带测度函数的连通支配集问题*)

马 俊 朱 洪

（复旦大学计算机科学与工程系智能信息处理开放实验室 上海 200433）

摘 要 连通支配集问题在网络广播上有着广泛的应用，本文引入测度函数的概念，提出了带测度函数的连通支配集问题(CDS(F))，使得它具有更广的应用范围。文中首先给出问题的形式定义，证明了它在各种情形下的NP完全性，并给出多项式时间的近似算法，它的近似度为 $\mathrm{Ln}\Delta+3$（Δ 为图中顶点的最大度数）。

关键词 支配集问题，组合优化，NP，NP完全，多项式时间归约，NP难

Connected Dominating Sets Problem with Measured Functions

MA Jun ZHU Hong

(Laboratory for Intelligent Information Processing, Fudan University, Shanghai 200433)

Abstract Connected dominating sets problem has widely used in network broadcast. This paper introduces the concept of measured function and defines connected dominating sets problem with measured functions (CDS (F)). A formal definition of the CDS (F) is firstly given, whose NP property is then proved in variant situations. We also present an approximation algorithm with approximation ratio $\mathrm{Ln}\Delta+3$(Δ is maximal degree of the vertex in our concerned graph).

Keywords Dominating set, Combinatorial optimization, NP, NP-complete, Polynomial reduction, NP-hard

1 引言

支配集问题是集合覆盖问题的一种特殊情况，在现实中有着很广泛的应用。如果把一个图中的点想象成需要警戒的点，每条边表示两个端点之间可以互相守望，那么在支配集中的每个点的位置放一个守卫就可以保证图中的每个点都被警戒到。近年关于支配集的近似算法和参数复杂性问题的研究很多，现在已经证明它是可以在 $1+\log_2|V|$（V 表示图的顶点集）内近似的[6]，而同时它的不可近似下界是 $c\log_2|V|$（c是0到1之间的某常数）[7]。这个问题的近似算法已经研究得相当成熟了，现在关于它的变种问题的研究也有很多。Baker对平面图研究支配集问题，给出了一个PTAS算法[8]。其他的变种问题包括独立支配集问题，完美支配集问题（一般情况下这不是一个NP完全问题），连通支配集问题。其中连通支配集在网络广播中有很重要的应用。比如说，图中每个节点是一个广播站或者接受站，两个点之间的边代表它们可以被互相广播到，于是，只要让连通支配集里边的每个点在收到广播的时候，都往自己的邻居再广播一次，就能保证网络中每个节点都被广播到。

但是以上提及的这些关于支配集的问题，都是假设一个点的支配范围是1（即它只能覆盖到跟自己直接相连的点），这在现实应用中显然是有局限性的。因此我们应该考虑点的支配范围超过1的情况，甚至更一般的，点的支配范围是一个函数的情况。本文基于连通支配集问题，把连通支配集问题推广到带测度函数的情况，在丰富支配集问题研究领域的同时也为其他图论问题的研究提供了一个有益的启发。

本文第2节对带测度函数的连通支配集问题进行分析，并证明该问题是一个NP-hard问题。第3节通过一般的连通支配集问题的近似算法给出带测度函数连通支配集的近似算法。第4节总结全文，并指出下一步的工作方向。

2 带测度函数的连通支配集问题

在本节中，我们先给出带测度函数的连通支配集问题（以下简称CDS(F)问题）的定义，并证明该问题是一个NP难优化问题。

定义1 支配集(DS)问题是指，给定连通无向图 $G(V,E)$，要求找到一个最小的点集 $D\subseteq V$，使得对任意 $v\in V$，或者 $v\in D$，或者存在 $u\in D$，并且 $(u,v)\in E$。

定义2 给定无向图 $G(V,E)$ 中的两个顶点 u,v 和函数 f，称 u,v 是 f-相邻的，如果 G 中存在一条 u 到 v 的路径 l，并且满足 $length(l)\leqslant f$。

定义3 给定无向图 $G(V,E)$ 中的一个顶点集 U 和函数 f，称 U 是 f-连通的（在 G 中），如果对于 U 中任意两个点 u，v，在 G 中存在一条 u 到 v 的路径 l，并且假设 l 上有 k 个 U 中的点，依次是 $v_1=u,v_2,\cdots,v_{k-1},v_k=v$，要求它们满足：$v_i$ 与 v_{i+1} 是 f-相邻的（$i=1,2,\cdots,k-1$）。

