

負荷分散枝被覆問題に対する最適性とアルゴリズム

原田 雄太* 小野 廣隆† 定兼 邦彦† 山下 雅史†

* 九州大学大学院システム情報科学府

† 九州大学大学院システム情報科学研究所

1 はじめに

グラフ $G = (V, E)$ においてマッチングとは互いに端点を共有しない枝の集合であり、グラフ理論において非常に重要な概念の 1 つとして捉えられている。特に最大マッチングについては本研究と深く関わりがあり、盛んに研究がなされている [2, 5]。マッチングではどの枝にも接続していない自由頂点の存在を許すが、それに対して全ての頂点が少なくとも 1 つの枝に接しているような枝の集合は枝被覆と呼ばれる。枝被覆はマッチングと密接な関係を持つことが知られており、最大マッチングから最小枝被覆への変換を行う多項式時間アルゴリズムの存在が示されている [3, 6]。この最小 (極小) 枝被覆は各連結成分がスターになる構造を持っている。そこで枝被覆の応用として、ユーザーが互いに直接通信し合うような P2P ネットワークなどにおいて、各連結成分をクラスタとみなすことによるクラスタリングなどの応用が考えられる。

しかしながら、最小枝被覆では各クラスタにおける負荷 (スターの枝数) の分散は考慮されておらず、負荷があるクラスタへ集中する恐れがある。負荷の偏りを避けるために、本研究では、各クラスタの負荷が均等化されるような枝被覆を求めることを目的とする。具体的にはグラフ G の各頂点 $v \in V$ に対する次数の 2 乗和が最小となる枝被覆 (負荷分散枝被覆) を求める。負荷分散枝被覆の最適性として、枝数が最小になる性質 (枝最小性) と各頂点の次数が均等化される性質 (負荷均等性) を持ち、これらの性質を証明

する。また負荷分散枝被覆を得る $O(|V||E|)$ 時間のアルゴリズムを提案する。

関連研究として、2 部グラフ $G = (U \cup V, E)$ におけるセミマッチングの負荷分散について研究がなされている [4]。セミマッチングとは U の各頂点が V の 1 つの頂点と対をなすような枝の集合 M であり、セミマッチング M において V の各頂点の次数の均等化が実現されている。セミマッチングの応用としてはタスクスケジューリングなどが挙げられる [1]。つまり、頂点集合 U と V をそれぞれタスクと計算機とみなすとき、セミマッチングはタスクの各計算機への割り当てに対応する。またセミマッチングも極小枝被覆と同様に各連結成分がスターになる構造を持っており、本研究で扱う問題はセミマッチングに対する一般グラフへの拡張として捉えられる。

以後の構成は次のようになっている。次節では各種の準備と本研究で扱う問題を与える。第 3 節では負荷分散枝被覆の最適性 (枝最小性と負荷均等性) に対する必要十分条件を示す。第 4 節ではアルゴリズムを提案し、その正当性と計算時間を与える。最後に第 5 節でまとめと今後の課題を述べる。

2 準備と問題

$G = (V, E)$ を頂点集合 V と枝集合 E を持つ単純無向グラフとする。 E の各枝は 2 頂点 $u, v \in V$ の対 $\{u, v\}$ で表される。各集合のサイズは $|V| = n$, $|E| = m$ とする。部分集合 $E' \subseteq E$ と頂点 v に対して、 $\delta_{E'}(v) = \{\{u, v\} \in E'\}$, $\deg_{E'}(v) =$

$|\delta_{E'}(v)|$ とする. $\delta_{E'}(v)$ は頂点 v と接続している枝の集合, $\deg_{E'}(v)$ は v の次数である. E に対しては $\deg(v) = \deg_E(v)$ を用いる.

枝集合 $E_c \subseteq E$ に対して, V の全ての頂点が少なくとも E_c の 1 つの枝と接続しているとき, E_c を枝被覆と呼ぶ. G に孤立点が存在すれば枝被覆は存在しない. したがって, 以後 G を孤立点を持たないグラフとする. また枝被覆 E_c に含まれる枝 $e \in E_c$ を被覆枝, 含まれない枝 $e \notin E_c$ を非被覆枝と呼ぶ. 枝被覆は Norman と Rabin により研究がなされており, 彼らは最大マッチングから最小枝被覆へ, またその逆の変換を行う多項式時間アルゴリズムが存在することを示している [6]. これは以下の命題が成立することを意味する.

