Vol.2016-AL-158 No.18
2016/6/25

距離限定部分グラフ探索問題に対する近似アルゴリズム

朝廣 雄一[1,a)]　土井 悠也[2]　志水 宏宇[2]　宮野 英次[2,b)]

概要：グラフ $G = (V, E)$ 中の頂点集合 $S \subseteq V$ を考える．S 中の任意の頂点 u, v の組に対して，G 中での u と v の間の距離が高々 d である場合に S は d-**クリーク**であると言う．また，S により誘導される G の部分グラフの直径が高々 d の場合には，S は d-**クラブ**であると言う．与えられた n 頂点グラフに対して，Max d-Clique 問題 (または Max d-Club 問題) の目的は，G 中の最大 d-クリーク (または最大 d-クラブ) を発見することである．任意の $\varepsilon > 0$ に対して，Max 1-Clique と Max 1-Club は，$\mathcal{P} = \mathcal{NP}$ でない限り多項式時間で $n^{1-\varepsilon}$-近似できない．なぜならば，これらの問題は Max Clique と同じ問題であり，Max Clique が $\mathcal{P} = \mathcal{NP}$ でない限り多項式時間で $n^{1-\varepsilon}$-近似できないからである ([5], [10])．さらに，任意の固定された $d \geq 2$ と任意の $\varepsilon > 0$ に対して，Max d-Club は $\mathcal{P} = \mathcal{NP}$ でない限り $n^{1/2-\varepsilon}$-近似できないことも知られている [2]．Max d-Club の近似上界に関しては，任意の偶数 $d \geq 2$ に対する多項式時間 $O(n^{1/2})$-近似アルゴリズムが著者らのグループにより提案されている [2]．しかしながら，奇数 $d \geq 3$ に対しては，このアルゴリズムの近似率は $O(n^{2/3})$ に悪化するため，近似下界 $\Omega(n^{1/2-\varepsilon})$ との乖離が依然として残っている [2]．本稿ではまず，Max d-Club の近似上界の改善を行う．すなわち，任意の奇数 $d \geq 3$ に対して多項式時間 $O(n^{1/2})$-近似アルゴリズムを提案する．そして同様のアイデアを用いて，任意の $d \geq 2$ に対して Max d-Clique も多項式時間で $O(n^{1/2})$-近似できることを示す．それとともに $\mathcal{P} = \mathcal{NP}$ でない限り，任意の $\varepsilon > 0$ に対して，Max d-Clique も $n^{1/2-\varepsilon}$-近似できないことについて述べる．

Approximation Algorithms for Maximum Distance-Bounded Subgraph Problems

Asahiro Yuichi[1,a)]　Doi Yuya[2]　Shimizu Hirotaka[2]　Miyano Eiji[2,b)]

Abstract: A *d-clique* in a graph $G = (V, E)$ is a subset $S \subseteq V$ of vertices such that for pairs of vertices $u, v \in S$, the distance between u and v is at most d in G. A *d-club* in a graph $G = (V, E)$ is a subset $S' \subseteq V$ of vertices that induces a subgraph of G of diameter at most d. Given a graph G with n vertices, the goal of Max d-Clique (Max d-Club, resp.) is to find a d-clique (d-club, resp.) of maximum cardinality in G. Max 1-Clique and Max 1-Club cannot be efficiently approximated within a factor of $n^{1-\varepsilon}$ for any $\varepsilon > 0$ unless $\mathcal{P} = \mathcal{NP}$ since they are identical to Max Clique [5], [10]. Also, it is known [2] that it is $\mathcal{NP}$-hard to approximate Max d-Club to within a factor of $n^{1/2-\varepsilon}$ for any fixed $d \geq 2$ and for any $\varepsilon > 0$. As for approximability of Max d-Club, there exists a polynomial-time algorithm which achieves an optimal approximation ratio of $O(n^{1/2})$ for any even $d \geq 2$ [2]. For any odd $d \geq 3$, however, there still remains a gap between the $O(n^{2/3})$-approximability and the $\Omega(n^{1/2-\varepsilon})$-inapproximability for Max d-Club [2]. In this paper, we first strengthen the approximability result for Max d-Club; we design a polynomial-time algorithm which achieves an *optimal* approximation ratio of $O(n^{1/2})$ for Max d-Club for any odd $d \geq 3$. Then, by using the similar ideas, we show the $O(n^{1/2})$-approximation algorithm for Max d-Clique for any $d \geq 2$. This is the best possible in polynomial time unless $\mathcal{P} = \mathcal{NP}$, as we can prove the $\Omega(n^{1/2-\varepsilon})$-inapproximability.

