『経済研究』（明治学院大学）第 169 号　2025 年

スケールフリーグラフ再考 I

浜　口　幸　弘

1. スケールフリー現象はめったにない

この 20 年間にわたって，現実世界に見られる多くのネットワークはスケールフリーの性質を持っている，すなわち，次数 k のノード（頂点）の個数はべき法則 $k^{-\alpha}$（$2<\alpha<3$）に従うと主張されてきた．ところが，最近の Broid and Clauset（2019）の膨大かつ詳細な調査によれば，このスケールフリーの特徴を示す現象はむしろ稀有であることが見いだされた．その概略は以下のとおりである．

Broid and Clauset（2019）は，ノードの次数分布が $k^{-\alpha}$ に従う（$2<\alpha<3$），すなわち，べき分布するというスケールフリーの古典的定義に基づき，ICON（Index of Complex Networks）に収められている 928 の様々なネットワーク（社会，生物学，技術，輸送および情報ネットワークなど）を選び，そのスケールフリー性を統計的に詳細に検証した．その結果，全体の約 49％は完全にスケールフリーではないと判明し，全ネットワークの中でスケールフリー性が強いと判断されたのは約 10％にすぎないというものであった．

そこで本稿では最初の段階として，Cooper and Frieze（2003）の考え方に基づき，スケールフリーグラフを生成する preferrential attachment，すなわち次数の大きさに応じた頂点の選択方法と一様ランダムな頂点の選択方法の 2 つを確率的に併せ持つような，Cooper and Frieze（2003）の理論を部分的に変更したモデルを考え，実際の現象をどの程度説明できるか検討する．

2. スケールフリーグラフの生成について

スケールフリーグラフを生成する preferrential attachment のモデル化は次のとおりである．既存のグラフ G に新しい頂点を 1 個ずつ加えるとき，その頂点から m 本の辺を頂点の次数の大きさに応じて確率的にグラフ G に接続する．このとき，多重辺とループも認めるものとする．以下，具体的に m の

大きさで場合分けして考える．

$m=1$ の場合について，上記のことを定式化すると以下のようになる．既存のグラフ G に追加していく頂点列を $v_1, v_2, \ldots$ と固定し，グラフ G における頂点 v の次数を $d_G(v)$ と記す（以下，$d(v)$ と略記）．そこでランダムグラフのプロセス $(G_1^t)_{t \geq 1}$ を次のように帰納的に定義すると，$\{v_i : 1 \leq i \leq t\}$ 上のグラフ G_1^t が構成される．すなわち，まず 1 個の頂点と 1 つのループを持つグラフ G_1^1 から出発する．そして，G_1^{t-1} まで構成されたとき，次に追加する頂点 v_t と G_1^{t-1} 上の任意の頂点 u を 1 本の辺で結ぶ．このとき，G_1^{t-1} におけるすべての頂点の次数和は $2(t-1)$ なので，t において確率変数 u は次の式を満たす．

$$P(u=v_s)=d_{G_1^{t-1}}(v_s)/2(t-1) \qquad 1 \leq s \leq t-1.$$

これは，Barabási and Albert（1999）のモデルを定式化したものであるが，本稿では，Cooper and Frieze（2003）と同様に，preferrential attachment においては，追加する頂点 v_t が次数の大きさに従って G_1^{t-1} 内の頂点をランダムに選択するものとし（ただし，v_t 自身を選択することはない），このとき頂点の内次数と外次数の区別はしないものとする．すなわち，生成されるグラフは無向グラフである．

また，$m>1$ の場合については，追加する頂点 v_t から m 本の辺が一度に G_1^{t-1} に接続される．よって，プロセス（G_m^t）については，プロセス（G_1^{mt}）において，最初の頂点から m 個の頂点ごとに統合して 1 個の頂点にすれば（G_m^t）が得られることになる．したがって，本稿では mt を t と置き換え，追加する頂点から 1 本の辺が既存のグラフに接続される G_1^t グラフを扱うものとする．