命題 1 ([6]) 全ての最大マッチングはある最小枝被覆に含まれる. また, 全ての最小枝被覆はある最大マッチングを含む. $\square$

この結果と同様の証明により, Gallai の定理が成り立つことが知られている. 即ち $\nu(G) + \rho(G) = n$ が成立する [3]. ここで, $\nu(G)$ と $\rho(G)$ はそれぞれ G における最大マッチングと最小枝被覆の枝数である. 実際に, 最大マッチング M において, M に被覆されていない各頂点 v に対して v と接続している任意の枝を 1 つ M へ加えることで $O(m)$ 時間で最小枝被覆を得ることができる. したがって, 最小枝被覆は最大マッチングアルゴリズムと同等の計算量で求めることができ, 最大マッチングを得るアルゴリズムとしては $O(\sqrt{nm})$ 時間のアルゴリズムが知られている [5, 7].

枝被覆の極小性についての命題を以下に与える. 証明は単純なものであり, 省略する.

命題 2 枝集合 E_c が全域森でありかつ各連結成分がスター (つまり, ある自然数 r に対して $K_{1,r}$) である時またその時に限り E_c は極小枝被覆である.

極小枝被覆 E_c の部分集合 E'_c において, 次数 1 の頂点と隣接している頂点を中心頂点, また次

数 1 の頂点を葉頂点と呼ぶ. 言い換えると, E'_c の各スター $K_{1,r}$ において, $r > 1$ のとき次数が 2 以上である頂点が中心頂点, 次数 1 の頂点が葉頂点である. $r = 1$ の場合, スターには 2 頂点のみ存在し, その両頂点とも中心頂点かつ葉頂点となる. また, E'_c のどの枝にも接続しない頂点を自由頂点と呼ぶ. 中心, 葉, 自由頂点の各集合をそれぞれ $C_{E'_c}$, $L_{E'_c}$, $F_{E'_c}$ とする. 極小枝被覆 E_c においては $F_{E_c} = \emptyset$ が満たされる. $L_{E'_c} \cup F_{E'_c} = V$ のとき, E'_c はマッチングである. 特に, $L_{E'_c} = V$ であれば E'_c は完全マッチングとなる.

枝被覆 E_c が与えられたとき, 頂点 v のコストを $\text{cost}(v) = \deg_{E_c}^2(v)$ と定義する. コストの総和は $c(E_c) = \sum_{v \in V} \text{cost}(v)$ で表す. このとき, 負荷分散枝被覆問題は以下のように与えられる.

負荷分散枝被覆問題

入力: 単純無向グラフ $G = (V, E)$,

出力: コストの総和 $c(E_c)$ が最小な枝被覆 E_c .

コストの総和が最小となる枝被覆を負荷分散枝被覆と呼ぶ. 負荷分散枝被覆を求めることが本研究の目的である. 負荷分散枝被覆は枝数と最大次数が最小化される性質を持ち, 次節で証明する.

3 負荷分散枝被覆の最適性

本節では, 負荷分散枝被覆を各種のパスにより特徴付けし, その最適性を示す.

枝被覆 E_c において, 被覆枝と非被覆枝を交互にたどるようなパス $P = (\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{k-1}, v_k\})$ を交互パスと呼ぶ. $v_1 = v_k$ である場合には交互閉路と呼ぶ. パス P に含まれる枝集合を E_P で表す. 便宜のため, 交互パス・閉路を頂点の列 $P = (v_1, \dots, v_k)$ として扱う. 交互パスは今後紹介していく各種パスの基本となる.

集合 A と B の対称差を $A \oplus B = (A \setminus B) \cup (B \setminus A)$ と表記する。パス P が枝被覆 E_c に対する交互パス・閉路であれば、 P に沿って被覆枝と非被覆枝の切り替えを行うことにより新たな枝集合 $E_P \oplus E_c$ が得られる。 P がパスである場合、 $v \in \{v_1, v_k\}$ に対して少なくとも $\delta_{E_c \cap E_P}(v) = 0$ または $\delta_{E_c \setminus E_P}(v) \neq 0$ のどちらか一方でも成り立つならば、パスの切り替え後も v_1 と v_k は被覆が保たれ、 $E_P \oplus E_c$ も枝被覆となる。その他の頂点 $v_2, \dots, v_{k-1}$ において被覆が保たれることは P が交互パスであることから明らかである。 P が閉路であるとき、 $E_P \oplus E_c$ が枝被覆である条件は $\{v_1, v_2\} \notin E_c$, $\{v_{k-1}, v_k\} \notin E_c$, $\delta_{E_c \setminus E_P}(v_1) \neq 0$ の条件のうち1つでも満たされることである。

3.1 枝最小性

最小枝被覆は枝減少パス・閉路によって特徴付けされることが証明されており [6], それらを紹介する。また、負荷分散枝被覆の枝数が最小になる性質 (枝最小性) を示す。

