◎理论与研发◎

内部节点受限的最小生成树问题算法研究

蒋小娟[1],张 安[1],陈 永[1],陈光亭[2]
JIANG Xiaojuan[1], ZHANG An[1], CHEN Yong[1], CHEN Guangting[2]

1.杭州电子科技大学 理学院,杭州 310018
2.台州学院,浙江 台州 317000
1.School of Science, Hangzhou Dianzi University, Hangzhou 310018, China
2.Taizhou University, Taizhou, Zhejiang 317000, China

JIANG Xiaojuan, ZHANG An, CHEN Yong, et al. Algorithm for minimum internal nodes constrained spanning tree problem. Computer Engineering and Applications, 2017, 53(10):35-37.

Abstract: The minimum internal nodes constrained spanning tree problem is considered. Give a metric graph $G=(V,E)$ with a cost function $w:E\rightarrow R^{+}$ and one subset R of $V\,(R\subset V)$, the minimum internal nodes constrained spanning tree problem asks for a minimum weight spanning tree such that every vertex in R is not a leaf. As the problem is NP− hard, a Pseudo-polynomial time optimal algorithm is first provided, then a simple polynomial time approximation algorithm with a performance ratio of 2 is designed and an instance is constructed to show the ratio is tight.
Key words: undirected weighted graph; spanning tree; approximation algorithm; performance ratio

摘 要:研究内部节点受限的最小生成树问题:给定一个赋权无向完全图 $G=(V,E)$,假定 $w:E\rightarrow R^{+}$ 为边集 E 的权重函数且满足三角不等式,给定点集 V 的一个子集 $R(R\subset V)$,目标是寻找图 G 的一个满足 R 中的点皆为内部顶点的权重最小的生成树。由于该问题是NP−困难的,提出了一个伪多项式时间最优算法,设计了一个近似比为2的多项式时间近似算法,并且给出例子以说明该近似比是紧的。
关键词:无向赋权图;生成树;近似算法;近似比
文献标志码:A **中图分类号:**O224 doi:10.3778/j.issn.1002-8331.1511-0349

1 引言

给定一个赋权无向连通图 $G=(V,E)$,其中 V 和 E 分别为图 G 的点集和边集,最小生成树问题是指在图 G 的所有生成树中寻找一个总权重最小的生成树。最小生成树问题是图论和组合最优化领域一个经典问题,其在计算机网络、通讯网络、交通运输网络及其他网络相关问题中具有极其广泛的应用。 目前,Kruskal算法[1]和Prim算法[2]是求解最小生成树问题最著名的两个高效算法。然而,随着基建和设备成本的逐渐增加,当今的通讯网络和交通运输网络建造对节点的结构或功能提出了更多要求或限制。例如,在无线传感器网络中,考虑到资源分配和设备成本,某些节点仅具有信息传送功能或者仅作为信息传输中转站,即具有特殊要求的这些节点只能作为网络中的内部节点[3]。本文所考虑的内部节点受限的最小生成树问题可描述如下:给定一个赋权无向完全图 $G=(V,E)$,假定 $w:E\rightarrow R^{+}$ 为边集 E 的权重函数且满足三角不等式,给定点集 V 的一个子集 $R(R\subset V)$,目标是寻找图 G 的一个满足 R 中的

基金项目:国家自然科学基金(No.11571252);浙江省自然科学基金(No.LY16G010008)。
作者简介:蒋小娟(1994—),女,研究领域为网络优化,E-mail:xj_jiang_math@hotmail.com;张安(1982—),男,博士,副教授,研究领域为算法设计与分析;陈永(1982—),男,博士,副教授,研究领域为组合优化;陈光亭(1965—),男,博士,教授,研究领域为网络优化、近似算法。
收稿日期:2015-11-27 **修回日期:**2016-04-18 **文章编号:**1002-8331(2017)10-0035-03
CNKI**网络优先出版:**2016-05-27, http://www.cnki.net/kcms/detail/11.2127.TP.20160527.1455.014.html

点皆为内部顶点的权重最小的生成树。

与本文所研究问题相关的一类问题是最小终端斯坦纳树问题[4]和最小内嵌斯坦纳树问题[5]。给定一个赋权无向完全图 $G=(V,E)$ 及点集 V 的一个子集 $R(R\subset V)$，最小终端斯坦纳树问题是指求图 G 的一个满足 R 中的点皆为叶子顶点的权重最小的斯坦纳树，最小内嵌斯坦纳树问题是指求图 G 的一个满足 R 中的点皆为内部顶点的权重最小的斯坦纳树。最小生成树要求扩张到 V 的所有顶点，而斯坦纳树仅需扩张 V 的部分顶点，这是两者的主要区别。 众多学者对上述两个问题进行了深入研究并均已得到较好结果（见文[6-10]）。目前尚无文献针对内部节点受限的最小生成树问题进行研究。 由于该问题是 NP－困难的，在本文中，首先提出了一个时间复杂度为 $O(n^{2k+1}\,\mathrm{lb}\,n)$ 的伪多项式时间最优算法，其中 k 为集合 R 中的顶点数目。接着设计了一个时间复杂度为 $O(n^2\,\mathrm{lb}\,n)$，近似比为2的多项式时间近似算法，最后对本文进行了总结和展望。

