A-012

Minimum Hub Cover Set 問題の近似困難性
Hardness of Approximating the Minimum Hub Cover Set Problem

鎌尾 祐汰 *　　下薗 真一 †
Yuta Kamao　　Shinichi Shimozono

概 要

単純無向グラフ $G=(V,E)$ の $S\subseteq V$ がすべての辺について, その端点のうち少なくとも 1 つを含むかまたは両端点に隣接する頂点を少なくとも 1 つ含むとき S を hub cover set という. Minimum hub cover set 問題は, 与えられたグラフ $G=(V,E)$ の hub cover set で要素数が最小なものを見つける問題である. この問題に, Minimum hitting set 問題からの近似率保存還元を与え, P=NP でない限り近似率が定数倍以内となる多項式時間アルゴリズムが存在しないこと, また NP が DTIME($n^{\log\log n}$) に含まれない限り, 任意の $\epsilon>0$ に対して近似率 $(1-\varepsilon)\ln\frac{|E|}{|V|}$ 以内で解ける多項式時間アルゴリズムは存在しないことを示す.

1　はじめに

本稿で取り上げるのは Minimum hub cover set 問題の近似困難性についてである. 与えられた単純無向グラフ $G=(V,E)$ に対し, すべての辺 $(u,v)\in E$ について, その端点 (u,v) のうち少なくとも 1 つを含むか, u と v の両端点に隣接する頂点 r を含む集合 $S\subseteq V$ で要素数が最小となるものを見つける問題を Minimum hub cover set 問題という.

この問題を効率よく解くことができるアルゴリズムは, グラフデータベースのクエリ処理において大幅なコスト削減につながると主張されている [1]. さらに, 物流ネットワークの輸送コストを減らすことを目的とした中継地点の位置選択に利用することができる [2].

Minimum hub cover set 問題とよく似た問題として, 最小頂点被覆問題がある. この問題は, 与えられた単純無向グラフ $G=(V,E)$ に対し, すべての辺 $(u,v)\in E$ について, その端点 (u,v) のうち少なくとも 1 つを含む集合 $S\subseteq V$ で要素数が最小となるものを見つける問題である. 3 角形の構造を持たないグラフに対しては, これらの問題は同等である.

文献 [3] では, k 個の外平面的グラフに分解できるグラフ対して, $\frac{k+1}{k}$-近似アルゴリズムが提案されている. さらに, 文献 [4] では, 任意のグラフに対して, Minimum hub cover set 問題は NP 困難であることが示されており, $\ln|E|$ の近似率を持つ多項式時間アルゴリズムが提案されている. しかし, 多項式時間近似アルゴリズムで近似率が $\ln|E|$ より小さい, たとえば定数倍の近似率を達成できる多項式時間アルゴリズムが存在するかどうかはわかっていなかった.

本稿では, hitting set 問題と呼ばれる問題から hub cover set 問題への多項式時間近似率保存還元を与え, hitting set 問題の近似困難性を用い, hub cover set 問題について, P=NP でない限り, 近似率が定数倍以内となる多項式時間アルゴリズムが存在しないこと, また NP が DTIME($n^{\log\log n}$) に含まれない限り, $\epsilon>0$ に対して近似率 $(1-\varepsilon)\ln\frac{|E|}{|V|}$ 以内で解ける多項式時間アルゴリズムは存在しないことを示す.

2 節では, 本稿で取り上げる 2 つの組合せ最適化問題（hitting set 問題, hub cover set 問題）の定義とともに, 組合せ最適化問題, 近似アルゴリズム, 多項式時間近似率保存還元などの定義を述べる. 3 節では, hitting set 問題から hub cover set 問題への多項式時間近似率保存還元を与え, その証明を述べる. 合わせて, hitting set 問題についての既知の結果から導き出される, hub cover set 問題の性質を記述する. 4 節では結果のまとめと今後の課題につい

* 九州工業大学大学院 情報工学府 情報創成工学専攻
† 九州工業大学大学院情報工学研究院

て述べる．

2 準備

この節では本稿の議論において，必要な知識をいくつか示す．まず，Minimum hub cover set 問題と Minimum hitting set 問題の定義を述べる．

定義 2.1 (**Minimum hub cover set 問題**)**.** 単純無向グラフ $G = (V, E)$ の $S \subseteq V$ が，すべての辺について以下の *3* つの条件のうち，少なくとも *1* つを満たすならば，S は G の *hub cover set* であるという．

1. $u \in S$.

2. $v \in S$.

3. $r \in S$ かつ $(r, u) \in E$ かつ $(r, v) \in E$.