枝減少パス 枝被覆 E_c に対する交互パス $P = (v_1, v_2, \dots, v_{2k-1}, v_{2k})$ において、両端の枝 $\{v_1, v_2\}$ と $\{v_{2k-1}, v_{2k}\}$ が被覆枝でありかつ $v_1, v_{2k} \in C_{E_c} \setminus L_{E_c}$ であるとき P を枝減少パスと呼ぶ。枝減少パスは単純パスに限らない。パスの例を図1に示す。 $v_1, v_{2k} \in C_{E_c} \setminus L_{E_c}$ より $\delta_{E_c \setminus E_P}(v_1) \neq 0$ と $\delta_{E_c \setminus E_P}(v_{2k}) \neq 0$ が満たされることから、枝減少パス P の切り替えにより得られる枝集合 $E_P \oplus E_c$ は枝被覆である。このとき $|E_P \setminus E_c| = |E_P \cap E_c| + 1$ であるので、枝被覆のサイズは1減少する。つまり $|E_P \oplus E_c| = |E_c| - 1$ である。また v_1 と v_{2k} の次数も1減少する。その他の頂点の次数に変化はない。

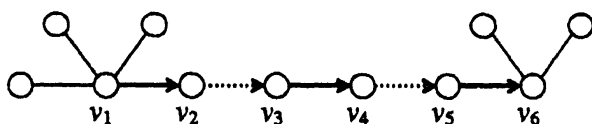


図 1: 枝減少パス

枝減少閉路 枝減少閉路 $P = (v_1, \dots, v_{2k})$ において、 $\forall \{v_{2i-1}, v_{2i}\} \in E_c, \forall \{v_{2i}, v_{2i+1}\} \notin E_c$ と $\delta_{E_c \setminus E_P}(v_1) \neq 0$ が満たされるとき、 P を枝減少閉路と呼ぶ。同様に、閉路 P の切り替えにより枝被覆 $E_P \oplus E_c$ が得られる。切り替えの際に、枝数は1減少し、 $v_1(v_{2k})$ の次数は2減少する性質を持つ。図2は枝減少パスの例である。

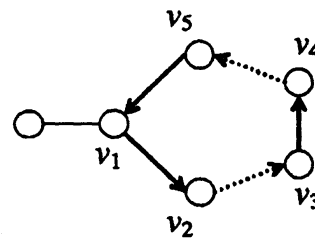


図 2: 枝減少閉路

次の定理は最小枝被覆の最適性を与えるものであり、Norman らによって証明されている。これにより負荷分散枝被覆の枝最小性が導かれる。

定理 3 ([6]) 枝被覆 E_c において枝減少パスと枝減少閉路がどちらも存在しない時またその時に限り E_c は最小枝被覆である。□

系 4 負荷分散枝被覆は最小枝被覆である。

証明 もし枝被覆 E_c に対する枝減少パスもしくは閉路 P が存在すれば、 P の切り替え操作により始点と終点の次数は減少されるので、コストの総和も減少することになる。したがって、負荷分散枝被覆 E_c には枝減少パス・閉路は存在せず、定理3より E_c は最小枝被覆である。□

3.2 負荷均等性

枝被覆 E_c のコストの総和 $c(E_c)$ が最小であるとき、各頂点の次数が均等化される、つまり最大次数が最小化されることを示す。この性質 (負荷均等性) は以下で与えるコスト減少パスにより特徴付け可能である。

頂点交互パス 極小枝被覆 E_c において、頂点交互パスを次のように定義する。交互パス $P = (c_1, l_1, c_2, \dots, c_{k-1}, l_{k-1}, c_k)$ であり、各 i において $c_i \in C_{E_c}$, $l_i \in L_{E_c}$, $\{c_i, l_i\} \in E_c$, $\{l_i, c_{i+1}\} \notin E_c$ を満たす。つまり頂点交互パスは中心頂点と葉頂点を交互にたどる交互パスである。 $\delta_{E_c \setminus E_P}(c_1) \neq \emptyset$ ($c_i \in C_{E_c} \setminus L_{E_c}$) を満たす場合、 P の切り替えにより得られる枝集合 $E_P \oplus E_c$ は枝被覆であり、また極小でもある。その理由は頂点交互パス P が中心と葉を交互にたどるパスであるので $E_P \oplus E_c$ に長さ3のパスが生成されないことにある。また、始点 c_1 の次数は1減少し、終点 c_k の次数は1増加する性質を持つ。

コスト減少パス 枝被覆 E_c に対する頂点交互パス $P = (c_1, \dots, c_k)$ において、 $\deg_{E_c}(c_1) > \deg_{E_c}(c_k) + 1$ を満たすパス P をコスト減少パスと呼ぶ。図3にコスト減少パスの例を挙げる。文字通り、コスト減少パス P に沿って枝の切り替えを行うとコストの総和が減少する。つまり $c(E_P \oplus E_c) < c(E_c)$ が成り立つ。これは以下の式(1)が成立することからも明らかである。

$$c(E_P \oplus E_c) = c(E_c) - 2(\deg_{E_c}(c_1) - \deg_{E_c}(c_k) - 1) \quad (1)$$

