数理解析研究所講究録 1325 巻 2003 年 181-184

平均次数の高いVC-PMの近似アルゴリズム

竹内 元気
名古屋大学情報文化学部
n9932@sis.nagoya-u.ac.jp

築地 立家
名古屋大学人間情報学研究科
tsukiji@math.human.nagoya-u.ac.jp

1 Introduction

最小頂点被覆問題とは、与えられた無向グラフ $G=(V,E)$ に対して、頂点数が最小となる G の頂点被覆を求める問題である。この問題は、取り得る解の集合から最もよい解を選び出す最適化問題のうち、最小の解を探すことを目的とする最小化問題である。また最小頂点被覆問題は NP 困難に属することが知られている。最適化問題が NP 困難な場合は、P = NP でない限り、最良の解をみつける多項式時間アルゴリズムは存在しない。したがって、最小頂点被覆問題に対して最良の解を求める多項式時間アルゴリズムは存在しないという立場で、この問題の近似アルゴリズムを考えることにする。ただし、本稿では入力されるグラフを完全マッチングを持つグラフに限定している。入力グラフが完全マッチングを持つグラフに限定されている最小頂点被覆問題は、VC-PM とよばれる。

Chen と Kanj [2] は VC-PM において近似度 r が得られるならば、一般のグラフに対しては近似度 $(r+2)/2$ が得られることを示している。また、VC-PM の近似アルゴリズムの近似度については、$1.069+0.069\bar{d}$ を求めた ($\bar{d}$ は平均次数)。そして、今村と岩間 [1] が Chen と Kanj のアルゴリズムを改良し、$0.0230\bar{d}+1.378$、$0.0115\bar{d}+1.534$、$0.0138\bar{d}+1.441$ (triangle-free の場合) の 3 つの近似度を示した。

本稿では、平均次数が高い場合を考える。そこで、まず今村と岩間のアルゴリズムを解析して、平均次数が任意に固定された時にも近似度を抑えられることを示した。さらに Opt 解のサイズを評価することにより、平均次数が 20 から 30 程度 (正確には $20.16<\bar{d}<31.75$) の場合に、今村と岩間のアルゴリズを解析した場合の近似度を改良することができた。その近似度は、$1.6010+0.0074\bar{d}$ となった。

2 VC-PM

無向グラフ $G=(V,E)$ が与えられたとき、「グラフ G の任意の辺 $(u,v)\in E$ に対して $u\in C$ または $v\in C$ が成立」をみたす $C\subseteq V$ をグラフ G の頂点被覆という (ここで、V は頂点集合、E は辺集合を表わす)。頂点被覆の中で頂点数が最小である C を探す問題を最小頂点被覆問題という。入力グラフ G が完全マッチング M を持つグラフに限定されている最小頂点被覆問題を VC-PM とよぶ。

ここで完全マッチングとは、グラフ G の全ての点を、2 点ずつ辺で向かい合わせる (マッチングさせる) 辺の部分集合である。ここで次の $G(V')$ と $N_G(v)$ の定義もしておく。$G=(V,E)$ の頂点部分集合 V' に対して、誘導部分グラフ $G(V')$ とは、$E'=\{(u,v)\in E|u,v\in V'\}$ とした時の $G=(V',E')$ のことである。また、$G=(V,E)$ において $v\in V$ の近傍 $N_G(v)$ とは、v に隣接する頂点集合 $\{u|(u,v)\in E\}$ のことをいう。以下ではグラフ G の頂点数を n、辺の本数を m とする。

3 今村と岩間の近似アルゴリズム

この節では Chen と Kanj の近似アルゴリズムを応用した今村と岩間の近似アルゴリズム [1] の概要を紹介する。入力するグラフ $G=(V,E)$ は完全マッチング M を持つものとする。入力されたグラフ G から任意の頂点 v に対して変数 x_v を、任意の辺 (u,v) に対して節 $(x_u\vee x_v)$ を、任意のマッチング辺 (u,v) に対して節 $(\bar{x}_u\vee\bar{x}_v)$ を対応させ、次のような論理式 F_G を定義する。

$$F_G=\bigwedge_{(u,v)\in E}(x_u\vee x_v)\wedge\bigwedge_{(u,v)\in M}(\bar{x}_u\vee\bar{x}_v)$$