ある辺 e に対して *hub cover set* S に含まれる $v \in S$ がその端点，あるいは両端点の隣接点であるとき，v は e をカバーするという．*minimum hub cover set* 問題は与えられたグラフ $G = (V, E)$ の *hub cover set* で要素数が最小のものを見つける問題である．

文献 [4] の中で，minimum hub cover set 問題は NP 困難であること，近似率 $\ln |E|$ の多項式時間近似アルゴリズムが存在することが述べられている．

定義 2.2 (**Minimum hitting set 問題**)**.** 有限集合 U, $C = \{c_1, c_2, \ldots, c_n\}$, $c_i \subseteq U$ とした (U, C) の $S' \subseteq U$ がすべての i について，c_i の要素の中で少なくとも *1* つを含むとき，S' を *hitting set* であるという．*minimum hitting set* 問題は与えられた (U, C) の *hitting set* で要素数が最小のものを見つける問題である．

この問題は minimum set cover 問題と同等な問題であることが [5] によって示されており，P=NP でない限り，近似率が定数倍以内となる多項式時間アルゴリズムは存在しないことが [6] によって示されている．さらに，NP $\subset$ DTIME($n^{\log\log n}$) でない限り，ε > 0 に対して $(1-$ ε$) \ln |U|$ 以内で解ける多項式時間アルゴリズムは存在しないことが [7] によって示されている．

次に，一般的な組合せ最適化問題と近似アルゴリズムの定義を述べる．

定義 2.3 (**組合せ最適化問題と近似アルゴリズム**)**.** 組合せ最適化問題 P は，有限アルファベット上で適切な符号化によって表現される *(1)* インスタンスの集合 $Inst_{\mathrm{P}}$, *(2)* 任意のインスタンス $x \in Inst_{\mathrm{P}}$ について定まる実行可能解の集合 Sol_{P}, *(3)* x とその解 s に非負整数（コストまたは利得）を与える測度関数 $m_{\mathrm{P},x}(s) : Sol_{\mathrm{P}}(x) \to \mathbb{Z}^+$, そして *(4)* 最大化または最小化の指定，の組である．

最適化問題 P のインスタンス x の実行可能解のうち，最小化問題においては最も測度（コスト）が小さい解を，最大化問題においては最も測度（利得）が大きい解を最適解という．

以後，本稿では最小化問題のみを扱う．

定義 2.4. ある多項式時間アルゴリズム A が最小化問題 P の任意のインスタンス x に対して $|x|$ の多項式時間内に解 $A(x) \in Sol_{\mathrm{P}}(x)$ を出力するとする．任意のインスタンス $x \in Inst_{\mathrm{P}}$ に対して A の出力する解 $s = A(x)$ のコスト $m_{\mathrm{P},x}(s)$ が x の最適解 s^* のコストに対する比が高々 $r > 1$ でおさえられる．つまり

$$\frac{m_{\mathrm{P},x}(s)}{m_{\mathrm{P},x}(s^*)} < r \tag{1}$$

となるとき，A を最小化問題 P の多項式時間 r-近似アルゴリズムであるという．

最後に APX-還元の定義を述べる．

定義 2.5. 組合せ最適化問題 A から組合せ最適化問題 B への *APX*-還元とは，以下を満たす関数の組 (f, g, δ) である：

1. A の任意のインスタンス $x \in Inst_{\mathrm{A}}$ から B のインスタンス $f(x) \in Inst_{\mathrm{B}}$ への多項式時間変換 $f : Inst_{\mathrm{A}} \to Inst_{\mathrm{B}}$.

2. $f(x)$ の解 $s \in Sol_B(f(x))$ から x の解への多項式時間変換 $g_{\mathrm{A},f}(x, s) : Sol_{\mathrm{B}}(f(x)) \to Sol_{\mathrm{A}}(x)$.

3. $f(x)$ の解 $s \in S_{\mathrm{B}}(f(x))$ が $(1 + \delta(\epsilon))$-近似であるとき $g(x, s)$ が $(1 + \epsilon)$-近似であることを保証する関数 δ.