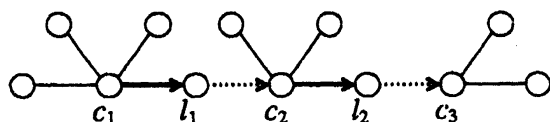


図 3: コスト減少パス

負荷分散枝被覆に対する必要十分条件を与える。まず準備として以下の補題を証明する。

補題 5 E_c と E_c^* を異なる最小枝被覆とする。但し、その対称差 $E_c \oplus E_c^*$ により構成されるグラフ $G' = (V, E_c \oplus E_c^*)$ において以下の2つの条件を満たすものとする。

1. $\deg_{E_c}(c_1) > \deg_{E_c^*}(c_1)$ を満たす頂点 $c_1 \in C_{E_c}$ が存在する。

2. 条件1の任意の c_1 に対して、 G' 上の E_c に対する全ての頂点交互パス $P = (c_1, \dots, c_k)$ は $\deg_{E_c^*}(c_k) \leq \deg_{E_c^*}(c_1)$ を満たす。

このとき、 G' 上に E_c に対するコスト減少パス $P = (c_1, \dots, c_k)$ が必ず存在する。

証明 E_c と E_c^* を上記2つの条件を満たす最小枝被覆とし、その対称差で構成されるグラフを G' とする。 G' において、 $\deg_{E_c}(c_1) > \deg_{E_c^*}(c_1)$ を満たす頂点 $c_1 \in C_{E_c}$ から次のような手順で頂点交互パス $P = (c_1, l_1, \dots, l_{k-1}, c_k)$ を構築する。構築中に既にたどった被覆枝の集合を Q とする。最初は $Q = \emptyset$ である。(1) 各中心頂点 $c_i \in C_{E_c}$ において、もし $\delta_{E_c \setminus E_c^*}(c_i) \setminus Q \neq \emptyset$ かつ $\deg_{E_c}(c_1) \leq \deg_{E_c}(c_i) + 1$ を満たしていれば、 c_i から任意の枝 $\{c_i, l_i\} \in E_c \setminus E_c^*$ をたどり、その枝を Q に加える。始点 c_1 は必ずどちらの条件も満たす。どちらか一方でも満足していなければ、 c_i をパス P の終点として終了する。(2) 各葉頂点 $l_i \in L_{E_c}$ において、 l_i に接続している任意の枝 $\{l_i, c_{i+1}\} \in E_c^* \setminus E_c$ をたどる。そのような枝は必ず存在する。もし $\delta_{E_c^* \setminus E_c}(l_i) = \emptyset$ であれば、 E_c^* は l_i を被覆しておらず、枝被覆であることに反する。また、葉への枝 $\{l_i, l'_i\} \in E_c^* \setminus E_c$ しか存在しなければ、 E_c が最小枝被覆であることに矛盾する。つまり、 l'_i と接続している E_c のある被覆枝を $\{l'_i, c'_i\} \in E_c$ とすると、パス $(c_1, \dots, l_i, l'_i, c'_i)$ は枝減少パスになる。なぜなら、 $l'_i \in L_{E_c} \setminus C_{E_c}$ 、即ち $c'_i \in C_{E_c} \setminus L_{E_c}$ であり、 $\deg_{E_c}(c'_i) > 1$ が満たされる。

上記の構築手順において各枝をたどる回数は高々1度であり、必ず終了する。構築した頂点交互パス $P = (c_1, \dots, c_k)$ において、手順(1)でパス P の構築が終了するケースは以下の(i)と(ii)に分けられる。(i) $\deg_{E_c}(c_1) > \deg_{E_c}(c_k) + 1$ を満たす場合。この場合、パス P が E_c におけるコスト減少パスであることは明らかである。(ii) $\delta_{E_c \setminus E_c^*}(c_k) \setminus Q = \emptyset$ の場合。もし $c_1 = c_k$ であれば $\deg_{E_c}(c_1) = \deg_{E_c^*}(c_1)$ となり、条件1に反する。よって $c_1 \neq c_k$ であり、パスの構築手順により最後に $E_c^* \setminus E_c$ の枝 $\{l_{k-1}, c_k\}$ をたどること