3. 既存理論の問題点の指摘

本稿では，Cooper and Frieze（2003）の提案したランダムグラフの生成，すなわち，preferrential attachment に基づく頂点の選択と一様ランダムな頂点の選択を確率的に融合した理論に基づいて議論を進める．まず，その結果の問題点のいくつかを以下に指摘する．

(1) 補題 1 の証明における不等式の評価

証明が成り立つには，前半の不等式の評価では $k>j_1$ の条件が必要であり，後半の不等式の評価では，$k>j_1 a$ の条件が必要である．ところが補題 1 を用いる定理 6 では，$k \geq 1$ としている．

(2) 補題 2 における d_k に関する不等式の評価

例えば，本稿と同じ条件（$\alpha=0$，$j_0=1$，$j_1=0$）に設定して d_k の不等式を具体的に求めると，$j_1=0$ にもかかわらず，上限の式中の分母が j_1 となる（詳細は後述）．これは，証明中の(16)式が(15)式から導出できないことによると思われる（例えば，$j_1=2$ として考える）．さらに，$\beta=1$ の場合に d_k は幾何分布となることはわかっているが（Bollobás, Riordan, Spencer and Tusnády（2001）），この不等式からは導けない．

(3) $j_1=1$ の場合の d_k の導出（P. 11）

$k > j_0$ の条件下で d_k を求めているが，結果を見ると，k は十分大きいと仮定して d_k の近似式を導出している．

本稿ではこれらの問題点を踏まえた上で，Cooper and Frieze（2003）のモデルを簡略化してランダムグラフの生成モデルを考察する．

4. 本稿のモデルの提案

ここでは，Cooper and Frieze（2003）の定理導出の流れに従って，いくつかの定理の修正を行う．スケールフリーグラフの生成においては，Bollobás, Riordan, Spencer and Tusnády（2001）の研究もあるが，LCD（Linearized Chord Diagram）に基づいて定理を導出しているので，Cooper and Frieze（2003）にあるような 2 種類の頂点選択方法を融合したグラフの生成には向いていない．

まず，前述の Cooper and Frieze（2003）の問題点を踏まえた上で，そのモデルを簡略化して以下のように考えるものとする．なお，できる限り Cooper and Frieze（2003）にある記号を用いるようにする．

既存の生成グラフにおいては，頂点同士の接続は起こらないものとする．よって，Cooper and Frieze（2003）において，$\alpha=0$，$q_j=0$（$j\geq 1$），すなわち，$j_1=0$ である．そして，Cooper and Frieze（2003）と同様に，追加する頂点は確率 β で生成グラフの 1 個の頂点を一様ランダムに選択，または，確率 $1-\beta$ で生成グラフの 1 個の頂点を次数の大きさに従いランダムに選択する（すなわち，preferential attachment）ものとする（ただし，$0\leq\beta\leq 1$）．この点が当モデルの本質的な特徴である．よって，Cooper and Frieze（2003）において，$p_1=1$ および $p_j=0$（$j\geq 2$），すなわち，$j_0=1$ である．したがって，Cooper and Frieze（2003）の記号と対応付けると本稿のモデルは以下のようになる．

$$a=1+\beta,\ b=\frac{1-\beta}{2},\ c=\beta,\ d=\frac{1-\beta}{2},\ e=f=0,\ \mu_p=1,\ \mu_q=0,\ \theta=2. \tag{1}$$

さて，以下の補題では，第 t 回目のステップで生成される G_1^t における次数 k の頂点数を $D_k(t)$ とし，その期待値を $\overline{D}_k(t)$ と記す．よって，$t\geq 1$ に対して，$\overline{D}_0(t)=0$ であり，また，$\overline{D}_2(1)=1$，および $k\neq 2$ に対して，$\overline{D}_k(1)=0$ である．そして，追加する頂点 v_t に選択される頂点 $u_t\in\{v_1,\dots,v_{t-1}\}$ の次数が接続前に k である事象を $\{d(u_t)=k\}$ と記す．また，F を一様ランダムに頂点を選択する事象とし，$\overline{F}$ を頂点の次数の大きさに応じて頂点を選択する事象とする．なお，事象 A が起こるとき 1，起こらないときに 0 の値をとる指示関数（indicator function）を I_A で表す．

補題 1.

