閾値グラフに対する標的集合選択問題

山下 智大[1]　小野 廣隆[2]

概要：標的集合選択問題 (Target Set Selection, TSS) とは, 無向グラフ $G = (V, E)$ と各頂点 $v \in V$ の閾値を定義する関数 $thr(v)$ が与えられたとき, 最小サイズの標的集合 $S \subseteq V$ を求める問題である. 初期状態において, ある頂点集合 $S \subseteq V$ に属する頂点は *active*, それ以外の頂点は *inactive* という状態を持っているが, [v の隣接点のうち少なくとも $thr(v)$ 個の隣接点が *active* であるとき, v 自身も *active* に変化する] という動的なプロセスを通して状態が拡散していく. このプロセスで最終的にすべての頂点が *active* になる S の中で, $|S|$ が最小のものを標的集合と呼ぶ. TSS は一般のグラフに対して NP 困難であることが知られているが, 本論文では閾値にある構造を仮定した場合には閾値グラフに対する TSS は多項式時間で解けることを示した.

Target Set Selection for Threshold Graph

Tomohiro Yamashita[1]　Hirotaka Ono[2]

Abstract: The NP-hard graph problem Target Set Selection (TSS) is defined as follows. Given an undirected graph $G = (V, E)$, each vertex $v \in V$ has a positive integer threshold value $thr(v)$, the task is to select a minimum cardinality target set $S \subseteq V$. A vertex set S is called a target set if it activates all vertices in G in course of a dynamic process where a vertex v gets activated when at least $thr(v)$ many of its neighboring vertices are activated. In this paper, we focus on threshold graph which is assumed some structure of the threshold values and show TSS can be solved in polynomial time.

1. はじめに

標的集合選択問題 (Target Set Selection, TSS) とは, 無向グラフ $G = (V, E)$ と各頂点 $v \in V$ の閾値を定義する関数 $thr(v)$ が与えられたとき, 最小サイズの標的集合 $S \subseteq V$ を求める問題である. 初期状態において, 与えられたグラフ $G = (V, E)$ の各頂点 $v \in V$ は *inactive* である. ある頂点集合 $S \subseteq V$ の各頂点 $v \in S$ を *active* にしたとき, *inactive* である頂点 $v \in V \setminus S$ に対して, [v の隣接点のうち少なくとも $thr(v)$ 個の隣接点が *active* であるとき, v 自身も *active* に変化する] という動的なプロセスを考える. ここで頂点集合 $S \subseteq V$ が, このプロセスによって, 任意の頂点 $v \in V$ の少なくとも $thr(v)$ 個の隣接点が, 動的なプロセスのある時点で *active* になり (つまり, その次の時点で v が *active* になり), 最終的にはすべての頂点が *active* になる場合に標的集合と呼ぶ.

標的集合選択問題は一般のグラフに対しては NP 困難であることが知られている ([1], [2], [4]). また, 直径が 2 のグラフに対して NP 困難 [4], 直径を 2 に限定したスプリットグラフに対しても NP 困難であることが分かっている [4]. しかし対象とするグラフクラスをうまく限定すると速く解けることがあることも知られている. 具体的には, 完全グラフ (直径が 1 であるグラフ)[4], 木グラフ [2], ブロックカクタスグラフ [3], 全点で閾値が 2 以下である弦グラフ [3] などは線形時間で厳密解を求められることが知られている.

本論文では, 閾値グラフに焦点を当てる. 閾値グラフは, 以下の規則に従って生成される.
(1) n 点の頂点 $v_1, v_2, \ldots, v_n$ を用意.
(2) 頂点 v_i に対して, 重み w_i を割り当てる.
(3) 全ての (i, j) の組に対して, $w_i + w_j \geq \theta$ である場合にのみ, v_i と v_j を辺で結ぶという操作をする. ただし, θ は閾

1　九州大学大学院経済学府経済工学専攻
Department of Economic Engineering, Kyushu University
2　九州大学大学院経済学研究院経済工学部門
Department of Economic Engineering, Kyushu Universuty

値である. このように生成される閾値グラフはその構成方法から, クリーク (部分グラフが完全グラフ) に独立点 (頂点間に辺を持たない頂点集合) が層的に接続しているように図示することができる (図 1).