から $|\delta_{E_c \setminus E_c^*}(c_k)| < |\delta_{E_c^* \setminus E_c}(c_k)|$ が満たされる。したがって $\deg_{E_c}(c_k) < \deg_{E_c^*}(c_k)$ が成り立つ。また条件 1 と 2 より, $\deg_{E_c}(c_1) > \deg_{E_c^*}(c_1)$ と $\deg_{E_c^*}(c_k) \leq \deg_{E_c^*}(c_1)$ が満たされる。これら 3 式の組み合わせにより $\deg_{E_c}(c_k) < \deg_{E_c^*}(c_k) \leq \deg_{E_c^*}(c_1) < \deg_{E_c}(c_1)$ が得られ, $\deg_{E_c}(c_1) > \deg_{E_c}(c_k) + 1$ が成立する。即ち P は E_c におけるコスト減少パスとなる。故に, G' に必ずコスト減少パスは存在する。□

上記の補題 5 により, 負荷分散枝被覆の最適性に関する以下の定理 6, 7 が示される。

定理 6 最小枝被覆 E_c においてコスト減少パスが存在しない時またその時に限り, E_c は負荷分散枝被覆である。

証明 グラフ G を負荷分散枝被覆問題の入力とし, E_c を G における最小枝被覆とする。 E_c が負荷分散枝被覆であればコスト減少パスが存在しないことは明らかである。よって E_c が非負荷分散枝被覆, つまりコストの総和 $c(E_c)$ が最小でなければ, 必ずコスト減少パスが存在することを示す。

E_c をコストの総和 $c(E_c)$ が最小でない最小枝被覆, E_c^* を E_c との対称差のサイズ $|E_c \oplus E_c^*|$ が最小である負荷分散枝被覆とする。また, その対称差 $E_c \oplus E_c^*$ の枝で構成される G の部分グラフを $G' = (V, E_c \oplus E_c^*)$ とする。このとき, G' が補題 5 の 2 つの条件を満たすことを示す。

E_c と E_c^* は両者とも最小枝被覆であるので, $\sum_{v \in V} \deg_{E_c}(v) = \sum_{v \in V} \deg_{E_c^*}(v) = 2\rho(G)$ である。加えて, $c(E_c) > c(E_c^*)$ であることを考慮すると, $\deg_{E_c}(c_1) > \deg_{E_c^*}(c_1)$ を満たす中心頂点 $c_1 \in C_{E_c}$ が必ず存在する (条件 1)。ここで, c_1 を始点とする G' 上の E_c に対する頂点交互パス $P = (c_1, \dots, c_k)$ を考える。 P と逆向きのパス $\bar{P} = (c_k, \dots, c_1)$ は E_c^* における頂点交互パスとなる。もし $\deg_{E_c^*}(c_k) > \deg_{E_c^*}(c_1) + 1$ であれば, $\bar{P}$ は E_c^* に対するコスト減少パスとなり, E_c^* のコストが最小であることに矛盾する。 $\deg_{E_c^*}(c_k) = \deg_{E_c^*}(c_1) + 1$ の場合, $\bar{P}$ に沿っ

て枝の切り替えを行うことにより, E_c との対称差がより小さな負荷分散枝被覆 $E_c^* \oplus E_{\bar{P}}$ が得られ, 再び E_c^* の選択に矛盾する。したがって $\deg_{E_c^*}(c_k) \leq \deg_{E_c^*}(c_1)$ が成立する (条件 2)。故に, 補題 5 の 2 つの条件を満たし, G' 上にコスト減少パスが存在する。□

定理 7 枝被覆 E_c が負荷分散枝被覆であれば, 最大次数 $\max_{v \in V} \{\deg_{E_c}(v)\}$ は最小である。

証明 非最小枝被覆には枝減少パス・閉路が存在し, 始点と終点の次数を減らすことが可能である。したがって, 最小枝被覆の中に最大次数が最小となる枝被覆が必ず存在する。ここで, 最大次数が最小でない最小枝被覆を E_c , 最大次数が最小でありかつ E_c との対称差が最小な最小枝被覆を E_c^* とする。また, $E_c \oplus E_c^*$ で構成されるグラフを $G' = (V, E_c \oplus E_c^*)$ とする。このとき, G' に必ずコスト減少パスが存在することを示せばよい。