$\overline{D}_k(t)$ は以下の等式を満たす．

$k=1$ の場合，$t\geq 1$ に対して，

$$\overline{D}_1(t+1)=\overline{D}_1(t)-\frac{\beta+1}{2}\cdot\frac{\overline{D}_1(t)}{t}+1.$$

$k \geq 2$ の場合，$t \geq 1$ に対して，

$$\overline{D}_k(t+1) = \overline{D}_k(t) - \left(\beta + \frac{(1-\beta)k}{2}\right)\frac{\overline{D}_k(t)}{t} + \left(\beta + \frac{(1-\beta)(k-1)}{2}\right)\frac{\overline{D}_{k-1}(t)}{t}.$$

証明

まず，前述のグラフの生成方法に基づき，$\overline{D}_k(t)$ に関する漸化式を求めると以下のようになる．$k \geq 2$ の場合，

$$\begin{aligned} D_k(t+1) = D_k(t) &+ I_F(I_{\{d(u_{t+1})=k-1\}} - I_{\{d(u_{t+1})=k\}}) \\ &+ I_{\overline{F}}\ (I_{\{d(u_{t+1})=k-1\}} - I_{\{d(u_{t+1})=k\}}). \end{aligned}$$

よって，条件付き期待値をとると，

$$\begin{aligned} E(D_k(t+1) \mid G_1^t) = E(D_k(t) \mid G_1^t) &+ E(I_F(I_{\{d(u_{t+1})=k-1\}} - I_{\{d(u_{t+1})=k\}}) \mid G_1^t) \\ &+ E(I_{\overline{F}}(I_{\{d(u_{t+1})=k-1\}} - I_{\{d(u_{t+1})=k\}}) \mid G_1^t). \end{aligned}$$

さらに，両辺の期待値をとることで以下を得る．

$$\begin{aligned} \overline{D}_k(t+1) = \overline{D}_k(t) &+ P(F \cap \{d(u_{t+1}) = k-1\}) - P(F \cap \{d(u_{t+1}) = k\}) \\ &+ P(\overline{F} \cap \{d(u_{t+1}) = k-1\}) - P(\overline{F} \cap \{d(u_{t+1}) = k\}). \end{aligned}$$

したがって，以下となる．

$$\overline{D}_k(t+1) = \overline{D}_k(t) + \beta\left(\frac{\overline{D}_{k-1}(t)}{t} - \frac{\overline{D}_k(t)}{t}\right) + (1-\beta)\left(\frac{(k-1)\overline{D}_{k-1}(t)}{2t} - \frac{k\overline{D}_k(t)}{2t}\right).$$

すなわち，

$$\overline{D}_k(t+1) = \overline{D}_k(t) - \left(\beta + \frac{(1-\beta)k}{2}\right)\frac{\overline{D}_k(t)}{t} + \left(\beta + \frac{(1-\beta)(k-1)}{2}\right)\frac{\overline{D}_{k-1}(t)}{t}.$$

$k=1$ の場合，同様にして

$$\overline{D}_1(t+1) = \overline{D}_1(t) - \frac{\beta+1}{2} \cdot \frac{\overline{D}_1(t)}{t} + 1.$$

□

この結果は，Cooper and Frieze（2003）の(7)式において，$\alpha=0$，$j_0=1$，$p_1=1$ および $p_k=0$（$k \geq 2$）である場合に対応する．次に，Cooper and Frieze（2003）の(2)式 d_k と対応付けると，以下のようになる．

$$d_0 = 0, \qquad d_1 = \frac{1}{a+b} = \frac{2}{3+\beta},$$

$k \geq 2$ に対して，$d_k = \dfrac{c+d(k-1)}{a+bk}\, d_{k-1} = \dfrac{(1-\beta)k+3\beta-1}{(1-\beta)k+2+2\beta}\, d_{k-1}.$

補題 2.