適当な分布の下で頂点に重みを割り当てる形で生成した閾値グラフは, 現実の複雑なネットワークの研究でしばしば取り上げられるスケールフリー性, と呼ばれる性質を持つことが知られている. 例えば, ウェブページ間のリンク構造を見ると, 少数の有名なページが多くのリンクを集めている一方, その他の多くのページはわずかなページとしか結ばれていない. このように, ネットワークにおいて点の次数が冪乗則に従うことをスケールフリー性という. 複雑ネットワークが必ずしも閾値グラフとして表されるわけではないが, このような現実的なネットワークとの関連性からも閾値グラフは注目されている.

閾値グラフは, 完全グラフの一般化でありスプリットグラフの特殊ケースである. 標的集合選択問題は, スプリットグラフに対しては NP 困難, 完全グラフに対しては線形時間で解けることが知られているが, 閾値グラフについては未知である. 本論文では, 各頂点の閾値にある構造を持たせた閾値グラフの特殊ケースに対しては, 閾値グラフの n 頂点のクリーク部分の標的集合のサイズを k としたとき, $O(n \log k)$ の計算時間で厳密解を求めることができることを示す.

本章の構成は以下の通りである. 2 章で対象とする問題の設定を行い, 3 章でアルゴリズムの正当性の証明の準備, 4 章でアルゴリズムとその正当性を見る.

2. 設定

閾値グラフは, クリークと独立集合で構成される. また, クリークに独立点が層的に接続しているように書き表せる. クリークは次数 (頂点に接続している辺の本数) の大きいものから右から並べ, 独立集合は次数が大きいものから右から並べることとする (図 1). このとき, クリークの点を右から $v_1, v_2, \ldots, v_n$ の n 頂点とし, 独立集合の点を右から $u_1, u_2, \ldots, u_m$ の m 頂点とする. また, 独立点 u_i $(i \in \{1, 2, \ldots, m\})$ が接続する点のうち, 最も次数の小さいもの (最も左にあるもの) を v_{u_i} とする. u_i は, $v_1, v_2, \ldots, v_{u_i}$ の点と接続している. 以上は, 閾値グラフに対しては一般性を失うことなく設定できる. また, 頂点 v の次数を表す関数を $deg(v)$ とすると, $thr(v) > deg(v)$ のとき, 動的プロセスを通して頂点 v が $active$ になることはないので, $thr(v) \le deg(v)$ とする. 同様に $thr(v) = 0$ の点は $active$ になるので, $thr(v) \ge 1$ とする. 本論文ではさらに, クリークに対して, $thr(v_1) \ge thr(v_2) \ge \ldots \ge thr(v_n)$ という設定を加える.

本論文の目標は次の定理を, アルゴリズムを作り構成的に証明することである.

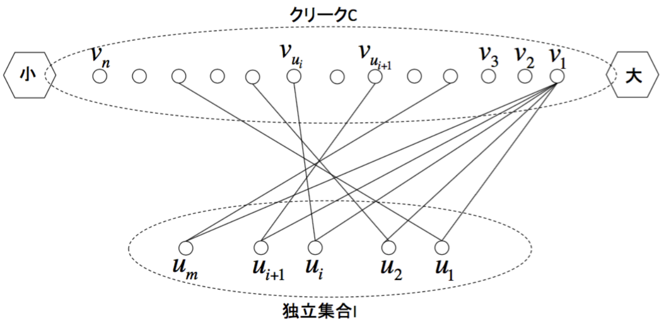

図 1 閾値グラフと記号の定義

定理 2.1 n 頂点 $v_1, v_2, \ldots, v_n$ のクリーク C と, m 頂点 $u_1, u_2, \ldots, u_m$ の独立集合 I で構成される閾値グラフに対しての標的集合選択問題は, $thr(v_1) \ge thr(v_2) \ge \ldots \ge thr(v_n)$ を満たすならば, C の標的集合を S とし $|S| = k$ とすると, $O(n \log k)$ 時間で解ける.

3. 証明の準備

定理を証明するために, まず補題を 4 つ証明する.