E_c において最大次数を持つ頂点を $c_1 \in C_{E_c}$ とすると, E_c と E_c^* の決め方により $\deg_{E_c}(c_1) > \deg_{E_c^*}(c_1)$ が成立する。 G' において, E_c に対する任意の頂点交互パスを $P = (c_1, \dots, c_k)$ とする。 $\bar{P} = (c_k, \dots, c_1)$ は E_c^* における頂点交互パスである。 $\deg_{E_c^*}(c_k) > \deg_{E_c^*}(c_1) + 1$ を満たす場合, もし $V \setminus \{c_k\}$ に $\deg_{E_c^*}(c) \geq \deg_{E_c^*}(c_k)$ を満たす頂点 c が存在すれば, $\bar{P}$ の切り替えにより得られる枝被覆 $E_c^* \oplus E_{\bar{P}}$ は最大次数が最小であり, かつ E_c との対称差がより小さな枝被覆となる。これは対称差の最小性に矛盾する。上記のような頂点 c が存在しなければ, $\bar{P}$ の切り替えにより最大次数が 1 小さな枝被覆が得られることになり, E_c^* の最大次数が最小であることに反する。 $\deg_{E_c^*}(c_k) = \deg_{E_c^*}(c_1) + 1$ の場合, $\bar{P}$ の切り替えによる最大次数の変化はなく, 対称差の最小性に反する。よって $\deg_{E_c^*}(c_k) \leq \deg_{E_c^*}(c_1)$ が成立する。したがって, 補題 5 より G' に必ずコスト減少パスが存在する。□

4 提案アルゴリズム

前節で負荷分散枝被覆の最適性が最小枝被覆におけるコスト減少パスの有無と関係があることを示した (定理 6). この性質を用いることで直ちに以下のアルゴリズム *BEC1* が得られる. 即ち, コスト減少パスが存在する限りパスの探索と切り替えを繰り返すという手法である. アルゴリズムの手順を以下に与える.

アルゴリズム *BEC1*:

Step 1: 最小枝被覆 E_c を求める.

Step 2: E_c に対するコスト減少パス P を探索する. もしパス P が存在しなければ E_c を出力する.

Step 3: コスト減少パス P に沿って被覆枝と非被覆枝の切り替えを行い, Step 2 へ.

図 4: *BEC1* の手順

第 2 節で紹介したアルゴリズムを用いることで, Step 1 で最小枝被覆 E_c が得られる. Step 2 ではコスト減少パスの探索するが, 探索は次のような手順で行う. まず, E_c において最大次数を持つ頂点 $c_1 \in C_{E_c}$ を選択し, c_1 を根とする交互探索木 T を構築する. 探索木 T は C_{E_c} から L_{E_c} への被覆枝とその逆向きの非被覆枝により構成される. 次に, 幅優先探索により $\deg_{E_c}(c_1) > \deg_{E_c}(c_k) + 1$ を満たす頂点 $c_k \in C_{E_c}$ を見つけ, c_k をパス P の終点とする. c_1 からのコスト減少パスが存在しない場合は, T に存在しない頂点の中で最大次数を持つ頂点から続けてパスの探索を行うことができる. T に存在する頂点からのパスを探索する必要はない. その理由は以下の補題 8 で示される. V の全ての頂点が調べられるとこのステップは終了である. このとき, コスト減少パスが存在しなければ E_c を出力する. 最後に, コスト減少パス P に沿って被覆枝と非被覆枝の切り替えることよりコストの総和 $c(E_c)$ を減少させ, 再びパスの探索を繰り返す.

補題 8 パス $P = (c_1, \dots, c_k)$ を $\deg_{E_c}(c_1) \geq \deg_{E_c}(c_k)$ を満たす頂点交互パスとする. このとき, c_1 を始点とするコスト減少パスが存在しなければ, c_k を始点とするコスト減少パスも存在しない.

証明 対偶を示す. c_1 から c_k への頂点交互パス P が存在するので, c_k を始点とする全てのコスト減少パスは c_1 を始点とするコスト減少パスへ拡張可能である. $\square$

定理 9 *BEC1* により $O(\min\{n^{3/2}, m\} m)$ 時間の計算量で負荷分散枝被覆が得られる.

証明 アルゴリズムの実行中, 枝被覆 E_c は枝の最小性を保持する. 加えて, アルゴリズムの終了時にコスト減少パスが存在しないことを考慮すると, 定理 6 より *BEC1* は負荷分散枝被覆を出力する. 続いて, アルゴリズムの計算量を与える. 計算量は Step 2 と 3 におけるコスト減少パス探索の反復に依存する. 各反復に対する計算量は, 幅優先探索により各枝を高々 1 度探索すれば十分であることから $O(m)$ 時間である. 各頂点 v において, v を始点とするコスト減少パスが切り替えられる回数は多くとも $\deg(v)$ 回である. なぜならば, 頂点 v が 1 度最大次数を持つと, 以後その頂点の次数は単調減少するからである. 全頂点对する次数の合計値は反復回数の上限であり, その値は $\sum_{v \in V} \deg(v) = 2m$ となる. したがって, Step 2 と 3 の反復回数は $O(m)$ 回で抑えられる. 同時に, セミマッチングの負荷分散問題に対して反復回数の他の上限が与えられている [4]. その値は負荷分散枝被覆問題に対しても適用可能であり, $O(n^{3/2})$ 回である. 故に, アルゴリズムの計算量は全体で $O(\min\{n^{3/2}, m\} m)$ 時間である. $\square$

改良アルゴリズム より高速なアルゴリズム *BEC2* を提案する. *BEC2* は次数の均等を保ちつつ, つまりコスト減少パスが存在しない状況を保ちながら, 最大マッチングを負荷分散枝被覆へ拡張していく手法である. アルゴリズムの動作を図 5 に示す.