$k \geq 1$ のとき，$d_k \leq \dfrac{2}{k}$ が成り立つ．

証明

k に関する帰納法で示す．$k \leq 2$ の場合，以下の不等式が成り立つ．

$$d_1 = \frac{2}{3+\beta} \leq 1.$$

$$d_2 = \frac{2(1-\beta)+3\beta-1}{2(1-\beta)k+2+2\beta} d_1 = \left(1-\frac{3-\beta}{4}\right)\frac{2}{3+\beta} \leq 1.$$

$k \geq 3$ のとき，$d_{k-1} \leq \dfrac{2}{k-1}$ が成り立つと仮定すると，以下となる．

$$d_k = \frac{(1-\beta)k+3\beta-1}{(1-\beta)k+2+2\beta} d_{k-1} \leq \left(1-\frac{3-\beta}{(1-\beta)k+2+2\beta}\right)\frac{2}{k-1}.$$

このとき，

$$\left(1-\frac{3-\beta}{(1-\beta)k+2+2\beta}\right)\frac{2}{k-1} \Big/ \frac{2}{k} = \left(1-\frac{3-\beta}{(1-\beta)k+2+2\beta}\right)\left(1+\frac{1}{k-1}\right)$$

$$= 1+\frac{1}{k-1}-\frac{3-\beta}{(1-\beta)k+2+2\beta}-\frac{3-\beta}{(k-1)((1-\beta)k+2+2\beta)}$$

$$= 1-\frac{(3-\beta)(k-1)-((1-\beta)k+2+2\beta)}{(k-1)((1-\beta)k+2+2\beta)}-\frac{3-\beta}{(k-1)((1-\beta)k+2+2\beta)}$$

$$= 1-\frac{2k-5-\beta}{(k-1)((1-\beta)k+2+2\beta)}-\frac{3-\beta}{(k-1)((1-\beta)k+2+2\beta)} \leq 1.$$

したがって，$d_k \leq \dfrac{2}{k}$ を得る． □

補題 1 および補題 2 を用いて次の定理 1 を示す．なお，$\overline{D}_2(1)=1$ および，$k \neq 2$ に対して，$\overline{D}_k(1)=0$ であることに注意する．また，Cooper and Frieze（2003）の式(13)と対比するために，記号 β の代わりに a，b，c，d を用いる．

定理 1.

$t \geq 1$，$k \geq 1$ のとき，ある定数 $M \geq 1$ に対して，以下の不等式が成り立つ．

$$|\overline{D}_k(t) - td_k| \leq M.$$

証明

まず，$\Delta_k(t) = \overline{D}_k(t) - td_k$ とする．また，$t \geq 1$ に対して，$\overline{D}_0(t)=0$ である．なお，以下の議論では，便宜上，$k \geq 0$ とする．そして，補題 1 および前述の d_k の漸化式から以下を得る．ただし，$p_1=1$ であり，$k \geq 2$ のとき，$p_k=0$ である．

$$\Delta_k(t+1)+(t+1)d_k$$

$$= \Delta_k(t)+td_k+p_k+\frac{1}{t}\Big((1-(a+bk))(\Delta_k(t)+td_k)+(c+d(k-1))(\Delta_{k-1}(t)+td_{k-1})\Big).$$

よって，

$$\Delta_k(t+1) = \left(1-\frac{a+bk-1}{t}\right)\Delta_k(t)+\frac{1}{t}(c+d(k-1))\Delta_{k-1}(t).$$

上の結果をもとに，t に関する帰納法で主張の不等式を示す．まず，$t\geq 1$ に対して，$|\Delta_0(t)|=0$ であり，$k\geq 0$ に対して，$|\Delta_k(1)|=|\overline{D}_k(1)-d_k|\leq 1$ である．