補題 3.1 クリーク C の標的集合を S $(|S| = k)$ とし, $thr(v_1) \ge thr(v_2) \ge \ldots \ge thr(v_n)$ を満たすとすると, $v_1, v_2, \ldots, v_k$ も C の標的集合である.

証明. $S = \{v_1, v_2, \ldots, v_k\}$ でない場合を考える. このとき, $v_{k+1}, v_{k+2}, \ldots, v_n$ の中の少なくとも一つの点が標的集合 S に含まれていることになる. その点を v とする. $v_1, v_2, \ldots, v_k$ の中で標的集合に含まれていない点を一つ選び, その点を v' とする. 点 v を点 v' に変えても標的集合であることを示す (図 2). まず, $S \leftarrow S \setminus \{v\} \cup \{v'\}$ としたとき, $thr(v)$ より小さい点を v^- とすると, v^- にとって v から受けていた影響が v' に変わるだけなので, 変わらず $active$ になる. また, 更新された S で v が $active$ になる場合, 更新前の S に点 v' が追加されているだけなので, 閾値 $thr(v)$ 以上を持つ点全体に拡散する. よって, 更新された S でも点 v が $active$ になることを確認すれば良い. $thr(v) - 1$ 以下の閾値を持つ点のうち, 最大の閾値を持つ点の集合を V^-_{max} とし, 閾値が t であるとする. 更新された S でも V^-_{max} に含まれる点は $active$ になる. $t + |V^-_{max}| < thr(v)$ と仮定すると, 更新前の S において thr(v) 以上の点が $active$ にならず, S が標的集合であることに反する. よって, $t + |V^-_{max}| \ge thr(v)$ であるので, 更新された S で, 点 v は $active$ になる.

以上より, $S \leftarrow S \setminus \{v\} \cup \{v'\}$ とした場合でも S が標的集合であると示された. この更新を繰り返し実行すると, $v_1, v_2, \ldots, v_k$ も C の標的集合であることが分かる.

補題 3.2 クリーク C の標的集合を S とすると, S はグラフ全体に拡散する.

証明. S はクリーク C の標的集合であるので, 任意の頂点

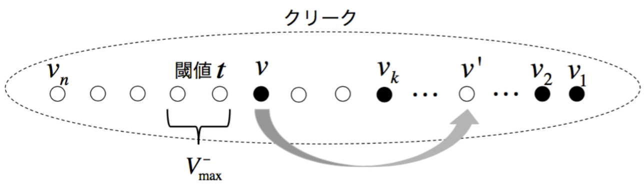

図 2

$v \in C$ は $active$ になる．C の全ての頂点が $active$ になった時点で，独立集合 I の点 u が $inactive$ であるとする．条件より $thr(u) \leq deg(u)$ であり，u は C にのみ接続していることより，I の頂点で $inactive$ なものは全て $active$ になる．

補題 3.3 グラフ全体の標的集合を S_1 とすると，S_1 を用いて，クリーク C の頂点のみを含み，全体の標的集合である S_2 を構成できる．

証明．クリーク C の頂点のみを含む全体の標的集合 S_2 を S_1 から構成する．S_1 が独立集合 I の頂点を含まない場合，$S_1 = S_2$ とすれば良い．S_1 が独立集合 I の頂点を含む場合，任意の頂点 $u \in I \cap S_1$ に対して，u をクリーク C の点で S_1 に含まれていない頂点に変更することを考える．まず，u の隣接点の集合を $N(u)$ とすると，標的集合の最適性より $N(u)$ に含まれる頂点全てが S_1 に含まれているということはない．つまり，$|N(u) \setminus S_1| \geq 1$ である．$N(u) \setminus S_1$ に含まれる頂点の中で最後に $active$ になる頂点を v としたとき，u を S_1 から削除して v を追加する，つまり $S_1 \leftarrow S_1 \setminus \{u\} \cup \{v\}$ と更新する．このとき更新された S_1 もグラフ全体に対する標的集合であることを示す．まず，頂点 u を除き頂点 v を追加しているので標的集合の最小性は保存される．次に，更新された S_1 もグラフ全体に拡散することを示す．頂点 u を標的集合から除いたときに直接の影響を受けるのは，頂点 u 自身と，$N(u) \setminus \{v\}$ の各点であるが，任意の $v' \in N(u) \setminus \{v\}$ について，条件より，頂点 v' は頂点 v が $active$ になる前に $active$ になるので，点 u の代わりに点 v を標的集合に入れても点 v_u は $active$ になる．つまり，$N(u) \setminus \{v\}$ に含まれる点は全て $active$ になり点 v は標的集合に入れているため，$N(u)$ に含まれる点は全て $active$ になる．$thr(u) \leq deg(u)$ であるので，点 u も $active$ になる．よって，S_1 を $S_1 \setminus \{u\} \cup \{v\}$ と更新したとき，更新された S_1 は標的集合である．この更新を，$S_1 \cap I = \emptyset$ となるまで実行し，そのときの S_1 を S_2 とすれば良い．