1 〒813-8503 福岡市東区松香台 九州産業大学 情報科学部
2 〒820-8502 福岡県飯塚市川津 九州工業大学大学院 情報工学研究院 システム創成情報工学研究系
a) asahiro@is.kyusan-u.ac.jp
b) miyano@ces.kyutech.ac.jp

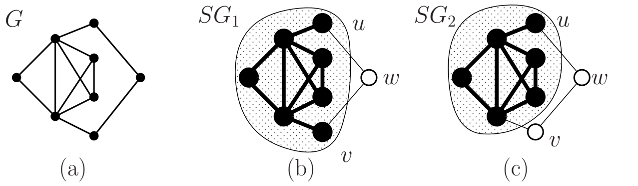

図 1 (a) グラフ G，(b) 最大 2-クリーク SG_1，(c) 最大 2-クラブ SG_2

1. はじめに

$G=(V,E)$ を重みなしグラフとする．ここで V と E はそれぞれ頂点集合と辺集合を表す．$V(G)$ と $E(G)$ で，G の頂点集合と辺集合を表すこともある．グラフ G 中の**クリーク**とは，お互いに隣接している頂点の集合 $Q \subseteq V(G)$ であり，Q により誘導されるグラフの直径は 1 である．グラフ理論と計算機科学において，最も重要かつ意欲的に研究されている問題の一つが最大クリーク問題 Max Clique である．Max Clique とは，与えられたグラフ G に対して，G 中のクリークのうち最大の頂点数を持つものを探す問題である．色々な分野の多くの応用問題がクリークを用いてモデル化される．

本稿では，Max Clique に対して一般化を行なった 2 つの問題である最大距離 d-クリーク問題 Max d-Clique と，最大直径 d-クラブ問題 Max d-Club を考える．ここで d は整数であり，これらの問題は [1], [7] により定義されている．グラフ G 中の**距離** d-**クリーク**(d-**クリーク**と略す) とは，頂点集合の部分集合 $S \subseteq V(G)$ であり，S に含まれる任意の頂点 u,v 間の G 中での距離が d 以下であるものである．グラフ G 中の**直径** d-**クラブ**(d-**クラブ**と略す) とは，頂点集合の部分集合 $S' \subseteq V(G)$ であり，S' により誘導される G の部分グラフの直径が d 以下となるものである．グラフ G が与えられた時，Max d-Clique と Max d-Club の目的は，それぞれ G 中の最大の d-クリークと最大の d-クラブを求めることである．$d=1$ のとき，Max 1-Club は Max 1-Clique と同じ問題であり，Max 1-Club は単に Max Clique と同じ問題でもある．しかしながら $d \geq 2$ の場合には，Max d-Club と Max d-Clique は明らかに異なる問題である．図 1-(a) にある，8 頂点のグラフ G を見て欲しい．G 中の最大 2-クリークは，図 1-(b) 中の 7 個の黒い頂点で誘導される部分グラフ SG_1 である．一方，最大 2-クラブは，図 1-(c) 中の 6 個の黒い頂点で誘導される部分グラフ SG_2 である．ここで，頂点 u と v は SG_1 や SG_2 の外にある頂点 w に隣接しているので，この 2 頂点の G 中での距離は 2 である．しかし，SG_1 自体の直径は 3 であり，SG_2 の直径は 2 である．

以下では，入力グラフ $G=(V,E)$ は，n 頂点と m 辺を持つとする．すなわち $|V(G)|=n$ かつ $|E(G)|=m$ とする．Max 1-Clique と Max 1-Club は Max Clique と同一の問題であるので，$\mathcal{P}=\mathcal{NP}$ でない限り，任意の $\varepsilon > 0$ に対して $n^{1-\varepsilon}$-近似が出来ない ([5], [10])．以下では $\mathcal{P}=\mathcal{NP}$ でないと仮定して議論を続ける．Max d-Club は，任意の $\varepsilon > 0$ と固定された $d \geq 2$ に対して，$n^{1/3-\varepsilon}$-近似ができないことを，Marinček と Mohar がギャップ保存帰着により示した [6]（ただし彼らの証明は $\mathcal{NP} \neq \mathcal{ZPP}$ を仮定して行われたものである）．その後，Asahiro らは [2] において，その近似下界 $\Omega(n^{1/3-\varepsilon})$ を改善し，$\Omega(n^{1/2-\varepsilon})$-近似が出来ないことを示した．また同じ論文で Asahiro らは近似上界について，偶数 $d \geq 2$ に対して Max d-Club は多項式時間で $O(n^{1/2})$-近似ができることを示した．しかしながら，彼らの提案したアルゴリズムは奇数 $d \geq 3$ に対しては，近似率が $O(n^{2/3})$ となってしまうため，Max d-Club に対する近似下界である $\Omega(n^{1/2-\varepsilon})$ と，近似上界である $O(n^{2/3})$ の間には依然として差がある [2]．本稿では，まず，Max d-Club に対するこの近似上界を改善する．すなわち，奇数 d に対しても近似率が $O(n^{1/2})$ となるような，単純な多項式時間アルゴリズムを Max d-Club に対して提案する．そして，同様のアイデアを利用することで，Max d-Clique に対しても $O(n^{1/2})$-近似アルゴリズムを与える．Max d-Clique に対しても，この近似率 $O(n^{1/2})$ が多項式の意味で最良である．なぜならば，本稿で述べるように，Max d-Clique の近似下界も $\Omega(n^{1/2-\varepsilon})$ となるからである．