次に，任意の $k\geq 0$ に対して，$t(\geq 1)$ の場合に，$|\Delta_k(t)|\leq M$ が成り立つと仮定する．$0\leq\beta<1$ の場合，$b>0$ なので，$k_0(t)=\left\lfloor\dfrac{t+1-a}{b}\right\rfloor=2+\left\lfloor\dfrac{2(t-1)}{1-\beta}\right\rfloor$ とする．このとき，$k_0(t)>0$ である．$k>k_0(t)$ の場合，$t+1$ までに次数 k の頂点に接続する辺の合計数は，そのときまでに追加される辺の合計数の 2 倍以下であるから，以下の式が成り立つ．

$$k_0(t)D_k(t+1)<kD_k(t+1)\leq 2(t+1).$$

よって，$D_k(t+1)\leq\dfrac{2(t+1)}{k_0(t)}=O(1)$ である．また，$k\geq 1$ のとき補題 2 から以下となる．

$$(t+1)d_k\leq(t+1)\frac{2}{k}<\frac{2(t+1)}{k_0(t)}=O(1).$$

そして $k=0$ では，$(t+1)d_0=0$ であるから，$|\Delta_k(t+1)|\leq M$ となる．$\beta=1$ の場合，$a=2$，$b=0$ なので，$1-\dfrac{a+bk-1}{t}=1-\dfrac{1}{t}\geq 0$ となり，以下の議論に含めることができる．$k\leq k_0(t)$ すなわち，$1-\dfrac{a+bk-1}{t}\geq 0$ の場合，上の $\Delta_k(t+1)$ の漸化式から帰納法の仮定を用いて以下となる．

$$\begin{aligned}|\Delta_k(t+1)|&\leq|\Delta_k(t)|\left(1-\frac{a+bk-1}{t}\right)+|\Delta_{k-1}(t)|\frac{1}{t}(c+d(k-1))\\&\leq M\left(1-\frac{a+bk-1}{t}+\frac{c+d(k-1)}{t}\right)\\&\leq M.\end{aligned}$$

最後の不等式は，$b=d$ および $c=a-1$ による．したがって，任意の $k\geq 0$ に対して，$|\Delta_k(t+1)|\leq M$ となる． □

この定理から $\overline{D}_k(t)$ は，十分大きな t に対して，おおよそ td_k になることが導かれた．次に，$k\geq 1$ に対して，d_k の大きさを β の値に応じて評価する不等式を与える．

補題 3.

$A_i\ (1\leq i\leq 5)$ はある正の定数とする．このとき，d_k についての以下の不等式が成り立つ．

$\beta=0$ のとき，

$$A_1\left(1-\frac{3}{k+2}\right)^{k-1}\left(\frac{e}{k+2}\right)^3\leq d_k\leq A_2\left(1-\frac{3}{k+3}\right)^{k}\left(\frac{e}{k+3}\right)^3.$$

$0<\beta<1$ のとき，

$$A_3\left(1-\frac{3-\beta}{(1-\beta)k+2+2\beta}\right)^{k+\frac{2(1+\beta)}{1-\beta}}\left(\frac{2e}{(1-\beta)k+3\beta-1}\right)^{\frac{3-\beta}{1-\beta}}\leq d_k$$

$$\leq A_4\left(1-\frac{3-\beta}{(1-\beta)k+3+\beta}\right)^{k+\frac{3+\beta}{1-\beta}}\left(\frac{2e}{(1-\beta)k+2\beta}\right)^{\frac{3-\beta}{1-\beta}}.$$

$\beta=1$ のとき，

$$d_k=A_5\left(\frac{1}{2}\right)^{k-1}.$$

証明

前述のように，d_k は以下の漸化式を満たす．

$$d_1=\frac{1}{a+b}=\frac{2}{3+\beta},$$

$k\geq 2$ に対して，$d_k=\dfrac{c+d(k-1)}{a+bk}\,d_{k-1}=\left(1-\dfrac{3-\beta}{(1-\beta)k+2+2\beta}\right)d_{k-1}$.