$(x_u\vee x_v)$ の形の節は positive clause、$(\bar{x}_u\vee\bar{x}_v)$ の形をした節は negative clause と呼ばれる。論理式 F_G の positive clause には各々1 の、negative clause には各々 $w(w>1)$ の重みがつけられている。この論理式を MAX 2-SAT の近似アルゴリズムで解き (このアルゴリズムの近似度を r とする)、その解を σ とする。この割り当て σ に対して、以下のような修正を行う。FALSE に割り当てられている変数 x_v を TRUE に割り当て直して、もし充足節の重みの和 $|\sigma|$ が増えるならば、変数 x_v を TRUE に割り当て直す。このような操作を、FALSE に割り当てられているどの変数 x_v も TRUE に割り当て直された時に、重みの和 $|\sigma|$ が減少してしまうというところまで繰り返して σ に極大性を持たせる。そして、$\mathcal{F}_\sigma=\{v|\sigma(x_v)=\mathrm{TRUE}\}$ の誘導部分グラフ $G(\mathcal{F}_\sigma)$ の頂点被覆 $C_\mathcal{F}$ と、$\mathcal{F}_\sigma=\{v|\sigma(x_v)=\mathrm{TRUE}\}$ との和集合を G の頂点被覆とする。ここで σ の極大性から $G(\mathcal{F}_\sigma)$ は次数が w 未満のグラフとなっている。次に、今村と岩間の論文で示されている補題を紹介する。

補題 3.1 $C_\mathcal{F}$ を $G(\mathcal{F}_\sigma)$ の任意の頂点被覆、G の頂点被覆 C を $C=\mathcal{T}_\sigma\cup C_\mathcal{F}$、$\sigma_C$ を $\sigma_C(x_v)=TRUE\Leftrightarrow v\in C$、$\delta=(|\sigma|-|\sigma_C|)/n$、$\bar{d}=2m/n$ と定義する。このとき以

下の不等式が成り立つ。

$$\frac{|C|}{Opt(G)} \leq 1 + \frac{2}{w}\delta + \frac{r-1}{wr}(\bar{d} + w)$$

ここで、δ の値が大きくなるにつれて上の補題の近似度も大きくなってしまうので、$w = 2, 3$ の場合に δ の値が大きくても良い近似度が与えられるような別の評価式を与えている。

ここでは $w = 3$ の場合の評価式の導き方を書くことにする。$w = 3$ の時、$G(\mathcal{F}_\sigma)$ は次数が 2 以下のグラフとなる。このグラフは独立頂点、閉路、パスからなる。

従って、この場合は最小頂点被覆が容易に求められる。というのも閉路、パスともに一つおきに頂点を選べば良いからである。このように頂点を選べば、充足節の重みの和が $|\sigma|$ から $|\sigma_C|$ へと δn 減少し、またいくつかの頂点が残る。この減少数と残る頂点数の関係を調べる。完全グラフ K_3 を例にとってみると、最小頂点被覆として 3 個の頂点のうち 2 個の頂点を選ぶことになる。この時の充足節の重みの和の減少数は、充足される positive clause の個数が 3 個、非充足となる negative clause（重み 3 がかかっている）の個数が 2 個であることから、充足節の重みの和は 3 減少する。残る頂点数については 1 個である。従って、残った頂点数と重みの和の減少数との比、(残りの頂点数)/(重みの和の減少数) は 1/3 となる。他の長さの閉路やパスの場合も同様にこのような比を考えることができるが、この比の値が最小になるのは K_3 の場合となる。従って、充足節の重みの和が $|\sigma|$ から $|\sigma_C|$ へと δn 減少した場合、少なくとも $\delta n/3$ 個の頂点が残る。つまり、$|C| \leq n - \delta n/3$ が言える。この不等式と $Opt(G) \geq n/2$ から

$$\frac{|C|}{Opt(G)} \leq 2 - \frac{2\delta}{3}$$

また補題 3.1 において $w = 3$ とすれば

$$\frac{|C|}{Opt(G)} \leq 1 + \frac{2}{3}\delta + \frac{r-1}{3r}(\bar{d} + 3)$$

が得られ、

$$\min\left(2 - \frac{2\delta}{3}, 1 + \frac{2}{3}\delta + \frac{r-1}{3r}(\bar{d} + 3)\right)$$

$$\leq 2 - \frac{1}{6r}(3 - (r-1)\bar{d})$$

が言えるから、このアルゴリズムの近似度は、$2 - (1/6r)(3 - (r-1)\bar{d})$ とすればよい。$r = 1.0741$ [3] とおくと、近似度 $0.0115\bar{d} + 1.534$ が得られる。