本稿の主な興味は，多項式時間近似であるが，近似に関する以外の関連研究をいくつか挙げておく．最適化問題である Max d-Club に対応する判定問題 d-Club が自然に定義できる．すなわち d-Club は，与えられた入力グラフに対して，大きさ k 以上の d-クラブが存在するか否かを問う問題である．Schäfer らは，計算時間が $O((k-2)^k \cdot k! \cdot kn + nm)$ と $O(2^{n-k} \cdot nm)$ である 2 つの固定パラメータアルゴリズムを提案した [9]．また，Chang らは，Max d-Club に対して計算時間が $O^*(1.62^n)$ である厳密アルゴリズムを提案した [3]．

2. 問題と既知の結果

2.1 定義

$G=(V,E)$ を無向連結グラフとする．頂点 u と v の間の辺を (u,v) と書くことにする．グラフ G 中の頂点の最大次数と最小次数をそれぞれ $\Delta(G)$ と $\delta(G)$ で表す．ある頂点 v に対して，v に隣接している頂点の集合を $N_1(v)$ で表す．また，v には隣接しておらず，$N_1(v)$ にも含まれないが，$N_1(v)$ に含まれる頂点に隣接している頂点の集合を $N_2(v)$ で表す．

グラフ S が，グラフ $G=(V,E)$ に対して，$V(S) \subseteq V$ かつ $E(S) \subseteq E$ となっている場合，S を G の部分グラフと呼ぶ．頂点の部分集合 $U \subseteq V$ に対して，U により誘導される G の部分グラフを $G[U]$ で表す．G の部分グラフ S が，

もし $E(S) = V(S) \times V(S)$ を満たしているならば，S (または $G[V(S)]$) と $V(S)$ をそれぞれ**クリーク**と**クリーク頂点集合**と呼ぶ．

道 P の長さは $|P|$ と書くことにする．本稿では，単純道だけを扱うものとする．すなわち，ある道に含まれる任意の 2 頂点 v_i と v_j に対して，$v_i \neq v_j$ とする．G 中の頂点 u と v に対して，u から v に至る最短路の長さ，すなわち u と v 間の**距離**を $dist_G(u,v)$ で表し，G の**直径**を $diam(G) = \max_{u,v \in V(G)}\{dist_G(u,v)\}$ と定義する．

正の整数 $d \geq 1$ とグラフ G に対して **d 次べきグラフ**$G^d = (V(G),(E(G))^d)$ は，頂点集合として $V(G)$ を持ち，G 中の頂点の組 u,v のうち $dist_G(u,v) \leq d$ である組に対する辺 (u,v) を全て持つグラフのことである．ここで $E(G) \subseteq (E(G))^d$，すなわち G に元あった辺の集合 $E(G)$ は $(E(G))^d$ に含まれることに注意されたい．

定義 1 G の部分グラフ S(または頂点集合 $V(S)$) は，もし任意の頂点の組 $u,v \in V(S)$ に対して $dist_G(u,v) \leq d$ が成立しているならば，d-**クリーク**（または d-**クリーク頂点集合**）と呼ばれる．

d-クリークの直径は d より大きいかもしれない．そして，直径が d 以下の d-クリークは d-クラブと呼ばれる：

定義 2 G の部分グラフ S（または頂点集合 $V(S)$）は，もし任意の頂点の組 $u,v \in V(S)$ に対して $dist_S(u,v) \leq d$ が成立している，すなわち $diam(S) \leq d$ であるならば，d-**クラブ**（または d-**クラブ頂点集合**）と呼ばれる．

文献 [7] では，d-クラブ（や d-クリーク）は，極大部分グラフとして定義されている．すなわち，d-クラブ (や d-クリーク）の頂点集合は極大でなければならないと定義されている．しかしながら，本稿では，極大性については気にしないことにする．なぜならば，文献 [8] において，与えられた d-クラブが極大かどうかを答える問題は $\mathcal{NP}$-完全であることが示されているからである．定義から明らかである通り，G の d-クラブは，G の d-クリークでもある．グラフ G 中の最大クリーク，最大 d-クリーク，最大 d-クラブの大きさをそれぞれ $\#clique(G)$ と $\#clique(G,d)$，$\#club(G,d)$ で表す．

