTP の最適なリーフ集合の補集合は，最小連結支配集合問題 (Minimum Connected Dominating Set Problem : MCDSP) の最適解である．MCSDP については，ある条件の下で，近似比 $(1-\varepsilon)\ln|V|$ ($^\forall\varepsilon>0$) の解法の非存在性が示されている [4]．

本稿では，[1, 2] で与えられた MLSTP の定式化及び緩和問題に基づく分枝限定法を提案し，数値実験の結果を示す．

2 定式化と緩和問題

$G=(V,E)$ の全域木 $T\subseteq E$ に関する特性ベクトル $\chi^T(\chi^T_e=1\Leftrightarrow e\in T)$ の集合を $\mathrm{ST}(G)\subseteq\{0,1\}^E$ とする．また，$i\in V$ に対して $\delta(i)$ を i に接続する枝集合とし，$F\subseteq E$ に対して $\boldsymbol{x}(F)\equiv\sum_{e\in F}x_e$ とする．このとき，MLSTP は次のように定式化される [1, 2]:

$$
\begin{aligned}
\text{(P)}\quad \text{最大化}\quad & \sum_{i\in V} y_i\\
\text{条件}\quad & \boldsymbol{x}\in\mathrm{ST}(G),\\
& \boldsymbol{x}(\delta(i))+(|\delta(i)|-1)y_i\le|\delta(i)| \quad (i\in V),\\
& y_i\in\{0,1\} \quad (i\in V).
\end{aligned}
$$

$(\boldsymbol{x},\boldsymbol{y})$ を (P) の実行可能解とする．このとき $\boldsymbol{x}$ は G の全域木 T を表現し，$\boldsymbol{y}$ は T のリーフ集合の部分集合 (以下，リーフ部分集合と呼ぶ) を表現している．

本稿では，変数 y_i $(i\in V)$ に関する連続緩和問題 $(\overline{\mathrm{P}})$ を考える．ここで，(P) の 3 種類の制約式のうち第 1 式と第 2 式から $y_i\le1$ が導かれるため，$(\overline{\mathrm{P}})$ は第 3 式を非負制約 $y_i\ge0$ に置き換えたものとしてよい．よって，$(\overline{\mathrm{P}})$ の最適解は第 2 式を等式で満たし，y_i $(i\in V)$ を消去した次の問題を $(\overline{\mathrm{P}})$ とすることができる:

$$
\begin{aligned}
(\overline{\mathrm{P}})\quad \text{最大化}\quad & \sum_{i\in V}\frac{d_i-\boldsymbol{x}(\delta(i))}{d_i-1}\\
&=\sum_{i\in V}\frac{d_i}{d_i-1}-\sum_{e=(i,j)\in E}\left(\frac{1}{d_i-1}+\frac{1}{d_j-1}\right)x_e\\
\text{条件}\quad & \boldsymbol{x}\in\mathrm{ST}(G).
\end{aligned}
$$

ただし $d_i=|\delta(i)|$ である．$(\overline{\mathrm{P}})$ は最小全域木問題であり，多くの解法が存在する．

3 子問題

グラフ上の最適化問題の多くは，G から頂点や枝を除去または縮約し，縮小されたグラフを子問題とする．しかし MLSTP の場合，このようなグラフを構成するのは難しい．また定式化 (P) において，固定される頂点に応じて y_i に 0 または 1 の値を代入した問題を考えるのは，その緩和問題が，(一般に NP 困難な) 次数制約付き全域木問題の形になる．そのため，本稿では次の方法をとった．

まず，MLSTP を “G の全域木の中で，リーフ部分集合の大きさが最大のものを求めよ” と書き直す．(P) はこの記述通りの定式化である．また，MLSTP の子問題を $\mathrm{MLSTP}(S_1,S_0,F)$ と表現する．ここで，(S_1,S_0,F) は頂点集合 V の分割であり，$\mathrm{MLSTP}(S_1,S_0,F)$ は “$S_1\subseteq S\subseteq S_1\cup F$ を満たすリーフ部分集合 S の中で要素数最大のものを求めよ” ということになる．MLSTP 自身 (ルート問題) は $\mathrm{MLSTP}(\emptyset,\emptyset,V)$ と等価である．以上の準備の下で，次の補題を示すのは容易である．

補題 3.1 $\mathrm{MLSTP}(S_1,S_0,F)$ の定式化は次のようにして与えられる:

$$
\begin{aligned}
&(\mathrm{P}(S_1,S_0,F))\\
\text{最大化}\quad & \sum_{i\in S_1}(|V|+1)y_i+\sum_{i\in F}y_i-|V|\cdot|S_1|\\
\text{条件}\quad & \boldsymbol{x}\in\mathrm{ST}_G,\\
& \boldsymbol{x}(\delta(i))+(|\delta(i)|-1)y_i\le|\delta(i)| \quad (i\in S_1\cup F),\\
& y_i\in\{0,1\} \quad (i\in S_1\cup F),\\
& y_i=0 \quad (i\in S_0).
\end{aligned}
$$