3 Minimum hub cover set 問題の近似困難性

定理 1. *Minimum hitting set* 問題から *Minimum hub cover set* 問題へ $\delta(\epsilon) = \epsilon$ で *APX*-還元可能である．

f は hitting set のインスタンス (U, C) から hub cover set のインスタンス $G = (V, E)$ への変換を以下のように行う．$C = \{c_1, c_2, \ldots, c_n\}$, $c_i \subseteq U$ である．

1. $V = U \cup \{\tilde{c_1}, \tilde{c_2}, \ldots, \tilde{c_n}\}$ とする．
2. c_i ごとに $V_i = c_i \cup \{\tilde{c_i}\}$ を頂点集合とし, V_i の全ての頂点間に辺を作り, その辺集合を E_i としたとき, $E = E_1 \cup E_2 \cup \cdots \cup E_n$ とする．このとき, $G_i = (V_i, E_i)$ は完全グラフである．

性質 1. 任意の $S \subseteq V$ について, すべての完全グラフ $G_i (1 \leq i \leq n)$ が $V_i \cap S \neq \phi \Leftrightarrow S$ は *hub cover set* である．

証明. 完全グラフ G_i の頂点集合 V_i のある頂点 u が S に含まれているとき, E_i に含まれる辺は, 端点のうち 1 つが u か, 両端点が u に隣接するので, G_i の辺はすべてカバーされている．これがすべての i に対していえるので, すべての完全グラフ G_i について $V_i \cap S \neq \phi$ ならば, S は hub cover set である．

次に, S が hub cover set であるならば, すべての完全グラフ G_i について $V_i \cap S \neq \phi$ であることを示す．完全グラフ G_i には, 頂点 $\tilde{c_i}$ を端点に持つ辺 e が存在する．S は, 辺 e の端点のうち少なくとも 1 つを含むか, その両端点と隣接する頂点を含む必要がある．しかし, $\tilde{c_i} \notin V_j \, (j \neq i)$ なので, それらの頂点は G_i 中にのみ存在する．よって, S が hub cover set であるならば, すべての完全グラフ G_i について $V_i \cap S \neq \phi$ である．

以上より性質 1 は成り立つ．

□

補題 1. $\tilde{c_i}$ を含む S から $\tilde{c_i}$ を含まない *hub cover set* S' へ要素数を増加させることなく変換可能である．

$\tilde{c_i}$ は (U, C) の U には含まれないので, 変換 g を以下のように行う．

1. S を走査し, $\tilde{c_i}$ を見つけると S から $\tilde{c_i}$ を除き, c_i の要素が 1 つ以上 S に含まれていれば, ステップ 2 に移る．そうでなければ, c_i の要素から $\tilde{c_i}$ を除く任意の要素を S に加え, ステップ 2 に移る．
2. S から $\tilde{c_i}$ がなくなると終了．そうでなければ, ステップ 1 に戻る．

証明. 変換 g では, $\tilde{c_i}$ を 1 つ除くごとに要素を 1 つ追加するか何もしないかなので, 集合の要素数が増加しないのは明らかである．また, $\tilde{c_i}$ を除いたとき $V_i \cap S = \phi$ ならば $\tilde{c_i}$ 以外の V_i の要素を S に加えるので, 変換後の S' はすべての G_i に対し, $V_i \cap S \neq \phi$ である．よって, 性質 1 より S' は hub cover set である．□

補題 2. S' は G の *hub cover set* を g によって変換した集合である.$\Leftrightarrow S'$ は (U, C) の *hitting set* である．

証明. まず, S' が G の hub cover set ならば, S' は (U, C) の hitting set であることを示す．S' は G の hub cover set であるが (U, C) の hitting set ではないと仮定する．このとき, $S' \cap c_i = \phi$ である c_i が存在する．さらに, $\tilde{c_i} \notin S'$ であるので, $V_i \cap S' = \phi$ である．これは, 性質 1 より S' が hub cover set であることに矛盾する．よって, S' が G の hub cover set ならば, S' は (U, C) の hitting set である．

次に, S' が (U, C) の hitting set ならば, S' は G の hub cover set であることを示す．S' は hitting set であるが hub cover set ではないと仮定する．性質 1 の逆より S' が hub cover set でないならば, $V_i \cap S' = \phi$ を満たす G_i が存在する． このとき, $S' \cap c_i = \phi$ であり, S' が hitting set であることに矛盾する．よって, S' が (U, C) の hitting set ならば, S' は G の hub cover set である．

以上より, G の hub cover set を g によって変換した S' は (U, C) の hitting set である．□

補題 2 より, hitting set $H \subseteq U$ と最小の hitting set H^*, g によって変換された hub cover set $C \subseteq V$ と最小の hub cover set C^* について,

$$\frac{|H|}{|H^*|} = \frac{|C|}{|C^*|} \tag{2}$$

であり, ある値 $\epsilon > 0$ について $\frac{|H|}{|H^*|} < 1 + \epsilon$ を満たすならば, $\frac{|C|}{|C^*|} < 1 + \epsilon$ であるので, Minimum hitting set 問題から Minimum hub cover set 問題へ $\delta(\epsilon)=\epsilon$ での APX-還元が可能である.