よって，$0<\beta<1$ のとき，$k\geq 2$ に対して，

$$d_k=d_1\prod_{j=2}^{k}\left(1-\frac{3-\beta}{(1-\beta)j+2+2\beta}\right)=\frac{2}{3+\beta}\exp\left(\log\prod_{j=2}^{k}\left(1-\frac{3-\beta}{(1-\beta)j+2+2\beta}\right)\right)$$

$$=\frac{2}{3+\beta}\exp\left(\sum_{j=2}^{k}\log\left(1-\frac{3-\beta}{(1-\beta)j+2+2\beta}\right)\right).$$

ここで以下の不等式が成り立つ．

$$\int_1^k\log\left(1-\frac{3-\beta}{(1-\beta)x+2+2\beta}\right)dx\leq\sum_{j=2}^{k}\log\left(1-\frac{3-\beta}{(1-\beta)j+2+2\beta}\right)$$

$$\leq\int_2^{k+1}\log\left(1-\frac{3-\beta}{(1-\beta)x+2+2\beta}\right)dx.$$

このとき，

$$\int\log\left(1-\frac{3-\beta}{(1-\beta)x+2+2\beta}\right)dx$$

$$=\int\log((1-\beta)x+3\beta-1)\,dx-\int\log((1-\beta)x+2+2\beta)\,dx$$

なので，以下となる．

$$\log\left(\frac{3+\beta}{2\beta}\right)^{\frac{3+\beta}{1-\beta}}\left(\frac{\beta}{e}\right)^{\frac{3-\beta}{1-\beta}}\left(1-\frac{3-\beta}{(1-\beta)k+2+2\beta}\right)^{k+\frac{2+2\beta}{1-\beta}}\left(\frac{2e}{(1-\beta)k+3\beta-1}\right)^{\frac{3-\beta}{1-\beta}}$$

$$\leq\sum_{j=2}^{k}\log\left(1-\frac{3-\beta}{(1-\beta)j+2+2\beta}\right)$$

$$\leq\log\left(\frac{4}{1+\beta}\right)^{\frac{4}{1-\beta}}\left(\frac{1+\beta}{2e}\right)^{\frac{3-\beta}{1-\beta}}\left(1-\frac{3-\beta}{(1-\beta)k+3+\beta}\right)^{k+\frac{3+\beta}{1-\beta}}\left(\frac{2e}{(1-\beta)k+2\beta}\right)^{\frac{3-\beta}{1-\beta}}.$$

これらのことから．以下を得る．

$$A_3\left(1-\frac{3-\beta}{(1-\beta)k+2+2\beta}\right)^{k+\frac{2(1+\beta)}{1-\beta}}\left(\frac{2e}{(1-\beta)k+3\beta-1}\right)^{\frac{3-\beta}{1-\beta}}\leq d_k$$

$$\leq A_4\left(1-\frac{3-\beta}{(1-\beta)k+3+\beta}\right)^{k+\frac{3+\beta}{1-\beta}}\left(\frac{2e}{(1-\beta)k+2\beta}\right)^{\frac{3-\beta}{1-\beta}}.$$

$\beta=0$ および $\beta=1$ のときも同様にして，主張の不等式を得る．　□

次に，$D_k(t)$ はどの程度 $\overline{D}_k(t)$ 付近，すなわち td_k 付近に集中するかを Azuma－Hoeffding bound を利用して考察する．以下の定理 2 の証明では，事象 H_i を第 i 回目のステップにおいて，追加する頂点 v_i とそれまでに生成されたグラフ上の任意の頂点 $u_i\in\{v_1,\ldots,v_{i-1}\}$ とのランダムな接続とする（$1\leq i\leq t$）．そして，一連の事象を $H=H_1,\ldots,H_t$ とし，ある関数 f に対して，$X_t=f(H)$ とする．このとき，Azuma－Hoeffding bound を利用するために，$1\leq i\leq t$ に対して，以下の式の上限を評価する．