さらに，変数 y_i ($i \in S_1 \cup F$) に関する連続緩和問題 ($\overline{\mathrm{P}(S_1, S_0, F)}$) は最小全域木問題になる． ■

4 数値実験

分枝限定法の実装は，品野らによって開発された汎用並列分枝限定法ツール PUBB を用いて行なった (PUBB の詳細は [7] に述べられている)．ただし，分枝限定法の並列化は行なっていない．

初期実行可能解の計算は，[6, 8] に与えられている近似解法に加えて，幅優先探索によって全域木を求める方法を実行した．また，子問題 MLSTP(S_1, S_0, F) の分枝は，未分枝頂点集合 F の中で次数が最大の頂点 v を選び，S_1 に v を加えた問題と S_0 に v を加えた問題を生成した．表1は頂点数を n，枝密度を p とするランダムグラフに対する計算結果である．ここで，各 (n, p) について 10 問生成し，その平均を示している．また，LB は初期実行可能解の値，UB はルート問題の上界値，OPT は最適値，子問題数は生成された子問題数，time は実行時間 (秒) である．使用した CPU は Pentium II (300MHz) である．

表 1: ランダムグラフに対する計算結果

n	p	LB	UB	OPT	子問題数	time
70	0.3	62.5	67.0	65.0	432481.0	462.96
	0.4	64.6	67.9	66.0	56040.4	75.13
	0.5	65.8	68.0	66.9	12955.8	19.86
	0.6	66.4	68.0	67.1	3929.0	7.22
	0.7	67.2	68.0	67.9	683.4	1.40
80	0.4	74.4	77.5	75.9	279511.6	462.33
	0.5	75.4	78.0	76.7	43589.0	89.51
	0.6	76.4	78.0	77.0	5686.8	13.78
	0.7	77.2	78.0	77.8	1692.2	4.70
90	0.5	85.1	88.0	86.4	121852.0	316.82
	0.6	86.2	88.0	87.0	7393.8	23.16
	0.7	86.8	88.0	87.4	4954.8	16.87
100	0.5	95.2	98.0	96.3	206056.0	667.95
	0.6	96.2	98.0	97.0	9225.2	34.72
	0.7	97.0	98.0	97.3	6658.6	28.55

表1が示すように，最適値は自明な上界値 ($|V|-1$) に近い傾向にある．実際，Kleitman and West[5] は，頂点数 n，最小次数が少なくとも k のグラフに対し，k が十分大きいならば最適値は少なくとも $(1 - b \ln k/k)n$ になることを示した (b は 2.5 より大きい定数)．そこで，小さい次数の頂点から成る例として $m \times n$ 格子グラフ $G_{m \times n}$ について考察した．そして次の結果を得た．

補題 4.1 $G_{m \times n}$ に対する ($\overline{\mathrm{P}}$) の最適値は $2mn/3$ である． ■

補題 4.2 $m, n \geq 4$ に対し $G_{m \times n}$ は，リーフの数が

$$mn - \min \left\{ \begin{array}{l} 2m + (n-4) + \left\lfloor \frac{n-4}{3} \right\rfloor (m-2), \\ 2n + (m-4) + \left\lfloor \frac{m-4}{3} \right\rfloor (n-2) \end{array} \right\}$$

に等しい全域木を持つ． ■

したがって，$G_{m \times n}$ の場合，最適値は $\Theta(mn) = \Theta(|V|)$ であり，また，上界値と下界値の差が大き過ぎることはない．しかし，本稿で提案した分枝限定法では $m \times n < 100$ までしか解くことができなかった．今後は，上界値・下界値とも最適値との差を縮めることが課題となる．

参考文献

[1] 藤江, 最大リーフ全域木問題について, 1997 年 OR 学会秋季研究発表会, 272-273.

[2] T. Fujie, "The Maximum-Leaf Spanning Tree Problem: Formulations and Facets," 投稿中.

[3] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, Freeman, New York (1979).

[4] S. Guha and S. Khuller, "Approximation Algorithms for Connected Dominating Sets," *Proceedings of the Fourth Annual European Symposium on Algorithms* (1996) 179–193.

[5] D. J. Kleitman and D. B. West, "Spanning Trees with Many Leaves," *SIAM Journal on Discrete Mathematics* **4** (1991) 99–106.

[6] H. Lu and R. Ravi, "Approximating Maximum Leaf Spanning Trees in Almost Linear Time," *Journal of Algorithms* **29** (1998) 132–141.

[7] 品野, 藤江, 汎用分枝限定法ツール PUBB による組合せ最適化問題の厳密解法, 統計数理 46 (1998) 411–431.

[8] R. Solis-Oba, "2-approximation Algorithm for Finding a Spanning Tree with Maximum Number of Leaves," *Lecture Notes in Computer Science* **1461** (1998) 441–452.