系 1. *Minimum hub cover set* 問題は $P{=}NP$ でない限り, 近似率が定数倍となる多項式時間アルゴリズムは存在しない.

系 2. *Minimum hub cover set* 問題について, $NP \subset DTIME(n^{\log\log n})$ でない限り, ε >0 に対して $(1-$ ε $)\ln\frac{|E|}{|V|}$ 以内で解ける多項式時間アルゴリズムは存在しない.

Hitting set において NP $\subset$ DTIME$(n^{\log\log n})$ でない限り, 任意の ε に対して $(1-$ ε $)\ln|U|$ 以内で解ける多項式時間アルゴリズムは存在しないことが示されている．ここで, hitting set のインスタンスである U を還元の際に, 生成した hub cover set のインスタンスである V, E で評価することを考える.

Hitting set のインスタンス c_i ごとに $V_i = c_i \cup \{\tilde{c_i}\}$ を頂点集合とし, 全ての頂点間に辺を作り, その辺集合を E_i としたとき, $E_1 \cup E_2 \cup \cdots \cup E_n = E$ となるので, すべての c_i の要素数が最大, つまり $c_1 = c_2 = \cdots = c_i = U$ であるとき, $|E|$ は最大になり,

$$\frac{|U|(|U|-1)}{2} + |C| \cdot |U| \geqq |E| \tag{3}$$

が成り立つ. さらに, $V = U \cup \{\tilde{c_1}, \tilde{c_2}, \ldots, \tilde{c_n}\}$ であるので,

$$|V| = |U| + |C| \tag{4}$$

である．(3) 式を簡単にした

$$|U|^2 + |C| \cdot |U| \geqq E| \tag{5}$$

と (4) 式より,

$$|U| \geqq \frac{|E|}{|V|} \tag{6}$$

となる.

よって, 系 2 は成り立つ.

4 まとめと今後の課題

Hitting set 問題から hub cover set 問題へ近似率保存還元を行うことで, NP $\subset$ DTIME$(n^{\log\log n})$ でない限り, ε >0 に対して $(1-$ ε $)\ln\frac{|E|}{|V|}$ 以内で解ける多項式時間アルゴリズムは存在しないことを示した.

今後は, hub cover set 問題について [4] によって示された近似率 $\ln|E|$ と, 本稿で示した近似率の限界値 $\ln\frac{|E|}{|V|}$ にはどれほどの差があるのかを複数の問題例に対して, 計算実験をすることで確認する.

具体的には, 問題例 $G=(V,E)$ に対して, $\ln|E|$ 近似アルゴリズムが求める hub cover set S と最小の hub cover set S^* を求め, $\frac{|S|}{|S^*|}$ と $\ln\frac{|E|}{|V|}$ を比較する.

この計算実験の結果によって, 本稿で示した近似率の下限 $\ln\frac{|E|}{|V|}$ をもっと上げることはできないかを考える.

参考文献

[1] Yelbay B, Birbil S, Bulbul K, Jamil H, Trade-offs Computing Minimum Hub Cover toward Optimized Graph Query Processing, arXiv 1311.1626, Cornell University (2013).

[2] Horst W, Hamacher, Tanja Meyer, Hub Cover and Hub Center Problems, Report in Wirtschaftsmathematik (2006).

[3] Yelbay B, Birbil S, Bulbul K, Jamil H, Approximating the Minimum Hub Cover Problem on Planar Graphs, *Optimization Letters*, vol.10, 33-45 (2016).

[4] Joel A.Trejo-Sanchez, Candelaria E.Sansores-Perez, Jesus Garcia-Diaz, Jose Alberto Fernandez-Zepeda, Approximation Algorithm for the Minimum Hub Cover Set Problem, Digital Object Identifier 10.1109/ACCESS.2022.3173615 (2022).

[5] Ausiello G, D'Atri A, Protasi M, Structure preserving reductions among convex optimization problems, *Journal of Computer and System Sciences*, vol.21, 136-153 (1980).

[6] Lund C, Yannakakis M, On the Hardness of Approximating Minimization Problems, *Joural of the ACM*, vol.41, No.5, 960-981 (1994).

[7] Uriel Feige, A Threshold of $\ln n$ for Approximating Set Cover, *Journal of the ACM*, vol.45, No.4, 634-652 (1998).