最大化問題に対して，アルゴリズム ALG が σ-近似アルゴリズムである，あるいは ALG の近似率が σ であるとは，全ての入力 G に対して，$OPT(G)/ALG(G) \leq \sigma$ が成立することを意味する．ここで $ALG(G)$ と $OPT(G)$ はそれぞれ，アルゴリズム ALG と最適アルゴリズムが出力する部分グラフの頂点数を表す．

最大 d-クリーク問題 (Max d-Clique) と，最大 d-クラブ問題 (Max d-Club) は次のように定式化される：与えられた無向連結グラフ G に対して Max d-Clique (または Max d-Club) の目的は，G 中の最大 d-クリーク（または最大 d-クラブ）を発見することである．$d = 1$ の場合，これら 2 個の問題は同じものであり，Max 1-Clique と Max 1-Club は，Max Clique と同じ問題である．このことから，これら 2 個の問題は，頂点間の距離とグラフの直径に関して Max Clique を一般化したものであると捉えられる．ここで，先に導入した記法を用いると，$\#clique(G) = \#clique(G,1) = \#club(G,1)$ が成立している．Erdös らが文献 [4] で示した通り，n 頂点を持つ無向連結グラフ G の直径は高々 $3n/(\delta(G)+1)+O(1)$ であるので，上記の 2 個の問題は $1 \leq d \leq 3n/(\delta(G)+1)+O(1)$ を満たす d の範囲において定義される．

2.2 d が偶数の場合の Max d-Club に対する近似アルゴリズム [2]

この節では，[2] で提案された，d が偶数の場合の Max d-Club に対する近似アルゴリズム $\texttt{ByFindStar}_d$ を簡単に紹介する．このアルゴリズムの基本的アイデアは，入力グラフ G の $\lfloor d/2 \rfloor$ 次べきグラフ中の 2-クリークを探すというものである．まず，アルゴリズム FindStar は，入力グラフ G の中から，最大次数 $\Delta(G)$ を持つ頂点 v を探し，v とその $\Delta(G)$ 個の隣接頂点により誘導される部分グラフ $G[N^+_G(v)]$ を出力する．このアルゴリズムの出力は，明らかに 2-クラブであり，よって，2 以上の d に対して $\Delta(G)+1$ 頂点の d-クラブを出力しているとも言える．このアルゴリズム FindStar は，かなり単純であり，線形時間で動作する．近似率は次の補題で与えられる．

補題 3 (**[2]**) 与えられた n 頂点グラフと，固定された整数 $d \geq 2$ に対して，FindStar は Max d-Club に対する $O(n^{1-1/d})$-近似アルゴリズムである．

上の補題 3 により，FindStar の Max 2-Club に対する近似度は，$O(n^{1/2})$ であることが分かる．論文 [2] において，近似度の下界が任意の $\varepsilon > 0$ に対して $\Omega(n^{1/2-\varepsilon})$ であることも示されているため，この FindStar の近似率 $O(n^{1/2})$ は，最良なものであると言える．

与えられた無向連結グラフ $G = (V,E)$ と整数 d から，d 次べきグラフ G^d を構成する単純なアルゴリズム PowerOfGraph が考えられる．PowerOfGraph は，まず最初に任意の頂点 $u,v \in V$ の組に対して $dist_G(u,v)$ を求め，もし $2 \leq dist_G(u,v) \leq d$ ならば，辺 (u,v) を追加することにより G^d を構成する．以上の 2 つのアルゴリズム PowerOfGraph と FindStar を組み合わせることで，Max d-Club に対する以下の多項式時間アルゴリズム $\texttt{ByFindStar}_d$ が提案されている．このアルゴリズム $\texttt{ByFindStar}_d$ の中では，入力グラフの $\lfloor d/2 \rfloor$ 次べきグラフが構成される．

アルゴリズム $\texttt{ByFindStar}_d$

入力: 無向連結グラフ $G=(V,E)$.

出力: G の部分グラフ S.

ステップ 1. グラフ G に $\texttt{PowerOfGraph}(G,\lfloor d/2\rfloor)$ を実行することで，$\lfloor d/2\rfloor$ 次べきグラフ $G^{\lfloor d/2\rfloor}$ を構成する.

ステップ 2. FindStar を $G^{\lfloor d/2\rfloor}$ に適用し, 最大の星グラフ T を得る.

ステップ 3. $S=G[V(T)]$ を出力する.

以下の補題により，$\texttt{ByFindStar}_d$ の出力は d-クラブであることが保証される.

補題 4 $\texttt{ByFindStar}_d$ の出力は $2\cdot\lfloor d/2\rfloor$-クラブである.

偶数 d に対する $\texttt{ByFindStar}_d$ の近似率は，次の定理により与えられる.