補題 3.4 グラフ全体の標的集合 S_2 を用いて，全体の標的集合である S_3 を構成できる．ただし，S_3 はクリーク C の頂点のみを含み，任意の $v \in S_3$, $v' \in C \setminus S_3$ に対して，$thr(v) < thr(v')$ となるものが存在しないものとする．

証明．$v \in S_2$, $v' \in C \setminus S_2$ に対して，$thr(v) < thr(v')$ となるものが存在しない場合は $S_2 = S_3$ であるので，そのような点 v, v' が存在する場合を考える．そのような点 v が $thr(v_{u_i}) \leq thr(v) \leq thr(v_{u_{i+1}})$ を満たす位置にあるとする．このとき，点 v を標的集合から除き，標的集合に含まれておらず $thr(v)$ 以上の閾値を持つ点のうち最も閾値の小さい点 v' を標的集合に追加することを考える．つまり，$S_2 \leftarrow S_2 \setminus \{v\} \cup \{v'\}$ と更新する．この更新された S_2 も標的集合であることを示す．まず，頂点 v を除き頂点 v' を追加しているので標的集合の最小性は保存される．次に，更新された S_2 もグラフ全体に拡散することを示す．条件より，$thr(v) < thr(v')$ である．また，点 v が接続しているのは，$C \setminus \{v\}$ の点と独立集合の点 $u_1, u_2, \ldots, u_i$ であり，点 v' が接続しているのは閾値グラフの特徴から，$C \setminus \{v'\}$ の点と独立集合の点 $u_1, u_2, \ldots, u_i, \ldots$ である．つまり，$N(v) \subseteq N(v')$ である．以上より，点 v を点 v' に代えても，$C \cup I \setminus \{v\} \setminus \{v'\}$ の点は $active$ になることが分かる．また，S_2 が標的集合とした場合，ある時点で点 v' が $active$ になることと，$thr(v) < thr(v')$ であることより，更新された S_2 では，点 v' より低い閾値を持つ点 v も $active$ になると言える．したがって，S_2 から点 v を除き点 v' を追加したものも標的集合である．この更新を $thr(v) < thr(v')$ となるものが存在しなくなるまで実行し，そのときの S_2 を S_3 とすれば良い．

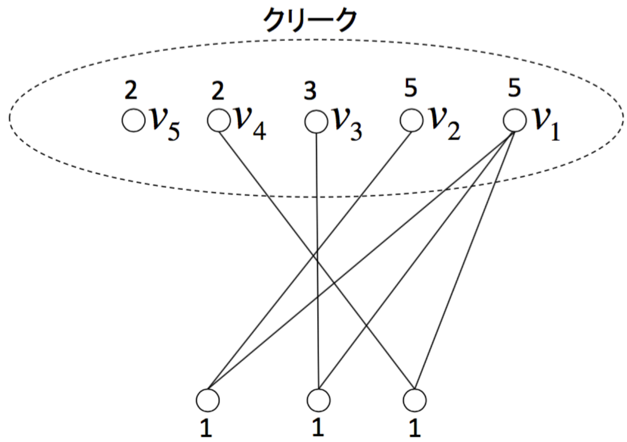

図 3 $|S'|$ が最小ではない例：クリークの標的集合は $\{v_1, v_2\}$ であるが，グラフ全体に対する標的集合は，例えば $\{v_1\}$ である．