4 今村と岩間の近似アルゴリズムの解析

この節では、今村と岩間の近似アルゴリズムにおいて negative clause にかける重みを $w = k + 1 (k \geq 1)$ と一般の形でおき、このアルゴリズムを解析する。重みが $w = k + 1$ の場合、$G(\mathcal{F}_\sigma)$ は次数が k 以下のグラフとなる。

4.1 K_{k+1} の頂点被覆について

K_{k+1} の頂点被覆としては $k + 1$ 個の点のうち 1 点を残して他の k 個の点を選ぶ。そしてその k 個の点を TRUE に割り当て直したときの $|\sigma|$ の変化は、K_{k+1} の $k(k+1)/2$ 本の辺に対応する positive clause が充足され、また k 個の negative clause が非充足になり $k(k+1)$ 減少する。差し引き $k(k+1)/2$、$|\sigma|$ は減少する。従って (残りの頂点数)/(重みの和の減少数) の比は $2/k(k+1)$ となる。

4.2 K_{k+1} 以外のグラフの頂点被覆について

K_{k+1} 以外のグラフにおいては次数が k よりも小さい点が存在する。そのような点の中でも最も次数が小さい点を v としその次数を l $(l < k)$ とおく。ここで $N_{G(\mathcal{F}_\sigma)}(v)$ の各々の点の次数は l 以上と仮定してよい。また $N_{G(\mathcal{F}_\sigma)}(v)$ の任意の点 u に対して、x_u =TRUE に割り当て、このグラフの頂点被覆の一部とする。同時にこの時の $|\sigma|$ の変化を考える。充足される辺の本数は少なくとも $l(l+1)/2$ 本はあるので充足される positive clause の個数は少なくとも $l(l+1)/2$ 個はある。また非充足となる negative clause の個数は l 個であるから、それにともなう重みの和の減少数は $l(k+1)$ であり、従ってこの時の $|\sigma|$ の減少数を β とすれば、$\beta \leq l(2k-l+1)/2$ となる。また、$\{v, N_{G(\mathcal{F}_\sigma)}(v)\}$ のうち残る頂点の個数は v の 1 個なので、(残りの頂点数)/(重みの和の減少数) は $1/\beta$ となる。ここで、$l(2k-l+1)/2 < k(k+1)/2$ $(l < k)$ と $\beta \leq l(2k-l+1)/2$ から

$$\frac{1}{k(k+1)/2} < \frac{1}{l(2k-l-1)/2} \leq \frac{1}{\beta}$$

が言える。以上のような操作を繰り返して $C_\mathcal{F}$ を出力する。$C_\mathcal{F}$ を出力するまでの流れは次のようになる。

1. $G = (V, E)$ を入力する。
2. $\mathcal{F}_\sigma$ を計算し、$X = \mathcal{F}_\sigma$、$C_\mathcal{F} = \phi$ とおく。
3. $G(X)$ において最小次数 l の頂点 $v \in X$ を選ぶ。
4. $l > 0$ ならば以下を実行する。
   - 4.1. 全ての $u \in N_{G(\mathcal{F}_\sigma)}(v)$ に対して x_u=TRUE にする。
   - 4.2. $C_\mathcal{F} = C_\mathcal{F} \cup N_{G(\mathcal{F}_\sigma)}$ とおく。
   - 4.3. $X = X - \{N_{G(\mathcal{F}_\sigma)}(v) \cup v\}$
   - 4.4. $G(X)$ において最小次数 l の頂点 $v \in X$ を選ぶ。
5. $C_\mathcal{F}$ を出力する。