$$|(E(X_t|H_1,\ldots,H_i)-E(X_t|H_1,\ldots,H_{i-1})|.$$

また，$X_t=D_k(t)$ とし，頂点列 $u_1,\ldots,u_{i-1},x,u_{i+1},\ldots,u_{t-1}$ を順次選択して，t 回目のステップで H_t が起こる確率 $P(H_t|H_1,\ldots,H_{i-1},H(x),H_{i+1},\ldots,H_{t-1})$ は（$H(x)$ は i 回目に x を選択する事象），

$$\begin{aligned}&P(H_t|H_1,\ldots,H_{i-1},H(x),H_{i+1},\ldots,H_{t-1})\\&=P(F|H_1,\ldots,H_{i-1},H(x),H_{i+1},\ldots,H_{t-1})P(H_t|H_1,\ldots,H_{i-1},H(x),H_{i+1},\ldots,H_{t-1},F)\\&\quad+P(\overline{F}|H_1,\ldots,H_{i-1},H(x),H_{i+1},\ldots,H_{t-1})P(H_t|H_1,\ldots,H_{i-1},H(x),H_{i+1},\ldots,H_{t-1},\overline{F})\end{aligned}$$

であり，F と $\overline{F}$ を含むことを意味している．

定理 2.

$D_k(t)$ について，以下の不等式が成り立つ．

$$P(|D_k(t)-\overline{D}_k(t)|\geq\sqrt{t\log t}\,)\leq\frac{2}{t^{1/8}}.$$

証明

第 i—1 回目のステップ迄で生成グラフ内の頂点列 $u_1,\ldots,u_{i-1}$ が順次選択され，第 i 回目のステップにおいて，追加頂点 v_i が生成されたグラフの頂点 a とランダムに接続する一連の過程を H_A とし，一方，頂点 $b\neq a$ とランダムに接続する一連の過程 H_B とする．また，$K=\{a,b\}$，$V_t=\{v_1,\ldots,v_t\}\setminus K$ とする．このとき，$K\cup\{v_i\}$ において両過程の辺の接続の仕方は異なるが，両 $K\cup\{v_i\}$ における頂点数と次数の総和は等しく，また，$\{v_1,\ldots,v_{i-1}\}\setminus K$ においては，両過程の頂点数，各頂点の次数および辺の接続の仕方は同じである．したがって，t 回目のステップまで考えると，同じ頂点列（ただし，K は次数 $d(a)+d(b)$ の一つの頂点 u とみなす）を選択する確率は両過程において一致する．そこで，第 i—1 回目のステップ迄で生成グラフ内の頂点 $u_1,\ldots,u_{i-1}$ が選択され，第 i 回目において頂点 x が選択されるとき，X_t の条件付き期待値 $E_i(x)$ を以下のように定義すれば，

$$E_i(x)=E(X_t|H_1,\ldots,H_{i-1},H(x)),$$

$|E_i(a)-E_i(b)|\leq 2$ となる．

また，頂点 $u_1,\ldots,u_{i-1}$ が選択される条件下で，第 i 回目において頂点 x の選択される条件付き確率を

$P_i(x)$ と記す．このとき以下を得る．

$$E_i(a) = E(X_t|H_1,\dots,H_{i-1},H(a)) = P_i(a) \times E_i(a) + \left(\sum_{b \neq a} P_i(b)\right) \times E_i(a).$$

ただし，b は第 i 回目においてランダムに選択される $b \neq a$ となる任意の頂点である．一方，期待値の性質から以下の式を得る．

$$E(X_t|H_1,\dots,H_{i-1}) = P_i(a) \times E_i(a) + \sum_{b \neq a} P_i(b) \times E_i(b).$$