補題 3.3 と補題 3.4 より，閾値グラフ全体の標的集合を S_{thr} $(|S_{thr}| = l)$ とすると，$\{v_1, v_2, \ldots, v_l\}$ もまた標的集合であることが分かる．つまり，クリークに対して $thr(u_1) \geq thr(u_2) \geq \ldots \geq thr(u_n)$ という条件のある閾値グラフの標的集合選択問題のどのような最適解でも，$v_1, v_2, \ldots, v_l$ という形に変形することができる．

一方，補題 3.1 と補題 3.2 より，クリーク C の標的集合を S $(|S| = k)$ とすると，$S' = \{v_1, v_2, \ldots, v_k\}$ もまた C に対する標的集合であり，S' は閾値グラフ全体に拡散すること

が分かる.

注. S' は閾値グラフ全体に拡散するが, 必ずしも $|S'|$ は最小であるとは言えないので, 標的集合であるとは限らない (図 3). 図 3 で, クリークの標的集合は $\{v_1, v_2\}$ であり, 確かに閾値グラフ全体に拡散することが分かる. しかし, 閾値グラフに対する標的集合は, 例えば $\{v_1\}$ である.

4. アルゴリズム

$thr(v_1) \geq thr(v_2) \geq \ldots \geq thr(v_n)$ を満たす, n 頂点 $v_1, v_2, \ldots, v_n$ のクリーク C と, m 頂点 $u_1, u_2, \ldots, u_m$ の独立集合 I で構成される閾値グラフに対しての標的集合選択問題を, C の標的集合を S とし $|S| = k$ としたとき, $kO(n)$ 時間で解くアルゴリズムを記述する. クリーク C の標的集合を求める部分のアルゴリズムは ([4], Nichterlein,Niedermeier,Uhlmann and Weller, 2012) を元に構成している.

まず, クリークの標的集合を求める ([4], Nichterlein,Niedermeier,Uhlmann and Weller, 2012) のアルゴリズム $\mathcal{A}$ を紹介する. 入力は, n 頂点からなるクリークと各点の閾値, 出力はクリークの標的集合のサイズ k である. また, 補題 3.1 を用いると, 閾値の大きい方から k 個選べば, それは標的集合になっていることが分かる.

アルゴリズム $\mathcal{A}$

0. n 頂点を $thr(v_1) \leq thr(v_2) \leq \ldots \leq thr(v_n)$ を満たすよう並べ替える.
1. $i = 1, k = 0$ とする.
2. ・$i \leq n - k$ の場合
   $k \leftarrow max\{k, thr(v_i) - i + 1\}$
   $i \leftarrow i + 1$ とし 2 の先頭に戻る.
   ・$i > n - k$ の場合
   k を返して終了.

次に, 閾値グラフに対する標的集合選択問題を解くアルゴリズムを記述する. 以下のアルゴリズムの入力は, $thr(v_1) \geq thr(v_2) \geq \ldots \geq thr(v_n)$ を満たす, n 頂点 $v_1, v_2, \ldots, v_n$ のクリーク C と, m 頂点 $u_1, u_2, \ldots, u_m$ の独立集合 I で構成される閾値グラフであり, 出力は, その閾値グラフの標的集合である.

アルゴリズム $\mathcal{A}'$

フェーズ 1. アルゴリズム $\mathcal{A}$ をクリーク C に適用し, 標的集合 S ($|S| = k$) を求める.

フェーズ 2. S から閾値グラフの標的集合を構成する.

0. $i = k, j = \lceil k/2 \rceil$ とする.
1. $S \setminus \cup_{l=j}^{k} \{v_l\}$ が閾値グラフ全体に拡散するか確認.
   拡散する場合 2, 拡散しない場合 3 へ.
2. ・$i - j = 1$ の場合
   $S \leftarrow S \setminus \cup_{l=j}^{k} \{v_l\}$ とし, S を返して終了.
   ・$i - j > 1$ の場合
   $i \leftarrow j, j = \lceil i/2 \rceil$ として 1 に戻る.
3. ・$i - j = 1$ の場合
   $S \leftarrow S \setminus \cup_{l=i}^{k} \{v_l\}$ とし, S を返して終了.
   ・$i - j > 1$ の場合
   $j = \lceil (i + j)/2 \rceil$ として 1 に戻る.