定义4 给定无向图 $G(V,E)$ 中的两个顶点 u,v 和函数 f，称 u 是 f-可覆盖 v 的，如果 u 和 v 是 f-相邻的。

在继续之前首先说明一下，下面为了描述简便，我们将假设 $m=|E|$，$n=|V|$。

定义5 CDS(F)问题是指：给定连通无向图 G 及测度函数 $f(n)$，要求图 G 的最小点集 $D\subseteq V$，满足：

1）可覆盖性：对任意 $v\in V$，或者 $v\in D$，或者存在 $u\in D$，使得 u 和 v 是 f-相邻的。

2）连通性：D 在 G 中是 f-连通的。

定理1 若测度函数 $f(n)\geqslant n-1$，则CDS(F)的优化问

*）本文工作得到科技部基金(No. 2001CCA03000)，国家自然科学基金(No. 60273045)，上海科学技术发展基金(No. 025115032)的支持。

马 俊 硕士，主要研究方向为计算复杂性理论，近似算法，参数复杂性；**朱 洪** 教授，博士生导师。

题是多项式时间可解问题。

证明：显然，任意取一个点都是图 G 的一个支配集。于是，图 G 的带测度函数 f 的最小连通支配集的大小为 1。 □

定理 2 若测度函数 $f(n)<n-1$，并且 $f(n)=\Theta(n)$，则 CDS(F)的优化问题是多项式时间可解问题。

证明：由 $f(n)=\Theta(n)$，则存在常数 c'，使得 $c'\geq n/f(n)$，又因为 $f(n)\geq 1$，从而存在常数 $c=2c'$，使得 $c\geq n/[f(n)]$。

下面我们证明，在这样的条件下，图 G 的带测度函数 f 的最小连通支配集的大小不会超过 c，从而我们可以在多项式时间内穷举所有小于等于 c 的点的子集来求得最优解。

首先假设 T 为图 G 的一个支撑树，显然 T 的一个连通支配集也必定是图 G 的一个连通支配集，于是只要能证明 T 有一个小于等于 c 的连通支配集就可以了。

求 T 的大小不超过 c 的连通支配集的算法如下：

1)初始化 T 的连通支配集的解为 $D=\phi$，任取一点 v 作为 T 的根。假设 $height(v)$ 为点 v 的高度，$L(u,v)$ 为点 u,v 之间的距离。

2)如果 $height(v)\leq[f(n)]$，则 $D=\{v\}$。算法结束。

3)令 $u=\text{argmax}_{k\in V}L(v,k)$。显然 $L(v,u)>[f(n)]$，于是在从树根 v 到 u 的路径上，存在一点 w，使得 $L(w,u)=[f(n)]$。令 $D=D\cup\{w\}$，从 T 中删去以 w 为根的子树(不包括 w)作为新的 T(因为 $L(w,u)=[f(n)]$，所以可以保证这一棵子树的所有结点都可以被 w 所覆盖)。转第二步。

从上面的算法可知，每次循环至少会从 T 中删掉 $[f(n)]$ 个点，同时往 D 中添加一个点，而 $n/[f(n)]\leq c$，于是 $|D|\leq c$。 □

下面将证明在某些特殊情况下 CDS(F)问题是一个 NP 难问题，从而我们只有寻找折衷办法(比如利用近似算法)来解决它。首先给出 NP 问题的相关定义。

定义 6 一个问题是 NP 完全问题，指这个问题是 NP 问题，且所有 NP 问题都能多项式时间归约到这个问题。

定义 7 如果说 η 是从判定问题 A 到判定问题 B 的多项式时间归约，是指 η 能以某种方式在多项式时间里把问题 A 的实例变换成问题 B 的实例，使得实例 x 是问题 A 的"是"实例当且仅当 $\eta(x)$ 是问题 B 的"是"实例。

定义 8 一个问题是 NP 难问题，指所有 NP 问题都能多项式时间归约到这个问题(或者说有一个 NP 完全问题能多项式时间归约到这个问题)，不论这个问题本身是否是 NP 问题。

定理 3 对于任意 $0<\varepsilon<1$，带测度函数 $f=n^{\varepsilon}$ 的连通支配集问题是 NP 难问题。

证明：已经知道顶点覆盖(VC)问题是 NP 难的[1]，下面将构造一个从它到带测度函数的连通支配集问题(CDS(F))的规约，从而证明 CDS(F)问题也是 NP 难的。