4.3 $w = k + 1$ の場合の近似度

以上のようにして $G(\mathcal{F}_\sigma)$ の頂点被覆を求めてやると (残りの頂点数)/(重みの和の減少数) の値の最小値は K_{k+1} の頂点被覆を考えた場合の $2/k(k+1)$ となる。従って、充足節の重みの和が $|\sigma|$ から $|\sigma_C|$ へと δn 減少すれば、少なくとも $(2\delta n)/k(k+1)$ 個の頂点が G の頂点被覆に選ばれないで残る。従って $|C| \leq n - 2\delta n/k(k+1)$ が成

り立つ。この不等式と $Opt(G) \geq n/2$、$k = w-1$ から

$$\frac{|C|}{Opt(G)} \leq 2 - \frac{4\delta}{(w-1)w}$$

また補題 3.1 [1] より

$$\frac{|C|}{Opt(G)} \leq 1 + \frac{2}{w}\delta + \frac{r-1}{wr}(\bar{d}+w)$$

が得られ、

$$\min\left(2 - \frac{4\delta}{(w-1)w}, 1 + \frac{2}{w}\delta + \frac{r-1}{wr}(\bar{d}+w)\right)$$
$$\leq 2 - 2\frac{(1-r)\bar{d}+w}{r(1+w)w}$$

が言えるから、近似度は $2-2((1-r)\bar{d}+w)/(r(1+w)w)$ とすることができる。$r = 1.0741$ [3]、$w = 4, 5$ とすると、近似度はそれぞれ $1.628 + 0.0069\bar{d}$, $1.690 + 0.0046\bar{d}$ となる。

5 改良したアルゴリズム

ここでは $w = 4$ とし、$Opt(G)$ のサイズの評価を行うことにより近似度を良くすることを考えた。その結果、$\bar{d}$ が 20 から 30 のあたりで、今村と岩間のアルゴリズムを解析した場合の近似度を改良することができた。

以下でも $G = (V, E)$ から論理式 F_G 構成し WEIGHTED MAX 2-SAT の近似アルゴリズムで近似解 σ を求め、$|\sigma|$ に極大性を持たせて V を $\mathcal{T}_\sigma$ と $\mathcal{F}_\sigma$ とに分けるところまでは同じである。これまでは Opt 解のサイズの評価として $Opt(G) \geq n/2$ という不等式をもちいた。ここでは頂点集合 $\mathcal{T}_\sigma$ をさらに U,W とに分割して、3 つの集合 U、W、$\mathcal{F}_\sigma$ に対してグラフ $G(U)$、$G(W)$、$G(\mathcal{F}_\sigma)$ それぞれの場合において頂点被覆に最低限必要な点の個数を考えることによって Opt 解の評価をする。このようにして評価された Opt 解のサイズの下界を n_V とする。ここで、U は与えられた完全マッチングのマッチング辺 (u, v) に対して $\mathcal{F}_\sigma$ の頂点とマッチしている頂点の集合、つまり $U = \{u|(u,v) \in M$ かつ $v \in \mathcal{F}_\sigma\}$ とする。W は残りの（$\mathcal{F}_\sigma$ とマッチしていない）頂点集合、つまり $W = V - \mathcal{F}_\sigma - U$ とする。また、$|\sigma|$ の極大性から $G(\mathcal{F}_\sigma)$ の辺にマッチング辺は存在しない。

5.1 $G(\mathcal{F}_\sigma)$ の分割

$w = 4$ の場合 $G(\mathcal{F}_\sigma)$ は次数が 3 以下のグラフになる。このグラフも閉路やパスや独立頂点に限らず様々な種類のグラフから構成されている。そこで我々のアルゴリズムでは、以下の 7 個の場合に即して $G(\mathcal{F}_\sigma)$ を分割する。それぞれの部分を順に $P_1, P_2, \cdots, P_7$ とし、その個数を $p_1, p_2, \cdots, p_7$ とする。$P1$ は独立頂点、$P7$ は完全グラフ K_4 とする。$P_2, P_3, \cdots, P_6$ はそれぞれ図 2 で描く。図 2 において、頂点に接続している点線の辺はあってもなくてもよい。また、辺と辺の間にある点線は、2 本の実線の辺がつながっていても、つながっていなくてもよいことを示す。ただし P_3 においては、右側の点線が存在する時は、他の 2 本の点線のうち少なくとも一方は存在するものとする。黒の頂点はこのアルゴリズムが TRUE に割り当てる頂点、白の頂点は残す頂点を表す。そして分けられた部分ごとに、頂点の個数、TRUE に割り当てられる頂点数、残る頂点の個数、頂点被覆に最低限必要な点の個数、重みの和の減少数を考える。頂点被覆に必要な頂点数については、それぞれの部分の外側の部分のグラフの形はどのようになるかは分からないので出る辺を除いて考える。重みの和の減少数についてはそれぞれの部分で考えられる最も大きな減少数を表 1 に書く。