2 伪多项式时间最优算法

为便于表述，记 $R=\{v_1,v_2,\cdots,v_k\}$，$|V|=n$ 和 $|R|=k$ 分别表示顶点集合 V 和 R 中的顶点数目，注意到任意一颗树中至少有两个叶子顶点，所以这里不妨假设 $1\leqslant k\leqslant n-2$。

首先给出该问题的复杂性结果，具体如下。

定理 1 内部节点受限的最小生成树问题是困难问题。

证明：考虑 $|R|=n-2$ 时（集合 R 中的顶点数目）的特殊情形。不妨假设，$s,t\in V/R$。由于要求集合 R 中的点均为内部顶点（度数至少为2），则此时 s 和 t 就成为所求生成树中仅有的两个叶子顶点，即该问题就等价于寻求图 G 中从 s 到 t 的一条Hamilton路，由于Hamilton路问题是 NP－困难的[11]，因此内部节点受限的最小生成树问题也是一个 NP－困难问题。

由于该问题是 NP－困难的，因此将着重从近似算法的角度考虑问题的求解，首先提出一个求解该问题的伪多项式时间最优算法，算法1具体步骤如下。

输入：赋权无向完全图 $G=(V,E)$，其中 $w:E\rightarrow R^+$ 为边集 E 的权重函数，点集 V 的一个子集 R（$R\subset V$ 且 $1\leqslant|R|\leqslant n-2$）。

输出：生成树 T_R（要求 R 中的点皆为内部节点）。

步骤 1 对每一个 $v_i\in R(1\leqslant i\leqslant k)$，任意选取 v_i 的两条边 e_i^{i1} 和 e_i^{i2}。 记

$$G_0=\left\{e_1^{11},e_1^{12},e_2^{21},e_2^{22},e_3^{32},\cdots,e_k^{k1},e_k^{k2}\right\}$$

步骤 2 在 G_0 基础之上，调用 Prim 算法或 Kruskal 算法求解图 G 的一个最小生成树，记所得生成树为 T_0。

步骤 3 在所有的生成树 T_0 中，选择一个权重最小的生成树记为 T_R 并输出。

显然，算法1的时间复杂度主要由步骤1和步骤3决定，由于图 G 为完全图，所以 G_0 的个数最多为 n^{2k} 个，即该算法的时间复杂度为 $O(n^{2k+1}\,\mathrm{lb}\,n)$。

定理 2 算法1是求解内部节点受限的最小生成树问题的最优算法。

证明：不妨假设 T_R^* 为该问题的一个最优解，由于要求 R 中的点皆为内部节点，所以在最优解 T_R^* 中任一顶点 $v_i\in R$ 的度数至少大于等于 2，即任一顶点 v_i 至少有两条边相连接。 由于算法1穷举了所有可能性并选取权重最小的一个生成树作为输出，显然算法1求解得到了该问题的最优解。

3 近似算法

本文近似算法设计的思想主要基于Christofides算法思想[12]：最小生成树和Double 技术，该算法是用来求解经典 TSP（Traveling Salesperson Problem）问题的。需要说明的是设计的近似算法最终的输出结果是一条Hamilton路（特殊的一颗生成树），而不是一颗常规意义下的生成树。算法2具体步骤如下。

输入：赋权无向完全图 $G=(V,E)$，其中 $w:E\rightarrow R^+$ 为边集 E 的权重函数，点集 V 的一个子集 R（$R\subset V$ 且 $1\leqslant|R|\leqslant n-2$）。

输出：生成树 T_R（要求 R 中的点皆为内部节点）。

步骤 1 记 $s,t\in V/R$，调用 Prim 算法或 Kruskal 算法求解图 $G/\{s,t\}$ 的一个最小生成树，记为 $T_{G/\{s,t\}}$。

步骤 2 记 $s',t'\in T_{G/\{s,t\}}$ 分别表示距离 s,t 最近的点，连接 s' 与 s，t' 与 t，记所得到的图为 T'。

步骤 3 复制树 T' 中除 $e_{s,s'}$ 和 $e_{t,t'}$ 以外的的每一条边，记所得图为 T''。

步骤 4 找到图 T'' 中从 s 到 t 的一条Euler路 EP_{st}，根据三角不等式把 EP_{st} 删减成一条 s 到 t 的 Hamilton 路 HP_{st}。

步骤 5 在所有可能的Hamilton路 HP_{st} 中选取权重最小的一个记为 T_R 并输出。

定理 3 算法2是求解内部节点受限的最小生成树问题的近似比为2的多项式时间算法。

证明：不妨假设 u,v 是最优解 T_R^* 中的两个叶子节点，由算法2可知，$w(T_R)\leqslant w(HP_{uv})$。由于 $T_{G/uv}$ 是图 $G/\{u,v\}$ 的一个最小生成树，所以 $w(T_{G/uv})\leqslant w(T_R^*/\{u,v\})$。根据图 T' 的定义可得，$w(T')\leqslant w\left(T_R^*\right)$。另一方面，依据三角不等式。显然有 $w(HP_{uv})\leqslant w(EP_{uv})=w(T'')\leqslant 2w(T')$，因此 $w(T_R)\leqslant 2w\left(T_R^*\right)$ 成立。