よって，以下の不等式を得る．

$$\begin{aligned}&|E(X_t|H_1,\dots,H_{i-1},H(a)) - E(X_t|H_1,\dots,H_{i-1})|\\ \le &\sum_{b \neq a} P_i(b) \times |E_i(a) - E_i(b)|.\end{aligned}$$

したがって，前述の式から以下を得る．

$$|E(X_t|H_1,\dots,H_{i-1},H(a)) - E(X_t|H_1,\dots,H_{i-1})| \le 2.$$

このとき，$Z_i = E(X_t|H_1,\dots,H_i)$ とすると（$H_i = H(a)$ とする），$|Z_i - Z_{i-1}| \le 2$（$1 \le i \le t$）であり，$\{Z_i\}_1^t$ は $H_1,\dots,H_t$ に関して Martingale なので，Azuma－Hoeffding bound によって，任意の $c > 0$ に対して，以下の不等式が成り立つ（例えば，Frieze and Karoňski（2016）参照）．

$$P(|Z_t - E(Z_t)| \ge c) \le 2\exp\left(-\frac{c^2}{8t}\right).$$

$Z_t = X_t = D_k(t)$ であり，$c = \sqrt{t \log t}$ とすれば，以下を得る．

$$P(|D_k(t) - \bar{D}_k(t)| \ge \sqrt{t \log t}\,) \le \frac{2}{t^{1/8}}.$$

よって，主張の不等式を得る．□

定理 2 によれば，十分大きな t に対して $D_k(t)$ の値はほぼ $\bar{D}_k(t)$，すなわち，td_k になることがわかる．

5. 議論

本稿で得られた結果と Cooper and Frieze（2003）の補題 2 を比較してみる．本稿の簡素化されたモデルを Cooper and Frieze（2003）と対比すると，$\alpha = 0$，$j_1 = 0$ であり，$b = d = \dfrac{1-\beta}{2}$，$e = f = 0$ である．よって Cooper and Frieze（2003）の補題 2 は以下となる．

$$C_1/k^{\frac{3-\beta}{1-\beta}} \le d_k \le C_1/k^{\frac{3-\beta}{(1-\beta)/j_1}}.$$

このとき，上限は不確定となる．また，本稿の結果では，一様ランダムな選択の割合が増えれば，すなわち，β が大きくなるにつれて d_k は幾何分布の性質を持ち始める，すなわち次数 k が大きくなるほど d_k はべき分布の場合に比べて相対的に小さくなる．この点において，Cooper and Frieze（2003）の補

題 2 と大きく異なっている．

さて，Broid and Clauset（2019）は，べき分布と比較するため 4 つの種類の分布を挙げており，これらとの比較の結果として，現実のネットワークのノード数はべき分布よりむしろ対数正規分布（Log-normal）に最も合致すると指摘している．したがって，次数 k が大きくなるほど，その次数をもつノードの数はべき分布の場合よりも少なくなる，すなわち，べき分布よりもしっぽ（tail）の部分が細くなるということになる．この点から見ると，本稿の示す d_k はその性質を相対的に満たすことになる．これは現実世界のネットワークの生成においては，次数の大きさに応じて単純に頂点が選択される訳ではなく，ある一定の割合で一様ランダムなノードの選択が行われる，あるいは他の選択が行われると推測できる．次の機会に，この点についてさらに考察を深め行く予定である．

参考文献

A. Barabási, and R. Albert, *Emergence of scaling in random networks*, Science 286（1999）.

B. Bollobás, O. Riordan, J. Spencer and G. Tusnády, *The degree sequence of a scale-free random graph process*, Random Structures & Algorithms, Wiley Online Library（2001）.

A.D. Broido and A. Clauset, *Scale-free networks are rare*, Nature communications（2019）.

C. Cooper and A. Frieze, *A general model of web graphs*, Random Structures & Algorithms, Wiley Online Library（2003）.

A. Frieze, A and M. Karoński, *Introduction to random graphs,* Cambridge University Press（2016）.