首先给出顶点覆盖的判定问题的实例定义，一个顶点覆盖问题的实例 $\langle G(V,E),K\rangle$，其中 $G(V,E)$ 是一个无向图，要求判定 G 中是否有小于等于 K 的顶点覆盖集覆盖所有边。由此构造带测度函数 $f=n^{\varepsilon}$ 的连通支配集实例 $\langle G'(V',E'),K\rangle=\eta(\langle G(V,E),K\rangle)$ 如下：

给定任意 $G(V,E)$，易知存在整数 δ 满足(证明见附录 A)：$1\leq\delta\leq(n+\delta m)^{\varepsilon}<\delta+1$，且 $\delta\leq 2n^{\frac{2\varepsilon}{1-\varepsilon}}$(也即 δ 是 n 的多项式大小)。

由 G 构造 $G'(V',E')$：首先对 E 中每一条边 e_i，添加一条长为 $\delta-1$ 的路径，并把这条路径分别与 e_i 的两端连接。即，对于 $e_i=(u,v)$，令 $V=V\cup\{u_i^1,u_i^2,\cdots,u_i^{\delta-1},u_i^{\delta}\}$，$E=E\cup\{(u,u_i^1),(v,u_i^1),(u_i^1,u_i^2),\cdots,(u_i^{\delta-1},u_i^{\delta})\}$。

其次，对于 V 中任意点对 u,v，如果 (u,v) 不属于 E，则把 (u,v) 添加到 E 中(即把原图 G 扩展成一个团)。

由上面描述的变换就有 $n'=|V'|=|V|+|E|\cdot\delta=n+\delta m$ 和 $m'=|E'|=n^2+m(\delta+1)$，测度函数 $f(n')=f(n+m\delta)$，于是 $1\leq\delta\leq(n+\delta m)^{\varepsilon}=(n')^{\varepsilon}=f(n')<\delta+1$。

例如，对于下面图 1 左所示的 5 元图，变换后如图 1 右所示。

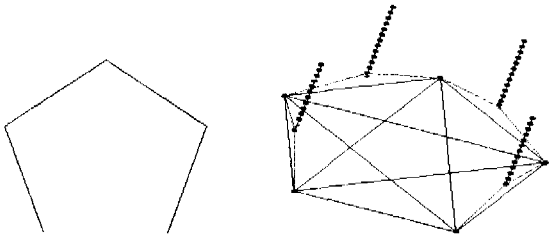

图 1

在继续证明之前先定义几个概念。定义 G' 中的由 G 中直接得到的点为基本点，G' 中新加的点定义为非基本点。

证明规约包含正反两方面的证明。首先我们证明 G 的顶点覆盖集 C 正好也是 G' 的带测度函数 f 的连通支配集 D'。这一点是显然的，因为在 G 中的任意两个点在 G' 中都是相邻的，而且由顶点覆盖集的定义，任意 $(u,v)\in E$，必有 $u\in C$ 或者 $v\in C$，也就是说对于一条新增的路径 $(u_i^1,u_i^2,\cdots,u_i^{\delta})$，必有 u 或 v 覆盖这条路径上所有的点(在 f-覆盖的意义下)。

其次，证明 G' 有小于等于 K 的带测度函数 $f=n^{\varepsilon}$ 的连通支配集 $D'\Rightarrow G$ 有小于等于 K 的顶点覆盖集 C。

两个基本点 u 和 v，如果 $(u,v)=e_i\in E$，定义 u,v 的导出点集 $I(u,v)=\{u_i^1,u_i^2,\cdots,u_i^{\delta-1},u_i^{\delta}\}$。

假设 D' 是 G' 的一个带测度函数 $f=n^{\varepsilon}$ 的连通支配集，我们分两种情况来证明：

1) 假设 D' 中的所有点都是基本点。

因为测度函数 $f=n^{\varepsilon}<\delta+1$，于是对任意 $(u,v)=e_i\in E$，u,v 中至少有一点属于 D'(否则无法 f-覆盖点 u_i^{δ})，于是 D' 也是 G 的一个顶点覆盖集。

2) 假设 D' 中存在非基本点。

假设 w 是 D' 中一个非基本点，$w\in I(u,v)$，$w\in D'$，因为能被 w 所 f-覆盖的点也都能被 u 所 f-覆盖，于是在 D' 中用 u 替换 w 并不会影响 D' 的覆盖性和连通性。

综上，证明了 VC 可以多项式时间规约到 CDS(F)，得出结论，对于测度函数 $f=n^{\varepsilon}(0<\varepsilon<1)$ 的情况，CDS(F)是 NP 难问题。