アルゴリズム BEC2:

Step 1: 最大マッチング M を求める.

Step 2: $F_M \neq \emptyset$ であれば, 自由頂点 $v_1 \in F_M$ を任意に選択する. そうでなければ, M を返す.

Step 3: v_1 を根とする交互探索木 T を構築する. 探索木 T は $L_M \cup \{v_1\}$ から C_M への非被覆枝と C_M から L_M への被覆枝により構成される.

Step 4: T 上で次数 $\deg_M(v_{2k})$ が最小となる頂点 $v_{2k} \in C_M$ への交互パス $P = (v_1, \dots, v_{2k})$ を探索する. そのようなパス P を増加パスと呼ぶ.

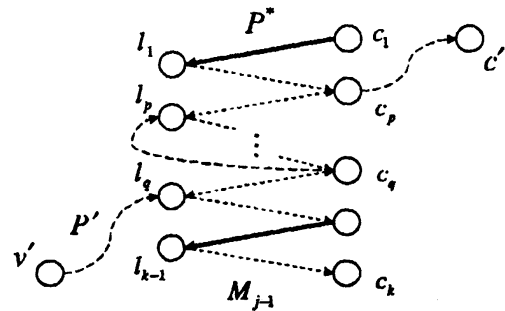
Step 5: 増加パス P の切り替えにより M を拡張し, Step 2 へ.

図 5: BEC2 の手順

BEC2 では初めに最大マッチング M を求め, 増加パスの探索による枝集合 M の拡張を繰り返す. 増加パスは自由頂点 $v_1 \in F_M$ を持つ非被覆枝 $\{v_1, v_2\}$ と頂点交互パス $(v_2, \dots, v_{2k})$ の結合からなるパスである. 各繰り返しにおける増加パス P の切り替えでは, $\{v_1, v_2\} \in M \oplus E_P$ と $|E_P \setminus M| = |E_P \cap M| - 1$ が満たされるので, 始点 v_1 は被覆され, 枝数は 1 増加する. このとき自由頂点数は 1 減少し, 自由頂点が存在しなくなれば終了する. よって出力される枝集合は枝被覆である. 一方, M の拡張においてコスト減少パスが生成されないことを次の補題 10 で示す. この補題はアルゴリズムの正当性に関わる.

補題 10 BEC2 の実行中にコスト減少パスが生成されることはない.

証明 背理法で証明する. BEC2 の各反復において構築される被覆枝の集合を順番に $M_0, M_1, \dots, M_i, \dots, M_{|F_{M_0}|}$ とする. M_0 は Step 1 において求められる最大マッチング, M_i は i 回目の反復により構築される被覆枝の集合であり, $|M_i| = |M_{i-1}| + 1$ が満たされる. ここで,

図 6: M_{j-1} における P^* と P' (補題 10)

M_j に初めてコスト減少パスが生成されると仮定し, そのパスを $P^* = (c_1, l_1, \dots, l_{k-1}, c_k)$ とする. M_{j-1} は P^* が生成される直前の被覆枝集合であり, M_{j-1} において探索される増加パスを $P' = (v', \dots, c')$ とする. 増加パス P' の切り替えにより M_{j-1} から M_j へ拡張される. このとき P^* 上で M_j に被覆枝として加えられる枝集合を Q とする. つまり, $Q = (M_j \setminus M_{j-1}) \cap E_{P^*}$ とする. P^* は生成される最初のコスト減少パスであるので $Q \neq \emptyset$ である. そうでなければ M_{j-1} に P^* が存在することになり, P^* の仮定に反する. また P^* は P' の切り替えにより生成されるパスであるので $Q \subseteq E_{P'}$ である. 一方, P^* の始点の次数を $\deg_{M_j}(c_1) = x$ とおくと, P^* はコスト減少パスであるので, 終点 c_k において $\deg_{M_{j-1}}(c_k) \leq \deg_{M_j}(c_k) \leq x - 2$ が満たされる. (各頂点 v の次数は単調増加であるので $\deg_{M_{j-1}}(v) \leq \deg_{M_j}(v)$ を満たす.) このような状況において矛盾を示す. 図 6 は M_{j-1} における P^* と P' の関係を表している.