考虑到 V/R 中最多有 $n-1$ 个节点，因此算法2最多构造出不超过 n^2 条Hamilton路，即该算法的时间复杂度为 $O(n^2 \operatorname{lb} n)$ 。

下面提供一个实例以说明该算法的近似比2是紧的。考虑图1所示的完全网络 $G=(V,E)$ ，$|V|=2n+2$，$R=\{s,t\}$ ，记 $A=\{1,2,\cdots,n\}$ 和 $B=\{1',2',\cdots,n'\}$ 。令 $w(i,s)=1$，$w(s,t)=1$，$w(t,j)=1(1\leqslant i\leqslant n, 1\leqslant j\leqslant n')$，其他任意两顶点之间的权重等于该两点间的最短路距离，即 $w(i,k)=2(1\leqslant i,k\leqslant n)$ ，$w(j,k)=2(1\leqslant j,k\leqslant n')$ ，$w(i,j)=3(1\leqslant i\leqslant n, 1\leqslant j\leqslant n')$ 。

显然在该实例中，图1就是要求顶点 s,t 为内部节点的最小生成树即最优解，此时 $w\left(T_R^*\right)=2n+1$ 。而算法2的最终输出结果是一条Hamilton路，由于顶点 s,t 均在该路中仅出现一次，所以任意一条Hamilton路最多包含4条权重为1的边，即算法解 $w(T_R)=4n$ 。于是，$w(T_R)/w\left(T_R^*\right)=4n/(2n+1)=2(n\to\infty)$ 。

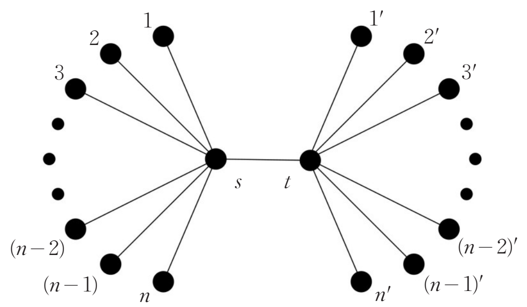

图1　顶点 s,t 为内部节点的最小生成树

4　结束语

近年来，随着设备及基建成本的逐渐增加，在无线传感器网络的构建过程中，对具有某些特殊功用的节点提出了更高要求。内部节点受限的最小生成树问题是经典最小生成树问题的一种新变形，本文首次对该问题进行探讨并尝试设计性能比较好且运算速度快的近似算法。虽然提出的近似算法无论在性能比还是计算时间上都较为令人满意，但是算法的最终输出是一条Hamilton路而不是一棵常规意义下的生成树。在后续的研究过程中，研究者可依此为突破口，尝试设计输出结果为常规意义下生成树的近似算法，相信类似近似算法的性能比将更优。此外，构造求解该问题计算速度较快的启发式算法也是一个研究方向。

参考文献：

[1] Kruskal J B.On the shortest spanning subtree of a graph and the traveling salesman problem[J].Proceedings of the American Mathematical Society，1956，7（1）：48-50.

[2] Prim R C.Shortest connection networks and some generalizations[J].Bell System Technical Journal，1957，36：1389-1401.

[3] Hsieh S Y，Yang S C.Approximating the selected-internal Steiner tree[J].Theoretical Computer Science，2007，381：288-291.

[4] Lin G H，Xue G L.On the terminal steiner tree problem[J].Information Processing Letters，2002，84：103-107.

[5] Huang C W，Lee C W，Gao H M，et al.The internal Steiner tree problem：Hardness and approximations[J].Journal of Complexity，2013，29：27-43.

[6] Fuchs B.A note on the terminal steiner tree problem[J].Information Processing Letters，2003，87：219-220.

[7] Drake D E，Hougardy S.On approximation algorithms for the terminal steiner tree problem[J].Information Processing Letters，2004，89：15-18.

[8] Martineza F V，Pinab J C D，Soares J.Algorithm for terminal steiner trees[J].Theoretical Computer Science，2007，389：133-142.

[9] Chen Y.An improved approximation algorithm for the terminal steiner tree problem[J].ICCSA，Lecture Notes in Computer Science，2011，6784：141-151.

[10] Li X Y，Zou F，Huang Y H，et al.A better constant-factor approximation for selected-internal steiner minimum tree[J].Algorithmica，2010，56：333-341.

[11] Garey M R，Johnson D S.Computers and intractability：A guide to the theory of NP-completeness[M].San Francisco：WH Freeman and Co，1979.

[12] Christofides N.Worst-case analysis of a new heuristic for the travelling salesman problem，Report 388[R].Graduate School of Industrial Administration，CMU，1976.