由定理 3 的证明过程可以看出，证明的关键在于 $[1,poly(n)]$ 在的范围内找到 $f(n+xm)-x$ 的零点($poly(n)$ 表示 n 的多项式大小)，并且要求这个零点出现在 $f(n+xm)-x$ 的递减阶段。按照这个思路，我们完全可以证明其他测度函数的 NP 完全性。

定理 4 带测度函数 $f=\lg n$ 的连通支配集问题是 NP 难问题。

证明：只需要找到 $\delta\leq poly(n)$ 使得 $1<\delta\leq\lg(n+m\delta)<\delta+1$ 即可，其他的证明同定理 3。

令函数 $y=\lg(n+xm)-x$，易得 $y'=\frac{m}{n+xm}-1$，y' 是一

个递减函数且 $y'(1)=\frac{m}{n+m}-1<0$，由此可知在 $[1,+\infty]$ 内 y 是一个递减函数。

当 $n_0+m\geqslant e^2$ 时，$y(1)\geqslant 2$，因此只需再证明在 $(1,poly(n))$ 内存在一点 x_{0-} 使得 $y(x_{0-})<0$，那么就可以推断在 $(1,poly(n))$ 内存在 y 的零点 x_0，也即 $\lg(n+x_0m)=x_0$，于是取 $\delta=[x_0]$ 即有 $2\leqslant\delta\leqslant\lg(n+m\delta)<\delta+1$。

首先，当 $2m\leqslant x$ 时，$\lg(n+xm)<\lg(2mx)\leqslant 2\lg(x)$；其次，当 $x\geqslant e^2$ 时，$2\lg x<x$。于是，取 $x_{0-}=\max\{2m,e^2\}$ 即可，显然 $x_{0-}\in(1,poly(n))$。证毕。 □

现在大家普遍认为 P≠NP，在这个假设下，NP 难优化问题不存在多项式时间的完全算法。因此我们下面将讨论这个问题的近似算法。

3 近似算法

命题 1 对于带测度函数 f 的连通支配集问题实例 $G(V,E)$，假设 G 的邻接矩阵为 M，计算矩阵 $M'=M^{f(n)}$，由 M' 生成新的图 $G'(V,E')$，其中 $(u,v)\in E'\Leftrightarrow M'(u,v)\neq 0$，那么 D 是 G 的带测度函数 f 的连通支配集当且仅当 D 是 G' 的带测度函数 $f'=1$ 的连通支配集。

证明：由邻接矩阵的定义易得。 □

由命题 1，首先由图 G 得到图 G'，计算 G' 的连通支配集 D，同时也就是 G 的带测度函数 f 的连通支配集。于是，可以直接利用连通支配集的近似算法得到带测度函数连通支配集的近似算法。目前已知的连通支配集的最好近似度为 $\mathrm{Ln}\Delta+3$[5]，于是我们可以直接得到下面的算法。

算法 1 带测度函数 $f(n)$ 的连通支配集算法：

输入：图 $G(V,E)$ 及测度函数 $f(n)$

输出：图 G 的带测度函数 $f(n)$ 的连通支配集 D。

步骤：

(1) 计算 G 的邻矩阵 M 及 $M'=M^{f(n)}$。

(2) 由 M' 构造图 $G'(V,E')$，满足：

$(u,v)\in E'\Leftrightarrow M'(u,v)\neq 0$。

(3) 对图 G' 应用连通支配集的近似算法，求 G' 的连通支配集的解 D'。

(4) 输出 D' 作为 G 的带测度函数 $f(n)$ 的连通支配集的解。

由命题 1 可知，算法 1 求出的解确实是满足 G 的带测度函数 $f(n)$ 的连通支配集，并且仍然保持 $\mathrm{Ln}\Delta+3$ 的近似度。

定理 5 对于带测度函数的连通支配集问题存在近似度为 $\mathrm{Ln}\Delta+3$ 的近似算法。

总结 本文提出了带测度函数的连通支配集问题，这个问题目前还没有人研究过。本文对多种情况研究了它的 NP 完全性，并给出近似算法。下一步的研究课题是对大于 n^{ϵ} 的测度函数研究以及把测度函数的情况推广到其他支配集的变种问题。此外，从实际应用的角度考虑，对测度函数的不同定义方式也是可研究的课题。

参考文献

1 Garey M R, Johnson D S. Computers and Intractability: A guide to the theory of NP-completeness, Freeman, San Francisco, 1978