定理 5 ([2]) 与えられた n 頂点グラフと，固定された偶数 $d\geq 2$ に対して，$\texttt{ByFindStar}_d$ は Max d-Club に対する多項式時間 $O(n^{1/2})$-近似アルゴリズムである.

しかしながら，d が奇数の場合には，$\texttt{ByFindStar}_d$ の近似率は次のように悪くなる.

定理 6 ([2]) 与えられた n 頂点グラフと，固定された奇数 $d\geq 3$ に対して，$\texttt{ByFindStar}_d$ は Max d-Club に対する多項式時間 $O(n^{2/3})$-近似アルゴリズムである.

より良い近似率を持つ多項式時間アルゴリズムを設計することが，本稿の目的である.

3. Max d-Club に対する $O(n^{1/2})$-近似アルゴリズム

第 2.2 節で述べたように，奇数 d に対する $\texttt{ByFindStar}_d$ の近似率は，偶数 d に対する場合より悪い．この原因は補題 4 に見ることができる．最悪の場合には，最適解である部分グラフの直径が正確に d である（すなわち最適解は d-クラブである）にも関わらず，$\texttt{ByFindStar}_d$ が出力する部分グラフの直径は $2\cdot\lfloor d/2\rfloor$ であり，これは d が奇数の場合には $d-1$ になる．例えば，$d=5$ の場合の例である図 2 を見て欲しい．$\texttt{ByFindStar}_5$ に対する悪い入力例を図 2-(a) に示している．このグラフは長さ 5 の道であり，すなわち 5-クラブでもある．よって，グラフ全体が Max 5-Club に対する最適解となっている．しかし，このグラフ G が入力として与えられた時，$\texttt{ByFindStar}_5$ は最初に G^2（図 2-(b)）を作成する．そして，$\texttt{ByFindStar}_5$ は，例えば図 2-(c) 中の実線で示している 5 頂点の星グラフを出力する．つまり，$\texttt{ByFindStar}_5$ の出力は G 中の 4-クラブ (長さ 4 の道) である.

上記の観察から，本稿では新しいアルゴリズム $\texttt{ByFindStar2}_d$ を以下の通り提案する．$\texttt{ByFindStar2}_d$ はまず最初に，各辺に対して 1 個ずつ新しい頂点を挿入する (図 2-(d) 参照)．そして図 2-(e) のように，そのグラフの d 次べきグラフを作成する（$\texttt{ByFindStar}_d$ は，入力グラフの $\lfloor d/2\rfloor$ 次べきグラフを作成していた)．この変更により，挿入した新しい頂点に接続する辺が，d が奇数の場合に長さ $d/2$ を表現するのに役立つことになる．そして，$\texttt{ByFindStar2}_d$ は，図 2-(f) 中の実線で示される星グラフを出力するが，この星グラフは G 中の d-クラブに対応する．以上のアイデアはとても単純であるが，これにより近似率は，定理 10 に示すように，以前の $O(n^{2/3})$ から $O(n^{1/2})$ へと改善される．以下が，Max d-Club に対するアルゴリズム $\texttt{ByFindStar2}_d$ の記述である.

アルゴリズム $\texttt{ByFindStar2}_d$

入力: 無向連結グラフ $G=(V,E)$

出力: G の部分グラフ

ステップ 1. 各辺 (u,v) に頂点 $w_{u,v}$ を挿入する．挿入された頂点を**白頂点**と呼ぶことにし，白頂点の集合を V_W で表す．一方，入力グラフ G に元からあった頂点を**黒頂点**と呼ぶ．この処理により得られたグラフを $H=(V\cup V_W,E')$ で表す．ここで $E'=\{(u,w_{u,v}),(v,w_{u,v})\mid(u,v)\in E\}$ である.

ステップ 2. $\texttt{PowerOfGraph}(H,d)$ を実行することで，H の d 次べきグラフ H^d を求める.

ステップ 3. H^d に対して，最多の黒頂点を含む星グラフ $T=(V',E')$ を探す（この手続きを FindStar2 と呼ぶことにする).

ステップ 4. $G[V'\cap V]$，すなわち T に含まれる黒頂点により誘導される G の部分グラフを出力する.

ステップ 1 で作成されたグラフ H 中の, 黒頂点の部分集合 B に対して $W(B)=\{w_{u,v}\mid u,v\in B$ かつ $(u,v)\in E(G)\}$ と定義する．すなわち $W(B)$ は，両方の隣接頂点が黒であり，かつ B に属しているような白頂点をすべて含む．次の補題は，G 中での (黒) 頂点間の距離と，H 中でのそれらの距離を関係づけるものである.

補題 7 G の d-クラブ頂点集合 S 内の任意の 2 つの黒頂点 u,v について，$dist_{G[S]}(u,v)\leq d$ となる必要十分条件は $dist_{H[S\cup W(S)]}(u,v)\leq 2d$ となることである．結果として，$G[S]$ が G の d-クラブになる必要十分条件は $H[S\cup W(S)]$ が H の $2d$-クラブになることである.