	P_1	P_2	P_3	P_4	P_5	P_6	P_7
頂点の個数	1	2	3	3	4	4	4
TRUE に割り当てられる頂点数	0	1	2	2	3	3	3
残る頂点の個数	1	1	1	1	1	1	1
重みの和の減少数	0	3	4	5	3	5	6
頂点被覆に最低限必要な点の個数	0	1	1	2	1	2	3

表 1: $P_1, P_2, \cdots, P_7$ の頂点数など

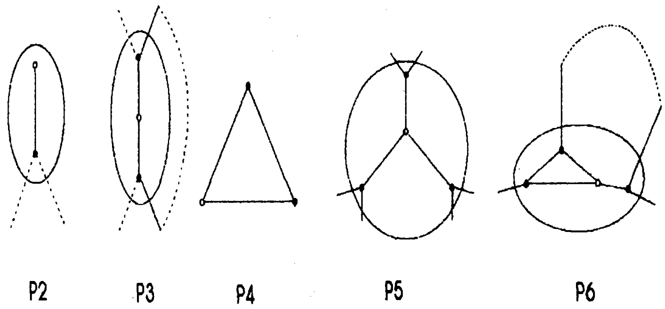

図 1: $G(\mathcal{F}_\sigma)$ の分割

5.2 得られる評価式

このように $G(\mathcal{F}_\sigma)$ を分割して頂点被覆 $C_\mathcal{F}$ を出力し、$\mathcal{T}_\sigma$ との和をとることにより G の頂点被覆 C を求める。表 1 を参考にすると、残る頂点の個数の合計を $n_\mathcal{F}$ とすれば、$n_\mathcal{F} = p_1+p_2+p_3+p_4+p_5+p_6+p_7$。すると $|C| = n - n_\mathcal{F}$ と書ける。また、$G(W)$ は完全マッチングをもつので、$G(W)$ を被覆するには少なくとも $|W|/2$ 個の頂点を必要することと、$G(U)$ を被覆するのに少なくとも $Opt(G(U))$ 個の頂点を必要とすることから、G の頂点被覆には $n_V = p_2 + p_3 + 2p_4 + p_5 + 2p_6 + 3p_7 + |W|/2 + Opt(G(U))$ 個以上の頂点を必要とする。従って $Opt(G)$ に関しては $\max(n/2, n_V)$ とすればよいのだが、$Opt(G(U))$ が分からないので、$\max(n/2, n_V)$ も分からない。よって、ここでは次の 2 つの評価式を与えておく。$|C| = n - n_\mathcal{F}$ と

$Opt(G) \geq n/2$ から、

$$\frac{|C|}{Opt(G)} \leq \frac{n - n_{\mathcal{F}}}{n/2}$$

また $|C| = n - n_{\mathcal{F}}$ と $Opt(G) \geq n_V$ から、

$$\frac{|C|}{Opt(G)} \leq \frac{n - n_{\mathcal{F}}}{n_V}$$

が得られる。

次に、充足節の重みの和の減少数 δn については、表を参考にすれば、最も大きくなる場合で $3p_2 + 4p_3 + 5p_4 + 3p_5 + 5p_6 + 12p_7$ である。これを α とおけば、$\delta n \leq \alpha$ であることと補題 3.1 [1]($w = 4$ とする) から

$$\begin{aligned}\frac{|C|}{Opt(G)} &\leq 1 + \frac{\delta}{2} + \frac{r-1}{4r}(\bar{d} + 4) \\ &\leq 1 + \frac{\alpha}{2n} + \frac{r-1}{4r}(\bar{d} + 4)\end{aligned}$$