2 Cormen T, Leiserson C, Rivest R. Introduction to Algorithms (Second Edition). the MIT Press, 2002

3 Vazirani V V. Appromixation Algorithms. Springer, 2001

4 Chen Ning, Meng Jie, Zhu Hong. Approximation for Dominating Set Problem with Measure Functions

5 Guha S, Khuller S. Approximation Algorithms for Connected Dominating Sets. In: European Symposium on Algorithms, 1996

6 Johnson D S. Approximation algorithms for combinatorial problems, J. Comput. System Sci., 1974

7 Raz R, Safra S. A sub-constant error-probability low-degree test, and sub-constant error-probability PCP characterization of NP. In: Proc. 29th Ann. ACM Symp. On Theory of Comp., ACM, 1997

· 附录 A

证明，对正整数 $m,n(1<n-1\leqslant m\leqslant n^2)$，存在正整数 δ 满足：

$1\leqslant\delta\leqslant(n+\delta m)^{\epsilon}<\delta+1$，且 $\delta\leqslant 2n^{\frac{2\epsilon}{1-\epsilon}}$。

证明：令函数 $y=(n+xm)^{\epsilon}-x$，易得 $y'=\frac{m\epsilon}{(n+xm)^{1-\epsilon}}-1$，$y'$ 是一个递减函数且有零点为 $\frac{(m\epsilon)^{\frac{1}{1-\epsilon}}-n}{m}$，由此可知 y 是一个凸函数。

又因为 $y(1)=(n+m)^{\epsilon}-1>0$，于是在 $(1,+\infty)$ 内有且只有 y 的一个零点，且这个零点必出现在 y 的下降阶段。记这个点为 x_0，令 $\delta=[x_0]$，于是 $y(\delta)\geqslant 0$ 也即 $\delta\leqslant(n+\delta m)^{\epsilon}$。再由 $(n+\delta m)^{\epsilon}\leqslant(n+x_0m)^{\epsilon}=x_0<\delta+1$ 可得证。

接着要证明 $\delta\leqslant 2n^{\frac{2\epsilon}{1-\epsilon}}$，只需证明 $x_0\leqslant 2n^{\frac{2\epsilon}{1-\epsilon}}$ 即可。

由于 $(n+x_0m)^{\epsilon}=x_0$，变形得 $x_0^{\frac{1}{\delta}}-mx_0=n$，由 $x_0\geqslant 1$ 可推出 $x_0^{\frac{1}{\delta}-1}-m\leqslant n$，于是 $x_0\leqslant(n+m)^{\frac{\epsilon}{1-\epsilon}}\leqslant 2n^{\frac{2\epsilon}{1-\epsilon}}$。

（上接第 204 页）

5 朱玉贤，等．现代分子生物学[M]．北京：高等教育出版社，1997

6 Liu Q, Liman W. Dna computing on surfaces [J]. Nature, 2000 (403): 175～179

7 Rooβ D, Wagner K W. On the Power of DNA- Computing [J]. Information and Computation, 1996, 131(2): 95～109

8 Roweis, Sam, Erik W, et al. A sticker based model for DNA computation [J]. Journal of Computational Biology, 1998, 5(4): 615～629

9 Wu H Y. An improved surface-based method for DNA computation [J]. Biosystems, 2001, 59(1): 1～5

10 Gatterdam R W. Splicing systems and regularity [J]. International Journal of Computer Mathematics, 1998(31): 63～67

11 Gibbons A, et al. DNA Computing [J]. Current Opinion in Biotechnology, 1997, 8(1): 103～106

12 Head T, et al. Computing with DNA by operation on plasmids[J]. Biosystems, 2000, 57(2): 87～93

13 Maley C, et al. DNA computation: Theory, practice and prospects [J]. Evolutionary Computation, 1998, 6(3): 201～229

14 Sakakibara Y. DNA computers: A new computing paradigm [J]. Journal of Photopolymer Science and Technology, 1998, 11(4): 681～686

15 Wu H Y. An improved surface-based method for DNA computation [J]. Biosystems, 2001, 59(1): 1～5

16 Takenaka Y, Hashimoto A. A proposal of DNA computing on beads and its application to SAT problem. In: 7th International Meeting on DNA Based Computers, 2001. 331～339

17 邹海明，等．形式语言、自动机和语法分析[M]．武昌：华中工学院出版社，1985

18 Kari L, et al. At the Crossroads of DNA Computing and Formal Languages: Characterizing Recursively Enumerable Languages by Insertion-Deletion Systems [A]. In: Proc. of 3rd DIMACS Workshop on DNA- Based Computers, Philadel-phia, June 1997. 318～333