証明 ($\Rightarrow$) $G[S]$ 中の u から v に至る最短路 P について考える．$H[S\cup W(S)]$ 中でこの道 P に対応する道 P' は，P に含まれる（黒）頂点の集合 $V(P)$ と，$V(P)$ により決定される白頂点の集合 $W(V(P))$ を含む．よって，P' の長さ $|P'|$ は，P の長さ $|P|$ のちょうど 2 倍になっている．すなわち，もし $dist_{G[S]}(u,v)\leq d$ であるならば，$dist_{H[S\cup W(S)]}(u,v)\leq 2d$ である.

($\Leftarrow$) 簡単のために $H'=H[S\cup W(S)]$ とおく．H' 中の任意の 2 つの白頂点 w_1 と w_2 は，H の構成方法により，隣

Vol.2016-AL-158 No.18
2016/6/25

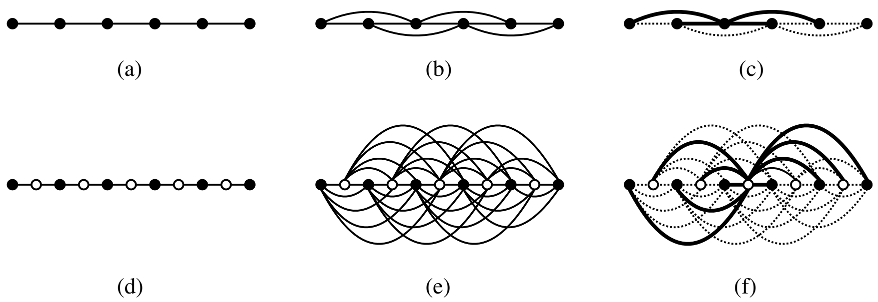

図 2 (a) $d=5$ の場合のグラフ例 G; (b) G^2; (c) ByFindStar$_5$ の出力 (実線); (d) グラフ H; (e) H^d; (f) ByFindStar2$_5$ の出力 (実線)

接していないことが分かる．同様に，H' 中の任意の 2 つの黒頂点 b_1 と b_2 も隣接していない．P' を H' 中の頂点 u から v へ至る最短路とする．P' に対応する G 中の道 P の長さは，$|P'|$ の半分である．ここで，P' に含まれる黒頂点はすべて S に含まれるので，道 P は $G[S]$ の中にあることに注意されたい．よって，もし $dist_{H[S\cup W(S)]}(u,v) \le 2d$ ならば，$dist_{G[S]}(u,v) \le d$ である． □

次に，d 次べきグラフを作ることにより，H の $2d$-クラブ中の黒頂点間の距離がどの程度短くなるかを考える．

補題 8 もし S が H の $2d$-クラブ頂点集合であるならば，S に含まれる任意の 2 個の黒頂点 u,v に対して $dist_{H^d[S]}(u,v) \le 2$ が成立する．すなわち，S は，H^d の 2-クラブ頂点集合である．

証明 $H[S]$ 中の頂点 u から v に至る，長さが高々 $2d$ である道 P を考える．x を P の中央 (あたり) にある頂点，すなわち $dist_H(u,x) \le d$ かつ $dist_H(x,v) \le d$ が成立する頂点とする．すると，$H^d[S]$ は，(u,x) と (x,v) の 2 つの辺を持つことになり，$dist_{H^d[S]}(u,v) \le 2$ となる． □

次の補題は ByFindStar2$_d$ の出力が，Max d-Club の実行可能解，すなわち d-クラブになっていることを保証する．

補題 9 ByFindStar2$_d$ の出力は，入力グラフ G の d-クラブである．

証明 Step 3 で得られた H^d 中の星グラフ $T=(V',E')$ の根を r とする．$v \in V'$ かつ $v \ne r$ である頂点 v を考える．$(r,v) \in E'$ であるので，r から v に至る長さ d 以下の道 P が H 中に存在するはずである．ここで，P 中の他の頂点も r からの距離がやはり d 以下であるので，それらも H^d 中では r と隣接している．つまり，それらの頂点も v と同様に V' に含まれる．よって u と v が両方とも V' に属しており $r \notin \{u,v\}$ でもある，すなわち (r,u) と (r,v) の 2 つの辺が T に含まれるならば，$dist_{H[V']}(u,v) \le 2d$ が成立する．なぜならば，$H[V']$ は，長さ d 以下の r から u に至る道と，同様に長さ d 以下の r から v に至る道の 2 個の道を含んでいるからである．r と $H[V']$ 中の他の頂点との距離も高々 d である．まとめると $diam(H[V']) \le 2d$，すなわち $H[V']$ は $2d$-クラブである．ここで $V' \cap V$ は V' 中の黒頂点の集合である．補題 7 と $H[V']$ が $2d$-クラブであるという事実により，任意の 2 個の黒頂点 $u,v \in V' \cap V$ に対して，$dist_{G[V'\cap V]} \le d$ が成立する．つまり，$G[V' \cap V]$ は d-クラブである． □