$\{c_p, l_p\}$ を Q に含まれる枝の中で P^* 上で c_1 に最も近い枝とする. $Q \subseteq E_{P'}$ より $\{c_p, l_p\}$ は P' に含まれるので, M_{j-1} において c_1 から c' への頂点交互パス $(c_1, \dots, c_p, \dots, c')$ が存在する. P^* が最初に生成されるコスト減少パスであることから, このパスはコスト減少パスではない. したがって $c' \neq c_1$ の場合, $\deg_{M_{j-1}}(c') \geq \deg_{M_{j-1}}(c_1) - 1 = \deg_{M_j}(c_1) - 1 = x - 1$ が成り立つ. $c' = c_1$ の場合は P' の切り替えにより $c'(c_1)$ の次数が 1 増加することから $\deg_{M_{j-1}}(c') = \deg_{M_j}(c') - 1 = x - 1$ が成立

する。つまり、増加パス P' は終点 d' において $\deg_{M_{j-1}}(d') \geq x-1$ を満たすパスである。

一方、 Q の枝の中で P^* において c_1 から最も遠い枝を $\{c_q, l_q\}$ とする。 $Q \subseteq E_{P'}$ よりこの枝も P' 上の枝であることから、 M_{j-1} に対する交互パス $(v', \dots, l_q, \dots, c_k)$ が存在する。このパスは M_{j-1} において構築される交互探索木に含まれるパスである。それに加えて、終点 c_k において $\deg_{M_{j-1}}(c_k) \leq x-2$ であることを考慮すると $\deg_{M_{j-1}}(c_k) < \deg_{M_{j-1}}(d')$ が満たされる。これは増加パス P' の終点として次数が最小である頂点が選ばれることに矛盾する。故に、アルゴリズムの実行中にコスト減少パス P^* が生成されることはない。 $\square$

定理 11 $BEC2$ により $O(nm)$ 時間の計算量で負荷分散枝被覆が得られる。

証明 アルゴリズムにより出力される枝被覆が最小であることは命題 1 から明らかである。即ち、増加パスを切り替えても連結成分 (スター) の数是不変であるので、出力される枝集合 M におけるスターの数は最大マッチングのサイズ $\nu(G)$ に等しい。上記と補題 10 を考慮すると、 $BEC2$ が出力する枝被覆はコスト減少パスの存在しない最小枝被覆、つまり負荷分散枝被覆であり、アルゴリズムの正当性が証明される。続いて、アルゴリズムの計算量の評価を行う。Step 3 と 4 における増加パスの探索に要する計算時間は前アルゴリズムと同様に幅優先探索により $O(m)$ 時間である。反復回数は丁度 $n-\nu(G)$ 回であり、これは Step 1 での最大マッチング M に対する自由頂点数 $|F_M|$ に等しい。したがって、 $BEC2$ により $O(nm)$ の計算時間で負荷分散枝被覆を求めることが可能である。 $\square$

5 おわりに

本研究では負荷分散枝被覆に対する最適性として枝最小性と負荷均等性を証明した。枝最小性は枝減少パス・閉路で、負荷均等性はコスト減少パスによって特徴付けを行った。さらに $O(nm)$

時間で動作するアルゴリズムを提案し、その正当性を示した。今後の課題として、他の手法を用いた計算量の改良を模索する。

参考文献

- [1] Bruno, J.L., Coffman, E.G. and Sethi, R.: Scheduling independent tasks to reduce mean finishing time, Communications of the ACM, vol. 17, pp. 382-387, 1974.
- [2] Edmonds, J.: Paths, trees, and flowers, Canadian Journal of Mathematics 17, pp. 449-467, 1965.
- [3] Gallai, T.: Über Extreme Punkt- und Kantenmengen, Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae, Sectio Mathematica 2, pp. 133-138, 1959.
- [4] Harvey, N.J.A., Ladner, R.E., Lovasz, L. and Tamir, T.: Semi-matchings for bipartite graphs and load balancing, Workshop on Algorithms and Data Structures (WADS), pp. 294-308, 2003.
- [5] Micali, S. and Vazirani, V.V.: An $O(V^{1/2}E)$ algorithm for finding maximum matching in general graphs, Proceedings of the 21st Annual IEEE Symposium on Foundations of Computer Science, pp. 12-27, 1980.
- [6] Norman, R.Z. and Rabin, M.O.: An algorithm for a minimum cover of a graph, Proceedings of the American Mathematical Society, 10:315-319, 1959.
- [7] Vazirani, V.V.: A theory of alternating paths and blossoms for proving correctness of the $O(\sqrt{V}E)$ general graph maximum matching algorithm, Combinatorica 14, pp. 71-109, 1994.