4.1 アルゴリズムの詳細

アルゴリズム $\mathcal{A}'$ の詳細について説明する. $thr(v_1) \geq thr(v_2) \geq \ldots \geq thr(v_n)$ を満たす, n 頂点 $v_1, v_2, \ldots, v_n$ のクリーク C と, m 頂点 $u_1, u_2, \ldots, u_m$ の独立集合 I で構成される閾値グラフが与えられたとき, アルゴリズム $\mathcal{A}$ を用いて, クリーク C の標的集合のサイズ (k) を求め, その標的集合を $\{v_1, v_2, \ldots, v_k\}$ とする. これが標的集合であることは, 補題 3.1 より保証される. また, 補題 3.2 より, これは閾値グラフ全体に拡散するが, 最小性は保証されない (図 3). 一方, 補題 3.4 より, 閾値グラフの標的集合のサイズが k' であるとき, $\{v_1, v_2, \ldots, v_{k'}\}$ は標的集合であることが分かるので, クリーク C の標的集合 $\{v_1, v_2, \ldots, v_k\}$ から, 二分探索で k' を求めるアルゴリズムである.

4.2 計算時間

アルゴリズム $\mathcal{B}'$ のフェーズ 1 は $O(n)$ で完了する [4]. また, S が閾値グラフ全体に拡散するか確認するのに高々 n ステップかかり, フェーズ 2 はクリーク C の標的集合 $\{v_1, v_2, \ldots, v_k\}$ から二分探索で閾値グラフの標的集合のサイズを求めているので $O(\log k)$ かかる. 以上より, このアルゴリズムの計算時間は $O(n \log k)$ と表せる.

4.3 アルゴリズムの正当性

アルゴリズム $\mathcal{B}'$ において, フェーズ 1 で求められる S は補題 3.2.2 より閾値グラフ全体に拡散する. ただし, 図 3 の例で見たように $|S|$ が最小であるとは必ずしも言えない. 一方, 補題 3.4 より, 閾値グラフに対する標的集合はクリーク C の閾値の大きなものから選んだように変形できる. 以上より, 閾値グラフの標的集合のサイズが k' であるとき, $\{v_1, v_2, \ldots, v_{k'}\}$ も標的集合であることが分かるので, クリーク C の標的集合 $\{v_1, v_2, \ldots, v_k\}$ から, 二分探索で k' を求めることができる.

本節と, 4.2 節より, 定理 2.1 は証明された.

5. おわりに

本論文では, NP 困難な問題として知られる標的集合選択問題を, スケールフリー性を持つことから社会ネットワークの文脈で注目される閾値グラフに対して考えた. このとき標的集合選択問題は, n 頂点からなるクリーク部分

の標的集合のサイズを k としたとき, 閾値グラフのクリーク部分に割り当てる閾値にはある大小関係を仮定すれば, $O(n \log k)$ 時間で解けることを示した．直径 1 であるグラフ (完全グラフ) は線形時間で解けることが分かっているが [4], スプリットグラフや, スプリットグラフの中でも直径 2 に限定したグラフでさえ NP 困難であることも知られているため [4], 完全グラフとスプリットグラフの間のグラフクラスである直径が 2 である閾値グラフでの本結果は意味のあるものであると言える.

参考文献

[1] O. Ben-Zwi, D. Hermelin, D. Lokshtanov, and I. Newman: Treewidth governs the complexity of target set selection. In: Discrete Optimization, 8(1):87-96 (2011)

[2] N. Chen: On the approximability of influence in social networks. In: SIAM Journal on Discrete Mathematics, 23(3),1400-1415 (2009)

[3] C.-Y. Chiang, L.-H. Huang, B.-J. Li, J. Wu, and H.-G. Yeh: Some results on the target set selection problem. In: Journal of Combinatorial Optimization (2012)

[4] A. Nichterlein, R. Niedermeier, J. Uhlmann, and M. Weller: On tractable cases of target set selection. In: Social Network Analysis and Mining (2012)