最後に，d を奇数に限定した場合の Max d-Club に対する ByFindStar2$_d$ の近似率について示す．

定理 10 ByFindStar2$_d$ は，d が奇数の場合の Max d-Club に対する多項式時間 $\lceil n^{1/2} \rceil$-近似アルゴリズムである．

証明 ステップ 3 の FindStar2 は H^d のサイズ（入力グラフ G の多項式サイズ) の線形時間で動作する．他のステップも G のサイズの多項式時間で動作するので，ByFindStar2$_d$ の計算時間は多項式である．

補題 7 と 8 により，G の d-クラブは，H^d の 2-クラブに対応する．ここで我々が求めたいのは，H^d 中の 2-クラブの中で，最多の黒頂点（すなわち入力グラフの頂点）を含むものであるが，それは H^d 中の最大 2-クラブとは必ずしも一致しない．そこで，別問題として，H^d 中から最多の黒頂点を含む 2-クラブを探すという問題を考える．ここで解の大きさは，頂点数ではなく黒頂点の数と定義する．以前のアルゴリズムで用いられていた手続き FindStar は，最多の頂点を含む 2-クラブを探そうとしていたが，今回の FindStar2 は，最多の黒頂点を含む 2-クラブを探すことを目的としている．この別問題の H^d 中の最適解 (2-クラブ) の頂点集合は，G 中の d-クラブ頂点集合とは必ずしもなっていない．しかしながら，補題 9 により，ByFindStar2$_d$ の出力は G の d-クラブとなっていることが保証されている．この証明の中では，頂点に隣接する黒頂点の数を**黒次数**と呼ぶことにし，H^d 中の頂点の最大黒次数を Δ_b で表す．

S^* を H^d 中で最多の黒頂点を含む 2-クラブとする．以下で，S^* 中の黒頂点の個数について上界を与える．もし，S^* が黒頂点しか含んでいないならば，その個数は $1+\Delta_b+(\Delta_b-1)^2=(\Delta_b)^2-\Delta_b+2$ で抑えられる．この値は，ある黒頂点から長さ 2 の道で到達できる黒頂点の個数を表している．次の場合として，S^* が白頂点 w を含んでいると仮定する．B_1 を $N_1(w)$ に含まれる黒頂点の集合とする．B_1 の大きさ $|B_1|$ は Δ_b 以下である．そして，$N_2(w)$

中の高々 $(\Delta_b)^2$ 個の黒頂点が B_1 の頂点に隣接している．このことは，B_2 を $N_2(w)$ 中のそのような黒頂点の集合とすると，$|B_2| \leq (\Delta_b)^2$ が成立していると言い換えられる．$N_1(w)$ の中にはいくつかの白頂点が存在するかもしれず，それらが $N_2(w)$ 中の黒頂点に隣接しているかもしれない．$N_1(w)$ に含まれる白頂点に隣接している $N_2(w)$ 中の黒頂点は，すべて B_2 に含まれることを示す．x を $N_1(w)$ に含まれる白頂点，y を $N_2(w)$ に含まれ，かつ x に隣接する黒頂点とする．H 中では $dist_H(w,x) \leq d-1$ が成立している．なぜならば，白頂点間の距離は偶数だからである．そうすると，x に隣接していて，かつ H 中で x から y に至る長さ d 以下の道に含まれる黒頂点が存在する．この頂点を z で表す．この z について，$dist_H(w,z) \leq d$ と $dist_H(z,y) \leq d-1$ が成立しているので，辺 (w,z) と (z,y) が H^d に含まれる．z は $N_1(w)$ に属する黒頂点であるので，y は B_2 の要素であると言える．つまり，B_2 に含まれる頂点以外が，$N_1(w)$ に含まれる白頂点と隣接し $N_2(w)$ に含まれる黒頂点として出現することはない．以上の議論により，S^* 中の黒頂点の数は高々 $\max\{(\Delta_b)^2 - \Delta_b + 2, |B_1| + |B_2|\} = \max\{(\Delta_b)^2 - \Delta_b + 2, (\Delta_b)^2 + \Delta_b\}$ 以下であることが分かる．$\Delta_b \geq 1$ であり，$(\Delta_b)^2 + \Delta_b$ の方が $(\Delta_b)^2 - \Delta_b + 2$ よりも大きいので，$(\Delta_b)^2 + \Delta_b$ を黒頂点数の上界と考えることができる．