となる。

一方、我々は次のような G の頂点被覆 C もまた同時に考える。G の頂点を $\mathcal{F}_\sigma$、U、W、に分割するところまでは同じである。G に対する新しい頂点被覆方法として、$\mathcal{F}_\sigma \cup W$ の任意の頂点 v に対し、x_v を TRUE に割り当て、U の任意の頂点 u に対しては、x_u を FALSE に割り当て直す。そして $\mathcal{F}_\sigma \cup W$ と $G(U)$ の頂点被覆 C_U との和をとることにより G の頂点被覆 C を求める。この時 C_U を求める際には $G(U)$ に対して既存の（多項式時間で終了する）頂点被覆アルゴリズムを施す。この時アルゴリズムが残した頂点の個数を n_U とする。このアルゴリズムの近似度を $s \leq 2$ とすると、このアルゴリズムの出力する頂点の個数は高々 $sOpt(G(U))$ である。よって、少なくとも $|U| - sOpt(G(U))$ 個の頂点を残す。従って、$n_U \geq |U| - sOpt(G(U))$ が言える。G においては、$|C| = n - n_U$ となる。この等式と $Opt(G) \geq n/2$ から

$$\frac{|C|}{Opt(G)} \leq \frac{n - n_U}{n/2}$$

また $Opt(G) \geq n_V$ とすれば

$$\frac{|C|}{Opt(G)} \leq \frac{n - n_U}{n_V}$$

となる。

これらの事実から、以上 5 つの近似度が得られる。実際は、これらの近似度のなかで最も小さい比を採用すればよい。

$$\min\left(1 + \frac{\alpha}{2n} + \frac{(r-1)(\bar{d}+4)}{4r}, \frac{n - n_{\mathcal{F}}}{n/2}, \frac{n - n_{\mathcal{F}}}{n_V}, \frac{n - n_U}{n/2}, \frac{n - n_U}{n_V}\right)$$
$$< 1.6010 + 0.0074\bar{d} \quad (\bar{d} > 20.16)$$

が言えるから、近似度 $1.6010 + 0.0074\bar{d}$ $(\bar{d} > 20.16)$ が得られる。

$21 \leq \bar{d} \leq 31$ における近似度の比較を表 2 に示した。近似度 1 が、今村と岩間のアルゴリズムの解析から求めた近似度である。$21 \leq \bar{d} \leq 26$ では重みは 4、$27 \leq \bar{d} \leq 31$ では重みは 5 としている。近似度 2 は我々が求めた近似度である。

d	21	22	23	24	25	26	27	28	29	30	31
近似度 1	1.772	1.780	1.786	1.793	1.800	1.807	1.814	1.818	1.823	1.828	1.832
近似度 2	1.756	1.764	1.771	1.778	1.786	1.793	1.801	1.808	1.815	1.823	1.830

表 2: 近似度の比較

6 まとめ

今村と岩間のアルゴリズムを重みの設定が 4 以上の場合を解析し、平均次数が任意に固定された時にも近似度を抑えられることを示した。また、*Opt* 解のサイズに依存した新しいアルゴリズムを提案することにより、平均次数が 20 から 30 程度の場合に、今村と岩間のアルゴリズムを解析した場合の近似度を改良することができた。今回、我々のアルゴリズムの評価は、重みの設定が 4 の場合に絞って行われた。今後は、重みが 5 以上の場合も解析していきたい。

謝辞

本稿を書くにあたり協力して頂いた築地 立家先生には深く感謝しています。

参考文献

[1] T. Imamura, K. Iwama, "完全マッチングを持つグラフに対する最小頂点被覆問題の近似解法", 2002 年度 夏の LA シンポジウム, 5-1-5-9, 2002.

[2] Jianer Chen and Iyad Kanj, "On approximating minimum vertex cover for graphs with perfect matching", *International Symposium on Algorithms and Computation*, pp.132-143, 2000.

[3] Uriel Feige and Michel X. Goemans, "Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT", *Proceedings of the Third Israel Symposium on Theory of Computing and Systems*, 1995.