H^d に対する `FindStar2` の出力である星グラフは，Δ_b 個以上の黒頂点を含む．よって `FindStar2` の出力解の大きさ $FindStar2(H^d)$ は，Δ_b 以上である．そして上記の議論から，最適解の大きさ $\#club_b(H^d, 2)$ は高々 $\min\{n, (\Delta_b)^2 + \Delta_b\}$ である．ここで H^d 自体は $n+m$ 頂点を持つが，黒頂点の個数は n 個であるので，この n が最適解の大きさの上界の一つとなる．よって，もし $\Delta_b \geq \lceil n^{1/2} \rceil$ であるならば，

$$\frac{\#club_b(H^d,2)}{FindStar2(H^d)} \leq \frac{n}{\Delta_b} \leq \frac{n}{\lceil n^{1/2} \rceil} \leq n^{1/2}$$

であり，反対に $\Delta_b \leq \lceil n^{1/2} \rceil - 1$[*1] であるならば，

$$\frac{\#club_b(H^d,2)}{FindStar2(H^d)} \leq \frac{(\Delta_b)^2 + \Delta_b}{\Delta_b} = \Delta_b + 1 \leq \left\lceil n^{1/2} \right\rceil$$

と求めることができる．補題 7 と 8 により，$\#club_b(H^d, 2) \geq \#club(H, 2d) = \#club(G, d)$ が成立する．最後に，$FindStar2(H^d) = ByFindStar2(G)$ であるという事実に基づき，

$$\frac{\#club(G,d)}{ByFindStar2(G)} \leq \frac{\#club_b(H^d,2)}{FindStar2(H^d)} \leq \left\lceil n^{1/2} \right\rceil$$

を得る．ここで $ByFindStar2(G)$ は，入力グラフ G に対する $\mathtt{ByFindStar2}_d$ の出力解の大きさである． □

注 11 上の議論は d が奇数であるという仮定のもとで行なっている．しかしながら，最適解に含まれる（黒）頂点数の上界 $(\Delta_b)^2 + \Delta_b$ は，d が偶数の場合も同様に示すことができる．これにより，偶数 d に対しても $\mathtt{ByFindStar2}_d$ は $\lceil n^{1/2} \rceil$-近似アルゴリズムであることが言える．

*1 Δ_b は整数である

4. Max d-Clique の近似

本節では Max d-Clique の近似上界と下界について得られた結果を述べる．ただし，それぞれの証明のアイデアは，定理 10 の証明と [2] によって示された Max d-Club に関する近似下界の証明と同様であるので省略する．最初の結果は，アルゴリズム $\mathtt{ByFindStar2}_d$ の Max d-Clique に対する近似率に関するものである．

定理 12 n 頂点グラフと固定された整数 d に対して，$\mathtt{ByFindStar2}_d$ は Max d-Clique の多項式時間 $\lceil n^{1/2} \rceil$-近似アルゴリズムである．(証明略)

近似下界については次の結果が得られる．

定理 13 任意の $\varepsilon > 0$ に対して，$\mathcal{P} = \mathcal{NP}$ でない限り，Max d-Clique は $n^{1/2-\varepsilon}$-近似できない．(証明略)

謝辞 本研究は科研費 25330018 と 26330017 の助成を受けたものである．

参考文献

[1] Alba, R.: A graph-theoretic definition of a sociometric clique. *J. Math Sociol.*, 3, pp.113–126, 1973.
[2] Asahiro, Y. , Miyano, E., Samizo, K: Approximating maximum diameter-bounded subgraphs. Proc. LATIN 2010, LNCS 6034, pp.615–626, 2010.
[3] Chang, M.-S., Hung, L.-J., Lin, C.-R., Su, P.-C.: Finding large k-clubs in undirected graphs. *Computing*, 95, pp.739–758, 2013.
[4] Erdös, P., Pach, J., Pollack, R., Tuza, Z.: Radius, diameter, and minimum degree. *J. Combin. Theory, Ser.B*, 47, pp.73–79, 1989.
[5] Håstad, J.: Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Mathematics*, 182 (1), pp.105–142, 1999.
[6] Marinček, J. and Mohar, B.: On approximating the maximum diameter ratio of graphs. *Discrete Math*, 244, pp.323–330, 2002.
[7] Mokken, R.J.: Cliques, clubs and clans. *Quality & Quantity*, 13, pp.161–173, 1979.
[8] Pajouh, F.M. and Balasundaram, B.: On inclusionwise maximal and maximum cardinality k-clubs in graphs. *Discrete Optimization*, 9, pp.84–97, 2012.
[9] Schäfer, A., Komusiewicz, C., Moser, H., Niedermeier, R.: Parameterized computational complexity of finding small-diameter subgraphs. *Optimization Letters*, 6(5), pp.883–891, 2012.
[10] Zuckerman, D.: Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3, pp.103–128, 2007.